96th Congress } JOINT COMMITTEE PRINT
2d Session }

SPECIAL STUDY ON ECONOMIC CHANGE

VOLUME 5

GOVERNMENT REGULATION: ACHIEVING SOCIAL AND ECONOMIC BALANCE

STUDIES

PREPARED FOR THE USE OF THE

SPECIAL STUDY ON ECONOMIC CHANGE

OF THE

JOINT ECONOMIC COMMITTEE
CONGRESS OF THE UNITED STATES

DECEMBER 8, 1980

Printed for the use of the Joint Economic Committee

U.S. GOVERNMENT PRINTING OFFICE
56-368 WASHINGTON : 1980

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

JOINT ECONOMIC COMMITTEE

(Created pursuant to sec. 5(a) of Public Law 304, 79th Cong.)

LLOYD BENTSEN, Texas, *Chairman*
RICHARD BOLLING, Missouri, *Vice Chairman*

SENATE

WILLIAM PROXMIRE, Wisconsin
ABRAHAM RIBICOFF, Connecticut
EDWARD M. KENNEDY, Massachusetts
GEORGE McGOVERN, South Dakota
PAUL S. SARBANES, Maryland
JACOB K. JAVITS, New York
WILLIAM V. ROTH, JR., Delaware
JAMES A. McCLURE, Idaho
ROGER W. JEPSEN, Iowa

HOUSE OF REPRESENTATIVES

HENRY S. REUSS, Wisconsin
WILLIAM S. MOORHEAD, Pennsylvania
LEE H. HAMILTON, Indiana
GILLIS W. LONG, Louisiana
PARREN J. MITCHELL, Maryland
CLARENCE J. BROWN, Ohio
MARGARET M. HECKLER, Massachusetts
JOHN H. ROUSSELOT, California
CHALMERS P. WYLIE, Ohio

JOHN M. ALBERTINE, *Executive Director*
LOUIS C. KRAUTHOFF II, *Assistant Director-Director, SSEC*
RICHARD F. KAUFMAN, *Assistant Director-General Counsel*
CHARLES H. BRADFORD, *Minority Counsel*

SPECIAL STUDY ON ECONOMIC CHANGE

LOUIS C. KRAUTHOFF II, *Director*
GEORGE D. KRUMBHAAR, Jr., *Counsel*
DOUGLAS N. ROSS, *Senior Economist*
TIMOTHY P. ROTH, *Senior Economist*
RICHARD D. BARTEL, *Economist*

LETTERS OF TRANSMITTAL

NOVEMBER 28, 1980.

To the Members of the Joint Economic Committee:

Transmitted herewith is a staff study, printed separately, and technical papers which together form Volume 5 of the Special Study on Economic Change (SSEC).

Volume 5 is entitled "Government Regulation: Achieving Social and Economic Balance" and is one of 10 areas on different aspects of the economy published by the SSEC. The SSEC was initiated in 1978 under the direction of the former Chairman of the Joint Economic Committee, Representative Richard Bolling, then Vice Chairman Senator Hubert H. Humphrey, and the former Ranking Minority Member, Senator Jacob K. Javits. It is intended to identify major changes in the economy and to analyze their implications for policymakers. The successful completion of this Study will, I believe, help provide an economic agenda for the United States for the decade of the 1980's.

The views expressed in the technical papers are exclusively those of the authors and do not necessarily represent the views of the Joint Economic Committee or of individual members. The staff study, which summarizes significant issues raised in the technical papers, was approved by the Chairman's Special Study Review Committee formed by the Chairman, Representative Bolling, Ranking Minority Member Representative Clarence J. Brown, and Senator Javits.

Sincerely,

LLOYD BENTSEN,
Chairman, Joint Economic Committee.

NOVEMBER 24, 1980.

Hon. LLOYD BENTSEN,
Chairman, Joint Economic Committee,
Congress of the United States,
Washington, D.C.

DEAR MR. CHAIRMAN: Transmitted herewith is a staff study, printed separately, and technical papers entitled "Government Regulation: Achieving Social and Economic Balance," which constitute Volume 5 of the Special Study on Economic Change (SSEC).

The SSEC was initiated under the leadership of former Chairman of the Joint Economic Committee, Representative Richard Bolling, Vice Chairman Senator Hubert H. Humphrey, and former Ranking

Minority Member, Senator Jacob K. Javits. The Study is divided into 10 substantive areas, which together chart major changes in the economy and analyze their implications for policymakers. Volume 5 comprises an economic analysis of the effects and benefits of government regulation and the costs borne by the private sector.

This study looks at the regulation in both general terms and with regard to individual cases, to provide policy analysts with a wealth of valuable information on this important subject.

It should be understood that the views expressed in the technical papers are exclusively those of the authors and do not necessarily represent the views of the Joint Economic Committee or of individual members. The staff study, which summarizes significant issues raised in the technical papers, was approved by the Chairman's Special Study Review Committee formed by the Chairman, Representative Bolling, Ranking Minority Member Representative Clarence J. Brown, and Senator Javits.

Sincerely,

JOHN M. ALBERTINE,
Executive Director, Joint Economic Committee.

CONTENTS

	Page
Letters of transmittal	III

GOVERNMENT REGULATION: ACHIEVING SOCIAL AND ECONOMIC BALANCE

Regulatory Change, 1960–79, in Historical Perspective—Thomas K. McCraw	1
The Political Economy of Wage-Price Policies: An Historical Review—Arnold R. Weber	18
The Role of Antitrust in a Deregulated Environment—Ronald R. Braeutigam	48
The Underground Economy: Estimates of Size, Structure, and Trends—Carl P. Simon and Ann D. Witte	70
The Effects of Foreign Government Intervention on the U.S. Competitive Position: A Perspective of Change—Mira Wilkins	121
The Regulatory Budget as a Management Tool for Reforming Regulation—Christopher DeMuth, Richard H. Shackson, Eric O. Stork, and Arthur W. Wright	141
The Impact of Regulation on the Performance of Industry—Paul W. MacAvoy and Dorothy M. Tella	176
Common Carrier Regulation and Technological Change: The New Competition in the Communications Industries—Nina W. Cornell and Douglas W. Webbink	197
The Impact of Regulation of Financial Institutions on Competition, and the Allocation of Resources—Paul M. Horvitz	225
Alternatives to Regulation in the Health Care Sector—Warren Greenberg	257
The Differential Impact of Government Regulation on Small Versus Large Business: A Review of an Inadequate Information Base—William E. Diamond	270
The Economic Effects of Federal Regulatory Activity on State and Local Governments: Concepts and Proposals—Robert E. Firestine	296
The Impact of Selected Federal Actions on Municipal Outlays—Thomas Muller and Michael Fix	326
Federal Regulations and Higher Education—Carl Kaysen and Crystal C. Lloyd-Campbell	378
Socially Imposed Costs of Higher Education—Howard R. Bowen	430
Regulation in the Federal System: What Do We Want From the Intergovernmental System?—Robert B. Hawkins, Jr	458

REGULATORY CHANGE, 1960–79, IN HISTORICAL PERSPECTIVE

By Thomas K. McCraw

CONTENTS

	Page
Summary	1
I. Change Since 1960	2
A. The Landis Report of 1960	3
B. The New Regulation and Its Meaning	4
C. The Growth of the Regulated Sector	5
D. Changing Perceptions of Regulation	6
1. By Scholars	6
2. By the Public	8
II. Why the Regulatory Changes Since 1960 Have Occurred	9
A. The Progressive Era, 1901–21	10
1. The Context of Progressivism	10
2. 1901–21 Compared with 1960–79	10
B. The New Deal, 1933–38	11
1. The Context of the New Deal	11
2. The Legacy of the New Deal	11
C. The Period 1960–79	12
1. The Basic Dissimilarity	12
2. The Reasons for the Dissimilarity: "Single-Issue Politics"	12
III. Implications for the Major Players	14
A. Academic Analysts of Regulation	14
B. The Government	15
C. The Regulators	16
D. Business Executives	17

SUMMARY

The purpose of this paper is to place the regulatory change of the last 20 years in historical perspective. The paper takes a broad view, generalizes freely, and overstates some of its arguments to highlight its principal themes.

The first theme is that the most recent era (1960–79) differs from the two preceding periods of regulatory hyperactivity in that it has been a time of two partially conflicting trends: one toward more regulation, the other toward deregulation. In neither of the two previous periods—the progressive era (1901–21) and the New Deal (1933–38)—was there such a strong trend toward deregulation at the same time additional regulatory powers were accruing to the State and Federal Governments.

Similarly, the two earlier periods differed from the period 1960–79 in that most of the regulatory legislation was sponsored by strong, activist, reform-minded Chief Executives—Theodore Roosevelt and Woodrow Wilson during the progressive era, Franklin D. Roosevelt during the New Deal. By contrast, during the most recent period it seemed to matter little who was President. Congress has issued regulatory legislation during the activist, reform-minded administration

of Lyndon Johnson, the noninterventionist, free market-oriented administration of Gerald Ford, and the in-between administrations of John F. Kennedy, Richard Nixon, and Jimmy Carter. Unlike the situation in previous eras, the source of most of the recent regulatory initiatives was Congress itself, a reflection of the growing independence of the legislative branch and the rise in the 1960's and 1970's of "single-issue politics." First triumphant in the civil rights movement, single-issue politics moved from noneconomic issues such as civil rights and the antiwar movement to partially economic issues such as consumerism, environmental protection, and health and safety. Closely related to the decline of party discipline within Congress, the emergence of single-issue politics represents a significant change in the way American Government works, and it has been a vital ingredient in both the growth of social and environmental regulation and the decline of economic regulation for industries such as airlines.

Intellectually, the most recent period differs from the earlier ones in that price theorists have become the leading scholars of regulation, displacing the lawyers and institutional economists who dominated discourse during the progressive era, and the political scientists who held primacy during the 1930's. This dominance by price theorists may be ephemeral, or it may be permanent. In either case, it is helpful to an understanding of the present period to speculate about the policy implications of the primacy at a particular time of one methodology over another.

I. Change Since 1960

One characteristic that sets off the recent period of regulatory change from earlier periods is that it has been a time of not one, but two distinct and partially conflicting trends. The first trend was toward greater government regulation. Every literate citizen is aware of the proliferation of such agencies as Environmental Protection Agency, Department of Energy, Occupational Safety and Health Administration, Consumer Product Safety Commission and National Highway Traffic Safety Administration.

The second trend, which at any given time during the period 1960–79 lagged behind the first by several years, is toward deregulation. Though not quite the cliche that "more regulation" has become, "deregulation" is nonetheless familiar to most Americans. Some have already experienced its pleasant effects, in the lower prices they pay for airline tickets and for securities transactions in the stock market. Others have felt its unpleasant effects, in the rising prices of decontrolled natural gas, petroleum products, and rents for apartments.

In any event, all of us—scholars, legislators, the general public—are at this moment confronting the peculiar spectacle of two powerful trends that seem to be directly at odds with each other. One is toward greater government regulation, the other toward less. The irony of the situation, the apparent paradox, is a theme of this paper, and much of what is said in all three sections of the paper is an attempt to inquire why this paradox exists, why it did not exist in similar form in the earlier periods of regulatory hyperactivity, and what it may augur for the future of regulation within the United States. The organization of the paper proceeds on these lines: the first section explores the

reasons why the two trends grew simultaneously during the 1960–79 period, the second delineates the ways in which the recent period differs from two earlier reform eras, and the third assesses the meaning of the conflicting trends for the American economy and the American polity. The aim throughout the paper is to provoke thought about regulatory change, and toward that end I have deliberately overstated some of the arguments and interpretations.

A. The Landis Report of 1960

One of the landmark documents in regulatory history is the so-called Landis Report to President-elect Kennedy, delivered in December 1960. The author of the report, James M. Landis, had credentials of expertise in regulatory matters that few Americans, if any, could match. A pupil and disciple of Felix Frankfurter, Landis was one of the bright young men of the New Dean who drafted and administered important reform legislation. He was a principal author of both the Securities Act of 1933 and the Securities Exchange Act of 1934. He served as a commissioner on three major agencies—the Federal Trade Commission, the Securities and Exchange Commission, and the Civil Aeronautics Board. In between regulatory assignments he saw duty as a White House troubleshooter and served as dean of the Harvard Law School.

In short, Landis knew whereof he spoke when he undertook to discuss affairs of law and regulation. Thus, when John F. Kennedy asked him to survey the current state of Federal regulation and report the results, a substantial proportion of the regulatory subculture in Washington—commissioners, congressional subcommittee members and their staffs, executives of regulated industries, and the regulatory bar—awaited the results with an anxious mixture of curiosity and fear (of radical change in the regulatory subculture which Landis might have recommended). Few members of this subculture had much to gain from regulatory change, and some took comfort in the fact that at least James Landis was one of their own, at least he had experience in the field. Furthermore, Landis was the author of what to this day remains the most persuasive theoretical treatise in favor of regulation ever written—"The Administrative Process," which he delivered at the Storrs Lectures at Yale in 1938.

By 1960, however, when the Landis report to the president-elect appeared, it was evident from the first page that its author has undergone a profound change in the years since the New Deal. Instead of a defense of regulation, or a temporizing analysis that evaded the real issues, the Landis report went boldly on the offensive. It offered harsh indictments of regulatory performance in agency after agency, detailing the inefficiency and cronyism that had invaded nearly all the commissions. The flavor and specificity of the report may be inferred from the following quotation:

> Inordinate delay characterizes the disposition of adjudicatory proceedings before substantially all of our regulatory agencies. In the Civil Aeronautics Board, for example, the average age of dockets closed by formal proceedings in 1960 was some 32 months. As of June 30, 1959, out of 464 proceedings then pending, 166 had been pending for more than 3 years. The Federal Trade Commission as of June 30, 1959, had 309 cease and desist orders pending, of which 118 had

been pending for more than 1 year and 30 for more than 3 years. In the Federal Power Commission the backlog of pending cases in 1959 was almost four times as great as in 1957. Only last September that Commission announced that it would take 13 years with its present staff to clear up its pending 2,313 producer rate cases pending [*sic*] as of July 1, 1960, and that within the contemplated 6,500 cases that would be filed during that 13-year period it could not become current until 2043 A.D. even if its staff were tripled. Contested proceedings before the Interstate Commerce Commission tend to run from 18 to 36 months, and numerous proceedings before the Federal Communications Commission and the Maritime Board have been pending for more than 3 years. The statutory period of 20 days during which a normal registration statement covering the issuance of new securities becomes effective under the Securities Act of 1933 has in practice been lengthened to some 40 to 60 days. Numerous similar statistics can be gathered from other agencies, including individual instances when even 10 and 14 years have elapsed before a final determination has been made. They all corroborate the fact of interminable delay.[1]

The Landis report went on in the same spirit for some 87 printed pages. It ended with a list of proposals that would promote the centralization of administrative responsibility in the chairmen of the agencies, so that they might wield authority commensurate with the complex tasks assigned to them. Landis emphasized the essential nexus between the president and the commission chairmen. He recommended that the chief executive be given wider authority over the behavior of the agencies and the appointment of their chairmen.

B. The New Regulation and Its Meaning

In terms of the paradox with which this paper began—that is, the irony today of parallel and simultaneous drives toward greater regulation and toward deregulation—the Landis report of 1960 may be seen to bear the seeds of both movements. In detailing the administrative delays characteristic of the existing agencies, Landis was in effect lamenting that the agencies' tasks were impossible to fulfill. Though he was too committed to the regulatory solution ever to put it in quite this way, the fact was that some of the agencies on which he focused—the Interstate Commerce Commission and the Civil Aeronautics Board for example—were engaged in economic regulation of an inherently inappropriate nature. Proceeding from an inapplicable premise (at least to the trucking and airline industries) of "natural monopoly," the agencies ruled on the fairness and reasonableness of rates, fares, and other fees. They controlled entry and exit for particular industries, and they often set the terms of competitive behavior. Insofar as Landis' indictment of economic regulation grew out of a recognition that something was deeply wrong not only with the commissions but also with the function itself, he may be seen as an early prophet—the John the Baptist, so to speak—of deregulation. That Landis was too tied to the old ways to make the final break was less important than that he saw so clearly that the system which he himself had helped to design was not working, and might never work.

At the same time, as an unreconstructed New Deal liberal, Landis was in full sympathy with the humanitarian thrusts of Kennedy's New Frontier, from which some important subsequent regulatory legislation grew. This "new" functional form of regulation, so different from the "old" economic regulation, focused on the rights of citizens

[1] James M. Landis, "Report on Regulatory Agencies to the President-Elect" (U.S. Senate, Committee on the Judiciary, 86th Congress, 2d session, 1960).

to a safe and healthy workplace, to equal employment opportunity, and to a clean environment. Though James Landis died in 1964, just as the "new social and environmental," or "functional, cross-industry" regulation was beginning, one senses that he would have given it his wholehearted approval and support.

Although some important regulatory legislation of the new variety grew from the initiatives of the Kennedy administration, its chief source was neither the Kennedy, Johnson, Nixon, Ford, nor Carter administrations. Indeed it was not the executive branch at all, but Congress, and the legion of congressional staff members on the lookout for issues through which their principals—particular congressmen and senators—could attain visibility and national prominence. Thus, Senator Muskie (D.-Me.) became identified with the environmental movement, Senator Magnuson (D.-Wash.) and Representative Moss (D.-Cal.) with consumer issues, and Senator Kennedy (D.-Mass.) with health care.

Another striking aspect of the new regulation was the rapidity of its rise, and the suddenness with which a consensus of the American electorate came to support it. This in turn reflected a rising consumer consciousness, a sophisticated (sometimes excessively so, to the point of misplaced cynicism) understanding of the way politics and business interacted, and an impatience with the ability of existing institutions to cope with new problems. It also reflected the rise of "single-issue politics," around which ad hoc coalitions are constructed without reference to existing party structures and in the face of attempts to impose party discipline. This theme of single-issue politics will be discussed in a later section of this paper.

C. The Growth of the Regulated Sector

It is unnecessary to elaborate in further detail on the growth of the regulated sector. It may be useful, however, to address three common misconceptions or exaggerated notions about the nature of regulatory change since 1960. The first is that the new social and environmental regulation has negatively affected American industry across the board.

The first misconception is that the new social and environmental regulation has negatively affected American industry across the board. This line of thought goes as follows: The "old" regulation was industry-specific: the CAB regulated the airlines, the ICC railroads and trucking, the FCC broadcasting and telecommunications. The new regulation, by contrast, is "cross-industry," and therefore—like the symbolic representation of justice—it is blind to differences among industries or among firms. A closer look would show that the impact of the new regulation has been very different for different industries (it has profoundly affected automobiles, chemicals, and utilities, for example, but has touched only lightly upon textiles, machinery, and agribusiness). Furthermore, the new regulation has affected big business in ways different from small business, and it has helped or hurt individual firms within the same industry to such a differential extent that it has changed the basis of competition within those industries.[2] Finally, some of the new regulation, particularly environmental

---

[2] See Robert A. Leone, "The Real Costs of Regulation," Harvard Business Review (November–December 1977), pp. 57–66.

and health and safety regulation, has promoted the rise of important new industries such as those manufacturing pollution-abatement equipment.

A second misconception is that the United States has become saddled with huge new parasitic Federal bureaucracies erected to administer the new rules. In fact, little overall growth has occurred in the Federal bureaucracy in the last 20 years, relative to the growth of the population in general; and such growth as has occurred in the public sector is mostly confined to State and local governments. It is true that several new agencies have appeared and grew at a rapid rate. The Environmental Protection Agency, for example, quickly grew from practically nothing to an organization employing more than 10,000 persons. This is an unusual example, however, and the general rule has been of modest growth, at least in terms of personnel. But personnel is a poor yardstick, and in terms of impact on the private sector, the growth on the new regulation has been of a much higher order.

A third misconception, or, to put it more precisely, a premature and oversimplified conclusion, is that the costs of the new regulation far outweigh the benefits. This conclusion inheres in practically all the scholarship of regulation, and in most of the journalism as well. The problem is not so much that it is an erroneous conclusion as that it mixes together two distinct points, each of them accurate but difficult to express as one. These two points are, first, that that the costs of regulation, though difficult to quantify, are not nearly so difficult to quantify as are the benefits, many of which have no value that can be expressed in market terms.[3] Thus, costs and benefits are sometimes incommensurable and unfortunately cannot be directly compared. The second point is that, given the stated aims of regulatory policy, the methods employed are inappropriate to the ends desired. More efficient means are available, so that the same benefits—however substantial they are—might be achieved at significantly less cost.

D. Changing Perceptions of Regulation

1. BY SCHOLARS

For the student of regulatory history, one of the arresting characteristics of regulation has been the shifting patterns of primacy over the "turf" manifested by different academic disciplines.

Prior to the 1930's, the dominant students of regulation were lawyers and institutional economists who drafted and administered regulatory statutes. One thinks of the first chairman of the ICC, Thomas Cooley, a leading constitutional lawyer whose assertion of turf rights over regulation typified his profession. Within a few years of its inception, the ICC's members dressed themselves in black robes and addressed each other as "brother," after the fashion of the judiciary.

---

[3] A wide'y quoted survey is Murray L. Weidenbaum and Robert DeFina, "The Cost of Federal Regulation of Economic Activity (Washington, D.C.: American Enterprise Institute, May 1978). This analysis is typical in that, as the authors say, "The focus of the study is only on costs; no judgments are expressed on the value of the many regulatory efforts" (p. 3).

The classical economists—the price theorists of the day—ignored the turf, leaving it to the lawyers and to a few of their own economics colleagues of the institutionalist school.

In the next decades, from roughly the 1930's through the 1950's, the leading writers on regulation came from the discipline of political science. They were typically interested in such questions as how the agencies fit into, and perhaps violated, the intricate system of checks and balances so carefully constructed by the Founding Fathers. Such leading political scientists as Robert E. Cushman, Merle Fainsod, and Marver Bernstein focused on the "quasi-legislative, quasi-executive quasi-judicial" nature of the agencies, and either damned the hybrid form as inherently unworkable or praised it as an adaptation essential to the government of complex modern economies.

Beginning in the 1960's, and peaking in the late 1970's, regulatory scholarship has been dominated by neoclassical economists. These scholars, using cost-benefit analysis at every turn, purported to quantify a series of almost unbelievable regulatory inefficiencies. In book after book, article after article, they hammered at regulatory performance, lampooned it, even devoted new journals (e.g., the Journal of Law and Economics and the Bell Journal of Economics) to the exploration of regulatory inefficiency. They changed the vocabulary of discourse, to the point that all who wish to understand the subject—lawyers, political scientists, and others—must become familiar with such formerly arcane terms as "externalities" and the principle of "second-best." Neoclassical economists made the field of regulation their turf in the 1970's, and they show little sign of retreat.

The implications of this latest conquest are two-fold. First, one must ponder whether the conquest is as permanent as it now appears, or whether—in view of the earlier and equally total dominance of the same field by lawyers and political scientists—one is instead witnessing a mere phase in a series of successive triumphs by different intellectual disciplines. The second point, a related one, is more important for the making of policy. Since each discipline has its own axe to grind—part methodological, part ideological—the question arises of how reliable are the prescriptions for reform of whatever intellectual sub-group happens to be king of the mountain at any given time. Lawyers and political scientists, for example, emphasize administrative structure, political feasibility, and due process of law. They hold these values dear, and they expressed them in such laws as the Administrative Procedure Act of 1946. The new price theorists of the 1960's and 1970's, on the other hand, hold the truth of economic efficiency to be self evident, and elevate it above all other considerations. Consequently, their prescriptions—by the nature of their methodology—call for the free play of market forces, and the removal of administrative constraint. In a word, they call for deregulation. Where the ends of regulation are attractive to society, as in the new social and environmental regulation, the neoclassists[4] call for incentive systems to replace the command and control techniques now preferred by public policy.

---

[4] I am oversimplifying a bit here, since by the late 1970's almost all economists agreed on certain aspects of regulatory inefficiency. It is not an exaggeration, however, to say that the neoclassicists led the way.

2. BY THE PUBLIC

At the present time, the American people are almost certain to be confused by headlines saying both that regulation is increasing but that deregulation is on the rise as well. Beyond this obvious irony, public perceptions of regulation are merely a part of the widespread perception that government in general is inefficient, wasteful, and corrupt. Virtually all public opinion polls show that faith in government is declining, along with faith in most other institutions.

What is not so obvious, however, is the continued desire on the part of the public for those services and social goals represented by the new functional and social-environmental regulation. To give one example, a 1978 poll showed the following: [5]

[In percent]

	Too much	Too little or about right
1. The Federal Government is spending	82	17
2. For improving and protecting the environment, it is spending	10	90
3. For improving national health, it is spending	7	93
4. For improving the Nation's educational system, it is spending	11	89
5. For improving the condition of blacks, it is spending	27	73

At first glance, these figures seem to represent conflicting signals. How can Americans be for more Federal expenditures for so many categories (including very important functional regulatory categories such as environmental cleanup and equal employment opportunity), while simultaneously for reduced Federal expenditures? The answer, insofar as there is one, may lie in the public perception that the means to these ends have been inefficient—that the job needs to be done, but that it must be done more efficiently. A poll conducted at about the same time found the following, similarly paradoxical results: [6]

[In percent]

	Yes	No
1. Do you approve of the job Congress is doing?	36	64
2. Do you approve of the job your Congressman is doing?	75	25

II. Why the Regulatory Changes Since 1960 Have Occurred

The date of establishment of almost every regulatory and quasi-regulatory agency in the American Government would be located falls within three periods, all in the 20th century. The first period was the progressive era, which lasted roughly from the start of Theodore Roosevelt's presidency until the end of Woodrow Wilson's, or from 1901 until 1921.

The second period of regulatory hyperactivity began with the inauguration of Franklin D. Roosevelt and ended with the last major piece of New Deal legislation, the Fair Labor Standards Act. This period—the briefest of the three—was from 1933 until 1938.

[5] Everett Carll Ladd, Jr., "What the Voters Really Want," Fortune (Dec. 18, 1978), p. 44.
[6] Ibid., p. 48.

The third period is the one under study in this paper, the period since 1960. The precise beginning of this era is a matter of arbitrary conjecture, of course. Perhaps the election of John F. Kennedy, with his slogan of "Let's get the country moving again," is an appropriate date. Or, maybe the first of the 1960's riots (October 1962, with the commitment of Federal marshals to assist integration of the University of Mississippi) makes a more logical substantive beginning. Or, if the test is the establishment of a new regulatory agency, the creation of EEOC in 1964 would be the date of choice.

Irrespective of precisely when the latest burst of regulatory activity began, this most recent period differs fundamentally from the earlier eras of reform legislation and regulatory initiatives. I will attempt through a brief comparative analysis to show why this has been the case, and then relate the analysis to present trends in American politics that may affect regulatory change in the future.

A. The Progressive Era, 1901–21

A remarkable number of regulatory and other government changes were institutionalized during the first generation of the 20th century, as even a superficial list shows:

1902 The Newlands Reclamation Act (conservation and irrigation).
1903 The first mandatory direct primary system.
1906 The Pure Food and Drugs Act.
1906 The Hepburn Act (which gave the ICC its first genuine power).
1908 The first city manager government (there were hundreds more by 1921, when the movement peaked).
1912 The first minimum wage law.
1913 The Federal Reserve Act.
1913 The 16th Amendment (legalizing a Federal income tax).
1913 The 17th Amendment (mandating direct popular election of U.S. Senators).
1914 The Federal Trade Commission Act.
1920 The 19th Amendment (women's suffrage).

This is only a brief list of the most important of hundreds of Federal, State, and local reforms. It omits such things as the first widespread enactment and use of direct democracy devices: the initiative, referendum, and recall. The list says nothing about the crusading zeal, often inspired by Protestant evangelical religion, which preceded and accompanied the enactment of many of the laws. Nor does the list include such illiberal measures as the 18th Amendment (prohibition), nor the restriction of immigration into the United States, nor the disfranchisement of blacks in many States, all measures sincerely viewed by their proponents as additional "reforms."

What the unadorned list does show, however, is a deep faith in creative tinkering with the procedures and institutions of the American political system. Most of the listed reforms addressed themselves less to the substance of what was happening than to the methods of dealing with rapid change. Even after the reforms, there was little direct involvement by government, and especially the Federal Gov-

ernment, in the everyday lives of American citizens. The size of the Federal establishment was tiny in these years, despite the temporary swelling brought on by American involvement in World War I. What the procedural emphasis of the reforms reflected more than anything else was a disillusionment with corruption and chicanery, a determination somehow to wrest control from the politicos and bring government "back to the people," and a deep concern with the rise of corporate giantism in American life. Hence the first really important antitrust cases in American history: (*Northern Securities*, 1904; *Standard Oil of New Jersey*, 1911; *American Tobacco*, 1911). And hence, too, the Federal Reserve Act of 1913 and the Federal Trade Commission Act of 1914.

1. THE CONTEXT OF PROGRESSIVISM

The entire 20-year period was one of almost uninterrupted prosperity. Virtually nobody paid income taxes, (the threshold of income at which one had to pay income taxes was so high that few "qualified"), that real wages rose steadily, and upward economic mobility was assumed to be available to all. Those who earned new wealth were determined to protect their rising stakes in American life, and those already well off were determined to remain so. There was a certain gentility—many historians call it naivete, especially during the pre-World War I period—in the assumptions of this generation of Americans. In particular, their assumption that tinkering with the mechanisms of government would automatically insure the continuance of democratic control seems in retrospect almost pathetic in its idealism and naivete.

"Welfare" measures were still assumed, by and large, to be the province of private, charitable institutions. There was a great deal of concern about morals, as reflected in the anti-saloon movement and the ultimately successful drive for a constitutional amendment prohibiting the sale of alcoholic beverages. Above all, there was a deep conviction that the growing uncertainties of life, brought about by a generation of heavy immigration, rapid industrialization, and growing urbanization, must be brought under control. The progressive generation had a sincere faith in the efficacy of expertise and of science in the abstract to deal with these problems, and with whatever other problems presented themselves. The coming of World War I did much to dash this mood of national innocence, of course, and the postwar generation became famous for its disillusionment with the hopeful mood of progressivism.

2. 1901–21 COMPARED WITH 1960–79

Both periods were characterized by sustained and serious inflation (during practically the entire progressive era, and for the second half of the more recent period). This hopefulness of the reform movements of the early 20th century was matched by the similar naivete of the "flower children" of the 1960's, though the progressives had great faith in institutions, while the 1960's generation was explicitly anti-institutional. The most striking difference in the two periods—and this is true of the New Deal period as well—is the unprecedented persistence in the 1970's of regulatory initiatives long after the national

reform mood had passed. This fact speaks directly to the changes that have occurred in American politics, and in particular within the party system.

B. The New Deal, 1933–38

The list of reform measures passed during this 5-year New Deal period is familiar, and—considering the brevity of the period—remarkable because the measures were more numerous than those of the progressive era or of the period 1960–79. So numerous are the items that might be listed as major reforms with a regulatory thrust in the New Deal that I will restrict the following list to those measures which established or strengthened agencies that are still active in 1979:

1933 The Agricultural Adjustment Act.
1933 The Tennessee Valley Authority Act (TVA).
1933 The Securities Act.
1933 The Glass-Steagall Banking Act (FDIC) (Federal Deposit Insurance Corporation).
1934 The Securities and Exchange Act (SEC) (Securities and Exchange Commission).
1934 The Communications Act (FCC) (Federal Communications Commission).
1935 The Motor Carrier Act (ICC regulation of trucking) (Interstate Commerce Commission).
1935 The Social Security Act.
1935 The Wagner Act (NLRB) (National Labor Relations Board).
1935 The Public Utility Holding Company Act (New functions for both the Federal Power Commission and the SEC).
1935 The Rural Electrification Act (REA).
1938 The Civil Aeronautics Act (CAB and FAA) (Civil Aeronautics Board and Federal Aviation Administration).

1. THE CONTEXT OF THE NEW DEAL

The central experience of the 1930's, of course, was the Great Depression, which lasted from 1929 until mobilization for World War II in 1941. This circumstance alone undermines analogies between the New Deal and other periods of rapid political change. A second unusual circumstance was the central role of President Franklin D. Roosevelt, one of the most popular politicians in American history. In retrospect, it seems remarkable that any person could have been elected four consecutive times to the presidency, and this fact too suggests that almost any parallels between our own time and the New Deal would be overdrawn.

2. THE LEGACY OF THE NEW DEAL

It is useful, however, to speculate about the meaning of the legacy of this idiosyncratic decade on our own time. For example, the enactment of so much economic regulatory law in the 1930's—in an extremely atypical economic context—should make us less reluctant to repeal those laws that might have seemed appropriate or essential

56-368 O - 80 - 2

during conditions of depression, but which seem now to make little or no sense. Two examples are the Civil Aeronautics Act of 1938 and the Motor Carrier Act of 1935. Each of these laws purported to rescue a particular industry from what was viewed at the time as "chaos." By chaos, contemporary analysts and lawmakers mean, in part, a higher degree of competition in airlines and trucking that seemed desirable during a period of deflation and depression.

Thus, there are powerful lessons to be learned from both the boldness with which the New Deal Congress enacted legislation, and from the profoundly different context of 1979 as compared to the 1930's. One lesson is that a similar boldness may be in order for undoing legislation that was appropriate for one context but is quite inappropriate for another.

C. The Period 1960–79

1. THE BASIC DISSIMILARITY

Aside from the several dissimilarities between the present period on the one hand and the progressive era and New Deal periods on the other, a fundamental dissimilarity is conspicuous with respect to regulation. During neither of the earlier periods was there anything like the two simultaneous movements of our own time toward deregulation and toward more regulation. With minor exceptions, both of the earlier periods were characterized by steady accretions of authority to the government, and steady expansion of the sphere of government. There were practically no backward steps, the single prominent exception being the repeal in the 1930's of the prohibition amendment passed in 1918.

2. THE REASONS FOR THE DISSIMILARITY: "SINGLE-ISSUE POLITICS"

In both the progressive and New Deal eras, strong presidents led the drive toward reform. In the period 1960–79, regulatory legislation often had the sponsorship of strong presidents (particularly Lyndon Johnson), but just as often it emerged without such sponsorship. Also in many cases, new regulation materialized or existing regulation was strengthened during the tenure of weak presidents. This characteristic, too, distinguishes the recent period from the earlier reform eras, and it raises the question of why and how it happened.

Perhaps the simplest answer is contained in still another Washington cliche of recent vintage: single-issue politics. The most striking evidence of this new phenomenon was the success during the 1960's of new coalitions built not around parties but around issues. The adherents to these issues were committed to them with a single-minded zeal that often baffled traditional party politicians and sometimes repelled them.

The spectacular success of single issue politics in issue after issue—civil rights, the women's movement, environmentalism—signaled a critical change in the way American politics was conducted. The lesson seemed to be that the way to achieve one's goal was to assemble a cadre of workers tirelessly devoted to the forwarding of its own programs, to

the exclusion of all others except insofar as such other programs might impinge.

What made the ground exceptionally fertile for single-issue politics was the concurrent decline of party discipline within Congress. Political analysts had long been aware of this danger, because the constitutional separation of powers left the American Government without the discipline built into, say, a parliamentary system such as that of Great Britain or Canada. In parliamentary governments, there is no formal separation between the executive and legislative branches. Thus, whereas the prime minister in a parliamentary government is by definition the leader of his party and a member of parliament, the American President might easily confront hostile opposition majorities in one or both houses of the legislature. This has happened repeatedly during the 20th century, with Presidents Wilson, Hoover, Truman, Eisenhower, Nixon, and Ford finding themselves stymied time and again by recalcitrant majorities of the opposition party.

Once the parties themselves began to fragment, the American Government began to experience periods of aimless drift, like a rudderless ship being tossed about by whatever wave it happened to encounter. The modern fragmentation of the parties began in the late 1930's, as the New Deal coalition of Democrats began to break up over the issue of whether to continue down the road toward reform, or whether to stop and consolidate gains already made. The onset of World War II disguised this incipient breakup of the parties, in two different ways. First, domestic politics was adjourned in favor of exclusive attention to the problem of winning the war. Second, the war produced the noteworthy bipartisan foreign policy designed by such Democrats as James F. Byrnes and Dean Acheson and such Republicans as Wendell Willkie and Arthur Vandenberg.

Furthermore, even though Democrats and Republicans might each experience quarrels within their parties, a new type of discipline emerged in the form of firm coalitions across parties, such as that between southern Democrats and northern Republicans. Though it did not appear so at the time, in retrospect it is clear that this type of coalition, though negative, was a stabilizing force in comparison with what was to come later. The southern Democrat-northern Republican coalition persisted in strength from about 1938 until the early 1960's, when new issues such as civil rights and economic competition between the Sunbelt and the Frostbelt split the coalition into many fragments.

The upshot of these trends was that American politics was in a weakened and vulnerable condition even before it was called upon to deal with two of the greatest crises in American history. These, of course, were the Vietnam war and the Watergate scandals. By about 1974, individual legislators had begun to see that the wisest possible course was one of independence, a political strategy of every person for himself. A surprising number of Senators and Representatives simply quit, in disgust or bewilderment. In particular, several senior legislators—including ousted committee chairmen—retired or resigned, overcome by an acceleration of change that had left them powerless to affect public policy.

With party discipline at a low ebb and with coalitions constantly shifting, power vacuums developed rapidly, and into the vacuums

stepped the multitudinous adherents of single-issue politics. Public interest groups such as Common Cause and the many organizations associated with Ralph Nader had already demonstrated that the techniques of mass action pioneered by the civil rights activists of the 1960's could be carried on with respect to economic issues. The acceleration of inflation, the emerging energy crisis, and the continuing breakup of party discipline within Congress created a confusing situation that could be exploited for the benefit of almost any interest group, left or right, pro or anti, reform or reactionary. If a group could mobilize or appear to mobilize public opinion, and if it could construct a powerful, intelligible, and factually supported case for a particular program, then the odds for success might be good, irrespective of the ideology behind a particular lobbying campaign.

One prominent example of the results of such single-issue lobbying was the rapid onset of regulation in the automobile industry. New and sometimes uncoordinated campaigns for fuel economy, safety, and emissions control all achieved spectacular success in the 1970's, with the net result that the automobile industry, which was virtually free of public regulation in the middle 1960's, become by the late 1970's one of the most pervasively regulated industries in the Nation. Other examples come readily to mind: the campaigns leading to the Occupational Safety and Health Act, and the Consumer Product Safety Commission. One of the most remarkable examples of the swiftness with which single-issue politics could succeed was the sudden reversal in 1978 of a long-established national policy toward mandatory retirement. This turnabout by Congress, which was probably justified but which was pushed through with practically no debate or exploration of the profound consequences, again signified that something new was happening on Capitol Hill. The episode also demonstrated that single-issue politics had the potential of getting the government out of a regulatory posture almost as easily as into it. For adherents of deregulation, this confusing event of 1978 offered possible encouragement, particularly since it preceded by only a few months the airline deregulation bill of 1978.

III. Implications for the Major Players

This concluding section will explore some of the implications of the existing regulatory paradox—that is, the simultaneous and powerful drives toward both more regulation and deregulation. I will take up these implications by suggesting how they relate to the major players involved: first, to the academic analysts of regulation; second, to the Congress and to the executive; third, to the regulators; and finally, to business executives. This section of the paper will be part commentary, part analysis, and part exhortation, and it will be much too brief to do justice to the subject.

A. Academic Analysts of Regulation

That regulation is dominated by specialists in microeconomics is neither a bad sign nor a good one, but rather a significant circumstance that should be recognized and addressed. Like all other academic disciplines, price theory in economics is to some degree a methodology in

search of a problem. In regulation it has found the perfect problem. Price theorists have presented such a thoroughgoing and convincing case against certain types of regulation—particularly price and entry regulation of airlines, trucking, and telecommunications—that they should now turn their energies to a pair of additional tasks too long neglected.

The first task is to quantify not only the costs of regulation but also the benefits. For this assignment, price theorists may be ill suited, since many of the benefits of regulation are not economic but psychological or cultural. There is the important question, for example, of legitimacy and perceptions of legitimacy, which are essential in democratic government. Regulation has often served the function of legitimation well, but legitimation is unquantifiable and therefore beyond the interest or sometimes the ken of the price theorists.

Even though they may be ill-suited for the task of quantifying the benefits of regulation, however, microeconomists should make the attempt, if only to justify their own insistence on the importance of cost-benefit analysis. Documentation of only one side of the cost-benefit relationship results in automatic overkill of the other side. Ironically, such a one-sided approach has diminished the credibility and influence of microeconomic analysis itself. Even so, price theory unquestionably has brought the greatest breakthrough in regulatory scholarship in the last generation.

The second assignment for all academic analysts, including the microtheorists, is to explore more fully and report more intelligibly the efficacy and appropriateness of a broadscale substitution of incentive systems as replacements for the existing command-and-control systems of regulation. Too often, the virtues of incentive systems appear so self-evident to academics that they ignore the administrative problems faced by both politicians and business executives in implementing such systems. The academics therefore devote far too little effort to empirical investigations of exactly how and where incentive systems would work best. And when they do report the results of their findings, they frequently do so in jargon-ridden language intelligible only to their own academic colleagues.

B. The Government

Neither the executive nor legislative branch needs reminding of the urgency of the tasks of regulatory reform. Both the White House and Congress have shown not only a healthy understanding of the relationships between regulation and inflation, but also the relationship between regulation and foreign competition. The final paragraph of the 1979 Economic Report of the President, for example, reads as follows:

> Perhaps the most important contribution the Federal Government can make to improving our trade position is to assure a more sensible regulatory environment. Too frequently, obstacles to production or investment have raised domestic costs or encouraged imports. If agencies are required to take into account the effects on trade and other costs of regulations, greater scope can exist for competitive forces, thereby allowing domestic producers to gain a greater share of domestic and foreign markets.[7]

---

[7] Economic Report of the President, 1979 (Washington, D.C.: Government Printing Office, 1979), p. 162.

It is this type of evidence of a growing awareness of the interrelated nature of regulation, inflation, and competitive economic performance that must continue if the government is to persist on the path of regulatory reform.

That path, however, will not be a smooth one, and it would be a serious mistake to suppose that all the rules and principles that apply to deregulation of airlines, trucking, and telecommunications apply also to cross-industry and environmental regulation of the EEOC and EPA variety. These newer forms of regulation represent consensual agreements among the executive, the legislature, and the electorate that certain goals of American society are worthwhile, and will remain worthwhile, in some cases irrespective of the cost.

C. The Regulators

Individual regulators should recognize above all that they are caught up in a dramatic historical moment. They might begin by explicitly acknowledging what the best of them already know to be the central truth of their situation: That the present regulatory apparatus does not represent a system constructed rationally to deal with present-day problems. Instead it represents the accumulated legacy of the three major periods of regulatory explosion mentioned earlier in this paper—the progressive era, the New Deal, and the period since 1960. In the broadest sense, the system represents an evolutionary merger of two historical processes. The first was the rise of big business in America, and the institutional response to it in the public sector. The second process has been the continual crises that 20th century governments have experienced or have perceived themselves to be experiencing. These forces combined to produce a system of ad hoc, competing governments that often cancel each other or promote mutual growth. The Environmental Protection Agency grows, for example, and thereby forces state and urban governments to increase their staffs of investigators and report writers. HEW pursues affirmative action programs, and in so doing compels universities and other institutions to complete innumerable questionnaires certifying compliance with HEW's guidelines. Sometimes all this activity serves the public interest at reasonable cost, and sometimes not.

For the regulators, a certain boldness and experimental spirit seem most appropriate for the present situation. As experience with deregulation accumulates in such areas as the airline industry, the regulators' justified discomfort with the unknown will abate. As one scholar noted, this process may be an interesting playback of the sequence that led to the present system:

> The early history of the Interstate Commerce Commission (ICC) provides us with an interesting example. Although the ICC was established in 1887, it took until about 1906 for the commission to produce really effective regulations for the railroad industry. Over that 20-year period, the commission built up "social capital"—that is, an atmosphere of public approval that other regulated industries could profit from later. I think we may be going through another historical watershed. If airline deregulation is achieved and is successful, "social capital" may be built up that will facilitate deregulation in other sectors. What the historical example really demonstrates is that change is difficult in either direction.[8]

---

[8] Merton J. Peck, quoted in Paul W. MacAvoy, ed., Unsettled Questions on Regulatory Reform" (Washington, D.C.: American Enterprise Institute, 1978), pp. 5–6.

D. Business Executives

Despite many notable exceptions, on the whole business executives have responded to the growth of regulation in Chicken Little fashion. They denounced practically every regulatory initiative, "stonewalled" for long periods, and complied with new regulations only with great reluctance, if at all. (Again, the automobile industry is an apt example.) Business groups commissioned numerous studies of the deleterious effects of regulation, with nearly every such study concluding—with a deceptive quantitative precision—that indeed the effects have been at least as deleterious as the sponsor believed, and often more so. The returns from such studies have been disappointing, on the whole.

More insightful business managers took a different approach. Recognizing the essentially transitional nature of the present situation, they searched for a modus vivendi that would draw on the best advice in both government and business, and work out through negotiation a compromise between the imperatives of regulation on the one hand and the realities of the balance sheet on the other. Foregoing their natural instinct to strike an adversary posture, they patiently added to their own corporate staffs a number of middle-level executives conversant with such matters as environmental impacts, occupational health and safety, and equal employment opportunity. Sometimes firms were forced to take these steps. More often than is commonly recognized, however, they took them voluntarily.

The result, whether forcibly or voluntarily initiated, was a substantial growth—especially within large manufacturing companies—of legions of executives expert in fields such as water quality, air standards, and pollution abatement. One can see in such industries as chemicals and pulp and paper the rapid progress of an attitude characterized less by stonewalling than by arm's-length negotiating, and in numerous instances by active cooperation between the firm and the relevant agency in setting realistic standards of compliance.

In the end, the public interest will be better served by this voluntaristic, cooperative approach than by litigation and other adversary proceedings, in which nobody can really win more than a transient victory. Furthermore, only through the cooperative approach can the incentive system so admired by academic analysts be given a fair trial and be compared with the command and control of regulation methods now in place.

THE POLITICAL ECONOMY OF WAGE-PRICE POLICIES: AN HISTORICAL REVIEW

By Arnold R. Weber*

CONTENTS

	Page
The emergence of wage and price policies	18
The rising plateau of wage-price policies	22
Designing wage-price policies	23
Coverage	27
Organizational arrangements	30
The problem of consensus	32
The development of standards	33
The use of sanctions	36
Models of wage-price policy implementation	39
The decoy model	39
The defensive model	40
The offensive model	41
The regulatory model	41
Whither wage-price policies?	42
The future of wage-price policies	44

Since the end of World War II, the United States has carried out a continuous flirtation with wage-price policies. In the Eisenhower administration, the courtship was cool and distant. In other instances, as during the Nixon administration, there was a deep infatuation that ended in estrangement. Despite the loss of innocence, President Carter has persisted in efforts to influence wages and prices through direct government intervention. Although wage-price policies lack the respectability of other economic measures, they still constitute a recognized policy alternative in the management of the national economy.

This analysis reviews the experience with various forms of wage-price policies instituted during "peace time" in the United States snce 1946. It does not attempt a direct assessment of the effectiveness of government intervention on wage and price movements. Rather, the analysis focuses on the political and economic context in which such policies have been adopted and the objectives they serve, the design of the programs, and the problems of implementation. In a broad sense, this review is concerned with the "political economy" of wage-price policies as revealed by the U.S. experience.

THE EMERGENCE OF WAGE AND PRICE POLICIES

The emergence of government policies to influence specific wage and price movements reflects a variety of factors. First, the intellectual triumph of Keynesian economics and its legislative expression,

*Provost and professor of economics and public policy, Carnegie-Mellon University. This article is a revision and update of an analysis initially published in Crauford D. Goodwin, *Exhortation and Controls: The Search for a Wage-Price Policy 1945-71*, Brookings Institution, 1975.

the Employment Act of 1946, reinforced governmental concern over inflation as a side effect of the treatment for stagnation and unemployment. During the 1960's the appeal of wage-price policies (or incomes policy) was strengthened by the widespread acceptance of the so-called Phillips curve. Derived from an analysis of historical data, particularly in Great Britain, the Phillips curve purportedly describes the relationship between levels of unemployment and the rate of change in wages and/or prices; other factors held constant, the lower the unemployment rate. the greater the increase in wages and prices.

Once the Nation became committed to full employment as a national goal, some method had to be found to cope with the inflation that so often was the concomitant of success. Wage and price policies, therefore, emerged as an attractive solution to this dilemma. Full employment would be achieved by manipulating fiscal and monetary aggregates, while price stability would be preserved by influencing key wage and price decisions in the overall structure of the economy. Thus, wage and price policies offered the missing piece in the puzzle whereby the country could simultaneously attain full employment and price stability.

This quest for full employment with price stability became the special responsibility of the President. As a consequence of the New Deal and the Employment Act of 1946, the President was now regarded as Chief Economist of the United States. The establishment of the Council of Economic Advisers institutionalized this role and gave it a strong executive character. It is not surprising that wage-price policies have been associated more with executive action than with legislative measures. Presidents Kennedy and Johnson, in particular, established an activist pattern that has strongly conditioned the expectations concerning the behavior of future chief executives. The Full Employment and Balanced Growth Act of 1979 (Humphrey-Hawkins) reconfirmed and expanded the role of the executive in national economic affairs.

If acceptance of the goal of full employment set the framework for wage and price policies, then *wars* gave these policies a high priority, the quality of familiarity and the aura of success. By diverting resources from civilian production, wars inevitably have created strong inflationary pressures while suffusing the economy with the glow of full employment. Theoretically, these pressures on prices could be contained by aggressive fiscal policies, but the political response usually was inadequate to the economic requirements and therefore other forms of government intervention were utilized. In this manner, wage-price policies (i.e., direct controls) had a high degree of acceptance at the conclusion of World War II. Direct controls, together with rationing, limited price increases to a tolerable level and acquired a coloration of patriotism.

Once full employment was validated as the dominant goal of postwar economic policy, it was inevitable that controls would be viewed as an appropriate device to deal with the inflationary perils of reconversion. When Congress blocked peacetime controls, President Truman converted the issue into a major political controversy that contributed to his upset victory in the 1948 presidential election. Although concern over inflation had subsided by 1948–49, wage-price policies regained high visibility with the onset of the Korean War. Again, in

Truman's battles with Congress over an effective stabilization program, the President elevated wage-price policies to a matter of national virtue.

Under the Johnson administration, the cumulative weight of the Vietnam war made wage-price poilcies a centerpiece of the administration's overall economic strategy. The use of "guideposts" and "jawboning" had started under President Kennedy as part of an effort to restore economic growth and full employment as a self-contained set of goals. During the Johnson administration, these policy instruments were transformed into defenses against inflation in an economy that was subject to increasing strains from the Vietnam war. When President Johnson rejected the advice of his economists to seek a tax increase in 1966, wage-price policies constituted the most appealing of the new weapons left in the arsenal.

The Nixon administration initially attempted to avoid policies aimed at direct intervention in wage and price decisions. Instead, the administration adopted a strategy of "gradualism" whereby monetary and fiscal restraint ideally would slow the inflation rate without precipitating large-scale unemployment. Despite these sanguine expectations, the economic momentum generated by the Vietnam war combined with the consequences of "gradualism" to create a politically untenable combination of high unemployment and rising prices during 1970–71. For this reason and other factors, the Nixon administration executed its historic volte face and imposed wage and price policies in their most draconian form with the freeze of August 15, 1971.

Thus, aside from the technical difficulties of achieving full employment and price stability, wars have generated the wide oscillations in economic performance that have made wage and price policies an appealing, if not irresistible, policy choice. Moreover, during the Vietnam period, the distinction between "wartime" and "peacetime" was blurred so that wage and price policies gave tacit recognition of the economic consequences of a war to which the country was unwilling to make a national commitment as a political matter.

The growth of trade unions and collective bargaining has been a less dramatic but equally significant factor in the evolution of wage-price policies in the United States. The effective application of wage-price policies is based on the assumption that many labor and product markets are characterized by noncompetitive conditions. If all wages and prices were promptly and fully determined by market forces, then presumably governmental intervention would have a counterproductive effect and result in distortions that ultimately would raise wages and prices. This line of reasoning, of course, is at the heart of the free market criticism of wage-price policies. However, if monopoly elements are present in the economy, then the appropriate form of "incomes policy" is the vigorous application of the antitrust laws. Wage-price policies as such then became only short-term supplements to long-term structural remedies.

This resort to the antitrust laws generally is inappropriate or infeasible in dealing with union "power." With the passage of the Wagner Act and the articulation of the "exclusive bargaining agent" concept, union organization became a protected activity and the success-

ful unions were ceded a significant degree of monopoly power. Hence, wage policies always have had a particular appeal in the United States as a device for dealing with the problems of economic stabilization arising from the lawful exercise of monopoly power through collective bargaining. Indeed, in the United States wage policy has always been the pivot of active efforts by the government to influence wages and prices. It is significant that throughout the Kennedy and Johnson administrations actions to restrict price behavior in the large oligopolistic industries usually began with preliminary efforts to obtain "responsible" wage settlements in collective bargaining.

With the acceptance of monopoly power in the labor market, wage policies have had to be fashioned and applied with a certain element of finesse, if not deviousness. On the one hand, various actions have been taken to induce unions to exercise their power responsibly—with price restraint as the quid pro quo. On the other hand, efforts have been made to identify egregious instances of monopoly power which appear to go beyond the pale defined by the Wagner and Taft-Hartley Acts. In this respect, the construction unions have been a favorite target. Collective bargaining in the construction industry received special attention from the Johnson administration and was the subject of a separate control effort during the Nixon administration. Under Nixon, the government responded to the "excessive" exercise of union power in construction by suspending the David-Bacon Act whereby the government itself has acted to reinforce the unions' power derived from the Wagner Act. Direct wage controls were applied in March 1971, a full five months before the comprehensive freeze was instituted. With the return of wage moderation to the construction industry, the Teamsters' union became the *bete noir* of collective bargaining. The most highly publicized instance of government pressure on private wage-price decisions during the Carter administration involved efforts to influence the outcome of the trucking industry negotiations in 1979.

Last, the quest for effective wage-price policies has been given impetus by the increased American sensitivity to the international economy. By the early 1960's, the economic relationships between the United States and the other Western industrialized nations changed significantly so that the U.S. balance of payments came under increasing pressure. Aside from threatening the position of the dollar, this deterioration of the U.S. balance of payments implied that other Western countries were able to "export" their unemployment to the United States. Indeed, in one of the earliest representations to President-elect Kennedy, W. W. Rostow called for wage and price policies in order to permit the administration to fulfill its goals for leadership in international affairs and to avert "stop-go" policies necessary to manage the balance of payments.

By 1971, continued adherence to fixed exchange rates, the devaluation of foreign currencies, and the massive transformation that had taken place in the world economy had made the U.S. balance of payments problem more acute and set the stage for the dramatic exercise of wage-price policies by President Nixon. The proximate cause of the New Economic Policy unveiled on August 15, 1971 was the sharp deterioration of the U.S. balance of payments and the heavy

pressure on the dollar in international money markets. Under these circumstances, it is not surprising that Secretary of the Treasury Connally, whose range of action was not limited by the same sense of tradition as many of his predecessors, was instrumental in persuading President Nixon to link the institution of wage-price controls to the closing of the gold window and the devaluation of the dollar. Wage-price controls would help the United States to preserve any trade advantage arising from the devaluation while also crimping the possible inflationary effect of this action. In the absence of direct controls, domestic prices were expected to rise because of the higher prices of foreign imports and the greater latitude these increases would give to domestic producers to raise the price of similar goods.

International monetary developments were also decisive in moving the Carter administration to adopt a formal system of wage-price guidelines in October 1978. For several months, the dollar had come under increasing pressure in foreign money markets because of persistent concern over the balance of payments deficit and the domestic inflation rate in the United States. Because the earlier program of price and wage "deceleration" had failed to allay these concerns, President Carter enunciated explicit wage and price standards backed by the threat of denial of government contracts as the penalty for noncompliance. As in 1971, the institution of formal wage-price policies was also associated with dramatic shifts in monetary policy, including a sharp increase in the discount rate and the mobilization of resources to defend the dollar.

The Rising Plateau of Wage-Price Policies

One of the striking consequences of the series of experiments with wage-price policies since World War II is the extent to which the framework for the discussion of these policies has been altered. For example, a mild passage in the President's Economic Report of 1957 calling for "shared responsibility" between government and business for price stability precipitated a sharp controversy that seems like an echo of innocence compared to the discussion of wage-price policies in the 1970's. Where President Eisenhower's call for "discipline" in 1958 appeared to be heretical, it is a cliche in 1979. Where President Eisenhower ignored demands from Senator Kefauver to force the steel companies to reduce their price increases in 1958, one of the first acts of President Ford was to casually criticize General Motors for a price increase and to express only passing appreciation when the company acceded to his request for a token price cut. President Carter, in turn, showed his resolve in fighting inflation by personally—and successfully—exhorting Sears, Roebuck & Company to roll back its catalogue prices.

Throughout the 1950's the debate over wage-price policies was expressed largely in ideological terms. In the 1900's, the discussion of wage-price policies focused on the pragmatic question of whether they would work rather than whether they would pose a threat to the Republic. The 34-month experience with direct wage and price controls during the Nixon administration focused the public discussion of wage-price policies more on tactics and design than on burning questions of principle. The most compelling argument levied

against comprehensive, mandatory controls during the Carter administration has not been that they violate the canons of the free market, but more directly that the 1971–73 experience demonstrated that they don't work and may have a counterproductive effect.

Designing Wage-Price Policies

The analysis of experiences with wage-price policies during the post-World War II period presents a kaleidoscope of slogans, organizations, heroes, and villains. If there is any common thread to the record of the last 33 years, it is that wage-price policy is a plastic concept whose shape is defined by economic circumstances and immediate political requirements rather than by any well-defined theories or system of implementation.

The plasticity of wage-price policies is most vividly revealed in the variety of objectives that they have been designed to serve. Aside from general strictures about the need for wage-price policies to help vanquish inflation, the objectives of wage-price policy never have been clearly defined and, more generally, have been overwhelmed by the tactical requirements of the moment.

First, wage-price policies sometimes have served a neopopulist objective by providing a forum for the government to confront powerful economic aggregations; i.e., monopolies who allegedly are unwilling to act in the public interest. This objective was expressed in the Truman administration's formulation of wage-price policies. Truman asserted that wage-price controls were necessary in the immediate postwar period to prevent business from reaping excessive profits during the reconversion period. Although this initial objective was transformed into a broader goal of preventing inflation to avoid a subsequent economic collapse, the populist stirrings were readily revealed by political events. Thus, rhetoric easily overwhelmed any sense of economic design during the election campaign of 1948. In a flurry of speeches, Truman reminded the electorate that he had urged that price controls be retained until production caught up with demand. The Republican Congress, controlled by special interests, had been unwilling to take this step that was so clearly in the public interest. In characteristic terms, Truman stated: [1]

> They (the Congress) have decided that the National Association of Manufacturers and the National Chamber of Commerce of the United States know all about prices and price controls. Well, now, we have price controls and rationing now, just as we have under government controls, only those price controls are controls so that only the man that has the money is able to get the necessities of life.

Wage-price policies also served as a platform for broadside attacks against "power" in the Kennedy administration. The most notorious exercise of presidential machismo came during Kennedy's controversy with the steel industry. When U.S. Steel raised prices following the 1962 labor negotiations, apparently in contravention of a pledge to the President, Kennedy launched a furious verbal assault that would have done justice to Theodore Roosevelt. The price hike was rescinded when Inland Steel was persuaded not to follow U.S. Steel's lead.

---

[1] As cited by Crauford Goodwin and R. Stanley Herron, "The Truman Administration: Problems and Policies Unfold," in Goodwin, op. cit. p. 55.

Even in the Nixon administration, wage-price policies served as a basis for identifying, if not stigmatizing, a particular set of villains. In this case, the Nixon administration cast the mauve light on the construction unions, which, in its judgment, were price conveyers of the inflationary virus. President Carter, an avowed populist, has substituted the oil companies for the steel industry as a prime target for political attacks on economic power.

The use of wage-price policy for such populist objectives essentially serves a political rather than an economic purpose. In macro-economic terms, it is misleading to believe that the assaults aimed at particular villains can serve a useful purpose. There is the hope that taking an individual company or union to task will set an example and restrain other businesses and unions in their wage and price decisions, but in most instances the impact will be ephemeral in view of the complexity and sheer magnitude of the U.S. economy. Moreover, wage-price policies which spring from populist ardor have limited staying power and are easily blunted by other political considerations. For example, the Kennedy administration's victory over the steel industry in 1962 was a Pyrrhic one; the hostility in the business community was so intense after "The Battle of the Running Blough" (after Roger Blough of U.S. Steel) that Kennedy had no stomach for another confrontation when the industry raised its prices again the next year.

The use of wage-price policies as a springboard for general attacks on economic power also means that the choicest targets are more likely to be business than organized labor. This reaction creates serious problems of inequity and, in addition, skirts a key problem of wage-price policy. To the extent that unions, in fact, contribute to inflationary pressures, this power should be dealt with directly in the framework of government policies, rather than relying on indirect and sporadic efforts to stiffen employers' resistance to union wage demands. In 1979 the wage guidelines were exceeded in the collective bargaining agreements reached with United Airlines and in the rubber and automobile industries. These settlements reflected the unwillingness of the unions to accept government-imposed restraints. The government issued a formal finding of "noncompliance" against the companies but this stricture had no impact at all on the unions involved. In the auto case, the union did reduce the size of the economic settlement with Chrysler. However, this adjustment was designed to strengthen the company's petition for financial aid from the Federal Government rather than to comply with the guidelines.

Second, wage-price policies have been used as a defensive adjunct to other expansionist policies. This objective was expressed during the Kennedy administration. Wage-price policies were held out to President Kennedy by his advisers to vouchsafe that a tax cut and other expansionist measures would not result in inflation. Ultimately, this counsel was accepted, particularly after the successful experiences in jawboning the steel and automobile industries persuaded an uneasy President that he would not run an undue economic—or political—risk in opting for growth and employment.

Similarly, the Nixon administration use wage-price policies to underwrite an aggressive expansionist policy. In addition to difficulties with the balance of payments, the major political problem from the Administration's point of view was the failure of "gradualism"

to achieve an acceptable trade-off between unemployment and price stability. Indeed, during 1970–71 the Administration appeared to have the worst of both worlds with high unemployment and rapidly rising prices. The defeats suffered by the Republicans during the election of 1970 precipitated a chorus of demands for action and the carefully orchestrated policy of "gradualism" was abandoned in favor of growth and employment. As in the Kennedy administration, wage-price controls held a temporary attraction because considerable excess capacity existed in the economy and it was believed possible to have controls without grossly impairing the operation of the market system. Moreover, public confidence in the Administration's economic policy had plummeted so that it was necessary to take dramatic steps to demonstrate that the President would be "tough" on inflation while stimulating a new surge of prosperity.

For this reason, the wage-price freeze became highly attractive and, beyond the arcane details of devaluation, symbolized the Administration's resolve to set things right. One consequence of the freeze, however, was to lock the Administration into a more rigorous and comprehensive program of controls than was necessary or desired by Nixon's top economic advisers. This system persisted through Phase II, but almost from the moment that Phase II was instituted the Administration initiated a strategy of decontrol that culminated in Phases III and IV. The second freeze in 1973 was a perverse political effort to restore confidence in a system that the Administration itself had undermined.

Third, wage-price policies have been used as a substitute for fiscal and monetary policies. One of the important arguments made in selling wage-price policies to President Kennedy was that a distinction should be made between demand-pull and cost-push inflation. Where there was excess capacity in the economy, inflationary tendencies were symptomatic of cost-push factors which could be restrained through wage-price policies. On the other hand, demand-pull inflation required broader fiscal and monetary measures. This text was adhered to through the Kennedy administration and in the early stages of the Johnson administration. But when strong inflationary pressures developed with the enlargement of the Vietnam war, it became obvious that strong fiscal measures were necessary. For political reasons associated with the conduct of the war, Johnson rejected the recommendations for a tax increase. Accordingly, the Administration had little recourse other than to broaden the application of wage-price policies.

A similar sequence of events took place during the Carter administration. In the first 18 months of his administration, President Carter aggressively pursued policies designed to reduce the level of unemployment. These efforts were marked by considerable success but resulted in a steady rise in the inflation rate. Because excess capacity ostensibly was present in the economy in early 1978, cost-push factors were identified as the dominant inflationary factor. Rather than reversing gears on fiscal and monetary matters, the Carter administration floated its first, tentative form of wage-price policy organized around the vague principle of "deceleration." This limited foray was calculated to prevent what was considered to be transient price increases from being reflected in wages and other factor prices that would sustain cost-push inflation.

A similar sequence of policy adjustments took place in late 1978. To deal with inflation and calm the international money markets, the guidelines program was strengthened significantly by the formulation of explicit wage-price standards and the threat of sanctions for noncompliance. Some tentative adjustments were made in monetary policy, but credible action was not taken on this front until one year later (October 1979) when it was clear that inflation was out of control.

Fourth, if wage-price policies sometimes have been used as a substitute for fiscal and monetary measures, they also have been brandished in an effort to permit fiscal and monetary measures to work by deflecting pressures for more direct government intervention. This paradoxical, if not cunning, approach was taken during the Eisenhower administration. To the extent that the Eisenhower administration had any wage-price policy at all, it was vague and hortatory. When the President included sermonettes on the need for restraint in his Economic Report, they were taken as evidence of his willingness to enter more strenuously into the battle against inflation. In fact, Eisenhower intended nothing of the kind. The only major case of direct government intervention undertaken by President Eisenhower involved the prolonged strike in the steel industry in 1959. In this instance the main issue was not wages, but work rules. The intervention was carried out by Vice President Nixon, who was showing his wares preliminary to the 1960 Presidential election campaign.

The Nixon administration also went through an extended period in which pseudo wage-price policies were formulated to divert pressure for government intervention in private decisionmaking. If "gradualism" was to have an opportunity to work, then demands for action had to be accommodated, at least superficially. Such a feint was executed in the spring of 1970 when the administration initiated a series of "inflation alerts" and established the Commission on Productivity. The "inflation alerts" were a series of ex post admonitions to business or unions which quickly subsided to inaudible levels. The Commission on Productivity, on the other hand, tried to focus the attention of the Nation on more fundamental factors affecting prices and wages.

A similar approach was adopted in the early stages of the Ford administration. Within the first month after he took office, President Ford forced a token rollback of prices in the automobile industry, established an agency to monitor wages and prices without attempting to clarify its mandate, strongly disclaimed any desire to impose mandatory wage-price controls on the grounds that they don't work, and kept the center of gravity of his policies on monetary and fiscal measures designed to grind down inflation over a longer haul.

There is, of course, a great disposition to criticize the lack of precision in defining the objectives of wage-price policies. But this criticism is beside the point. The overriding generalization is that only rarely have these policies been intended to control wage and price movements in any rigorous or comprehensive sense. Even in those cases that have involved the favored targets of such policies—the steel and automobile industries—it was not obvious whether the government was courting the industry or controlling it. When Nixon took the plunge during Phases I–II, there was a systematic effort to control wages and prices, but the initial exultation soon turned to a flush of

embarrassment and the half-hearted program of Phases III–IV. In fact, wage-price policies have been employed primarily as a tactic in the general political maneuvering over inflation rather than a program with explicit objectives that have been sustained in their own right.

Coverage

The coverage of wage-price policies has reflected the variety and imprecision of the program objectives. Indeed, the dominant features of the coverage of wage-price policies have been selectivity and uncertainty. The only times that an explicit delineation of coverage was made were during the Nixon administration's exercise in direct wage-price controls and as part of President Carter's second-generation guidelines program. In the Nixon case, the need for a clear demarcation was dictated as much by the legal requirements of the program as by the objectives and administrative methods. While the Carter guidelines were formally characterized as "voluntary" in nature, the use of government procurement authority as a sanction against noncompliance during the period October 1978 to October 1979 required an explicit delineation of actual coverage. With these exceptions, wage-price policies have been applied more or less on a "reserve power" concept whereby the executive branch selects its target to support the program objective that is appropriate at the time.

If there is any generalization to be derived from the experience of the past 33 years, it is that the broader the ostensible coverage of the wage-price policies, the more narrowly these policies were applied. This paradox is perhaps best explained by relating coverage to the catalogue of policy objectives. Thus, the neopopulist objective of wage-price policies enunciated during the Truman administration presumably included any monopoly or center of economic power that was able to exact a toll from the economy; i.e., members of the U.S. Chamber of Commerce and the National Association of Manufacturers. The significant fact, however, is how infrequently this highly activist and sometimes bellicose President attempted to translate this objective into action in specific cases. The only significant case arising during the Truman administration was the seizure of the steel industry in 1951 and this incident was not so much a direct attack on prices as to force the steel companies to accept a wage settlement and to avert a strike during wartime.

When the objective of wage-price policies has been to provide a protective mantle for expansionist policies, the coverage of the policies has been shaped by tactical considerations. Thus, in the Kennedy administration, wage-price policies were largely brought to bear in highly visible situations where it could be demonstrated that they were an effective defense against the exercise of market power. The requirements for this demonstration of muscle flexing were two-fold: That the industry be "basic" in the sense that price movements in that industry had wider reverberations throughout the economy, and that it be oligopolistic so that the wage-price policies could effectively be applied on an ad hoc basis through executive daring. The industry had an additional attraction when it also involved a large union so that there could be parallel interventions for wages and prices. From the outset,

Rostow and other Kennedy advisers had expressed the desirability of arranging "wage-price treaties" whereby the union would agree to a moderate wage settlement in return for a pledge that the companies with whom it bargained would respond with similar restraint. The concept of a "wage treaty" was to emerge many years later in more grandiloquent terms as a "social compact" or "national accord." [2] In 1979 President Carter unsuccessfully pressed for "real wage insurance"—a variant of the social contract concept whereby wage increases would be kept within the limits of the government's guidelines in return for a statutory pledge to reduce income tax rates in the event that the annual increase in the Consumer Price Index exceeded a designated level.

For these reasons, the most frequent targets of the policies historically have been the steel and automobile industries. When the Kennedy administration appeared to enjoy some success in moderating wage and price movements in these industries, various task forces moved onto other metal industries. In these cases, the wage agreements did not quite meet the technical requirements for price stability, but they were close enough to the current guidelines to justify demands that the companies absorb the increases in wage costs without seeking relief through price adjustments. On the other hand, the Kennedy administration generally failed to induce price decreases, consistent with the current theory of wage-price policies, in cases where productivity increases exceeded the national average.

Why the union leaders played this game is not clear. However, the implementation of such accords was most effective where the union leaders were personally swayed by Presidential blandishments or saw broader social goals in such cooperation. In any case, when efforts were made to control wages independent of price decisions in the same industry, they generally were unsuccessful. In 1966, Secretary Wirtz made a limited attempt to restrain union wage demands in the construction industry. The government girded for a battle with Local 825 of the Operating Engineers and came out second best when the union achieved its demands through a series of arbitration awards and supplementary arrangements. A more important setback was incurred in the case involving the International Association of Machinists and the air transport industry. Although air fares presumably were firmly under the thumb of the Civil Aeronautics Board, the Machinists' union was adamant in its demands and finally settled for a contract with United Airlines calling for a 6-percent wage increase, far in excess of the current version of the guidelines. (Ironically, the settlement reached between the IAM and United Airlines in 1979 was also stigmatized as a violation of the wage guidelines by the Council on Wage and Price Stability.)

The only situation in which efforts to control wages independent of prices were marked by success involved the Construction Industry's

---

[2] The term social compact is somewhat ambiguous. In the United Kingdom it has been used to describe a general agreement between the government and the national labor federation under which the trade unions will moderate otherwise justifiable wage demands in return for a government commitment to reduce the erosion of real earnings and to expand various social programs. In the United States, the AFL–CIO and the Carter administration in late 1978 negotiated a National Accord. The document covered general economic policy objectives but did not include a pledge of wage restraint. The federation did otherwise agree to support the proposed Pay Advisory Committee designed to buttress the wage guidelines program.

Stabilization Committee (CISC) during the Nixon period. For three years, the construction unions actively cooperated with a special wage stabilization effort. The circumstances involved were unique and cannot be readily duplicated in other situations. The unions perceived an active threat from nonunion competition which had established new beachheads in the industrial and commercial construction sectors. The suspension of Davis Bacon in March of 1971 had helped to overcome the union's initial reluctance to participate. When CISC was established, the suspension was rescinded. The effort was engineered by John Dunlop, then a professor at Harvard University (and later Secretary of Labor under President Ford) who called upon three decades of experience and personal relationships in the industry.

When wage-price policies have served as a supplement to or substitute for fiscal and monetary measures there has been some effort to establish comprehensive coverage. After President Johnson had rejected pleas for a tax increase, various task forces were organized to develop an "early warning system" to identify and influence price movements in specific industries. Joseph Califano established this special monitoring group and made direct government representations in more than 40 cases. Bereft of statutory power over wages and prices, the Johnson administration tried to cope with the tide of rising prices engendered by the Vietnam war by broadly expanding the scope of its informal activities. It is noteworthy that following the United-IAM case, minimal efforts were made to influence union wage demands directly. In fact, many of the wage agreements negotiated by the major unions during this period left their members in a relatively favorable position with respect to real wages despite the increases in the price level that took place during 1969–71.[3]

The most extensive coverage of wage-price policies was, of course, associated with the period of direct controls from August 1971 to April 1974. This experiment began with a universal wage-price freeze and was trimmed back over three years by a series of exceptions and formal acts of decontrol. To a large extent this experience can be viewed as an aberration. Because the wage-price freeze was part of a broad-gauged economic policy change, some short-term, dramatic measures were desirable. Moreover, the fact that statutory authority to impose direct controls had been enacted by the Congress to harass the President made such a step feasible. In view of the euphoria associated with the freeze, the President could not immediately trim back the program even though there were severe misgivings within the Administration over so comprehensive and rigorous an effort. In fact, Phase II represented a subtle effort to sharpen the coverage of controls by adopting the three-tier system. Controls would be rigorously enforced only in the first two tiers involving firms with more than $50 million in sales on the price side, and those with 1,000 or more employees on the wage side. Once the election of 1972 was passed, a major step toward de facto decontrol was taken with the initiation of Phase III in January 1972. From then on, the administrators steadily retreated to a smaller perimeter and by April 1974 only about one-quarter of the economy was under the jurisdiction of the Cost of Living Council.

---

[3] Marvin Kosters, Kenneth Fedor and Albert Eckstein, "Collective Bargaining Settlements and the Wage Structure," Labor Law Journal, vol. 24 (August 1973), pp. 517–25.

A similar progression took place during the Carter administration. In 1977, a comprehensive and somewhat vague program of wage-price "deceleration" was launched under the aegis of the Council on Wage and Price Stability. The effort was initiated in the context of an economic expansion encouraged by the administration to whittle down unemployment. This program languished and was replaced by the guidelines in October 1978 when inflation and pressure on the dollar caused increased concern in the U.S. and abroad. As in the Nixon days, the program was declared to be universal in coverage but in practice was focused on a few very large cases.

Organizational Arrangements

Although considerable attention has been given to the organizational arrangements appropriate to the maintenance of wage-price policies, this issue really has been subsidiary to the general problem of devising an effective program to control inflation. For example, a great deal of energy was expended in the last stages of the Johnson administration in developing organizational options for the implementation of wage-price policies. While these issues were not trivial, it is apparent that they masked the growing frustration of the President's economic advisers over Johnson's reluctance to adopt the appropriate fiscal measures.

Nonetheless, some significant distinctions may be indicated in the nature of the organizational arrangements used to formulate and implement wage-price policies. Again, these distinctions are linked to the objectives of the program and its coverage. Where wage-price policies remain at the level of exhortation, organizational arrangements generally are limited to policy-oriented groups which pull together various combinations of top officials in the executive branch. In virtually every administration since World War II there has been a chameleon-like progression of Cabinet Committees and special Task Forces that are concerned with price stability, economic growth, and other euphemisms designed to elevate a concern over price and wage developments. The composition of these Cabinet Committees has varied from case to case, but usually they have included the Secretary of the Treasury and the Secretaries of the major constituent-oriented departments, such as Commerce, Labor, and Agriculture. In each case, the Chairman of the Council of Economic Advisers (CEA) has played a major role. Under the Nixon administration, the designation of the Secretary of the Treasury as the primary economic spokesman for the Administration somewhat reduced the influence of the CEA. This pattern was maintained by President Carter. However, to show further resolve in the battle against inflation, Carter also appointed Alfred Kahn, former Chairman of the Civil Aeronautics Board, as his special anti-inflation assistant.

As wage-price policies become more explicit and their implementation more aggressive, supportive organizational arrangements become necessary simply to carry the administrative burden. During the latter stages of the Johnson administration, the Cabinet-level policy mechanisms were supplemented by an Interagency Price Committee (The Nelson Committee) to identify problem areas on the price side

and to coordinate the strategy for negotiating with management. The actual negotiations apparently were carried out by John Douglas, an Assistant Attorney General, and members of the CEA. After some months, the level of involvement was raised to involve Joseph Califano, the powerful Assistant to President Johnson. Where the organizational arrangements are informal and lack statutory authority, the level of executive involvement tends to escalate as the deference inspired by lower echelon officials diminishes.

Beginning with the Nixon years, wage-price policies have been associated with formal organizational structures. Because the Nixon program operated from a legal base and was comprehensive at the outset, the usual array of special task forces would not suffice. The wage-price freeze was administered by the Cost of Living Council, (CLC) a Cabinet-level group. Acting through its full-time Director, the Council commandeered the staffs of existing agencies, particularly the Office of Emergency Preparedness and the Internal Revenue Service. These arrangements were elaborated during Phase II to include a separate, all-public Price Commission and a tripartite (labor-business-the public) Pay Board. The latter two were independent agencies set up to insure the "fairness" of the control system. CLC, however, continued to direct the field structure which was centralized in a special unit of the IRS. The Pay Board and Price Commission were abolished at the beginning of Phase III to improve coordination, but also so that the Administration could control the application of the wage-price policies and the orderly dismantling of the program.

The CLC was laid to rest when its statutory authority expired in April 1974. The agency was resurrected later in the year, however, as the Council on Wage and Price Stability (COWPS). COWPS provided the organizational bona fides of President Ford's commitment to price stability. Although it has had a checkered career, the agency has continued in existence since 1974. During the Ford administration, COWPS largely served as a public scold, reacting to ostensible inflationary actions taken by both the government and the private sector. The agency was the obvious candidate to administer the Carter "deceleration" and guidelines programs and was given expanded resources (but not statutory authority) to do the job. Throughout its existence, COWPS has been within the Executive Office of the President, bespeaking presidential concern over its mandate.

One additional organizational strategy should be noted. Under Eisenhower, decoy organizations were set up to demonstrate the government's concern over inflation while abjuring active involvement in wage-price decisions. Thus, the Cabinet Committee on Price Stability was established in 1958 to monitor and review price developments. The Committee was chaired by Vice President Nixon and aside from Allan Wallis, the executive director, did not have any staff. In a steady stream of press releases and speeches by Wallis, the Committee waved a red flag with great energy so that the demands for a more active policy abated. Similarly, as noted above, one of the first actions of President Ford was to ask Congress to establish the Council on Wage and Price Stability. The legal authorization was probably superfluous, but it dramatized the new President's involvement and, by enlisting formal congressional approval, muted demands for the reinstitution of direct controls.

The Problem of Consensus

While the various organizational arrangements have been closely tailored to the objectives and administrative requirements of the specific wage-price policies, they have skirted what is the most vexing problem of such efforts—the development of mechanisms to define a consensus acceptable to the major interest groups in the economy. In a political democracy, wage-price policies cannot be effective unless they are accepted, tacitly or otherwise, by the major groups that will be affected by the policies. Even when the program is based on statutory authority, as in the Nixon administration, its viability must stem from acceptance, if not support, rather than coercion. This is true for the simple reason that in a political democracy the management of the economy cannot be carried out by throwing union leaders and businessmen in jail. When wage-price policies remain in the realm of general exhortations and few specific actions, the question of consensus is not important. But when jawboning or legal authority are used to restrict the latitude of the parties and efforts are made to control their decision directly, then the question of defining and preserving a consensus becomes critical.

In this respect, all of the postwar administrations have been singularly ineffective over time. The usual device for attempting to define a consensus has been some variant of a labor-management committee. The members of the committee generally have been drawn from the ranks of organized labor—and big labor at that—and the top executives of large firms. These labor-management committees were formed in the early days of the Kennedy administration and again during Phase III of the Nixon program. A similar function was carried out during the Ford administration by the Labor-Management Group, an informal private body with government blessing. The Group brought together representatives of big business and big labor and was chaired by the ubiquitous Professor Dunlop. It continued to meet after the election of President Carter. Earlier, the Truman administration had adopted a parliamentary approach by convening a large-scale, labor-management conference shortly after the end of the war.

In every case, these devices failed to create an acceptable framework for wage-price policies. Despite intensive cultivation by the CEA, the Kennedy Advisory Committee on Labor-Management Relations never could agree on operating principles. An equally inauspicious record was made by the Labor-Management Advisory Committee during Phases III and IV. The major pronouncement of this committee was a resolution calling for the end of controls. The later Labor-Management Group stopped meeting as a consequence of the AFL–CIO resentment over business opposition to the Labor Law Reform bill. In October 1979 the Carter administration made another, more formal effort to establish a consensus-forming mechanism when it created the Pay Advisory Committee.

When the Nixon administration initiated comprehensive, statutory controls, the program designers resorted to the traditional tripartitism to deal with the question of consensus, at least with respect to the regulation of wages. The Pay Board was formed as a tripartite body and operated on this basis for five months. It provided representation

for only a narrow sector of the economy and for individuals who represented a limited constituency, or no constituency at all, and failed to grapple with the question of a national consensus. Rather, the Board sought to mitigate the opposition of powerful groups by permitting them to participate in the decisionmaking process even though this participation might disadvantage other groups in the society. This objective was only partially achieved and after five months four of the five labor representatives resigned as an expression of their dissatisfaction.

The Carter administration has been equally unsuccessful in achieving a consensus on wage-price policies, notwithstanding the initial congeniality between the new President and organized labor. The AFL–CIO generally chose to ignore the deceleration program and has actively opposed the guidelines both at the bargaining table and in the courts. While attacking the constitutionality of the use of government procurement authority to enforce the guidelines, the AFL–CIO has called for direct, mandatory controls on prices, wages, and other forms of income. In contrast, the guidelines did gain the support of segments of the business community, notably General Motors. To bolster a revised version of the guidelines, Secretary of the Treasury G. William Miller initiated direct negotiations with high labor officials. These discussions resulted in the affirmation of a National Accord concerning a wide range of economic policies. A side agreement to the Accord provided for the establishment of the Pay Advisory Committee. In form and mission, the Committee showed many of the genetic characteristics of the Nixon Pay Board. The new Committee was chaired by the ubiquitous John Dunlop.

The Development of Standards

The general objective of wage-price policies is, of course, to control inflation. However, this lofty goal must be cast into operational terms if the policies are to have any influence in specific cases. When wage-price policies have been limited to broad appeals for restraint, as in the Eisenhower and Ford administrations, little effort has been made to develop a meaningful standard. Indeed, in view of the dog-in-the-manager strategy associated with such policies, there is no need for more precise standards.

The heavy reliance on the productivity standard has strongly shaped the application of wage-price policies. The productivity criterion tends to link changes in the price level to wage movements. Presumably, if wage increases are limited to the long-term trend in national productivity, the Nation will be able to attain stable prices for the economy as a whole. Although this approach is generally credible, it may suffer in application to specific cases. To the extent that wage-price policies are aimed at developments in particular firms and industries, and to the extent that price behavior is a function of many factors other than productivity, it is not clear that putting all the money on this one horse is the best strategy. If the ultimate purpose of wage-price policies is to thwart "excessive" price increases, the operational standard should permit the administrators to move directly against prices rather than to set a standard that makes "responsible" wage changes a precondition

for intervention. The "market power" of a firm or industry may be—and frequently is—quite independent of its wage policies. It is significant to note that as wage-price policies took shape during the Kennedy administration almost every intervention was precipitated by the onset of collective bargaining negotiations. The strategy was clear: Persuade the union to limit its demands to the range of national productivity and then use this occurrence to force the companies to minimize their price increases.

This approach probably has had several consequences. First, it focuses on collective bargaining as a causal factor of inflation even though the real culprit may be overly expansive fiscal and monetary policies or exogenous factors such as cartelized oil prices. Second, it narrows the range of analysis of "acceptable" price behavior. An equally effective procedure for dealing with prices might be to examine a broad range of economic data including profits, rate of return on investment, costs, and excess capacity. Third, the use of productivity as the dominant criterion for determining acceptable wage and price movements sidesteps the problem of the distribution of income between labor and capital. Although the policy administrators hope to be neutral, the effects of the policy may be to shift income from labor to other groups. The application of the productivity criterion requires that labor subordinate its historical aspirations for a redistribution of income. Whether the position of organized labor is correct or incorrect is unimportant; the critical factor is that it requires labor to deny the validity of an overriding ideological goal. For this reason, organized labor strenuously fought for, and achieved, special exemptions from the wage standard for low income workers during both the Korean war episode and Phase II of the Nixon administration Stabilization Program.

The problems of using productivity as the intellectual basis for wage-price policies are rendered more acute when a specific numerical figure is applied. President Kennedy endorsed the guidepost-productivity concept and an arithmetical range, but he did not agree to a specific number. President Johnson's economists, however, were emboldened to measure the size of the guidepost, were determined that 3.2 percent was the defensible point estimate, and set it as a standard for responsible wage behavior. The fine print prepared by the CEA recognized that 3.2 percent might not be applicable in many cases. But the news media inevitably projected the 3.2 standard without equivocation and it was used as the yardstick to measure the effectiveness of the program. If the government "succeeded" in its efforts in the steel and automobile industries by limiting wage increases to the productivity standard, then it abjectly failed in dealing with the IAM.

Significantly, once the guidepost was shattered in the IAM dispute and the administrators were relieved of the anxiety of measuring every case against the 3.2 standard, the system of policy implementation appeared to be widely transformed. Using the Nelson Committee and the Califano-Robeson task force, the emphasis shifted to prices directly. On the basis of a variety of data, decisions were reached on acceptable price behavior in negotiations between management and government officials. Under this approach, the tactical relationships between wage

and price policies were reversed. If price increases were limited by government pressure, then employer resistance would be stiffened to union demands above the wage standard.

Quite a different process took place during the Nixon administration. Here, the administration did not gradually immerse itself in controls, but dove in with little preparation. Once the administration had determined to establish independent price and wage regulating agencies, it was confronted with the problem of whether to define a goal to guide these agencies. The initial judgment was to reject an explicit goal or standard. This judgment was based on a respect for the complexity of factors determining wage-price movements and short-term political considerations; to the extent that a specific goal was established, it would be easier to identify the failures of the administration, especially in an election year. It was ultimately determined that some goal was necessary to provide a measure of discipline for the program.

These conflicting considerations were accommodated by establishing a goal that, in effect, had a 50-percent variance. The Cost of Living Council ordained that the objective of Phase II was to bring the rate of price increases down to a range of 2 to 3 percent by the end of 1972. After a process of bargaining within the Pay Board, this translated into a wage standard of 5.5 percent, providing for a 3 percent increase for productivity and a 2.5 percent price increase—which was halfway between the 2 to 3 percent goal established by CLC. The standards established a general framework for evaluation but individual decisions were made on a much wider range of factors. On the price side, the greatest emphasis was given the concept of "allowable" costs, with the profit margin rule establishing a secondary defense. Similarly, the 5.5-percent wage standard was subject to various modifications in the light of equity and "established bargaining relationships."

The Carter administration demonstrated both a capacity for innovation and a respect for past practices in fashioning the standards for its wage-price policies. The initial standard implemented in early 1978 called for the general deceleration of wage and price increases. Businesses (and unions) were asked to reduce the rate of increase of wages and prices by 0.5 percent below the level that had been experienced during the previous year. The articulation of a general but highly relative standard permitted the definition of a goal without getting into the messy task of determining the applicability of a specific numerical standard in individual cases. If the various economic decisionmakers would all slow down a little, the country could move closer to price stability.

The deceleration standard proved to be too ingenuous for the second, more rigorous exercise in wage-price policies initiated by the Carter administration in October 1978. If employers who failed to comply with the program were to be denied government contracts and employees who did comply were to be protected by Real Wage Insurance, then the standard had to be developed in more precise, numerical terms. Thus, an arbitrary (but then reasonable) price goal of 6 percent was established for business. This price goal translated into a wage standard of 7 by adding the trend rate of national productivity gains (1.5 percent) to the price target and subtracting .5 percent to reflect a scheduled increase in Social Security contributions. Both the

wage and price standards were subject to elaborate exceptions, but the rules governing prices proved to be more flexible by permitting cost pass-throughs and the alternate use of a profit margin test for compliance. In any case, the productivity concept was once again used as the keystone in the architecture of the wage-price policies.

The lesson of these experiences is compelling. Although productivity may have theoretical credence as a standard at the level of the general economy and helps to define the goals of wage-price policies over time, it has limited usefulness in the administration of these policies. The wider the system of wage-price policies and the more "serious" the efforts to implement these policies, the more likely that productivity will have reduced relevance. This was the experience during both the Korean war, the later phases of the Nixon program, and with the wage-price policies of the Carter administration.

One modest attempt has been made to use a wage standard that is not linked directly to productivity trends. During the Nixon period, the Construction Industry Stabilization Committee eschewed any fixed standard and based its decisions on an administrative theory of wage relativities. That is, the wage adjustment permitted for a given craft group in a particular labor market was founded on an identification of the historical wage relationship among the relevant crafts (and unions) in the construction industry. The objective of this approach was to slow down the overall rate of construction wage increases while preserving a stable wage structure. By inference, "excessive" wage increases developed when one craft union registered gains which disturbed the existing structure and provoked similar, outsized increases by other craft groups. This approach enjoyed some success in slowing dampening construction wage trends in the period 1971–74. However, the conceptual and administrative difficulties of applying a theory of wage relativities to the economy as a whole would be staggering. In addition, such a framework does not afford a basis for linking the wage and price standards in a logical and consistent manner.

The Use of Sanctions

The implementation of wage-price policies requires that they be accompanied by methods of enforcement appropriate to the ends that the policies presume to serve. When wage-price policies are restricted to global statement of restraint, the methods of enforcement are equally tentative. Those unions and businessmen who act in an "excessive" manner will be subject to the scorn of their peers, a flurry of critical editorials, and other manifestations of public disfavor. These measures are likely to be passing shots without lasting effects on wages and price movements. Because the definition of wage-price policies at such a general level normally serves political rather than economic purposes, the situation should not be viewed as evidence of these policies' intrinsic deficiencies. For example, President Eisenhower was scarcely concerned that his sotto voce pleas for discipline were not heeded. Even during the Truman administration, there was suspicion that President Truman was experiencing such pleasure in flagellating corporations and the Congress that he would have been disappointed if, say, General Motors had seen the errors of its ways and announced that it was foregoing all price increases for the next year.

The rules of the game change significantly when efforts are made to apply wage-price policies in specific cases. In this context, the decisive factor is whether there is some statutory authority. With the excep- war), wage-price policies have been administered in an extra-legal tion of the stabilization program under Nixon (and during the Korean framework. (The Carter administration elected to use the government's procurement authority to explore the wage-price guidelines. This approach involved the exercise of statutory powers established for purposes other than directly regulating wages and prices.) Therefore, it has been necessary to marshall an array of ad hoc weapons that reflect the current strengths of government and the vulnerabilities of the parties which are the focus of the policies. Indeed, to some extent the choice of targets under an informal system of controls is determined almost as much by the vulnerability of the parties to government action as for rational economic reasons. In this respect, firms in oligopolistic industries have been the customary targets of wage-price policies and one of the favorite weapons in bringing about compliance has been the threat of prosecution under the antitrust laws. President Kennedy pulled this card out of the deck almost as an instinctive reaction during the great controversy with the steel industry in 1962.

The government, of course, can use other tactics to bring about compliance by businessmen. In an aluminum case, Alcoa revoked a price increase following high-level negotiations with Secretary of Defense McNamara concerning the disposition of the government's stockpile of strategic materials. Obviously, the disposition of the stockpile could have a major effect on prices. Beyond this measure, the government has a wide range of subtle, or not so subtle, weapons including the application of the tax laws, decisions with respect to imports and tariffs, and the provision or withholding of direct financial aid.

A legal system of enforcement was established during the Nixon administration. The regulations of the Price Commission and the Pay Board were duly published in the Federal Register and noncompliance was subject to fines, injunctions, and criminal penalties. Despite the availability of these sanctions, it is important to note that any system of enforcement has to be based on a selective strategy. In the case of Phases II–IV, the so-called three-tier system created degrees of probability of legal prosecution. The most vulnerable were those firms and unions in the first tier which required prior approval of price and wage increases; units in the second tier had periodic reporting requirements, and units in the third tier generally were left with their consciences and the theoretical possibility of a visit from an IRS agent. The system of enforcement was designed to cage what Lloyd Ulman has called the "rogue elephants" while assuming that the rest of the herd would behave decorously because of fear, self-interest, and a concern for the national welfare.[4]

The Carter administration program lacked the specific statutory authorization of the Nixon effort but was administered as if it were a bona fide legal system of regulation. COWPS issued regulations, established reporting requirements for large firms, made findings of "compliance" and "noncompliance," and laid out appeals procedures

---

[4] Lloyd Ulman and Robert J. Flanigan, "Wage Restraint: A Study of Income Policies in Western Europe," University of California Press, 1971, pp. 240–243.

for malefactors. However, this administrative system had a fanciful quality because of COWPS' unwillingness or inability to apply stringent sanctions. In three major cases, COWPS declared the parties were in noncompliance. These involved price decisions by Hess Oil, and the labor agreements between United Airlines and the IAM and the United Rubber Workers and the Big Five rubber companies. Shortly after Hess' fall from grace, it was awarded a contract to provide fuel to the Defense Department after agreeing to take compensatory actions. In addition, the use of government purchasing authority engendered little concern on the part of the unions involved. Although strike pressure from the URW resulted in a guideline-breaking agreement, the companies bore the brunt of the mild sanction, requiring the offset of the "excess" wage increment from the prices of specified goods.

Aside from assuring the integrity of the program, the greatest importance of the system of sanctions is the extent to which it helps to maintain a position of evenhandedness in dealing with labor and management. In fact, there appears to be a marked difference in the effectiveness of sanctions for labor and management depending on whether an informal or formal system of wage-price policies is involved. In an informal system, such as that which prevailed during the Kennedy-Johnson period, business is more vulnerable to the application of sanctions than is organized labor. In fact, most of the Presidents' big guns were wheeled out against business in the form of threats of antitrust prosecution, stockpile disposition and other forms of persuasion. In addition, because business historically has been the target of criticism in American society and must continue to interact with "the public" through the market system, it has shown a greater sensitivity than labor to expressions of disapproval from the President. During the Johnson administration one of the more plaintive incidents occurred when Joseph Block, the chief executive of Inland Steel, complained to President Johnson of the pressures that had been brought to bear when, in 1964, the company had raised the price of galvanized sheets and coils. Block said:[5]

> Last December my company raised its price on a relatively minor steel product—galvanized sheets. We did this because the cost of the coating material—zinc—had gone up considerably and we regarded our profit as inadequate. It was in reality a minor matter, yet based on the government reaction, one would have thought we had dropped an atomic bomb. We were told that we might trigger inflation. We were told that we might induce Mr. Abel here to increase their wage demands. We were told that we would lose business to foreign steel. . . . All this would seem to indicate that we must have been pretty dumb not to consider such matters in advance of our action. And perhaps we were. . . ."

In contrast, the government strategy toward unions in the application of informal or extrastatutory wage-price policies has been more honeyed and less effective. Apparently, the normal approach was to bring the union leader in question (usually David McDonald of the Steelworkers and Walter Reuther of the United Autoworkers) to the White House where they were offered deep draughts of presidential flattery and confidential chats on the need for responsibility. When

---

[5] James L. Cochrane. "The Johnson Administration: Moral Suasion Goes to War," in Goodwin, op. cit., p. 217 (footnote 47).

push came to shove, however, as in the Operating Engineers and IAM-airlines negotiations, there appeared to be little that the government could do to make intransigent union leaders cooperate. Under informal controls, it has been easier to bring the President of U.S. Steel to bay than a business agent in New Jersey.

In contrast, the system of enforcement under a formal, legally based program probably is more effective against labor than it is against business. Especially when there is some slack in the labor market, there is a high degree of assurance that individual managers will try to apply the wage standards with fidelity. All the incentives arising from economic self-interest press in this direction. And to the extent that one employer adheres to the standard in dealing with his employees, he will make it easier for other employers by reducing competitive pressures to move wages above the standard. During Phases I–IV, George Meany, President of the AFL–CIO, complained with some credence that there were four million enforcement agents (employers) on the wage side but only a handful of IRS agents on the price side. Similar complaints were raised during the Carter administration. Although many large unions ignored the wage guidelines without penalty, most employers in nonunion situations appeared to generally conform to the wage standard, invoking the call of patriotism or the prospect of loss of government business as the justification for their position.

The effect of this asymmetry in the use of compliance was powerfully illustrated in determining the fate of wage-price policies under the various administrations. For example, it was organized labor that kicked over the traces during the Johnson administration as the Consumer Price Index rose rapidly in 1966. In contrast, the demise of the Nixon stabilization program was largely brought about by the inability of the administrative machinery to contain price increases. The Carter guidelines program has undergone similar stresses resulting in at least two, equally unsuccessful, modifications.

Models of Wage-Price Policy Implementation

From this welter of experience, it is possible to identify four models of wage-price policies. These models link together objectives, elements of coverage, organizational arrangements, the nature of the standard, and sanctions. It cannot be said that the models are "rational" in the sense that they represent some efficient combination of variables to achieve the objective of controlling inflation. They do have a retrospective logic, however, to the extent that they describe a consistent relationship between objectives, organization and the other elements of wage-price policies.

The Decoy Model

The first model may be characterized as the "decoy model." Wage-price policies are articulated to divert political pressure for government action rather than as part of a serious effort to influence wage and price behavior. Under this approach, the objective of the program is stated in the most general terms so that they apply to every economic unit while having relevance for none. There is comprehensive

coverage in principle, but no effective coverage in fact except when some egregious incident takes place. The standard is formulated in the broadest possible terms and may embrace slogans such as "responsibility," "discipline," or "restraint." Normally, a Cabinet Committee will commemorate the administration's dedication to price stability, although in some instances special organizations will be established to "monitor and review" price and wage developments. Sanctions are limited to exhortation and expressions of concern. To a large degree, the "decoy" model is an anti-wage-price policy model.

The decoy model was first unveiled during the Eisenhower administration, was further refined in the early part of the Nixon administration, and was revived by President Ford. In each case, the President and his economic advisers were reluctant to go beyond expressions of concern and calls for responsibility. No specific targets for government intervention or criteria for such actions were identified. Well-publicized organizations were established in the form of the Cabinet Committee on Price Stability, the Commission on Productivity, and the Council on Wage and Price Stability. There was no hint of pressure other than presidential browlifting.

The Defensive Model

The second combination of elements may be characterized as the "defensive model." Here, the general objective is to provide a reserve capability to intervene in particular wage and price decisions that are highly visible or which are identified as engendering inflationary pressures. The defensive model is especially suited to those circumstances when wage-price policies are utilized in support of expansionist policies. Coverage is universal in principle; but, in fact, the policies are applicable primarily to large economic units in basic industries. A criterion for responsible behavior is defined but left in imprecise form so that the executive may retain discretion in determining when to intervene or not to intervene. In implementing the program, the President will call upon the existing agencies of government whose efforts will be coordinated on an ad hoc basis by some central unit, normally the Council of Economic Advisers. The sanctions employed may encompass the full range of government influence and authority, such as threats of prosecution under the antitrust laws, stockpile disposal, etc.

The defensive model of wage-price policies was employed during the Kennedy administration, the early stage of the Carter administration, and with a few modifications, properly characterized Phases III and IV during the Nixon administration. The Nixon administration case is especially interesting because here a formal regulatory system was turned into a defensive program by executive action. Although the stabilization program was founded on statutory authority, Phase III was unveiled as an exercise in self-regulation. Firms and unions were expected to conform to the existing regulations, but the "stick in the closet" was substituted for an assertive enforcement procedure. To the extent that there was tension between the underlying philosophy of Phases III–IV and the statutory framework within which the program was administered, efforts were made to resolve the tension by pursuing an accelerated schedule of decontrol.

The Offensive Model

The third model of wage-price policies is the "offensive model." In this case, the policies are viewed as part of a serious effort to contain inflation even though they are not part of a legally based system. Coverage is broadly defined and the policies applied to sectors with a significant impact on wage and price levels. Special administrative units are organized to link the process of review to the exercise of influence and sanctions. An effort is made to define a standard in more precise terms so that it establishes a trip point for government action. In effect, the executive branch becomes committed to act when decisions in designated sectors of the economy exceed the standard. This "offensive" approach was employed during the Johnson administration when wage-price policies came to serve as a substitute for fiscal and monetary policies.

The Regulatory Model

The most robust variety of wage-price policy is associated with the "regulatory model." Under this arrangement, there is an organized effort, supported by statutory authority, to directly regulate wage and price movements. Because it is founded on law, the model does not depend on positive intervention by the government. Instead, there is an expectation that there will be general compliance with the published standards and regulations. Coverage tends to be broad, and formal adjustments may be made from time to time in response to changing economic circumstances. The program is administered by self-contained agencies that usually enjoy an independent status. Conventional legal sanctions for compliance such as fines, injunctions, and criminal penalties are available. Historically, the regulatory model is most closely associated with wartime situations.

This model was used in peacetime during Phases I–II of the Nixon administration to deal with the consequences of extraordinary events and policies that transformed the position of the United States in the world economy. Specific objectives, if any, were to provide the protection for a high stimulative policy and to suppress inflationary pressures as they built up in the economy. When the true test of the system of controls came at the beginning of 1973, the administration retreated to the "defensive model" before many shots had been fired.

Although it does not strictly meet all the requirements, the Carter program of wage-price guidelines is most appropriately classified as an example of the regulatory model. In effect, President Carter created his own legal basis for the program by brandishing the government's procurement authority as a sanction for noncompliance. Moreover, COWPS effectively introduced all the legal paraphernalia of the regulatory model including reporting requirements, appeals procedures, and the codification of standards in the Federal Register. Despite sharp criticisms by labor and business, the program withstood legal challenge in the Federal courts.

These models, of course, are not prescriptive in nature. Rather, they describe a configuration of responses that has characterized past efforts at the implementation of wage-price policies. The primary impression is one of improvisation, with the blending of short-term economic re-

quirements and political pressures, and resort to jerry-built organizational arrangements. If the purpose of wage-price policy has been to regulate the flow of economic events in the economy, it has been more analogous to sandbag embankments thrown up to resist flash floods than to dams which try to regulate systematically the flow of wage and price decisions.

Whither Wage-Price Policies?

Wage-price policies in the United States have now undergone 33 years of trial and error and it is fair to say that there has been as much of the latter as the former. Even without a systematic (let alone econometric) assessment of their effectiveness, several deficiencies have been cast in sharp relief by this review.

First, there has been a consistent failure to develop arrangements for defining a national consensus concerning the objectives and rules of the game governing wage-price policies. The preferred technique has been some variant of a labor-management committee. On more venturesome occasions, broadly constituted, one-time conferences have been convened bringing together representatives of diverse interest groups. The 1974 National Conference on Inflation was perhaps the most sophisticated of these endeavors. Unfortunately, there is no evidence that these advisory committees or pseudo-parliamentary conclaves achieved anything approaching a durable consensus. Indeed, the record indicates that they have served more to underscore differences in interest than to create a common framework for the harmonization of such interests. Nor can there be any optimism concerning the ability to define such a consensus in the foreseeable future. Where the economy is organized on market principles of self-interest and these principles are extended into the political process, there is little prospect that there will be an agreement on the operational goals of wage-price policies other than in wartime or when there is a universal concern over impending catastrophe. In addition, unlike many Western European countries, the United States does not have the broadly representative economic organizations that can strike a comprehensive consensus. Organized labor represents only about 20 percent of the labor force while business representation is fragmented in a variety of organizations from the Business Roundtable to the Chamber of Commerce. Also, it is unlikely that other vocal interest groups such as consumerists and environmentalists would stand idly by while business, labor, and the government engaged in closed negotiations over wage-price guidelines.

Second, wage-price policies have suffered from an inability to achieve an evenhanded treatment of wages and prices. The imbalance has not been a matter of conscious design but has been a consequence of the particular administrative arrangements for the implementation of wage-price policies and intrinsic differences in the mechanisms for controlling wages and prices. It is true, as George Meany has repeatedly stated, that the employer is often a willing enforcer of wage standards but that the control of prices must be left to the less than fully effective actions of government.

As noted previously, wages have been most severely restrained under formal systems of regulation and prices have borne the brunt of government actions under informal programs. The political conse-

quences of this asymmetry have been magnified by the fact that wage-price policies usually have been cast in national terms although no one realistically expects them to be applicable to all situations. Thus, each instance of differential treatment comes to be viewed as evidence of class oppression rather than as an effort to deal with economic power or market deficiencies in particular cases. Differential treatment may be accepted when it is related to surgical efforts to deal with problems in individual industries, but not when it is viewed as part of some national scheme for shifting power relations among economic groups.

Third, the quest for a stand that is comprehensive, equitable, and sufficiently precise for effective administration has been less than successful. The productivity concept serves a useful purpose in establishing the goals of wage-price policies, but it affords only limited guidance for their attainment and, indeed, may be mischievous in individual cases. By fastening on productivity as the dominant standard, price restraint has frequently been linked to wage restraint, although the relationship between these occurrences should not be determinate. Aside from problems of measurement, the relevance of productivity as a sensible basis for wage and price decisions in the short run is diminished as you move from the economy as a whole to untidy market for specific goods and categories of labor. To a large extent, productivity has been an attractive operational standard for wage-price policies because of its convenience in casual systems of administration rather than its applicability in specific cases.

Fourth, there has never been a sensible theory of coverage of wage-price policies. Presumably, wage-price policies emerged as an attractive alternative, because they could bridge the gap between the macroeconomic policies that would sustain high levels of growth and employment for the economy as a whole and the microeconomic wage-price decisions in particular cases. These linkages have never been carefully identified. Consequently, wage-price policy administrators have turned almost reflexively to the same set of industries. If the steel industry didn't exist, it probably would have been invented for the convenience of the Chairman of the Council of Economic Advisers. In this manner, wage-price policies invariably have been brought to bear on steel and autos (and now oil), although the health services and food distribution industries may now have a more consequential effect on the general price level.

Fifth, efforts to coordinate wage-price policies with fiscal and monetary measures have seldom been explicit or successful. The combination of the tax cut and wage-price guidelines did appear to have a salutary effect during the Kennedy administration but under Johnson, Nixon and Carter incomes policies generally were substitutes for fiscal and monetary restraint. This element of dissimulation was clearly perceived in October 1979 when the latest iteration of wage-price policies failed to calm international money markets and the Federal Reserve was forced to administer harsh measures to dampen the increase in credit and the money supply.

Last, the organizational arrangements for the implementation of wage-price policies have had all the continuity of a pick-up volleyball team. For the greater part, wage-price policies have been administered by a cast of thousands drawn from different agencies at different times.

56-368 O - 80 - 4

Because the objectives, coverage and legal authority associated with wage-price policies have never been clearly established on a continuing basis, the organizational arrangements have had a consistent quality of improvisation. This has been true even in the case of the Council on Wage and Price Stability which has been in existence for five years.

The Future of Wage-Price Policies

Where does this analysis leave us and what does it imply for the future of wage-price policies? At the outset, it should be recognized that wage-price policies in one form or another will be a continuing option in the management of the economy. For better or worse, such policies are perceived as a least-worst alternative for dealing with inflation without incurring the widespread unemployment implicit in the application of restrictive fiscal and monetary policies. Within a highly politicized arena for economic policymaking, assertions that direct government intervention in wage and price decisions have not and cannot deal effectively with inflation, elicit counterarguments that aggressive administration, or "presidential leadership," or a "social contract" between labor, management, and the government will remedy past deficiencies.

In addition, if experience in other Western industrialized nations is any guide, the increased vulnerability of the U.S. economy to international developments is more likely to increase rather than diminish the appeal of wage-price policies. In both 1971 and 1978, wage-price policies were instituted as part of a broad strategy to cope with problems of the dollar in international money markets. Also, a contemporary twist to the changing theory of incomes policy is that restraint is necessary to prevent increases in the price level arising from exogenous factors, such as OPEC, from being translated into wage increases that in turn will generate additional pressure on prices.

The post-World War II record clearly confirms the more frequent resort to wage-price policies over time. Following the strident jawboning of the Truman administration, no serious effort was made to institute explicit wage-price policies until 1962. The Kennedy-Johnson guidelines program was more or less continued in effect until 1966. However, in the 1970's, direct wage-price controls, including two freezes, were maintained from August 1971 until April 1974. Voluntary guidelines were revived by President Carter in 1977 and were strengthened in 1978 and 1979. Despite increased criticism and even scorn, it is almost certain that the program will be with us at least into the first year of the 1980's. Moreover, there is little evidence that the public has become widely disenchanted with wage-price policies. Although there has been a general political reaction against government regulation, a Gallup poll in the summer of 1979 revealed that 58 percent of all respondents and 48 percent of those with college educations favored wage and price controls as an antidote for inflation.

If it is the case that some form of government intervention in wage and price decisions will recur in the future, the operational question is what approach promises to be most constructive (or least destructive) in light of the postwar experience. First, policymakers should avoid the compulsion to fashion some comprehensive consensus or "social contract" to sustain wage-price policies. If such a consensus is to

be forthcoming in the United States during peace time, recent history indicates that it is unlikely to be the product of advisory committees, summit conferences, or tripartite bodies. Such devices may be used to facilitate prior consultation with the major interest groups, but the task of defining a consensus is properly the obligation of the Congress and the executive. In 1967, procedures for developing a consensus on guideposts were proposed whereby Congress would review and approve or disapprove the President's statement of the guideposts for any year. This approach holds some promise and has been implicitly used during the Carter administration when Congress renewed the authorization and budget for COWPS after the guidelines program had been initiated. The political system and not some artificial assemblage of economic interest groups must bear the burden for establishing a consensus, however fragile it may be. The somewhat romantic notion of divining the popular will must be tempered by the fact that organized labor, for example, has ultimately balked at every effort to create or preserve a working consensus on wage-price policies.

Second, it should be recognized that wage-price policies cannot be administered effectively on a global basis. The government may promulgate a general standard for wage and price behavior which is applicable to the economy at-large on a voluntary basis. But the arsenal of weapons to induce compliance should be applied only selectively in particular product and labor market situations. Many of the deficiencies of wage-price policies can and should be dealt with by trimming back the scope of administrative efforts so that they are less likely to become enmeshed in broad social and political conflicts or fruitless attempts to restrain prices that are determined by forces out of reach of the regulators.

Within the more discriminating framework, selective policies will have two broad functions. On the wage side, they would be concerned primarily with preventing distortions in the national wage structure, rather than attempting to control the general level of wage increases directly. If wage movements have an autonomous inflationary effect, it is usually manifested through structural distortions as one union attempts to leapfrog another in its wage demands or as nonunion employers strive to maintain "traditional wage differentials" with the unionized sector. Major upward shifts in the general level of money wages usually are more symptomatic of inflationary pressures that already suffuse the economy than the independent exercise or augmentation of "union power" in some comprehensive sense. Wage policies are not likely to be an effective barrier against broadside inflationary forces, but they can help to promote the effective adjustment of intra- and inter-industry wage structures to rapidly changing economic circumstances. A wage "target" may be established for the economy as a whole; however, it should be flexibly applied in the context of particular market and collective bargaining situations. If Phase II and the Pay Board had any salutary effect, it was to preside over the restabilization of the national wage structure in the wake of dislocations induced by outsized agreements in the construction, transportation, and retail food industries.

On the price side, administrative efforts would be concentrated on situations in which competition is limited by the organization of the market (as in health services), where a firm or industry can exploit

temporary imbalances in supply and demand to reap economic rents (as in energy) and large, oligopolistic industries. Continued attention would also be given to those industries in which prices are strongly influenced by government actions and policies. Governmental price policies should not be viewed as an adequate substitute for the vigorous enforcement of the antitrust laws.

The selective approach also would help to relieve the sense of inequity arising from the inability to afford evenhanded treatment of labor and business within a global framework. To be sure, specific interests will feel either advantaged or disadvantaged by the imposition of selective wage-price policies. But as a tactical matter and political matter, it is easier to deal with special interests than with class interests. Experiences with the development of selective wage policies in the food distribution and construction industries in 1972-73 indicate that they can be maintained without arousing hostilities. In addition, by narrowing the focus of wage-price policies to specific industries, full weight can be given to all relevant economic data rather than fixing on a single criterion such as productivity. This framework would also deemphasize the political necessity of balancing restraints on wages and prices in the same industry even though the economic forces influencing the two income shares are different.

Third, the task of developing a framework for wage-price policies and implementing them in specific cases should be given to a permanent wage-price commission. In the past, such a commission has been viewed with distaste. However, the record of wage-price policies in six administrations indicate that these bodies will always be present in some form, either as a Cabinet committee, a subterranean interagency task force or as some ad hoc unit. It is best to recognize that efforts to influence wage and price decisions will be a durable element in national economic policymaking and establish a commission on a continuing basis. A permanent commission will have a higher degree of public accountability for its actions, an accountability that has been blurred or ignored in past exerises of wage-price policies. Last, to the extent that expertise counts, a permanent wage-price commission would be an institutional depository for expertise in devising and administering these programs. Otherwise, each crisis precipitates a frantic search for the few tired bureaucrats who were involved "the last time around."

To a limited extent, COWPS has assumed the role of a continuing wage-price commission. It is significant that COWPS became the lineal descendent of the Cost of Living Council five months after the latter agency was quietly laid to rest. Thus, a wage-price agency has been on the scene more or less continuously since 1971. But COWPS has always been viewed as a temporary bureaucratic contrivance that must justify its existence on a year-to-year basis. This attitude has been reflected in staff and budget support.

The wage-price commission would have the authority to review wage and price developments in individual industries and to develop procedures for fact finding and public hearings. Consideration may also be given to authorizing the commission to initiate manatory controls in specific cases. This aproach has been taken on an ad hoc basis in the past in construction and the retail food on the wage side and, of

course, in the oil industry on the price side. The Carter administration's hospital cost containment proposal is a variant of selective controls. Also, the wage-price commission can (and has) acquire(d) a broader mandate by turning its attention to government actions, particularly in the area of regulation, which contributes to inflation.

It may be argued that if the President has such authority, the political pressures to exercise it on a broad scale would be irresistible. This is especially likely if Congress can badger the President without any involvement in the decisions. This problem can be resolved by giving the President the right to impose selective controls subject to approval by Congress within 15–30 days through the process of negative legislation; i.e., if Congress does not act in the prescribed time, the authority goes into effect. Also, where the authority exists on a permanent, legal basis, it is less likely to be the subject of promiscuous use. In too many cases wage-price policies have been applied by employing the economic equivalent of political "dirty tricks," ignoring rudimentary standards of due process. If these arguments are not persuasive, then an additional proviso may be added limiting the exercise of the commission's authority to impose direct controls in any individual case to one year. The authority would expire after 12 months unless the commission demonstrated to Congress that the wage and price behavior of the units involved posed a continuing threat to economic stability.

All of this is rather unheroic, if not prosaic. But the experience of the past 30 years clearly indicates that the heroic concept of wage-price policies has not been realized. To some extent, our attitude toward such policies is still colored by notions of populist retribution or innocent visions of full employment and price stability. It is time that we placed wage-price policies in a more modest, operational framework where their contributions may be more limited, but their failures less dispiriting.

THE ROLE OF ANTITRUST IN A DEREGULATED ENVIRONMENT

By Ronald R. Braeutigam*

CONTENTS

	Page
Summary	48
1. Introduction	50
2. Antitrust and Regulation: A Study in Conflict	51
The Interface Between a Regulated and an Unregulated Sector	52
3. The Trend Toward Regulatory Reform	54
Airline Service	54
Natural Gas	54
Oil Prices	55
Motor Carriers	55
Railroads	55
Telephones	56
4. Monopoly	56
How Will Antitrust Deal With Firms That Have Gained Large Market Shares Under Regulatory Sanction?	57
Will a "Thrust Upon" Argument Be a Valid Defense?	57
Partial Deregulation: Immunity and the Interface	58
How Will Antitrust Deal With Refusals To Serve?	59
5. Horizontal Restraints and Oligopoly	60
How Effective Will Antitrust Be In the Wake of Rate Bureaus?	60
Will Regulatory Reform Lead to Free Entry?	62
Will Regulatory Reform Introduce New Incentives for Predatory Pricing?	63
6. Other Areas: Vertical Restraints and Mergers	65
The Merger Problem	66
7. Conclusion	67
Bibliography	68

SUMMARY

This paper focuses on new trends and problems that will confront antitrust enforcers as a result of regulatory reform. It emphasizes those problems that are either new or take on more significance because of regulatory reform, rather than reiterating well-known problems often treated in textbooks on antitrust. The paper also draws numerous examples from industries most likely to be affected by regulatory reform—including, among others, the airline, stockbrokerage, railroad, telephone, cable television, and hydrocarbon production in-

* Northwestern University and California Institute of Technology.

EDITOR'S NOTE.—Regulatory reform is not necessarily synonymous with deregulation. If regulatory reform is viewed as separate from deregulation—which is one subset—then regulatory reform in no way suggests *a priori* the need for vigorous antitrust action. In some instances, however, Congress may determine that deregulating an industry may be more conducive to overall economic efficiency.

dustries. Since no single form of regulatory reform is typical, the role of antitrust will vary from one industry to another. In cases where regulatory reform measures alone do not assure that markets will perform in a competitive fashion, some attention must be paid to antitrust policy.

If any single theme has emerged as dominant, it is this: The most complex problems will arise in those industries in which deregulation is partial. In these cases, the social control of an industry creates policy problems that may find neither mutual exclusion nor collective exhaustion in the course of regulation and antitrust. In short, there is a danger that regulators and antitrust enforcers will fight over jurisdiction in some important matters, while other important problems receive the attention of neither.

This uncomfortable possibility may occur in a number of areas. (1) Who will scrutinize the price at which one firm sells goods or services to an affiliated firm, especially when only one of the firms is regulated? (2) Who will determine when a price in a regulated market is predatory or otherwise anticompetitive, particularly if the firm charging that price also serves a regulated market? (3) What will be the boundary of antitrust immunity? (4) Under what conditions will a merger involving a firm serving both regulated and unregulated markets be allowed, and who specifies these conditions? (5) Who will decide when a refusal to serve is illegal, particularly if the sale in question involves both regulated and unregulated firms? In the text are circumstances in selected industries under which each of these dilemmas might actually occur.

Whether deregulation is partial or not, the first task of antitrust enforcers will be to determine whether structural change is required to prevent the exercise of unchecked economic power by firms now unaccountable to regulators. Structural change may be necessary to foster competitive markets, particularly where regulation has created highly concentrated markets. Antitrust enforcers must expect two new types of defense: (1) That the large market shares for which structural relief is sought were thrust upon existing firms by regulators; and (2) that structural relief is unnecessary, since deregulation will by itself naturally erode the market shares of larger firms.

The paper also emphasizes that antitrust enforcement will encounter a number of practices antithetic to the creation and maintenance of competitive markets, practices that are deeply ingrained in the fabric of the industries being deregulated. It will not be easy for antitrust to overcome the inertia of decades of sanctioned collusion and monopoly. Some of the institutions at the heart of the regulated system must be eliminated with deregulation, including domestic rate bureaus in transportation industries where price and entry are decontrolled. Even then antitrust enforcers must watch closely to insure that behavior is independent, especially where other institutions, such as international conferences, continue to exist.

Finally, antitrust enforcers must constantly watch for any obstacles that impede free entry where free entry is desirable. Regulatory reform by fiat does not guarantee free entry in fact.

1. Introduction

In recent years there has been a wave of effort to lessen the extent to which certain industries have been regulated. The broad label applied to this movement is deregulation. It has affected airlines, stockbrokers, railroads, motor carriers, telephone companies, cable television, and hydrocrabon producers. To date deregulation has been widely implemented in some industries (e.g., stockbrokerage), and quite limited in others (e.g., railroads). It has been formalized in some cases (e.g., natural gas), and only proposed in others (e.g., motor carriers). In short, there is no single form or extent of deregulation that can truly be called typical.

In the United States antitrust and regulation are two important policy instruments for controlling industries that do not perform well absent government intervention.[1] Where intervention is required, the tools of antitrust are typcially the first selected if competitive markets can be forged with their use. Where that is not possible, regulation provides a second line of control.

While deregulation may result in a number of benefits, it may not always lead to the initiation and maintenance of effective competition. New and innovative approaches to antitrust may be required, even in those areas of antitrust that are rather traditional in presently unregulated sectors.

Virtually every major aspect of regualtion is antithetic to antitrust policy. Under some forms of regulation, firms meet through institutions such as rate bureaus to discuss tariff proposals openly. They also often agree to market sharing or market splitting arrangements for which they seek regulatory sanction. They have frequently sought and received permission to effect mergers that would most surely not have been allowed in unregulated markets.

These kinds of interfirm activities are well entrenched after decades of regulation. With deregulation, antitrust may be called on to take an especially hard stance against these activities to foster the independent behavior that will be required for competitive markets. Even that may not be enough. For example, the creation of a competitive market environment in a given industry may not only require a cessation of a trend toward mergers, but even a reversal. The task of antitrust will not be easy here, especially since firms may enter a defense of prior regulatory sanction against such an action.

The role of antitrust in a deregulated environment is by no means a clear one. It will obviously depend on the form that deregulation takes in each industry, an issue discussed in section three. The role of antitrust is not clear even now, as the deregulatory movement unfolds, nor was it before the wave of deregulation began. History shows that the boundary between regulatory and antitrust jurisdictions has never been completely delineated, which is described more fully in section two. Deregulation will not eliminate the narrow and awkward description of that boundary; it will merely shift the battleground in every case in which a portion of the industry remains regulated. In section three this issue will require resolution in a number of industries.

---

[1] Compare this with, for example, the United Kingdom, in which nationalization is an often employed form of social control of industry.

The paper then concentrates on those areas where the burden on antitrust will probably be the greatest with deregulation. Section four addresses some of the problems that can be expected in the areas of monopoly. Section five examines the kinds of problems that will confront antitrust enforcers in the areas of horizontal restraints and oligopoly, including pricing issues. Section six addresses other aspects, including vertical restraints and mergers, and section seven briefly summarizes some of the most important findings.

The paper does not emphasize many of the traditional issues that have been addressed in the vast literature on antitrust. A repetition of these issues would not further the purpose here, especially given the numerous excellent treatises in the field, including Areeda (1974), Kaysen and Turner (1959), Scherer (1970), Weiss (1967), and Bork (1978).

No attempt is made to reiterate the many arguments for and against deregulation in the various industries, except where those arguments specifically relate to the issues to be encountered in the enforcement of antitrust policy. Rather, the focus is on issues which are either new or taken on a larger significance as a result of deregulation. This task alone will prove challenging enough.

2. Antitrust and Regulation: A Study in Conflict

It is true that Government activities influence even unregulated markets in many ways. For example, the Government controls import tariffs and quotas, regulates the money supply, levies taxes, controls government expenditures on goods and services, enforces contracts, and determines minimum wages. These actions affect virtually all markets.[2]

By contrast, in markets within the regulated sector of the economy, the Government intervenes as a referee to affect the heart of the mechanism that allocates resources.[3] The levels of prices, quality of service, investment in plant or profits may be controlled. Firms may legally disseminate data about prices and levels of output, and may engage in joint efforts to influence Government sanctions of the same. Entry into and exit from markets may be limited. Price discrimination may be sanctioned by regulatory authorities, and, as mentioned earlier, mergers that might not be allowed in unregulated markets may be approved under regulation. A number of kinds of activities may be allowed under regulation that would otherwise be illegal.

The rationale for regulation has been described in many places in the literature, and need not be reiterated in any detail here.[4] The reason most often cited from an economic prespective is the natural monopoly argument. A natural monopoly is said to exist in markets in which "the minimum optimal scale of production is so large that there is room in a given market for only one or at most very few firms realizing all production and distribution economies of scale."[5] Thus, the argument goes, a single supplier (or a few suppliers) would be able to serve the entire market at a lower cost per unit of output than if there were many competing suppliers. When the preservation

---

[2] See Kahn (1970), Vol. I, especially pp. 2–3, for a description of the regulated sector.

[3] For a good review of the literature in regulation, see Joskow and Noll (1978).

[4] See Kahn (1970), Vol. I, pp. 5–11 for a discussion of the legal rationale for regulation, and pp. 11–12 for a summary of the economic rationale.

[5] Scherer (1970), pp. 519–520.

of a competitive market is made difficult by the nature of production technology, an exclusive franchise is granted and monitored under regulation.

There are other potential justifications for regulation.[6] Regulation may be used to: (1) dampen the effects of economic fluctuations on certain markets; (2) subject the effects of changes in the economic environment to approval by administrative process instead of an impersonal market mechanism; or (3) deal with conditions that might arise from incomplete information in a market. Other possible reasons include the redistribution of income by controlling the extent to which price discrimination is allowed, or by requiring one service to subsidize another. These redistribution schemes often require limited entry so that firms cannot enter only the lucrative parts of regulated markets and thereby reap the rewards of cream skimming. In addition, it is sometimes argued that regulation prevents windfall profits, and allows regulators to adjust for externalities that may exist when firms and consumers do not base their actions on all of the social costs and benefits associated with a market.

For whatever reasons an industry may have been regulated, enforcement of the antitrust laws in this country has historically been quite limited in regulated industries. As Areeda has noted, ". . . where there is natural monopoly there is little reason for antitrust policy except insofar as: (1) the maintenance of monopoly ceases to be inevitable; or (2) power in the monpoly area radiates outward into areas where competition is both possible and desirable."[7]

The Interface Between a Regulated and an Unregulated Sector

If an entire industry were a natural monopoly, then the tasks confronting both regulators and enforcers of antitrust would be simpler than they often are. A typical example of such a monopoly would be a local electric utility, whose services are provided largely without competition from other services offered by unregulated companies. Regulators have more complete control over such a monopoly than they would if unregulated rivals provided a service that would be a substitute; antitrust concerns itself less with injury to the nonexistent rivals, since regulators have jurisdiction over the entire existing industry.

The boundary between regulated and unregulated sectors, however, is not always so clear. For example, regulated railroads often face competition from unregulated barges or from an unregulated sector of the motor carrier industry.[8] Regulated telephone companies now face competition in the manufacturing of terminal equipment and in the provision of domestic long distance private line communications services.[9] And until the Natural Gas Policy Act of 1978, regulators of in-

---

[6] For a good summary of these potential reasons for regulation, see the "Study on Federal Regulation," United States Senate Committee on Governmental Affairs (1978), Volume VI, pp. 270–291.
[7] Areeda (1974), p. 106.
[8] See Kahn (1971), Vol. II, Chapter 1.
[9] See Owen and Braeutigam (1978), Chapter 7.

terstate wellhead sales of natural gas had no jurisdiction over intrastate sales, and could not force producers to direct gas supplies to the less lucrative interstate markets.[10] In such cases as these, regulators have found that their control over the industry is much less extensive because of competition from an unregulated sector. Similarly, enforcers of anitrust may be concerned that the performance of the unregulated sector is somehow impaired by the regulated sector.

Increased interdependence among markets at the boundary between regulated and unregulated sectors has expedited the movement toward deregulation in instances such as those just mentioned.[11]

Where competition at the fringe has proven viable, the natural monopoly argument for regulation has been questioned by those who suggest that many of the resource allocation decisions previously made by regulators might better be made through an unregulated market.

Since deregulation will move, but not eliminate the interface between regulated and unregulated sectors, questions of implied immunity, primary and exclusive jurisdiction, and state action will continue to await resolution. Historically, the arm of antitrust has found jurisdiction in some regulated industries,[12] not so in others,[13] and has found ambiguous stature in yet other cases.[14]

At still another level the extent of the role of antitrust remains unresolved. Specifically, when does the (legal) use of the administrative process of regulation differ from the (illegal) abuse of that process? [15] Under what circumstances, if any, can the antitrust statutes be used to limit the extent to which a particular interest group engages in lobbying or other activities to delay proceedings or to deter the interests of other groups? In two cases decided in the early 1960's, the Supreme Court appeared to eliminate abuse of process as a Sherman Act violation, based on immunity implied by the First Amendment.[16] More recently, the Court has held that there are some abuses of the administrative process that do constitute antitrust violations, including the knowing submission of false data to a regulatory authority,[17] and the concerted and repeated effort of an interest group to use litigation to deter entry.[18]

In short, a number of legal issues remains undecided. While it is obvious that deregulation will move some of the issues from the domain of regulation into that of antitrust, the nature of the problems at the interface will remain.

---

[10] See Owen and Braeutigam (1978), Chapter 3.
[11] See Nelson (1978), and Braeutigam (1978).
[12] *Cantor* v. *Detroit Edison*, 44. L.W. 5357 (1976), *U.S.* v. *AT&T*, Civil Action No. 74–1698 (District D.C.), Memorandum and Order on Jurisdictional Issues, filed Nov. 24, 1976.
[13] *U.S.* v. *NASD*, 422 U.S. 694 (1976).
[14] *McLean Trucking Company* v. *United States*, 321 U.S. 67 (1944). For an interesting discussion of this case, see Areeda (1972), p. 45.
[15] See Owen and Braeutigam (1978), pp. 32–35, for a discussion of use and abuse of administrative process.
[16] See *Eastern Railroad President's Conference* v. *Noerr Motor Freight* 365 U.S. 127 (1961) and *United Mine Workers* v. *Pennington*, 381 U.S. 657 (1965).
[17] *Woods Exploration* v. *Alcoa*, 438 F.2d 1286 (5th Circuit), *cert. den.*, 404 U.S. 1027 (1971).
[18] See *California Motor Transport* v. *Trucking Unlimited*, 404 U.S. 508 (1972) and *Otter Tail Power Company* v. *U.S.*, 410 U.S. 366 (1973). In the first case the Supreme Court prohibited a group of truckers from collectively planning to exhaust process by opposing all new trucking applications for entry certificates. In the second case Otter Tail attempted to discourage municipalization of power production by repeated litigation, and the Supreme Court proscribed this practice.

3. The Trend Toward Regulatory Reform

As noted in section one, the term deregulation has taken on a number of different meanings, depending on the industry in question. This section gives context to its meaning for some industries that have recently been or may be the target of some type of deregulation. The exact form of deregulation is not yet known fully in any of these industries. Even those which now have statutes (for example, natural gas) will require a number of regulatory judgments and procedural specifications not known at the present time. Nevertheless, the general spirit of the deregulation movement in each case can be described in this section, allowing us to discuss some of the potential problems and issues relevant to the future of antitrust in subsequent sections.

Airline Service

Starting in 1975, the Civil Aeronautics Board (CAB) began to relax its tight grip on the levels of air fares. It followed a gradual path toward rate freedom to the maximum extent consistent with its mandate to regulate, the Civil Aeronautics Act of 1938.[19] A wave of tariff reductions swept across the industry beginning in 1977 and accelerating in 1978.[20] At the conclusion of 1977, both entry into and tariffs on commercial air freight transportation were decontrolled by statute, with a proviso that future tariffs for freight service would not be predatory. The notion of a predatory price was not defined by statute The Airline Deregulation Act enacted deregulation of air passenger service. The Act was designed to decontrol both rates and entry.[21] In section five, important aspects of entry are described that will determine whether a deregulated air industry will lead to vigorous competition. Free entry by fiat need not imply that free entry will in fact exist; as a corollary it follows that deregulation by fiat may not automatically lead to a vigorously competitive market.

Natural Gas

Following a number of proposals to deregulate natural gas during the 1970's, Congress successfully enacted the Natural Gas Policy Act of 1978.[22] The NGPA gradually deregulates a category of natural gas called "new" natural gas produced at the wellhead, with total decontrol of wellhead prices occurring after 1985.[23]

Regulation will remain in force for large segments of the natural gas industry, including the prices of gas not designated as new gas (i.e., "old" gas), the pipelines that transport gas from the field to local

---

[19] Stat. 102, 72 Stat. 740, USCA 1303.

[20] For a discussion of these reductions in tariffs, see Keeler (1978), pp. 135–136.

[21] For a discussion of the events leading to the Airline Deregulation Act, see Keeler (1978).

[22] See the Federal Register, Federal Energy Regulatory Commission, Natural Gas Policy Act of 1978, Friday, December 1, 1978, Part VIII.

[23] New natural gas produced from onshore wells must come from new reservoirs, or new wells no closer than 2.5 miles from the nearest marker well (a marker well is any well from which natural gas was produced in commercial quantities after January 1, 1970, and before April 20, 1977, with the exception of wells whose surface drilling began after February 19, 1977), or if closer than 2.5 miles to a market well, 1000 feet deeper than the deepest completion location of each market well within 2.5 miles.

markets, and the local utilities that distribute the gas to customers.[24] In section five we set the interface between unregulated producers and the regulated pipelines may pose some interesting and difficult problems to both regulators and enforcers of antitrust.

Oil Prices

Like prices for natural gas, domestic oil prices have been held below the price on the world market for a number of years. At the start of 1979, approximately 30 percent of the domestically produced oil was subjected to a ceiling of about $6 per barrel of crude oil; the balance was to be priced at about $14 per barrel, and that was well below the current price of oil on the world market.[25]

In his nationally televised speech on energy on April 5, 1979, President Carter announced his intention to deregulate the price of all domestic oil, with successive steps of decontrol being completed by perhaps 1981. This form of deregulation differs markedly from the deregulation of natural gas, since old natural gas will remain regulated even after 1985. The President has announced plans to propose to Congress a windfall profits tax to prevent oil producers from realizing the large supernormal profits that would result with deregulation absent the tax. At this writing the exact form of the proposed windfall profit tax is not clear.

Motor Carriers

Although there exists no deregulation statute for motor carriers, there is a clear movement afoot to seek deregulation of that portion of the interstate motor carrier industry that is now regulated. The Interstate Commerce Commission (ICC) regulates approximately 46 percent of the ton-miles of intercity freight carried by the trucking industry. Intrastate agencies regulate another 10 percent, with the balance being unregulated.[26] When deregulation is discussed in connection with the motor carrier industry, it usually refers to freedom of entry in the carriage of any commodity over any route at unregulated tariffs.

Railroads

Although the complete deregulation of railroads, including the removal of all tariff and entry restrictions, has not been a part of the recent wave of deregulation, a new law has relaxed certain elements of railroad regulation. The passage of the Railroad Revitalization and Regulatory Reform Act of 1976 (the Quad-R Act) is most notable in this respect. Among other things, this Act in principle allows a railroad to vary its tariffs within a "zone of reasonableness" without obtaining approval from the Interstate Commerce Commission, absent a finding of the ICC that the railroad has "market dominance."

While one might have expected railroads to alter their rates given this new flexibility, in particular by lowering rates where they face

---

[24] For a detailed discussion of the various segments of the natural gas industry, see Braeutigam (1978).
[25] Montgomery (1978), p. 825.
[26] See Roberts (1978), p. 473.

intermodal competition, no new rash of rate readjustments has followed the passage of the Act. In fact, as Nelson shows, ". . . railroads have been very cautious about taking advantage of the . . . Act during the first 18 months or so of its validity." [27]

Telephones

Two major areas of the telephone industry previously monopolized by regulated telephone companies have been opened to entry in the last decade. The two areas include the supply of terminal equipment and the provision of long distance private line transmission. The courts have sustained the decision of the Federal Communications Commission (FCC) to allow competition in the supply of terminal equipment. The primary remaining restriction is that the equipment be a type certified as acceptable by the Commission.[28]

The long distance private line market cannot yet be described as competitive, although some entry has occurred. Both rates and entry remain regulated in this market.

Congress is now engaged in an effort to rewrite the Communications Act of 1934 to reflect the existence of a number of new technologies (e.g., fiber optics, satellites, and microwave systems) that have arisen since the Act was passed over 40 years ago. It may very well be that more competition in various areas of the telephone industry will result from that effort.

Several other industries marked by some relaxation in regulatory restraint could be added to this list, including cable television,[29] banking,[30] securities markets,[31] and water carriage.[32] The main point of even this partial enumeration is to show how deregulation can vary across industries, and to provide a background against which to assess the role of antitrust in the future.

4. Monopoly

One of the most often cited reasons for the implementation of a regulatory scheme, as described in section two, is the prevention of the unfettered exercise of monopoly power, especially when technology precludes the competitive coexistence of a large number of firms. If deregulation is to succeed, it must do so at least in part because the industry or part of an industry that is deregulated is not a natural monopoly. In some of these industries a single firm or a few firms have managed to achieve large market shares under regulation. This leads to the first of the issues that antitrust must face with deregulation:

---

[27] Nelson (1978), p. 64.

[28] See Owen and Braeutigam (1978), Chapter 7.

[29] See Owen (1978), pp. 347–389.

[30] See the United States Senate Committee on Governmental Affairs, Vol. V (1978), pp. 197–228.

[31] See Stoll (1978), pp. 589–656.

[32] For example, in 1973 the ICC abolished the so-called "barge mixing rules," thereby better enabling water carriers to compete with other modes in transporting certain commodities. See Lieb (1978), p. 92. Only a small percentage of intercity water carriage is regulated in any case.

How Will Antitrust Deal With Firms That Have Gained Large Market Shares Under Regulatory Sanction?

The application of antitrust to firms with large market shares in unregulated industries has changed over the years, and is still an issue without clear resolution. As recently as January 1979, the National Commission for the Review of Antitrust Laws and Procedures recommended to the President and the Attorney General that "a proviso be added to the end of Sherman Act Section 2 in order to clarify the appropriate standards" for determining whether a firm has attempted to monopolize an industry.[33]

The purpose here is not to attempt a treatise on the comparative merits of *per se* and rule of reason interpretations of the Sherman Act, a traditional issue of antitrust, but rather to ask whether firms in industries just deregulated or in a transition to a deregulated state will be treated in the same way as firms in historically unregulated industries. At the two extremes antitrust agencies could attempt to break up firms with large market shares as soon as deregulation is enacted (consistent with the structural approach of antitrust since *Alcoa*),[34] or they could adopt a temporary wait-and-see attitude to find out whether deregulation measures alone are sufficient to induce a competitively performing industry. They could also pursue an intermediate stance, depending on the political climate, including the extent of antitrust enforcement activity tolerated by Congress, the availability of suitable remedies under existing legislation, and the possibility of obtaining new remedies with new legislation. If antitrust agencies do attempt to break up large firms, they may have to deal with the following issue.

Will a "Thrust Upon" Argument Be a Valid Defense?

In the *Alcoa* case the courts acknowledged the possibility of a "thrust upon" defense in a monopolization case. In *Alcoa* the issue was whether Alcoa had achieved a monopoly in the ingot market by actions to exclude its competitors, or whether monopoly had been thrust upon Alcoa by virtue of its "superior skill, foresight, and industry." Although the Court found that Alcoa was not a monopoly, the Court did leave open the possibility that a thrust upon defense might be valid against a charge of monopolization.

A variation of this defense may well occur with deregulation, since a previously regulated firm might argue that its large share of a market was thrust upon it by regulation. While such an argument is not only possible, but perhaps inevitable with deregulation, it will be of utmost importance that antitrust enforcement overcome this defense. Otherwise, the functioning of these markets will be checked by neither regulation nor antitrust.

---

[33] National Commission for the Review of Antitrust Laws and Procedures (1979), pp. 150–151.
[34] *United States* v. *Aluminum Co. of America, et al.* 148 F.2d 416 (1945).

Partial Deregulation: Immunity and the Interface

As noted in section three, in several industries the present movement is toward partial rather than complete deregulation. In those cases the distinction between the regulated and unregulated activities in an industry may be narrow and awkward. It may not be obvious where antitrust immunity exists under the regulatory umbrella, particularly if an industry is continually introducing new services or products that require a determination of jurisdiction.

Two examples may help to illustrate this point. First, consider the present movement in the telephone industry to allow competition in the provision of long distance private line telecommunication services. At present this activity remains regulated, although the Federal Communications Commission has decided to allow entry at regulated tariffs.[35]

Although the FCC allowed entry into private line markets, it wanted to retain a monopoly status for the traditional long distance "message toll service" (MTS) markets. In fact when the entrants into the private line markets attempted to introduce new services that, in the opinion of the FCC, too closely resembled the MTS services of the established telephone carriers, the FCC attempted to reject those offerings.[36] On appeal, however, the courts have reversed the FCC and denied the notion that MTS should be granted the standing of a statutory monopoly.[37]

The point is this. With deregulation, antitrust enforcers may find themselves confronted with unregulated markets that have significant interactions with regulated markets. The classic questions arise. What is the relevant market? Does regulatory action in one area supersede antitrust action in a closely related but unregulated market?

It is natural to hope that, with well-designed deregulation measures, problems like the one above will be minimal. But it would be naive to believe they will be nonexistent.

The railroad industry provides the second example of the kinds of problems that may arise at the interface between a regulated and unregulated sector. As described in section three, the Railroad Revitalization and Regulatory Reform Act of 1976 allows railroads to change prices within zones of reasonableness without ICC approval, as long as the ICC does not determine the firm to have market dominance over a particular commodity. This suggests that under partial deregulation, a regulatory commission may take on the role of an antitrust enforcer. The recent United States Senate Committee on Governmental Affairs' "Study on Federal Regulation" commented on this as follows:

> In other words, if the ICC determines that a railroad is dominant over a particular commodity, full rate regulation would be maintained. The statute attempts directly to answer one of the main concerns of those opposed to deregulation: The possible abuse of monopoly power. The ICC therefore assumes the role of antitrust enforcer.[38]

---

[35] See FCC Final Report and Order, Docket 18920 (Specialized Common Carrier Services), Federal Register, June 9, 1971, paragraphs 103, 120.

[36] Most notably, in 1975 the FCC rejected a proposal of MCI to offer its so-called Execunet service (see FCC Order 75-799, July 2, 1975). For a discussion of this see Owen and Braeutigam (1978), pp. 229-230.

[37] Ibid., p. 230.

[38] "Study on Federal Regulation" (1978), Vol. VI, p. 93.

Thus the role of antitrust enforcement may take on a new character with deregulation. The cast of public representatives may be expanded beyond the Department of Justice and the Federal Trade Commission to include regulatory agencies themselves. This increased division of responsibilities may actually increase the immunity of a partially deregulated industry from antitrust attack from sources outside a regulatory agency.

How Will Antitrust Deal With Refusals To Serve?

Where regulation grants an exclusive franchise, it usually imposes a common carrier obligation on the recipient of that franchise. This obligation typically states that the firm is required to serve all customers who demand service under the conditions stated in existing tariffs. Thus, customers will be assured of receiving service even though a single company; or in a case like the airlines, a few companies have charge of producing that service.

As deregulation occurs in airlines, parts of the telephone and railroad industries, and in the motor carrier industry, the common carrier obligation is likely to be removed from the deregulated portions of these industries. In some cases, customers may claim that they based their decisions to enter some major enterprise, for example, the building of a plant, on the expectation of the continued provision of a common carrier service such as railroad transportation.

Refusals to deal (or serve) and group boycott are well defined antitrust offenses. Where partial deregulation occurs, however, the role of antitrust may be delicate, particularly if a regulatory agency such as the ICC is acting as an antitrust enforcer in some areas as suggested earlier. Even without that complication the question is not easy to answer. Should provision service be required over some time until an otherwise deprived customer is able to make other arrangements? What if other arrangements are extremely costly? Will the court system be inundated with many antitrust grievances previously brought before regulators?

Refusals to serve have long been an issue within regulated industries. When firms other than telephone companies attempted to gain the approval of the FCC for the attachment of customer terminal equipment that they manufactured, A.T. & T. opposed this strongly. Customers using such terminal equipment encountered a great deal of resistance over a number of years before they secured the right to network service.[39] As we noted in section three, the FCC decision to allow non-telephone companies to manufacture customer terminal equipment was in itself a form of partial deregulation. Enforcers of antitrust can expect more of this with the broadening of the deregulation movement.

To be realistic, antitrust can no more hope to eliminate all pockets of monopoly power from deregulated markets than it has in historically unregulated markets. At best it can be hoped that the use of antitrust tools will minimize the extent of such economic power, and several characteristics of deregulated markets may make this task more difficult than in historically unregulated markets. These characteristics

---

[39] See the discussion of the Hushaphone and Carterfone cases in Owen and Braeutigam (1978), pp. 230–231.

include initially highly concentrated markets, the anticipation of the thrust upon defense, and the potential barriers that limit the reach of antitrust where deregulation is partial.

5. HORIZONTAL RESTRAINTS AND OLIGOPOLY

In terms of relative numbers, the most common types of antitrust cases are those involving horizontal restraints.[40] Where no single firm can monopolize an industry, firms may have an incentive to conspire to monopolize an industry, using such tactics as price fixing and dissemination of data on output, market shares, and prices. Where firms can collude, they may be able to extract extra-normal profits if they can restrict output to bring higher prices, just as an unregulated monopolist might be expected to do.

The courts have not been sympathetic toward overt attempts at collusion. In fact, price fixing is illegal *per se*, as are trade association activities designed to facilitate the exchange of market share and price data.[41]

Most oligopoly cases, however, do not involve overt collusion, and the courts have struggled for years over the circumstances under which it may be possible to infer collusive action from such evidence as parallel pricing behavior. Judge Medina in the *Investment Banker's Case* summarized the nature of the dilemma confronting the courts:

> True it is that conspiracies . . . are often hard to detect. No direct proof of agreement between the wrongdoers is necessary; circumstantial evidence of the illegal combination is here as elsewhere often most convincing and satisfactory. But, when all is said and done, it is the true and ultimate fact which must prevail. Either there is some agreement, combination or conspiracy or there is not. The answer must not be found in some crystal ball or vaguely sensed by some process of intuition, but in the evidence adduced in the record of the case which must be carefully sifted, weighed, and considered in its every aspect. This is an arduous but necessary task.[42]

One cannot expect the difficulties in detecting collusive behavior to be overcome soon. Regulation has created a number of structural conditions and practices that may prevent the emergence of effective competition, including equipment standards, credit terms, maintenance standards, and output quality standards, that may contribute to parallel behavior that is most difficult to prevent. In addition, in some cases regulation has created and sanctioned institutions, like rate bureaus, whose primary function is to assure conformity in some aspects of performance.

How Effective Will Antitrust Be in the Wake of Rate Bureaus?

One of the interesting features of domestic transportation is the overt collusion among carriers in a given mode in recommending tariffs to regulators. These collective efforts are accomplished through nonprofit organizations of the carriers known as rate bureaus or conferences.

---

[40] See Posner (1970) for a statistical summary of the types of antitrust cases brought by the Department of Justice and the Federal Trade Commission from 1890 to 1969.

[41] See Areeda (1974), Chapter 3. However, not all data dissemination practices are illegal. See *Tag Manufacturers Institute* v. *FTC*, 174 F.2d 452 (First Circuit, 1949).

[42] *United States* v. *Morgan*, 118 F. Supp. 621, 634 (S.D.N.Y. 1953), quoted in Stelzer (1976), p. 122, as a part of the decision in *United States* v. *Chas. Pfizer and Co., Inc.* 367 F. Supp. 91 (S.D.N.Y. 1973).

These organizations date back into the last century, and they have historically been viewed dimly by antitrust enforcers. Indeed as Weiss and Strickland note, the first collusion case brought under the Sherman Act and reaching the Supreme Court involved the Trans-Missouri Freight Association.[43] This organization consisted of 18 railroads that controlled traffic west of the Mississippi River, and the association attempted to set rates for all its members. In striking down this arrangement, the Court first enunciated its *per se* interpretation of the Sherman Act, that every restraint of trade was illegal.[44]

The Supreme Court again struck down the practices of rate bureaus in *Georgia* v. *Pennsylvania Railroad Co.*, finding that Congress had not empowered the ICC to exempt railroad carriers from the Sherman Act.[45] Following extensive hearings on the matter, however, Congress found much support for bureaus from both shippers and carriers. In 1948 Congress granted rate bureaus statutory exemption from the antitrust laws by passing the Reed-Bulwinkle Act.[46] This exemption has continued until today. The Railroad Revitalization and Regulatory Reform Act of 1976 requires that bureaus cannot vote on rates for services provided by only one line and that only carriers that could engage in a joint line movement can vote on a joint line rate. Within this structure, then, rate bureau activities are immune from antitrust.

Domestic rate bureaus are typically organized by geographic regions. There are 10 railroad rate bureaus, 11 major rate bureaus for motor carriers, and several others for domestic water carriers. Additionally, international air carriers have their own organization to coordinate international fares, the International Air Transport Association (IATA),[47] and the Federal Maritime Commission has the power to grant antitrust immunity to conferences among ocean carriers.[48]

The potential problems that rate bureaus pose amidst a movement toward deregulation are strongly apparent. The power to fix prices is antithetic to the functioning of a competitive market. The National Commission for the Review of Antitrust Laws and Procedures has recommended to the President and the Attorney General that the Reed-Bulwinkle Act should be repealed,[49] and that the antitrust exemptions granted to ocean shipping conferences should be examined closely and removed where there is excessive and unnecessary restraint of trade.[50] The abolition of rate bureaus is especially important where the total deregulation of pricing and entry is contemplated, as the case may be in the motor carrier industry.

The existence of international conferences poses a particularly difficult problem in public policy. On the one hand, these conferences may facilitate the achievement of diplomatic and national defense objectives.[51] On the other hand, they may provide a means for participants

---

[43] *United States* v. *Trans-Missouri Freight Association,* 166 U.S. 290 (1897).

[44] The Supreme Court abandoned the *per se* interpretation of the Sherman Act in favor of a rule of reason in 1911 in *Standard Oil Company of New Jersey* v. *United States*, 211 U.S. 1 (1911). The courts received the *per se* interpretation in 1945 with *United States* v. *Aluminum Company of America,* 148 F.2d 416 (2d Cir. 1945).

[45] *Georgia* v. *Pennsylvania Railroad Co.,* 324 U.S. 439 (1945).

[46] 62 Stat. 472 (1948).

[47] Lieb (1978), pp. 172–173.

[48] "National Commission for the Review of Antitrust Laws and Procedures" (1979), p. 275.

[49] Ibid., p. 197.

[50] Ibid., p. 273.

[51] Ibid., p. 273.

to disseminate data and to otherwise engage in activities that have effects on the domestic markets that some of the participants may serve, for example on the airlines industry. The role of antitrust will no doubt depend crucially on the circumstances particular to each industry.

In short, then, our discussion of rate bureaus emphasizes that deregulation measures alone will not necessarily lead to independent behavior, especially since established firms may have well developed mechanisms for communicating information detrimental to a competitive market performance. Similarly, antitrust enforcers must recognize that free entry by fiat need not lead to free entry in fact, a subject taken up now.

Will Regulatory Reform Lead to Free Entry?

Free entry is often cited as necessary for perfectly competitive markets. Of course, in many of the industries in which deregulation is occurring, it is important to recognize that it may be desirable to require entrants to meet certain standards. For example, no one has seriously argued that airline deregulation should include an abolition of safety standards for aircraft or in the use of airways. Similarly, few would suggest that standards of financial responsibility should be abolished for insurance companies and brokers.

In one sense, the existence of any such standards means that entry is not truly free. Yet in another, if someone who is willing to satisfy these standards is allowed to enter by law, then a form of free entry exists. The important point to make here is that the legislation of free entry need not lead to free entry in fact. Where free entry is desirable, antitrust enforcers should pay particular attention to all features of a market that might deter entry, including the existence of large entry costs, excess capacity, and the potential of a multiproduct firm to extend monopoly power from a market that is regulated into another market that is deregulated.

The notion of large entry costs is usually applied to situations in which large capital requirements exist to enter, particularly if this results in economies of scale in production. Kahn has noted that in industries of this sort (for example, the local distribution of electric power or the local telephone exchange) destructive competition might result in the absence of regulation.[52] Segments of industries that exhibit these characteristics are not good candidates for deregulation of both prices and entry, a statement largely reflected in the nature of the deregulation movement for the various industries discussed in section three.

While the absence of large entry costs and economies of scale, however, is a necessary condition for a competitive market performance, that absence is not sufficient. There may be institutional or historical reasons for which entry might not be truly free. For example, in the airlines industry existing firms have secured choice gate locations and time slots at major airports and established well-developed schedules for connecting flights. Although the technological barriers to entry are relatively low in this industry, effective entry requires that landing

---

[52] Kahn (1971), Vol. II, pp. 172–178.

rights can be obtained by new rivals, a problem most likely to arise in congested airports.

At the present time it is not clear just how these landing rights will be made available to firms in the industry. It is not sufficient to dismiss this caveat by simply saying that air slots, including landing rights, will be auctioned off in some undefined manner. Studies have shown that the structure of a market can be strongly influenced by the kind of auction that is conducted.[53] The tenure of the landing rights that are purchased, the relative sizes of the bidders, and the nature of the transportation network served by each bidder will also affect the performance induced by a particular type of an auction.[54] While a detailed discussion of auction processes is well beyond the scope of this paper, antitrust enforcers should take an interest in the development of these institutional arrangements with deregulation, since the structure and performance of certain markets may be strongly affected by whatever approach is ultimately chosen. For example, at the time of this writing, an issue which remains unresolved is the mechanism by which air slots are to be allocated in congested airports.

Will Regulatory Reform Introduce New Incentives for Predatory Pricing?

Any incentives for anticompetitive pricing that have existed in historically unregulated markets will also appear in unregulated markets. Already discussed is one form of anticompetitive pricing, price collusion. Now addressed will be a second form, predatory pricing.

Pricing that is predatory is not easy to define, as Areeda has noted. "It connotes conduct that has the purpose or effect of destroying or weakening a rival. But, of course, fair competition has the same objective: To prevail in the marketplace relative to rivals." [55] The debate over what constitutes predatory pricing has been tortuous. Does it mean pricing below marginal cost, average variable cost, or where profits are negative? If a measure of profits is to be used, how does one calculate the profits associated with a particular product, particularly if some of the costs incurred by the firm are shared by the product in question and other products? Even if a particular notion of costs or profits is deemed appropriate as a benchmark, it is often difficult to measure the relevant entity.

Finally, a determination of predatory pricing often turns on whether the pricing practice prevails for a long time or a short time. If the latter, is a low price viewed as simply promotional, or as an attempt to eliminate competitors, perhaps with the intent to raise prices later and to deter entry by a threat of a repeated introduction of a low price? [56]

It is well beyond the scope of this paper to examine these issues for unregulated industries in any detail. There is generally less reason for concern over predatory pricing in markets in which entry

---

[53] See, for example, Ferejohn, Forsythe, and Noll (1977), Hong and Plott (1977), Plott and Smith (1978), and Isaac and Plott (1979).

[54] For a discussion of competitive bidding for franchises, see Demsetz (1968) and the United States Senate Committee on Governmental Affairs' "Study on Federal Regulation" (1978), Vol. 5, pp. 99–102.

[55] Areeda (1974), p. 669.

[56] For a good discussion of these issues, see Areeda (1974), pp. 669–673.

is relatively easy. Suppose the aim of predatory pricing is to drive out competitors to create market power, with the intent to raise the price later to generate supernormal profits. Then if entry is easy, supernormal profits cannot long prevail without signaling entry.

This argument sets forth an additional reason why antitrust enforcers should make sure that no large barriers to entry remain in deregulated industries. Incentives to engage in predatory pricing are greatly reduced when entry is easy.

Rather than focusing on these well-known issues, the main purpose of this section is to address a new kind of pricing issue that may be introduced with deregulation. Specifically of interest is the case in which a firm serving a newly deregulated market also provides services in another regulated market. The major potential consequence of this situation is that a number of the important pricing dilemmas that have long confronted regulators may now be transferred to the courts, requiring resolution by antitrust procedures.

A seminal article by Averch and Johnson suggests reasons for this concern.[57] That article examined a situation in which a regulated firm provides service both to a regulated monopoly market and to a second market that might be opened to entry at unregulated prices. For example, a telephone company might provide service as a regulated monopolist in one market, and provide other services in unregulated markets (e.g., the supply of customer terminal equipment or the provision of long distance private line service). Averch and Johnson have shown that if such a firm is regulated by a rate of return constraint applied collectively to all of the products of the firm, then the firm may have an incentive to price the competitive services below marginal cost to expand the rate base and thereby generate larger profits.

This is notable because a completely unregulated firm would not have an incentive to incur a long run loss in a market.[58] Yet a partially deregulated firm, one of whose markets is totally deregulated, might well have an incentive to sustain a long run loss in a competitive market.

Both the Interstate Commerce Commission and the Federal Communications Commission have spent nearly a decade trying to decide what criteria should be used to describe prices that are fair,[59] a task that has been complex even when all of the markets involved fell within the jurisdiction of a regulator. With partial deregulation, the task will now be split between regulators and antitrust enforcers.

To summarize the point, economic theory has suggested that there may be long-run incentives to price below marginal cost where a single firm serves both regulated and unregulated markets. Under virtually any definition, this would be viewed as predatory pricing.

---

[57] Averch and Johnson (1962).

[58] The assumption is that there are no strong demand complementarities among the services of the partially regulated firm in making this argument for the unregulated firm. Even if there are strong demand complementarities, however, the basic point that a partially regulated firm may have an incentive to set price below marginal cost in a competitive market remains valid even if there are such complementarities.

[59] Federal Communications Commission, "Revisions of Tariff FCC No. 260 Private Line Services, Series 5000 (TELPAK)," Docket 18128, 61 FCC 2d 606, Nov. 26, 1976, and Interstate Commerce Commission, "Rules to Govern the Assembling and Presenting of Cost Evidence," Docket 34013, 337 ICC 298, July 30, 1970.

6. Other Areas: Vertical Restraints And Mergers

The various areas of antitrust cannot always be neatly separated into mutually exclusive areas. Accordingly, already addressed are certain issues relevant to both vertical restraints and mergers. For example, section four on monopoly dealt at some length with one form of vertical restraint, refusals to deal, and I need not repeat those issues here.

This paper recognizes that the traditional issues of vertical restraint addressed in antitrust cases for historically unregulated markets will surely remain relevant in markets that are deregulated, including problems with tying arrangements, exclusive dealing, and exclusive franchising.

The paper focuses on other issues that are perhaps less obvious, and are more directly associated with the deregulation movement.

One of the most important problems of vertical restraint will arise at the interface between regulated and deregulated markets. Section 10 of the Clayton Act prohibits common carriers from purchasing inputs without competitive bidding from companies with whom they have interlocking directorates. There will be at least two ways in which enforcement of this provision will be important with deregulation.

First, antitrust authorities should be alert to the possibility that a firm that is regulated in one of its markets may refuse to deal with any of the firms in a deregulated market other than its own affiliate. For example, if a telephone company maintains a monopoly in local exchange, and has an affiliate that produces telephone equipment, there may be an incentive for the local exchange company to restrict its purchases of equipment to its own affiliate. Of course, the restriction need not be complete. Any such restriction forecloses a portion of the equipment market to competing supply firms.

The potential problems of vertical restraint may go beyond foreclosure of the market. If the local exchange company is regulated by a rate of return, then it might not object to paying higher-than-competitve prices for equipment since these higher prices will be reflected in an inflated rate base and ultimately in higher profits for the local exchange company. The equipment supplier would also realize extranormal profits at these higher prices. This suggests that vertical relationships can lead to extranormal profits in partially deregulated industries; the problem has been formally analyzed by Dayan (1972).

A similar problem could arise in other industries. For example, consider the case of vertically integrated pipelines and suppliers of oil and gas. A pipeline might be willing to pay a higher-than-competitive price for, say, natural gas purchased from its own affiliate. The pipeline could pass these higher fuel costs along to customers under automatic fuel price adjustment mechanisms often used by regulators, and the producers of gas would realize supernormal profits on such sales. Thus, the existence of many competing producers in wellhead markets may not guarantee that actual wellhead sales take place at competitive prices.[60]

---

[60] Braeutigam suggested this (1978), pp. 711–712.

The warning signals from these examples are clear enough. If regulators do not scrutinize these transfer prices closely, antitrust enforcers may be saddled with that responsibility. The task will not be easy. To take only the cited examples, natural gas supply contracts are often complicated so that a comparison of prices from one contract to another is not easy. In the telephone example, there are many different types of equipment whose prices would have to be examined. Since the current direction of the deregulation movement will apparently lead to vertical affiliates that straddle the interface between regulated and unregulated markets, the suggested problems appear to be both important and inevitable.

The Merger Problem

Mergers are an area that will undoubtedly require the increased attention of antitrust enforcers with deregulation. Through their direct determination of market structure, mergers affect market performance. Many of the market structure decisions previously made by regulators will be made in the antitrust arena with deregulation. While merger decisions made by regulatory authorities have not been totally immune from antitrust attack,[61] there can be no doubt that the role of antitrust regarding mergers will be expanded with deregulation.

The most difficult aspect of the merger problem, at least in some industries, is that the structure of the industry sanctioned under regulation may already be oligopolistic. Since at least 1950, with the passage of the Cellar-Kefauver Act, merger rulings in historically unregulated industries have largely attempted to nip oligopoly in its incipiency. It is much easier to prevent a merger that could lead to an oligopolistic development of an industry than to break up firms after an oligopolistic structure has been reached. Unfortunately, in historically regulated industries, the structure of the industry may have long ago become highly concentrated. Thus the role of antitrust may be heavily oriented toward undoing the damage done by past mergers to create competitively structured markets. In some cases this may be difficult to do, particularly where firms involved have some parts that remain regulated while other parts are participants in deregulated markets.

While these tasks may be difficult and somewhat different from the ones involving mergers in historically unregulated markets, the issues antitrust enforcers appear to be largely the same. The central question remains: How big do firms have to be to realize economies of scale in production, and will the size of the market permit enough of these efficiently sized firms to coexist so that a competitive structure can be reached? Because this central issue has not changed with deregulation, the paper does not dwell on the volumes of literature that have attempted to answer this question for each of the industries described in section three.[62]

---

[61] See, for example, *U.S.* v. *El Paso Natural Gas,* 376 U.S. 651 (1964).

[62] See, for example, the United States Senate Committee on Governmental Affairs' "Study on Federal Regulation" (1978), Appendix to Volume 6, for a number of such studies.

This section closes by drawing attention to one rather interesting structural possibility that may arise with deregulation. It is the possibility of the integrated transportation company.[63] It has most often been discussed in the transportation industry, but also may arise in the energy department. Regulation has restricted the extent to which a transport firm can offer service using more than one mode, particularly within the same geographic area. With deregulation, transport firms might try to diversify by forming integrated, multimodal companies, offering perhaps rail, barge, and motor carrier services simultaneously. Such companies could provide transport services more cheaply, especially since they would have incentives to choose the most efficient form of transportation required to render a service. This would involve an expansion of the production activities of existing firms, some of which may be attempted through mergers. Antitrust enforcers may be confronted by a decison whether this type of diversification is consonant with a competitive market structure, and if so, whether a move toward such a structure can be accomplished by the entry of existing firms into other modes without mergers.

7. Conclusion

The paper has described the role of antitrust in a deregulated environment. It has focused on new trends and problems that will confront antitrust enforcers, drawing numerous examples from the industries most likely to be affected by deregulation. An examination of the trend of deregulation for a number of industries shows that no single form of deregulation can truly be viewed as typical. Accordingly, the role of antitrust will vary from industry to industry.

If any single theme has emerged as dominant, it is this: The most complex problems will arise in those industries in which deregulation is partial. In these cases, the social control of an industry creates policy problems that may find neither mutual exclusion nor collective exhaustion in the course of regulation and antitrust. In short, there is a danger that regulators and antitrust enforcers will fight over jurisdiction in some important matters, while other important problems receive the attention of neither.

One important policy question that will have to be settled is whether existing agencies have the requisite jurisdiction and powers to create an effectively competitive environment. If not, perhaps existing agencies will require new authority, more resources, and new remedies. While it is not currently obvious whether and to what extent such changes will be needed, the paper points to the types of questions that will be most important in signaling the need for change. First, who will scrutinize the price at which one firm sells goods or services to an affiliated firm, especially when only one of the firms is regulated? Second, who will determine when a price in a regulated market is predatory or otherwise anticompetitive, particularly if the firm charging that price also serves a regulated market? Third, what will be the boundary of antitrust immunity? Fourth, under what conditions will

[63] For a discussion of transportation companies, see Friedlaender (1969), pp. 155–162.

a merger involving a firm serving both regulated and unregulated markets be allowed, and who specifies these conditions? Fifth, have regulations created structural conditions and practices that contribute to consciously parallel behavior? And sixth, who will decide when a refusal to serve is illegal, particularly if the sale in question involves both regulated and unregulated firms? The text suggests circumstances in selected industries under which each of these dilemmas might actually occur.

Whether deregulation is partial or not, the first task of antitrust enforcers will be to determine whether structural change is required to prevent the exercise of unchecked economic power by firms now unaccountable to regulators. Structural change may be necessary to foster competitive markets, particularly where regulation has created highly concentrated markets. Antitrust enforcers must expect two new types of defense: (1) That the large market shares for which structural relief is sought were thrust upon existing firms by regulators, and (2) that structural relief is unnecessary, since deregulation will by itself naturally erode the market shares of larger firms.

The paper also emphasizes that antitrust enforcement will encounter a number of practices antithetic to the creation and maintenance of competitive markets, practices that are deeply ingrained in the fabric of the industries being deregulated. It will not be easy for antitrust to overcome the inertia of decades of sanctioned collusion and monopoly. Some of the institutions at the heart of the regulated system must be eliminated with deregulation, including domestic rate bureaus in transportation industries where price and entry are decontrolled. Even then, antitrust enforcers must watch closely to insure that behavior is independent, especially where other institutions, such as international conferences, continue to exist.

Finally, antitrust enforcers must constantly watch for any obstacles that impede free entry where free entry is desirable. Deregulation by fiat does not guarantee free entry in fact.

Bibliography

Areeda, P., *Antitrust Analysis*, Little, Brown, and Company, 1974.

——— "Antitrust and Utility Regulation," *The Bell Journal of Economics and Management Science*, Vol. 3, No. 1, (Spring 1972), pp. 42–57.

Averch, H., and Johnson, L. "Behavior of the Firm Under Regulatory Constraint," *American Economic Review*, December 1962, *52*, pp. 1053–69.

Bork, Robert H., *The Antitrust Paradox*, Basic Books, 1978.

Braeutigam, R., "An Examination of Regulation in the Natural Gas Industry," in *Study on Federal Regulation*, United States Senate Committee on Governmental Affairs, appendix to Vol. VI (1978), pp. 659–727.

Dayan, D., *Vertical Integration and Monopoly Regulation*, Ph.D. dissertation, Princeton University, 1972.

Demsetz, H., "Why Regulate Utilities?," *Journal of Law and Economics*, April 1968, *11*, pp. 55–65.

Ferejohn, J., Forsythe, R., and Noll, R., "An Experimental Analysis of Decision-making Procedures for Discrete Public Goods: A Case Study of a Problem in Institutional Design," California Institute of Technology, Social Science Working Paper 155, October 1977.

Friedlaender, A., *The Dilemma of Freight Transport Regulation*, Brookings, 1969.

Hong, J. T., and Plott, C. R., "Implications of Rate Filing for Domestic Dry Bulk Transportation on Inland Waters: An Experimental Approach," California Institute of Technology, Social Science Working Paper 164, July 1977.

Isaac, M., and Plott C. R., "The Opportunity for Conspiracy in Restraint of Trade: An Experimental Study," California Institute of Technology, Social Science Working Paper 255, April 1979.

Joskow, P., and Noll, R., "Regulation in Theory and Practice: An Overview," California Institute of Technology, Social Science Working Paper 213, May 1978.

Kahn, A. E., *The Economics of Regulation* Vol. I (1970) and Vol. II (1971), Wiley.

Kaysen, C., and Turner, D., *Antitrust Policy*, Harvard University Press (1959).

Keeler, T., "Domestic Trunk Airline Regulation: An Economic Evaluation," in *Study on Federal Regulation*, United States Senate Committee on Governmental Affairs, appendix to Vol. VI (1978), pp. 75–160.

Lieb, R., *Transportation: The Domestic System*, Reston Publishing Company, 1978.

Montgomery, W. D., "Case Study of Regulatory Programs of the Federal Energy Administration," in *Study on Federal Regulation*, United States Senate Committee on Governmental Affairs, appendix to Vol. VI (1978), pp. 731–833.

National Commission for the Review of Antitrust Laws and Procedures, *Report to the President and the Attorney General*, January 22, 1979.

Nelson, J. R., "Abstract of Regulation of Overland Movements of Freight," in *Study on Federal Regulation*, United States Senate Committee on Governmental Affairs, appendix to Vol. VI (1978), pp. 1–92.

Owen, B., "Cable Television: The Framework of Regulation," in *Study on Federal Regulation*, United States Senate Committee on Governmental Affairs, appendix to Vol. VI (1978), pp. 347–389.

——— and Braeutigam, R., *The Regulation Game: Strategic Use of the Administrative Process*, Ballinger (1978).

Phillips, A. (ed.), *Promoting Competition in Regulated Markets*, Brookings, 1975.

Plott, C. R., and Smith, V. L., "An Experimental Examination of Two Exchange Institutions," *The Review of Economic Studies*, *XLV* (1), February 1978, pp. 133–153.

Posner, R., "Statistical Summary of Antitrust Enforcement," *Journal of Law and Economics*, October 1970.

Roberts, P., "Some Aspects of Regulatory Reform of the U.S. Trucking Industry," *Motor Carrier Economic Regulation*, National Academy of Sciences (1978), pp. 470–502.

Scherer, F. M., *Industrial Market Structure and Economic Performance*, Rand-McNally, 1970.

Stelzer, I., *Selected Antitrust Cases: Landmark Decisions*, fifth edition, Irwin, 1976.

Stoll, H., "Regulation of Securities Markets: An Examination of the Effects of Increased Competition," in *Study on Federal Regulation*, United States Senate Committee on Governmental Affairs, appendix to Vol. VI (1978), pp. 589–656.

United States Senate Committee on Governmental Affairs, *Study on Federal Regulation* (1978), Volumes V, VI, and Appendix to Volume VI.

Weiss, L., *Economics and American Industry*, Wiley (1967).

——— and Strickland, A., *Regulation: A Case Approach*, McGraw-Hill (1976).

THE UNDERGROUND ECONOMY: ESTIMATES OF SIZE, STRUCTURE, AND TRENDS

By Carl P. Simon and Ann D. Witte*

CONTENTS

	Page
Introduction	71
I. Unreported production and trade of legal goods and services	73
A. Unreported income due to attempts to evade taxes or benefit loss	73
(i) Size and trends	73
(ii) Costs and benefits of tax evasion	75
(iii) Suggestions for research and policy	77
B. Nonreporting due to seller's status	78
(i) Size and trends	78
(ii) Cost and benefit to illegal immigration	79
(iii) Suggestions for research and policy	82
C. Nonreporting due to illegal status of goods	83
(i) Size and trends	83
(ii) The costs and benefits to individuals and organizations directly involved	85
(iii) Suggestions for research and policy	86
II. Production and distribution of illegal goods and services	87
A. Illegal gambling	88
(i) Size and trends	88
(ii) Suggestions for research and policy	91
B. Arson	93
(i) Size and trends	93
(ii) Suggestions for research and policy	95
C. Heroin	96
(i) Structure of the industry	96
(ii) Size and trends	98
(iii) Costs to society	100
(iv) Other drugs	101
(v) Suggestions for research and policy	102
III. Summary and conclusions	103

APPENDIXES

A. The method of estimating income unreported for tax evasion purposes	113
B. The method of estimating the national income of the stolen goods market	116
Bibliography	110

*We would like to acknowledge the financial and other support received from staff members of various congressional committees. In particular, we would like to thank Douglas Ross and Michael Lockerby of the Joint Economic Committee and Robert Cozart of the House Ways and Means Committee. We would also like to thank the many other individuals both in the public and private sector who have helped us extensively, especially Phillip Cook of Duke University. The statements and conclusions in this paper are, of course, the authors alone and should not be taken to reflect those of the Joint Economic Committee.

INTRODUCTION

In any study of economic change, the phenomenon of the underground economy is an intriguing and important factor to consider. Numerous articles on this mysterious economy have appeared in recent years. Some authors claim that the underground economy is now approximately 10 percent of reported GNP and growing rapidly. Although most authors compare the size of the underground economy with widely published measure of reported economic activity such as GNP, we know of no previous effort that actually uses the same national income accounting methods to compute the size of underground activity as one uses to compute reported economic activity. We make such an effort in this report.

Of what does this underground economy consist? Definitions vary, but it is generally agreed that the economy has two major divisions—production and distribution of illegal goods and services, such as drugs, and the nonreporting of transactions in legal goods and services such as tax evasion. Researchers have only recently tried to estimate the size of this economy and have come up with widely different estimates—$65 to $176 billion in 1976. Reports indicating an underground economy of this size caused understandable concern among those concerned with tax revenue. Recently, at the request of the House Ways and Means Committee, the Internal Revenue Service has estimated that between $100 and $135 billion of taxable income was not reported by individuals on their tax returns in 1976. As a result, the IRS claims that the Federal Government lost $19 to $26 billion in income taxes.

Why have estimates of size varied so widely? The basic reason is twofold. First, research on the underground economy is in its infancy and methods of estimating its size vary markedly. Second, different researchers estimate different measures of size. For example, while the Internal Revenue Service estimates taxable income, other researchers estimate "GNP".[1] There is no reason why estimates of different variables even for a single year should be the same. It is the purpose of this paper to examine the underground economy and attempt to estimate its size by use of the now well-developed national income accounting techniques. These techniques suggest that we estimate the level of output of any sector by two methods (summing the value of all final goods and services—the value added approach—and summing the values of income received for productive services—the incomes approach) and attempt to reconcile any differences.

In this report, we will only be able to use a single approach (generally the incomes approach) to estimate the size of any given sector of the underground economy. It is important to note that GNP and National Income figures include only income earned for productive, generally market oriented, activities.

---

[1] Interestingly, IRS, like a number of other researchers, estimates different measures of income for different sectors, and then adds the noncomparable results obtained. IRS makes extremely careful and valuable estimates of legal source "taxable income" which is not reported to IRS. IRS then estimates income from illegal sources using the "value added" approach of national income accounting. As noted below the two "incomes" are quite different, and cannot be meaningfully combined.

They do not include transfer items such as the interest on government bonds or welfare payments nor do they include gains from the sales of stocks, bonds, mortgages, and similar debt and equity instruments. Note that these latter payments would be included in taxable income. Further, standard national income accounting techniques exclude most home production (e.g., housecleaning, painting one's own house) from National Income. To make our estimate of National Income for the underground economy consistent with National Income estimates for reported activities, we will employ these same types of accounting rules. For example, when considering the stolen goods market, we include the income of fences in our estimate of National Income because the fence is being paid for services rendered. However, we do not include the value of goods stolen and kept by the thief for his own use: a type of nonmarket production. Nor do we include the value of money stolen by thieves: a forced transfer. Furthermore, in our discussions of tax evasion, we do not include in our estimates unreported income which was derived from sources such as alimony payments, tax refunds, or capital gains. As a result of this approach, our estimates tend to be lower than other estimates of the amount of illegal activity but hopefully offer more valuable comparisons with the reported National Income.

Specifically, in this paper, we will attempt to estimate carefully the size of six sectors that we believe either dominate or represent important trends in the underground economy. These sectors are the production and distribution of heroin, the provision of gambling services, the destruction of building and equipment for profit (fraud arson), tax evasion, illegal aliens and stolen goods markets. According to the 1967 President's Commission on Law Enforcement and Administration of Justice, these areas would encompass virtually all of the underground economy.

As we studied these six sectors, it became increasingly obvious that, unlike the situation in the legitimate economy, there is limited economic interaction among the sectors of the underground economy. For example, the output of one sector is rarely an input to another. However, some types of interaction are present. For example, most individuals who produce illegal goods and services do not report their "illegal" income on their tax returns and thus are part of the tax evasion sector. Many addicts support their habits by stealing property which is ultimately fenced on the stolen goods market. However, the major thing which these six sectors seem to have in common is their illegality. As a result, there is some artificiality in using the phrase underground "economy."

In the first part, we examine the first three of these sectors which fall in the division—unreported transactions. In the second part, we examine the last three sectors which fall in the division—production and distribution of illegal goods and services. We estimate the size (National Income) and trends in each of these sectors, and wherever possible, we consider costs and benefits from the point of view of the individuals involved, the government, and society as a whole. Finally, we offer suggestions for policy and research in each area.

The final part of the paper contains our summary and overall suggestions. In this part, we estimate the size of the underground economy

based on the previous two parts and other available evidence. We contrast this estimate with previous estimates and suggest reasons for differences. Finally we make suggestions for research and policy which cut across the individual sectors studied.

I. Unreported Production and Trade of Legal Goods and Services

This part of our paper examines the production of legal goods which are not properly reported while part II examines the production of goods and services which we as a society have declared illegal. The first section below deals with the most important area of non-reporting and indeed the dominant section of the underground economy—tax evasion and the evasion of "public benefit loss," such as social security and AFDC. While undoubtedly the vast majority of income earned in producing the illegal goods and services discussed in the next part goes unreported, we, like IRS, do not believe that current estimates of tax evasion include much of such income. This is true because current estimates of incomes escaping tax rely mainly on "paper trails" (various non-tax sources of income reporting) for discovery. Incomes earned in the production of illegal goods and services generally leave no such "trails," and, thus, are not generally included in estimates of tax evasion. This is important since if substantial amounts of income from illegal production were included in tax evasion estimates, we would be "double counting" underground National Income when we add estimates of the income escaping tax to the estimates we obtain in the next part.

Later sections will treat non-reporting due to the illegal status of either the seller or the property sold. Although such non-reporting often involves tax evasion, tax evasion is usually not the main reason for non-reporting in these latter two areas. In addition, incomes not reported in these latter two areas, like incomes earned by producing illegal goods and services, often leave no "paper trails" and so are not completely included in current estimates of tax evasion.

In each of the areas discussed below, we will consider not only the size and trends of the activity but also the costs and benefits of the activity for the individuals directly involved in the transaction, for government units, and for society as a whole. Each section will end with suggestions for potentially fruitful topics for policy consideration and for future research.

A. Unreported Income Due To Attempts To Evade Taxes or Benefit Loss

(1) SIZE AND TRENDS

In estimating the size of this sector, one is not interested in tax evasion or benefit fraud per se. Rather one is interested in the amount of income which is underreported because of tax evasion and benefit fraud efforts. With this in mind we will consider the benefit and income tax programs with the broadest possible coverage. These are the Federal Income Tax, Social Security, the major national welfare programs

(AFDC and Food Stamps) and unemployment compensation program. Theoretically, there should be some overlap here with underreporting of income for benefit fraud purposes often reflected in income tax evasion figures. In practice, we are not at all convinced that current estimates of tax evasion reflect the majority of the income which is not reported for benefit fraud purposes. However, conversations with personnel of the Quality Assurance Program of the Social Security Administration have indicated that no estimates have been made of the extent of the income which is not reported for purposes of benefit fraud. Consequently, we will have to rely on two sources to estimate the extent of income underreporting in this section: (1) Information compiled by the IRS and GAO in connection with attempts to measure taxpayer compliance; and (2) differences between estimates of adjusted gross income contained in national income accounts and reports of adjusted gross income on individual tax (IRS) returns.

Using these two sources and methods detailed in appendix A, we estimate that between $60 and $75 billion of income earned by legal activity was unreported for tax evasion purposes in 1974. Further, we believe, for reasons detailed in appendix A, that figures in the upper range of our estimates are likely to be more accurate than those in the lower range. If pressed to narrow the range of our estimates, we would estimate that between $70 and $75 billion of income from legal sources was unreported primarily for tax evasion purposes in 1974.

Using limited time series data available from IRS's Tax Compliance Monitoring Program (TCMP) and more extensive time series data available from the Department of Commerce's Bureau of Economic Analysis (BEA), we estimate that unreported income for tax evasion purposes grew at an average annual rate of 3 to 11 percent in the 1965 to 1969 period, 8 to 10 percent in the 1969 to 1973 period and by approximately 5 percent in the 1973 to 1976 period. These estimates are quite similar to the growth rate of reported personal disposable income for comparable periods. Comparing the growth rates of reported disposable personal income to those of income unreported for tax evasion purposes, we conclude that this sector of the underground economy may have grown slightly more rapidly than reported income in the 1965–69 period, at approximately the same rate in the 1969 to 1973 period and possibly somewhat more slowly than reported income in the 1973 to 1976 period. However, we, like IRS in its report, can make no definitive statements concerning recent trends.

As a whole, our results indicate that the failure to report income for tax evasion purposes is a formidable problem and, indeed, that this sector of the underground economy accounts for the bulk of underground activity. However, it is wise to put the size and growth of this sector of the underground economy in proper perspective. Even if we accept the larger of our estimates ($70 to $75 billion) for the size of this underreporting in 1974, unreported income for tax evasion purposes was only approximately 6 percent of reported personal income in that year. Moreover, it appears that the size of this sector of the underground economy has only grown moderately relative to reported economic activity during the last decade and may have actually declined relative to reported activity during the mid 1970's.

(II) COSTS AND BENEFITS OF TAX EVASION

Economists have been actively building models detailing the costs and benefits of tax evasion to the taxpayer evading taxes. Most of these models define the costs and benefits involved quite narrowly. According to these models, the benefit to the taxpayer who successfully avoids taxes is the increase in personal wealth. The cost to the taxpayer who fails is the penalty imposed. Other economists have recently suggested that one must consider the moral and ethical position of the taxpayers as well as the monetary gains and losses involved. This suggestion seems particularly relevant in light of the finding of Schwartz and Orleans (1967). These authors in cooperation with the IRS were allowed to conduct an experiment to determine if sanctions or moral appeals were more effective in improving tax compliance. The results of the experiment indicate that moral appeals may be more effective than sanction threats in obtaining tax compliance.

A second line of research, simulation studies, provides further interesting insights into the way in which individuals react to different tax strategies. A recent study (Friedland, Maital and Rutenberg, 1978) found that "the fraction of earned income reported becomes very elastic with respect to the tax rate". In other words, as tax rates become higher and higher, the fraction of income unreported increases even more quickly. Furthermore, these researchers found that large fines tend to be more effective deterrents than frequent audits. This is contrary to IRS's beliefs: "IRS considers the audit of returns to be the greatest stimulus to voluntary compliance" (Comptroller General of the United States, 1976a, p. 2). Finally, this study found that the determinants of the decision to underreport and the actual amount of underreporting are quite different and that personal characteristics are important in the determination of both decisions. This latter conclusion is further supported by preliminary IRS research and by sample surveys. In the most recent survey finding for the U.S., Spicer and Lundstedt (1976) found that perceptions of inequity, the number of tax evaders known personally, and a previous audit experience tended to increase tax evasion. This latter finding is quite startling. Spicer and Lundstedt hypothesize that this may be the result of a negative reaction to the audit experience. Strümpel (1969) has also noted that stringent assessment may lower compliance and willingness to cooperate. This seems to contradict the results of a recent GAO survey which indicated that 70 percent of the audited taxpayers surveyed reacted favorably to their audit experience. At the least, the above points up the need for additional research on the determinants of the extent and amount of tax avoidance. This is also a recommendation of a recent GAO study of IRS (Comptroller General of the United States, 1976a, p. 56). The IRS recent report is an excellent first step in this direction.

The non-personal factors that make tax evasion more or less appealing vary markedly by source of income. Both the size of evasion possible and the probability of detection are important variables. Wages and salaries subject to withholding are generally quite fully reported, while other sources of income not subject to withholding are more incompletely recorded. Information from IRS's TCMP effort found rates

56-368 O - 80 - 6

of voluntary compliance varying from 99.6 percent for wages to 67.2 and 79.9 percent for farm and non-farm business income, respectively. Voluntary compliance for rental income was even lower, 45.7 percent.

A relatively recent development with major tax evasion possibilities is the large scale reintroduction of barter in the United States. The reintroduction which goes under the formal name of the "reciprocal trade business" facilitates the exchange of large quantities of goods and services between businesses and individuals. The size of this sector is difficult to determine, but *Purchasing World*, a trade publication, estimates that 48 percent of purchasing agents in the United States engaged in some form of barter. The service sector is the major user of reciprocal trade services although manufacturers are increasingly utilizing this sector for stock liquidation purposes. The self-proclaimed largest firm in this new and growing industry is Atwood Richard, Inc., which estimates that it handled over $100 million worth of goods and services during fiscal year 1976–77. While large firms such as Atwood Richard's undoubtedly pay all taxes due, the growth of the industry and participation of individuals as well as firms in the business make one suspicious that one of the attractions of this area is the ease with which such transactions can be kept from the eye of prying tax agents. While the emergence and growth of this sector is cause for concern, we believe that at present barter transactions account for less than 5 percent of tax evasion.

The benefits to the taxpayer of avoiding taxes vary directly with the tax rates applicable on the income evading taxes. The increasing progressivity of tax rates coupled with the movement of families into higher marginal tax rate brackets as a result of inflation have undoubtedly increased incentives for tax evasion for many individuals. As mentioned above, research seems to indicate that the rate of tax avoidance increases at an increasing rate with the overall tax rate. This provides an additional argument against very high tax rates. Not only do they decrease the incentive to work and thus GNP, but they also may lead to rapid declines in tax compliance rates.

The cost of tax evasion to the government is the loss of revenue plus the cost of compliance programs. The revenue loss to the Federal Government will depend on the amount of income that is avoiding taxes and the applicable tax rate. IRS estimates (U.S. Department of the Treasury, 1979) that underreporting and failure to report legal income resulted in between $13 and $17 billion income tax revenue loss in 1976. Compliance costs added approximately $1 billion to this cost in 1976. Thus costs to the Federal Government of income tax evasion was probably around $16 billion in 1976, an amount which is only slightly below the Federal budget deficit in 1974.

The social costs of tax evasion are far larger and much more difficult to judge. As a result of tax evasion, taxpaying citizens must either pay higher taxes or forego the public services that would otherwise be available. Assuming that tax evaders like those who comply with tax laws are equally members of "society," this is not a social cost, but only a transfer. Tax evasion does generate efficiency loss, but this loss comes through effects on net work incentives, the cost of compliance programs, the cost of evasion itself, and most importantly the effect on

social mores. As the President's Commission expressed it, tax evasion affects the "moral climate of our society" (President's Commission on Law Enforcement and Criminal Justice, 1967, p. 104). Attitudes toward taxes in some European countries (e.g., France, Italy) warn us that public attitudes toward taxes are extremely important.

(III) SUGGESTIONS FOR RESEARCH AND POLICY

One is struck by the dearth of good, empirically based, policy oriented research on unreported income and tax evasion in the United States. We were able to find only two such studies in our literature review (Schwartz and Orleans, 1967; and Groves, 1958). The first of these studies which was discussed above was conducted in cooperation with IRS while the latter was conducted in cooperation with the Wisconsin Department of Revenue. The lack of such empirical work in recent years probably is due to the increased stringency of privacy statutes, and increased sensitivities of taxing agencies. While the secrecy of individual returns must be maintained, we feel strongly that much additional research could be usefully carried out. This research could be based on aggregate data or on individual data from which all personal identifiers had been deleted. We believe that it is important that such research be possible for independent researchers as well as IRS and its contractors. IRS is currently conducting a compliance study although we were unable to obtain any results.

We would like to voice our support for two recent GAO suggestions. The first suggestion is that IRS "expand and accelerate its research into factors which influence compliance" (Comptroller General of the United States, 1976a, p. 56). This research should carefully assess the rather extensive theoretical (psychological, sociological and economic) literature and be careful to specify models as completely as possible. This will require integration of census and other data sources. The second suggestion is that IRS "initiate action to periodically estimate the size and analyze the characteristics on the non filer population". IRS data matching programs with the Social Security Administration and State and local government agencies would be one potential source of data. Alternatively, IRS could attempt an estimate tax evasion for a "typical area" using the total population lists compiled by the Bureau of Census in connection with the 1980 census.

We would like to suggest that future IRS research work be subject to extensive external review and comment. In the past IRS research has been circulated only to a limited extent although it generally contains no individual data. We feel that increased external review could improve both the quality and breadth of IRS research.

Existing theoretical research on tax evasion has tended to be quite narrow in perspective and could be usefully broadened. Integration of insights from more traditional deterrence work and sociological and psychological theories could prove most useful. Recent work by Spicer and Lundstedt (1976) discussed above is interesting in this regard.

Much of the research cited above has potential policy implication. Perhaps, most importantly this material seems to suggest that a

broader range of policy instruments be used to encourage tax compliance. Education and moral appeal may have as great if not greater effect than high audit rates and penalties. The negative effect of audits on taxpayers' attitudes toward taxes is particularly interesting in this regard. Improving the public image of government may also prove a relatively cheap method of improving compliance. Finally, tax form simplification, which IRS is currently exploring, could substantially increase reporting by low income groups. Optimal enforcement policies should be broadly and carefully assessed.

Another set of policies with great potential is the expansion of withholding and reporting at source (i.e., by the payers of the income). When reporting at source for incomes and dividends was begun in 1964, reported income from these sources increased by 45 percent. The marked jump in unreported interest income noted above may make this an area ripe for withholding or at least closer computer checking of reported information on interest income. Before recommending such increased withholding, the Government should consider both the costs (e.g., increased recordkeeping) and benefits involved. A good example of the type of study needed is a recent Joint Committee on Taxation report on independent contractors (1979). This report concluded that "it is doubtful whether additional tax revenues obtained from nonfilers would justify the administrative complexity and expense which withholding on self-employment earnings would entail" (p. 38).

Another intriguing policy alternative is novel sentencing practices. For example, part of the sentence imposed in a recent antitrust case against firms and individuals involved in price fixing in the paper label industry was that individuals involved make speeches to public bodies concerning the nature of their offenses. A similar sentence is imposed on those convicted of tax fraud cases in Germany. When one considers the "respectability" of many tax evaders, this type of sentence may have large deterrent effects.

B. Nonreporting Due to Seller's Status

(1) SIZE AND TRENDS

In this section, we discuss the underreporting of income by those illegally in this country. Estimates of the number of illegal aliens in this country have varied markedly. Lancaster and Scheuren (1977) estimated that there were 3.9 million undocumented aliens in the age group 18 to 44 years old (prime working years) in April 1973. Other studies have suggested that there was no increase in the size of the undocumented population employed in the non-agricultural sector between 1969 and the mid 1970's. Although estimates of the size of the undocumented population vary considerably with the underlying death rate assumed for undocumented persons, all such careful estimates calculated fall below four million for 1975.

Only those non-resident aliens who are working and whose product or income is not reported form part of the underground economy. No reliable estimates of the number of illegal aliens working in the United States are currently available, but some educated "guesses" have been

made. As can be seen in table 1, employment of illegal aliens appears to be concentrated in the agricultural, service and light industries—with the most rapid growth in light industry.

It is not known what proportion of the earnings of the employed illegal aliens goes unreported and thus adds to the size of the underground economy. IRS estimates that approximately 25 percent of illegal aliens are part of the underground economy. Using the Social Security's Exact Match File and information reported in documents filed by employers, IRS (Department of the Treasury, 1979) estimates that illegal aliens earned between $5 and $6.6 billion of unreported income in 1976. This income of illegal aliens is included in our estimates of tax evasion (see section I–A above). Our reading of the literature on illegal aliens lead us to believe that neither the Exact Match File nor employer documents completely capture the income of illegal aliens. Having no solid evidence on which to base an estimate of the income of aliens who are completely off the books, we will assume that it was approximately $1 billion in 1974.

TABLE 1.—ESTIMATES OF EMPLOYED ILLEGAL ALIENS BY CATEGORY OF EMPLOYMENT

Industry	Number estimated in 1974 [1]	Number estimated in 1977 [2]	Percent increase 1974–77
Agriculture	335,000	1,200,000	258
Heavy Industry	105,000	176,000	68
Light Industry	214,000	1,000,000	367
Service	301,000	990,000	229
Construction		300,000	

[1] Estimate submitted to Congress by Leonard F. Chapman, Commissioner, Immigration and Naturalization Service (INS), Sept. 18, 1974, as reported in "Special Report: Illegal Immigration," The Environmental Fund, 2 (November 1978), p. 7.
[2] Estimated by Janet Graham, Public Affairs Office, INS, February 1977 as reported in "Special Report . . ."

Given the estimated increase in illegal alien employment between 1974 and 1977, it seems likely that this portion of the underground economy has grown markedly in recent years. Indeed the "best guesses" in table 2 indicate an average annual growth rate in employment of illegal aliens of approximately 50 percent in the 1974 to 1977 period. This growth rate should be viewed with considerable caution since much more carefully done population studies show much slower rate of increase in the illegal alien population.

(II) COST AND BENEFIT TO ILLEGAL IMMIGRATION

The benefits to the illegal alien of participation in the U.S. underground economy depend on a number of factors. Consider how a rational individual might view the migration decision. Such an individual would first compare the net income streams (s)he might receive in his or her native country and in the United States. The relative value of these streams would depend upon the average income and unemployment rates which the individual would face in the two countries as well as movement costs and costs of living differentials between the two countries. This rational, highly simplified, model would lead us to expect more immigration to the United States the larger the amount by which U.S. wages exceed wages in the potential immigrant's

native country, the larger the amount by which the immigrant's native country's unemployment rates exceed those in the United States, the closer the immigrant's native country was to the United States (and thus the smaller moving costs) and the smaller the differences in cost of living between the two countries. Although the rational model is unrealistic in many ways, its predictions have generally been confirmed. In addition, the migration literature has indicated that migration from any country increases with the size of the migrant stock already in that country. Researchers have suggested that this effect is due to improved knowledge of opportunities as well as the reduced psychic and real costs of movement when friends and relatives are in the country to which one immigrates. Thus, one would expect more immigration to the United States from countries where large numbers of individuals have migrated to the United States in recent years.

To continue our model of rational migration behavior, the individual who has decided that immigration is potentially beneficial has to next decide whether legal immigration is possible and what are the costs involved. In general, legal immigration is most feasible for highly skilled individuals from Western European nations. Immigration for other nationalities and for less skilled workers has been increasingly restricted. For example, changes in the immigration laws in 1976 reduced the number of immigrant visas available to Mexican nationals from 62,205 in fiscal year 1975 to 57,863 in fiscal year 1976, and 44,079 in fiscal year 1977 (Jasso, 1979, p. 3). This is clearly a case where government regulation has increased the size of the underground economy.

Given that immigration is potentially beneficial and that legal immigration is perceived as impossible or too costly, the rational individual will consider illegal immigration. The net present value of the income available in the United States for such individuals is reduced by the decrease in wages due to wage discrimination against illegal aliens in the United States and by costs involved in avoiding detection and deportation. In addition, the income of the individual in any period after immigration is now subject to additional uncertainty. This additional uncertainty is due to the fact that there is some possibility that in any given period the individual will be deported and thus not earn the income expected in the United States.

Once the immigrant has decided upon illegal immigration, (s)he faces a final decision—whether to participate in the regular or underground economy. The choice is constrained (like the type of immigration decision) by the employment opportunities available to the illegal alien. From the illegal alien's point of view, participation in the underground economy offers a number of advantages. First, the probability of deportation should be lower for individuals in the underground economy than those in the regular economy since records are minimal and the employer as well as the employee has an incentive to maintain secrecy.[2] Second, illegal aliens may forego payment of taxes and other statutory deductions. The evasion of such tax is often desirable to the illegal alien as (s)he will probably not be able to enjoy many of the public services made available by such taxes due to the fear of discovery and deportation.

---

[2] Since there is currently no law in United States that prohibits a person from hiring undocumented aliens the incentive to secrecy in the regular sector of the economy is much less.

From the employer's point of view, not reporting illegal alien employees has become increasingly attractive in recent years. By not reporting such hirings the employer can forego employer contributions to private health and retirement plans as well as social security, unemployment insurance, and workmen's compensation payments. The size and coverage of these payments have grown markedly in recent years making evasion increasingly attractive. The potential benefits to the employer of not reporting employment of illegal aliens is higher than for regular employees because detection is less likely due to the alien's desire for anonymity. Here again government action that increases required benefit payments and broadens the coverage of such payments stimulates the growth of "off-the-book" employment including that of illegal aliens. Further government regulations of hours, wages, and working conditions which increase employer costs will also stimulate such "off-the-books" employment.

The costs to the U.S. Government of illegal immigration are hard to judge. Those illegal aliens who work in the regular economy pay all normal taxes. However, illegal aliens are less likely to take advantage of many of the services provided by these taxes due to fear of discovery and deportation. For example, the Quality Assurance Program of the Social Security Administration has estimated that the use of such public assistance programs as Aid to Families of Dependent Children (AFDC) and Supplemental Security Income (SSI) by illegal aliens is negligible (Jasso, 1979b, p. 12). After surveying a number of empirical studies the House Select Committee on Population concluded that there was only "a moderate level of utilization of free medical services by illegal immigrants" (Report of the House Select Committee on Population, 1978, p. 35.) In addition, illegal aliens are less likely than other individuals to file tax returns due to fear of discovery. Given the generally low wage employment under consideration this probably results in an overpayment of taxes due in most instances.

Illegal aliens working in the underground economy, on the other hand, are probably a net drain on U.S. Government bodies since they would use some public services (e.g., schools, transportation) but would generally pay only sales tax, user charges, and property tax.

Another cost to U.S. society from the work of illegal aliens in the underground economy in addition to the likely tax burden noted above is the displacement of U.S. workers as a result of competition with illegal aliens. We were unable to find any reliable estimates of either of these effects although INS has developed a number of theoretical approaches to the problem.

Similar questions have been addressed in at least three other settings: (1) Legal immigration to the United States; (2) the brain drain from less developed countries; and (3) the effect of intranational migration on regional disparities. In all of these areas, both theoretical and empirical conclusions have been ambivalent. Immigration may be beneficial to native residents if increases in demand and the effect of economies of scale (effects which tend to lower costs with increased output) outweigh increased labor supply effects. It may be detrimental if the reverse is true. Since many illegal aliens in the United States send large proportions of their earnings to friends and relatives in their native country and live quite meagerly themselves, increases in demand for U.S. products due to illegal immigration are

likely to be small. Further, most employment of illegal aliens tends to be concentrated in industries (see table 1) with few economies of scale. Thus it seems likely that the net effect of illegal immigration on U.S. citizens is probably detrimental. The wages of unskilled U.S. workers are probably depressed as a result of this immigration. U.S. consumers may benefit somewhat through lower prices for goods (e.g., produce, clothing) produced by illegal aliens. However, in the absence of perfect competition such price effects are likely to be smaller than the effect on unskilled workers wages. In addition, illegal immigration may serve to transfer income from relatively poor, unskilled U.S. workers to relatively well off U.S. consumers and producers.

(III) SUGGESTIONS FOR RESEARCH AND POLICY

Reviewing the research literature on illegal aliens one is struck by its comparatively recent vintage. Most work has been done in the last few years and has been concerned with estimating the size of the illegal alien population. This work appears to be of high quality and is being pursued relatively vigorously.

The breadth of research on illegal aliens could be usefully broadened. Existing migration models could be usefully married to models of the decision to migrate illegally. Such models would have to take into account changes in immigration policy and changes in tax-benefit ratio for illegal aliens.

Studies of the nature of labor markets for both illegal aliens and the legal aliens they resemble might prove quite illuminating. Such research could usefully be combined with current efforts to obtain better and more frequent estimates of the size of the illegal alien population.

Migration policies should be formulated with full consideration of their likely effect on the size of the illegal alien population. Less stringent migration policies combined with increased enforcement of illegal alien statutes might be quite beneficial.

Policies for dealing with illegal aliens already in this country are more difficult to formulate. A policy designed to uncover and deport large proportions of this population would seem socially disruptive as well as costly. Perhaps, legitimizing the long term population, as suggested by President Carter, may be the best policy. This policy could be combined with stringent enforcement of laws against more recently arrived illegal aliens. Such enforcement might be carried out jointly with stepped up enforcement of the minimum wage laws, social contribution statutes (e.g., social security) and IRS action against non-filers.

Enforcement efforts should be directed at employers who hire illegal aliens and the individuals who provide passage, documentation and connections as well as the aliens themselves. A Federal law may well be useful to make the known hiring of illegal aliens illegal. The burden of performing a reasonable check on status could usefully be put on employers.

The two-pronged enforcement effort outlined above should also serve to stem the flow of new aliens by making it both less lucrative and more costly to immigrate illegally. If the prospect of legal immi-

gration is also increased, a rather substantial dent might be made in the illegal immigrant flow. Potential benefits of development aid which generates new jobs in developing countries might be increases in wage rates and decreases of unemployment rates in countries from which many aliens come. Such positive changes in proximate developing countries should at least slow the flow of illegal aliens.

In making decisions to increase payment levels and to expand the coverage of social insurance, policymakers would be wise to consider the fact that such increases and expansions provide incentives for both employers and employees to go "off the books." In addition, stringent health and safety regulations and wage and hours laws set up similar incentives.

C. Nonreporting Due to Illegal Status of Goods

In this section we will consider the failure to report transactions in goods because of the status of these goods—particularly stolen goods. Although closely related to transactions in illegal goods and services, the goods and services discussed in this section are different because it is only their status which makes the transactions illegal and not the nature of the goods themselves.

(I) SIZE AND TRENDS

There is a large number of transactions of stolen goods in the United States at both the wholesale and retail levels. At the wholesale level, most activity is carried out by large rings of organized professional thieves who specialize in a particular type of goods, such as appliances. It is often alleged that many of these large specialized theft and distribution operations are dominated by organized crime and have developed complex organizational structures somewhat analogous to the structures found in the heroin industry. At the retail level, diversity appears to prevail with operations running from the thief fencing his own stolen goods to highly specialized and sophisticated international operations involved in the sale of jewelry and art objects.

It is very difficult to obtain estimates of the size of this sector of the underground economy because of the limited amount of previous research and the diversity of the sector. Some economists would not even count this sector as part of the underground economy (e.g., Henry, 1976) because this sector adds no value, but rather represents a mechanism for illegal transfer of ownership. While we are sympathetic to this position and, indeed, adhere to it when estimating the social costs of stolen goods markets, we believe that the incomes earned in this sector should be included in the "National Income" of the underground economy. We believe that the correct analogy with the reported economy would be with government workers employed in welfare agencies. The National Income figures reflect the incomes of such workers and by analogy we believe the "National Income" figures for the underground economy should include the incomes of those involved in the stolen goods markets. Reselius and Benton (1973) suggest an alternative reason for including stolen goods markets in underground "National Income". They see the thief's operations as akin to legitimate production and the fence's operations as akin to legitimate retailing and wholesaling.

Our calculations of the National Income of the stolen goods market are consistent with this approach. However, after extensive discussions about national income accounting principles, we have chosen to exclude the value of money stolen and the value of goods stolen for the thief's own use from the National Income of the stolen goods market. As noted in appendix B, we estimate that the value of such money and goods was between $5.8 and $7.6 billion in 1975.

Given that we feel that incomes generated in this sector should be included in underground National Income, we are left with the problem of estimating these incomes. Both direct and indirect approaches to the problem are possible and have been used in the past. Direct approaches estimate the incomes of thieves and fences by considering the value of property stolen, the purchase and sale prices of fences, and the costs of doing business for thieves and fences. Indirect approaches seek to estimate trends in stolen goods markets by observing the movement of series (e.g., value of goods stolen, youth unemployment rates, truck hijackings) that are believed to be related to the size of the stolen goods market. We believe that the direct approach provides more reliable estimates of the size of stolen goods markets and will employ it to obtain estimates of incomes earned in these markets in 1974. Using the incomes approach to National Income estimation, this gives us our estimates of the amount of underground National Income generated by stolen goods markets. We will use the indirect approach to estimate trends in this market.

Using the methods and data described in detail in appendix B, we estimate that $28.8 billion worth of property was stolen in 1975. Most of this property was not stolen by traditional strong arm tactics (e.g., robbery, burglary), but rather represents business loss as a result of shoplifting and employee theft. Indeed, only approximately $6 billion of the total $28.8 billion loss occurred as a result of the traditional property offenses (robbery, burglary, larceny). Further, it appears that while the rate of growth in traditional theft has slowed considerably since the late 1960's, the rate of increase in such non-traditional areas as employee theft has actually accelerated.

In order to estimate National Income for the stolen goods industry, we need to know not the value of goods stolen, but rather the incomes earned by the thieves and fences who operate in this industry. Using the methods and data described in appendix B, we estimate that thieves selling their goods in the stolen goods market earned net incomes of between $3 and $3.4 billion in 1975. In addition, fences operating in these markets in 1975 earned estimated net incomes of between $2.4 and $5.6 billion. Thus our estimate of the National Income for the stolen goods industry in 1975 is between $5.6 and $9.4 billion. Moving this 1975 figure backward in time using the consumer price index as a deflator, one obtains an estimated 1974 "National Income" for the stolen goods industry of between $5.1 and $8.6 billion.

Trends in the size of the stolen goods industry are hard to estimate. However, assuming that the size of this industry moved in approximately the same manner as the value of stolen goods reported to the police, we estimate that this industry grew at approximately 11 percent per annum in the 1960–69 period and at a rate of approximately 8 percent per annum during the 1970–75 period. The growth rate of

the stolen good industry exceeds that for reported GNP by approximately 4 percent per annum in the 1960–69 period and was below that for reported GNP by approximately 1 percent per annum in the 1970–75 period. It appears that the stolen goods sector of the underground economy grew relative to the reported economy during the sixties, but has declined slightly relative to that sector during the 1970's.

(II) THE COST AND BENEFITS TO INDIVIDUALS AND ORGANIZATIONS DIRECTLY INVOLVED

The major benefit to buyers in the stolen goods industry is the lower prices available. As noted in appendix B, price discounts are substantial with both retail and wholesale buyers receiving discounts as high as 80 percent of the corresponding legitimate price. The lower prices available on stolen good markets mean that reported measures of the level of prices such as the Consumer Price Index overstate the true level of prices. Given the rapid growth of stolen goods markets in the 1960's we may have overstated our rates of inflation during that period. However, the tapering off of growth in the size of the stolen good markets in the 1970's means that this source of overstatement has not been present in more recent years.

Risk to buyers in stolen goods markets appears minimal; thus the price discounts available in these markets are not subject to an extensive risk premium. At least some buyers appear to be attracted to the stolen goods dealers to participate vicariously in "the illegitimate" (Klockars, 1974).

Professional thieves and fences appear attracted to the industry because the real income available (both monetary and nonmonetary) are higher than incomes available in legitimate alternatives available to them. Interestingly, the stolen goods industry (thievery and fencing) offers one of the more dynamic opportunities in our economy for individual entrepreneurship. While most legitimate entrepreneurship requires relatively substantial human and financial capital, entrepreneurship of the old "blood and guts" kind still seems possible in the stolen goods sector of the economy.

Intermittent or occasional blue collar thieves seem driven to theft by financial or personal necessity. For this group rewards are small and participation seems to come from dire financial need, addictive problems (e.g., drugs, gambling, alcohol) or a need to strike out.

Part-time employee thieves and fences, on the contrary, appear basically to be easily supplementing their incomes. The identification of such individuals with legitimate lifestyles means that nonmonetary rewards must be small.

Government at all levels incurs extensive enforcement costs as a result of the stolen goods markets. Federal, State and local police have extensive anti-theft and less extensive anti-fencing programs. Courts and corrections deal with large numbers of property offenders each year. Overall criminal justice expenditures to combat property crime exceeded $10 billion in the mid 1970's. In addition, all levels of government lose taxes as a result of untaxed illegal transactions and incomes. For example, assuming a 17 percent tax rate (the rate assumed for unreported income by IRS), Federal income tax receipts were $1 to

$1.5 billion less than they should have been in 1975 because of the stolen goods industry.

Social costs of the stolen goods industry are substantially greater. While, as noted above, the actual theft of goods has no social costs but represents from an economic perspective an equity problem, the external economic effect of such forced transfers is substantial. Business feels the costs in increased insurance premiums, inventory costs and increased expenditures on prevention. Business prevention expenditures alone amounted to almost $6 billion in 1976 (U.S. Department of Commerce, Bureau of Domestic Commerce, 1976, p. 7). We as consumers feel these increased business costs in higher prices. In addition, we incur prevention expenditures ourselves (e.g., locks, guards) and alter our daily living habits. Our lives are generally less rich due to fear of property crime. As a society, we lose further from the work disincentive generated by our society's inability to enforce property rights.

(III) SUGGESTIONS FOR RESEARCH AND POLICY

As far as we are aware, there have been no previous attempts to estimate the overall size of stolen goods markets in this country. We feel that additional research in this area could be very beneficial. One intriguing possibility would be for the relevant agencies of the Departments of Justice and Commerce to cooperate in expanding the victimization surveys to include a much wider spectrum of property crime against business. The Bureau of Domestic Commerce could then try to reconcile its figures for business loss due to property crime obtained from the trade literature with estimates obtained from victimization surveys.

The Bureau of Domestic Commerce or an alternative agency (e.g., Department of Housing and Urban Development, Department of Agriculture) should be encouraged to obtain estimates of property crime loss for construction and agriculture. More importantly, we need estimates of property crime loss for all levels of government. Data for such estimates could be obtained in connection with existing surveys such as the Census of Governments.

There has been a good deal of research on traditional or blue collar thieves; however, this research is poorly integrated at present and could be usefully synthesized. Our knowledge of white collar thieves is much sparser. Research in this area using criminal justice and business records could be very informative. In addition, surveys using nontraditional questioning methods (e.g., random response) could provide valuable insights to the methods used in this sector. Biographies and autobiographies of white collar thieves could provide valuable insights as they have for blue collar thieves.

There has been a good deal of recent research on fencing, but most of it concentrates on relatively small time dealers. In addition, this research has generally taken a legal or sociological approach. Additional research on "legitimate" fences and broker/fences could prove valuable. In addition, it seems likely that substantial insights could be obtained from a business or economic approach to the fencing industry. Roselius and Benton (1973) produced some interesting insights using a marketing approach.

Traditionally, law enforcement efforts have focused on the traditional or blue collar thief when attempting to control property crime. As has been pointed out by a number of authors (e.g., Walsh, 1976) this may be a short sighted approach since replacement in these occupations appears to occur quite rapidly. An alternative approach would be to focus property crime abatement effort on the fencing industry. Walsh (1976) provides a long list of potential strategies. However, even most of these strategies focus on the traditional or blue collar portion of the industry. The dynamic growth of loss to white collar thieves and, thus, presumably fences who sell predominately stolen goods from such thieves, would seem to make this area ripe for some innovative law enforcement efforts.

As has been pointed out in some detail by Blakey and Goldsmith (1976), successful large scale prosecution of fences requires changes in existing law. Blakey and Goldsmith (1976, pp. 1620–1626) provide a model theft and fencing law in their recent paper; and, thus we feel it unnecessary to go into needed legal changes here.

One potential long run policy is to use and record identifying insignia on products. This marking combined with computerized lists of stolen goods could be quite helpful in increasing the recovery rate for stolen property. In a recent article Roumasset and Hadreas (1977) suggest that we develop a computerized "telecrook" system modeled after existing systems for detecting bad checks.

In the area of employee theft, a two-pronged approach would seem potentially quite successful. First, business needs to increase security through more careful inventory control and pre-employment screening. Second, business must be willing to punish employee thieves when they are caught. Recent evidence suggests that at least retail businesses could shift some enforcement resources to employee theft. Consider the following findings of the Bureau of Domestic Commerce:

> In discount stores, it is estimated that for every dollar lost to a shoplifter, three are lost to employees. Although apprehension of shoplifters outnumber those of employees by 10 to 1, one company reports that dollar losses from employee pilferage are more than seven times as great as shoplifting losses (U.S. Department of Commerce, Bureau of Domestic Commerce, 1976, p. 19).

II. Production and Distribution of Illegal Goods and Services

While the previous sections examined the production of legal goods and services which are not properly reported, the following three sections will deal with the production of illegal goods and services. The first section below deals with illegal gambling, the activity which many feel provides the largest source of revenue among illegal activities. The other sections treat arson motivated by fraud and the well-studied heroin industry. In each section, we will estimate the "national income" of the sector and, where possible, discuss the trends in its growth, its mode of operation, and its cost to society. We close each section with a discussion of some unanswered research questions and of the available policy options. We are unable to include detailed analyses of sectors such as loan-sharking and prostitution for which no reliable size estimates are available.

A. Illegal Gambling

(I) SIZE AND TRENDS

The 1967 Task Force report of the President's Commission on Law Enforcement and Administration of Justice stated that "there is universal agreement among law enforcement officials that gambling is the greatest source of revenue for organized crime and the crime that involves by far the largest amount of money." In their comparison of the economic impact of crimes involving illegal goods and services, the Commission estimated that the seven billion dollar GNP of the gambling industry made up 87 percent of the total value-added for the production of all illegal goods and services—far outstripping the combined figures for narcotics, loan-sharking, prostitution, and illegal liquor activity. Thus, one is left with the impression that an accurate estimate of the size of the GNP of the illegal gambling industry will go a long way to estimating the total impact of all illegal production.

The Commission asked the National Organization of Racing Commissions (NORC) to estimate the amount which was illegally bet on horse racing each year, believing that this amount would be the major component of the amount bet for all illegal gambling. However, the President's Commission rejected the NORC estimate of illegal horse race betting of $3.3 billion, preferring instead the general statement that "estimates by experts of the annual amount of illegal gambling vary from $7 to $50 billion . . . Total annual profits are estimated at $6 to $7 billion."

To estimate the percentage of the population which gambles illegally each year and to get an idea of how much is actually wagered, we will study and compare three surveys of America's gambling habits—one local and two national.

In early 1972, Oliver Quayle and Company conducted two surveys of the betting habits and attitudes of adult residents of New York City for the Fund for the City of New York. The Fund was mainly concerned with the feasibility and implications of legalizing the numbers game and sports betting, two of the largest illegal gambling enterprises in the City and in the State of New York. The following table summarizes the participation rate, total amount wagered, payout percentage, and the gross receipts after the payout for each of the four types of illegal gambling which the Fund for the City of New York studied. Since New York City already has a legal off-track betting system, the Fund did not perform as complete an analysis of illegal horserace betting as it did for the other games.

TABLE 2.—ILLEGAL GAMBLING IN NEW YORK CITY

Activity	Percentage of New York City residents that participate	Takeout rate (percent)	Total amount wagered	Total takeout-total wager less payoff
Numbers	24	47.0	$576,204,700	$270,816,200
Sportsbetting with bookmakers		4.5	428,360,000	19,276,200
Sports card betting	7	60–80.0	35,410,000	24,787,000
Horserace betting with bookmakers	[1] 24	17.0	83,643,000	14,219,300
Total	(?)	29.0	1,123,617,700	329,098,700

[1] Before OTB.

A year and a half after the Quayle survey, in November of 1973, Mr. Alfred King, a member of the Criminal Division of the United States Department of Justice, estimated the extent of illegal gambling in the United States for the Department of Justice. Adapting some figures from the Quayle survey, Mr. King estimated that 37.8 percent of all horserace bets not placed at the racetrack were placed with bookies in New York City. He then used the financial report of New York's legal Off-Track Betting (OTB) operation to estimate the dollar value of illegal bets on horseracing in New York for the first 6 months of 1973. He calculated the weekly amount from this by averaging and compared this weekly handle with the weekly average of arrests for 1971 and 1972 by the Federal Government's strike forces in New York City. Having arrived at a factor of expansion, he used this same factor to project not only horse bets in New York but all types of illegal gambling in every urban area of the country on the basis of those bets uncovered by the strike forces nationwide.

The Department of Justice originally announced that their estimate for the total volume of illegal gambling in 1973 lay somewhere between $29 and $39 billion. No breakdown of this figure into different types of gambling or different geographic regions has been made available. However, Reuter and Rubinstein (1978) state that the Department's estimate of illegal sports betting in New York City was $2.8 billion (compared with the estimate of $428 million by Quayle for such gambling).

TABLE 3.—TAKEOUT AND HANDLE FOR U.S. GAMBLING, 1974

Activity	Takeout rate (percent)	Total takeout	Estimate of total amount wagered
Legal:			
Horses at track	16.0	$1,247,000,000	$7,930,000,000
OTB, New York	21.0	171,000,000	967,000,000
Legal casinos	15.0	1,004,000,000	6,076,000,000
Bingo	33.0	551,000,000	1,735,000,000
Lotteries	55.0	374,000,000	639,000,000
Total legal	19.3	3,347,000,000	17,347,000,000
Illegal:			
Sports books	4.5	105,000,000	2,341,000,000
Horse books	16.6	227,000,000	1,368,000,000
Numbers	54.0	575,000,000	1,064,000,000
Sports cards	60.0	115,000,000	191,000,000
Casino games	15.0	19,000,000	110,000,000
Total illegal	20.5	1,039,000,000	5,074,000,000
Total, legal and illegal	19.6	4,385,000,000	22,421,000,000

Source: National Gambling Commission, 1976, p. 64.

The most extensive survey of the extent of gambling (both legal and illegal) in the United States was performed in the summer of 1975 by the University of Michigan Survey Research Center (SRC) for the Commission on the Review of the National Policy Toward Gambling. The Commission, which we will call "the National Gambling Commission" for brevity's sake, was charged with surveying "what is known about each form of gambling," "who should regulate it and how," and what are "the possible consequences of its legalization." Table 3 summarizes the findings of the survey with regard to the extent of gambling participation in the United States.

As a check on these figures, the SRC compared their estimates for the "handle" (total amount wagered) of *legal* gambling in the United

States with the official 1974 financial reports of all legal gambling establishments. They found that their estimate fell within 0.1 percent of the total reported officially by the operators of legal games and suggested that this concurrence pointed to the validity of their estimates of the illegal handle.

How does one then explain the difference between the $29 billion estimate of the Department of Justice and the $5 billion estimate of the SRC survey for the total handle of illegal gambling?

The Justice Department estimates are based on the following assumptions: (1) The rate of arrests for illegal operations in New York City is the same as the arrest rate for the rest of the country; (2) the rate of arrest of (illegal) horseracing betting is the same as the rate of arrest for all other kinds of illegal betting; (3) the 37.8 percent rate computed by the Quayle survey for illegal horserace betting is an accurate figure.

There are obvious difficulties with each of the first two assumptions. The use of apprehension data is problematic, especially if the efforts of law enforcement agencies are unevenly distributed across the nation or among different forms of gambling. Reuter and Rubenstein (1978, pp. 60–61) suggest that the Department of Justice had a higher sampling fraction in New York than elsewhere since federal antigambling efforts there were "built around wiretapping, and the rate of surveillances per capita was higher in New York than in other major cities." This could give "an upward bias" to the Department of Justice figures.

The Department's use of the Quayle survey also raises some questions since it assumed that all horse bettors use bookmakers to the same extent as horse bettors who are also sports bettors. Because sports betting is the main service of bookmakers, any sports bettor who wants to bet on horses with a bookie is likely to have already established contact with one.

However, there are indications that the SRC survey erred on the low side. The Quayle survey found that most of the volume of sports wagering was concentrated among a rather small number of sports bettors. For example, 27 percent of all football bettors—those betting $500 or more a year—accounted for 85 percent of the dollars bet. The sample design of the SRC survey "did not account for the possibility that most illegal gambling is concentrated among a relatively small group of people. If there were as many as 500,000 people in the Nation who bet an average of $50,000 annually, the sample used by the SRC has a very small chance of producing good results of their gambling activity." (Melnick and Crocker, 1976, p. 10) If such a group were missed in the sample and the SRC estimate were accurate for the rest of the population, "the true amount of illegal betting might exceed $30 billion a year." (Melnick and Crocker, 1976, p. 11)

Conversations with researchers familiar with the New York City sports betting environment indicate that there is a substantial number of bettors who wager over $1,000 a week with bookmakers in that city. Furthermore, there are games just about every week which involve 100 to 125 bettors and handles of $250,000.

The SRC survey estimated that the total handle of illegal gambling in the United States was $5,074 million in 1974. Since there is sufficient

cause to suspect that this estimate is low, we will double the SRC estimate to obtain a figure that should be an upper bound on the total illegal handle. Thus, we feel safe in estimating that between $5 and $10 billion were wagered in illegal bets in 1974. Assuming that the takeout rate of 20.5 percent calculated for the aggregate illegal gambling is correct—the takeout rate for legal gambling was a rather similar 19.3 percent—one arrives at a takeout somewhere between $1 and $2 billion for illegal gambling in 1974.

(II) SUGGESTIONS FOR RESEARCH AND POLICY

The above discussion on gambling and its impact on American society suggests a number of questions, whose answers would clarify the impact of illegal gambling in American society. All three studies mentioned in this section took for granted that organized crime derives substantial revenue from illegal bookmaking and numbers games, although the National Gambling Commission's final report cast some doubt on the connection between illegal gambling and national crime syndicates like the Mafia. More recently, Reuter and Rubenstein (1978) have argued convincingly after a long and careful analysis of police files that the fragmentation and low profitability of illegal gambling in New York City suggests minimal syndicate control of such gambling there. Additional studies of police files, similar to the Reuter-Rubenstein study, should be conducted to clarify the relationships between gambling and organized crime in all parts of the country. If the connections are weak, much of the work on legalized gambling will have to be redone.

The authors feel comfortable with the estimate that the total handle of illegal gambling in 1974 lay between $5 and $10 billion. If there are sufficiently many gamblers who bet an average of $1,000 a week throughout the year, then the total handle would be much closer to the upper limit. It would be interesting to discover how much of such gambling exists and what percentage of the total illegal handle it comprises.

There is no way at present to estimate the trends in the intensity of illegal gambling. The only convincing national figures we have are the 1976 SRC data. If three or four different areas of the country, including New York City, made careful periodic estimates of the intensity of gambling, we could keep track of the National Income of this rather large industry and possibly investigate the effect of different policies on the economic success of this industry.

As will be seen in section II.C, researchers have conducted a careful analysis of the costs to society of heroin consumption. What are the costs to society of illegal gambling? Comparison of these figures would help set priorities for police activities. In the case of illegal gambling, these costs would have to be balanced by a computation of the benefits from such gambling. For example, illegal gambling provides employment for large numbers of people who may be counted in our unemployment figures.

There appear to be three alternatives to the present method of dealing with illegal gambling: (*i*) Increase police activity against illegal

56-368 O - 80 - 7

gambling; (*ii*) legalize or at least decriminalize most gambling activities; and (*iii*) encourage and support legal (possibly state-run) alternatives to illegal gambling. However, it is not clear that any of these options has major advantages over the other options or over present policy. No reasonable amount of police activity will eliminate all illegal gambling. As the National Gambling Commission concluded in its final report, "gambling is inevitable." For example, the numbers game is played heavily in many black and Puerto Rican neighborhoods of New York. "Its lore is part of the community fabric. Runners and controllers are well known figures." (Fund for New York City, 1972, p. 9.) Somewhere between 10,000 and 100,000 New York City residents worked for the numbers business in 1970. In addition, Canes (1976, p. 113) estimated that "85,000 people in the United States derive income from bookmaking activities." Before one increases police activity against illegal gambling at the expense of potential police activity against other illegal activities, one should measure the social costs and benefits of illegal gambling and compare these figures with those of other illegal activities. In fact, the National Gambling Commission reported that the New York City police actually relaxed their activities against small-time bookmakers in 1971 in an effort to cut down on police corruption.

The National Gambling Commission reported that there is broad support for the legalization of some gambling activities. It concluded that gambling should be dealt with at the State and local levels, where the mores of the people of a particular area play an important role. It cautioned that a significant number of Americans believe that any form of gambling is wrong and that legalized gambling may undermine the work ethic for many citizens. Indeed, legalized gambling appears to create new gamblers. Furthermore, if gambling revenues are taxed, those who have easily avoided paying income taxes on their profits from illegal gambling will continue their tax evasion when gambling is legalized, unless stringent reporting requirements are implemented. If such reporting requirements are not implemented, most gambling income will be unreported whether it is from legal or illegal sources.

Finally, there is the option of allowing gambling only for state-run games. In 1976, states earned some $1.2 billion from legalized gambling with pari-mutuel wagering accounting for $719.3 million and state lotteries for $462 million. States have expressed three motives for their legalization of certain gambling activities: revenue, competition for illegal gambling, and in the case of present and proposed casinos in New England, a stimulus for the urban renewal and economic growth of certain areas. The National Gambling Commission emphasized that the first two of these goals are incompatible since games which are truly competitive with illegal games must be highly flexible, inexpensive to run, retain a fairly small take-out, and be virtually un-taxed. The Commission also found that if State-run gambling is viewed as a voluntary form of taxation, it is a regressive tax and one that is much more expensive to collect than income and sales taxes in terms of its administrative overhead. The Commission also noted that despite the introduction of the state-run Off-Track-Betting system in New York, illegal bookmaking in New York City continued to increase.

B. Arson

(I) SIZE AND TRENDS

The American Heritage Dictionary of the English Language defines arson as "the crime of maliciously burning the building or the property or another, or of burning one's own for some improper purposes, as to collect insurance." Arson is a unique crime in that an investigation must be conducted before one knows whether a crime was committed. Such investigations are not as common and as careful as they could be and there are a number of reasons why arson will not be detected even after a careful investigation. Most police, fire, and insurance experts take it for granted that one-half of those fires whose causes are classified as unknown and probably a number of those classified as accidental are actually caused by arson.

Fire reports classify the causes of fire into five main categories: (*i*) Accidental; (*ii*) natural (lightning, etc.); (*iii*) incendiary (intentionally set fires, including fraud fires); (*iv*) suspicious (suspected of being incendiary); and (*v*) unknown cause (no cause established). Table 4 adapted from the annual report in Fire Journal (September 1975), summarizes the statistics for building fires in 1974. The statistics were compiled by the National Fire Protection Association. Note that 9 percent of the building fires and 17 percent of the building fire losses in 1974 can be attributed to arson. If one includes 50 percent of the fires of unknown cause which are thought to be arson induced, then one finds that arson accounted for 15 percent of the building fires and 36 percent of the building fire losses for 1974.

TABLE 4.—BUILDING FIRE STATISTICS FOR 1974

Cause of fire	Number of fires	Percent of all fires	Loss value	Percent of all dollar losses
1. All causes	1,270,000	100	$3,260,000,000	100
2. Accidental or natural	996,400	78	1,460,000,000	45
3. Incendiary or suspicious	114,400	9	563,000,000	17
4. Unknown causes	159,200	13	1,237,000,000	38
5. Incendiary or suspicious plus ½ unknown cause	194,000	15	1,181,500,000	36

In our study of the underground economy, we are concerned solely with arsons from which some profit is made. "In common type of insurance fraud, a person may buy a property—generally a vacant building in an economically depressed section of the city—and insure it for more than it's worth. A fire will then result in a substantial profit on the investment. . . . For example, in 1969 a man bought two properties in central St. Louis for $6,000 and placed the deeds in the names of two straw parties. Within two years, there had been a serious fire of suspicious origin in each property, with the insurance payments totalling $33,424. This same property-owner had received over $415,000 in insurance payments for 54 fires occurring within a 2-year period (he was indicted for arson for one of these fires in early 1972)." (Boudreau et al., 1977, p. 20.) Businesses may resort to arson when, for example, they build up a large inventory of unsalable seasonable goods at the end of the season or their plant becomes outmoded

or requires extensive renovation. The foreclosure of a mortgage, adverse market conditions, or obsolete merchandise may induce a businessman to try to profit from his situation through the use of arson.

In some of this country's larger cities, professional arson rings have operated to defraud the insurance companies of millions of dollars. One such ring was uncovered in Detroit in 1974 when 57 persons were charged with 186 counts of arson. Detroit police reported that top arsonists received an average of $1,500 per fire, and that the best of the professional "torches" are involved in up to 100 fires a year.

There appear to be few reliable figures on what percentage of arson fires are set to defraud insurance companies and what the costs associated with such fires are. Of 1,703 fires established as arson in Ohio, North Carolina, and Pennsylvania during the years 1950 through 1955, 275 (or 16 percent) were found to have been insurance fraud fires. The Metropolitan Chicago Loss Bureau, in a careful study of 80 percent of arson-suspected fires in Chicago, found that one of every seven arson fires was set to defraud insurance companies. Since Chicago's fire problems are typical of the Nation's problems, we will use the "1-of-7" figure in our arson-for-fraud computations. On the other hand, arson-for-fraud constitutes a much higher percentage of the property loss figures as compared to other arson fires. Conversations with arson researchers have suggested that the property damage in arson-for-fraud fires is 30 percent to 40 percent of the property damage due to all arson fires.

Table 4 indicates that 194,000 arson building fires caused $1.182 billion in property damage in 1974. Using the above rates (1-of-7 and 35 percent, respectively), we are led to estimate 27,714 building fires set to defraud insurance companies, resulting in a property loss of $440.65 million.

There are two basic questions in the arson-for-fraud investigation for which no one appears to know the answer. First of all, how much profit is being made by those who are burning their own homes or businesses to defraud insurance companies? Sophisticated statistical analysis of one month's sample of arson fires in Tennessee has led J. D. Icove (1979) to the conclusion that the ratio of insured value to actual value of property destroyed by arson fire is 1.5 to 1. We'll use this ratio in our calculations. Under this assumption, the insurance payments for arson-for-fraud fires would total $620.55 million in 1974 and $660.98 million in 1975. The net profits for the property owners would be $206.85 million in 1974 and $220.32 million in 1975.

A second important unanswered question in the investigation of fraud arson is: How much are the professional torches making each year? Since arson is a crime done in secret and easily disguised, we may never have a good estimate on this important figure. The investigations of Karchmer (1977, 1978) into several recent prosecutions place the range of the torch's fee at between 10 percent and 25 percent of the insurance settlement. The torch's fee was 10 percent of the insurance payment in an arson conspiracy described by Battle and Weston (1954, p. 224). However, articles in *Psychology Today* (Horn, 1975) and *Newsweek* (1974) describe estimates by some arson investigators that "top arsonists make an average of $1,500 a fire." We can compute a gross upper bound for the torches' fee if we assume a fee of

$1,500 for each fraud arson incident. This computation leads to a figure of $41.55 million for 1974 and $45.6 million for 1975. If on the other hand, we assumed that torches received 10 percent of the insurance settlements and were in all the fraud arson fires, we would arrive at profit figures of $41 million for 1974 and $44 million for 1975. Since these two sets of figures are similar, we will use them as our upper estimates on the torches' profits for these years.

Finally, table 5 below describes the increase in the amount and costs of arson for each year beginning with 1964. These long-run comparisons have a number of shortcomings. Not only must one take inflation into consideration when studying the trends in the cost of arson, but one must also keep in mind the effect that more sophisticated arson-detection techniques and more zealous anti-arson police activities can have on the number of arson-fires reported in each year. Still, the 15 percent annual rate of increase in the number of arson fires and the 23 percent annual rate of increase in the nominal value of property destroyed by these fires indicate that arson is one of the more rapidly growing components of the underground economy.

TABLE 5.—ESTIMATED "INCENDIARY, SUSPICIOUS" BUILDING FIRES, UNITED STATES, 1964–75

	Number (thousands)	Property damage (millions)		Number (thousands)	Property damage (millions)
Year:			Year—Continued:		
1964	31	$68	1970	65	$206
1965	34	74	1971	72	233
1966	38	94	1972	85	285
1967	45	141	1973	94	322
1968	50	131	1974	114	563
1969	57	179	1975	144	634

(II) SUGGESTIONS FOR RESEARCH AND POLICY

The rapid growth of the fraud-arson segment of the underground economy has begun to catch the attention of the general public. Unfortunately, we know very little about the fraud-for-arson industry, and we have at the present meager resources to discourage its rapid growth.

There are a number of straightforward questions about the size and impact of the arson industry which need to be answered. We need to take the guesswork out of the questions of just how many fires each year are caused by arson and what percentage of these are set for financial gain. In particular, insurance companies should be encouraged to estimate carefully the number of suspicious claims they receive each year, the total of the payments made for these claims, and the market value of the property involved. In addition, detailed studies should be made to determine the social and economic costs to society of arson, especially fraud arson. These studies should take into account the vast loss of life and shutdown of firms which have been attributed to arson. Even though the National Income of the fraud arson industry is small compared to that of gambling and drug use, the actual costs to society from fraud arson are probably greater than those of illegal gambling and heroin. Once we have the answers to these questions,

we can better decide how to increase our commitment to the fight against fraud arson and we can establish realistic goals in this battle.

As part of this study, efforts should be made to establish the role of professional torches in the fraud-arson picture. How many are there, how do they operate, what fees do they charge, do they operate across state lines, and are they connected to national crime syndicates? Armed with this information, Congress can encourage and, if necessary initiate, cooperative efforts of law enforcement agencies against fraud-arson rings.

Especially urgent needs in the war against arson are a substantial increase in the number of trained arson investigators and in the effectiveness of the detection devices they use. Many well-populated areas of the country have only one arson investigator to examine all the suspicious fires which arise. Ohio and Illinois have 10 investigators each to cover the entire state. As a result, about 40 percent of building fire losses are in fires of unknown origin. There are a few technical devices such as gasoline vapor detectors, which can help arson investigators. However, the sensitivity of these devices still leaves a lot to be desired. (Arson investigators find it especially difficult to determine the causes of fires which start in the electrical wiring system. Boudreau et al. (1977, pp. 91–96) compiled a set of recommendations of needs and strategies in the struggle against arson.) Congress should make sure that adequate funding is available to encourage the engineering of more sensitive arson detection devices and to train an adequate number of investigators in the use of these devices.

Insurance companies need assistance with the dilemmas that arson presents. On the one hand, they are required to pay fire claims expeditiously; on the other hand, they should be able to delay payment in cases where fraud arson is suspected without concern for punitive legal suits. In addition, they tread a fine line between the responsibility of sharing information with law enforcement authorities and the legal constraint of preserving the privacy of information concerning their clients. "Insurance companies can help to eliminate the profit motive from arson through more selective underwriting, greater avoidance of over-insurance, not paying claims until the investigation has been concluded, more defense of fraudulent claims in civil court, and providing more information on insurance coverage to arson investigators." (Boudreau et al., 1977, p. 94.) If necessary, Congress can pass laws which guarantee that insurance companies have the protection, flexibility, and responsibility they need to play their central role in the war against fraud arson.

C. Heroin

(1) STRUCTURE OF THE INDUSTRY

The heroin industry has been the most thoroughly studied segment of the underground economy. The drug problem attracted national attention in the early 1970's, especially as incidence of narcotic use reached the high school population. This notoriety resulted in increased support for research on the causes and consequences of drug abuse, with an emphasis on the study of heroin.

Until the early 1970's, most of the heroin which reached the U.S. market was produced from poppies grown in Turkey with the transformation to the morphine base accomplished by syndicates in or near Istanbul. It was then shipped to covert laboratories in Southern France where it was refined into heroin.

During the early 1970's, intense international efforts stopped the Turkish-French connection through the arrest of the major French and Corsican drug traffickers and the ban on the cultivation of the opium poppy by the Turkish Government. However, Mexico quickly picked up the slack and now supplies 90 percent of the heroin available in the United States. Enforcement is especially difficult because so little pure heroin is needed to supply the entire U.S. heroin market for one year. The Bureau of Narcotic and Dangerous Drugs (BNDD) estimated that the U.S. supply was 10,000 to 12,000 pounds of opium in 1970, just 2 percent of the estimated world production of illicit opium. This amount could easily be grown on an area of 12 square miles and transported in a single medium-sized airplane.

A number of researchers have provided rather careful descriptions of the heroin-distribution industry. The most complete study may be found in Mark Moore's book, *Buy or Bust.* Moore estimated that there were 25 heroin importers in New York City in 1974. Each importer stood at the top of rather long, vertical organizational systems, with each unit maintaining a type of monopoly control over the units below it. Moore estimated that there were between 70,000 and 150,000 heroin users in New York City in 1974 and that they consumed about 940 kilograms of pure heroin at a retail value of $470 million. After breaking down the profits and value added at each level, he arrived at a total value added of $324 million for annual consumption in New York City.

Moore's model estimated that the "average user" consumed 4½ five-dollar-bags of heroin a day, derived 44 percent of his income from thefts and 32 percent from narcotics dealing, and spent 22 percent of each year in jail or in a treatment center.

In another interesting study, Silverman and Spruill (1977) used a time series model to estimate the way in which demand for heroin and property crime change with changes in the price of heroin in Detroit in the early 1970's. They estimate that when the price of heroin increases by 10 percent, demand for heroin decreases by 2.7 percent. The relative unresponsiveness of heroin demand to price increase means that as prices rise heroin expenditures actually increase. Silverman and Spruill also found that a 10 percent increase in the price of heroin leads to a 3.1 percent increase in total property crime against the poor nonwhite population. They suggested that 30 percent of Detroit's property crime is committed to support the consumption of heroin.

Finally, it should be noted that the initial use of heroin follows a distinctive pattern. It usually occurs between the ages of 15 and 21 and is initiated by a well-meaning friend, usually in the setting of a previously established peer group activity. The initiator is usually new to the use of heroin, perhaps not yet addicted. The epidemiologists Hunt and Chambers (1976) argue that the most effective way to intervene in any on-going heroin epidemic may be to identify users who have only recently become involved with heroin and to develop treatment alternatives that will attract them away from its use.

(II) SIZE AND TRENDS

A number of researchers have used different techniques to estimate the number of heroin users. Joseph Greenwood, who has been estimating the size of the population of heroin users every year since 1969 for the Bureau of Narcotics and Dangerous Drugs (BNDD) made the following estimates for the years 1969 to 1974:

Year	Estimate
1969	315,000
1970	524,000
1971	559,000
1972	626,000
1973	602,000
1974	725,000

Source: Hunt and Chambers, 1976, p. 112.

Greenwood used the arrest and rearrest figures of the BNDD register to compute his estimates. In a given period he knew how many of heroin users were on the register. He used the rearrest data of these "marked" users to estimate the size of the user population. This method is analogous to counting balls in an urn by marking a sample of balls in the urn (say m balls), mixing them completely back into the sample, and drawing out a random sample of balls from the urn. The fraction of marked balls to total number of balls in the second sample should approximately equal m divided by the total number of balls in the urn.

Epidemiologists Leon Hunt and Carl Chambers (1976) argue that Greenwood's method grossly underestimates the size of the heroin-user population because the second sample that Greenwood counted (the rearrest group) was far from random; for a person who has been arrested once has a higher probability of being rearrested than others in the user population. They further argue that later findings showed that Greenwood's 1969 estimates were grossly underestimated. Using statistical comparative analysis, they estimated that 3,772,000 Americans used heroin in 1974. This data suggests that there were 500,000 to 750,000 compulsive, regular heroin users in 1974 and another two to three million noncompulsive, irregular users.

To estimate the national income of the heroin industry in 1974, we will use a couple of approaches. First, we will extrapolate Moore's figures for New York City to arrive at national figures. Moore's estimate of 70,000 to 150,000 users in New York City included some irregular users ("joy poppers" and "drug dabblers"). Using the data of the previous paragraph, we will assume that one-tenth of the Nation's heroin use occurs in New York and therefore that the national income of the heroin industry is 10 times the New York figure of $324 million, i.e., $3.24 billion. Second we will use information in the 1978 Annual Report of the Office of Drug Abuse Policy (O.D.A.P.) which estimated the amount of heroin imported into and consumed in the United States in 1976. Assuming that there were 500,000 daily users and 1,500,000 less-than-daily users and that each daily user required 50 mg. a day for 365 days a year, O.D.A.P. estimated that daily users consumed 9.1 metric tons in 1976. Combining this estimate with one that less-than-daily users consumed 0.9 metric tons in 1976, O.D.A.P. suggested that 10 metric tons were consumed. Assuming a cost of $1.43

per pure mg. led to an estimate that $14.3 billion of heroin were consumed in 1976. However, O.D.A.P.'s estimate of the total supply of heroin available to the U.S. market was 6.16 metric tons—well below their estimate of the 10 metric tons consumed. Believing that: (1) The supply figures are more reliable; (2) most heroin addicts used less than 50 mg. a day and could only consume heroin 80 percent of the year; and (3) the average addict spent less than $1.43 per mg., we will multiply the above total cost figure by (6.16/10.0) to arrive at an adjusted estimate of $8.6 billion. This adjusted estimate derives further suport from the fact that the National Narcotics Intelligence Consumers Committee estimated that $8.8 billion of heroin was consumed in 1977. Finally, to change this total cost estimate to a national income estimate, we multiply it by the ratio of value-added to total cost (324/470) that Moore found in his detailed study to arrive at a national income of $6.5 billion for the heroin industry in 1976. Finally, we move this figure to 1974 by multiplying it by the appropriate price ratio (100/130) (the supply and addict population were rather steady in this period) to obtain an estimate of $5 billion for the national income of the heroin industry in 1974.

We close this section with a discussion of the trends in the size of the heroin-use population. As noted in the discussion of the estimates of Greenwood and of Hunt and Chambers, the number of users has been increasing in the period between 1969 and 1974. However, all studies indicate that the number of those who begin their use of heroin in any given year peaked in 1969 and has been declining rapidly since then. Epidemiologists who study the incidence of heroin use all drew graphs similar to the one below, relating each year to the number of persons who first used heroin in that year. We have not labeled the vertical axis in units because different researchers considered different populations and different measurement techniques in describing the incidence of heroin abuse.

As noted earlier, it is the new user who probably contributes most significantly to the spread of the heroin epidemic. Thus, the above graph certainly leaves the clear impression that the war on heroin has met with success, at least between the years 1969 and 1975, and that the uncontrolled increase of the heroin epidemic of the late 1960's has been overcome. One can take some hope in the fact that the number of narcotic overdose deaths, the hepatitis rate, the number of addicts seeking treatment, and the crime rate for property crimes associated with heroin addiction all decreased between 1971 and 1973. (Office for Drug Abuse Prevention, 1973.) A few caveats are in order here. First of all, in the mid-1970's, the Mexican Connection took over as the chief supplier of heroin in the United States. Researchers will have to monitor the new distribution system carefully to see whether it sends the incidence of heroin use climbing again. For example, the incidence graph for Los Angeles, a major center in the new distribution system, is quite different from the shape of the previous graph. The total number of reported voluntary admissions to treatment centers was 50 percent higher for Los Angeles than it was for New York City in 1976. Articles in the *Washington Post* in August of 1979 point to a sudden marked increase in heroin in our nation's capital.

(III) COSTS TO SOCIETY

In order to understand the magnitude of the problem of heroin abuse, to compare its impact with that of other problems facing society, and to evaluate alternative courses for action for dealing with the heroin problem, society needs a careful estimate of the total social costs of heroin abuse. Since 1971, a number of researchers have attempted to measure these costs.

One of the most recent and most careful reports in this direction was that of Rufener, Rachal, and Cruze (1976a, 1976b), researchers at North Carolina's Research Triangle Institute, as part of a cost-benefit evaluation of the drug abuse treatment programs sponsored by the National Institute of Drug Abuse. This report estimated the costs as $10.3 billion for the fiscal year 1975 (July 1, 1974 through June 30, 1975). Because it used the most careful estimation procedures and the most recent data available, its results will be summarized below.

The researchers first enumerate the tangible costs borne by society because of the existence of drug abuse. For convenience, these costs are divided into two groups: direct costs and indirect costs. "An explicit or direct cost is one in which resources are used and a formal payment is made in cash or in kind (i.e., through the direct provision of some commodity or service). When resources are used to treat the medical consequences of drug abuse, the labor is paid for in wages, the materials used are paid for and the capital used receives a return . . . An indirect or implicit cost occurs whenever resources are used for which no formal payment is made, that is, these resources are not priced by any market mechanism. When a drug abuser is treated in a hospital, his time is being used, but no formal payment is made for its use. In order to place a value on this indirect cost, one looks at the value of what the individual gave up, or the opportunity cost." (Rufener et al., 1976b).

Table 6 summarizes the costs to society of drug abuse. It differs from the tables in Rufener, Rachal and Cruze (1976b) only in that it does not include their direct cost for "nondrug crime." This adjustment is made to avoid a double-counting pointed out by Gillespie (1978).

TABLE 6.—ECONOMIC COSTS OF DRUG ABUSE

[Fiscal year 1975, assuming 500,000 addicts; in millions of dollars]

	All drugs	Heroin
DIRECT COST COMPONENT		
Medical treatment	494.0	111.2
Emergency room visits	12.0	1.2
Inpatient care	314.0	72
Mental hospital inpatient care	35	8
Miscellaneous medical costs	133.0	30
Law enforcement costs	1,342.0	780
Drug laws	667.0	105
Nondrug laws:		
Public expenditures	419.0	419
Private expenditures	256.0	256
Judicial system costs	296.0	151
Drug laws	172.0	27
Nondrug laws	124.0	124
Correction costs	294.0	161
Drug laws	158.0	25
Nondrug laws	136.0	136
Drug traffic control	93.0	46.5
Drug abuse prevention	995.0	535
Housing stock loss	84.0	84
Total, direct costs	3,598.0	1,868.7
INDIRECT COST COMPONENT		
Unemployability	2,478.0	2,478
Emergency room treatment	.4	.04
Inpatient hospitalization	20.0	4.6
Mental hospitalization	8.0	1.8
Drug-related deaths	12.5	2.3
Absenteeism	1,594.0	161
Incarceration	1,205.0	494
Drug laws	845.0	134
Nondrug laws	360.0	360
Work loss due to treatment	88.0	47
Total:		
Indirect costs	5,405.9	3,188.74
Direct costs	3,598.0	1,868.7
Total costs	9,003.9	5,057.44

If ones assumes that there were 750,000 addicts in fiscal year 1975, then the indirect cost component for unemployability would be $3,716 million instead of $2,478 million. This would increase the total cost to society as a result of drug abuse to $10,241.9 million and the total cost to society as a result of heroin abuse to $6,295.44 million.

(IV) OTHER DRUGS

As the 1979 IRS Report points out, "The difficulty in developing more precise national estimates for (cocaine and marijuana) mainly stems from the wide variations in quantities and frequencies of their use, which make consumption-based estimates of the volume of traffic in these drugs much more uncertain than the corresponding estimates for heroin trafficking." (Department of the Treasury, 1979, p. 134.) The National Narcotics Intelligence Consumers Committee, as quoted in the IRS report, estimates that $14.4 billion dollars worth of cocaine and $11 billion dollars worth of marijuana were consumed in 1977.

Both figures are larger than their $8.8 billion estimate for the 1977 retail value of heroin consumed. On the other hand, the 1978 Report of the Office of Drug Abuse Policy (ODAP) estimates 13.2 metric tons of cocaine were consumed and that 14 to 19 metric tons of cocaine entered the United States in 1976. At a cost of $.49 per mg., O.D.A.P. estimates the total cost for 13.2 metric tons at $6.46 billion—somewhat lower than the corresponding estimate of the cost of heroin consumption in 1976. Since no consensus exists for 1976 and no reliable estimates have even been made for 1974, we will assume that the national income of the cocaine and marijuana industries in 1974 were each approximately equal to the mean of our estimates for the heroin industry in 1974; namely, $4 billion.

(V) SUGGESTIONS FOR RESEARCH AND POLICY

There are a number of research questions whose answers would aid our understanding of heroin usage today. First of all, how does the fact that Mexico has replaced Turkey as the main source of heroin affect our description of the New York-based heroin industry? Is the actual number of heroin users in 1974 closer to 400,000 or to 4,000,000? What are the current trends in the number of heroin users, especially in light of the new Mexican connection?

In discussing effective policies against the spread of heroin use, one must keep in mind two socially and geographically distinct populations—the potential users who have not yet been introduced to heroin and the current regular users, many of whom are addicted to the drug. As mentioned above, there is strong agreement that 80 to 90 percent of heroin users were first introduced to the drug by a close friend in an informal setting. Such transactions among friends are relatively impregnable to narcotic enforcement efforts. Nor are they affected by the existence of treatment centers since the introducer usually has only recently begun using heroin and feels neither the pangs of addiction, nor the need for treatment at this early stage. Epidemiologists Leon Hunt and Carl Chambers (1976) argue that the only way that a program aimed at preventing new use of heroin can be successful is that it anticipate the locales of new use, focus on the individual new users, identify these new users early, and discover some way to induce the new unaddicted user to give up heroin. Their studies of the national diffusion of heroin use indicate that this diffusion has followed population density and city size, moving from densely populated large cities to sparsely populated smaller cities, with a movement that is unaffected by police activity, the curtailment of supply, or any other intrinsic condition. However, Moore (1977) uses simulation studies to argue that strangers do play a role in the propagation of heroin use—especially as a vector between different groups of friends. He believes that even if narcotics enforcement strategies are limited to deterring stranger-to-stranger transactions, narcotics enforcement can have a significant effect on the rate at which heroin use spreads.

However, as Moore points out, the enforcement of narcotics laws against regular heroin users can have strong adverse effects on these users—direct effects such as police harassment, jail, poor self-image, and a stigmatization by society; plus indirect effects such as high prices, variable toxicity, and the general difficulties of holding a regular job and affording the necessities of life. Enforcement programs

which are aimed at discouraging experimentation and encouraging the abandonment of usage by new users should not inflict too great a cost on current users, because of the indirect effect (e.g., increased property crime) generated. One way of making the effective price of heroin different for new and old users is to insure that connections are as difficult to make as can be. For example, there should be no area to which an inexperienced user can come and expect to find heroin. In addition, a wide variety of treatment programs should be available along with creative diversion systems which keep arrested users out of jail. Jail terms do little to improve the behavior and condition of users, at the same time they make finding reasonable employment at a later stage much more difficult.

With regard to the narrow, vertical heroin distribution system, Moore suggests that the most effective place to attack this system and confiscate supplies is at the middle level—that of the "weight-dealers." This level is low enough in the system that frequent transactions are made without the elaborate security strategies of higher levels. It is high enough that such dealers work with large quantities of heroin and would be difficult to replace. To attack this level effectively, police will have to rely on extended undercover operations, including search warrant and wiretap investigations.

The policies described above require well-trained leadership which can arrive at creative solutions to some of the problems posed. Congress can make sure that there are funds to train and coordinate such leadership and to carry out the above programs, including the necessary undercover operations. Congress should also insure that Federal funds are not being wasted on unplanned, counterproductive drug-law enforcement practices which may have only a temporary cosmetic effect. Finally, encouraged by the recent disintegration of the Turkish-French connection, Congress should work to provide incentives for Mexico to shut down its production of illegal opium poppies and should encourage our intelligence agencies to trace the current flow of opium from Mexican fields into American streets.

III. Summary and Conclusions

In this paper, we have assessed size, structure and trends in the underground economy. We have done this by surveying the literature in the six areas that we believe make up the major sectors of this economy. We divided these six areas into two broad groupings: (1) Production of illegal goods and services (drugs, gambling and arson); and (2) the unreported production and trade in legal goods and services (tax evasion, illegal aliens, and stolen goods markets). For each of the six areas studied, we have attempted to determine a sectoral "National Income" either by the incomes or the value added approach. This contrasts with the estimates of other papers on the underground economy which have used and combined such figures as gross revenues, profits, social costs, valued added, and taxable income in their estimates. We feel that the National Incomes approach is appropriate because it allows valid comparisons with the Commerce Department's regular statements on the National Income of the American economy.

Due to lack of data and divergent existing estimates, our National Income estimates should be viewed with considerable caution. With

this caveat in mind, we estimate the following National Incomes for our six sectors in 1974: (1) drugs—$11.2 to $13.0 billion;[3] (2) gambling—$1 to $2 billion; (3) arson—$.2 billion; (4) income tax evasion—$70 to $75 billion; (5) illegal immigrants—$1 billion;[4] and (6) stolen goods markets—$5.1 to $8.6 billion. Summing these six figures and adding approximately $2 billion for sectors omitted from our survey[5] one obtains an estimated National Income of between $90.5 and $101.8 billion in 1974. This is a sizable if not staggering amount being 8 to 9 percent of reported National Income in 1974. We gain added confidence in our estimate by noting that it is relatively consistent with estimates of the relative size of underground activity in other mixed economies. For example, recent estimates suggest that the activity in Germany's economy in 1974 was roughly 7 percent of the reported economic activity. As might be expected, estimates of the relative size of the underground economy in more highly taxed economies such as Sweden and Great Britain are higher (10 and 8 percent, respectively). *U.S. News and World Report*, 1979).

In terms of size, the underground activity reflected in non-reporting of legal income is by far the most prominent component. In terms of growth in recent years, the employee theft portion of the stolen goods markets, and the arson fraud industry, appear most dynamic. We, like IRS, can draw no definitive conclusion regarding trends in the entire underground economy in recent years. However, based on our research and IRS's, it appears that the underground economy has not grown markedly more rapidly than reported National Income in the mid 1970's.

The estimates we report above, obtained from a bottoms-up, national income accounting approach, to estimating the size of the underground economy, is substantially below recent estimates of Peter Gutmann (1977, 1978, 1979) who, using a coincident series approach, estimated a "GNP" for the underground economy of $176 billion in 1976. Gutmann obtains his estimate by considering the ratio of currency to demand deposits. He assumes that the ratio of currency to demand deposits in the 1937–41 period (21.7 percent) was normal and assumes that all growth in this ratio since that time was due to increased underground transactions. Since the ratio was 34.4 percent in 1976, Gutmann estimated that 12.7 percent of the currency outside banks in 1976 ($28.7 billion) was used to "lubricate the underground economy." Assuming further that only cash was used in underground transactions and that the number of times money changed hands in the underground economy equaled that in the legal economy (6.15), Gutmann obtains his estimate of $176 billion for the size of the underground economy ($28.7 billion x 6.15 = $176 billion).

Consider Gutmann's major assumptions, beginning with the assumption that the currency/demand deposit ratio in 1937–1941 was normal. As has been reported by Cagan (1958) and Anderson (1966), the ratio of currency to money supply experienced a number of short

[3] Our estimates for drugs other than heroin are quite tenuous.

[4] This is an estimate of completely unrecorded income paid to illegal aliens. An additional $5 to $6.6 billion of unreported income paid to illegal aliens, but recorded in Social Security records, employer reports or interviews is included in our estimate of tax evasion.

[5] There are no believable estimates for sectors such as loan sharking.

run trends during the first half of this century including a decline in the 1933 to 1941 period and a sharp rise from 1941 to 1945. Thus, Professor Gutmann chooses a period of low currency holdings for his comparisons. Further, 1937–1941 could hardly be considered a normal period for this country since we were just recovering from the most severe depression in our history and preparing to enter the longest war of this century.

If 1937–1941 is not a normal period and thus cannot provide us with an estimate of normal currency holdings, how might one obtain such an estimate? The best way would be to consider the reasons why people hold liquid assets and the incentives likely to cause them to shift such holding among alternative media. Decisions to hold currency and other liquid assets are part of an individual's portfolio decision. The amount of liquid assets an individual seeks to hold will depend on income, the relative yield on liquid versus nonliquid assets, and tastes. An individual will hold some proportion of income liquid to carry out day to day transactions. The proportion of this transaction demand for liquid assets which consists of currency will depend on the availability and attractiveness of substitutes. The more attractive substitutes available, the lower the percentage of cash one would expect individuals to hold. Dividing currency in circulation by personal consumption expenditures for various years in the post World War II period, one finds, as one would expect, a downward trend with currency in circulation being 16 percent of personal consumption expenditures in 1947 and 7 percent in the 1976 to 1978 period. See table 7. Thus, one potential alternative to Professor Gutmann's explanation for the growth in currency is that this growth is merely a reflection of increased income and prices which have caused consumers to hold a larger absolute amount of currency in order to carry out their increased volume of transactions.

TABLE 7.—CURRENCY OUTSIDE BANK AS A PROPORTION

	Currency outside banks (billions)	Personal consumption expenditures	Personal saving	Demand deposits	Time/saving deposits (at commercial banks)	Yield on corporate bonds (Moody's Aaa)	Unemployment rate (percent)
1947	$26.4	$0.16	$5.39	$0.30	$0.75	$2.61	
1950	25.0	.13	2.31	.27	.68	2.62	5.3
1955	27.8	.11	1.87	.26	.56	3.06	4.4
1960	29.0	.09	1.69	.25	.40	4.41	5.5
1961	29.6	.09	1.47	.25	.36	4.35	6.7
1962	30.6	.09	1.50	.25	.31	4.33	5.5
1963	32.5	.09	1.73	.26	.29	4.26	5.7
1964	34.3	.09	1.31	.26	.27	4.40	5.2
1965	36.3	.08	1.20	.27	.25	4.49	4.5
1966	38.3	.08	1.16	.28	.24	5.13	3.8
1967	40.4	.08	.99	.28	.22	5.51	3.8
1968	43.4	.08	1.14	.27	.21	6.81	3.6
1969	46.1	.08	1.31	.28	.24	7.03	3.5
1970	49.1	.08	.97	.29	.21	8.04	4.9
1971	52.6	.08	.92	.29	.19	7.39	5.9
1972	56.9	.08	1.15	.29	.18	7.21	5.6
1973	61.5	.08	.87	.29	.17	7.44	4.9
1974	67.8	.08	.94	.31	.16	8.57	5.6
1975	73.7	.08	.88	.33	.16	8.83	8.5
1976	80.8	.07	1.18	.35	.17	8.43	7.7
1977	88.6	.07	1.32	.35	.16	8.02	7.0
1978	97.5	.07	1.27	.37	.16	8.73	6.0

Sources: "Economic Report of the President." Washington, D.C.: U.S. Government Printing Office, 1977; "Federal Reserve Bulletin," various volumes

This leaves us still to explain, however, why consumers increased their holdings of currency relative to demand deposits. Two trends probably provide much of the explanation. First, the growth of credit card use allows persons to hold their noncurrency liquid assets in interest bearing time and savings accounts rather than in non-interest bearing checking accounts, at least for a time. Second, the increased liquidity of many time and savings accounts means that they are much closer substitutes for demand deposits than they were 15 or 20 years ago. Both of these trends would suggest that the ratio of currency to time deposits would be going down over time. As can be seen in table SC1 this has indeed been the case. It appears that households have rearranged liquid portions of their portfolios to hold relatively more time deposits and fewer demand deposits. Indeed, this fact is a major reason why many economists now define the money supply to include such assets.

Professor Gutmann also assumes that the velocity of circulation of money in the underground economy is equal to that in the legitimate economly. This seems tenuous since in the legitimate economy the velocity of circulation of demand deposits greatly exceeds that of currency. Indeed, Paul Anderson (1977) has recently estimated that as much as two-thirds of the currency in circulation is in savings hoards. On the basis of this information, it appears that Professor Gutmann overestimates the velocity of circulation of illegal currency. Recent articles suggest that 3 to 4 may be a more reasonable velocity of circulation for illegal currency.

The recent IRS report (U.S. Department of the Treasury, 1979) and testimony of Nancy Teeters of the Board of Governors of the Federal Reserve System (U.S. House of Representatives, 1979) provide further reasons for rejecting the estimates of Professor Gutmann and other researchers who use the currency approach to estimating the size of underground economy.

While we and others do not agree with Professor Gutmann's estimates, the attention he has drawn to the underground economy has certainly been useful. It may be worthwhile to pursue the currency approach to estimating the size of the underground economy; however, future work should attempt to estimate much more carefully legal holdings of currency and to understand the way in which transactions take place in the underground economy. One sector of the underground economy which this approach cannot hope to address is barter. Our own feelings are that studies of individual sectors such as we present above will probably provide the most accurate estimates possible of the size of the underground economy.

Interestingly, one author has recently attempted to combine the sectoral and currency approaches (Henry, 1976). Henry estimates that tax evasion financed by large denomination bills amounted to $30 billion in 1976, a figure well below our estimate for total tax evasion in 1974. Henry further estimates that the 1976 annual volume of "profit motivated" crime (drugs, gambling, auto theft, etc.) was $30 billion. This substantially exceeds our own estimate for 1974 ($10 to $15 billion), Our sectoral estimates appear to diverge most markedly from Mr. Henry's in the area of gambling where as we note above estimates are widely divergent and "solid" evidence has only recently become

available. Mr. Henry appears to use the estimate of the 1967 President's Commission on Law Enforcement that the profit due to illegal gambling was $7 billion. He revises these figures upwards to reflect 1976 prices and adds on the profit of the loan sharking industry, "which might easily be almost as large as the annual level of illegal gambling" (Henry, 1976, p. 31). Henry's estimates of the overall size of the underground economy in 1976 was $60 billion, well below our own estimate for 1974. However, Mr. Henry's estimates excludes a number of sectors (e.g., tax evasion not using big bills, arson, and illegal immigrants) and thus is not really an estimate of total underground activity. Mr. Henry's use of currency information appears intriguing although we have as yet been unable to receive details of his approach. Available evidence seems to indicate that Henry estimated excess currency in circulation by attempting to net out legitimate uses. This approach could be fruitful and although Mr. Henry's equations do not appear to have been thoroughly specified, they make an interesting beginning.

After the second draft of this paper was prepared, we received the IRS's *Estimates of Income Unreported On Individual Income Tax Returns*. The IRS report estimates that individuals failed to report from $100 to $135 billion in taxable income in 1976. Our estimates and those of the IRS are fairly consistent, especially when one considers the differences in income concept (National Income versus taxable income), the differences in base year (1974 versus 1976), and the differences in sectors covered. Because the IRS had more detailed data on the amount of off-the-books informal economic activity, we incorporated this IRS estimate in our report. The largest difference between our report and that of the IRS is in the estimates of the profit of the illegal gambling sector. While we use a modification of the national survey results which were compiled by the National Gambling Commission in 1976 to reach our estimates of $1 to $2 billion, it appears that the IRS relied solely on the rough estimates of the 1967 President's Task Force on Organized Crime. In addition, we provide an estimate of the very large and important stolen goods sector, which the IRS fails to do.

Our survey leads to a number of policy and research suggestions which cut across individual sectors of the underground economy. Feeling more comfortable with research, we will begin there. Surveying the various sectors of the underground economy, we were struck by the vastly varying quantities and qualities of research in the different sectors and by the degree to which research techniques and approaches vary by sector and ignore studies in other sectors. Per dollar of underground National Income generated, the heroin industry has been most extensively studied. In addition, work on the heroin industry has tended to use more sophisticated research techniques than work in many other areas. It has made extensive use of real data and has been relatively policy oriented. This situation contrasts markedly with research in the area of tax evasion which has been sparse per dollar of underground GNP and has been largely theoretically oriented. If one looks for causes of these differences, one is struck by two factors. First, many Federal agencies directly interested in the heroin industry are relatively open and research oriented. Indeed, many of these agencies are associated with National Institutes of Health (NIH) and National

Institutes of Mental Health (NIMH) complexes and much research on the heroin industry has been done by people in public health professions. Contrast this with IRS which is quite secretive and far from research oriented. Second, the articulate public became concerned with and wished to understand the heroin industry in which many of their sons and daughters became involved during the late 1960's and early 1970's. Contrast this with the tax evasion situation where many members of the articulate public are directly involved. Other intersectoral contrasts are similar if less dramatic.

As an example of the more thorough research that has occurred in the study of heroin use, consider the large number of studies of the total economic cost (both explicit and implicit) to society because of heroin use. We cite one of these studies at the end in part II above. Comparable studies do not exist in other areas of the underground economy and should be undertaken. Accurate comparison of the costs to society in these areas would be valuable guides in decisions on the allocations of law enforcement efforts.

Overall, research effort has focused on the traditional or blue collar areas of the underground economy that are dominated by "deviant" members of our society and has been relatively sparse in sectors where "honest" citizens predominate. We believe that research could be beneficially reoriented to focus more on such areas as employee theft and income tax evasion. We do not in any way wish to downplay the importance of research on more traditional areas of the underground economy, but rather to suggest that our overall understanding of this economy would probably be increased most if some research funds were to be reoriented. In the area of tax evasion, increased research is not likely to be highly productive until IRS is more cooperative with researchers. The intriguing results produced by the study of Schwartz and Orleans (1967) cited above make a strong case for the potential benefits of such cooperation.

General policy implications are more difficult to achieve, and are probably best understood when the reasons for concern about the size of the underground economy are considered. We believe that concern about the underground economy stems from three sources: (1) Concern about public revenue loss; (2) concern about distortion of reported economic indicators; and (3) concern about government laws and regulations driving more and more economic activity "underground." We will consider each reason for concern in turn and the policy and research suggestions emanating from each concern.

Concern about public revenue loss is serious and, indeed, the surfacing of underground income would alleviate, to varying degrees, the budgetary problems of all levels of government. The IRS estimates (Department of the Treasury, 1979) that between $19 and $26 billion of income tax revenue was lost to the Federal Government in 1976 as a result of the "underground economy." Other levels of government lose lesser although substantial amounts of revenue. Obviously, for revenue purposes, it would be desirable to surface as much underground activity as possible. How might we do this? Most simply by decreasing the incentives to participate in this illegal economy. Nonreporting of legal activities can be most simply decreased by a two-pronged approach which decreases the benefits and increases the cost

of participating in the underground economy. By pushing people into higher marginal tax brackets, our inflationary economy continually makes tax evasion more lucrative. As we point out in the tax evasion section, levels of noncompliance appear to increase at an increasing rate as tax rates increase. This increase in benefits of tax evasion might be counteracted either by decreasing inflation (the more desirable course) or by indexing tax rates so that inflation does not provide a built in incentive to "go underground." Government regulation (e.g., health and safety regulations, pollution regulation, social insurance payments, wage and hour laws, minimum wage laws) which increase the cost of legitimate business activity also stimulates the growth of underground activity. We as a society would be well advised when establishing such regulations to consider the incentives they provide to "go underground."

Concern about distortion of economic indicators takes two forms: (1) Concern about the magnitude of indicators; and (2) concern about trends in indicators. Because there is substantial underground activity, we understate the size of economic activity (GNP, National Income) and overstate the levels of unemployment and inflation in our economy. Further, as has been recently pointed out by Edward Denison (1978), we understate the productivity of our productive resources. Distortion of trends in indicators of economic activity, employment, inflation and productivity are, if anything, more important than distortions of the size of these indicators. Distortion of the trend in overall economic activity (e.g., GNP, National Income) will only occur if the underground economy grows at a different rate than the legitimate economy. It appears that the underground economy grew more rapidly than the legitimate economy in the late 1960's and possibly the early 1970's; thus, traditional indicators of the size of legitimate activity (e.g., GNP) were understated for these periods. However, the underground economy appears to have grown more slowly than legitimate activity in the mid 1970's; thus, we may have overstated the growth of our economy in this period.

Distortions of the unemployment and inflation rates due to selective employment and growth of the underground economy are more important. Our research leads us to believe that some proportion, perhaps a sizable one, of minority teenage unemployment is overstated as, relatively large numbers of young minority group members find employment in the underground economy. We wish to express our support of the recent recommendation of the American Economic Association's Census Advisory Committee: "That the Census Bureau give very high priority in its current and future planning to increasing our understanding of the nature of the measure employment and unemployment of teenagers, minorities, women, and other groups with special employment problems of particular importance to policymakers." To the extent that people make more purchases in the relatively cheap underground economy as inflation rages, we overstate our inflation rate.

Denison has estimated that measured output per unit of input in the nonresidential business sector was 1.8 percent smaller in 1975 than it would have been if 1967 conditions had prevailed. The need for large protection expenditures caused by property crime and the rise in "off

the books" activity are major causes of this understatement of productivity.

Concern about government laws and regulations driving economic activity underground are well founded although we don't believe there is any simple solution to this problem. It is only laws and regulations which make certain types of production and transactions illegal. By making drugs, gambling and prostitution legal and abolishing our immigration laws we could decrease the size of the underground economy by as much as, perhaps, $10 billion. We as a society may want to seriously consider legalization in the case of gambling (already legal in a number of areas) and prostitution.

Other types of government regulations (health and safety regulations, pollution regulations, means tests for social benefits payments and services, income restrictions for Social Security recipients, social insurance payments, wage and hour laws, and minimum wags laws) make "off the books" activity relatively more attractive. In establishing and strengthening such laws and regulations, we, as a society, should consider the incentives they provide "not to report." We would probably be better off with fewer "means tests" and fewer regulations on small business activity.

Currently law enforcement efforts, like research, concentrate on traditional crime areas (e.g., burglary, drugs). One might justify this concentration on equity grounds although it seems odd in terms of the relative dollar magnitudes involved. The effects of a traditional crime usually fall on relatively few individuals and often have a major impact on them (e.g., the individual who is robbed or has his or her house burglarized). In contrast, the effects of a nontraditional crime (e.g., an employee theft or tax evasion) is much more diffuse. We all feel the effect in higher prices, and lowered public services or higher taxes, but the effects for any single individual at any given moment in time are relatively small. If we as a society feel that our criminal justice system should be primarily equity oriented, we may not wish to reallocate law enforcement efforts. As economists, we are not prepared to judge.

One possible alternative suggested by Senator Bible (in the area of cargo theft) it to use civil law and large fines to combat such things as tax evasion and employee theft. Since it seems likely that both of these crimes are primarily economically motivated, large fines associated with a high probability that such fines are imposed may prove quite effective deterrents. There is at least one important difficulty with this suggestion. It has the potential of creating a dual system of justice with obviously nonconforming lower class individuals being tried largely in criminal court and the apparently conforming, primarily upper or middle class individuals being tried primarily in civil court. Needless to say such a prospect is distressing on equity grounds. However, civil trials are certainly better than no trials at all.

Bibliography

Anderson, P. S. 1966. "The Review of Currency in Circulation During World War II," *National Banking Review*, (March) 359–367.

——— 1977. "Currency in Use and in Hoards," *New England Economic Review*, (March/April) 21–30.

Battle, B. P. and P. B. Weston. 1954. *Arson, a Handbook of Detection and Investigation.* New York: Arco Publishing Company.
Bequai, A. 1978. *White Collar Crime: A 20th-Century Crisis.* Lexington, Mass.: Lexington Book.
Blakey, G. R. and M. Goldsmith, 1976. "Criminal Redistribution of Stolen Property," *Michigan Law Review* 74 (August) 1512–1626.
Boudreau, John F., Quon Y. Kwan, William E. Faragher, and Genevieve C. Denault. 1977. *Arson and Arson Investigation: Survey and Assessment.* Report to U.S. Department of Justice, Law Enforcement Assistance Administration. Washington, D.C.: G.P.O.
Cagan, Philip. 1958. *The Demand for Currency Relative to the Total Money Supply.* New York: National Bureau of Economic Research.
Canes, Michael E. 1976. "The Market for Pro-Football Betting." in Eadington. *Gambling and Society.* Springfield, Ill.: C. C. Thomas.
Commission on the Review of the National Policy Toward Gambling. 1976. *Gambling in America. Final Report.* Washington, D.C.: G.P.O.
——— 1976b. *Gambling in America. Appendix 2. Survey of American Gambling Habits and Behavior.* Washington, D.C.: G.P.O.
Comptroller General of the United States. 1976a. *How the Internal Revenue Service Selects Individual Income Tax Returns for Audit,* report to the Joint Committee on Internal Revenue Taxation. Washington, D.C.: General Accounting Office.
——— 1976b. *Audit of Individual Tax Returns by the Internal Revenue Service,* report to the Joint Committee on Internal Revenue Taxation. Washington, D.C.: General Accounting Office.
——— 1979. *Who's Not Filing Income Tax Returns? IRS Needs Better Ways To Find Them and Collect Their Taxes.* Washington, D.C.: General Accounting Office.
Denison, E. F. 1978. "Effects of Selected Changes in the Institutional and Human Environment Upon Output Per Unit of Input." *Survey of Current Business* LVIII, (January) 21–43.
Emerson, F. E. 1971. "They Can Get it for You Better than Wholesale," *New York Magazine* 22 (November) 34–37.
Friedland, N. J., J. Thibaut and L. Walker. 1973. "Some Determinants of the Violation of Rules," *Journal of Applied Social Psychology,* 2, 265–270.
Fund for the City of New York. 1972. *Legal Gambling in New York: A Discussion of Numbers and Sports Betting.*
Fuqua, Paul. 1978. *Drug Abuse: Investigation and Control.* New York: McGraw-Hill.
Gillespie, Robert W. 1978. "Heroin Addiction, Crime and Economic Cost," *Journal of Criminal Justice,* 6, 305–313.
Gottfredson, M. R., et al., 1978. *Sourcebook of Criminal Justice Statistics—1977.* Washington, D.C.: U.S. Government Printing Office.
Groves, H. M. 1958. "Empirical Studies of Income Tax Compliance," *National Tax Journal,* 11 (December) 291–301.
Gutmann, P. M. 1977. "The Subterranean Economy," *Financial Analysts Journal,* 33 (November/December) 26–27, 34.
——— 1978. "Off the Books," *Across the Board,* 15 (August) 9–14.
——— 1979. "The Subterranean Economy," *Taxing and Spending* (April) 4–11.
Henry, James. 1976. "Calling in the Big Bills," *Washington Monthly* (May) 27–33.
Holahan, John. 1972. "The Economics of Heroin," in Wald, P. M. and P. B. Hutt, *Dealing with Drug Abuse.* New York: Praeger Publishers, 255–299.
Horn, J. 1976. "Building-burners for Hire." *Psychology Today,* 9 (February).
Hunt, L. and C. Chambers. 1976. *The Heroin Epidemics.* New York: Spectrum Publications.
Icove, D. J. 1979. Ph. D. Dissertation, University of Tennessee, Department of Engineering Science and Mechanics.
Jasso, Guillermina. 1979a. "Attitudes toward International Migration Among Texans: Advance Report," draft Immigration and Naturalization Service.
——— 1979b. "Deportable Aliens in the United States: A Brief Summary of Research Questions and Research Findings," Working draft U.S. Immigration and Naturalization Service.
Joint Committee on Taxation. 1979. *Issues in the Classification of Individuals as Employees or Independent Contractors.* Washington, D.C.: U.S. Government Printing Office.

Karchmer, Clifford. 1977. "The Underworld Turns Fire Into Profit." *Firehouse Magazine* (August), reprinted in *Congressional Record-Senate*, October 4, 1977.
Karchmer, C. et al. Battelle Law and Justice Study Center. 1978. "Development of a Technical Assistance Manual on Strategies to Combat Arson-for-Profit."
Klockars, Carl, 1974. "The Fence: Caveat Emptor Caveat Vendor," in T. P. Thornberry and E. Sagaris (ed.). *Images of Crime: Offenders and Victims.* New York: Praeger.
Lancaster, C. and F. J. Scheuran. 1978. "Counting the Uncountable Illegals: Some Initial Statistical Speculations Employing Capture-Recapture Techniques," *Proceedings of the American Statistical Association, 1977.* Washington, D.C.: American Statistical Association.
Lehman, S. C. 1976. "Relationship Between Personal Income and Taxable Income, 1947–74," *Survey of Current Business*, 56 (December), 17–19, 27.
Lemkau, Paul V. et al. 1975. *Social and Economic Costs of Drug Abuse.* Baltimore: Johns Hopkins University.
Little, A. D., Company. 1975. *Social Costs of Drug Abuse.* Cambridge, Mass.: A. D. Little Co.
Melnick, Daniel and Royce Crocker. 1976. "A Review of Two Studies of Gambling in the United States." in: Commission on the Review of the National Policy Toward Gambling (1976b).
Moore, Mark. 1977. *Buy and Bust.* Lexington, Mass.: Lexington Books.
New York Times. 1979. "Illegal Aliens in New York: A Life of Fear, Costly to All," March 18, 1979, 1, 40.
Newsweek. 1974. "Detroit: the Torches' Game." November 25, 1974. 42–43.
O'Donnell, Joyn A., H. L. Voss, R. R. Clayton, G. T. Slatin, and R. G. Room. 1976. *Young Men and Drugs—A Nationwide Survey,* National Institute on Drug Abuse Research Monograph Series, No. 5, Washington, D.C.: U.S. Government Printing Office.
Office for Drug Abuse Prevention. 1973. *1973 Annual Report.* Washington, D.C.: U.S. Government Printing Office.
Office of Drug Abuse Policy. 1978. *1978 Annual Report.* Washington, D.C.: U.S. Government Printing Office.
Park, Thae. 1978. "Reconciliation of BEA Personal Income with SOI–AGI" Working Paper Government Division, Bureau of Economic Analysis, U.S. Department of Commerce.
President's Commission on Law Enforcement and Administration of Justice. 1967. Task Force Report: *Crime and Its Impact—An Assessment.* Washington, D.C.: U.S. Government Printing Office.
Report of the House Select Committee on Population. 1978. *Legal and Illegal Immigration to the United States.* Washington, D.C.: U.S. Government Printing Office.
Reselius, T. and D. Benton. 1973. "Marketing Theory and the Fencing of Stolen Goods," *Denver Law Journal* 50 (No. 2) 177–205.
Reuter, Peter and Jonathan Rubinstein. 1978. "Fact, Fancy and Organized Crime." *The Public Interest* 53 (Fall). 45–67.
Roumasset, J. and J. Hadreas. 1977. "Addicts, Fences, and the Market for Stolen Goods," *Public Finance Quarterly,* 5 (April) 247–272.
Rufener, Brent L., J. V. Rachel, and A. M. Cruze. 1976. "Management Effectiveness Measures for NIDA Abuse Treatment Programs, vol. I: Cost Benefit Analysis." Research Triangle Park, N.C.; Research Triangle Institute.
———. "Management Effectiveness Measures for NIDA Drug Abuse Treatment Programs, vol. II: Costs to Society of Drug Abuse." Research Triangle Park, N.C.; Research Triangle Institute.
Schwartz, R. D. and S. Orleans. 1967. "On Legal Sanctions," *University of Chicago Law Review,* 34 (Winter), 274–300.
Silverman, L. P. and Nancy Spruill. 1977. "Urban Crime and the Price of Heroin," *Journal of Urban Economics,* 4, 80–103.
Spicer, M. W. and S. B. Lundstedt. 1976. "Understanding Tax Evasion," *Public Finance/Finances Publiques,* 31 (No. 2), 295–305.
Strategy Council on Drug Abuse. 1975. *Federal Strategy for Drug Abuse and Drug Traffic Prevention 1975.* Washington, D.C.: U.S. Government Printing Office.
Strümpel, B. 1969. "Contribution of Survey Research," in A. Peacock (ed.), *Quantitative Analysis in Public Finance.* New York: Praeger, 29–32.

U.S. Department of Commerce, Bureau of Domestic Commerce. 1976. *The Cost of Crimes Against Business*. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of the Treasury, Internal Revenue Service. 1979. *Estimates of Income Unreported on Individual Income Tax Returns*. Washington, D.C.: U.S. Department of the Treasury, Publication 1104 (9–79).

U.S. House of Representatives, Subcommittee on Oversight, Committee on Ways and Means. 1979. "Statement" by Nancy H. Teeters, Member of the Board of Governors, Federal Reserve System, *Underground Economy Hearings*, September 10.

U.S. News and World Report. 1979. "Cheating on Taxes—A World-Wide Pursuit. October 22. 53, 56.

Walsh, Marilyn, 1971. *The Fence—A New Look at the World of Property Theft*. Westport, Ct.: Greenwood.

Washington Post. 1976. "Nader Says IRS Is Soft on Fat Cats." November 15, 1976.

Appendix A. The Method of Estimating Income Unreported for Tax Evasion Purposes

Table A–1 contains IRS's estimates of "factor income" unreported by individuals *who filed tax returns* in 1965, 1969, and 1973. These estimates were derived by IRS's Tax Compliance Monitoring Program (TCMP). By "factor income", we mean, of course, income derived from the provision of the services of factors of production (land, labor, capital). This is the only income which enters reported National Income and which we will enter in our estimates of underground National Income to ensure comparability. Such income specifically excludes transfers, such as alimony and tax refunds and income from asset sales (a stock) such as personal capital gains. It also excludes income from most nonmarket production.

Comparing these estimates of underreported factor income for 1965, 1969, and 1973, one notes rapid growth in estimated unreported income during this period. Indeed, the percentage increases for underreported factor income exceeded the percentage increase for reported personal income by 15 percent in the 1965 to 1969 period and by 5 percent in the 1969 and 1973 period. These estimates support the claims that this sector of the underground economy has been growing rapidly in recent years and is increasing even relative to reported economic activity. During the 1969 to 1973 period, the rate of increase in underreported income was greatest for interest and estate and trust income. Perhaps, most disturbing was the marked increase in unreported wage income during this period. It is interesting to note that the rate of growth of the estimated underreported income slowed down in the early 1970's.

TABLE A-1.—ESTIMATED UNDERREPORTING OF INCOME BY TYPE, TCMP ESTIMATES

[Dollar amounts in millions]

Type	Year			Percent increase 1965–69	Percent increase 1969–73
	1965	1969	1973		
Wages	$1,683.0	$1,954.8	$2,891.0	16.1	47.9
Dividends	369.7	544.6	908.0	47.3	66.7
Interest	264.6	471.7	917.3	78.3	94.5
Nonfarm business	4,001.3	7,131.0	9,805.8	78.2	37.5
Farm business income	1,352.3	2,071.6	3,353.5	53.2	61.9
Rents	867.4	1,235.6	2,102.2	42.4	70.1
Royalties	62.2	62.1	85.2	0	37.2
Partnership income	1,081.5	1,460.8	2,060.3	[1] 66.1	[1] 45.6
Estate and trust income	25.2	82.4	152.5	327.0	85.1
Small buiness corporation income	(2)	335.5	555.6	(2)	65.6
Other	1,147.7	1,378.2	1,620.2	20.1	17.6
Total	10,854.9	16,728.3	24,451.6	54.1	46.2

[1] Rate of change is for Partnership Income and Small Business Corporation Income.
[2] Not available. In this case this particular item was included in Partnership Income for this year.

Source: TCMP phase III, cycles 2, 3, and 5. This data was supplied to the Subcommittee on Oversight in the House Ways and Means Committee by IRS and was kindly made available to us by Robert Cozart, a staff attorney for that subcommittee.

In its recent report, the Internal Revenue Service (U.S. Department of the Treasury, 1979) noted that a small portion (0.054) of unreported income uncovered by TCMP comes from illegal sources. To avoid double counting when adding estimates from part I and part II we exclude this proportion of underreporting from TCMP estimates to otain our estimate of wage and salary underreporting for 1973, $2,735 million.

The TCMP unreported income estimates cited above substantially underestimate the actual amount of unreported income for a number of reasons. First of all, these estimates do not include the non-reporting and underreporting by large corporations. Many believe that such underreporting is very large, and perhaps increasing due to audit declines. Audits of multinational corporations have decreased by as much as 40 percent. One former IRS agent has charged that the IRS lacks the capability to audit giant multinational firms, claiming that its inability to audit A.T. and T. completely resulted in a failure to collect more than $2 billion in taxes. (Bequai, 1978, p. 126; *Washington Post*, November 15, 1976).

Second, the TCMP estimates are low because they reflect only unreported income turned up by a complete audit. Such an audit is unlikely to uncover undocumented income such as that paid in cash or in bartered goods. Third, the TCMP estimates cited above are only estimates of unreported income for tax payers who actually filed and thus do not reflect unreported income of nonfilers.

To remedy this last omission, the U.S. General Accounting Office (GAO) conducted a careful study of the "delinquency gap" between the 68 million Americans who were required to file Federal income tax returns in 1972 and those who actually did file returns. In its report, issued in July of 1979, (Comptroller General of the United States, 1979) the GAO estimated that about 5 million people, with aggregate taxable incomes of $26 to $35 billion, did not file returns in tax year 1972. Most of these nonfilers were in the low income range with 52 percent of the total nonfiler population making an income of less than $5,000. However, many high income self-employed managers, administrators and crafts persons also failed to file.

As we did when estimating underreported income, we want to include only income from the sale of factors of production and to exclude such income as alimony and pension payments, state tax refunds, and capital gains. Since we have no breakdown into income components, as we did for those who underreported their incomes, for nonfilers, we assume the same ratio, 0.86, of factor income to total income holds for unreported income as it does for underreported income. This leads to an estimate of $21 billion to $30 billion for unreported taxable factor income for nonfilers in the tax year 1972.

To remedy the second omission, the International Revenue Service in its recent report estimated underreporting not uncovered by TCMP to be $12 billion in 1976. To obtain estimates of total tax evasion associated with legal income sources in 1974, the year we are attempting to estimate total underground National Income, we must: (1) move 1973 TCMP figures for unreported income of filers and 1972 GAO estimates of income of nonfilers forward to 1974; and (2) move 1976 IRS figures for non-TCMP underreporting backward to 1974. We do this by assuming that growth in underreported and unreported income between 1972 and 1976 was approximately equal to the growth in underreported income between 1969 and 1973—10 percent per annum—as indicated in table A–1. This leads us to an estimate of $36.7 billion for underreported 1974 income and an estimate between $25.4 billion and $36.3 billion for unreported 1974 income, and a sum between $62.1 billion and $73 billion. We feel that the larger number is likely to be closer to the truth due to the failure to include corporate underreporting.

To double-check these figures, we turn now to our second source of the amount of income which is unreported to the IRS, the Department of Commerce's Bureau of Economic Analysis (BEA) comparison of National Income Account estimates of Adjusted Gross Income (AGI) and AGI actually reported on Federal (IRS) income tax returns. Table A–2 reports the difference between these two figures for the 1947–1976 period.

TABLE A-2.—ESTIMATED UNDERREPORTING OF INCOME BY TYPE, BEA ESTIMATES

	Total (billions)	Type of income			Total
		Interest	Dividends	Other	
Year:					
1947	$23.1				
1948	22.9				
1949	23.3				
1950	23.1				
1951	26.3				
1952	25.8				
1953	27.0				
1954	24.8				
1955	25.6				
1956	26.7				
1957	26.6				
1958	30.3				
1959	29.0				
1960	30.9				
1961	29.5				
1962	30.0				
1963	29.7				
1964	34.4				
1965	37.4				
1966	40.6				
1967	37.0				
1968	41.5				
1969	41.5				
1970		13.9	0.9	30.8	45.6
1971		13.5	.9	31.9	46.3
1972		13.3	2.4	31.5	47.2
1973		14.0	3.1	39.5	56.6
1974		18.4	3.6	35.5	57.5
1975		20.5	3.3	37.5	61.3
1976		22.6	5.8	37.1	65.4

Note: Percent increase 1965–69, 11 percent; percent increase 1969–73, 36 percent; percent increase 1970–76, 43 percent; percent increase personal income 1970–76, 72 percent.

Source: Lehman (1976) and Park (1978).

BEA warns that one should not consider this difference a measure of tax evasion for three reasons: (1) It includes amounts earned by low income individuals not required to file tax returns; (2) it includes income subsequently uncovered by IRS audit programs; and (3) it includes errors and omissions in the reconciliation items in personal income. Henry (1976) has estimated that the first source of discrepancy probably amounted to between $6 and $8 billion in 1975. Item (2) shifts unreported income in one year to the reported category in a subsequent year, but does not really affect the relative distribution of productive activities at the time of production. Thus, we feel we can safely ignore this adjustment. The overall statistical discrepancy in the national income accounts in 1974 was $6.6 billion. What part of that resides in non-reported adjusted gross income is difficult to say. Assuming that unreported adjusted gross incomes contains its proportionate share (4 percent) of this discrepancy, the unreported income figure is inflated by $.264 billion in 1974. Adjusting 1974 unreported income for items (1) and (3) one obtains an estimate of unreported income resulting in tax evasion of $50.3 billion in 1974.

Estimates of the amount of income unreported for tax evasion purposes obtained from BEA data reflect only income which leaves a "paper trail." Thus, we would expect this figure to be below that obtained by using IRS and GAO figures. Indeed our estimate of income unreported for the purpose of tax evasion in 1974 using BEA data is approximately $12 below our estimates obtained using IRS and GAO data. Considering both information sources we feel confident that the true amount of income unreported for tax evasion purposes and not estimated in other sections of our paper is between $60 and $75 billion with estimates in the low $70 billion range having the greatest probability of being correct. It would seem possible to obtain much more accurate estimates if IRS and BEA were to cooperate and pool their data, an effort on a long term basis as they did for portions of the recent IRS report.

One interesting difference in trends between BEA and IRS estimates of unreported income is that IRS estimates indicate growth rates of unreported income above that of personal income while BEA estimates the opposite. Interestingly, both BEA and IRS estimates point to dramatic increases in failure to report interest income.

Appendix B. The Method of Estimating the National Income of the Stolen Goods Market

In this appendix, we will estimate the National Income of the stolen goods market. First, however, we must decide what are the factors of production in this market, what activities in the stolen goods market should be considered in our attempts to use National Income accounting methods to compute its National Income, and finally how do we value the results of various "productive" activities in this market. Some authors, e.g., Henry (1976) would consider thievery a transfer, albeit an involuntary one, and not include any of it in the National Income of the underground economy under the usual accounting methods. However, the activity of a fence in buying stolen goods from a thief and selling them to other customers within and outside the underground economy is clearly analogous to that of the owner of a legitimate wholesale or retail store and accordingly should be included in underground National Income. Similarly, it is almost as natural to include the economic activity which a thief performs when he sells his stolen goods to a fence. We will include both of these activities in our computations of the National Income of the stolen goods market. However, we will not include the value of the goods which a thief steals and keeps for his own consumption. Nor will we include the money which thieves steal. These latter two seem more analogous to (involuntary) transfers, and therefore are not contributions to current output, i.e., are not a component of National Income. For those who would like to include these excluded activities, we compute a value for them at the end of this appendix.

Our approach will be to first determine the number of property offenses of each type and then to estimate the average value of items stolen. We will multiply these numbers together to determine a value for certain goods and money stolen in 1975. We will make separate calculations using different sources for property crimes against businesses and for property crimes against individuals and their residences. Finally, we will try to estimate what proportion of these stolen goods are sold to fences, what percentage of market value fences pay thieves for these goods, and at what percentage of market value fences sell these goods to their customers. Finally, we estimate the costs of doing business for both thieves and fences.

The most available information on the number of property offenses is the information obtained from the Federal Bureau of Investigation's Uniform Crime Report (FBI's UCR). While this information is extremely valuable and provides the only national information available for an extended time period, the UCR has been criticized on a number of grounds. Perhaps, most importantly for present purposes. UCR information is believed to underestimate seriously the amount of property crime due to non-reporting of offenses.

The crime rate as measured by UCR data increased rapidly during the 1960's by 8 percent a year in the 1960–1972 period. Property offenses of robbery and larceny over $50 increased most rapidly during this period, by approximately 10 percent per year. Thus, based on UCR information alone, one would estimate that this sector of the underground economy grew rapidly during the 1960–1972 period.

Beginning in 1973 the results of national victimization surveys became available. These surveys allow one to estimate the actual amount of all crime, not just reported crime. Results of these surveys for the 1973–75 period indicate that the rate of increase in robbery has declined markedly. At the same time, the rate of increase for larceny remained high (22 percent increase for the two year period) and business burglaries increased at a moderate rate (by 9.6 percent during the 2-year period). Thus, the growth rate of traditional sources of stolen goods may actually have decreased in the early 1970's.

The word traditional is important. Victimization surveys cover only robbery and burglary in their data on commercial establishments. However, by far the majority of business losses come in other areas such as cargo theft, employee theft and shoplifting. The Bureau of Domestic Commerce (BDC), U.S. Department of Commerce, has since 1971 made estimates of the cost of "ordinary" crime to business for certain sectors. Table B–1 contains these estimates for

the 1971–1976 period. As can be seen in this table, the estimated loss to the business sectors studied grew from $12.2 billion in 1971 to $25 billion in 1976, an annual average rate of growth of over 15 percent. These figures seem to indicate that property crime against business may be one of the more dynamic sectors of the underground economy in recent years. The BDC figures exclude construction and agriculture. While no dollar figure is available for construction, a 1972 survey indicated that between 21 and 24 percent of building contractors suffered losses due to theft in 1971 (U.S. Department of Commerce, Bureau of Domestic Commerce, 1976). As far as we are aware, there exists no estimate of theft loss by governments.

Table B–1 indicates that the relative theft loss is highest in retailing and wholesaling and lowest in manufacturing. In recent years, the *rate* of growth in theft loss has been the highest in the service and wholesaling industries. It is interesting that loss in the transportation industry has tapered off in recent years.

TABLE B–1.—ESTIMATED LOSS DUE TO PROPERTY OFFENSES AGAINST BUSINESS—BY SECTOR

[Dollar amounts in billions]

Business sector	1971	1973	1974	1975	1976	Percentage increase 1971–76	1974 loss as a percent of 1974 industry GNP
Retailing	$4.8	$5.2	$5.8	$6.5	$8.1	69	4 (Retailing and Wholesaling combined)
Wholesaling	1.4	1.8	2.1	2.4	3.4	143	
Manufacturing	1.8	2.6	2.8	3.2	4.3	139	1
Services	2.7	3.2	3.5	4.3	6.7	148	2
Transportation	1.5	1.7	1.9	2.3	2.5	67	2
Total	12.2	14.5	16.1	18.7	25.0	104	

Source: U.S. Department of Commerce, Bureau of Domestic Commerce, 1976, p. 7 and Economic Report of the President, 1977, p. 193.

There is one final source of information on the amount of property stolen with which we can observe the trends in the stolen goods market: FBI reports on stolen and recovered property. Table B–2 presents this data for the 1960–1975 period. As can be seen in this table, the overall rate of increase in reported property theft has been very large although it has tapered off somewhat in recent years. The largest increase has been in the area of miscellaneous goods which are believed to make up a large portion of the merchandise on the stolen goods market.

TABLE B–2.—STOLEN PROPERTY IN DOLLARS PER 100 PEOPLE IN CURRENT DOLLARS

	Total	Auto	Miscellaneous [1]	All other [2]
Year:				
1960	$502	$246	$112	$144
1961	508	249	112	147
1962	535	267	124	144
1963	679	346	159	174
1964	824	445	190	189
1965	840	445	190	205
1966	831	457	190	184
1967	991	535	276	180
1968	1,152	588	305	259
1969	1,287	656	375	256
1970	1,356	637	445	275
1971	1,483	653	525	305
1972	1,349	588	490	271
1973	1,375	558	549	268
1974	1,587	579	664	344
1975	1,979	737	812	428
Annual rate of growth (percent):				
1960–69	11.3	11.5	14.4	6.6
1970–75	7.8	3.0	12.8	9.2

[1] Includes all property not included in other categories such as office equipment, firearms, household goods, consumable goods and livestock.

[2] Includes clothing, furs, currency, and jewelry.

Source: Blakey and Goldsmith, 1976, p. 1617.

Let us now put some of this information together to estimate the size of the stolen goods market. We need to know the volume of goods entering this market and the value of these goods at all levels of the distribution chain. We will use estimates of: (1) The total number of certain property crimes from the victimization survey; (2) the dollar loss to certain business sectors from the BDC study; and (3) the value of property stolen and reported to the police. For the offenses that it covers, we believe that the victimization survey provides the most complete coverage. However, this data gives only the number of offenses, not the value of goods stolen. Fortunately, the FBI in its Uniform Crime Reports gives estimates of the average loss incurred in various types of property offenses. If one multiplies these figures by the estimated number of offenses from victimization surveys, one obtains estimates of loss for each offense covered. Summing these, one obtains an estimate of approximately $6 billion property loss in 1975 for the offenses covered by the victimization survey (see table B-3). Since this survey excludes white collar crimes, employee theft, and shoplifting against businesses, we will use the victimization surveys to estimate only the number of crimes against individuals and their residences. We will then use the BDC data to estimate the losses of business because of property crimes.

TABLE B-3.—ESTIMATED VALUE OF STOLEN GOODS IN 1975

	Estimated number	Average loss	Estimated loss
Offense:			
Robbery	1,383,099	$133	$189,952,167
Burglary	8,207,303	422	3,463,481,866
Larceny (except motor vehicles)	9,670,663	166	1,605,330,058
Motor vehicle theft	508,472	1,457	740,843,704
Total			5,993,607,795

Source: Gottfredson et al., 1977, pp. 303, 463.

To exclude the amount and value of robberies and burglaries against businesses from the victimization study data in table B-3, we assume that individuals and residences accounted for 77 percent of the victimizations for these two offenses and 60 percent of the value lost. Multiplying the figures in the first two rows of the last column of table B-3 by 0.6 and adding this result to items 3 and 4 in the last column yields an estimate of $4.5 billion in property theft against individuals in 1975. Adding this to BDC estimates of business loss due to property theft for five industries in table B-1 ($18.7 billion), one obtains a conservative estimate of the value of property stolen of $23.2 billion. Assuming further that government loses the same proportion of its GNP to theft as the service industry (2 percent), that construction loses the same proportion of its GNP to theft as the wholesale and retail sector (4 percent) and that agriculture loses the same proportion of its GNP as manufacturing (1 percent), one obtains an estimated loss for these three industries of $5.6 billion in 1975. Adding this to our previous figures, one obtains an estimate of $28.8 billion worth of property stolen in 1975.

We now have an estimate, however rough, of the value of goods stolen in 1975; however, we want an estimate of the value of goods changing hands on the stolen goods markets. To obtain this estimate four adjustments are necessary. First, one must adjust for the loss of value which goods suffer when being traded in illegal markets. Second, one must adjust for the amount of stolen goods retained for personal use and thus not entering the illegal market. Third, one must adjust for the proportion of loss from property crime in the form of money rather than goods. Finally, one must adjust for the proportion of goods which do not enter stolen goods markets because they are recovered.

The President's Commission on Law Enforcement and the Administration of Justice (1967) estimates the amount received by thieves for stolen goods may be between 20 and 33 percent of market value. This seems quite consistent with the thieves' proverb: "When you take something to a fence you should try to get a third of the value of the goods" (Klockars, 1974, p. 114). Fences then mark up the prices of goods received. The extent of the mark-up will depend on the relevant (wholesale or retail) legitimate price for the goods, market conditions, the length of the distribution system, and the degree of legitimacy established.

Table B–4 illustrates the prices charged by a small part-time New York thief/fence in 1971. Since this fence seemed to specialize in the sale of clothing to others in his neighborhood without the use of a legitimate front, the percent of retail price charged for these items may be considered a lower bound on the proportion of retail price that a fence will charge a retail customer. Another fence, studied by Klockars (1974), used a second-hand merchandise store as a front and charged approximately 50 percent of retail price to retail clothing customers and 34 percent of retail price to wholesale customers on the same merchandise.

TABLE B–4.—SELECTED PRICES CHARGED BY A THIEF/FENCE, NEW YORK CITY, 1971

Item	Legitimate retail price	Stolen market price	Percent stolen market of legitimate price
Mack tractor-trailer	$35,000	$9,000	26
1971 Cadillac Fleetwood	9,500–9,700	2,500	26
Black diamond mink coat	3,000	900–1,300	[1] 37
1.85-carat diamond ring	975	200	21
RCA color TV (23-in screen)	497	100	20
Minolta SRT 101 single lens reflex camera	265–385	65	20
Man's wool overcoat	200	60–75	34
Man's wool suit	185	50	27
Man's suede overcoat	165	60	37
Olivetti Lettera 36 portable electric typewriter	180	40–50	25
Woman's suede midi coat	115	45	40
Bulova Accutron watches	110–195	45–60	34
Man's leather jacket	110	45	41
Singer sewing machine (touch and sew)	349	65	19
Woman's fur-trimmed coat	175	60	34
Whirlpool 13,500-Btu air-conditioner	319	100	31
6 ft by 9 ft Rya rug	159	45	28
Johnnie Walker Red (case)	[2] 79.92	40	50
American Tourister 3-suiter	63	20	32
Sunbeam Mixmaster electric mixer	60	15	25
Jonathan Logan dresses	24–34	7–10	29

[1] When ranges are given the midpoint is used.
[2] Plus tax.

Source: Emerson, 1971, p. 39.

Considering all of the above, we estimate that thieves on the average receive 25 percent of the retail value of the goods they steal, that small time fences with no cover resell these goods for approximately 30 percent of the retail price, and that "professional fences" charge approximately 50 percent of retail to retail customers and 30 percent of retail to wholesale customers. We could find no information on prices further down the stolen redistribution system, but given that most individuals who bought wholesale from Klockar's fence were "legitimate" wholesalers and retailers, it seems likely that prices quite like the prevailing legitimate price were charged by these individuals.

To continue our calculations on the National Income of the stolen goods market, we want to estimate what proportion of the $28.8 billion of goods and currency stolen in 1975 actually reached the level of the fences. Using FBI data, Walsh (1976, p. 123) estimates that 16 percent of all property stolen was currency in 1974. Applying this percentage to our 1975 data leads one to believe that $24.2 billion in goods and $4.6 billion in currency was stolen in 1975. Again using FBI data, the classification system of table B–2, Blakey and Goldsmith (1976) estimate that 37 percent of all stolen autos, 41 percent of miscellaneous goods, and 22 percent of all other goods were recovered in 1975. We could find no estimates of the proportion of stolen property kept by thieves for their personal use; however, 10 to 25 percent does not appear an unreasonable estimate.

Based on the literature cited above, we will now make some heroic assumptions to obtain estimates of incomes to thieves and fences who participate in the stolen goods market. First, we estimate aggregate 1975 income for thieves. Recall that we estimated that $28.8 billion in property was stolen in 1975. Assuming that 16 percent of this was currency, one arrives at an estimate of $24.2 billion in tangible property stolen in 1975. Assuming that thieves kept 10 to 25 percent of this for personal use, that 30 to 35 percent of goods are recovered, and that 90 percent of the goods recovered have not passed through the stolen goods market, we estimate that $13 to $15 billion worth of goods

entered the stolen goods market in 1975. Assuming that thieves receive 25 percent of the value of the goods they steal, we would estimate that thieves had a gross income of $3.25 to $3.75 billion dollars from the stolen goods market in 1975. To obtain net income, we must subtract the thieves' costs of doing business which we assume to be approximately 10 percent. This gives thieves an estimated net income of $3 to $3.4 billion in 1975.

Next, we estimate the income that fences receive from stolen goods sales. This is more difficult because fences come in a wide variety of shapes and styles. Assuming that on the average stolen goods are eventually sold to consumers for between 50 and 75 percent of retail price, one obtains an estimate of total gross fence income of $3.25 to $7.5 billion in 1975. Assuming business expenses equal to 25 percent of gross income, one obtains an estimate of net fencing income of $2.4 to $5.6 billion in 1975. Adding this to the income of thieves as estimated above, one obtains an estimate of "National Income" for the stolen goods sector of the underground economy of between $5.6 and $9.4 billion in 1975. Moving this 1975 figure backward in time using the consumer price index as a deflator, one obtains an estimated 1974 National Income for the stolen goods industry of between $5.1 and $8.6 billion.

As noted above, we have not included in our estimate of the National Income of the stolen goods markets either the amount of currency stolen or a value for the goods stolen by thieves and kept for their personal use. The above calculations estimate that 16 percent of all property stolen was currency in 1975, leading to our estimate that $4.6 billion in currency was stolen in 1975. We also estimated that thieves kept 10 to 25 percent of all property stolen for their own use. This would amount to $2.4 to $6 billion worth of stolen goods in 1975. Such goods should not be given their market value, however, but should be valued somewhere between their market value and the price which fences would pay for these goods. If we value these items at 50 percent of their market value—twice the amount which we've assumed a fence would pay for them—we are led to value such goods at between $1.2 and $3.0 billion in 1975.

THE EFFECTS OF FOREIGN GOVERNMENT INTERVENTION ON THE U.S. COMPETITIVE POSITION: A PERSPECTIVE OF CHANGE

By Mira Wilkins*

CONTENTS

	Page
Summary	121
I. The Issues	122
II. An Historical View	124
III. General Effects of Foreign Government Interventions	126
IV. Specific Effects of Foreign Government Interventions	130
V. Options for U.S. Government Response	136

SUMMARY

Policies of foreign governments affect the U.S. competitive position. Because we are part of an interdependent world economy, the interventions of foreign governments can have profound, but at times ambivalent, consequences in relation to our competitive position. Interdependence has been increasing in the post-World War II years. In these years the United States has voiced its commitment to an open international economy and its actions have generally adhered to this goal. Do economic changes, specifically the regulations of foreign governments, require new policies?

Promotional and regulatory actions of governments are not new—nor are some of the U.S. problems in the world economy. America's competitive position is associated with its status as a mature nation and as a global leader.

Foreign government interventions can strengthen foreign economies relative to the United States, while in absolute terms both the U.S. and the foreign economies gain. Other foreign government policies, such as those of O.P.E.C. countries, incur direct damage on the ability of U.S. industry to compete. Foreign government interventions that affect the value of the dollar have consequence on the level of inflation at home, and in turn on the competitive position of U.S. industry. Typically, when we think of foreign government actions and the domestic economy we consider trade restraints. The latter seem less crucial than the earlier more basic effects discussed.

In textiles, steel, shipbuilding—traditional industries—the United States is meeting major competition from businesses aided by foreign government interventions. The same is true in automobiles, in which industry in 1950 we had 76 percent of world production, and now have

*Professor of economics, Florida International University, Miami, Fla. I want to thank Douglas Ross for his stimulating ideas and suggestions.

30 percent. Japan has become a formidable factor in the world automobile industry; its automobile industry was protected from competition by its government for two-and-a-half decades. Elsewhere around the world, government involvement in national automobile industries is the norm. Likewise, in the modern data processing industry, America is being challenged by foreign competition, which has been nurtured by foreign government interventions. Japan particularly is showing competitive strength. The Japanese government has been an important factor in creating a favorable environment for Japanese industrial growth. O.P.E.C. government interventions affect the levels of world oil production and the price of America's basic energy source. Never before in history has the U.S. competitive position been so captive to foreign government intervention. The extent of U.S. inflation and the condition of the U.S. dollar are affected by foreign government interventions that in turn affect the U.S. competitive position.

The appropriate responses for the United States to these circumstances are distinct from those of other nations, because of the still pivotal role of the U.S. in the world economy. While options include doing nothing, protectionism, protest (and counterbalancing measures), positive action at home to bolster the U.S. competitive position, international solutions, and awareness of vulnerability, the first two of these must be rejected. Protests to urge others to lower trade restraints and counterbalancing measures against "unfair" competition are possible, although they have not been successful in coping with O.P.E.C.'s monopolistic pricing. Action in the United States to stimulate competitive industries and to encourage productivity growth are desirable. These must be taken in an international context. America must assert world leadership. Interdependence brings vulnerability, and we need to be aware of vulnerabilities. In short, foreign government intervention does have a profound impact on the U.S. competitive position. The U.S. response should involve more and new international initiatives and a renewal of the U.S. leadership in the world economy. Global interdependence is the dominant characteristic of America's changing economy.

I. The Issues

Policies of foreign governments affect the U.S. competitive position. The purpose of this paper is to explore some of the effects. In the post-World War II years, the United States has been the undisputed leader in the world economy. It remains the country with the greatest gross national product. Yet, when once we talked about the United States and the rest of the world, we saw the ramifications of the international interactions spreading outward. A West German could comment, "If the United States caught cold, we got pneumonia." [1] Today, by contrast, Americans must look at the impact of foreign government action on our economy. In this paper, I am not so much interested in comparisons of foreign and U.S. Government measures, but instead with the interactions.

---

[1] See quote in Wall Street Journal, Dec. 5, 1978, referring to an outdated view.

We are part of an interdependent world economy. As Lewis Solomon has put it, "a world economic system has evolved beyond present national regulatory mechanisms." [2] As Richard Cooper has written, interdependence enlarges freedom, since it permits nations to make "a more economical use of limited resources." [3] But, as Solomon, Cooper as well as Robert Keohane and Joseph Nye insist, interdependence also restricts national autonomy.[4] Indeed, Marina v. N. Whitman sees interdependence as threatening national economic autonomy.[5] However we weigh the interactions, we must be aware that because of the interdependence more frequently now the conduct of foreign governments, acting in their perceived national interest, has profound consequences in relation to the U.S. competitive position. It is a truism that the U.S. competitive position can only be understood vis-a-vis the position of our competitors. The interactions are not black and white. There are ambivalences. Measures of foreign governments can clash or coincide with U.S. interests or at times have mixed outcomes.

In the last three decades, U.S. participation in the world economy has deepened. Faster communications and transportation, more travel, migration, trade, capital flows and technology transfers, the expanded presence of United States and foreign-based multinational corporations, as well as the proliferation of governmental transnational organizations (from the International Monetary Fund to the World Health Organization) propel our country into a tangled web of global interconnections. In 1978, in current dollars, U.S. gross national product increased 12 percent, U.S. foreign trade (merchandise exports plus imports) rose 16.6 percent.[6] Compare those figures with the foreign direct investment. That year United States direct foreign investment was up 12 percent. Foreign direct investment in the United States (expanding from a smaller base) mounted 18 percent.[7] The United States is not only the most important home (headquarters) for multinational enterprise, it is also the most important host for (recipient of) foreign-based enterprises.[8] To a great extent, foreign direct investment becomes a good proxy for interdependence. Business regulations of foreign nations that affect competitive vigor spillover into the U.S. economy in part through direct foreign investment.

Throughout the post-World War II years, the United States has accepted the view that the economic health of the world is to the U.S. advantage, that lower tariffs and the removal of barriers to commerce are desirable, that "market distortions" resulting from foreign governmental subsidies and taxes should be minimized, and that the free flow of capital and expansion of enterprise internationally should be encouraged.[9]

---

[2] Lewis D. Solomon, "Multinational Corporations and the Emerging World Order" (Port Washington, N.Y.: Kennikat Press Corp., 1978), 40.
[3] Richard N. Cooper, "The Economics of Interdependence: Economic Policy in the Atlantic Community (New York): McGraw-Hill, 1968), 4.
[4] Robert O. Keohane and Joseph Nye, "Power and Interdependence, World Politics in Transition" (Boston: Little, Brown, 1977).
[5] Marina v. N. Whitman, "A Year of Travail," *Foreign Affairs*, 57:3 (1979), 528.
[6] These percentages are based on numbers given in the "Economic Report of the President 1980," 203, 316. We have used current dollars to keep these figures comparable to the direct foreign investment figures. Exports rose 17.6 percent, imports 15.6 percent.
[7] Survey of Current Business, August 1979, 15, 38.
[8] Federal Reserve Bank of New York Quarterly Review, 4 (Winter 1979–80), 25.
[9] See, for example, Stephen D. Cohen, "The Making of United States International Economic Policy," (New York: Praeger, 1977). The view is still held; at the Bonn Summit in

(Continued)

While this country has generally followed policies in accord with its rhetoric, in fact the U.S. government has, at times, deviated from its ideals. The United States has lowered its trade barriers, but not eliminated them. This nation has subsidized agriculture. Communities in this country are not reprimanded by the Federal government when they offer "lures" to business that distort free market choice. The United States has provided tax subsidies for exports. For balance of payments reasons in the 1960s the nation put restrictions on capital outflows. Domestic regulation of business has increased costs, affecting U.S. business's competitive position.

These measures are mentioned, because although this paper looks at foreign government interventions that have impact on America's competitive position, foreign states are far from alone in acting to provide for their domestic interests—actions with broad consequences in the international arena. While Elizabeth Yaeger points out that "the United States is not the protectionist villain in the open world" [10]—nor is it the saint.[11] What then are the consequence of foreign government interventions on the U.S. competitive position? Do economic changes, specifically those affected by the regulations of foreign governments, require new policies in the United States?

II. An Historical View

In Europe, the years from 1500 to 1800 were called a time of mercantilism. Mercantilism had different aspects in various countries, but in each it was accepted that the state ought to act in a positive fashion to promote its nation's welfare.

When Adam Smith published the "Wealth of Nations" (1776), he countered that the market mechanism worked for the general welfare far better than state intervention. With England in world leadership in the nineteenth century, that nation's advocacy of free trade and reduced government involvement emerged.

The era of minimal state intervention passed rapidly. By the late nineteenth century, around the world, government after government adopted regulatory rules to promote its own national interest. Sometimes regulatory action was limited (tariff protection for infant industry, for example). At times, it included subsidization of transportation or new industry (in the U.S. case, large land grants to the railroads). Other times, it meant government ownership: in much of Europe, for example, public utilities came to be government-owned. The government of each nation that sought to "catch up" with industrialized England developed means to assist its own industry. The

(Continued)

July, 1978, the participants—including the United States—pledged not to allow "sluggish growth, sectoral difficulties, or trade imbalances to serve as pretexts for actions that would undermine the framework of free trade among nations." "Economic Report of the President, 1979," 142. During the Tokyo Round of trade negotiations both the Ford and Carter Administrations were committed in principle to free-trade policies. See William R. Cline, et al., "Trade Negotiations in the Tokylo Round" (Washington: Brookings Institution, 1978), 3.

[10] In comments on a preliminary draft of this paper.

[11] Robert Baldwin, "Nontariff Distortions of International Trade" (Washington, D.C.: Brookings Institution, 1970), 154–165 is useful but dated on "effective protection" in the United States.

United States and Germany with their high tariff policies were no exceptions.[12]

Economic historians of late nineteenth century England dwell on what they call the "climacteric"—a time when the once industrial leader showed itself as a mature economy.[13] Characteristic of that mature economy were large capital exports,[14] an absence of or decline in entrepreneurial vigor (although this proposition has been questioned),[15] a loss of technological initiative,[16] and a rise of "collectivism" (that is, a growing reliance on the intervention of the state).[17] More important, Britain's dominance was challenged by the emergence of the United States and Germany as major powers; each aggressively invaded Britain's domestic and overseas markets. Before World War I American businessmen made direct investments in Great Britain in producing such new products as sewing machines, cigarettes and automobiles, and U.S. business seemed to achieve a competitive edge.[18]

While Britain to be sure experienced economic growth, her progress relative to her rivals seemed less significant. There were, however, Britons who decried and those who denied her loss of status. The sun afterall never set on the British Empire. Britain remained ascendant. Yet, in the interwar years, the "mature" Britain failed in her efforts at world leadership. Increasingly, Britain turned protectionist when faced with foreign competition.[19]

In the post-World War II years that same label "mature economy," which was applied to late nineteenth century-early twentieth century England, now has come to be affixed to the U.S. economy.[20] The parallels are disturbing. The United States is a large capital exporter (although we are also importing capital, the United States remains the world's most important creditor nation).[21]

Questions are asked about what has happened to the entrepreneurs of past times; at the same time, the absence or decline in entrepreneurial vigor is sharply contested. We worry over the loss of technological advantage and note the large number of foreign patents being registered in Washington.[22] We are concerned with a fall in productivity growth. Then, too, the phrase "reluctant collectivism" surfaces as de-

---

[12] Alexander Gerschenkron, "Economic Backwardness in Historical Perspective" (New York: Praeger, 1965) argued that late industrialization involved more government involvement.

[13] Barry E. Supple, ed. "The Experience of Economic Growth" (New York; Random House, 1963), 203–225; and Charles P. Kindleberger, "Economic Response" (Cambridge, Mass.: Harvard University Press, 1978), 222–236.

[14] A. R. Hall, ed., "The Export of Capital from Britain 1870–1914" (London: Methuen & Co., 1968); A. K. Cairncross, "Home and Foreign Investments 1870–1913" (New York: Cambridge University Press, 1953).

[15] Kindleberger, "Economic Response," 232–236.

[16] See Mira Wilkins, "The Emergence of Multinational Enterprise, American Business Abroad from the Colonial Era to 1914" (Cambridge, Mass.: Harvard University Press, 1970), 215, 217.

[17] See A. V. Dicey, "Law and Public Opinion in England" (London: Macmillan 1948), 64–65, 259–302.

The first edition of this book was published in 1905.

[18] Wilkins, "The Emergence," 212–213.

[19] See Charles P. Kindleberger, "The World in Depression, 1929–1939" (Berkeley: University of California Press, 1973).

[20] For a start on the usage, see W. W. Rostow, "The United States in the World Arena" (New York: Harper & Brothers, 1960), 412, and W. W. Rostow, "Stages of Economic Growth" (Cambridge, Eng.: Cambridge University Press, 1960), Chaps. 5 and 6.

[21] See data in "Survey of Current Business."

[22] For example, the number of patents issued by the U.S. patent office to residents of the United States peaked in 1971. By 1977, fully 35 per cent of all new patents issued were to residents of foreign countries. Commissioner of Patents and Trademarks, "Annual Report 1977," 12, 15.

scriptive of this era.[23] Indeed, one of the most vivid parallels to the Great Britain of earlier years is that the United States seems challenged by the rise of other strong economies.

Senator Frank Church has been quoted as saying:

> We alone seem to be transfixed.
>
> The rest of the world has moved into a far stronger economic position. Our own has eroded away to the point that we are no longer able to compete in the international market place successfully . . .
>
> The countries that have done us the greatest economic damage are not the Russians and Chinese . . . but rather those countries we customarily call our allies and trading partners.[24]

By contrast, Ben. J. Wattenberg, a senior fellow at the American Enterprise Institute, argues:

> There is, in my view, a profound sense in this country that we are a nation that remains No. 1, whose sun is ascendant. At the same time, Americans feel that we have been governed for a decade now by men who are afraid to say that or act upon it, by leaders who are seen to be sitting by impotently while American status around the world diminishes.

Wattenberg rallied his audience with: "Don't erode our values. Don't erode our money. Don't erode our status." [25]

The words of Church and Wattenberg give the economic historian an uneasy sense of *deja vu.* The ghost of the late nineteenth century—early twentieth century England seems there to haunt us. In each case, the dominant economy—the leader—was (is) challenged by outside competition. In neither case was that outside competition a creature exclusively of the market place, but rather a consequence of government-business relations.

III. General Effects of Foreign Government Interventions

A range of interventions by foreign governments can affect the U.S. competitive position, including general monetary and fiscal policies, along with specific policies on prices, employment, economic growth, balance of payment, competition, energy, migration, safety and health, ecology and congestion, labor relations, social welfare, as well as consumer and stockholder protection. Foreign government promotion, regulation, and ownership of business have impacts on foreign economies, which in turn in an interdependent world can have significant impacts on the competitive position of the United States. The consequences are not of a piece. Some are negative, others positive, others mixed.

What are the effects? First, in broad terms, measures of foreign governments can result in the enhancement of foreign economies. If foreign economies prosper, what is the result in the United States? Mercantilist thought argued that the rise of competitive, powerful states meant the loss, in position, of other states. This "their gain, our loss" model implied a worldwide economic pie that had to be divided. If one country got more, another received less. But, responded opponents of this view, in an expanding world economy, if there was

[23] See Fred Hirsch, "Social Limits to Growth" (Cambridge, Mass.: Harvard University Press, 1978).
[24] Miami Herald, Dec. 10, 1978.
[25] *Ibid.*, Dec. 17, 1978. Business Week, Mar. 12, 1979, devoted an entire issue to "the decline of U.S. power" and principally to the "failure of leadership."

economic growth, if the economic pie grew in size, everyone would gain. Indeed, the commitment of the United States in the post-World War II years to worldwide economic growth assumed that all would benefit.

Yet if the enhancement of foreign economies through the interventions of their governments means their gain, despite our absence of loss, in fact we do lose in relative terms. This is a very simple proposition. Even if other economies gain and we also gain, if the gain of a foreign nation exceeds ours, then the United States experiences a loss relative to that other country.

In effect, many Americans ask, when the Brazilian government policy favors import substitution (the substitution of local industry for imports), is such Brazilian industrialization good for the American economy? When the Canadian government subsidizes Michelin to build a plant there to fill Canadian and United States market needs, is the result U.S. unemployment? When Mexico requires a foreign investor to export, does this governmental requirement mean "job destruction" in the United States? When the nine-member governments of the European Economic Community allow free trade within the customs union, does that reduce U.S. exports to those European nations? In each of the cases above, actions taken by Brazilian, Canadian, Mexican, and European governments to develop their own national economies and to provide national employment have had specific effects on U.S. industry and employment.

If the United States introduces environmental restrictions that raise the cost of producing goods here and other nations chose not to impose similar regulations, then—all other things remaining equal—those nations automatically achieve a competitive advantage. We are not suggesting that the United States drop, or even reduce, environmental controls; rather that the absence of parallel foreign government regulations places producers in the United States at a competitive disadvantage. Similarly, U.S. rules vis-a-vis the Arab boycott of Israel may open the way for foreign firms to move into what would otherwise have been markets for U.S. goods. Foreign governments have not imposed the same restrictions. So too, the Securities and Exchange Commission's insistence on disclosure and accounting for bribes and the Foreign Corrupt Practices Act of 1977 may be highly desirable; nonetheless, lack of comparable scruples on the part of foreign states may once again result in the diminution of the competitive abilities of American business.

While Brazilian industrialization may lower certain U.S. exports and raise U.S. imports, while Canadian subsidies to industry, Mexican export regulations, and the working of the European Economic Community may have a similar outcome, it can be argued the principal advantage of such governmental regulations to the actor-state is to enhance the competitive strength of its individual nation. Likewise, a comparable consequence results from the absence abroad of cost-creating or moralistic rules that impose burdens equal to those in the United States.

At the same time, it seems likely that the steps taken abroad to industrialize further other economies provide new opportunities for U.S. exports and U.S. direct investments that would not otherwise have

existed. The major trade expansion in the post-war period has been among advanced nations. The more industrialized the countries of the world, the greater will be U.S. exports (as well as imports); [26] likewise, U.S. direct foreign investments are largest in the developed nations.[27]

Recently, Secretary of Labor Ray Marshall told reporters that as the economies of other nations grow, the United States will export more products and this will create added jobs.[28] The same theme was repeated in the "Economic Report of the President, 1979": "policy choices by one country directly affect economic performance in others. If some countries grow very slowly, their trading partners will have to abandon dynamic export industries. . . ." [29] In addition, U.S. foreign investments bring sizable returns to the United States.

The jury is still out on whether the rise of competitive states is in net harmful or helpful to the U.S. economy. Paradoxically, it could well be that it is far better for this country "to gain" by such foreign economic growth and to lose relative position, than "to lose" by acting to curb foreign economic expansion and maintain or gain relative position.

There is another consideration, however. If in their desire to attract investment for national ends, other countries bring in U.S. multinational corporations—as they have done—does this mean U.S. public policy options in enhancing the U.S. competitive position are curtailed. Of course, it does. As Richard Cooper pointed out as early as 1968, "If business domains exceed governmental jurisdiction, effective regulation is greatly limited . . ." [30] While I would prefer "limited" to greatly limited in this formulation, Cooper is right. The limits are several, and include the fact that the multinational firm can (1) draw on different funding sources, using monies available in world rather than national markets (it can in effect bypass national credit restraints); (2) utilize its options in investment strategies to raise investments abroad rather than at home; and (3) through intrafirm trading decisions offset the consequences of a devalued currency.[31] That multinational enterprise may limit U.S. antitrust options is perhaps the least of the concerns.

Nonetheless, despite such limits on U.S. regulatory power, I would suggest that the negative consequences to the United States of most foreign government regulations that serve to bolster foreign economies by attracting multinational enterprise seem more in terms of relative competitive strength than in direct damage to the American economy. In fact, as noted, the increased vigor of the world economy means that the economy of the United States seems to end up better rather than worse off in absolute terms.

It is perhaps worthy of consideration to reflect that in our new world economy of multinational enterprises, strong foreign national economies nurture foreign-owned multinational enterprises and the U.S. economy is stimulated by new investments from abroad. Indeed, there are

---

[26] Jan Tinbergen, coordinator, "Reshaping the International Order" (New York: E. P. Dutton & Co., 1976), 34.
[27] See statistics on U.S. direct investments abroad published annually in "Survey of Current Business."
[28] Miami Herald, Jan. 13. 1979.
[29] "Economic Report of the President, 1979," 142.
[30] Cooper, "Economic Interdependence," 6.
[31] See Solomon, "Multinational Corporations," 39–44.

favorable effects on the competitive position of the United States of "Made in America" Volkwagen cars and Yamaha pianos.

By contrast, when nations in the Organization of Petroleum Exporting Countries unilaterally hike prices, these interventions—also designed to promote the welfare of the nation states—have an immediate, direct, undermining impact on the U.S. competitive position. To be sure, there may be compensating advantages to our economy: OPEC countries become large importers, purchasing U.S. exports. Their nationals invest in the United States, adding to the U.S. capital base. Nonetheless, these advantages far from compensate for the all-negative domestic inflationary effects.

The United States has developed an energy-intensive industry and style of life. Accordingly, the rise in the price of oil has had devastating consequences. The negative inflationary effect on the U.S. economy, which reduces our competitive position, is an immediate outcome of deeds of foreign states. Here is a vivid example of how governments of foreign nations can take regulatory measures that directly and unquestionably affect the competitive position of this country in a deleterious manner.

Interventions by foreign central banks affect the value of the dollar and, in turn, the level of U.S. inflation. Policies of foreign governments that influence the dollar supply have obvious and immediate, domestic consequences. Not only the actions of European governments can have impact, but also, for example, those of the government in the newly-independent Bahamas. To the extent that U.S. banks can avoid U.S. Federal Reserve requirements by using branches in the Bahamas, the power of U.S. monetary authorities is limited by the options provided under the laws and regulations of the Bahamas.[32]

Foreign government interventions need not have mixed consequences—or be harmful to the U.S. competitive position. They can be to our direct advantage. Thus, while the dollar's decline from its 1971 level may have at first improved the competitive position of the United States, in recent years its decline has—most economists would agree—gone too far. Yet, it takes more than the U.S. Government to support the dollar. The cooperation of foreign central banks has been required.

The value of the dollar vis-a-vis other currencies is important to our coping with problems of inflation and to our competitive position. A declining dollar makes U.S. exports more competitive only when not accompanied by serious inflation at home.[33] A declining dollar makes U.S. imports more expensive and that serves to stimulate inflation. Marina v. N. Whitman cites estimates that a 10 percent decline in the dollar can be expected "to raise domestic prices by about 1.5 per cent for any given level of unemployment." [34] Here again, as a participant in a world economy, the regulatory actions of foreign

[32] Andrew F. Brimmer, "Scope and Expansion of American International Banking," Speech before International Banking Symposium, Miami, Nov. 14, 1978, points out that the Nassau "shell branches" have channelled funds from the Euro-dollar market to their head offices in the United States to allow the latter "to cushion the impact of domestic monetary restraint."

[33] These comments are, of course, an oversimplification. This is not, however, the place to go into J-curve effects, lags, and so forth. What is important is that the actions of foreign governments affect the value of the dollar and thus our competitive position.

[34] Whitman, "A Year of Travail," 535.

central banks in aiding the dollar reverberates on the domestic economy.

Indeed, so linked has the world economy become, so much a global village, that while our competitive position and certainly our trade deficits are affected by foreign government restraints on U.S. exports, such barriers to trade so often discussed seem less crucial than foreign government interventions to enhance their domestic economies vis-a-vis the U.S. economy, to raise oil prices, and to alter the value of the dollar. These measures have truly significant impact on the U.S. competitive position. They will command our attention.

IV. Specific Effects of Foreign Government Interventions

In certain basic industries (textiles, steel, and shipbuilding), developing nations are competing with traditional enterprises in the United States and Europe. The British government in Hong Kong and the government of South Korea have offered regulatory environments conducive to industrial development. Of Hong Kong, a Price, Waterhouse "Information Guide" could state that the "government's role in the economy remains one of providing a stable framework within which commerce and industry can function efficiently and effectively with a minimum of interference." To encourage development, that government modified its industrial land policies to be sure shortages of land did not deter new industry.[35] The South Korean government also does more than maintain stability; it actively assists industrial development and exports. For the textile industry, Hong Kong and South Korea offer low wages, plus a favorable regulatory environment.

In steel, America in the post-war period has been disturbed by European and Japanese competition. Now, concerns over imports are present in Europe as well. Both Americans and Europeans are alarmed about Japanese steel imports. In response, the Commission of the European Communities (EC Commission) is pursuing plans to aid the European steel industry, by establishing minimum prices, introducing production limits, and restructuring the industry and its work force.[36] These interventions seem defensive, designed to halt decline. By contrast, Japan has in the post-World War II years developed a steel industry under regulatory conditions that have stimulated efficient, low-cost production. Johannes Hirschmeier and Tsunehiko Yui explain how MITI (Ministry of International Trade and Industry) supported Japan's steel industry investment programs in the 1950s. When the 1957 recession came, MITI helped form a cartel of 33 Japanese steel companies, which reduced output 30 to 50 per cent. Then, in 1959 the Japanese steel industry again expanded,

> . . . not heeding the warnings of MITI to exercise restraint. . . . When the recession began in 1964 the steel makers agreed to limit their output, yet Sumitomo assigned a low quota, refused to go along. Much furor was raised, and eventually President Hiuga of Sumitomo was called to the Ministry of Trade and Industry for consultation, while MITI openly lashed out against the "enterprise egoism" of Sumitomo. Agreement was of course reached.[37]

---

[35] Price, Waterhouse, Information Guide, Sept. 1976, 11. Earlier regulations created today's conditions.

[36] Europe, January-February, 1980, 49.

[37] Johannes Hirschmeier and Tsunehiko Yui, "The Development of a Japanese Business 1600–1973" (Cambridge, Mass.: Harvard University Press, 1975), 262–263.

MITI also aided in "orderly" exports of Japanese steel, promoting exports with tax measures and assigning export quotas. MITI encouraged mergers in the steel industry, on the grounds that the larger the firm the more competitive.

By the end of the 1960s the Japanese steel industry did not need government protection. Liberalization occurred in the 1970s. Then, with the yen increasing in value, the imported inputs for Japanese steel (iron ore, coke) became cheaper and Japanese steel remained competitive. Now, however, Brazil and Mexico (whose steel industries were within our lifetime in their "infancy") are beginning to export.[38] In Brazil, most of the basic steel industry is government-owned.[39] In Mexico, on July 4, 1970, the government required that new firms in steel be 51 per cent nationally-owned.[40] U.S. steel industry officials believe Brazilian and Mexican production will in time become fully-competitive. Do such government-owned companies have special advantages?

Raymond Vernon and Yair Aharoni have two volumes forthcoming on state enterprises. They note that the "interactive process" between managers of state-owned enterprise and their governments carries with it obligations. But, it also generates rewards that include "access to subsidized capital, guarantees against bankruptcy exemptions from import restrictions, preference in government purchases, and relief from regulation. . . . in their international dealings, they (state-owned companies) often get special support from their governments."[41] Such a general statement should be kept in mind, as world markets encounter competition from Mexican and Brazilian steel. One of the key concerns about competition with state enterprises is that they can price without regard to profits.

In shipbuilding, the United States is at a competitive disadvantage. South Korea is becoming significant. Brazil, which was not a factor in shipbuilding a few years ago, has become an active exporter. Foreign government subsidies in shipbuilding have been common for years.[42]

Not only in our oldest industries has foreign government supportive intervention stimulated competitors that challenge the position of the United States In 1950, 76 per cent of the world's automobiles were produced in the United States. In 1978 the figure was 30 per cent.[43]

In the 1970's, Japanese automobile production has risen at a phenomenal pace. Using the latest technology, Japan's passenger car output is now second only to that of the United States (in 1978, 6 million units output compared with 9.2 million U.S. production).[44] Japan has reached first place in world passenger car exports.[45] The Japanese automobile industry grew, protected for two-and-a-half decades from foreign competition by substantial tariff and nontariff barriers. That government gave preferential treatment to Japanese automobile produc-

---

[38] Time, Dec. 4, 1978.

[39] See Richard S. Newfarmer and Willard F. Mueller "Multinational Corporations in Brazil and Mexico." Report to the U.S. Senate, Committee on Foreign Relations, Subcommittee on Multinational Corporations, 94th Cong., 1st sess. (1975), 108, 112.

[40] *Ibid.*, 58.

[41] See report on Vernon and Aharoni's forthcoming work in Explorations, 17 (Feb. 1980), 5.

[42] Baldwin, "Nontariff Distortions", 113–115.

[43] MVMA Motor Vehicle Facts and Figures '78," 20.

[44] *Ibid.*, 21.

[45] *Ibid.*, 80.

ers in its purchases.[46] For many years, the Japanese government blocked multinational corporations from bypassing the barriers to trade and investing behind the tariff walls in the Japanese automobile industry. The Japanese have long imposed steep gasoline taxes that encouraged Japanese producers to design small and light-weight automobiles with fuel-efficient engines. The Japanese government moreover had stringent regulations on quality and emission control, which have helped make Japanese cars more competitive in world markets. The state support to the Japanese steel industry also contributed to that nation's automobile industry's competitive strength.

In the early 1970's Japan began to liberalize its barriers to foreign investment and in March 1978, Japan reduced its automotive tariff to zero. By then its automobile industry did not need protection.[47] In 1977, the latest year for which figures are available, passenger car imports into Japan totaled a mere 41,395 units or about 1 per cent of a 2.9 million unit car market.[48] The reason for the small imports lay in the competitive strength of the Japanese automobile industry.

European automobile production has expanded in the post World War II years. American multinational corporations have participated in this European industry through investments. Government regulation has shaped the growth, influencing product design. Umberto Agnelli, President of Fiat, has declared that he is:

> . . . firmly convinced that the external system of restrictions we have had in Europe, which tend to reward limited consumption [of oil], has stimulated innovation in the search for solutions which are both cost efficient and effective in meeting the requirements and the expectations of the authorities and of the consumers. But each time policy-makers outside the industry have tried to influence technical decisions (for example, the Italian fiscal system based on engine capacity and the number of cylinders), this has had a negative effect, preventing the manufacturer from combing the market for the most suitable solutions to his problems.[49]

European governments have sought to encourage mergers in the automobile industry. Governments have attempted to aid domestic producers. Key European automobile companies have come to be in full, or in part, government owned. In France, Renault became government-owned after World War II. Volkswagen has German government ownership. BL in England is a nationally owned company. Fiat, while private, has always been aided by the Italian state.

Canada too has sought to encourage automobile makers to locate plants there. Canada refused to give U.S.-made Volkswagens duty reductions under the Canadian Automobile Agreement, unless Volkswagen increased its purchases of Canadian-made automobile parts.[50]

Even more, in less developed countries, foreign governments have attempted in the 1960's and 1970's to help build domestic automobile industries. Governments have sought to compel the substitution of domestic output for imports. They have insisted on increasing local content and encouraging exports. In many cases, they have virtually

---

[46] Baldwin, "Nontariff Distortions," 66, and Cline, "Trade Negotiations," 190.

[47] Mira Wilkins, "Multinational Automobile Enterprises and Regulations: An Historical Overview" Douglas H. Ginsburg and William J. Abernathy eds. "Government Technology, and the Future of the Automobile" (New York: McGraw-Hill, 1980), 246–247.

[48] Based on figures in "MVMA Facts and Figures," 80.

[49] Umberto Agnelli, "Policies for the Promotion of Effective Technological Change: A European Point of View," in Ginsburg and Abernathy, "Government, Technology", 219.

[50] Wilkins, "Multinational Automobile Enterprises and Regulation," 245.

barred imports. In the decade 1957–1967, American car makers were forced (the alternative was to lose the market) to raise local content of automotive products to more than 60 percent in Mexico, 95 percent in Argentina, and 100 percent in Brazil. Once these goals were achieved, Argentina permitted new product lines only to those firms that raised exports. Mexican regulations matched the import of components, dollar for dollar, to the value of exports. Brazil gave permission for added expansion only to companies that agreed to a ten-year export program. American multinationals, faced with competition from Volkswagen (in Mexico and Brazil), Fiat (in Argentina), and Toyota and Datsun, complied.[51] In 1977, Brazil reached number 10 in world production of passenger cars—still a long way from the United States (Brazil: 464 thousand units, U.S.: 9.2 million units). The automobile industry in Brazil would never have existed had it not been for Brazilian government interventions to create the conditions for its development.[52]

Thus, although the United States is still the clear world leader in automobile production, its premier role of 1950 has been eroded by the growth of competitive automotive industries in other countries. Such industries have emerged, assisted by foreign government interventions.[53] The United States is now a large importer of automobiles. Our "domestic" industry is becoming transformed. Volkswagen will soon be America's fourth largest domestic producer. Renault is buying shares in American Motors. Honda will assemble cars in Ohio. Toyota and Nissan are studying the possibilities of U.S. production. A recent news item reported that Satra Corp. would build a $1.5 million plant in Savannah, Georgia, to add accessories to the Soviet-built Lada, an economy car! [54] That Volkswagen has partial German government ownership and Renault is 100 percent French government owned *may* make these companies more subject to interventions by their governments. In any case the consequence of their governments' regulations—through the enterprise—can be introduced into our changing U.S. economy. For example, Volkswagen and Renault products have long been designed to be gas-frugal (to meet high gasoline costs, beefed up by European government taxes). They obviously introduce these products in their U.S. business. Douglas Lamont sees dire consequences to the United States of direct foreign investment by foreign government-owned companies.[55] More study is needed of the effects on the U.S. competitive position of direct investments by foreign government-owned entities.

The United States still excels in the world's data processing industry—with IBM as the global leader. Yet by the second half of the 1970's, IBM's technological superiority had narrowed vis-a-vis Europe and had virtually disappeared vis-a-vis Japan. A major reason, IBM argues, is that European and Japanese governments have promoted

---

[51] C. Fred Bersten, Thomas Horst, Theodore H. Moran, "American Multinationals and American Interests" (Washington, D.C.: Brookings Institution, 1978), 376.

[52] For the early history of Brazil's attempts to create an automobile industry, see Mira Wilkins and Frank Ernest Hill, "American Business Abroad: Food on Six Continents" (Detroit: Wayne State University Press, 1964), 414–418.

[53] Wilkins, "Multinational Automobile Enterprises and Regulation."

[54] Miami Herald, Mar. 12, 1980.

[55] Douglas F. Lamont, "Foreign State Enterprise: A Threat to American Business" (New York: Basic Books, 1979).

their computer industries. European governments gave subsidies to national companies and favored the latter in their purchasing. The European Economic Community encouraged mergers in the computer industry, seeking to aid European Community firms in international markets. The Japanese government has protected its national data processing industry with tariffs, sped inflows of technology, financed research and development, provided preferential procurement arrangements, subsidized marketing of Japanese computers, rationalized the industry, offered tax incentives, and in effect, provided an environment for the growth of sophisticated data processing companies that are now beginning to gain a position in the U.S. market.[56]

In short, in many industries—from the traditional to the high-technology modern ones—foreign government interventions have aided the emergence of competitive firms. In Europe and Japan mergers have met with government approval so as to create giant enterprises that can compete with American big business. A critic of a preliminary draft of this paper objected that to emphasize government support neglected "labor costs, technology and productivity gains, exchange rate policy, trade barriers, and export promotion." [57] The last three items are governmental in nature and part and parcel of the governmental assistance that is being offered abroad. Labor costs in Zaire (for example) are very low, but ineffective government policy makes for an absence of competitiveness. Low labor costs by themselves do not make for competitive strength. Technology and productivity gains are, of course, critical to competitiveness (that almost goes without saying), but why does one country have such achievements and not another? My argument is that the new competitiveness of many foreign countries has been enhanced by a supportive role of government. The newly competitive enterprises in these nations have not only contributed to the strengthening of their home economics, but are or are becoming active participants in world markets—through trade and investment.

Ezra Vogel has predicted that if current trends continue Japan "will easily surpass the U.S. in absolute terms as an industrial power in the mid-1980s." [58] The forecast should concern American policy makers. Contributing to Japan's success has been Japanese government promotion of its dynamic industrial sectors. The Japanese government has not hesitated to encourage mergers with an eye to efficiency, to underwrite industrial research and technological innovation, to provide low interest loans to business, and to aid in Japanese market promotion at home and abroad. An official of the Japanese Ministry of International Trade and Industry was recently quoted as explaining the Japanese view. Japan—like other nations—has a few industries where it "will have a comparative advantage over other competitors, for a limited time." It "must invest strategically where it can exploit that comparative advantage rapidly, manufacture on a very large scale, and gain world-wide dominance. Planning is

[56] Information from IBM Corp., November 1977.
[57] Comments of Richard Bartel, May 7, 1979.
[58] See letter from Ezra F. Vogel, in "New York Review of Books," Apr. 3, 1980 and his "Japan As Number One" (Cambridge, Mass.: Harvard University Press, 1979).

crucial." [59] Japan does not expect the market to function freely. The planning is to be by government and business, in tandum.

Foreign government intervention in the oil industry has had an even more profound and direct impact on the U.S. competitive position. It is appalling how unprepared the United States was for the measures of O.P.E.C states in raising pricing in October 1973, and January 1974. "The Economic Report of the President, 1980," begins "Last year world oil prices more than doubled." [60] The Council of Economic Advisers are unequivocal in stating "oil price increases dominated economic developments in 1979." [61] A review of U.S. government policies bears witness to the initial total absence of any recognition that interventions of O.P.E.C. nations could have detrimental economic consequences in the United States.[62] Levels of oil production and prices are determined by foreign government decisions. This means the price and quantity of the basic energy resources for American industry is not within this nation's control. Fully 50 percent of U.S. energy consumption is petroleum.[63]

The oil crisis of the 1980s involves our relations with governments of oil producing states. Will Iran be able to restore and desirous of restoring full oil production and exports (it depends on Iranian government policy); will Mexico sell us adequate gas and oil at prices that we are willing to pay (it is a Mexican government decision); will Saudi Arabia continue to act to control production to meet world requirements (again a foreign government decision)? Sovereignty is not at bay. The actions of these states directly affect U.S. energy supplies. The price of oil is a political one.[64] Never before in history has the U.S. competitive position been so captive to foreign government intervention.

Inflation at home is directly related to our energy costs. Inflation, however, becomes difficult to control—and not only because of the regulatory actions of foreign governments in the sphere of oil prices. Control over our monetary policies is also cushioned by international interdependence, and government regulatory policies abroad.

It is frightening, but certainly realistic, that in April, 1978, German chancellor Helmut Schmidt told European Community leaders that but for German support, the American dollar "would drop like a stone." Europeans, he declared, should act to protect dollar.[65] When President Carter in November, 1978, took steps to rescue our currency, he did so with the cooperation (and intervention) of German, Swiss, and Japanese central banks. He could not have done otherwise. The consequences of President Carter's actions were that when the foreign central banks acquired dollars, they invested them in U.S. Treasury securities, creating further economic interdependence. The declining-dollar crisis has been abated only through the cooperation between foreign central banks and the U.S. government. "The Economic Report of the President, 1980," notes that the dollar is the principal inter-

---

[59] Quoted in New York Times, Mar. 16, 1980.
[60] "The Economic Report of the President," 1980, 3.
[61] *Ibid.*, 156.
[62] Raymond Vernon, ed., "Oil Crisis" (New York: W. W. Norton, 1976). In this volume, see particularly, Mira Wilkins, "The Oil Companies in Perspective," 173–174.
[63] Robert Stobaugh and Daniel Yergin, eds. "Energy Future" (New York: Random House, 1979), 15 (figures are 1977).
[64] See the perceptive views of Daniel Yergin, Miami Herald, Jan. 16, 1979.
[65] Europe, January–February 1979, 6.

vention currency. "For example, if Switzerland wishes to intervene to support the Swiss franc in relation to the German mark, it is likely to do so by selling dollars to buy francs. This can put pressure on the dollar if the seller of the Swiss francs is unwilling to hold the new dollar balances." [66] As noted earlier, there is evidence that the declining dollar has contributed to U.S. inflation. Inflation at home directly affects the competitiveness of U.S. industry. The value of the dollar, likewise, has an important impact on the abilities of the United States to compete in world markets.

Trade deficits reflect America's competitive position. U.S. merchandise imports have exceeded U.S. merchandise exports in every year since 1976.[67] The status of U.S. industry vis-a-vis that of other countries is in part revealed in that deficit.[68] The high price of oil is, of course, a significant factor. So too, inflation and the value of the dollar are associated with our trade position. In all those matters, we have seen that foreign government intervention has had impact. In our interdependent world the interventions of foreign governments are well in evidence at home.

V. Options for U.S. Government Response

The appropriate U.S. Government responses are distinct from those of any other nation. The reason lies in the still pivotal role of the United States in the world economy. Yet because of global interdependence, American policy makers' freedom to make independent decisions is limited. National decision-making must be in the context of America's world role.

In the historical section of this paper, I considered the parallels with the Great Britain of an earlier era. The parallels were not designed to be a prediction of future decline. No determinism was implied. Rather the parallels were provided as a warning. America is at a critical point in its history. Can it as a "mature" nation retain its leadership? Is the United States able to maintain and better still to improve its competitive position? Those who anthropomorphize societies would argue that maturity is followed by inevitable decline. The challenge to a changing economy is not to let the word "mature" turn into a self-fulfilling prophesy.

The response in the United States to the interventions of foreign governments that promote their national economies at the expense of the U.S. competitive position can be one of "do nothing." We can avoid the problem, deny its seriousness, ignore what is occurring. This is hardly an intelligent approach.

A second response is to turn inward, to develop protectionist measures, to keep other countries' industries at bay. This can have the devastating consequences of inducing retaliatory actions on the part of foreign governments, which in turn can cause severe damage to the U.S. economy. While we remain the world leader, the harm to the United States far exceeds that to other countries. Indeed, it can be

---

[66] "Economic Report of the President, 1980", 177.

[67] *Ibid.*, 316. There has been a negative merchandise trade balance every year since 1971, except 1973 and 1975.

[68] President Carter attributed the deficit "in part" to "a loss of American vitality." "Economic Report of the President, 1979", 13.

argued that Britain's turn to protectionism was the crowning symbol of its inability to continue its role as the major world power. Moreover, there is precedent for counterproductive protectionism in American economic history.

With the Smoot-Hawley Tariff of 1930, we raised duties high to protect threatened industries and to safeguard employment, but the effects on the domestic economy proved miserable. Other countries responded in kind. There followed a sharp decline in international trade and a rise in international loan defaults that contributed to the worsening of the Great Depression at home. Employment was not safeguarded. Employment dropped. We are now far more involved in the international economy than in 1930. In the 1980s U.S. protectionism is far more dangerous.

A third reaction is to protest foreign nations' barriers to trade and to counterbalance their subsidies. Protests on Japanese trade restrictions have resulted in substantial liberalization. By contrast, protests on O.P.E.C. nations' interventions in relation to oil prices have fallen on barren ground. Protests characterized the U.S. posture at Geneva during the Tokyo Round of the Multilateral Trade Negotiations. The approach appears to have had some success. The Tokyo Round agreements of 1979 do help stem the tide of protectionism. The Tokyo Round developed codes on government procurement that reduce national preferential treatment; on standards that accelerate the certification of foreign goods; on customs valuations that provide more consistency; on subsidies and countervailing duties that seek to restrict the former and clarify when the latter are appropriate; and on antidumping duties that offer specifications on their valid use.[69] The United States has used countervailing and antidumping duties; while these verge on protectionism, they can on occasion be justified as a reaction to market distortions.

They are consistent with U.S. antitrust policy that militates against predatory pricing. While the foreign company is out of the direct reach of U.S. policy, its pricing in the domestic market is within such reach. The difficulty lies in the calculation of the antidumping or countervailing duty. If foreign costs are lower than U.S. costs (and there is no subsidy), the American consumer ought to reap the benefits. Hopefully, the Tokyo Round agreements will help in clarifications on appropriate uses. By contrast, thus far we have not devised any effective counterbalance to the high oil prices!

A fourth approach of the United States is to recognize foreign industrialization, to be aware that foreign governments in developing countries, and in developed ones, are and will in the future encourage national industrialization and will act in what they perceive to be their domestic economic interests. If the United States is to cope with this reality, its regulatory environment must be designed to augment the U.S. competitive position—and not, I would suggest, in traditional, but instead in the newer industries. More attention needs to be paid to policies toward industry that raise U.S. productivity and encourage technological innovation.

In times past when skilled labor was expensive, American industry substituted capital for labor. Now, when because of foreign govern-

---

[69] "Economic Report of the President, 1980", 180.

ments' intervention, energy costs have soared, the need is to aid industries in innovating, so as to lower the cost of this input. The United States ought to seek to enhance the country's competitive position *in* high-technology, skill-intensive, education-intensive industries. In these, America retains a significant advantage. We ought to review our antitrust policy in an international rather than national context, being sure that this policy is being used to aid rather than to thwart American industry in world markets. The U.S. government should have a role in stimulating economic change, in strengthening American industry. Indeed, if this country is to maintain and to improve its competitive position, an extremely appropriate response to foreign governments' support of their industries is to consider U.S. regulatory policies in a global context, to accent through public policies the advantages that we have, to spur productivity increases, to encourage expenditures designed to reduce the cost of energy for U.S. industry, to develop policies to encourage investment in high-technology industries, and to do so in full recognition that such approaches serve the United States in its leadership role in the international community. The United States must have an energy policy that is reconciled to foreign national actions, that emphasizes domestic conservation, while diversifying supplies of energy, both geographically and technologically (government encouragement and sponsorship of new research in alternative sources of energy).

A fifth approach turns to international solutions. In July, 1978, in Bonn, Germany, the United States and the leaders of the six other largest non-Communist industrial nations met at a fourth annual summit meeting. The goal was to devise a "Concerted Action Program" to coordinate macroeconomic strategies. The United States recognized that its policies must not be divorced from those of other nations.[70] In June, 1979, the leaders of the same seven nations met in Tokyo, again pledging cooperation with one another. These summits are steps in international harmony. Such interactions need to be enlarged at all levels of government. The United States initiative in tariff cuts brought reductions by our major trading partners; at the Tokyo Round, we agreed to reduce tariffs by 30 percent, the European Economic Community by 27 percent, Japan by 22 percent.[71] Regrettably our trading partners did not follow our lead 100 percent, but the results are positive. We need to intensify international cooperation in the energy sphere. There is a mutuality of interest among consumer nations in reducing the costs of energy. The United States participates in the International Energy Agency; it seems, however, that I.E.A.'s solutions to energy problems are more defensive than creative (that is, the I.E.A. has plans to respond to crisis, but not to avert crisis).

The United States should take the lead in encouraging the I.E.A. to seek solutions that result in lower energy costs. Is there a way of satisfying the desires of oil producing nations for revenues needed for their development without the consequence wounds inflicted on the United States and indeed the world economy? In an article in Foreign

---

[70] "Economic Report of the President, 1979," 13–14, 141–143, and Whitman, "A Year of Travail," 537–539. Participants were the United States, Germany, Japan, France, Italy, the United Kingdom, and Canada.

[71] "Economic Report of the President, 1980," 180.

Affairs, Marina versus N. Whitman commented on the trade-off between U.S. "policies that would promote the achievement of specific national economic or political goals" and "those likely to promote the viability of a coherent international economic system. . . ." [72] Is there really a trade-off? If the international economic system fails to be "coherent," can specific domestic economic goals be met? I would suggest that there is now so much interdependence within the world economy that a coherent domestic policy on energy or inflation (or on any economic issue) can only be accomplished based on a profound awareness of the international economy—and particularly of the actions of other national governments. We must see to it that our policies and those of our trading partners have positive effects on one another, that is, are mutually reinforcing. America needs to be in the forefront in seeking out international solutions to inflation and the value of the dollar. Both are international problems. International cooperation—and more important international leadership—is the only way to avoid beggar-thy-neighbor policies that would hurt the economic health of the entire world community, our own economy included.

We must ask why the United States has been less than successful in asserting a leadership position in international economic affairs in recent years? What can be done to improve that role, if anything? [73] Can it be that since the Nixon shock of 1971 (that effectively ended the Bretton Woods system), since the oil crisis of 1973–74 that revealed our paralysis, and since the stagflation of the 1970s, America has been on the defensive? This is not the place to discuss the internal constraints on U.S. leadership, for instance, the relationship between the Executive and Congress, or the jurisdictional disputes between agencies responsible for domestic and foreign economic policies. Clearly, however, if America is to carry forth its leadership role, there is a need for thoughtful consideration of these issues.[74] Efforts should be made to enhance the U.S. leadership role.

A sixth approach, associated with, and inseparable from the last two involves the importance of U.S. awareness of how often foreign state intervention can have serious U.S. consequences. We need to identify areas of vulnerability and develop constructive strategies to cope with these before the negative effects are felt. It seems unlikely that any vulnerability will be as great as that in oil, but studies should be made in respect to other essential inputs. Our competitive position should not be held hostage to foreign state action.

In sum, a number of options are available to the United States as we feel the effects of foreign government interventions on our competitive position. To respond in an ostrich-like fashion by ignoring the impacts or to be complacent is totally inappropriate. Protectionism is counterproductive. To some extent, it is possible to protest others' barriers to trade and to a very limited degree it may be desirable to take counterbalancing measures against "unfair" competition. We have

[72] Whitman, "A Year of Travail," 528.

[73] See Letter from Raymond Vernon to Louis C. Krauthoff, Oct. 9, 1977, on the first draft of this paper.

[74] For a start see Richard F. Kaufman, "Reorganization of Foreign Economic Policy," U.S. Congress, Joint Economic Committee, "The U.S. Role in a Changing World Political Economy", 96th Congress, 1st session, 1979, 647–668. Kaufman argues for a need for a systematic review of how foreign economic policy is made and the possibilities for structural improvement.

56-368 O - 80 - 10

not, however, learned how to extend our antitrust policies to the monopolistic pricing of O.P.E.C.

The fourth, fifth, and sixth options outlined above seem more appropriate and viable paths. Creative, innovative policies to accent American industry's advantages are required. We need policies to encourage productivity growth. Such approaches must take into consideration the economic health of the world community. We must assert leadership in the world economy. We need to be aware, however, that interdependence does bring dependence, that is, vulnerability, and we must attempt to anticipate serious vulnerabilities.

The United States is still the world leader. Neither Germany with its strong mark, nor Japan with its strong yen and its dynamic economy is able to offer world leadershhip. No other country is prepared to do so. We must prove that national maturity is not followed by senility.

The response to interventions of foreign states that affect our competitive position should involve more, rather than less, international initiatives, involvement, coordination, cooperation, and awareness. We must recognize that global interdependence is, in fact, the dominant characteristic of America's changing economy.

THE REGULATORY BUDGET AS A MANAGEMENT TOOL FOR REFORMING REGULATION

By Christopher DeMuth, Richard H. Shackson, Eric O. Stork, and Arthur W. Wright*

CONTENTS

	Page
Summary	142
Introduction	142
I. The Rationale for a Regulatory Budget	143
A. Regulatory Budget Defined	143
B. Regulation as an Economic and Political Activity	144
C. The Case for Budget Control of Regulatory Compliance Costs	146
1. The Regulatory Budget as an Element of Regulatory Reform	146
2. The Need for More Effective Management of the Regulatory Process	147
3. A Regulatory Budget as a Management Tool	148
4. Coverage of a Regulatory Budget	149
a. Compliance Costs vs. Benefits and Indirect Costs	150
b. Agency Coverage	151
II. The Legal and Political Aspects of a Regulatory Budget	151
A. Administrative vs. Legislative Establishment of a Regulatory Budget	151
B. The Legal Aspects of a Regulatory Budget System	154
III. Organization and Management for a Regulatory Budget	159
A. Organization for a Regulatory Budget	159
B. The Regulatory Budget Process	160
1. Formulation of a Regulatory Budget	160
2. Execution of a Regulatory Budget	161
C. Enforcing a Regulatory Budget	162
1. Overstatement of Compliance-Cost Estimates During Budget Formulation	162
2. Understatement of Compliance-Cost Estimates During Promulgation	163
3. Controlling Overspending of Regulatory Budget Allocations	163
a. Bookkeeping	163
b. Sanctions	163
D. Congress and a Regulatory Budget	164
E. Public Participation in the Regulatory Budget Process	165
1. In the Formulation Process	165
2. In the Execution Process	166
F. Special Problems of Introducing a Regulatory Budget System	167
1. Phase-in	167
2. Initial Agency Coverage	167
3. New and Revised vs. Total Compliance Costs	168

*Christopher DeMuth is director of the Faculty Project on Regulation, Harvard University; Richard H. Shackson is director for transportation, Energy Productivity Center, Carnegie-Mellon Institute for Research; Eric O. Stork is visiting professor of technology and public policy, Purdue University; and Arthur W. Wright is associate professor of economics, Purdue University.

	Page
IV. Measuring Compliance Costs for a Regulatory Budget	168
A. Definition of Compliance Costs	169
B. Criteria for Evaluating Measures of Compliance Costs	169
1. Precision	170
2. Consistency	170
3. Transparency	170
4. The Cost of Making the Estimates Themselves	171
C. Compliance Costs of Existing vs. Proposed Regulations	171
D. Assigning Compliance Costs to Appropriate Fiscal Years	172
E. The State of the Art in Estimating Regulatory Compliance Costs	173

Summary

Regulation is one method by which the Federal Government claims resources to achieve its goals. The share of the Nation's resources claimed by regulation has grown rapidly in recent years.

At present, there are only weak constraints on the government's use of resources through regulation, and there is no procedure for incorporating the full cost of regulation into government decisionmaking. A system for budgeting regulatory compliance costs has significant potential as a management tool for controlling and shaping the economic impact of Federal regulation.

Although on strictly legal grounds it might be possibe to establish a regulatory budget system by administrative action, the political importance of such a system suggests that it should be established by legislation. There appear to be no constitutional barriers to including the so-called independent regulatory commissions in a regulatory budget system, along with the regulatory agencies in the executive branch.

The organization and management of a regulatory budget system could be similar to that currently utilized for the fiscal budget. Enforcement of regulatory budget ceilings would pose no unusual problems. For start-up it would appear expedient to focus on the compliance costs of new and revised regulations. Coverage of existing regulations could be added subsequently.

Without estimates of the cost of regulation there could be no regulatory budget. While existing methods of cost estimation need improvement, the very existence of a regulatory budget system would stimulate new or improved methods. Use of a methodology recently developed by Arthur Andersen and Co. could at acceptable cost provide the precise, consistent, and transparent estimates of compliance costs that would be required for a workable regulatory budget.

Introduction

The last several years have seen a marked increase in protests against regulation by the Federal Government. One reason for the protests is concern over the costs that regulation imposes on the private sector of the economy and on non-Federal governments. Those costs arise from efforts to comply with regulations, from lags or uncertainties in government procedures, and from distortions in the incentives of both regulators and regulatees.

In response to mounting political pressure, various attempts at reform have already been made. Thus far, however, those attempts have

been tentative and piecemeal. Calls for more comprehensive reform measures have come from both inside and outside the Federal Government.

One such reform measure is the regulatory budget. Secretary of Commerce Juanita Kreps suggested the idea in April 1978 in congressional testimony. Senator Lloyd M. Bentsen of Texas introduced a bill, S-3550, to establish a regulatory budget in the second session of the 95th Congress (1978), and a similar bill, S-31, in the first session of the 96th Congress. Professor Murray Weidenbaum urged inclusion of a regulatory budget in a comprehensive reform program outlined in the New York Times of December 17, 1978. In October 1978, interested parties from Congress, the administration, business, labor, universities, and public-interest groups discussed the regulatory budget as part of an all-day seminar on reforming regulation, sponsored by the Department of Commerce and chaired by Secretary Kreps.

The interest in a regulatory budget reflects the promise that it appears to hold as a tool of public management. With the promise, however, go a number of problems of design and execution. This paper explores both the promise and the problems of a regulatory budget.

The purpose of this paper is to analyze rather than to advocate; to explore rather than to conclude. The paper does not address the question of whether the Federal Government should adopt a regulatory budget. Rather, it considers how a regulatory budget could work, what the economic and other properties of a workable regulatory budget would be, and what difficulties would be encountered in trying to make the idea work.

I. The Rationale for a Regulatory Budget

This chapter details the promise that a regulatory budget holds for reforming Federal regulation. The first section defines a regulatory budget for purposes of this paper. The next section analyzes the current Federal regulatory process: how it operates, and why it has come in for so much valid criticism of late. Finally, the chapter suggests that a regulatory budget system would be an effective response to that criticism.

A. *Regulatory Budget Defined*

It is essential at the outset to define certain terms used in this paper:

> A *regulatory budget* would set limits for a given period on the *compliance costs* that the executive branch could impose, through *regulation*, on the private sector or on other governmental units.
>
> *Compliance costs* refer to the increase in outlays necessary to bring products or procedures into line with the requirements of Federal regulations.[1]
>
> The term *regulation* refers to executive actions, other than general taxes or subsidies, that are intended to alter specific private or non-Federal government decisions.[2]

[1] Examples of compliance costs include outlays for filing mandatory forms, hiring extra production workers to meet safe-manning rules, and adding new plant or equipment to comply with emissions standards.

[2] Examples include mandatory specifications for goods and services; ceiling or floor prices in particular markets; outright bans on specific goods or activities; and charges on efluents from production or consumption.

The budgeting of compliance costs is the meaning of a regulatory budget as it has been proposed in recent legislation and in most recent writings on reforming Federal regulation. But the term *regulatory budget* has also been used in a broader sense that would explicitly incorporate into the budget process the full social costs, or even both the costs and benefits, of regulation. The full social costs of regulation include, in addition to compliance costs, government administrative expense and indirect costs in the form of reductions in the value of social output. The benefits of regulation consist of increases in the value of social output and thus are generically similar to the indirect costs.[3]

The analysis of a regulatory budget in this paper is confined to the narrower definition, encompassing only compliance costs. The discussion below in Section C.4.a, p.10, explains why the other costs and the benefits of regulation are better excluded from a regulatory budget system.

B. Regulation as an Economic and Political Activity

Regulation is one method that the Federal Government uses to claim the economic resources that it devotes to its programs and operations. Other methods include taxation, the creation of new money, tax credits or deductions, and loan guarantees. The methods may differ mechanically (e.g., in whether the resources enter and leave the Treasury), but in all cases the Federal Government in effect acquires the means with which to pursue its goals. In so doing it alters the allocation of resources and the distribution of income. With some approximation, the government's claims to economic resources can be measured and stated in common (dollar) terms, regardless of which method is used.

The amount of resources claimed by the Federal Government in a given period is mainly the result of the political, not the market process. Public disputes over taxes, spending, the national debt, the money supply, so-called tax expenditures, and (recently) regulation reflect this fact. Because resources are scarce, at the heart of such disputes is a set of allocation questions: How many resources should the Federal Government claim? How should the total resources claimed be divided among the various agencies and programs? How should the government manage the resources it claims to use them most effectively?

Until well into the 20th century, the Federal fiscal budget was run almost informally (from an organizational standpoint) out of the Executive Office. The Bureau of the Budget and its successor, the Office of Management and Budget, are relatively recent developments, as are the now-elaborate procedures used to gather accurate, reliable fiscal data for use by the executive branch.

The present movement to reform Federal regulation may also be usefully viewed in historical perspective. Since the mid-1960's, Federal

---

[3] Expressing indirect costs and benefits in terms of the *value* of social output does not presume that either magnitude must be measured solely in monetary or other quantifiable form.

Government regulatory activity has expanded rapidly and on a wide front. Economists explain the burst of activity as the result of increased demands for Federal regulation by people or organizations who stand to benefit from it. Now, however, the costs of regulation have reached the point where those bearing them find it worthwhile to spend time and money opposing new regulations and lobbying for the repeal or revision of existing ones. In short, just as there are demands for regulation, so there are now demands for regulatory reform.

Current efforts to reform Federal regulation are not without precedent:

In 1971, the Office of Management and Budget established what became known as the Quality of Life Review. The purpose was to allow affected Federal agencies to comment on proposed regulations that were intended to enhance the quality of life. Although the review was supposed to apply broadly to all Federal agencies dealing with public health and safety, it was applied almost exclusively to regulations proposed by the U.S. Environmental Protection Agency.

President Ford instituted an Inflation Impact Statement program in November 1974. This program provided for an evaluation of the anticipated impact of all major new regulations upon prices, productivity, and competition.

The Ford program was supplanted in March 1978 by President Carter's more ambitious Improving Government Regulations program. Under that program, executive agencies were required to publish semiannual agendas of contemplated regulations, and to prepare "regulatory analyses" of all regulations having "an annual effect on the economy of $100 million or more." [4] Those analyses were to include "a succinct statement of the problem; a description of the major alternative ways of dealing with the problems that were considered by the agency; an analysis of the economic consequences of each of these alternatives and a detailed explanation of the reasons for choosing one alternative over the others." [5]

President Carter's program also included the establishment of the Regulatory Analysis Review Group (RARG), chaired by the Chairman of the Council of Economic Advisors (CEA), and including representatives of each of the principal economic and regulatory agencies of the executive branch. Analytical staff support to RARG is provided by the Council on Wage and Price Stability. RARG each year makes a detailed review of 10 to 20 regulatory analyses. Upon completion of such a review, the Chairman of CEA decides whether to file written comments on the regulatory proposal, meet with the head of the agency involved, or submit a report to the President.

In October 1978, President Carter directed the creation of a Regulatory Council,[6] composed of 35 departments and agencies, to help coordinate Federal regulatory activities. The Council is

---

[4] Executive Order 12044, Mar. 23, 1978.
[5] Ibid.
[6] "Memorandum for the Heads of Executive Departments and Agencies," The White House, Washington, Oct. 31, 1978.

required to publish a semi-annual calendar of proposed regulatory activities and—to the extent that they have been estimated by the initiating agency—their anticipated costs. The first such calendar was published February 28, 1979.

No doubt the past efforts at regulatory reform have had some influence. That influence has been largely due to better information generated by the new procedures or to the persuasiveness of particular individuals participating in those procedures. On the whole, however, the past efforts have not systematically brought economic constraints to bear on the regulatory process. For that reason, they have had at best a limited effect on inducing Federal decisionmakers to husband the economic resources claimed through regulation.

C. The Case for Budget Control of Regulatory Compliance Costs

A regulatory budget system would introduce economy into the Federal use of regulation to claim resources. In that context, it might be tempting to think of a regulatory budget simply as a way to reduce Federal regulation. The temptation, however, should be resisted. The true role of a regulatory budget would be as a tool for managing Federal regulation.

1. THE REGULATORY BUDGET AS AN ELEMENT OF REGULATORY REFORM

Regulatory reform is an imprecise term that means different things to different people. One dimension of reform concerns the personal stakes that individuals or groups have in the regulatory process. At one extreme of this dimension are the government officials who pledge to reform regulation by cleaning their own houses. At the other extreme are the business leaders who equate reform with the dismantling of the entire regulatory apparatus.

A different dimension of regulatory reform concerns the management of the process of promulgating Federal regulations. This dimension has to do with how the regulatory process operates rather than with who stands to benefit from it. Proposals for the operational reform of regulation are in the tradition of the fiscal reforms in the executive branch (1960's) and in the Congress (1970's).

The idea of a regulatory budget belongs on the management dimension of reforming Federal regulation:

> The regulatory budget system would be a management tool for use by politically responsible officials.
>
> Under such a system regulatory officials could be left free to make detailed decisions on how best to implement the broad provisions of laws passed by Congress; however, they would also be limited in the compliance costs they could thereby impose.
>
> A regulatory budget would provide no guarantee that the result would be fewer or less severe regulations. It should not be viewed as part of an anti-regulatory strategy. Rather, it would be a tool for better managing the regulatory process that has over time become so large a component of our national economic life.

2. THE NEED FOR MORE EFFECTIVE MANAGEMENT OF THE REGULATORY PROCESS

Proposals to reform the management of the Federal regulatory process presume a diagnosis that there is a problem needing correction. The diagnosis in this paper focuses on the incentives that the Federal Government faces in passing laws and in promulgating the regulations to put the laws into effect. The crux of the diagnosis is that, under the existing system, there are at best weak incentives to consider the full costs of Federal regulation. As a result, there are effective incentives both to overregulate and to choose particular forms of regulation that may be excessively costly.

As noted earlier, regulation claims scarce resources for government use. Sound management practice would require that decisions on regulation take into account all resources so claimed. This would help set a limit on the total amount of resources used and induce decisionmakers to allocate the total to the most effective uses. It would further cause regulatory officials to consider costs in selecting specific targets of regulation and discipline them to choose the least-cost ways of attaining given regulatory objectives.

Under the present arrangement, however, the Congress and the executive branch are held accountable for only a small fraction—the administrative costs, which are included in the fiscal budget—of the total resources claimed by Federal regulation. This provides little incentive for regulatory officials to consider compliance costs or possible reductions in social output when decisions involving regulation are made.

In effect, the present Federal regulatory process contains what in the study of market allocation is called an "externality": pertinent information is omitted from decisionmaking, with the result that the full cost of regulation exceeds the cost as perceived by decisionmakers. As a consequence, there is probably more regulation, and its composition is different, under the existing arrangement than if Federal decisionmakers were held accountable for all the resources claimed by regulation.

The incentives just outlined operate at two distinct levels of the Federal Government. At the policymaking level—that is, in the Congress and at the top echelons of the administration where legislative proposals are initiated—the benefits and costs of programs involving the use of regulation will be weighed with only a partial accounting of the full costs. It is even possible that there is a positive incentive to resort to regulation, in preference to other government methods of claiming resources that are subject to fuller accountability. For example, current efforts to compel a balanced fiscal budget might have little impact on the true economic scope of government, as opposed to the mere size of expenditures, unless attention is also paid to the compliance costs being imposed by regulation.[7]

The second level of the Federal Government at which the above incentives operate is policy execution. Typically, the laws as passed by Congress state policy goals only in broad, idealized terms. It is left

---

[7] For example, with some ingenuity, the government could probably establish a comprehensive national health insurance program entirely through regulation, with scant effect on the public budget.

to the writers of the regulations in the bureaucracy to provide the details of implementation. This means that regulatory officials have considerable latitude in choosing both the particular regulatory targets and the specific kinds of regulatory methods for assuring compliance. In the existing regulatory process, those officials have only weak incentives—through fiscal-budget control of administrative costs—to choose targets and methods that would achieve the congressionally mandated goals at minimum incremental cost. They have much greater incentive to select targets and methods that—regardless of the cost of complying with the resulting regulatory requirements—minimize the risk of failure to achieve the assigned social goal, and thus minimize the risk of personal criticism for having failed to achieve their goals.

A related point concerns the life histories of individual regulations pertaining to a given law. The broad, idealized tasks embodied in a law are inherently unattainable in practice. Thus, even an ambitious, mission-oriented set of initial regulations will not achieve all the possible objectives. As the initial program approaches success, however, the agency will turn its attention to other, as yet unmet objectives—of which there is an inexhaustible supply. Moreover, the agency has little if any incentive under the present system to retire existing regulations. The result is that the number and scope of regulations under a given law tend to grow steadily with time.[8]

3. A REGULATORY BUDGET AS A MANAGEMENT TOOL

The preceding diagnosis suggests that a serious defect of the present Federal regulatory process is that it produces excessive, and excessively costly, regulation. The source of the problem is the failure of the current process to take into account the full costs of regulation. There are a number of possible methods for restructuring regulatory incentives to make decisionmakers aware of the full costs and to force them to incorporate them into their decisions. The method examined in this paper is the familiar management tool of the budget.

One student of the Federal fiscal process characterized a budget as a "series of goals with price tags attached" and (because resources are limited) as a "mechanism for making choices among alternative expenditures." [9] A regulatory budget would put price tags (in the form of compliance costs) on the pursuit of Federal goals through regulation. It would also place limits on the total of compliance costs that may be imposed on the national economy, and on their allocation among individual agencies.

Under a regulatory budget system, the President and Congress would have to decide, explicitly and in advance, what the total Federal regulatory burden would be for a given period. They would also have to determine the relative importance of regulation in different areas to allocate the individual agency budgets. The Federal officials

[8] This analysis assumes normal human self-interest on the part of loyal government employees, and does not impute to them any venality or vindictiveness. Recent work by Niskanen, Tullock and others has shown that the bureaucratic counterpart of market competition will replace officials who fail to serve their own self-interests with ones who do.

[9] Aaron Wildavsky, "The Politics of the Budgetary Process," 2d editon (Boston: Little, Brown, 1974), p. 2.

who actually write regulations would be given an effective incentive to design new regulations so as to economize on the limited resources assigned to them. It would even be possible to encourage the timely removal of marginally effective or obsolete existing regulations, by providing agencies with regulatory budget credits for the resulting cost reductions.

The regulatory budget can thus be seen as a tool for establishing management control over the economic impact of regulation. *Control* is used here in its generic sense. The word, which is derived from the accounting profession, refers to a higher level of abstraction against which subordinate matters can be evaluated without having to examine them in detail.[10]

It is the lack of such a higher level of abstraction that has limited the effectiveness of previous efforts to control the economic impact of Federal regulation. Instead, those efforts have been based on two mistaken assumptions:

> That government officials outside an agency proposing a regulation can know enough details about the specific issue at stake to prevail in a debate with the far more knowledgeable proponent-agency's officials; and
>
> That the outside officials will be as determined and persistent as the proponent agency.

That is why the Quality of Life Reviews and the Inflation Impact Statements frequently amounted to little more than annoying ankle-pecking of the proponent agencies. In the end, the proponent agencies usually prevailed, even if after significant delays.

The regulatory budget would make it possible to do away with fruitless and enervating second-guessing of the judgments of politically responsible agency heads. As long as an agency remained within its budget allocation, higher levels of government would not have to worry about its regulatory requirements causing unacceptably large adverse economic impacts. Agency heads who failed to get the most out of their regulatory budget allocations (in terms of their assigned goals) would be disciplined through normal political channels: pressure from the interest groups that support the goals in question. Indeed, the tightened constraint of a regulatory budget to husband compliance costs would give such groups even more incentive to apply pressure than the weak constraint of current procedures.

4. COVERAGE OF A REGULATORY BUDGET

Two related aspects of what a regulatory budget would cover require attention. The first aspect is the inclusion of compliance costs and the exclusion of benefits and indirect costs. The second aspect is the range of Federal agencies whose regulations would be subject to budget limits.

---

[10] When banking first began, the proprietor of a counting house in the Italian city-state maintained a *contra rolus* against which the subordinate accounts (maintained by assistants who might not be trustworthy) had to balance. In that way the proprietor was able to tell—without having to review every detailed transaction—whether his assistants were stealing from him and to pinpoint areas that required his managerial attention. Over the centuries the word was anglicized to *counter roll,* and subsequently contracted into *countrol* and eventually *control.*

a. Compliance costs vs. benefits and indirect costs

As noted at the outset, the regulatory budget analyzed in this paper would cover only the direct costs of complying with Federal regulations. A possible objection to a regulatory budget so defined is that it would exclude two important economic effects of Federal regulation—namely, benefits and indirect costs (i.e., output losses).[11] This would appear to violate the goal for a regulatory budget suggested in Section C.2, above: to make the Federal Government more accountable for the overall economic consequences of its regulatory decisions.

There is some merit in this objection. The weighing of costs and benefits in government decisionmaking can certainly stand improvement. In spite of all the efforts that have been devoted to benefit-cost analysis and similar techniques, Federal decisionmaking remains highly imprecise and qualitative—one could even say impressionistic—when it comes to assessing the impacts, positive and negative, of Federal programs on the U.S. economy.

A regulatory budget system, however, would not be the right vehicle for attempting to introduce the needed improvement, except where compliance costs are concerned. The reason is not the desirability but the difficulty of estimating the benefits and indirect costs of Federal regulation on a sufficient scale and with enough reliability to be practical. The immense task of analysis, data collection, and computation would be prohibitively costly even if it were possible to reach consensus on the quantified results. Thus, under a regulatory budget, consensus on benefits and on indirect costs would have to be reached as it is now under the fiscal budget: implicitly through the political process.

The exclusion of benefits and indirect costs from the formal regulatory budget process would not prevent the useful application of benefit-cost analysis to individual problems. For instance, benefit-cost analysis could be used to decide whether regulation of a specific product or activity was worthwhile, or to choose between alternative forms of regulation. Neither would the exclusion mean that benefits and indirect costs could not be weighed in the political and legislative debates on particular regulatory programs. All it would mean is that the data used directly to control the economic impact of regulation would be limited to the direct costs of compliance.

A comparison with the fiscal budget is useful here. Benefits and indirect costs are not specifically included in the fiscal budget, either. But they are incorporated implicitly, most often in qualitative or conjectural form, in the political debates that precede decisions to raise or spend funds. While quantitative benefit-cost analyses are frequently conducted to support or oppose specific projects or programs, the only data that actually enter the fiscal budget are revenue and expenditure estimates.

---

[11] Administrative expenses would also be excluded. However, they are now covered by the fiscal budget, and there would be little point in transferring them to a regulatory budget. Moreover, administrative expenses account for such a minor fraction of the total costs of Federal regulation that it would scarcely be worth complicating the regulatory budget by including them.

b. Agency coverage

The question of which Federal agencies to cover in a regulatory budget could be answered in terms of the common distinction that is drawn between *social* and *economic regulation. Social regulation*, which covers such broad areas as health, safety, and welfare, is said to impose mainly compliance costs on the economy. In contrast, *economic* regulation, which applies to prices and quantities in specific markets, is said to impose mainly indirect costs on the economy. One could therefore argue that the budgeting of compliance costs should be confined to social regulation and not applied to economic regulation.

There is increasing evidence, however, that this common distinction between social and economic regulation is blurred. The indirect costs of social regulation—for example, in workplace and product safety, environmental quality, and drugs—are now viewed as substantial and growing. By the same token, the compliance costs associated with economic regulation—for example, in trucking, agriculture, crude oil, and natural gas—are widely recognized as imposing heavy burdens on business firms. It would be difficult, on grounds of compliance *vs.* indirect costs, to classify the auto fuel economy standards of the Department of Transportation as either solely social or solely economic.

A different criterion for settling the question of agency coverage is provided by the logic of a regulatory budget developed in this paper. Unless all agencies' regulations were included in the budget, Federal policymakers would have an incentive to evade budget limits by shifting programs to agencies left outside the system. The avoidance of opportunities to evade budget discipline would be central to the proper functioning of a regulatory budget. By this criterion, agency coverage should be total, not partial.

II. The Legal and Political Aspects of a Regulatory Budget

Two of the most basic practical problems in implementing a regulatory budget system concern politics and legality. The first problem has to do with whether a regulatory budget could, or should, be established administratively rather than through legislation. The second problem concerns the authority of the President to enforce the constraints of a regulatory budget system.

A. Administrative vs. Legislative Establishment of a Regulatory Budget

A question frequently asked by those in the Federal Government (especially in the executive branch) who are interested in the idea of a regulatory budget is whether it could be adopted administratively without new legislation.

There appear to be two principal reasons for the interest in an *administrative* regulatory budget:

First, as noted in chapter I, the budgeting of regulatory compliance costs is viewed as a logical next step in the evolution of the reform efforts that began with the Quality of Life Review under President Nixon and evolved into President Ford's Inflation Impact Statement program and President Carter's Improv-

ing Government Regulations program. All of those efforts were instituted by executive action without legislation.

Second, the legislative politics of enacting a regulatory budget into law appear daunting. The budgeting process would affect programs under the supervision of virtually every committee of the Congress. It could also affect the division of political authority between the executive and legislative branches.

On strictly legal grounds, a strong brief could be written for the position that the President has the constitutional authority to institute a regulatory budget by executive action. While the matter might ultimately have to be settled in the courts, the regulatory budget as outlined in chapter I would pertain to procedures and practices that are well within the scope of executing the laws passed by Congress—traditionally, the exclusive preserve of the executive branch of the Federal Government.

On closer scrutiny, however, an *administrative* regulatory budget would suffer from a defect so grave as to render it not worth the effort. It would ignore the crucial *political* function of the regulatory budget—that of compelling general agreement on an overall limit to Federal regulatory activities, quite apart from the merits or demerits of particular regulatory actions. Such a political function must perforce include Congress in a systematic manner. By extension, a workable regulatory budget system would have to be developed jointly by the administration and Congress, and then enacted into law. Thus, the legislative politics of a regulatory budget, however daunting, would have to be confronted.

The late 1970's have been a time of retrenchment in American politics, characterized by efforts to consolidate the activities of government after the boom years of the 1960s and early 1970s. During the past decade many of the traditional institutions and ideas that had limited the role of government in American life—for example, the congressional seniority system; the coalition of Republicans and southern Democrats; and the dominance of political parties in selecting and promoting public officials—were greatly weakened or collapsed altogether. There is today little or no disposition either to revive the discarded political institutions or to repeal the legislation that followed their demise. Nevertheless, the current era is one of searching for new—and more formal—institutions of political discipline.

The most conspicuous and controversial attempts to establish new forms of political discipline have been directed at fiscal limitation. An early instance, in 1973–74, was President Nixon's policy of selective impoundment of congressional appropriations. The policy provoked furious opposition in the Congress and elsewhere, and it fared poorly in the courts. Even so, it did spur Congress to take major steps, such as the establishment of budget committees and a Congressional Budget Office to exert more control over legislative appropriations.

More recently, there have been numerous attempts to place explicit institutional restrictions on the size and scope of government. A number of referendums and proposed constitutional amendments to limit state taxing or spending has passed, such as California's Proposition 13. Currently there are efforts to amend the U.S. Constitution to tie Federal expenditures to economic growth, or to require a balanced fiscal budget.

The regulatory reform movement is perhaps less controversial than the fiscal limitation movement, but it is motivated by the same quest for new forms of governmental restraint. Efforts to reform the Federal regulatory process reflect a desire to consolidate and better manage the enormous growth of regulation since the late 1960's. There is also a concern to reduce the impact of regulation on the U.S. economy.

Thus, regulatory reform is the policy complement of fiscal limitation. Regulation differs from fiscal action in that it promotes policy objectives not by the spending of public funds, but rather by causing private funds to be spent differently than they otherwise would have been. As the Federal Government's direct administrative expense on regulation is relatively small, and as at present the other costs of regulation are not accounted for in government decisionmaking, the Federal Government has a built-in incentive to increase its reliance on regulation for the pursuit of social goals. Success in imposing fiscal limitation (such as the proposed constitutional amendments currently being debated) would serve to sharpen that incentive. Thus, the efforts to tighten fiscal discipline in the absence of a corresponding effort to tighten regulatory discipline could give the paradoxical result of *reducing* rather than increasing the political accountability of government.

One need not, however, favor tighter fiscal discipline to favor regulatory reform in general, or a regulatory budget in particular. A regulatory budget would be a counterpart of the fiscal budget. Both are means of working toward agreement on the overall size of the public (Federal) sector, and on the allocation of expenditures to particular uses or programs.

As noted earlier, a system of budgeting regulatory compliance costs is in one sense an outgrowth of the current regulatory review program in the executive branch. It would, however, have fundamentally different purposes: (a) accounting for publicly mandated expenditures resulting from Federal regulations; (b) requiring agreement on an overall ceiling (more or less flexible, as in the case of the fiscal budget) on such expenditures; and (c) allocating regulatory expenditures among programs in accordance with prevailing views about the relative social benefits of the programs.

It is important to recognize that the second and third functions of a regulatory budget would be a supplement to—not a substitute for—the benefit-cost analyses of individual regulatory decisions that the current regulatory review program calls for. Regulatory budgeting would implicitly acknowledge the impossibility of measuring precisely the benefits of Federal regulatory programs, and would thus leave the decision on relative benefits to be made in an explicitly political way, through the allocation of the agreed-upon total regulatory expenditures.

The balancing of competing social demands in that manner is not simply an executive management function. Neither is it just a matter of deciding whether a particular regulatory proposal is necessary, unnecessary, or excessive under a particular statutory directive. Rather, it is a matter of deciding how much of the Nation's resources to devote, in the aggregate, to the pursuit of all of the legislatively mandated social goals, as well as how much of the total to devote to

each of the particular goals set forth in legislation (e.g., environmental quality or occupational safety).

The closest functional analogy to a regulatory budget is not the current regulatory review program, but rather the fiscal revenue and appropriations process. Thus, Congress must be involved in regulatory budgeting not because of any particular line of legal precedent, but because it is a *policymaking* process which under our Constitution is performed jointly by the executive and legislative branches. This is the basic reason why it is pointless to debate whether the President could legally impose a regulatory budget without congressional authorization.

B. The Legal Aspects of a Regulatory Budget System

An effective regulatory budget system—one that would motivate regulatory officials to set clear priorities and to choose cost-effective measures in particular cases—would have to be enforceable. As a practical matter this means that the President would have to have unambiguous executive authority to dismiss regulatory officials who failed to live within their regulatory budgets or who otherwise refused to cooperate in the budgeting process. For the program to be complete, the President's authority would have to extend to the independent regulatory commissions as well as to the executive branch agencies.

For political if not for strictly legal reasons, the President probably could not unilaterally assert such authority over either type of regulatory agency. The President plainly could, however, as a constitutional matter, exercise such authority according to statutory mandate, and he could do so in a way that would not compromise any independence that the Congress might wish to maintain in the independent agencies.

To illustrate, suppose that a President were to go beyond the occasional reconciliation of regulatory disputes within the Executive Branch (as in President Carter's action in the cotton dust dispute), and embark upon a systematic policy of ordering substantial reductions (or increases!) in proposed regulatory actions. Immediate congressional complaints and court challenges would be a certainty. A legal and political precedent would be President Nixon's executive fiscal impoundment program, mentioned earlier. The courts, in cases such as *Train, Administrator, EPA* v. *City of New York*, 420 U.S. 35 (1974), *State Highway Commissioner of Missouri* v. *Volpe*, 479 F.2d 1099 (8th Cir. 1973), *American Federation of Government Employees* v. *Phillips*, 358 F. Supp. 60 (1973), and *Gnadamuz* v. *Ash*, 368 F. Supp. 60 (1973), held that the President was without authority to impound funds appropriated by Congress to be spent for particular purposes. This was so even if in the President's judgment a lower level of spending was required for purposes transcending those of the statutes involved (such as reducing inflation).

It could be argued that those cases should be distinguished from *regulatory* impoundment. Not only did they involve appropriated Federal funds rather than revision of executive branch decisions, but also the impounded funds were grants to specific private groups rather than funds for general management and enforcement activities. Congress' statutory reaction to the fiscal impoundment challenge, however, sug-

gests that it would react similarly to any systematic Presidential intervention in specific regulatory decisions.[12]

It would be possible to resolve this issue through a statute explicitly authorizing the President or his subordinates to take part, on a case-by-case basis, in final decisions on formal rulemaking or informal regulatory proposals. Indeed, the *Exposure Draft* of the American Bar Association's current study of the regulatory process proposes "enactment of a statute authorizing the President to direct certain regulatory agencies to take up, decide, or reconsider, critical regulatory issues within a specified period of time, and thereafter to modify or reverse certain agency actions relating to such issues." [13] It is doubtful, however, that regular participation in detailed regulatory decisions (as envisioned by the ABA *Exposure Draft*) would be a useful expenditure to the President's time, or that any President would wish to have such formal authority.[14]

The legal precedent is quite thin regarding the President's authority to govern the activities of independent and executive branch regulatory officials, either directly as proposed by the American Bar Association or indirectly through a regulatory budget process. The two most important decisions, both of which concern the President's authority to dismiss Federal officials, are *Myers* v. *United States*, 272 U.S. 52 (1926), and *Humphrey's Executor* v. *United States*, 295 U.S. 602 (1935).

In *Myers*, the President unilaterally removed a postmaster before the expiration of his term, although the postal statute stipulated that removals required the advice and consent of the Senate. The postmaster's administratrix sued to collect her husband's salary for the balance of his term, arguing that the President had exceeded his executive powers under the statute. She lost, the Court holding that the statute itself violated the President's constitutional authority as chief executive officer. The court noted that (272 U.S. at 135):

> The ordinary duties of officers prescribed by statute come under the general administrative control of the President by virtue of the general grant to him of the executive power *and he may properly supervise and guide their construction of the statutes under which they act* in order to secure that unitary and uniform execution of the laws which Article 2 of the Constitution evidently contemplates in vesting general executive power in the President alone. (Italics supplied.)

The executive power (including the power to remove as well as to supervise and guide) being a constitutional one, it could not be compromised by congressional action. The court noted, however, that the President's authority over Federal officials was not unlimited (*id.*):

> Of course, there may be duties so peculiarly and specifically committed to the discretion of a particular officer as to raise a question whether the President may overrule or revise the officer's interpretation of his statutory duty in a particular instance. Then there may be duties of a quasi-judicial character imposed on executive officers and members of executive tribunals whose decisions

---

[12] The Congressional Budget and Impoundment Control Act of 1974 required the President to follow specific procedures whenever impounding funds and provided for a swift legislative veto.

[13] American Bar Association, "Federal Regulation: Roads to Reform," 1978, p. 101. The final report will be issued in the summer of 1979.

[14] President Carter recently said in defense of his regulation-review program, "I have not interfered in [the regulatory] process. I have a statutory responsibility and right to do so, but I think it would be a very rare occasion whenever I would want to do so." (Press Conference, Feb 27, 1979.)

after hearing affect interests of individuals, the discharge of which the President cannot in a particular case properly influence or control.

The *Humphrey's Executor* case was similar to *Myers* in almost every respect except that the official involved was a member of the Federal Trade Commission rather than a postmaster. The Commissioner, a Hoover appointee with several years left in his statutory term, was removed by President Roosevelt on grounds of political incompatibility. His executor later sued to collect the Commissioner's pay for the remainder of his term. The executor won, on grounds that the FTC being "predominantly quasi-judicial and quasi-legislative . . . occupies no place in the executive department and . . . exercises no part of the executive powers vested by the Constitution in the President" (295 U.S. at 624). The President's executive authority under the Constitution did not extend to officers whose functions were more judicial or legislative than executive. As to such officers, the Court held that "no removal can be made during the prescribed term for which the officer is appointed, except for one or more of the causes named in the applicable statute." (The FTC statute provided for removal by the President only in cases of "inefficiency, neglect of duty or malfeasance.")

Taken together, the *Myers* and *Humphrey's Executor* cases establish the following distinction: On the one hand, the President may unilaterally remove appointees whose functions are purely executive—who serve as an arm of the President in his role of chief executive officer. On the other hand, the President may not remove appointees whose functions are essentially legislative or judicial—who serve to render their judgments based on the merits of particular claims. Under those cases, the President's exclusive power to remove executive officers is grounded in the Constitution, so that Congress may not compromise it by contrary statutory provisions governing removal. It is important to note, however, that the President's lack of power to remove quasi-judicial officers is *not* a constitutional requirement. Rather, it is based upon (a) judicial assessment of legislative intent regarding removal of quasi-judicial officers, combined with (b) judicial opinion that the President does not have the power, on constitutional grounds, to remove Federal officials whose functions are non-executive.

Legislative intent may be clear when Congress provides by statute that officials may be removed by the President only for "inefficiency, neglect of duty or malfeasance," as it has in the case of most of the independent regulatory commissions. But where the Congress fails to specify the nature of the removal authority, the courts will decide the matter according to the statutory function of the official involved. In *Weiner* v. *United States*, 357 U.S. 349 (1959), the Supreme Court held that the President lacked the power to remove a hold-over appointee whose function was quasi-judicial (he was a member of the War Claims Commission), even though the law in question (the War Claims Act) did not restrict the removal power in any way.

Thus, the issue of the President's ability to enforce a regulatory budget cuts across the regulatory programs of both the independent commissions and the executive branch agencies. As the *Myers* decision noted, even within the executive branch the President's constitutional authority to "supervise and guide" Federal officials is limited. It may not include the authority to (a) "overrule or revise"

interpretations of statutory duties "peculiarly and specifically committed to the discretion of the particular officer," or (b) "influence or control" duties of a "quasi-judicial character"—those involving decisions after agency hearings affecting interests of individuals. Again, however, there is no *constitutional* barrier to the broader Presidential authority if statutes so provide; for example, the President is authorized by statute to intervene in the award of international airline routes.

While a regulatory budget system would certainly affect agencies both in establishing priorities and in deciding particular cases, it would do so largely indirectly through the allocation of budget ceilings, rather than through direct Presidential involvement in particular cases. As a legal matter, the regulatory budget would be preferable to the case-by-case Presidential involvement in the regulatory process as proposed by the American Bar Association. A statute authorizing the President to intervene in particular regulatory decisions would raise acute problems of judicial review that would be difficult to resolve in advance by statute. In every case where a proposed regulation was modified by the President and then, after final publication, was challenged in court, the court would have to decide not only whether the responsible official had correctly interpreted the requirements of the statute in the first instance, but also whether the President himself had correctly interpreted the requirements of both the statute and the regulatory-review statute.

A regulatory budget would avoid these problems while achieving the same goal of placing overall discipline on the regulatory process. The discretion of regulatory officials in interpreting their statutory duties would not be constrained in particular cases (at least not by the budget process itself), but such officials would be obliged to live within a compliance-cost budget formulated by the President and Congress, just as they must at present remain within their appropriations. The President would not be pressured more, and could well be pressured less, than he is at present to take action in particular regulatory controversies. Challenges to regulations would not be complicated by judicial review of Presidential decisions that weighed the requirements of a regulatory statute in a particular case against, say, inflation or other countervailing considerations permitted by a regulatory-review statute.

As under any management control system, difficult decisions would have to be made under a regulatory budget system. Consider the case of a regulatory agency charged with enforcing a statute that gave it little or no discretion in a specified factual situation (for example, the Delaney Clause requiring FDA prohibition of food additives found to cause cancer). The agency might be forced to take action that would cause it to exceed its regulatory budget for a given period. Cases such as this, however, could be handled by supplemental authorizations of authority to impose compliance costs, analogous to the supplemental appropriations used in the fiscal budget. It does not appear that adding an economizing restraint to the regulatory process would have legal implications different from those of the current fiscal budget process.

The *Myers* and *Humphrey's Executor* decisions suggest that the independence of the so-called independent regulatory commissions is a matter of congressional determination rather than constitutional requirement. Presumably this extends not only to the President's authority to remove officials, but also to his authority to oversee their conduct while they remain in office. If Congress may constitutionally permit the President wide discretion in removing members of the regulatory commissions, a *fortiori* it may permit him to "supervise and guide their construction of the statutes under which they act" just as closely as is his prerogative, according to *Myers*, in the case of executive officials.

Indeed, there is good reason to doubt that *Humphrey's Executor*, standing alone, obliges the President to show as much deference to the independent commissions as is common (viz., President Carter's exclusion of the commissions from his "Improving Government Regulations" program). *Humphrey's Executor* was concerned with backpay. It did not present the Supreme Court with the more difficult practical issue of whether a commission member could continue to exercise authority against the opposition of the President—much less the issue of whether the President may direct the commissions' general management and procedures apart from deciding particular cases. It should be noted in this connection that the *Humphrey's Executor* decision was one of statutory interpretation, and that the statutes under which the independent commissions operate nowhere prohibit the President from influencing their general policies.

The *Exposure Draft* of the American Bar Association's report on regulatory reform has this to say concerning the legal applicability of executive branch regulatory review to the operations of the independent commissions (p. 108):

> As originally proposed the Carter Order (Executive Order 12044) would have imposed its discipline on independent agencies as well as executive branch agencies. The final Order, however, leaves the "independent" agencies untouched. A majority of the Commission regrets this omission. In the Commission's views, the President has constitutional power, in the present absence of any statute to the contrary, to prescribe housekeeping or procedural requirements for an independent Federal agency that leave intact the policy-making and adjudicatory authority of the agency. Should the contrary point of view on this issue prevail, the Commission supports enactment of a statute expressly authorizing the President to impose such disciplines on the independent agencies.

Nevertheless, the legal necessity of the President's deference toward the independent commissions is certainly less important than its roots in political custom. President Carter, it should be noted, excluded the independent commissions from his regulation-review program not on the advice of lawyers, but after being strongly admonished to do so by numerous Members of Congress. It seems safe to assume that, if the President attempted on his own to include the independent commissions in a regulatory budget system, the result would be an intense political controversy and prompt court challenge that would seriously compromise the success of the system from the start. The important point, then, is not only that a management tool such as a regulatory budget *may* be extended to the independent commissions by statute, but that doing so would put an end to the current bifurcated nature of regulation review under President Carter's administrative program.

In summary, if Congress elected to establish a regulatory budget system, it could give the President the legal authority to require full adherence to the constraints of such a system. In contrast, a Presidential attempt to impose a regulatory budget without a legislative mandate would encounter enormous political obstacles and strenuous legal challenges.

III. Organization and Management for a Regulatory Budget

The next practical problem in implementing a regulatory budget system is how such a system would actually operate. This chapter examines the organizations that would be involved in operating a regulatory budget, and how a budget could be workably formulated and executed. The chapter concludes with a discussion of the special problem of initially setting up a regulatory budget system.

A. Organization for a Regulatory Budget

A central administrative body would be required to manage the overall regulatory budget system. For convenience, that body is referred to in this paper as the Office of Regulatory Budget (ORB).

The role of ORB would be broadly analogous to that of the Office of Management and Budget (OMB) in the fiscal budget process. It would be the responsibility of ORB to develop and administer the detailed procedures needed to operate a regulatory budget, and to manage the formulation and execution of the budget. While the functions of ORB could be performed by some other entity of the Executive Office of comparable rank with OMB, the similarity in responsibilities would argue for establishing ORB as a part of OMB.

A key decision would be whether to merge the regulatory budget operations into the existing fiscal estimates groups of OMB. The fiscal estimates staffs possess a wealth of knowledge about the regulatory reponsibilities of the agencies that they supervise, and are experienced in bringing to bear on the activities of those agencies the point of view of the Executive Office. Hence the fiscal estimates groups could significantly enhance the effectiveness of a regulatory budget process, especially at the outset.

Opposing considerations, however, might make it preferable to establish ORB as a separate element within OMB. Dispersing regulatory budget operations among the various fiscal estimates groups would relegate ORB to the role of a central staff providing mainly procedural guidance. Especially in the initial stages of a regulatory budget, such a role could impede the prompt revision of procedures as experience was gained, and could hamper the dissemination of the revised procedures to the regulatory agencies. Also, dispersion of budget operations would also make it more difficult to achieve consistency of treatment among regulatory agencies.

Furthermore, the fiscal estimates groups already have a large and demanding workload that is tied to the inexorable demands of the fiscal budget calendar. Thus, there would be a risk that a new function such as regulatory budgeting would not be able to compete effectively for managerial and staff attention if it were merged into the fiscal estimates divisions. Finally, regulatory agencies would not neces-

sarily deal with a regulatory budget process through their fiscal budget offices, with which OMB's fiscal estimates groups customarily deal. If so, the existing lines of communication between OMB's fiscal estimates groups and their client agencies would not be the same as the lines of communication needed for the regulatory budgeting process.

On balance, it would appear to be preferable to establish ORB as a separate element of OMB, with responsibility and authority to deal directly with the agencies in the formulation and execution of the regulatory budget. When the regulatory budget system had been operating long enough to have settled down, consideration could be given to integrating the operation of the system into the fiscal estimates groups, retaining for ORB the role of specialized staff to deal with across-the-board procedures and with the aggregation of annual agency authorizations.

B. The Regulatory Budget Process

For ease of exposition, the discussion of the regulatory budget process follows the budget cycle for a single period, from formulation through execution. In practice, of course, the cycles for successive periods would overlap just as they do for the fiscal budget. In broad outline, the regulatory budget process described here parallels the existing Federal fiscal budget process. The regulatory budget period is assumed to be the traditional Federal fiscal year,[15] which begins on October 1 of the preceding calendar year and ends on September 30 of the current calendar year.

1. FORMULATION OF A REGULATORY BUDGET

ORB would initiate the regulatory budget cycle about 21 months before the fiscal year for that budget began. The cycle would start with the issuance to regulatory agencies of detailed procedural guidelines for submitting their requests for compliance-cost budgetary authority. Those guidelines would include tentative ceilings for total agency requests.[16]

The responses of the agencies would be due at ORB about 14 months before the applicable fiscal year began. The agencies' requests would describe the proposed new regulations and the estimated costs of complying with *all* the regulations that would be in effect during that fiscal year.[17] The expected benefits of the agencies' regulations would also be included in their submissions to ORB, in support of their requests for budget authorizations. As discussed in chapter 1, however, the benefits of regulations would *not* enter explicitly into the regulatory budget process.

ORB would review the agencies' requests and recommend modifications to conform with government policy on the total amount and

---

[15] Consideration was given to possible alternative periods for regulatory budgeting. However, no persuasive arguments for a different period were found, and the potential for confusion from non-aligned fiscal and regulatory budgetary periods would be great. One possible problem—that a regulation would not be ready for promulgation in the period planned—could be handled by making regulatory budget authority the equivalent of so-called no-year fiscal appropriations that can be carried forward until used.

[16] For the first several regulatory budget cycles, tentative ceilings might not be possible because of a lack of data on compliance costs.

[17] As discussed below, the compliance-cost estimate for all regulations could not be required initially. Such estimates could be added to the system only after baseline data had been developed.

the composition of regulatory activity. The modifications would almost certainly be downward, as most agencies would submit requests for above-ceiling compliance-cost authorizations. About 11 months before the beginning of the applicable fiscal year, ORB would submit to the President its recommendations on both aggregated, government-wide budget totals, and on the agency-by-agency dollar amounts of authorized compliance costs.

The President, aided by his staff, would review ORB's recommendations prior to making the final decision on his proposal to Congress for the next year's regulatory budget. As part of the review, individual agencies could appeal ORB's modifications of their requests for budget authorization to the President himself, as happens occasionally in the fiscal budget process.

The results of the President's decision would be communicated by ORB to the agencies, which would then prepare detailed submissions to send to Congress. Early in the new congressional session, the President's regulatory budget for compliance-cost authorizations for the next fiscal year, both total and agency-by-agency, would be forwarded to Congress. Shortly thereafter, the individual agencies would submit their detailed budget requests to Congress.

The formulation of a regulatory budget would end with the enactment into law of the President's proposals as modified by Congress. More detailed discussion of congressional involvement in the regulatory budget process is presented in section D below.

2. EXECUTION OF A REGULATORY BUDGET

Once the fiscal year had begun, the regulatory activities of Federal agencies would be constrained by the compliance-cost ceilings set in the budget for that year. The constraints would need to be sufficiently flexible to permit the agencies to adapt to changing circumstances in a timely manner. Thus, a regulatory agency would not be limited to the promulgation of only those regulations that were specifically included in its original request to ORB, nor would the agency be required to issue every regulation it had included in its request.[18] Also, if an agency revoked some of its existing regulatory requirements, it could be allowed to increase its regulatory budget authorization by the amount of compliance costs thereby saved.

The procedures for proposing and promulgating new or revised regulations under a regulatory budget (or for revoking current regulatory requirements) would resemble the existing procedures, but would expand on them in significant ways:

At the time of proposal, an agency would publish a draft economic impact statement, much as it is required to do now. The draft statement would include a preliminary estimate of the anticipated cost of complying with the new or revised regulations. The comments received on the proposal would aid in preparing the subsequent comprehensive estimate of compliance costs.

When the agency formally promulgated the new or revised regulation, it would be required to publish its final economic impact

[18] The agency would, however, have to achieve a reasonable correlation between the regulations specified in its submission to ORB and those in fact acted upon, if it wished to retain credibility for future budget cycles.

analysis. That analysis would contain a comprehensive estimate of the additional compliance costs that would result from the regulations.

The final economic impact analysis would be open to public comments for a period of 90 days. The comments would be submitted to ORB, with a copy to the promulgating agency.[19]

After the comment period ended, ORB would resolve differences between the agency's estimated compliance costs and those in the comments received. ORB would reach a final decision on the compliance cost figures to be charged to the agency's regulatory budget authorization within 90 days of the close of the comment period (six months after promulgation of the regulations). That decision would be the official estimate of compliance costs and would be so certified by ORB.[20]

The certified estimate would be charged against the promulgating agency's regulatory budget authorization.

C. Enforcing a Regulatory Budget

The existence of a regulatory budget would impose a new constraint on Federal agencies that would not be welcomed by regulatory officials. Efforts to avoid the constraint should therefore be anticipated.

Three principal problems in enforcing a regulatory budget merit attention: (1) an incentive to overstate compliance-cost estimates during budget formulation; (2) an incentive to understate those estimates in promulgating regulations;[21] and (3) preventing the overspending of regulatory budget allocations. These three problems are discussed in order.

1. OVERSTATEMENT OF COMPLIANCE-COST ESTIMATES DURING BUDGET FORMULATION

In preparing their requests to ORB in the early part of the budget cycle, Federal agencies would have an incentive to overstate their estimates of the compliance costs of new or revised regulations to secure the largest possible budget authorizations.

ORB's ability to identify overstatements at this stage of the regulatory budget process would be limited because even the agencies would rarely know so far in advance how the subjects of a proposed new requirement would actually respond to it. Only after the agencies had further developed their regulations and received comments on them would they possess the data to support more reliable estimates.

Overstatement of estimated compliance costs during budget formulation would not, however, be a fatal defect in a regulatory budget system. ORB's recommendation to the President for allocations of budget authority would depend only in part on preliminary estimates of compliance costs by the agencies. Also important would be high-

---

[19] Some potential inequities in the public-comment process are discussed below in section E.

[20] Where ORB would acquire the expertise for reviewing comments and revising compliance cost estimates—from its own staff or from outside consultants—is an empirical question that cannot be answered definitely in this paper. It is likely that both sources would be used, with the mix varying from case to case.

[21] An agency's incentives for estimating compliance costs for regulations to be removed would be just the opposite of the incentives for new or revised regulations.

level policy guidance given to ORB examiners on how much in additional national resources could be devoted to the various social goals represented by the regulatory agencies' missions. In addition, the agencies themselves would not want to make outrageous initial requests, lest they weaken their subsequent credibility with ORB.

2. UNDERSTATEMENT OF COMPLIANCE-COST ESTIMATES DURING PROMULGATION

In promulgating regulations, it would be in an agency's interest to understate estimated compliance costs to minimize the share of the available authorization devoted to any one regulation. During promulgation, however, ORB would be in a much better position than during budget formulation to evaluate the agencies' estimates. For one thing, ORB would have available the supporting data in the detailed economic impact analyses, as well as the data submitted as part of the public comments.

Secondly, agencies would have an offsetting incentive not to understate their compliance-cost estimates. By systematically understating the costs of all new regulations for a given period, the agency would run the risk that ORB would systematically certify estimates greater than the agency's figures. This would increase the agency's chances of exceeding the regulatory budget authorization for that period. The resulting deficit would prompt sanctions (as discussed below), or would at least be deducted from the agency's next fiscal-year budget authorization, which would already have been established. An agency would presumably have difficulty justifying supplemental allocations to cover deficits caused by its own poor compliance-cost estimates.

3. CONTROLLING OVERSPENDING OF REGULATORY BUDGET ALLOCATIONS

Enforcing the budget ceilings of regulatory agencies would require both appropriate bookkeeping procedures and sanctions that could be imposed when a budget ceiling was exceeded.

a. Bookkeeping

The bookkeeping procedures for a regulatory budget would be conceptually similar in form to those for the fiscal budget, but would be simpler and less costly to carry out:

Authorized expenditure limits and (certified) conpliance costs saved by retiring old regulations would be entered as credits on the agencies' ledgers. Compliance costs spent on new or revised regulations would be entered as debits.

The scale of bookkeeping needed for a regulatory budget would be far smaller than for the fiscal budget. The huge number of individual fiscal transactions each year would dwarf the sum total of the separate regulatory actions taken each year by Federal agencies.

b. Sanctions

It would not be possible under a regulatory budget to impose sanctions for overspending parallel to those used under the fiscal budget.

Thus, new kinds of sanctions would have to be devised to make a regulatory budget work.

The authority to commit appropriated fiscal funds is extensively delegated to relatively low administrative levels. That is necessary because of the vast number of individual transactions involved. The traditional enforcement tool is to hold individual certifying officers financially liable for funds committed in excess of their allocations. The threat of personal financial liability has proven widely effective in achieving observance of fiscal budget limits.[22] In addition, modern management information systems have made possible secondary cross-checks of whether fiscal obligations exceed authorizations.

It has been suggested that a serious defect in a regulatory budget system would arise from the inability to hold anyone financially liable for exceeding authorized compliance cost allocations. The implication is that a regulatory budget could not be effective because it would deal only with funny-money, not the real money with which the fiscal budget deals.[23]

In fact, the so-called funny-money problem is more apparent than real. The heart of the issue of controlling the overspending of any budget authority is the level at which the discipline takes place. Under the fiscal budget, as already noted, obligating authority is delegated to low administrative levels. In contrast, the regulatory budget equivalent of fiscal obligating authority would not be nearly so widely dispersed. The reason is that most regulations are promulgated over the names and by the authority of Presidentially appointed heads of Federal agencies.[24]

Thus, enforcing regulatory budget ceilings would be primarily a political, not an administrative, problem. Without exception, the officials who would authorize regulatory-budget expenditures—by promulgating regulations—would be at politically responsible levels. Regulatory budget discipline would have to be imposed through the political process.

In starkest terms, to prevent the overspending of regulatory-budget allocations, the President would have to stand ready to dismiss an agency head who failed to stay within his or her budget authorization. Since most officials who promulgate regulations serve at the pleasure of the President, regulatory-budget discipline should pose no problem for a President who was determined to obtain it.[25]

D. Congress and a Regulatory Budget

A regulatory budget would be a major change in Federal policy toward regulation. For that reason, it was argued in chapter 2 that the President should not attempt to establish a regulatory budget sys-

[22] The discipline of personal financial liability can, of course, break down if it loses credibility—e.g., if the sums get larger than an individual could possibly pay. A rule of thumb among bureaucrats is to overobligate big if one is going to overobligate at all.

[23] It is possible to devise schemes that would use personal financial liability to control the overspending of regulatory budget allocations. However, such schemes appear to hold little promise of being either operational or effective.

[24] In cases in which promulgating authority is delegated, invariably it is only to a few high-level officials.

[25] It is frequently commented that imposing budget descipline on the independent regulatory commissions (such as the Federal Trade Commission or the Consumer Product Safety Commission) would pose problems for the President. As noted in chapter 2, however, such need not be the case.

tem unilaterally by executive order, even though constitutionally it probably could be done. Rather, such a system should be set up under the authority of a law passed by Congress.[26]

Beyond merely authorizing a regulatory budget, Congress would need to participate regularly in the budgeting process itself. This would require that Congress be kept informed of three key decisions regarding the regulatory budget: (a) the aggregate compliance-cost authority proposed for the budget year; (b) the proposed budget allocations to individual agencies (which the Congress would review, modify, and approve); and (c) proposed increases during the current year in specific agencies' budget allotments.

How Congress organized itself to deal with the President's regulatory budget proposals would be vitally important. The primary purpose of the regulatory budget system would be to set overall limits on the compliance costs resulting from regulatory requirements. The purpose would *not* be to improve decisionmaking on individual regulations. To accomplish the primary purpose, Congress in its action on the regulatory budget would need to take an overview of the broad economic impact of regulation and avoid being distracted by individual regulatory requirements.

It would be important, therefore, that Congress deal with the President's budget proposals on a unified basis—for example, through its budget committees or the Joint Economic Committee. While the other, more specialized congressional committees could advise on specific matters, it would not be appropriate for those committees to set the regulatory budget authorizations for their client agencies. As the originators of legislation giving rise to regulations, the specialized committees could not reasonably be expected to be any more objective than the regulatory agencies about the overall limits of total compliance costs in the areas of their particular interest.

To the extent that Congress elected to become involved in individual regulatory issues, it could do so through its existing substantive appropriations and oversight committees. But the regulatory budget could not be successful as a management tool if Congress were to inject itself into substantive regulatory details when it acted on the President's requests for authority to impose compliance costs.

E. Public Participation in the Regulatory Budget Process

Providing an opportunity for public participation in government decision processes has in the past several years become a matter of major concern and effort. Thus the issue of public participation in the regulatory budget process needs consideration. As before, it is useful to distinguish between the formulation and execution stages of that process.

1. IN THE FORMULATION PROCESS

There would be no more reason to have the public participate in the formulation of a regulatory budget, prior to its submission to the

[26] The authors are aware that legislation to establish a regulatory budget system has been introduced in Congress. As the purpose of this paper is to present a comprehensive analysis of the entire regulatory budget approach, no effort has been made to make the discussion consistent with articular provisions in any draft legislation, nor are specific comparisons made between the draft legislation and the ideas set forth in this paper.

Congress, than in the formulation of the fiscal budget. Traditionally, the fiscal budget has been kept confidential until the President submits it to Congress. Members of the public then have their say during the congressional deliberations on the final form of the budget.

A regulatory budget would best be handled in the same manner. Until summary data were available on the compliance costs to be imposed on the economy during a future fiscal year, there could be little useful public discussion of whether the relative and absolute levels of those costs were appropriate. Such data would be available for the first time when the President's regulatory budget proposals were sent to Congress. The public could then submit comments at the hearings that Congress would hold on the proposed regulatory budget.

2. IN THE EXECUTION PROCESS

Public participation in the execution of a regulatory budget would involve two key issues:

Which regulations would be acted upon when the budget limit forced a choice; and

Whether the compliance-cost estimates for the proposed or promulgated regulations were valid.

The choice among competing new regulations would be as amenable to public participation under a regulatory budget system as it is currently without it. Public interest groups already have ample ways to communicate their concerns on new regulations to politically responsible regulatory officials. Thus, no special further procedures for public participation would seem necessary.

Public comments on the validity of compliance cost estimates, in contrast, would involve at least new forms, if not new channels, of public participation. Under a budget for regulatory compliance costs, various segments of the public would have a stake in the cost estimates finally certified by ORB. The Federal agencies' own estimates would tend to appear, on the one hand, too low to those likely to bear the costs and, on the other hand, too high to those advocating more stringent regulation.

There is a possible problem of equity in public access to government under a regulatory budget. Most of the public comments on the compliance-cost estimates would come from groups with relatively large stakes in the certified cost figures. Groups with smaller stakes—for instance, small businesses or widely dispersed groups such as consumers or the poor—might be discouraged from participating by the costs of submitting comments.

This is not the place to debate the validity of the foregoing argument.[27] It is pertinent to note, however, that the same problem exists under the present system of government regulation. The existence of a regulatory budget would add a dimension to the problem, in that compliance-cost estimates would play a larger role in determining

---

[27] There are various ways in which groups whose individual members cannot represent their own interests effectively can nevertheless be heard in government decisionmaking processes. Virtually every industrial and business group has an association that can prepare and submit comments on behalf of its members. In some cases, the government itself subsidizes the comments—for example, through intervention by the Small Business Administration.

what is regulated and what is not. But having a regulatory budget could also reduce the costs of access to the public decisionmaking process by providing a formal structure, where now only informal procedures are used.

F. Special Problems of Introducing a Regulatory Budget System

Three special, interrelated problems would be encountered in starting up a regulatory budget system:

How should a regulatory budget be phased in?

Which agencies should be included in the introductory phases?

Should a regulatory budget initially cover only the compliance costs of new and revised regulations, or the total costs of complying with all regulations?

1. PHASE-IN

A regulatory budget could not be installed overnight. A large amount of preparatory work would be needed to make even a skeleton regulatory budget viable. Lead time would be required to build staffs and train operating personnel in both ORB and the agencies. ORB would need staff to develop the budget procedures, without which the budget formulation process could not begin. The agencies would need to acquire and train staff to develop the estimates of compliance costs.

The lead time required before agencies could be subjected to regulatory budget constraints would be at least two years—none too generous an amount of time to allow ORB to be formed and initially staffed, and then to develop guidelines for the agencies on how to prepare their budget requests. To illustrate: If the final decision (represented, for example, by enabling legislation) to proceed with a regulatory budget were made in July 1980, the first period during which the new constraint applied would be fiscal year 1983 (October 1, 1982–September 30, 1983). The agencies would send their initial requests for regulatory budget authorizations to ORB in the summer of 1981. The first budget allocations would be issued to the agencies about September 1982, to take effect October 1, 1982.

Once an initial schedule were adopted, interim procedures would be needed to head off agency attempts to promulgate as many regulations as possible before the regulatory budget went into effect. One procedure would be to count the compliance costs of regulations promulgated during the lead-time period against the budget allocation for the first year of operation. In the above illustration, all regulations issued after October 1980 would be charged against an agency's fiscal 1983 budget authorization.

2. INITIAL AGENCY COVERAGE

The philosophy that it is easier to start something new by trying it out first on a limited scale would argue for beginning a regulatory budget with only a few agencies. Its coverage could be expanded later, once some experience had been gained.

It would be several years, however, before any experience with regulatory budgeting could be analyzed and lessons drawn from it.

Considering the lead times required, it would be six or seven years after the initial decision to try regulatory budgeting before additional agencies would be included (two years to implement the initial trial; two or three years of operating experience and evaluation; and two years to implement the expansion itself).

Starting small, moreover, could lend an air of experiment that would encourage the agencies initially covered to do everything in their power to make the system fail. In addition, officials in the trial agencies would feel discriminated against; the resulting resentment would further motivate them to sabotage the system.[28] Finally, partial initial coverage would provide an opportunity for both legislative and executive policymakers to shift regulatory actions to the exempt agencies.

On balance, therefore, it would appear preferable to apply a regulatory budget simultaneously to all regulatory agencies right from the start.

3. NEW AND REVISED VS. TOTAL COMPLIANCE COSTS

As discussed in chapter I, a regulatory budget operating at full scale would deal with the *total* compliance costs imposed by Federal regulations. The magnitude of the task of estimating the total continuing compliance costs of all regulatory requirements, however, would make it impractical to cover them in the first few years of regulatory budget operation.

A workable compromise would be to begin regulatory budgeting with only the compliance costs of *new and revised* regulations, and to shift to a *total* compliance-cost base after a period of a few years. The cost data for new and revised regulations that would be generated during the first few years of budget operation would ease the eventual shift to total compliance-cost budgeting.

IV. Measuring Compliance Costs for a Regulatory Budget

Without estimates of the costs of regulation there could be no regulatory budget. Yet the art of estimating regulatory costs is not fully developed. The existing detailed cost studies encompass only a handful of agencies, and the cost estimates in those studies apply only to a specific industry or group of industries. There is as yet no generally accepted convention regarding the cost elements to be included or excluded, nor is there a single methodology for cost calculation which has been generally applied.

Nevertheless, the budgeting of regulatory compliance cost would be possible within the present state of the art. The history of the Federal fiscal budgeting process demonstrates that budgetary control can be achieved with far less sophisticated techniques than are being used today. The existence of a regulatory budget would, of course, provide an incentive to develop better procedures for estimating the costs of regulation.

[28] The effects of such resentment were illustrated by President Nixon's Quality of Life Review. The review was intended to apply to regulations proposed by all health and safety regulatory agencies. In fact, it was applied almost exclusively to regulations proposed by the Environmental Protection Agency (EPA). EPA officials and their clientele continually railed at what they perceived to be discrimination. Early in 1977 the acting Administrator unilaterally refused to continue to subject EPA to the review.

This chapter examines the measurement of regulatory compliance costs.[29] The concept of compliance costs is defined, and the desirable criteria for evaluating measures of compliance costs are discussed. The process of estimating compliance costs for existing vs. proposed regulations, and the assignment of compliance costs to appropriate fiscal years are then explored. The chapter concludes with a survey of the current state of the art of estimating regulatory costs.

A. Definition of Compliance Costs

Compliance costs consist of expenditures made expressly to meet the requirements of Federal regulations. The expenditures may be made by the private sector, by State and local governments, or by other Federal agencies. In principle, only incremental costs due to regulation—that is, costs in excess of what would have been spent without regulation—should be counted in compliance costs. As noted below, in practice the precise isolation of incremental costs is difficult.

Both capital and operating costs may be incurred in complying with regulations. Examples of capital costs of compliance include outlays for extra construction or new equipment. Examples of operating costs of compliance include expenditures on added research and development, extra variable inputs (such as labor and raw materials) additional supporting services, and further administration (such as paperwork).

Weidenbaum and De Fina estimated that in 1976 total compliance costs in U.S. industry were some 20 times greater than Federal administrative costs. For some regulations specific to certain industries, the estimated ratio was greater than 50:1.[30] To date, no estimates have been made of the corresponding ratio of compliance to indirect costs.

Determining exactly what to include in the compliance costs to be charged against agencies' regulatory budgets would not be a straightforward matter. To illustrate the difficulty:

> Capital outlays on girls' locker rooms under Title IX and the expense of seeing that truck drivers keep proper time logs for the Interstate Commerce Commission should obviously be counted as compliance costs.
>
> Less clear would be the cost of employing extra workers as counterparts of Federal inspectors, as the oil companies claim they must do to comply with crude-oil price-control rules.
>
> Controllable expenses like lobbying in Congress or maintaining an office in Washington to monitor the regulatory agencies probably would not be allowed as compliance costs.

B. Criteria for Evaluating Measures of Compliance Costs

A regulatory budget would play an important role in the policymaking process. For that reason, the quality of the measures of compliance costs used in the budget would not be a matter of indifference.

---

[29] Chapter I noted that of the three components of regulatory costs (administrative, compliance, and indirect), only compliance costs could be included explicitly in a workable regulatory budget system. For that reason, this chapter discusses only the measurement of compliance costs.

[30] Murray L. Weidenbaum and Robert De Fina, "The Cost of Federal Regulation of Economic Activity," (Washington: American Enterprise Institute, reprint No. 88, May 1978).

What is acceptable quality is ultimately a political matter that cannot be resolved in this paper. It is useful, however, to suggest certain criteria that would affect the quality of compliance-cost measures and that would therefore provide points of reference against which to evaluate particular measures and the methods used in obtaining them. The following list of criteria also indicates possible directions for further work on the methodology of measuring compliance costs.

1. PRECISION

A workable regulatory budget would have to employ point estimates of specific compliance costs. The closer those point estimates were to actual costs, the greater the success in husbanding the resources claimed by Federal regulation—the very purpose of a regulatory budget. In part, precision would be a function of time: more precision would be possible the longer a given regulation had been in effect. Perfect precision, of course, would be unattainable; perfect precision would be far too costly to strive for. Note, however, that even the easily quantifiable dollar magnitudes in the fiscal budget, after years of evolution, are not precise: estimating errors of millions of dollars are common.

2. CONSISTENCY

In a workable regulatory budget, individual compliance costs would have to aggregate into totals (by agency, sector, or total budget) that reflected the resources claimed by efforts to conform with regulatory requirements. A million dollars' worth of costs should represent the same claim on economic resources in one agency or sector as in the next. If that were not the case, the incentives created by a regulatory budget would be distorted, and the rationale for such a budget would be weakened.

The methodology used to estimate compliance costs would have to be consistent across different regulations and across the various regulated sectors of the economy. Note that striving for greater consistency might lead to a loss of coverage in compliance-cost estimates if one element of costs could be measured in some but not in others.

3. TRANSPARENCY

An important quality of compliance-cost measures would be that anyone using the same data and methodology could duplicate the estimates used in a regulatory budget. To permit duplication, both the data and the methodology would have to be transparent to all parties concerned. The procedures and assumptions of the methodology would need to be visible and well documented.[31] The data, once decided upon, would have to be readily available and subject to outside, independent

[31] Federal policymakers and others might want on occasion to examine regulatory compliance costs broken down by specific economic sector or geographic area, or divided into individual cost components. To permit this, it would be necessary to retain the detailed characteristics of the disaggregated compliance cost data that would be summed to get the agency and overall totals. This is because aggregation is an informationally irreversible operation. The precise degree of detail retained (e.g., 2 vs. 4-digit Standard Industrial Classification codes, region vs. state, or capital vs. construction costs) would depend on the desired uses and could be periodically adjusted.

audit. Note that the audit requirement could pose problems of privacy, particularly for business firms.

4. THE COST OF MAKING THE ESTIMATES THEMSELVES

The process of obtaining measures of compliance costs would itself entail costs. Data collection, processing, analysis, and dissemination all require the use of scarce resources. If the cost estimation itself were judged burdensome (a matter for political judgment), it would constitute a powerful argument against adopting a regulatory budget. Obtaining compliance-cost measures would probably be most burdensome when a regulatory budget process was first introduced.

The cost of making the estimates would not, of course, be independent of the degree of precision or consistency required of measures of compliance costs. For example, it might be necessary to sacrifice some precision to reduce the burden of making the cost estimates.

C. Compliance Costs of Existing vs. Proposed Regulations

The logic of a regulatory budget would require measures of the compliance costs of both existing and proposed (new or revised) regulations. Only with both existing and proposed regulations covered would the discipline of the budget cover the total costs of complying with Federal regulation. It is important to recognize that the decision to leave an existing regulation in place for another budget period has the same cost impact as the decision to introduce a new regulation with equal compliance costs. Also, the costs of complying with existing regulations would have to be known if agencies were to receive budget credits for regulations that they removed (as was suggested in chapters I and III).

It was noted earlier, however, that obtaining compliance-cost measures for all the regulations promulgated prior to the adoption of a regulatory budget would be a formidable task. It would be inordinately expensive to achieve in the first few years of operation, at least with any degree of precision, consistency, and transparency. Thus, a gradual approach would be required—perhaps one that attempted estimates only for broad categories such as major existing programs, or even entire agencies, in the first few years after adoption. With time, however, more detailed measures of the compliance costs of existing regulations could be developed. Once acceptable base-cost measures had been obtained, annual updates would be relatively straightforward.[32]

Measuring the compliance costs of proposed (new or revised) regulations would pose much less of a problem than existing regulations, in terms of the volume of information and computation involved. Proposed regulations would, however, have estimation problems of their own. Unlike existing regulations, for which empirical track records could be compiled (given enough effort), the costs of complying with

[32] The history of Federal fiscal budgeting is instructive. Only after World War II was systematic attention paid to existing, or base, expenditures as well as to increments or decrements to that base. Even in the late 1970's, the full transition to *zero-based* budgeting, in which all of an agency's expenditures are in principle open to review, is not yet complete.

56-368 O - 80 - 12

proposed regulations would by definition have to be estimated on the basis of few hard data.[33]

Preparing estimates of the compliance costs of proposed regulations would require two distinct steps. First, it would be necessary to forecast the probable methods of compliance. Second, constructive cost estimates would need to be prepared for those methods. Of the two steps, the first would appear by far the more challenging and difficult. It would probably require some active involvement by the parties affected by the regulations. Such involvement could take the form of voluntary comments on the agencies' preliminary forecasts, or of more formal analysis or audit.

In the case of larger business firms and governmental units, there is precedent for such involvement under the present system. For small businesses and governments, and for consumers, however, there is little precedent. The exact nature of active involvement by fragmented, unorganized individuals is not easy to envision. Representation of consumers by public-interest groups would be one possibility, although it would be fraught with possible objections. And as suggested in chapter III, small business could be represented by trade associations.

D. Assigning Compliance Costs to Appropriate Fiscal Years

The costs of complying with a regulatory requirement are typically incurred over a number of years. Moreover, the time pattern in which compliance costs are incurred bears no simple relationship to the fiscal year in which the requirement is promulgated. Indeed, some costs (especially investments) may actually anticipate (and thus precede) promulgation of a new regulation.

A successful regulatory budget system would have to include *all* compliance costs within its authorization limits regardless of when they were incurred. In theory, this could be accomplished with a system which charged agencies with the compliance costs of all their regulations in force in the fiscal year in which the costs were incurred.

In practice, however, charging compliance costs as incurred could undermine the goal of using a regulatory budget system as a management tool. Politically appointed heads of regulatory agencies—who can reasonably expect to be in their jobs for relatively short periods—would have little incentive to worry about the future budgetary claims from regulations that they take the credit for promulgating. This could lead to neglect of future compliance costs in decisions on new regulatory requirements.

An alternative to charging compliance costs only as incurred would be to charge an annualized compliance-cost value of each regulation that was in effect. At the time a regulation was promulgated, an agency would estimate the time pattern of compliance costs, calculate its present discounted value, and convert that value to an annualized (i.e., annuity-equivalent) amount.[34] In that manner the regulatory budget process would better take into account the total compliance costs over the expected lifetime of a regulatory requirement.

---

[33] Parallels in the fiscal budget are the estimates of outlays on new programs or of revenue from tax changes.

[34] By the principles discussed in ch. III, ORB would take into account public comments as well as the agency's estimates when it certified the time pattern of costs.

The annualized cost approach would have the disadvantage of masking the economic impact of extraordinarily large costs incurred in one or two years (for example, automobile-industry retooling required for fuel economy and emission standards). By the same token, the failure to use annualized compliance costs would tend to discriminate against regulations that require large outlays in one or two years, but provide benefits over long periods with relatively low additional costs.

A variant that would combine elements of both of the above approaches would be to require agencies to submit multi-year regulatory cost estimates along with their annual requests for compliance cost allocations. There is precedent for this variant in fiscal budgeting for costly programs extending over several years (e.g., weapons systems). The multi-year cost estimates would at least raise the issue of future costs in the formulation of a given year's regulatory budget. It would, however, impose relatively weaker discipline on future costs than the annualized cost approach.

The present analysis suggests that the use of annualized compliance costs would be the best approach to capture total costs over the lifetime of a regulatory requirement.[35] However, the issue merits further study, both of the defects of charging compliance costs as incurred and of the actual calculation of annualized costs.

E. The State of the Art in Estimating Regulatory Compliance Costs

It is useful to conclude the discussion of measuring compliance costs with a survey of existing studies of the cost of regulation. The methodologies used in the 70-odd studies examined range from informed guesses to detailed, highly structured cost accounting. The existing studies address primarily administrative costs and certain elements of compliance costs. Although some of the studies comment on the existence of indirect costs, only a few quantitative estimates of indirect costs have been made.

Table 1 gives a breakdown of the studies examined by affiliation of authors and time period analyzed (historical or future). A study that addresses both historical and future costs is counted in both columns. Independent investigators were responsible for more than half of the historical studies, but this group attempted almost no future estimates. This is not surprising in view of the previously mentioned need to forecast the method of compliance before estimates of future costs can be made. Business firms are better able than independent investigators to make forecasts of compliance methods, and the Government can require such forecasts from industry.

TABLE 1.—EXISTING REGULATORY COST STUDIES

Affiliation of author	Historical costs	Future costs
Government (or contractor)	17	15
Industry (or contractor)	6	5
Academic institutions	27	2

[35] Annualized cost also appears to be best way to handle the phase in of regulatory budgeting that is suggested in chapter III. In that approach only new and revised regulations would initially be subject to the constraint of a regulatory budget allocation, and the compliance costs imposed by regulatinos that were effective prior to the start of regulatory budgeting would become a formal part of the system only some years later.

There is a glaring lack of uniformity in the costs included in or excluded from these studies, as well as in the methods used to account for those costs. Hence it is difficult to draw comparisons. Nevertheless, several pairs of studies conducted by the two opposing parties to a specific debate suggest the range of discrepancies that can occur. Here are three examples:

(1) In 1978 the Consumer Product Safety Commission estimated the direct compliance cost of a proposed fabric flammability regulation for the furniture industry at $57 to $87 million a year; the market impacts would be a 2 to 3 percent increase in the wholesale price of furniture and added consumer costs of $144 million per year. The American Textile Manufacturers Institute, in contrast, estimated direct compliance cost at $1.3 billion per year.

(2) In 1978 the Environmental Protection Agency (EPA) estimated the annual direct compliance costs of a proposed ambient air quality standard for ozone at $6.9 to $9.5 billion per year. The Council on Wage and Price Stability, using a different but equally logical methodology, estimated those direct compliance costs at $14.3 to $18.8 billion per year.

(3) In February 1979, the Environmental Protection Agency estimated that a proposed 1981 standard for diesel engine particulates would actually result in a net saving of $160 per ton of particulates removed. The Council on Wage and Price Stability (COWPS) estimated that the same standard would impose a cost of $4,740 per ton. For a more stringent proposed 1983 standard, the estimated costs per ton were $3,200 (EPA) *vs.* $7,650 (COWPS).

There are two explanations for such wide discrepancies in estimated compliance costs. First, there are currently no generally accepted conventions for choosing the specific costs to be included or for the methodology of computing total compliance costs. Second, because each of the above examples concerned the impact of a proposed regulation, it was necessary to make assumptions about the methods to be used for compliance. The assumptions of the opposing parties were quite different in each case. The existing studies of historical costs, where the methods of compliance were known, show much smaller discrepancies.

An important contribution to the literature on regulatory costs was made by the first issue of the *Regulatory Calendar*, which appeared in the *Federal Register*. February 28, 1979. The calendar (which will be published semiannually by the Regulatory Council) listed 109 major rules being considered by 20 Federal departments; compliance cost estimates were included for about one-third of the entries.

In a statement at the time of publication, Douglas Costle, the chairman of the Regulatory Council, acknowledged that "agencies presently calculate costs in different and sometimes conflicting ways," and he cautioned against attempting to aggregate the costs published in the calendar. He called for development of better cost-estimating methods and for dissemination of those methods among the agencies. Nevertheless, the work currently being done by the Council may help lay the groundwork for making the cost estimates that would be needed to implement a regulatory budget system.

Significant progress toward resolving the problems of estimating regulatory compliance costs, particularly settling upon a uniform methodology, was made in a study commissioned and recently released by the Business Roundtable.[36] That study, conducted by Arthur Andersen and Company, estimated the costs incurred in 1977 by 48 cooperating companies in complying with all the regulations of six agencies or statutes:

(1) Environmental Protection Agency.
(2) Occupational Safety and Health Administration.
(3) Equal Employment Opportunity Commission.
(4) Department of Energy.
(5) Employee Retirement Income Security Act.
(6) Federal Trade Commission.

The 48 participating companies represented 23 two-digit SIC industry groups; each group contained a minimum sample of three company divisions or other operating units.

While the numerical results of the Arthur Andersen study themselves are of interest, much the more interesting for present purposes is the methodology developed for the study. The procedures for consistently determining compliance costs across firms and regulations, together with the supporting documentation from the companies, would provide the kinds of precise, consistent, and transparent cost estimates required for a workable regulatory budget. Of particular note is the assiduous care taken to determine the *increment* in company costs due solely to complying with regulations. The methodology has been criticized by some economists for failing to include indirect as well as compliance costs; for purposes of a regulatory budget, however, the criticism is beside the point, as argued in chapter I of this paper.

Arthur Andersen and Company estimated that the study itself cost 0.4 percent of the computed compliance costs. Subsequent studies made by the *same* participants would probably reduce that percentage because of the experience gained in the initial study.

The studies surveyed above illustrate a point made at the outset of this chapter—namely, that the art of estimating regulatory costs is still undergoing development. At the same time, a handful of those studies—in particular, the *Regulatory Calendar* and the Arthur Andersen study for the Business Roundtable—also illustrate a second point made at the outset: a workable regulatory budget would be achievable within the present state of the art.

---

[36] The Business Roundtable, "Cost of Government Regulation Study" (New York, 1979).

THE IMPACT OF REGULATION ON THE PERFORMANCE OF INDUSTRY

By Paul W. MacAvoy and Dorothy M. Tella*

EDITOR'S NOTE.—The paper "The Impact of Regulation on the Performance of Industry" is drawn from the volume "Government Regulation of Business: Its Growth, Impact, and Future". Chamber of Commerce of the United States, 1979, and is reprinted with permission.

CONTENTS

	Page
Price Regulation in the Utilities and Transportation Industries	178
Regulation in a Noninflationary Economy	179
Regulation in an Inflationary Economy	180
The Interaction of Regulation and Inflation	182
Health, Safety, and Environmental Regulation	185
Anticipated and Realized Effects	187
Quality-of-Life Improvements	189
A New Era of Effectiveness	191
Prospects for Reform	192

Until the past decade, state and federal regulatory agencies had concentrated mainly on setting prices and establishing levels of service in public utilities and interstate transportation. As recently as the mid-1960s, the regulated sector of the economy comprised a relatively few industries—electric and gas services, finance and insurance, transportation, communication, and natural gas extraction—that together accounted for about one-tenth of GNP.

Today the landscape looks very different indeed. Price controls have been extended to petroleum production, refining, and marketing, and health and safety and environmental quality regulation has brought forth a new set of agency controls that constrain company operations in much of the rest of industry as substantially as direct price controls do in the public utilities. Several industries—including mining, construction, chemicals, paper, primary metals, petroleum refining, motor vehicles, and stone, clay, and glass—now devote a substantial percentage of their investment to complying with environmental and occupational safety and health standards. With regulation affecting their capacity growth, and with their major products or production processes subject to regulatory agency specifications, the performance of these industries is as much influenced by agency controls as that of the utilities. If one counts this group of industries as effectively regulated, the regulated sector of the U.S. economy now accounts for close to one-fourth of GNP—more than double its size less than 15 years ago.

As controls have spread to new parts of the economy, antagonism to regulation also has increased, at what seems to be no less rapid a

*Paul W. MacAvoy, a former member of the Council of Economic Advisers, is professor of organization and management and economics at Yale University. Dorothy M. Tella is director of the Trends and Perspective Center, Chamber of Commerce of the United States.

rate. Not surprisingly, the most vocal of the antagonists have been companies and trade associations in the industries most affected by health and safety and environmental controls. However, dissatisfaction with regulation and the way it has been working is by no means confined to those bearing the costs of new controls. Environmental and public interest groups that initially pressed for legislation creating new agencies to protect health, safety, and the environment are critical of the way the agencies have been operating. Companies traditionally protective of price regulation in the natural gas, airline, and railroad industries have increasingly been calling for reduction or elimination of controls. Consumer groups have also been supporting deregulation in areas, such as air transport and trucking, where 10 or 15 years ago opposition to regulation was confined in the main to academic economists. Dissatisfaction with some part of the control apparatus now seems to come from nearly every quarter.

Growth and dissatisfaction in regulation have come together almost as a matter of course. The tasks set for the agencies have been immense and complicated—perhaps beyond the capacity of the regulatory process. And the results have, in the administrative process, fallen far short of the goals set. This essay is concerned with the results of regulation: what has happened, and why, and what is likely to happen if there are no significant changes in the present practices of the regulatory agencies.[1] The first part of the essay deals with price regulation in the public utilities and transportation industries. The second deals with safety and environmental regulation by the three agencies whose activities to date should have had an important impact on the operations of industry: the Environmental Protection Agency (EPA), the Occupational Safety and Health Administration (OSHA), and the National Highway Traffic Safety Administration (NHTSA).

As we trace the effects of these regulatory activities, it will be apparent why there is such widespread dissatisfaction with controls today and why it is of fairly recent origin. Some part of the responsibility lies with the economic conditions of the times. Price regulation has had quite different results since the late 1960s than in the period up to then. Operating under conditions of high general inflation and even more sharply rising input costs in the late 1960s and early 1970s, the regulatory process caused profitability in the regulated industries to decline and investment growth to be slowed. As a consequence, consumers have encountered declining service quality and, in extreme cases, capacity shortages. Inflation and the business cycle have not been the only determinants of regulatory performance, however. In the case of safety and environmental regulation, controls themselves have had the predictable effect of raising prices and reducing output. At the same time, they have not generated the hoped-for benefits in terms of improved health and safety conditions. As we shall see later in our paper, the evidence to date is that there have been few if any nationwide improvements in accident rates or environmental quality that can be attributed to the regulatory activities of EPA, OSHA, or NHTSA.

Since regulation across the board has apparently worked poorly, one might expect that the prospects for broad-scale reform would be as-

[1] For a more detailed treatment of the effects of economic and social regulation during the 1960s and 1970s, see Paul W. MacAvoy, "The Regulated Industries and the Economy" (New York: W. W. Norton & Co., Inc., 1979).

sured. There has indeed been some lessening of controls in the past five years or so. Legislation to deregulate the airlines has been passed and there have been other, less spectacular deregulatory measures as well. There has been some legislated reduction in agency controls in banking and the railroads, and changes in rate-setting procedures instituted by regulatory agencies themselves have resulted in electricity and natural gas prices more closely approximating market prices. Both the Ford and Carter Administrations have attempted to introduce cost-benefit or economic impact analysis into the regulatory process as a step toward making health and safety and environmental quality regulation cost-effective and beneficial. But these initiatives on the whole have been quite limited and the prospects for widespread improvement of the process are not good.[2]

In the transportation industries and the utilities, legislation either removing the price-setting authority of regulatory bodies or specifying rate-making practices that keep prices in line with current costs is needed to assure appropriate rates of capacity growth. However, where current regulatory practices have had their worst effects, such reforms would mean the sharpest price increases, and this is bound to be a stumbling block for Congress and the state legislatures. Institutionalizing cost-benefit analysis in the regulatory process to make health and safety and environmental regulation effective would require that Congress give regulatory agencies statutory authority to trade off conflicting economic and social goals. This is an extremely broad and significant reform initiative which the reluctant congressional committees are unlikely to take, since they themselves do not make such trade-offs. Therefore, our assessment of what is likely to happen if there is no change in the current practices of regulatory agencies may also represent the best prediction of what will actually occur.

Price Regulation in the Utilities and Transportation Industries

The privately owned electric and gas companies, telephone companies, airlines, railroads, and motor carriers have been regulated for somewhat different reasons and with somewhat different expectations about what agency controls would accomplish. The utilities in general were regulated because the companies were local monopolies. Regulation was expected to hold prices and profits to competitive levels, thus giving consumers the benefit of lower prices and more service than would presumably be forthcoming if the companies were left to exercise their monopoly power. The transportation industries, particularly the airlines and trucking companies, had none of the characteristics of natural monopolies. Although the argument for regulating them cited the need for controlling arbitrary and discriminatory pricing by a single or a few suppliers, it was meant also to emphasize levels and quality of service: Regulation was rationalized on the grounds that consumers would benefit from more or higher-quality service than was likely to be offered in unregulated markets. The agencies regulating these industries were expected to hold prices to cost-of-service levels; however, to the extent that they required the companies to provide

[2] Recent reform initiatives and their limitations are described in *ibid.*, Chapter 4.

more or better service, it was expected that the cost of service, and hence the price, would also have to be higher.

Despite some differences in the rationale for regulating and in the practices of the individual regulatory bodies, the regulatory process has turned out to be quite similar in all the industries. Partly this reflects the fact that all the regulators are subject to court review of their procedures and decisions, and partly it came about as the newer of the regulatory bodies borrowed from the practices of the older. In most cases, the regulatory agency or commission establishes the dimensions of the market—who may serve it and what services are to be provided—and determines the level and structure of prices. The rate-setting process involves the regulatory body's hearing and deciding on company requests for price increases. In general, increases are allowed to the extent that they can be justified on cost-of-service grounds—that is, to the extent that the increase in revenues generated by the higher prices does not exceed increases in operating costs, depreciation on plant and equipment, and the "fair" or "reasonable" return on investment.[3]

This common regulatory process can produce quite different results depending on the structure of the different regulated industries and the objectives of different regulatory bodies. It can also produce different results under different conditions in the overall economy. As we shall see, regulation had quite different effects on both the utilities and the transportation industries in the 1950s and early 1960s, when prices were stable and industry costs falling, than in the late 1960s and the 1970s, when inflation was high and industry costs in general were rising.

Regulation in a Noninflationary Economy

In the 1950s and early 1960s, regulation appeared to do its share to establish constant prices in the regulated industries. However, there was little incentive for companies to increase prices in that period, since costs were falling and demand for service at constant prices was expanding rapidly. Therefore, the regulatory constraint on price increases was not very severe. Natural gas pipelines were found to be charging no less in the early 1960s in their regulated sales to retailers than in their unregulated sales to industrial consumers. The rates of regulated electric companies were not substantially lower than the rates of unregulated companies. Some differences between the charges of regulated and unregulated companies to industrial and commercial customers were in evidence in the 1940s and 1950s; however, no differences in rates for residential customers (the group presumably most in need of protection from monopolies) showed up until close to 1960, and, even then, regulated electricity prices were only slightly lower.

---

[3] A general formula for the regulated limit on revenues is $p \cdot q = c + d + (r \cdot B/(l-t))$ where p is the price or rate level on sales of q; c is total operating costs; d is depreciation; and $r \cdot B/(l-t)$ equals total allowed profit returns before taxes at the tax rate t and at the fair rate of return r on rate base B of undepreciated investment. The estimates of c and d are based on accounting data for some recent test period of operations and as such are not significantly controversial. But $r \cdot B$ contains two subjective estimates: the fair rate of return r and the company's undepreciated capital base B, which because they are subjective and judgmental could exceed costs and then be compounded by multiplication. It is in determining the r and B factors in capital returns that the regulatory review effectively constrains company decisions.

Residential telephone customers were indeed the beneficiaries of substantially lower rates brought about by regulation. However, this was not because the regulatory bodies successfully forced monopoly prices to cost-of-service levels, but rather because the Federal Communications Commission (FCC) and the state public service commissions used their pricing authority to establish a pattern of cross-subsidization whereby above-cost rates for interstate calls provided the revenues needed to keep local rates well below the cost of service. In all these industries, the commissions and companies worked to expand service at constant prices.

The results of regulation in the transportation industries were along lines consistent with the original dual rationale for regulating—to fix prices and to expand service. Regulation of the airline, railroad, and trucking industries fixed prices for high-volume service but at the same time raised prices in these industries by very substantial margins, in many cases by as much as 30 to 50 percent.

The higher prices did not generally show up in higher profitability for the regulated transportation companies. Mainly, this seems to have been because the Interstate Commerce Commission (ICC) and the Civil Aeronautics Board (CAB), in line with their legislative charters to promote expansion and improvement of service, required the companies to provide higher quality, and hence more costly, service or to provide service at less than cost to particular categories of customers. Both the CAB and the ICC used their rate-making authority to promote expansion of service. Under the rate structure prescribed by the CAB, profits from long-distance service were to be used to expand short-distance flights and service to small communities, where fares would be low but per mile costs were high. Under the ICC's rate policies, the railroads' above-cost rates on high-density lines provided revenues to subsidize service on light-density feeder lines and to shippers in small communities.

With public utility and transportation regulation apparently producing constant prices and expanded service, the prevailing view of politicians and the public seemed to be that regulation was working quite successfully. The regulated industries were outperforming the economy as a whole in real GNP growth in the late 1950s and early 1960s, and prices were increasing more slowly in the utilities and transportation industries than on average in the unregulated service industries. Price stability, however, had little to do with regulation, since productivity was increasing rapidly in all the regulated industries—more rapidly in every case than the average for the unregulated service sector—and unit costs were declining. From 1958 to 1965, unit labor costs fell in all the regulated industries except trucking and, in all except the railroads and the telephone companies, unit capital costs were also either declining or stable. With costs decreasing, prices could have been expected to fall or at least remain stable in the absence of regulation.

Regulation in an Inflationary Economy

The effects of regulation on prices quickly changed when inflation set in and industry costs began to rise in the late 1960s. The first surge

of costs left regulated prices far behind. In the electric and gas utilities, the telephone companies, and the airlines, regulated prices barely rose at all from 1965 to 1969. Meanwhile, labor and capital costs were increasing at rates greater than one percent a year in the utilities and upward of five percent a year in the airlines.

One's instinct might have been to see this as a sign of increased regulatory effectiveness—to believe that the commissions were, with a lag, bringing regulated prices into line with costs of providing service. Had this been the case, however, output and investment should have increased. The lower prices forced by regulation should have increased demand, and enlarged demand should have induced the companies to add to capacity. What actually happened was quite different: The lag of regulated prices behind costs in the latter 1960s in due course pushed investors' rates of return below capital costs and, as a consequence, investment and capacity growth were reduced.

The pattern of reduced profitability is shown in table 1. As indicated, declining profit margins in the late 1960s caused declines in investors' rates of return, pushing returns in all the regulated utilities and transportation industries down to, if not below, the equivalent market rate of return. The predictable effects on investment began to show up in the 1970s (as shown in table 2). In the electric and gas utilities and the airlines, the rate of growth of real net investment fell sharply in 1969–77, and in the gas utilities, airlines, and railroads it was actually negative. In the unregulated service sector, investment growth was affected by the cyclical downturns in the economy in the early and mid-1970s, but it was still slightly higher in the 1970s than in the late 1960s.

TABLE 1.—RATES OF RETURN IN THE REGULATED INDUSTRIES

Industries	1958–65	1966–69	1970–77
Annual rate of return on sales (percent):			
Transportation	10.3	14.3	9.8
Public utilities	33.0	36.6	28.4
Unregulated service industries [1]	7.1	7.6	7.1
Annual rate of return to stockholders (percent):			
Transportation	[2] 10.0(+2.2)	5.5(−0.5)	3.3(−1.4)
Public utilities	16.5(+9.9)	1.9(−3.8)	6.3(+1.0)
Unregulated service industries	11.9(−1.1)	9.7(+2.6)	3.6 (−.7)

[1] Includes wholesale and retail trade; insurance agents, brokers, and service; hotels and other lodging places; personal services; business services; auto repair, services, and garages; motion pictures; and amusement and recreation services not elsewhere counted.

[2] The differences between stockholders' return and the return on other stockholdings of equal "risk" (as measured by adding calculated "beta" risk premium to the return on the S & P 500 stock average).

Source: As estimated from Standard & Poor's "Compustat" data on individual company financial inputs. The return on sales is estimated as sales minus cost of goods sold divided by sales and the return on investment, as stock appreciation plus dividends divided by stock price.

TABLE 2.—INVESTMENT IN THE REGULATED INDUSTRIES

Industries	1958–65	1965–69	1969–77
Annual rates of change (percent):			
Transportation	10.6	2.2	−3.0
Public utilities	10.0	14.2	5.1
Unregulated service industries	7.5	1.0	1.8

Source: U.S. Department of Commerce, Bureau of Economic Analysis.

As a consequence of lower investment growth, output growth slowed in all the regulated industries. (See table 3.) Reduced growth rates were explainable in part by the 1974–75 recession, but the decline in growth in the regulated industries was far more than the decline in the rest of the service sector of the economy. Whereas in the 1960s growth rates in all the regulated industries except the railroads were substantially higher than the average for the rest of the service sector, by the mid-1970s rates were lower in the regulated than in the unregulated service industries. Thus, the same industries that through the mid-1960s had been pushed by regulation to expand saw regulation force down rates of return in the late 1960s and put an end to their history of expansion and superior performance.

The Interaction of Regulation and Inflation

The important question is whether regulation and inflation interacted in ways that could only be detrimental to the regulated industries. To answer it requires looking in some detail at how the regulatory process works. The two adjectives most used to characterize this process are "slow-moving" and "backward-looking." Whereas market prices change in response to changes in costs and demand, regulated prices change only after past cost changes are recorded and then only at the end of what is inevitably a lengthy administrative process. Regulatory bodies have traditionally required companies to justify price increases on the basis of historical costs, that is, costs in some past "test year."

This is of little consequence when costs are relatively stable or falling, but when costs are rising, as they have been since the late 1960s, the slowness of the regulatory process and the requirement that prices be justified on the basis of past costs mean that regulated price changes lag behind inflationary changes elsewhere in the economy. In theory, regulatory bodies should constantly monitor rates of return, permitting the companies to raise prices when profit rates fall below the current cost of capital and forcing price reductions when rates of return exceed the cost of capital. In practice, regulatory bodies as a rule have looked at rates of return only when companies file for price increases. For their part, companies as a rule have filed for increases only when their rates of return were declining and were below what the regulatory body had allowed in its most recent cases.

TABLE 3.—OUTPUT GROWTH RATES IN THE REGULATED INDUSTRIES

Industries	1958–65	1965–69	1969–77
Average annual rates of change in gross product originating by sector (percent):			
Transportation	4.1	3.8	7.0
Public utilities	6.1	6.5	2.4
Unregulated service industries	4.7	4.2	3.4

Source: U.S. Department of Commerce, Workfile 1205–02–03, 1978 revision.

In the late 1960s, when inflation set in and industry costs were rising, rates of return on equity fell, and eventually most regulated companies were obliged to file for rate increases and thus to confront

the regulatory process. From 1964 through 1967, the number of formal rate of return reviews of private gas and electric companies by state regulatory commissions was between eight and 10 a year. In 1969, the number was in the 30s and by 1972, in the 80s. With a larger number of companies included in the request process, the agencies' case loads were enlarged and the time lag between application and granting of higher prices was lengthened. Since new prices were based on costs in the year or so before the companies filed their applications, and since inflation in costs continued in the case period, the commission-determined prices fell short of current costs and the companies again requested price increases. This pattern of "case-stacking" was set in the utilities and, to a lesser extent, in the transportation industries beginning in 1968–69.

Inflation created yet another special problem for the regulated industries that in theory should not have occurred. In principle, the regulated price level should be the level that currently yields a rate of return just equal to capital cost. In practice, regulatory bodies—especially the state public utility commissions—have been as much concerned with preventing price increases as with finding the price level that provided a sufficient rate of return. When utility prices were stable or declining, as they were in the late 1950s and early 1960s, the commissions were not overly concerned if company rates of return exceeded what had originally been allowed or what was being allowed similar companies in current rate cases. But by the same token, the commissions have been reluctant in recent years to grant large price increases even where they have been justified by the customary cost criteria.

By the late 1970s, a decade of high inflation and more stringent regulation had left its mark on all the regulated industries. In the industries most affected by regulatory constraints on profitability, investment had been cut back and there were already signs of capacity shortage. In the least affected industries, there were no immediate problems, but regulatory policies had set the stage for similar reductions in profitability, investment, and service quality farther out on the horizon.

The most adversely affected industry to date has been natural gas production, where capacity shortages developed in the 1960s and production shortages showed up in the early 1970s.[4] The natural gas producing industry is very different in structure from the utilities and the transportation industries and has been regulated in a somewhat different way. Regulation of gas producers represents an extreme variant of public utility regulation. The Federal Power Commission (FPC) started out in the late 1950s to treat gas producers as public utilities and set prices on the basis of individual costs of service. However, because there were thousands of individual producers, the FPC soon turned instead to setting ceiling prices for all gas sold from each major gas-producing region. In the early 1960s, the FPC set interim ceiling prices on new sales at late-1950s levels and pro-

[4] While cost conditions of field gas supply were rather special, the economic conditions that made the natural gas industry the exception in the early 1960s spread to the other price-regulated industries in the late 1960s and early 1970s. These patterns are documented in MacAvoy, *op. cit.*, Chapter 2.

ceeded with hearings to establish final prices. By the end of the 1960s, final prices had been set for only two of the five gas-producing regions, and these were at levels reflecting average production costs in 1960. Since the final prices were thus still equal to or only slightly above the interim ceilings, for all intents and purposes natural gas prices were fixed for a decade at levels approximating the average cost of production in 1960.

Basing regulated prices on average costs in an industry characterized by rising costs of development has the same qualitative effect as setting prices on the basis of historical costs in a period of high inflation. Prices are too low to cover the costs of capital expansion, so capacity growth falls short of increases in demand. Because of the peculiar nature of the natural gas market, the predictable growing gap between demand and capacity did not become visible to consumers until the winter of 1971–1972, although capacity shortage had been building throughout the 1960s. This was so because the natural gas producers' market is a market for reserves, in which the buying pipeline companies do not purchase gas for delivery but rather contract for the right to draw on inground reserves over a specified period. In entering into contracts, pipelines generally sought 10 to 20 years reserve backing for current production levels. With such reserve backing, there was no immediate shortage when the demand for new reserves exceeded the supply; rather, the pipelines increased production from existing reserves. By 1968, the demand substantially exceeded the supply of additional natural gas reserves, yet both new and old customers of the pipeline companies continued to get the supplies of low-price gas they wanted (although with increasingly less reserve backing to assure future service). By 1971, reserves had been so drawn down that during the severe winter that year the major pipelines could not meet some delivery commitments to customers, and the reserve shortage became a production shortage.

After natural gas, the industry most adversely affected by regulation has been the electric power industry. Because of regulatory lag and the commissions reluctance to grant large price increases, allowed rates of return on equity have been substantially below what has been estimated as needed to prevent power shortages in the 1980s, and earned rates of return have generally been below allowed rates. Regulated prices have not been permitted to rise enough to equate capacity expansion and the growth of demand.

There has been limited improvement recently. The average time between filing and the effective date of rate increases in this industry has dropped somewhat since 1976 because of a growing practice of allowing companies to put proposed price increases into effect before the outcome of the regulatory hearing. Commissions have been allowing higher rates of return in recent years, and the rate of return to investors in this industry has increased more rapidly than the market rate of return. But investment in new capacity has been constrained by low rates of return and, barring reduced consumer demand because of accelerating fuel costs, capacity shortages are likely to result in the late 1980s.

In the other regulated industries, the overall lag of prices behind costs has been considerably less severe. For some, however, the failure of regulatory bodies to allow changes in the structure of prices to re-

flect changed market conditions has set the stage for industry problems destined to affect service quality.

Problems are closest on the horizon for the railroads. In the face of rising fuel, labor, and capital costs, the railroads were granted substantial rate increases after 1967. However, the ICC has been inflexible in allowing the railroads neither to discontinue unprofitable service to small shippers and on short-distance feeder lines nor to raise rates for these categories of service sufficiently to cover their costs. As a consequence, the railroads have had to subsidize short-distance service and less-than-carload shipments out of revenues earned from large volume, long-distance services. Since costs for short-distance service have increased rapidly, this pattern of cross-subsidization could only be maintained by raising margins rapidly on the profitable categories of service. However, the high margins on the long-distance, large-volume freight have induced competition from trucking and barge lines. The more rate-cost margins have been increased, the greater the risk that the railroads lose their profitable business to competition. Thus far, the southern and western railroads have been able to maintain both subsidized services and positive overall profitability. However, if costs continue to rise in the next decade, this will no longer be so. Any loss of profitability now is likely to have serious consequences because of the railroads' history of deferred maintenance. If sizable investments are not made in the next several years, the quality of rail service will decline, and the railroads will not be able to handle the growing demand for such specialized services as coal transport, which have the potential for restoring railroad profitability and growth in the long run.

The telephone companies, too, face problems arising from patterns of cross-subsidization fostered by the FCC and the state public service commissions. As the companies' costs rose in the late 1960s the state commissions did not allow rate increases that kept up with costs, but instead minimized increases in residential and other local charges. Most of the overall increases in company revenues came from long-distance service, where costs were still decreasing because of new technology and economies of scale. Keeping prices on long-distance service constant as costs were declining provided the telephone companies with the revenues to subsidize local service. At the same time, however, the resulting increase in rate-cost margins for long-distance service greatly encouraged specialized telecommunications companies to seek entry into long-distance markets, and the FCC allowed and even fostered this new competition within the industry. While the telephone companies face no immediate capital problems as a result, they cannot—any more than the railroads—continue for long to charge prices substantially above costs in their newly competitive markets. Regulatory policies that resort to cross-subsidization instead of allowing prices to adjust to rising costs and changing market conditions will ultimately produce the same pattern of declining profitability, reduced investment, and deterioration in service that has shown up in the electric and gas industries.

Health, Safety, and Environmental Regulation

The rapid growth of health, safety, and environmental quality regulation in the late 1960s and early 1970s is not easy to explain. The

"market failures" cited to justify new regulation did not show up then for the first time, and at least some of the indicators of decline in the quality of life that prompted Congress to regulate—the rise in on-the-job injury rates and in highway death rates, for example—were explained almost wholly by demographic factors beyond the reach of the regulatory process. To be sure, there was greater awareness of environmental and product quality issues, and the federal government was certainly more inclined to intervene on any front than in the past. But even so, stronger policies to protect individual health and safety and the quality of the environment need not have meant a new generation of industry regulation. In every case where Congress chose to regulate, there were alternatives—court penalties for polluters, tax penalties for employers with poor safety records, or government-funded information programs—and, in general, better arguments could have been made for the alternatives than for agency controls.

Why Congress chose to regulate, however, may be less important than what the new regulatory agencies chose to do with the authority Congress gave them. In every instance, Congress laid out broad goals for improvement of conditions affecting health and safety and gave the agencies broad authority to set standards toward these ends. In principle, their grants of standard-setting authority left the agencies many options, ranging, for example, from establishing maximum pollution levels for air corridors to setting forth equipment requirements for individual plants. In practice, the new agencies separately but uniformly turned to specifying particular physical conditions of production and physical characteristics of equipment.[5]

The Clean Air Act Amendments of 1970 gave EPA the authority to set national air quality standards reflecting acceptable levels of exposure to designated pollutants and also to set performance standards for electric power plants and automobiles. Despite the emphasis on regionality and performance in the legislation, EPA from the beginning set performance standards in terms of emissions per unit of production that were based on use of particular control methods, so that air quality performance standards became, in effect, requirements for very specific kinds of equipment. The Occupational Safety and Health Act of 1970 said nothing of the kinds of standards that the Secretary of Labor should set, only that they were "to assure as far as possible every working man and woman in the nation safe and healthful working conditions and to preserve our human resources." However, within a few months of its creation, OSHA had set several thousand separate standards prescribing in detail safe physical conditions in various workplaces. The National Traffic and Motor Vehicle Safety Act of 1966, whose declared purpose was "to reduce traffic accidents and deaths and injuries to persons resulting from traffic accidents," authorized standards for motor vehicle performance that "shall be practicable, shall meet the need for motor vehicle safety, and shall be stated in objective terms." In practice, NHTSA's performance standards, like EPA's, turned out to imply requirements for very specific design features and pieces of equipment on new-model automobiles.

There is no simple explanation for the uniform behavior of the new agencies—for their having taken broad legislative goals and

---

[5] The processes of social regulation, and the issue of design versus performance standards in particular, are discussed in *ibid.*, Chapter 3.

standard-setting authority only to end up focusing on the details of equipment and the physical conditions of production. The reasons are probably similar to the reasons that the old regulatory bodies fell into the pattern of rate setting based on historical costs. The new and the old agencies both needed to find operating procedures that would enable them to handle their case loads in a reasonably efficient way and to arrive at decisions that would withstand judicial review. Cost-of-service data provided a seemingly objective basis for setting prices and thus met these broad requirements. Similarly, equipment standards were tangible and provided an "objective" basis for setting and enforcing health, safety, and environmental regulation.

From the outset, the new agencies were under pressure to demonstrate to Congress and the public that they were taking action to improve the quality of life. Both OSHA and NHTSA were required to put comprehensive sets of safety standards into effect within the first few months of their operation. Both responded by adopting existing standards of government agencies and private safety groups that included, but were not exclusively, equipment standards. Beyond providing a response to pressure for a first wave of standard setting, however, equipment and design standards were appealing because they tended to speed up the regulatory process and thus allowed the agencies to show more evidence of regulatory activity. They had the appearance of objectivity and the clear advantage of specificity, so that there was less occasion for lengthy dispute over the implications of general times and conditions and thus less risk of legal challenge that would slow the regulatory process.

Compelling as it may have been for EPA, OSHA, and NHTSA to settle into a pattern of regulating equipment and production processes, it was a fateful course on two counts. On the one hand, it gave the agencies a degree of control over production and investment in several industries that they could not have had under a regime of performance standards or regional air and water quality standards. On the other hand, it meant that regulation dealt with matters once-removed from the aims of the legislation and, probably largely for this reason, was destined to have far less impact than was hoped on accident rates or on the quality of the environment.

Anticipated and Realized Effects

The goal of regulation, of course, was to improve health and safety and the quality of the environment—to lower the incidence of occupational injury and disease, reduce the rates of highway deaths and injuries, and make the air and water cleaner. It was understood from the outset that regulation would be costly as well. The standards set by the new agencies imposed operating and capital costs on producers that had to be reflected eventually in higher prices for consumers, and it was predictable that the higher prices would reduce demand and, ultimately, production in the industries most subject to regulation.

The industry effects have shown up much as expected. The industries most subject to health and safety and environmental quality regulation—automobile manufacturing; paper; chemicals; stone, clay, and glass; primary metals; non-fuel mining; and construction—have generally experienced larger price increases in the 1970s than other,

56-368 O - 80 - 13

less-regulated industries. From 1969 to 1973, the period in which OSHA safety standards and the first wave of pollution-control standards were being put in force, prices in the heavily regulated industries rose at more than twice the average rate for the less-regulated manufacturing industries.* As a result, the realized rates of growth of real GNP in the most-regulated industries were lower than the average for less-regulated manufacturing. Whereas output increased by an average 2.7 percent a year in the less-regulated manufacturing sector during 1970–73, growth was less than one percent a year in all the heavily regulated industries except chemicals. (The comparative changes for the two groups of industries are shown in table 4.)

TABLE 4.—PRICE AND PRODUCTION CHANGES IN THE INDUSTRIES SUBJECT TO ENVIRONMENTAL AND HEALTH AND SAFETY REGULATION

[Average annual rates of change (percent)]

	1958–69	1969–73	1973–77
Price changes:			
Most regulated (includes mining; construction; paper; chemicals; stone, clay, and glass; and primary metals)[1]	2.2	6.6	8.8
Unregulated manufacturing (total manufacturing with the exception of the most regulated industries and automobile manufacturing)	1.3	2.8	7.9
Production changes:			
Most regulated	4.6	1.6	−.7
Unregulated manufacturing	5.4	2.7	.6

[1] Automobile manufacturing, which is one of the most-regulated industries, requires detailed treatment because of exceptional difficulties in evaluating the gross product originating data series for this industry. A detailed treatment of price and production changes in this industry is found in Paul W. MacAvoy, "The Regulated Industries and the Economy" (New York: W. W. Norton & Co., Inc., 1979).

Source: U.S. Department of Commerce, Workfile 1205–02–03, 1978 revision. Price changes are derived from the gross product implicit deflator. Production changes are derived from gross product originating series.

One can search for other explanations than regulation for the pattern of performance in the heavily regulated industries after environmental and safety standards were put into effect. However, nothing in the prior experience of these industries suggests that the higher rate of price increase is explained by the high rate of general inflation since 1969. The price increases in the heavily regulated industries in the 1970s quite clearly do not reflect larger profits. Profit margins on sales declined during 1970–73, but by more in the heavily regulated industries than in the less-regulated manufacturing sector. Rising prices and reduced output in the heavily regulated industries must have been due primarily to cost increases, and part of these cost increases have to be attributed to the new regulations.

A second round of economy-wide effects from regulation also may have materialized, although this cannot be shown from the data on recent economic behavior, given the ups and downs of prices and production in the middle and late 1970s. Diversion of investment to meet regulatory equipment requirements could be expected to lower the rate of growth of the output-generating portion of the total capital stock, thereby reducing productivity growth. Most likely there has been investment diversion—not all the expenditures for required equipment can have come from unemployed resources or from con-

*By the mid-1970s, industry had largely absorbed the costs of the new controls; hence, marked differences between the rates of price increase in the most-regulated and the less-regulated industries do not show up for the 1973–77 period. The effects of a second wave of pollution-control standards following from the 1977 Clean Air Act Amendments and of OSHA health standards may show up for the late 1970s and the 1980s.

sumption outlays. If so, the level of GNP by the late 1970s was lower than it would have been had part of the outlays for pollution and safety controls gone into output-generating investment.

A third round of effects may be more pervasive but even less evident. Controls not only reduce current production and investment for future production but also constrain the development of new technologies. Regulatory equipment specifications have been backward-looking for the most part. Where forward-looking, they have specified only the safety or pollution-control characteristics of new equipment. The aggregate effects on research and development performance are unknown. Cases have turned up of "forced" or "defensive" development that added little to the quality of products but rather held the line on current production. However, there have been other cases where new "mandated" equipment has been both safer and more productive. On the whole, we suspect that there has been both considerable restriction and considerable diversion of development effort resulting from the application of safety and environmental standards. In the 1980s, these effects may reduce the level of output and the growth of the economy more than the first- and second-round effects.

Quality-of-Life Improvements

As for the other results of regulation—the hoped-for improvements in health and safety and the quality of the environment—there have indeed been improvements in some quality-of-life indicators since regulatory programs began, and the regulatory agencies have taken credit for them. Where there have not been improvements, the claim has generally been made that without regulation things would have been worse. However, there is little evidence to support such claims. To date, regulation appears to have had little effect in improving safety or the quality of the environment.

In the case of OSHA safety regulation, there has been no significant improvement in national rates of on-the-job deaths and injuries since the agency has been operating. Nor is there any evidence that in the absence of OSHA standard setting things would have been worse. OSHA's inability to improve the national record of on-the-job injuries has been blamed on an insufficiency of agency resources to inspect and to enforce standards. But a study of OSHA's own target industry program, in which the agency targeted extensive additional resources on inspection and enforcement of standards in a few industries, found that injury rates were insignificantly different in the target industries and in other industries where only a small percentage of firms were inspected.[6] Further analysis along those lines, this time of differences among firms in rates of injury involving lost workdays, indicates that OSHA inspection or lack of it makes no significant difference in explaining injury rate differences.[7] OSHA is the agency with the largest number of regulations covering specific items of equipment; yet it has done nothing to affect the rate of injuries that regulation was intended to prevent.

---

[6] Robert Stewart Smith, "The Occupational Safety and Health Act (Washington, D.C.: American Enterprise Institute for Public Policy Research, 1976), Appendix C.

[7] Aldona DiPietro, "An Analysis of the OSHA Inspection Program in Manufacturing Industries, 1972–1973." Draft Technical Analysis Paper, U.S. Department of Labor (August 1976).

In the case of NHTSA, the details of what has happened are different, but the conclusion is generally the same. It is true that highway death rates have fallen substantially in the period since NHTSA began setting standards, and there is evidence that the safety devices mandated for new cars between 1966 and 1970 have reduced rates of death and serious injury for drivers and other occupants of automobiles involved in accidents. On the basis of these observations, NHTSA has claimed success in reducing highway death and injuries; however, a deeper probing suggests that this agency has been little more effective than OSHA. The decline in the overall highway death rate since 1966 appears to be explained fully by factors other than required changes in the safety equipment on automobiles—factors that include a reduction in the proportion of young drivers in the total driving population and lower vehicle speeds. The reduced fatality rate per accident among drivers protected by the new safety equipment, which should have led to a reduction in the overall highway death rate, was offset by an increase in the total number of accidents per vehicle mile—an increase that seems to have been due to greater risk taking by drivers of the more "crash-proof" cars.[8]

Moreover, even where NHTSA standards do appear to have had beneficial results—in improving crash-survivability for drivers—their effectiveness may have been confined to NHTSA's early years. A 1976 study by the General Accounting Office found that, while the required safety equipment added to new cars between 1966 and 1970 seemed significantly to reduce the risk of death and serious injury to the occupants of cars involved in accidents, the required safety design features added after 1970 have not appeared to bring about a further reduction in risk.[9]

The record of standard setting does not look a great deal better for environmental regulation. There have been some significant reductions since 1969 in emissions of air pollutants per unit of industrial output and per automobile mile traveled, but for the most part these appear to have occurred independent of EPA regulation and on the whole have not resulted in substantial nationwide improvement of air quality. Of the five major air pollutants for which EPA publishes nationwide emissions trends, only one—suspended particulates—showed substantially lower emissions levels in 1977 than in 1970, and two—nitrogen oxides and carbon monoxide—showed higher levels. Emissions levels for the four pollutants other than particulates were higher in 1977 than in 1973.[10]

The picture looks considerably better if pollution levels are adjusted for changes in economic activity. For the four pollutants other than particulates, reductions in emissions per unit of real GNP between 1970 and 1977 ranged from five to 30 percent. There are cases where emissions per unit of output in individual industries have been reduced by fairly substantial margins. For example, emissions of sulfur oxides per unit of output in metal processing dropped by about

---

[8] Sam Peltzman, "Regulation of Automobile Safety," (Washington, D.C.: American Enterprise Institute for Public Policy Research, 1975).

[9] The Comptroller General of the United States, "Effectiveness, Benefits, and Costs of Federal Safety Standards for Protection of Passenger Car Occupants." Report to the Committee on Commerce, United States Senate. CED-76-121, July 7, 1976.

[10] U.S. Environmental Protection Agency, "National Air Quality, Monitoring, and Emissions Trends Report, 1977." EPA-450/2-78-052, December, 1978.

half between 1970 and 1977 and in electric utility power generation, by about 20 percent. The decrease in emissions of volatile organic compounds and carbon monoxide per motor vehicle mile also has been substantial, as has the reduction, by more than one-third, in emissions of volatile organic compounds per unit of output in the chemical industry.

However, the fact that emissions of air pollutants may have decreased since 1970 says nothing of the role that EPA standards may or may not have played. Standards covering all pollutants and all industries have not been in force during the whole period since the passage of the Clean Air Amendments of 1970, so that, obviously, not all changes in the period can be credited to regulation. Just as obviously, other factors have also been at work reducing pollution levels. In fact, regulation appears to have been effective only in reducing automobile and electric utility emissions. In the chemical, primary metals, and petroleum refining industries, EPA standards appear to have had no significant effect. In these industries, changes in pollution levels per unit of output have been due to trends in technology that were already evident before the advent of EPA regulation. Moreover, regulation appears to have had no significant effect on emissions levels for industry as a whole. The reductions that have occurred since 1969 are, again, explainable by technology and the business cycle rather than regulation.[11]

A New Era of Effectiveness?

These first discouraging findings are certainly not voluminous enough to warrant a declaration that EPA, OSHA, and NHTSA standard setting has been a failure. It is possible that we have not waited long enough and that evidence of benefits will show up in due course. However, there are probably stronger grounds for the more pessimistic judgment that regulation as now practiced by these agencies is incapable of generating the improvements in safety and environmental quality that were the objectives of legislation.

In the case of safety regulation, the suspicion is that the present mode of standard setting does not begin to come to grips with the complex causes of either occupational or highway accidents. Industrial accidents have multiple and subtle causes that have little to do with the position of fire extinguishers or the size and shape of doors. Safety specialists estimate that only 20 to 30 percent of industrial accidents have been due to identifiable physical hazards susceptible to inspection. Thus, even if OSHA were capable of writing optimum standards for safe physical conditions at workplaces and of making sure that these standards were observed everywhere at all times, 70 to 80 percent of all industrial accidents would still lie beyond reach. In the case of motor vehicle standards, the relationship between equipment conditions and safety is even more complex. Regulation itself may have induced behavioral changes that offset the beneficial effects of new safety design features—by lowering the risk of death or serious injury in the event of an accident, the improved safety features of automobiles may have encouraged drivers to take more risks. To the

[11] MacAvoy, op. cit., Chapter 3.

extent that this has led to more accidents, it suggests that even the best equipment standards may be incapable of reducing harm.

Stricter and more widespread enforcement of standards probably could have produced larger reductions in air pollution than have been realized to date, although it would almost certainly have been at much higher cost. It is not coincidental that the only two places where regulation has affected pollution loads—stationary electric power sources and automobile emissions—are the only two where the federal EPA has enforcement powers. Elsewhere, EPA sets standards and leaves enforcement to state environmental agencies. EPA's practice has been to set uniform standards on the basis of what is achievable with current or prospective technology, and the state agencies have pursued flexible enforcement policies, allowing variations when affected companies have been able to argue that the cost of compliance would be so high as to constitute confiscation of assets. Since the cost of compliance tends to be highest where harm to the environment is greatest, the inability of agencies to enforce standards where the cost to individual companies is very high raises the question whether regulation can produce substantial improvements in environmental quality.

In recent years, there have been major shifts in emphasis in both safety and environmental programs that might give rise to hopes for a new era of effectiveness. OSHA has sharply shifted its emphasis in the past three years from industrial safety to health regulation and, as a consequence, is moving away from preoccupation with the detailed physical conditions of workplaces. However, the problem of regulating exposure to harmful substances in the workplace is very much akin to that of regulating exposure to harmful substances in the ambient air. In health regulation, OSHA is likely to follow EPA in setting uniform standards for all industries and types of plants. Where the harm is greatest, the cost of compliance with these standards is likely to be so high as to make enforcement politically difficult. Yet it is only there that standards would be effective in changing environmental or health conditions.

Thus, there are two problems that seem to stand in the way of effective regulation: first, standards that turn out to be unrelated to major causes of harm and, second, standards that cannot be enforced where the resulting cost would be high and reduction of harm greatest. This does not mean that there may not be quality-of-life improvements in the future or that regulatory agencies may not claim these gains as evidence of their effectiveness. On-the-job injury rates may turn down in the 1980s because of the increasing average age of the labor force, and improvements in air and water quality may occur because changes in demand or in the prices of materials and energy lead to shifts in technology or product mix that reduce pollution. However, it seems predictable that the process of administering health and safety and environmental regulation will be so poorly directed and so full of gaps that little more can be gained from using controls in the 1980s than has been gained in the 1970s.

Prospects for Reform

Although agency controls have been applied to very different kinds of industries for very different purposes, there has been a sameness in

the results. Controls in the 1970s have reduced investment and output growth in all the regulated industries. In the case of the utilities and transportation industries, regulatory bodies have brought on these results by failing to allow prices to keep up with rising costs and failing to allow services offered to reflect changing conditions of demand. In the case of the industries that are the major targets of EPA, OSHA, and NHTSA standards, the results have come about through cost increases, price increases, and reductions in investment other than that called for to meet regulatory equipment requirements. In neither case have there been offsetting benefits. The reduced ability of the electric and gas utilities and the railroads to provide service is a negative benefit to the consumer. Higher prices and reduced output due to safety and environmental regulations have apparently been accompanied by few if any quality-of-life improvements.

Thus, industries which together produce close to one-fourth of GNP now operate under agency controls that reduce production without generating substantial benefits. Improvement of this state of affairs unfortunately lies beyond the reach of most of what is being proposed and discussed under the title of "regulatory reform." Reform that focuses on administrative processes and the management of agency rulemaking has little potential for restoring profitability and growth to the regulated sector of the economy.

There is no single formula for restoring growth to all regulated industries. In the case of the electric and gas utilities, a rate-setting process should be developed that takes into account only current and future costs of providing service, including capital costs. The most urgent reform in regulation of the railroads and the telephone and petroleum companies is to allow them greater freedom—along the lines now available to the airlines—to respond to changing market conditions. In safety and environmental quality regulation, what is most needed is a process that generates only cost-effective, beneficial regulation.

Whether or not one calls such measures "deregulation" is a matter of preference, although all would involve a substantial lessening of agency controls on the regulated industries. Deregulation is understood to mean removal of the agencies' statutory authority to set prices and standards, and this is not necessary in all cases to improve the performance of the industries. In recent years, a fair amount of salutary reform has come from regulatory bodies themselves, some of it involving changes in procedures that come close, in fact, if not in law, to deregulation. In an effort to cope with their heavier case loads and alleviate industry problems caused by inflation, state public utility commissions have experimented with a variety of new practices. In a number of states, the commissions have allowed requested price increases to go into effect immediately, pending the outcome of hearings and subject to later repayment of overcharges. On occasion, commissions also have allowed companies to project costs and revenues on the basis of future, rather than past test years. This represents quite a radical departure from the traditional regulatory practice of basing prices on "objective" data and known (past) conditions and comes sufficiently close to market pricing to make one suspect that it is *de facto* deregulation.

In the airline industry, too, the first deregulatory steps were taken by the regulators. When attempts at legislative deregulation during

the Ford Administration failed, the CAB itself began to institute reforms, the most important being to allow experimentation with discount fares. Permitted greater flexibility in pricing, the airlines were able to reduce fares substantially on some heavily traveled routes, take advantage of rapidly rising demand, and increase overall profitability. The CAB's experiment in the two years of internal reform that preceded legislative deregulation proved to reduce fares and increase industry returns—a set of results not likely from the same experiments in other regulated industries.

Natural gas is still another case where a regulatory agency moved more quickly than Congress to deal with problems caused by regulation. In 1976, while Congress was bogged down in controversy over the details of phasing out fixed price controls on gas, the FPC conducted a comprehensive rate proceeding that resulted in roughly a 100 percent increase in new contract prices. Since then—with unregulated intrastate prices at the same time increasing by equal or somewhat greater margins—there has been substantial new exploration and development and a slowing in the growth of demand, both working to reduce the capacity shortages that had built up by the early 1970s. The method by which the FPC set new prices in 1976 is noteworthy because that body, too, abandoned enough of the substance of regulation to be credited with having in good part deregulated. Downplaying historical average costs, the FPC instead estimated the current cost of obtaining additional supplies from unregulated sources. Its estimates of marginal costs were based on contract practices for new reserves in the unregulated intrastate market. The commission in effect set regulated prices on the basis of unregulated prices.

That regulatory bodies themselves have been the source of what has amounted to some important deregulation is surprising if one pictures these agencies as entirely inflexible, unresponsive bureaucracies. In fact, the older federal agencies and state utility commissions have considerable incentive to manage their case loads efficiently and keep the industries they regulate performing well, because these seem to be the criteria, by and large, by which Congress and the state legislatures have evaluated the agencies' own performance in the long run. (This is the case even when agencies have been given high marks in the short run for rejecting proposed rate increases.) Moreover, since the regulatory process is in large measure the creature of the agencies themselves, they have considerable latitude within their existing statutory authority to change established rate-setting procedures when these procedures stand in the way of agency objectives.

However, one should not overstate what can be expected in the way of significant reform from the regulatory bodies by their own case and rule making. Commissions vary considerably in their commitment to deregulation and their concern for industry problems. How the courts will respond to new agency procedures is always unpredictable. And, in many cases the agencies simply cannot do what is needed to improve industry performance without new statutory authority. Binding legislative requirements for holding down gas or oil price increases cannot easily be relaxed by federal commissions. It seems doubtful as well that the kinds of reform needed to make health and safety and environmental quality regulation cost-effective and beneficial can come about without new legislation. EPA, OSHA, and NHTSA do

not have the same stake in the performance of the industries they regulate as the older regulatory bodies, and on that account alone they are not likely to be the source of innovations designed to mitigate the effects of current practices.

The major initiative to reform health and safety and environmental regulation by using cost-benefit analysis in the formulation of standards has come from the White House and has been founded on concern for the impact of controls on the general price level. Cost-benefit analysis done at the behest of the Executive Branch may weed out the extreme cases of new industry health, safety, and environmental standards that are very costly and clearly nonbeneficial. However, it is likely to fall short of producing the result that is wanted: regulation where and only where controls can improve upon on-going private operations. EPA, OSHA, and NHTSA do not now have the statutory authority to balance safety or environmental goals with conflicting economic goals or to choose the most effective method of regulating if that method is anything other than standard setting. Health and safety and environmental regulation is not likely to be any less costly or more beneficial than it has been up to now until there is new legislation that allows the agencies not to regulate in cases where regulation promises to produce results no better than the market does and unless the agencies are given authority to use taxes or subsidies in lieu of standards when the former promise to be more cost-effective.

Prospects for the sort of legislation that could restore profitability and growth to the regulated sector do not seem very good at present. In the case of most industries operating under price regulation, decontrol would mean price increases, either for most consumers or for certain groups that now enjoy subsidized service. Reforms to make health and safety and environmental regulation cost-effective and beneficial would mean less regulation than now, and this would have to be explained in one of two ways: that the regulatory experience to date had been a failure, or that some of the established objectives of regulation had to be abandoned. Either way, there would be enormous personal and political capital to be lost, and legislative resistance would probably be strong.

Then too, deregulation of this kind is, for both legislators and industry, a leap into the unknown. It may be agreed that existing regulation has not worked and that industry could perform better without controls, but the transitional effects of decontrol are hard to predict. The success of legislation deregulating the airlines appears to have been due in good part to there having been prior experimentation with deregulation by the CAB and substantial analysis by the Department of Transportation of the potential problems of transition from regulation to deregulation. The results in both cases helped to reassure Congress and the industry that there would not be severe disruptions of service or market structure in the wake of deregulation. Similar experiments and analysis have not taken place in the other industries that would be prime candidates for reduced controls.

There is another barrier to reform of health and safety and environmental regulation—that Congress would be called upon to delegate power to independent agencies or agencies in the Executive Branch to set priorities among competing social and economic goals.

This would give the agencies power that subcommittees in Congress themselves do not have to trade off gains and losses in production and price stability against safety gains and losses, and it might well give away power to the President to make these trade-offs. Such changes might work for the better in the economy, but, given their political implications, they seem unlikely to come about.

In the end, however, one cannot help feeling that the prospects for lessening controls on industry would be a great deal better if the case for it were better argued. Making a persuasive case for decontrol means talking less about inflation, paperwork, and big government and instead focusing on what controls on industry are really doing. What they are doing is causing a number of major industries to perform poorly. What they are not doing is producing the promised social benefits. These are things that business leaders may not feel comfortable saying—the less so, the greater their concern for getting along with government. However, they are the two halves of the case for deregulation, and they are what the public at large is likely to care about.

COMMON CARRIER REGULATION AND TECHNOLOGICAL CHANGE: THE NEW COMPETITION IN THE COMMUNICATIONS INDUSTRIES

By Nina W. Cornell and Douglas W. Webbink*

CONTENTS

	Page
Summary	197
I. Introduction: Economic theories of regulation	199
Organization of the paper	200
II. Natural monopolies	201
The regulatory response to natural monopolies	202
Natural monopoly in common carrier telephone communications	202
The impact of future changes on natural monopolies in common carrier communications	206
Regulation of common carrier communications as a natural monopoly	207
Rate of return regulation under conditions of static technology	208
Rate of return regulation under conditions of technological change	210
Rate of return regulation of competitors	212
III. Externalities	213
The universal service argument	214
Technical harm to the network argument	215
Technological change and externalities	216
IV. Income redistribution and internal cross subsidies	216
Internal subsidies are received by persons other than the intended recipients	217
Internal subsidies give the wrong signals to potential competitors	218
Internal subsidies hide what is really happening	218
Internal subsidies are less efficient than direct cash payments	219
V. Conclusions	220
Common carrier deregulation and antitrust enforcement	220
The macroeconomic impact on the economy of the rate of return regulation under conditions of technological change	222
Externalities	223
Income redistribution	223

Summary

This paper discusses three major classes of arguments for regulating telephone common carriers: (1) Certain telephone services are said to exhibit natural monopoly characteristics; (2) there are said to be significant externalities; and (3) regulation may be used as a

---

*Chief Cornell and Deputy Chief Webbink, Office of Plans and Policy, Federal Communications Commission. The views expressed are those of the authors and do not necessarily reflect the views of the Commission. It is not the intention of the authors to draw conclusions about the merits of any specific action of any specific company, but rather to discuss general tendencies concerning the interaction of regulation, competition, and technological change by communications firms. The authors wish to thank Stanley Besen, Thomas Casey, Jerry Duvall, Paul Fox, William Ginsberg, Peter Greenhalgh, Daniel Kelley, and Jeffrey Krauss for helpful comments on an earlier version of this paper.

device to redistribute income. This paper discusses the impact of rapid technological change on each of these arguments and the need for traditional common carrier regulation such as regulation of rate of return, specific rates, and entry of firms into the industry.

The primary reason for regulation of telephone companies is the assumption that at least some of their services have natural monopoly characteristics. However—even in a period of static technology—terminal equipment, interexchange transmission, local transmission and switching obviously exhibit different characteristics with regard to economies of scale and the possibility of being natural monopolies. When technology is static, regulation has a number of undesirable effects: It may create incentives for firms to shift assets and expenses to different parts of the firm in inefficient ways, as well as to over capitalize and even to expand all expenses. Moreover, regulatory agencies are even more concerned about bankruptcy than excess profits, even though bankruptcy may increase economic efficiency. Finally, regulatory agencies are backward looking, when firms need to be forward looking.

With technological change, the impact of regulation may be even more adverse. Natural monopoly assumptions may no longer be valid. Moreover, regulation may slow the rate of depreciation and hence the rate of introduction of new technology. Regulation may also make prices less flexible, and will surely slow the rate of entry and exit into new services. When markets with a competitive structure are regulated, legal arguments and procedures will replace competition in the market place.

A second set of reasons for regulating telephone services concerns externalities. Some argue that there are externalities from ensuring that each subscriber can reach every other subscriber, of knowing that every individual can get help in an emergency, and from preventing harm to the telephone network. It is unclear, however, that there are substantial external benefits that are not taken into account and paid for by subscribers from knowing that one can reach all persons.

While there may be real external benefits in ensuring emergency help to all individuals, such an externality argument only suggests a need to provide access to the telephone network; it does not imply that people should get telephone *usage* below direct costs. Moreover, there are also growing substitutes for wireline telephone service such as citizens band radio.

The most convincing argument for the existence of an externality is that users of telephone equipment are capable of doing harm to the telephone network and other telephone users. Such an argument does not imply a need for traditional rate of return regulation, however, but only a need for certification or type approval or type acceptance of equipment, and the installation of circuit breaker equipment on each subscriber's premises.

The third argument that has been raised in favor of common carrier regulation is that regulation can be used to redistribute income. For example, wealthy people could cross subsidize poor people, urban areas could cross subsidize rural areas and business users could cross subsidize residential users. However, there are many problems with cross subsidies. Much of the subsidy goes to persons other than those

to whom it is intended. Subsidies give the wrong market signals to potential entrants and innovators. Internal subsidies also hide from the public the identities of gainers and losers, and the true size of the subsidy. Finally, direct payments and direct taxes are bound to be more efficient.

For all of these reasons this paper concludes that traditional common carrier regulation should be used as little as possible. First, competitively structured markets should not be regulated at all. Second, common carrier regulation may be especially costly to society in a period of rapid technological change. Third, traditional common carrier regulation is not an appropriate way to deal with externalities or income distribution.

Nevertheless, there may still be a need for some minimal government intervention. Certification, type acceptance, and type approval of equipment may be a legitimate government function. So too are the use of taxes and direct subsidies to redistribute income to certain groups. Finally, with less regulation there may well be a need for increased antitrust enforcement. There may also be a need to be more concerned about the relationships between subsidiaries of companies that operate partly in unregulated competitive markets and partly in regulated monopoly markets.

I. Introduction: Economic Theories of Regulation

In the past decade, economists have scrutinized and theorized about the reasons, goals, and results of Government regulation of economic activity.[1] These economists have attempted to explain why economic activity has been or should be regulated. There is widespread but by no means universal agreement among economists that regulation may be appropriate when the unregulated market fails to produce certain desired outcomes and when the benefits of regulation exceed the costs.

---

[1] See for example: Bruce M. Owen and Ronald Braeutigam, "The Regulation Game: Strategic Use of the Administrative Process" (Cambridge, Mass.: Ballinger Publishing Co., 1978); Paul L. Joskow and Roger G. Noll, "Regulation in Theory and Practice: An Overview," Social Science working paper number 213, California Institute of Technology, Division of the Humanities and Social Sciences, May 1978; Sam Peltzman, "Toward a More General Theory of Regulation," Journal of Law and Economics, 19 (August 1976), pp. 211–240; Richard A. Posner, "Theories of Economic Regulation, "Bell Journal of Economics and Management Science, 5 (autumn 1974) pp. 335–358; Roger G. Noll, "The Behavior of Regulatory Agencies," Review of Social Economy, 29 (March 1971). pp. 15–19; Richard A. Posner, "Taxation by Regulation," The Bell Journal of Economics and Management Science, 2 (spring 1971), pp. 22–50; and George J. Stigler. "The Theory of Economic Regulation" Bell Journal of Economics and Management Science, 2 (spring 1971. pp. 3–21). On the impact of regulation, see for example; Paul W. MacAvoy and John W. Snow, editors, "Regulation of Entry and Pricing in Truck Transportation" (Washington, D.C.: American Enterprise Institute, 1977); Paul W. MacAvoy and John W. Snow, "Regulation of Passenger Fares and Competition Among the Airlines" (Washington, D.C.: American Enterprise Institute, 1977); Paul W. MacAvoy and John W. Snow, editors, "Railroad Revitalization and Regulatory Reform" (Washington, D.C.: American Enterprise Institute, 1977); Sam Peltzman, "Regulation of Automobile Safety" (Washington. D.C.: American Enterprise Institute, 1975); Stephen G. Breyer and Paul W. MacAvoy, "Energy Regulation by the Federal Power Commission" (Washington, D.C.; The Brookings Institution. 1974); George W. Douglas and James C. Miller III, "Economic Regulation of Domestic Air Transport: Theory and Practice" (Washington, D.C.: The Brookings Institution. 1974); Roger G. Noll, Merton J. Peck and John J. McGowan, "Economic Aspects of Television Regulation" (Washington, D.C.: The Brookings Institution, 1973); Thomas Gale Moore. "Freight Transportation Regulation: Surface Freight and the Interstate Commerce Commission" (Washington, D.C.: The American Enterprise Institute, 1972).

On the question of changing and improving regulation, see: Roger G. Noll, "Breaking Out of the Regulatory Dilemma: Alternatives to the Sterile Choice," Indiana Law Journal, 51 (spring 1976), pp. 686–699, and Roger G. Noll, "Reforming Regulation" (Washington, D.C.: The Brookings Institution, 1971).

Other important economic tools such as taxes and subsidies have also been suggested to correct market failures but this paper is primarily concerned with direct Government regulation.

Economists have sought regulation when the market otherwise would: (1) be monopolistic because the technology of production is most cheaply done by only a single firm; (2) impose costs or benefits on people who do not buy or sell the product; or (3) charge different classes of customers prices higher or lower than those society wants to have charged.[2] To put these three situations into the language of economics, the three situations have led some economists to urge regulation in cases where the industry is likely to be a natural monopoly, cases where there are significant externalities, and cases where market prices run counter to societal goals for income distribution.

There are also two different kinds of regulation used in these circumstances: (1) Price and entry (and exit) controls; and (2) performance requirements. The first kind of regulation includes tariff regulations and equipment construction or installation authorizations required of common carriers under Title II of the Commission's Act.[3] The second kind of regulation includes such programs as the Federal Communications Commission's type acceptance and type approval programs for equipment.[4] Standards are set for equipment, but no controls are set on who may buy or sell it or on the price of the equipment.

In the past, economists saw price and entry controls as the appropriate type of regulation when there was a natural monopoly. In contrast, when regulation is appropriate at all to correct for externalities, economists argue for performance standards.[5] Some non-economists, however, have argued for price and entry controls in this latter case, as well as cases where they wish to use prices to redistribute income.

Organization of the Paper

Each of these three arguments has been used for adopting or continuing price and entry control regulation of communications, particularly common carrier communications. This paper will look at each of these rationales for this type of regulation of communications and at the consequences of such regulation in a period of technological change, using examples drawn from common carrier communications.

Part II of the paper discusses regulation of natural monopolies and the impact of changing technology, growing demand and increased competition on the need for traditional common carrier regulation. Part III discusses the appropriate government response to externali-

---

[2] There are, of course, other economic theories of regulation represented in the literature cited above. Included in those theories are the ideas that regulatory agencies may seek to avoid conflict, or have been captured by some part of the regulated industry. The paper will not discuss those theories in greater detail.

[3] 47 U.S.C.A. 201–22. See also: 47 CFR 61 and 63.

[4] See 47 CFR 15.1–15.423.

[5] See: Nina W. Cornell, Roger G. Noll and Barry Weingast, "Safety Regulation," Chapter 11 of "Setting National Priorities: The Next Ten Years," edited by Henry Owen and Charles L. Schultze (Washington, D.C.: The Brookings Institution, 1976), pp. 457–504. For a different view, see: Robert Steward Smith, "The Occupational Safety and Health Act: Its Goa's and Its Achievements" (Washington, D.C.: American Enterprise Institute, 1976). See also: Anthony C. Fisher and Frederick M. Peterson, "The Environment in Economics: A Survey," Journal of Economic Literature, 14 (March 1976), pp. 1–33.

ties, with several examples taken from telephone regulation. Part IV describes the impact of using common carrier regulation as a method for redistributing income or wealth. Finally, Part V provides the conclusions.

II. Natural Monopolies

When economists say that firms or particular services or even whole industries tend to be natural monopolies, they are suggesting that there are such large economies of scale or economies of non-duplication that a single firm would be able to offer a particular service or services in a particular area at a lower real average cost than could two or more competing firms. Thus, it has been assumed that if two or more firms attempted to compete in providing local telephone service they might duplicate each other's facilities and therefore operate at such a small size that each would have higher average costs than a single larger company.[6] If it is true that any service or industry is a natural monopoly, it follows that companies will find it profitable to merge or share services through joint ventures or other arrangements unless they are prevented from doing so.

The determination that a particular industry is or is not a natural monopoly rests on two considerations: The specific technology used to produce the output of the industry; and the size of the market potentially available for that output. If the technology exhibits decreasing average costs over the long run for a single firm as the quantity of output rises, it may be inefficient to divide production between or among suppliers[7]. This result depends on the level of output that can be produced at the minimum average cost being roughly equal to or greater than the size of the potential market. If the size of the potential market expands significantly beyond the quantity that a single firm can produce at minimum average cost, the industry is no longer a natural monopoly.

For an industry to remain a natural monopoly over time, the size of the potential market must remain roughly equal to or below the level of output corresponding to the minimum average cost with any new technology. As these two factors are unlikely to change in the same way over time, the concept of a natural monopoly is in fact more useful in analyzing an industry with little change than when analyzing an industry experiencing rapid change.[8]

---

[6] Paul J. Garfield and Wallace F. Lovejoy, "Public Utility Economics," (Englewood Cliffs: Prentice Hall, Inc., 1964), pp. 15–27, 44–259.

[7] The notions of economies of scale and decreasing average cost are frequently (and incorrectly) blurred. Although a techno'ogy characterized by economies of scale will exhibit declining average costs (prices assumed fixed), there is no simple relationship between the two concepts. See: W. Ginsberg, "The Multiplant Firm with Increasing Returns to Scale," Journal of Economic Theory, 9 (November 1974), pp. 283–292, and references cited therein.

[8] There is a growing economic literature on whether economic welfare would be increased if a natural monopoly firm that is unsustainable in the face of new entry should be protected against that potential entry. See, for example: Robert D. Willig, "Multiproduct Technology and Market Structure," *American Economic Review Papers and Proceedings*, 69 (May 1979), pp. 346–351; John C. Panzar and Robert D. Willig, "Free Entry and the Sustainability of Natural Monopoly," *The Bell Journal of Economics*, 8 (Spring 1977), pp. 1–22; William J. Baumol, "On the Proper Cost Tests for Natural Monopoly in a Multiproduct Industry," *American Economic Review*, 67 (December 1977), pp. 809–822; William J. Baumol, Elizabeth E. Bailey, and Robert D. Willig, "Weak Invisible Hand Theorems on the Sustainability of Multiproduct Natural Monopoly," *American Economic Review*, 67 (June 1977), pp. 350–365; and Gerald R. Faulhaber, "Cross-Subsidization: Pricing in Public Enterprises," *American Economic Review*, 65 (December 1975), pp. 966–977.

The authors of the paper remain unimpressed about the policy relevance of this literature. First, the sustainability analysis is comparative-static analysis; it does not consider

(Continued)

If a firm is a natural monopoly or if it is given monopoly status by law or regulation, it has market power. This means that it will have the capability to set the price of its services above marginal (and average) costs.[9] Whenever the price exceeds the costs of providing the service, society loses some output that it values more than it values other outputs that in fact get produced.[10] Thus, the failure of a monopoly market to maintain prices equal to the costs of producing the last unit of output results in economic inefficiency.[11]

The Regulatory Response to Natural Monopolies

To try to prevent such economic inefficiencies, regulatory agencies attempt to make monopoly firms earn no more profits than they would if they were operating in a competitive market. Thus, regulatory agencies attempt to determine what is a "fair" or "equitable" or "competitive" rate of return on assets. They then strive to prevent a firm from setting prices so as to exceed that rate of return. To do this, they require the firm to file its prices, or tariffs, with the agency. The agency is then supposed to check those prices against the firm's costs to ensure that no more than the fair return is being earned.[12]

Along with tariff requirements, the regulatory agency also requires advanced approval for all installations of facilities by the firm. This process is supposed to ensure that the firm does not overconstruct the system and thereby unnecessarily increase the rate base. Prior approval over facilities also effectively controls whether other firms can enter the industry.

Natural Monopoly in Common Carrier Telephone Communications

The common carrier telecommunications industry provides three interrelated but distinct services: Communications between two or more local exchange switching areas (interexchange communications)[13]; communications within local switching areas (the "local loop"); and terminal equipment. These three functions are obviously interrelated. Communications between two subscribers may be routed through one or many switching centers. It is more economical to use

---

(Continued)

the impact of competitive entry on the rate of technological change. Second, even in the static case, the welfare economics implications are ambiguous because under certain conditions consumer welfare may increase rather than decrease when entry takes place in a natural monopoly market. Finally, it seems doubtful whether the information is available or ever will be available to determine if any real world firm fits the sustainability example. See: Vinson Snowberger, "Government Preservation of A Regulated Monopoly," *Quarterly Review of Economics and Business*, 18 (Winter 1978), pp. 81–89.

[9] Marginal cost means the cost of producing an additional unit of output.

[10] Whenever price is higher than marginal cost, lowering the price would increase the amount produced and sold. In that process, additional resources are used to provide the service. The price consumers would pay is greater than the value of the additional resources used. Thus, when those resources are used to produce some other output instead, society has given up output that it values more for output that it values less.

[11] For a discussion of the problems with using marginal cost pricing, see: R. H. Coase, "The Theory of Public Utility Pricing and Its Application," *The Bell Journal of Economics and Management Science*, 1 (Spring 1970), pp. 113–128

[12] Of course the process is really more complicated than this. For example, regulatory agencies are more likely to scrutinize requests for rate increases than rate decreases, regardless of the impact of the change on a firm's rate of return. See: Paul C. Joskow, "Pricing Decisions of Regulated Firms: A Behavioral Approach," *The Bell Journal of Economics and Management Science*, 4 (Spring 1973), pp. 118–140.

[13] Interexchange communications is sometimes called intercity transmission, even though it may take place between two centers located only a few miles apart or several thousand miles apart.

centralized switching than to connect directly each terminal with every other terminal when there are many subscribers. The choice of where to put the switches (the hierarchy of switching centers) determines the number of switches a local or a long distance call goes through. The tradeoffs made in the design of the hierarchy of switching centers have important implications for the cost of providing telephone service, the question of whether costs are charged against interstate or intrastate rates, and the ability of potential entrants to compete in providing communication services.

Despite these interrelationships, the three segments have been separated by both the firms that provide them and the agencies that regulate them.[14] Nevertheless, under the 1934 Communications Act and various state regulations all three segments have been subjected to price and entry control provisions.[15]

The Federal Communications Commission is attempting to determine in a number of proceedings the extent to which each of these segments today is or is not a natural monopoly.[16] Since the FCC is still receiving evidence on the existence of economies of scale, no definitive conclusions can currently be drawn. The desire of firms not now providing communications services to enter and compete with the established carriers, however, offers some suggestions about the nature of the industry.[17] The firms that seem to be aiming at competing with a wide array of currently provided services or equipment indicate that they do not see those markets exhibiting such economies of scale that one firm could most cheaply supply the market.[18] It may

---

[14] "Historically, Bell System Operations have been organized along 'functional lines.' In 1978, we restructured our organization along with the lines of the market sectors we serve [viz., business, residence, network]." American Telephone and Telegraph Co., *1978 Annual Report*, p. 12. Despite this restructuring the Bell System insists that, "the exchange and interexchange portions of the core network should continue to be planned, owned and operated as a single entity. . . ." ("The Telecommunications Network and Universal Service" p. 2, a paper provided to the staff of the Senate Subcommittee on Communications.) There has been no evidence that the coordination necessary among the different services requires either a single network or unified ownership. (Indeed, there are currently over 1600 firms in the "network partnership"). Allowing entry and ensuring fair but free competition would provide a market test of this proposition. Should the Bell System be correct, a single unified network would result.

[15] 47 U.S.C.A. (201–222. Some of this regulation, however, has occurred at the state level rather than at the Federal level.

[16] See, for example: *Report and Order* in the Matter of Allocation of Frequencies in the Bands Above 890 Mc., Docket 11866, 27 FCC 359 (1959); *Notice of Inquiry and Proposed Rulemaking* in the Matter of MTS and WATS Market Structure, Docket 78–72, released March 3, 1978; *Supplemental Notice of Inquiry and Proposed Rulemaking* in the Matter of MTS and WATS Market Structure, Docket 78–72, released August 30, 1979; *First Report* in the Matter of Economic Implications and Interrelationships Arising from Policies and Practices Relating to Customer Interconnection, Jurisdictional Separations and Rate Structures, Docket 20003, 61 FCC 2d 766 (1976).

[17] On specialized common carriers, see: *Report and Order* in Docket 11866 27 FCC 359 (1959: on telephone instrument registration see: *First Report and Order* in Docket 19528, 56 FCC 2d 593 (1975). *Second Report and Order* in Docket 19528, 58 FCC 2d 736 (1976), *Aff'd sub. nom,* North Carolina Utilities Comm'n v. FCC. 522 F2d 1036 (4th Cir., 1977). *cert. denied,* 434 U.S. 874 (1977); *Third Report and Order* in Docket 19528, 67 FCC 2d 1255 (1978). On satellite competition see: *Second Report and Order,* In the Matter of Establishment of Domestic Communications Satellite Facilities by Non-Government Entities, Docket 16495, 35 FCC 2d 844 (1972). See also: Xerox Corporation Petition for Rulemaking in re Amendment of Parts 2, 21, 87, 89 and 91 of the Rules for the establishment of a new common carrier Electronic Message Service (EMS) in the band 10.55–10.68 GHz, November 16, 1978; and U.S. v. FCC, et al., No. 77–1249, (D.C. Cir., August 29, 1978), ("SBS"). For a more detailed history of many of these events up to 1976, see: *Report by the Federal Communications Commission on Domestic Telecommunications Policies,* September 28, 1976.

[18] Or, as Alfred Kahn titled a section of his book, "If Competitors Want to Enter, How Natural Can Monopoly Be?" Alfred E. Kahn, *The Economics of Regulation: Principles and Institutions, Vol. II, Institutional Issues* (New York: John Wiley and Sons, Inc., 1971), p. 146. It is of course possible that even though there is a natural monopoly firm, it sets

(Continued)

be, in addition, that some firms are only entering on the fringes of the general communications market by taking advantage of economies of specialization and product differentiation, rather than responding to an overall absence of the necessary conditions for a natural monopoly.

TERMINAL EQUIPMENT

The market evidence that suggests an absence of large economies of scale is perhaps strongest in the case of the provision of terminal equipment. Regulatory rules permitting (but not requiring) customer ownership of terminal equipment have allowed firms other than phone companies to offer such equipment.[19] Competition has developed in the sale of terminal equipment to all classes of customers from the largest business users to the smallest households. Firms have entered into competition with the traditional providers in every location where there is telephone service, either by opening stores, selling to existing stores, or through catalogs.

It is still too soon to see how widespread such equipment penetration will be as the rules have not yet been fully implemented everywhere, nor have they been in existence for long.[20] But, the willingness of firms to put a significant array of terminal equipment on the market indicates that the economies of scale in manufacture of such equipment are not perceived as being significant. This conclusion is strengthened by the fact that the bulk of the equipment that is being most widely offered for residential use is virtually identical to the equipment being provided under tariff. If the largest markets were perceived as being for new, fancy terminal equipment not widely available from the carriers under tariff, there would be little if any duplication of the existing telephone types. A search of the retail catalog offerings shows otherwise.

The emerging competition in the terminal equipment market does not yet demonstrate whether the techniques by which most terminal equipment is produced have changed or are changing.[21] Therefore, there is no way of knowing whether the provision of terminal equipment once was a natural monopoly that has been eroded by technological change, or more probably whether it never was a natural monopoly in the first place.

(Continued)

price so far above marginal costs that new firms with substantially higher marginal costs decide to enter. That is not, however, the usual case. In most cases new entrants will expect price responses from the existing firm and would only enter if they believe that they will be able to provide a service at a *cost* that is equal to or below that of the existing monopoly firm.

[19] *First Report and Order* in Docket 19528, 56 FCC 593 (1975).

[20] Technically, customer provided terminal equipment has been allowed since the Carterphone decision in 1969. Until the implementation of the terminal equipment registration program, however, the Carterphone decision was largely meaningless because phone companies required use of a tariffed interconnect device, ostensibly to protect the network from technical harm. The tariffs on the interconnect devices served as an economic barrier to widespread customer provision of terminal equipment. Customer provision of terminal equipment still is slowed wherever state utility commissions have not required local telephone companies to offer a reduced tariff for service alone, as opposed to service plus the use of the phone instrument.

[21] Many new telephone instruments are being developed. For example, there is Northern Telecom's E phone, being introduced in Canada, ITT's electronic phone being installed in Denmark, and GTE's new flip phone. It is too early to tell, however, if these newer instruments will replace most existing equipment.

INTEREXCHANGE TRANSMISSION

Unlike the competition in the terminal equipment market, the growing competition in the interexchange transmission market (between switching centers) does reveal the impact of technological change on economies of scale. Originally interexchange service went by wire, and then by cable. Given adequate demand there are probably substantial economies of scale in laying one cable rather than in laying two or more of lesser capacity. The development of point-to-point terrestrial microwave facilities provided a lower cost substitute for long distance transmission by twisted pairs of wire and by coaxial cable in many cases. While the economies of scale in laying a high capacity cable rather than two or more low capacity cables may be substantial, the economies of scale may be much smaller in the case of terrestrial microwave facilities.[22] In that case, it may not be costly to have competing microwave facilities at least on many high density routes.[23]

The reduced scope of economies of scale in the newer interexchange transmission technology plus a growing market for intercity services has lured suppliers to this market to compete with the existing carriers. This competition was made possible when, in 1959, the Federal Communications Commission approved the establishment of private point-to-point microwave systems.[24] Later, in 1971, the Commission approved the development of specialized common carriers.[25]

The development of geostationary satellites marked a further step in reducing the economies of scale in transmission relative to the size of the market. Communications firms are increasingly using satellites in place of both microwave towers and coaxial and twisted pair cable. For long distance transmission, satellite systems do exhibit diminishing average cost up to capacity of the individual satellite since, once the satellite is in place, its costs can be spread over all users. Nevertheless, it may be no more costly to have multiple satellites than a single satellite, particularly if the satellites can share standby or backup facilities. Again, this lack of overwhelming economies of scale was recognized by the regulatory process when in 1972 the Federal Communications Commission decided to allow competition and multiple entry in domestic satellites.[26]

LOCAL TRANSMISSION

The process of substituting radio for wire and cable in interexchange transmission began in the late 1940's and continues to the present. Companies are now requesting permission to make the same substitution in at least some cases involving local transmission.[27] Once

---

[22] It is difficult to tell whether there are economies of scale in cables, because the higher capacity cables have always been newer cables which used a different technology than the older, lower capacity cables. It may be less costly to purchase land for microwave facilities than the right of way for cable, however, and the facilities themselves may be less costly than cable.

[23] See: Leonard Waverman, "The Regulation of Intercity Telecommunications," Chapter 7 of Promoting Competition in Regulated Markets, edited by Almarin Phillips (Washington, D.C.: The Brookings Institution, 1975), pp. 201–239.

[24] Report and Order in Docket 11866, 27 FCC 359 (1959).

[25] First Report and Order in Docket 18920, 29 FCC 2d 870 (1971), aff'd (1971); 513 F. 2d 1142 (9th Cir. 1975), *cert denied,* 423 U.S. 838 (1975).

[26] Second Report and Order in Docket 16495, 35 FCC 2d 844 (1972).

[27] See for example: Xerox XTEN Petition, November 16, 1978.

again the willingness of these companies to enter into competition with the existing carriers suggests that the would-be service providers do not see the new technologies as meeting the test of a natural monopoly. This may be either because the size of any economies of scale in radio technology are smaller than for wire line or cable, or because the size of the market has now outgrown the economies of scale in all these technologies. In either case, the would-be service providers are openly contemplating a competitive rather than a monopolistic market structure.[28]

SWITCHING

As noted earlier, both local and interexchange transmission involve both the "lines" used to connect customers and the switches that permit the company to use fewer total lines. Technological change in switches has matched the change in the technology of lines, particularly with the development of digital switches. Here, too, the size of the least cost switch may have changed as any potential economies of scale in hardware have been balanced against the potential lack of such economies in software.

DATA PROCESSING AND COMMUNICATIONS

Since 1934, technological changes have not merely altered the technology available within the separate segments of the industry. Over the past couple of decades, data processing and communications industries have been merging.[29] Today the same device that can serve as a switch can also do a variety of data processing activities. Moreover, the same terminal that can do a variety of data processing activities on the customer's premise can also serve as a communications terminal. The merging of these two industries means that intelligence in the system can be placed either inside the transmission path or inside the terminal.

The merging of the two industries has not only affected the choice of technology and therefore any economies of scale but it has also enlarged the size of the potential market for communications. The capability to store individual messages until the recipient is ready to receive them and the ability to convert messages so they are comprehensible by many different terminal devices and central computers increase significantly the ways and the amount individuals will want to communicate, whether for personal or business purposes.

The Impact of Future Changes on Natural Monopolies in Common Carrier Communications

The changes outlined above are almost certainly not the end of change in the technology of common carrier communications. There are certain to be many changes in the future that will affect the cost structure of the provision of services. For example, at the current time it is quite expensive to lay wire throughout sparsely populated rural

[28] See: Xerox XTEN Petition, November 16, 1978.
[29] See Notice of Inquiry and Proposed Rule Making in Docket 20828, adopted July 29, 1976, FCC 76-745.

areas because of the cost of stringing twisted wire cable many miles to service a single home. The development of rural radio telephone in place of copper wires could substantially reduce the costs of serving rural areas. With radio telephone the average costs of serving rural users might be similar, whether one firm served all the homes in a rural area or several competing firms each served a smaller number of homes. Thus, the importance of economies of scale might be diminished and the strength of any arguments in favor of natural monopolies for rural service might decrease.

Similarly, technical developments in citizens' band radio equipment might make that service a lower cost but complete substitute for wired telephone service. If citizens' band radios are developed that can signal the desired recipient and then lock out other users of the channel for the duration of the conversation, the resulting service would be like radio telephone but with the switching inside the terminal equipment. And, citizens' band radio equipment has been to date a sharply competitive market.

Still another change on the horizon is an increase in the rate at which messages are sent. Particularly for non-voice messages, technology that speeds up the rate of transmission will not only potentially alter any existing economies of scale, but also would greatly expand the potential market for communications.

It is also possible that technological change may go the other way in some cases: Technology may increase as well as decrease the significance of economies of scale and hence natural monopolies. For example, glass fiber using optical transmission (and switching) methods is a higher capacity communications medium than typical coaxial cable systems.[30] Fiber optics might bring into a home or office telephone, telegraph, computer data, broadcast television and radio signals all on a single cable. This might diminish, rather than increase, the possibilities for competition in the provision of the communications networks.[31]

Regulation of Common Carrier Communications as a Natural Monopoly

The previous discussion of technological changes that have already occurred, namely in terminal equipment and interexchange transmission, pointed out that over time, regulatory changes have followed technological change. Each of the major changes in the relationship between the scope of economies of scale and the size of the market ultimately has brought a regulatory change that has permitted competition and ended a presumption of monopoly. With the end of the presumption of monopoly due to the impact of technological change or even sheer growth in the size of the market, comes the question of the

[30] See, for example: E. J. Claire, C. Richard Patisaul and J. C. Wyatt, "Broadband Fiber-Optic Systems," paper presented at the IEEE International Conference on Communications, Toronto, Ontario, June 1978.

[31] While there may be economies of scale for broad band communications systems, the economies may be less important for simple telephone service. Hence, a consumer who does not want those additional services may find a competitive alternative in the coaxial system. The one other technological change that may result in regulatory movement from competition to monopoly is the development of cellular mobile radio. See: *Second Report and Order* in the matter of An Inquiry Relative to the Future Use of the Frequency Band 806–960 MHz, in Docket 18262, 46 FCC 2d 752 (1974).

appropriateness of rate of return regulation. For, if the major rationale that gave rise to rate of return regulation is gone, namely the existence of a natural monopoly, should rate of return regulation be maintained over the long run?

Rate of Return Regulation Under Conditions of Static Technology

It was noted above that rate of return regulation in theory was supposed to make the monopolist earn no more than a competitive rate of return and thus operate more like a competitive firm. It was hoped that such regulation would promote economically efficient operation of the industry. In reality no regulatory agency can ensure that a monopolist operates in an economically efficient manner.

INCENTIVES TO EVADE REGULATION

First, because regulatory agencies attempt to force firms to earn less than they are capable of earning, regulated firms have incentives to evade regulations. Thus, if some of the firm's activities are not regulated, firms may have incentives to shift revenues and profits into the unregulated part of the business, and conversely to shift costs and assets into the regulated part.[32]

Second, because the regulatory agency calculates a rate of return on assets rather than a rate of profit per unit of output, firms have incentives to expand their asset base or rate base both through extensive use of capital (the so-called Averch-Johnson effect) and by expanding the size of their business.[33] In this process, the firm is likely to use more capital-intensive techniques than those that would produce the output at lowest cost, including capitalizing any expenses that the regulatory agency will allow it to capitalize.[34]

In recognition of these incentives, regulatory agencies have been given the power to require prior authorization for facilities.[35] The difficulty with such power, however, is that regulatory agencies can almost never know the technology of production better than the firms in the industry, and no regulatory agency will be able to judge if all requested expenditures are reasonable or necessary. If there were many roughly equal-sized independent firms serving geographically different markets regulators could compare their technological judgments to try to find the least cost method of providing service. When there is only one firm, or one giant firm and several or many small ones, regulators have no market test to use.

Third, even if the firm has chosen a more capital intensive technique than might have been chosen had the industry been a competitive one, the firm may also use more labor and materials per unit of output and thus be less efficient than would have occurred in a

---

[32] See, for example: William G. Shepherd "The Competitive Margin in Communications," Chapter 4 of *Technological Change in Regulated Industries*, edited by William M. Capron (Washington, D.C.: The Brookings Institution, 1971), pp. 86–122; and Roger G. Noll, "Regulation and Computer Services," Stanford University Graduate School of Business, Research Paper No. 330, September 1976, p. 37.

[33] Harvey Averch and Leland L. Johnson, "Behavior of the Firm Under Regulatory Constraints." *American Economic Review*, 52 (December 1962), pp. 1052–1069.

[34] The American Telephone and Telegraph Company capitalizes the labor costs of installation of telephone equipment.

[35] Communications Act of 1934, 47 U.S.C.A. p. 214.

competitive situation. Under certain circumstances, regulated firms that are already earning their allowed rate of return may have little incentive to produce goods and services at least cost, since the value of any cost-saving improvement must be passed on to customers, rather than to stockholders. Similarly, any cost increasing action (such as higher management salaries) may be passed on to customers, rather than on to stockholders.[36]

The recognition of these incentives to evade regulation or at least diminish the effect of regulation has forced regulatory agencies to become embroiled in all sorts of accounting controversies such as: What is the "correct" rate of return; which revenues should or should not be counted; what expenses are "legitimate"; what assets should be included in the rate base; what is the appropriate rate of depreciation; and how joint and common costs between regulated and unregulated parts of the firm should be treated.[37] To many of these questions there is no single "correct" answer. Rather there may be many different answers all of which are equally arbitrary and equally "correct".

What is certain is that arriving at any of the possible arbitrary resolutions of each of these issues requires considerable resources both on the part of the regulatory agency and on the part of the regulated firm.[38] The costs of at least the firms' resource use also get added to the prices it charges. The problem with trying to use rate of return regulation to emulate (or simulate) the outcome of a perfectly competitive market is that regulatory agencies may be unable to *force* firms to do what competitive markets *cause* them to do in their own self interest.[39]

ECONOMIC INCENTIVES: PROFITS AND LOSSES

A central tenet of microeconomic theory is that firms react to economic incentives: They produce products and services that are or are expected to be profitable, and they cease to produce products and services that are not profitable. The attempt of regulated firms to act efficiently in this way may come into conflict with the goals of regulatory agencies. If regulated firms make large profits, even temporarily, the regulators feel they must act to lower the profits. Conversely, if firms discover that certain products and services are unprofitable, they may wish to abandon them. In an unregulated market, a firm can just shut down its unprofitable line whereas in a regulated market, the firm may not do so without permission of the regulatory agency, which may not grant it.

If there is one thing that regulatory agencies dislike as much or more than "excess profits" it is bankruptcy of regulated firms. Many regulatory decisions to prohibit new entry or new competition into an

---

[36] Harvey Leibenstein, "Allocative Efficiency Versus X-efficiency," *American Economic Review*, 56 (June 1966), pp. 392–415.

[37] See, for example: Alfred E. Kahn, *The Economics of Regulation: Principles and Institutions, Volume I: Principles*, (New York: John Wiley and Sons, 1970), pp. 20–57; and Richard A. Posner, "Natural Monopoly and Its Regulation," *Standford Law Review*, 21 (February 1969), pp. 548–643.

[38] Apparently the long lines division of AT&T spent $325,000–$350,000 on the disaggregated embedded cost study to implement Docket 18128.

[39] In addition, of course, before a regulatory agency can even *attempt* to force a firm to behave competitively, the agency must determine what would happen in a competitive market. This is probably an impossible task.

industry are based on the "need" to protect an existing firm from bankruptcy.[40] Bankruptcy serves a useful purpose in competitive markets: If some firms are more efficient than others, the economy may be better off if the inefficient ones go out of business or are acquired. Legal requirements for due process and against confiscation of property push regulatory agencies to attempt to block losses and bankruptcy.

PAST VS. FUTURE-LOOKING DECISIONS

The same concern for due process and the need to establish a record that can withstand court challenge prevents regulatory agencies from adopting pricing rules that mirror those that would prevail in a competitive market. In competitive industries firms make decisions about what to produce, what costs they believe are reasonable and what prices they should set by making estimates about the future demand for their services and products. Often prices and costs are calculated based on some estimate of future volume of sales.[41]

In contrast, regulators must use past (historical) data on costs, assets and revenues to determine if rates and the rate of return are "fair" or "reasonable." This would approximate competitive prices only if both technology and demand for the output were static. Even without technological change, if demand is growing the regulatory process is unlikely to ensure economic efficiency.[42]

Rate of Return Regulation Under Conditions of Technological Change

When technology is changing rapidly, information on past costs may be even less relevant to future pricing decisions than it is when technology is static. For example, information about the past average cost of long distance communications using twisted wire pairs or coaxial cable may be totally irrelevant to determining the appropriate future prices for communications via microwave or satellite. The only way in which past costs are relevant for regulating new technology is that they may serve as a benchmark to indicate that the cost of adding units of the new technology is lower than the cost of adding additional units of the old technology.

Hence, if regulators do their job in the traditional way, no matter how hard they work, they will be looking at the "wrong" data.

Another area of pricing where regulatory processes and market processes diverge is in determining the appropriate depreciation rate for capital. Markets set these rates based on three factors: (1) Estimates of the rate of technological change in the industry; (2) the willingness of customers to absorb a particular schedule of depreciation expenses; and (3) the IRS. Regulators, on the other hand, needing evidence on which to base decisions and preferring not to have to

---

[40] See Owen and Braeutigam, "The Regulation Game," pp. 1–36. See also the other studies cited in footnote 1.

[41] Patrick Conley, "Experience Curves as a Planning Tool," IEEE Spectrum (June 1970), pp. 63–68; Perspective on Experience (Boston: The Boston Consulting Group, Inc., 1972).

[42] If regulated firms can retain any excess profits earned up until the next rate hearing, regulatory lag may actually add to economic efficiency. For a brief discussion on this question, see: Alfred E. Kahn, "The Economics of Regulation: Principles and Institutions, Volume II: Institutional Issues," pp. 59–60.

grant rate increases, are likely actively or tacitly to encourage regulated firms to set depreciation rates equal to the physical life of the equipment.[43] Under certain conditions, the regulated firm likewise faces an internal incentive to want long lives for equipment. The slower the equipment is depreciated, the longer the equipment remains in the rate base, and thus the longer it is bringing in the allowed rate of return.[44] Moreover, durability is one method of adding to the capital "requirements" for the firm.

In a period of rapid technological change, firms that wish to keep up with the change not only need to base prices on future costs but also must be able to modify quickly their mix of products and services and perhaps also change their organizational structure quickly. Such firms must be able to develop new products and services, enter new markets, abandon old markets, raise or lower prices, acquire new organizations, or sell parts of the old organization. In an unregulated industry such as the semiconductor industry, companies can make changes as rapidly as the company's own internal bureaucratic organization will allow.[45]

In contrast, regulatory processes are slow. For example, certificates of public convenience and necessity (214 certificates in the case of interstate communications common carriers, station licenses in the case of microwave facilities) may slow down the process of starting new services—especially if the regulatory agency allows competing applicants or companies to protest the filing—and then holds hearings on the protests.[46]

Furthermore, under conditions of rapid technological change, the factors that tend to make both the regulated firm and the regulators choose long depreciation periods worsen the ability of the industry and the regulated firms to adopt new technologies. Long depreciation periods leave the regulated firm in a position of having a relative overabundance of plant and equipment that has been in use for a long time but still has not fully recovered its capital cost. Firms in this position justifiably feel that they have been assured that the plant will be paid for. The fact that it has not, however, leaves them unable to replace it with the new technology rapidly. Perhaps more important, the under depreciation is used as an argument for not allowing new firms to enter the industry with the new technology because it would threaten the ability of the regulated firm to recover the cost of its old plant.

The same regulatory processes that slow the adoption of new technologies and new services work to hinder abandonment of existing services even when few or no customers are now using them. Regulators feel obliged to offer due process to actual or potential customers

[43] Note that even using long depreciation periods for equipment may not produce the lowest prices over time. If those depreciation periods produce long delays in adopting much lower cost technology, consumers may be hurt, not helped, by the stretch out of depreciation.

[44] In reality, of course, while accumulated depreciation is subtracted from the gross value of plant and equipment, and thus tends to lower the rate base, current depreciation is also an expense that increases revenue requirements. Hence, the rate of depreciation affects both revenue requirements and the rate base and it is the net .esult of these two factors that determines whether a regulated firm has an incentive to depreciate more slowly than an unregulated firm.

[45] U.S. Federal Trade Commission, Bureau of Economics, "Staff Report on The Semiconductor Industry: A Survey of Structure, Conduct and Performance," by Douglas W. Webbink (Washington, D.C.: Government Printing Office, 1977), pp. 40–61, 92–114.

[46] See Owen and Braeutigam, "The Regulation Game," pp. 1–36. See also the other studies cited in footnote 1.

before any service offering can be withdrawn.[47] Competitive markets offer customers no such opportunities if firms have fulfilled their contracts.

Rate of Return Regulation of Competitors

The discussion above highlighted economic inefficiencies created by rate of return regulation of a monopolist under conditions of technological change. When a regulatory agency recognizes that technological change has eliminated natural monopoly and then allows in competition, the continuation of rate of return regulation will create additional economic inefficiencies. Assume that the transition for the previous monopoly firm has occurred: It has fully modernized its plant and equipment; it has adopted a more productive technology than if it had remained a monopoly; and it has become more efficient in its use of labor and materials given the technology adopted. Nonetheless, the continuation of requirements that all firms in the industry undergo both the tariff filing processes and the prior approval of facilities imposes economic costs larger than just the costs of preparing and filing the necessary applications.

In the tariff area, the requirement that firms post their prices makes it difficult for those same firms to bargain with their customers over rates or to adjust them quickly and flexibly to market conditions. This in turn means that the kind of price discounting that often occurs in a workably competitive market cannot take place. Particularly affected are the kinds of discounts that occur when a fairly large potential customer seeks service during a time of otherwise slack demand for the supplier. In competitive markets suppliers are often willing to offer such customers a lower price because the slack period business saves the supplier the costs of either closing part of his operations or continuing to pay his labor force despite the lack of demand. Because such periods cannot always be predicted, it is often not possible to file tariffs that cover this situation.

The requirement that firms file tariffs, moreover, immediately subjects them to the competition not of the marketplace but of fancy legal arguments over why the regulatory agency should reject, or suspend and investigate, the original firm's tariffs. Alleging price discrimination if the tariff provides for volume discounts (another common form of price discount in competitive markets) is considered a powerful legal argument that can be used to try to prevent the discounting. Should a competitor be successful in persuading the regulatory agency to order a hearing, the firm can face substantial legal costs in defending its proposed prices. As a result, it may find it less expensive to withdraw the new rate.[48] Conversely, the tariff filing requirement may also lead firms to file tariffs for services they are not yet ready to provide. Since such legal competition takes place

---

[47] For example, on August 19, 1976, Data Transmission Company requested emergency authority to discontinue all service on August 26, 1976. The Commission ordered DATRAN to continue service until September 15, 1976. 60 FCC 2d 958 (1976).

[48] On June 2, 1978, MCI filed a transmittal (No. 91) to modify one of its existing tariffs. See: *Memorandum Opinion and Order* in Docket 78–241, 69 FCC 2d 848. After the Commission designated the tariff for hearing, MCI proceeded to withdraw the tariff. On October 3, 1978, MCI filed a new transmittal (No. 99) to amend the same tariff. On December 21, 1978, the Commission adopted this revised tariff modification. *Memorandum Opinion and Order*, FCC 78–883, 70 FCC 2d 666 (1978).

more often over proposed rate decreases than proposed rate increases, the consumers are the losers in such battles.

The requirement for prior approval of facilities construction, with the requirement that the regulatory agency obtain public comment on such requests, similarly imposes costs beyond the costs of preparing the applications. If the application calls for new technology or a new service, the proposing firm not only is required to give its competitors an early blueprint of such technological innovations but also its exact plan for deployment. Because such early warning does not take place in markets not subject to price and entry controls, technological innovation is discouraged in regulated markets compared with unregulated ones.

III. Externalities

A second major economic argument for regulation is that the private market may not take into account externalities: Both benefits and costs that are imposed on parties other than the buyers and sellers of particular goods or services.

The traditional examples of externalities come from fields other than telecommunications. For example, if an automobile creates air pollution it imposes costs (such as the need to buy air purifying equipment and to paint buildings more often) on persons other than the buyer and seller of the automobile and the buyer and seller of the gasoline it consumes. Similarly, many kinds of information, such as information about candidates for public office, provide benefits to society in general which go far beyond the immediate buyer of the book, newspaper, or magazine containing the information.

Economists who have studied externalities in fields such as pollution, occupational health and safety, and information use have suggested many alternative methods to compensate for such externalities.[49] For example, economists have suggested pollution and health hazard taxes, subsidies on the provision of information, testing, registration, and labeling of energy using devices, and performance standards for housing construction materials.[50] Economists never recommend full common carrier type regulation to correct for externalities. Control over entry, exit, rate of return, and specific rates will never be the most efficient regulatory response to externalities, because the existence of externalities does not depend upon whether a firm has any monopoly power.

There are two different alleged externalities in common carrier communications. The first, and most important one, is that there are large external benefits from ensuring that every home and business has a telephone. The second alleged externality is the potential harm to other users of the network that one user's malfunctioning equipment can cause.

---

[49] Nina W. Cornell, Roger G. Noll and Barry Weingast, "Safety Regulation," Chapter 11 of *Setting National Priorities: The Next Ten Years*, edited by Henry Owen and Charles L. Schultze, (Washington, D.C.: The Brookings Institution, 1976), pp. 457–504. For a review of much of the relevant literature through 1975, see: Anthony C. Fisher and Frederick M. Peterson, "The Environment in Economics, A Survey," *Journal of Economic Literature*, 14 (March 1976), pp. 1–33. See also: Richard Zeckhauser and Albert Nichols, "The Occupational Safety and Health Administration. An Overview," and Larry R. Ruff, "Federal Environmental Regulation," in U.S. Senate, Committee on Governmental Affairs, *Study on Federal Regulation*, Appendix to Volume VI, Framework for Regulation, 96th Cong., 1st Sess., December 1978, Senate Document No. 96–14, pp. 163–344.

[50] See, for example: Allen V. Kneese and Charles L. Schultze, *Pollution, Prices and Public Policy*, (Washington, D.C.: The Brookings Institution, 1975)

The Universal Service Argument

The benefits to society from each individual having a telephone may be far greater than the benefits to any individual subscriber. Thus, if subscribers had to pay the full cost of a phone, not everyone would choose to pay for a phone, and too few phones would be used from the point of view of social welfare.

The externality argument behind universal service actually has two components. The first is that one person benefits from another person's subscription to telephone service, because each new subscriber increases the number of people any other subscriber can reach. The second component of the externality in universal service is that most people in our society value knowing that all other people in our society can call or be called for help in an emergency.

The failure of telephone rates to collect directly the value to an individual of the near universality of other subscribers may be overstated in part because of the way in which phone rates are charged. The most widespread charging system at present is a single fee that covers the instrument, access to the phone system and unlimited actual local usage. The externality lies only in the access portion of the service, not the instrument or actual usage portion. The flat fee system of pricing makes it more difficult to charge users a price that matches the cost and value of the access portion of the service. Unbundling of rates, which some common carriers are in the process of carrying out, helps to minimize the extent of this part of the externality behind universal service. To the extent that any remaining externality is seen to reduce significantly the number of people who subscribe to basic telephone service, a direct subsidy for access would compensate for the problem with less cost to society than providing for increased access by continuing to require rate of return regulated monopoly provision of *all* telephone service.[51]

The second part of the external benefit from "universal service" is the knowledge that everyone can call for help in an emergency. To ensure this, however, all that is needed is the minimum level of service to allow communications in an emergency. Such a minimum level of "universal service" might be satisfied by a single basic dial telephone on a party line accessible to each house and place of business that is available for emergency use only.

Once the regulatory agency determines the minimum level of service needed to ensure communications in an emergency, moreover, rate of return regulated monopoly is unlikely to be the most effective way of ensuring that it is provided. Because a firm facing rate of return regulation has an incentive to try to expand the rate base beyond the most efficient size, such a firm may be reluctant to stop at the minimum level of service or of equipment. Thus, there is a significant chance that emergency communications capabilities will be more sophisticated and hence more expensive than necessary to achieve the goal. In addition, a regulated firm may argue incorrectly in favor of the need for the firm to provide this emergency service at a price below cost to

---

[51] See the discussion in the next section of this paper of the relative merits of direct subsidies vs. rate of return regulation.

all customers, including those who would be willing and able to pay for it.

If the minimum level of service that society wants is in fact more than each individual subscriber is willing to pay for, the appropriate economic technique for dealing with that externality should be used. For example, low income customers could be given a direct subsidy, or the telephone company could be given a government grant just large enough to provide a minimum level of service to some of its subscribers. Two types of service that most meet the definition of providing for communications in an emergency, namely rural service and service in poor neighborhoods, are both directly subsidized. Rural service is partially paid for by REA financing, and service to the poor is paid for by welfare. In both cases, the direct subsidy is probably more effective than an indirect cross subsidy since it is more likely to ensure that the service provided is only what the market would not provide, and no more.[52]

Technical Harm to the Network Argument

A second externality argument used by some parties favoring regulation of common carrier communications is that without one monopoly provider of the service from one customer's equipment all the way to the other customer's equipment, a malfunctioning piece of equipment that was not promptly taken care of could cause harm to the network or to another user's equipment and hence impose external costs on those other users. This argument would apply particularly to the terminal equipment portion of the industry.[53]

Because the current local interconnection technology physically connects each customer's premises to a local central office switch, and ultimately to other customers, there is a possibility that malfunctioning equipment could cause problems elsewhere in the system besides at the customer's location. Rate of return regulation, however, is not the only, nor is it likely to be the best method for dealing with the problem.

For example, a program which requires testing and type approval or type acceptance such as the Federal Communications Commission uses for telephone terminal equipment and radio transmitters may be the least costly way of insuring that potentially destructive or dangerous equipment is not attached to the telephone network.[54] General performance standards are always likely to be more efficient than design or device standards, since the former allows firms flexibility in the way they meet the standards, whereas the latter could freeze into place obsolete technologies.

In addition to equipment performance standards, there may be a need to require the installation of low cost telephonic equivalent of circuit breaker panels to ensure that if type-approved equipment is installed incorrectly, it still cannot harm the network and other users.

---

[52] The question of indirect cross subsidies will be discussed further in the following section of this paper.

[53] For a discussion of the arguments for and against the so-called primary instrument concept (the idea that at least one telephone company-owned instrument should be on each subscriber's premises), see: Report and Order in the Matter of Implications of the Telephone Industry's Primary Instrument Concept, Docket 78–36, 68 FCC 2d 1157 (1978).

[54] 47 CFR 68.1–68.504. One problem with putting circuit diagrams on the public record, however, is that they immediately become available to existing and potential competitors, even before the manufacturers of the instrument can begin marketing it.

Technological Change and Externalities

To a lesser extent than was the case with a natural monopoly, technological change can reduce specific external effects. Thus, technological changes that result in the development of terminal equipment that can be prevented from imposing damage on other parts of the system (as with circuit breakers, for example) may end the need for regulation to prevent that kind of harm.[55]

Also, while a minimum level of universal telephone service might not have been profitable and hence possible in the past without a subsidy, new developments in lower cost radio communications technology may make it profitable in the future. Similarly, there may be available increasing numbers of low-cast alternative ways to obtain help in an emergency, such as citizens' band radio.

To the extent that technological change brings new services and products for which there is a greater demand than there had been for the earlier services and products, consumers may be more willing to pay for such products. This in turn may reduce the gap between the amount and kind of such services provided by the market and the amount and kind that society wants to have available. Hence, the externality arguments in favor of regulation may disappear, at least in some cases. And in any event, the regulation that may be needed to correct any externalities is not the rate of return regulation of a governmentally imposed monopoly, but direct explicit subsidies or performance standards to ensure the purchase of sufficient communications services of the desired kind.

IV. Income Redistribution and Internal Cross Subsidies

A third major reason some parties advance price and entry control regulation is the desire to use such regulation as a method for redistributing income or wealth.[56] Regulatory agencies may require regulated firms to engage in cross subsidies, that is to provide certain services at a loss and make up those losses from other services. Regulated firms may also choose to provide certain services at a loss without the permission (or even the knowledge) of the regulatory agencies. Finally, regulatory agencies may limit or prevent competition in the provision of one service by a regulated firm in exchange for the continued operation (at a loss) of some other service.

For example, some people argue that nationwide averaging of telephone rates is desirable precisely because it may create internal cross subsidies.[57] That is, such rate averaging may force urban customers to subsidize rural customers, assuming that the real cost of serving a rural customer is higher than for an urban customer.[58] Similarly, those who

---

[55] Another reason given by some parties for performance standards is to ensure that all devices use the same protocols, or can be interconnected. The market itself, however, will create such protocols because it is in the best interests of all unless transactions costs are too high. Only in the case of high transactions costs would such standardization possibly need to be imposed by the Government.

[56] See especially: Richard A. Posner, "Taxation by Regulation," The Bell Journal of Economics and Management Science, 2 (Spring 1971), pp. 22–50.

[57] See, for example: Statement of John D. deButts, Chairman of the Board and Chief Executive Officer, American Telephone and Telegraph Co., September 28, 1976, in U.S. Congress, House of Representatives, Committee on Interstate and Foreign Commerce, Subcommittee on Communications. *Competition in the Telecommunications Industry*, Hearings, 94th Cong., 2d Sess., September 28, 29 and 30, 1976, Serial 94–129, pp. 13–16.

[58] This paper does not address whether the present nationwide averaged MTS rates do in fact contain cross subsidies and if so which group of customers is subsidizing what other groups.

support the concept of "universal service" often argue that it is, in effect, a cross subsidy from high income to low income customers. There may also be subsidies from business to residential customers, and even from telephone users at low-use hours to users at peak-use hours. There are, however, at least four reasons why regulation is an inefficient method of redistributing income or wealth.

First, regulation is an imprecise method for helping certain specific individuals or groups since much of the subsidy goes to persons other than those for whom it is intended. Some persons may receive no benefit at all, and many people will be made worse off by the regulation. Second, such internal cross subsidies create signals to potential entrants and competitors that cause them to enter the "wrong" markets and produce the "wrong" services and products. Third, such internal subsidies hide from the public the true magnitude of the subsidy as well as the identity of the true winners and losers from the subsidy. Finally, those who are helped by the regulation would almost always be made better off and never worse off by a direct cash grant or subsidy.

Internal Subsidies Are Received by Persons Other Than the Intended Recipients

Consider, for example, the arguments in favor of pricing local residential telephone service below cost and making up any loss through higher rates to business users or higher long distance rates.[59] If all residential telephone subscribers were charged a rate below cost, then all would receive a subsidy. If *all* local services were priced below cost, not only would a family with a single dial phone on a party line receive a subsidy, but so would a family with six push button phones on three separate private lines. Thus, all telephone users, both high income and low income, would receive a subsidy, and depending upon its structure, the high income users might receive a far larger subsidy than the low income users. Of course, those individuals and families who do not have telephone service at all would receive no subsidy. Consequently, much of the subsidy would go to individuals other than those for whom it was intended while others would receive no subsidy at all.[60]

At the same time, any losses incurred by a telephone company in providing this cross subsidy must be made up somehow. In this ex-

---

[59] A separate and non-trivial question is whether such subsidies do in fact take place. For example, a study by the staff of the New York State Public Service Commission found that terminal equipment charges did *not* contribute to keeping down the cost of basic residential service. Instead, the staff study found that basic service *subsidizes* terminal equipment. See the Testimony of Alfred E. Kahn, Chairman, New York State Public Service Commission, in U.S. House of Representatives, Subcommittee on Communications, "Competition in the Telecommunications Industry," op. cit., pp. 1001–1020.

In general, it is difficult to determine if subsidies exist due to the existence of substantial joint and common costs between local residential service and both commercial and long distance service. For the purposes of this discussion, however, the authors of this paper simply wish to discuss the undesirable effects of such a cross subsidy if it does exist.

[60] In California, a limited telephone service at subsidized rates has been provided to subscribers thro gh the use of so-called "lifeline" rates. This subsidized service is available to *all* custo ners in the service area, however, not just low income families. In 1976, 24 percent of the families who subscribed to the subsidized lifeline service had incomes above $15,000. See: William S. Fockler, "Lifeline: Welfare Pricing of Local Telephone Service," in Assessing New Pricing Concepts in Public Utilities, Proceedings of the Institute of Public Utilities Ninth Annual Conference, (East Lansing: Michigan State University, Graduate School of Business Administration, The Institute of Public Utilities, 1978), pp. 458–473.

ample, local business subscribers or long distance telephone users must make up the difference. Such a subsidy scheme is inherently inefficient because some users are paying less than the real costs to society of the services they use, and others are paying more than the real costs. Hence, there will be a tendency for the first group to use too much of the service, and the other group to use too little.

Internal Subsidies Give the Wrong Signals to Potential Competitors

The second reason that internal subsidies through regulation may be inefficient is that in a period of rapid technological change they may create incentives for new entry, new competition and the development of new products and services—all in the "wrong" markets. To continue with the previous example, suppose that long distance and local commercial telephone service are priced above cost to subsidize local residential telephone service. If that were true, competing manufacturers and communications service companies would have an incentive to enter the commercial and long distance markets and provide lower priced equipment and service, even if it were true that their costs of producing that equipment and service were *higher* than those of the original telephone company.

Similarly, if residential telephone rates were set below cost by the existing telephone company, potential entrants might not enter that market, even if they were capable of providing such equipment and service at lower real cost than the phone company.

It does not matter which way the internal cross subsidy goes: Whether commercial service subsidizes residential service, or whether residential service subsidizes commercial. As long as some prices are set above cost and others below cost, potential competitors will have an incentive to develop substitutes for the services where price exceeds cost. Hence, they may undertake research and development and spend money on the "wrong" kind of investment.

Furthermore, if such a situation exists, the efficient regulatory solution is never to prevent competition or new entry in the service where price exceeds cost. Instead, the solution is to repeal any requirement and any justification for cross subsidies by the telephone company. In addition, efficiency will be improved if other companies are allowed to buy services when cost exceeds price, and if feasible to resell those services in competition with any services where price exceeds cost. Such reselling will tend to cause any internal cross subsidies to disappear.

Internal Subsidies Hide What Is Really Happening

A third reason why internal cross subsidies, whether required by regulatory agencies or done without the knowledge of regulatory agencies, have undesirable effects is that subsidies hide what is really happening. Because there are no records of direct payments of cash subsidies or collection of taxes, internal subsidies make it impossible for the public to know the magnitude of the subsidy or the identity of the group of gainers and losers and thus to hold either the regulatory

agency or the regulated firm accountable.[61] Thus, it is difficult today to know whether commercial telephone service subsidizes residential service, and even more difficult to be certain of the dollar magnitude of any such subsidy.[62] It is even less conceivable that one could determine which groups by income, race, and location gain how much and which groups lose how much from any such subsidies. For example, how many dollars per year a family in Chicago with an income of $4,000 is receiving from reduced local phone rates is probably unanswerable. Indeed, it may never be known if they are receiving any subsidy at all, despite arguments that they "need" low priced local telephone service.[63]

Similarly, because of the lack of accountability for internal cross subsidies, it is difficult to determine how big a subsidy is desirable or "necessary." For example, those who speak in favor of "universal telephone service" generally neglect to define precisely that term. However strongly one favors "universal service," the way one defines that term will have a huge impact on the cost of such service and the method for attaining it.[64] Would universal service be achieved if everyone had a single black telephone instrument on a four-party line, or is universal service only achieved if every family had three color push button telephones and a single party line? Similarly, does universal imply that *everyone* should have a reduced local telephone rate, or does it imply that families with an income under $4,000 should have a reduced local telephone rate, or does it only imply that everyone should have sufficient income that he could purchase local phone service if he chooses to? Could a goal of universal service be met by a low rate for access to a single telephone instrument line, but a higher charge for every minute of use?

Internal Subsidies Are Less Efficient Than Direct Cash Payments

The final point to be made about such an internal cross subsidy is that the targeted group would almost always be made better off, and never worse off, if it received a direct cash subsidy rather than an internal cross subsidy. Thus, if the ultimate purpose of requiring low

---

[61] One particular case of an alleged cross subsidy in common carrier communications in which it seems important to make the common carrier accountable is the alleged contribution to our national defense. Currently, decisions about plant and equipment are made by suppliers who have national defense needs in mind. But extra hardening of generally used facilities has been paid for by ratepayers not taxpayers. The result has been a lack of debate over the relative merits or relative costs of, for example, extra hardening versus government subsidized surplus capacity as a way of ensuring communications after a nuclear attack.

[62] On the optimal way to price telephone service so users would pay for the costs of services they want and not other services, see: Bridger M. Mitchell, "Optimal Pricing of Local Telephone Service," American Economic Review, 68, (September 1978), pp. 517–537.

[63] Of course, even direct cash subsidy programs do not guarantee that subsidies have the anticipated results including going to those for whom they were intended. See: Gerald R. Jantscher, "Bread Upon the Waters: Federal Aids to the Maritime Industries" (Washington, D.C.: The Brookings Institution, 1975); and Charles L. Schultze, "The Distribution of Farm Subsidies: Who Gets the Benefits?" (Washington, D.C.: The Brookings Institution, 1971). But at least it is easier to determine the identity of recipients with direct cash subsidies.

[64] There may not be a reason for the government to prescribe any form of universal telephone service, even if the government is strongly committed to other goals such as diminishing poverty and discrimination. This paper, however, does not address the pros or cons of universal service. Rather, it addresses the best method of achieving the goal of Section 1 of the Communications Act, "to make available, so far as possible to all the people of the United States a rapid, efficient, Nation-wide, and world-wide wire and radio communication service with adequate facilities at reasonable charges. . . ."

priced local service is to help poor families, those families would be better off if they received cash. The reason they would be better off is that if they were given the cash, they could decide whether they preferred to spend the money on telephone service or food or shelter or whatever. Hence, they could purchase those goods or services that had the greatest value to them.[65]

Thus, for all of these reasons, look with great suspicion on any arguments that price and entry regulation should be used to redistribute income or wealth by creating or maintaining internal cross subsidies.[66]

V. Conclusions

This paper discussed three categories of reasons some parties have advanced for applying price and entry controls to regulate common carrier telecommunications firms: (1) They may be natural monopolies; (2) there may be significant externalities in this industry; and (3) the desire to redistribute income through regulation. In each case the discussion indicated that price and entry controls were either costly or inappropriate for the goals sought.

Common Carrier Deregulation and Antitrust Enforcement

This paper began with a discussion of natural monopolies. It indicated that traditional common carrier regulation may have adverse effects on the efficiency of firms. For example, rate of return regulation creates incentives for firms to be less careful about cost minimization. Regulation is also backward looking, whereas firms need to be forward looking. Similarly, firms need to be able to change rapidly, especially in response to technological change, whereas regulation with its requirements for administrative procedures and due process is inherently slow. Moreover, regulation may become a barrier to change and new or additional competition. For that reason, full common carrier regulation (including regulation of entry and exit, prices and profits) is unambiguously worse than unregulated competition in markets that are or are rapidly becoming competitive. In those markets, our recommendation is clear: Get rid of any regulatory restrictions on entry and exit, prices and profits.[67] If there are externality or income distribu-

---

[65] Note, however, one important fact about subsidy by regulation. Often different government agencies and different legislative committees control subsidy and taxing powers on the one hand and regulatory powers on the other hand. Thus, one reason legislative committees and executive departments may choose regulation over a direct subsidy or a tax is that the regulatory organization may be more favorable to the subsidy than is the taxing or the direct subsidy organization.

[66] Of course, some people do not accept the notion of complete consumer sovereignty. Such people might fear that if low income families were given a cash subsidy they would spend that money on "undesirable" goods or services (such as liquor and gambling). While the authors of this paper would prefer to allow complete consumer sovereignty, even if consumers were only allowed to spend the subsidy on telephone service, they would still be better off if they received a direct subsidy rather than an indirect internal cross subsidy through regulation. With the direct "telephone subsidy" consumers would at least have the choice of which telephone services among those available from the telephone company they would prefer to buy.

[67] While the recommendation of this paper is clear, any change to an unregulated market cannot occur overnight. Several transition problems require resolution, among them establishing access arrangements between intercity and local segments of the industry, establishing effective depreciation policies, and determining what rules will govern participation in unregulated markets by firms that also participate in regulated markets. These issues, however, must be faced under present conditions and are reflected in ongoing FCC proceedings. See for example, the Supplemental Notice of Inquiry and Proposed Rulemaking in the Matter of MTS and WATS Market Structure, Docket 78–72, released August 30, 1979; and Tentative Decision and Further Notice of Inquiry and Rulemaking in the Matter of Amendment of Section 64.702 of the Commission's Rules and Regulations (Second Computer Inquiry), Docket 20828, released July 2, 1979.

tion problems in those markets, deal with them directly, but do not use common carrier regulation.

In markets that appear to have the characteristics of natural monopolies, it may be less clear that competition is possible. It is not obvious, however, that many (or any) markets are truly natural monopolies, especially in the absence of a market test. More important, markets that exhibit natural monopoly characteristics today are likely not to exhibit them tomorrow. And, in any case, even in natural monopoly markets, it is not obvious that the benefits of common carrier regulation outweigh the costs. As we indicated previously, common carrier regulation creates incentives for inefficient production (such as overcapitalization), is backward looking when firms need to be forward looking, and has difficulty dealing with rapid changes in demand for services, technology, and associated changes in costs and kinds of services available. Thus, such regulation may substantially slow the rate of innovation and technological change.

Nevertheless, one may still choose to treat monopoly markets differently from competitive markets: The former subject to some minimal regulation, and the latter subject to no regulation. In addition, as long as some markets are regulated and others are not, firms that operate in both markets will have an incentive to evade regulation by trading off one market against the other, that is, by shifting profits to the unregulated market and costs to the regulated market. Thus, in a transition period from more to less regulation, there may be a particular need to worry about firms that operate in both regulated and unregulated markets.

These factors suggest a large number of conclusions. First, as indicated above, competitive or soon to be competitive markets should be deregulated. Totally free entry and exit and pricing should be allowed. To the extent there is legitimate worry about a potential future oligopoly or monopoly structure in those currently competitive markets, the Justice Department should be relied on to enforce sections 1, 2, 4 and 7 of the Sherman Act and section 7 of the Clayton Act, and the Federal Trade Commission relied upon to enforce section 5 of the Federal Trade Commission Act.[68]

The next conclusion is that while competitive markets should not be regulated, it would be desirable to limit the ability of monopoly firms to cross subsidize between competitive and monopoly markets. The ideal solution might be to separate totally any competitive services from a monopoly firm. This would imply first a prohibition on monopoly firms entering any new competitive markets, and second divestiture of any existing competitive parts of a monopoly firm. Such divestiture might include both horizontal and vertical subsidiaries, including manufacturing subsidiaries, value added subsidiaries, certain long distance transmission facilities and computer services.

To the extent that such prohibitions against entry into new markets and divestiture of existing parts of a common carrier are not considered feasible for legal or public policy reasons, any such competitive services could be done by fully separate arms-length subsidiaries.

---

[68] On antitrust enforcement, see: F. M. Scherer, Industrial Market Structure and Economic Performance, (Chicago: Rand McNally and Company, 1970), pp. 422–517.

Such subsidiaries would not only be legally separate companies but totally physically separate entities that do not share any common facilities. Physical separation appears to be necessary to minimize the problems of costs and their impact on cost accounting, particularly if there still exists the rate of return regulation of the monopoly part of the firm.

The Macroeconomic Impact on the Economy of the Rate of Return Regulation Under Conditions of Technological Change

The discussion of common carrier regulation implies that industry has invested in more capital than may have been necessary to achieve the same output and has held on to older technology longer than would have occurred in a competitive market. First, the regulated firms face a series of incentives to expand their rate bases, and to substitute more capital for other inputs than they would in the absence of regulation. Second, the relatively guaranteed rate of return that resulted from past restrictions on competitive entry has given these companies high credit ratings and hence first choice in the capital markets.

It is also possible that the regulatory structure with its built in barriers to entry and change has resulted in fewer communications services being produced and sold than might have come from a more competitive market. Such a conclusion may never be able to be proven or disproven quantitatively. Nonetheless, such an outcome is consistent with the likelihood that despite regulation, prices were not as low as they might have been had the market structure of the industry been more competitive. The same incentives to expand the rate base and not use the least-cost techniques indicate a likelihood that costs were not the lowest possible. Moreover, the demands of the regulatory processes themselves impose costs, sometimes large ones. Thus, with costs and prices higher than they would have been under a competitive structure, fewer communications services would have been produced and used than would have occurred if prices consumers faced were lower.

For the economy as a whole, if these conclusions are true, two impacts are suggested. First, communications may be relatively more capital intensive than might have been the case under an unregulated and a more competitive structure.[69] At the same time, communications firms may have been able to obtain more financial capital at lower interest rates that they might otherwise have done.

The possible failure to produce as much output in the communications industry as might otherwise have been produced may also have had another macroeconomic effect. Communications are widely used to make other production processes more efficient.[70] To the extent that

---

[69] Despite the excess use of capital that may have occurred, the impact on employment cannot be specified. On the one hand, Leibenstein's X-Inefficiency argument would lead to greater use of labor than necessary given the technology adopted. This would suggest that a competitive market would have used less labor than was in fact used. On the other hand, if a competitive market would have used a less capital-intensive technology and would have produced more output, employment might have been higher than it was. What the impact of competition on past employment would have been depends upon which effect was stronger, and this cannot be determined empirically.

[70] See: U.S. Department of Commerce, Office of Telecommunications. The Information Economy: Definition and Measurement by Dr. Marc Uri Porat, OT Special Publication 77–12(1), (Washington, D.C.: Government Printing Office, 1977).

less expensive, more widely available communications would have been used by businesses, productivity would have been lower than it might otherwise have been. Again, these second order effects are uncertain and would be hard to measure even if it could be conclusively and quantitatively shown that communications have been overcapitalized and underproduced, and by how much.

The possibility of these effects having occurred because of rate of return regulation of common carrier communications raises again the fundamental question of whether to continue such regulation. One possible argument in favor of retaining price and entry controls is a possibility that technology in the future may become more subject to economies of scale. Such a possibility exists, for example, if fiber optic cable becomes economically competitive with the various forms of radio systems for intercity transmission. If fiber optics, however, becomes a low cost substitute for twisted pair cable, coaxial cable, and microwave systems, the cost of those alternative technologies would set a ceiling on the price of fiber optic systems. Given that, and given all the costs, monopoly regulation seems unnecessary.

Externalities

The next part of this paper discussed regulation of externalities. Even if monopoly regulation is unnecessary, the present need for externality regulation may well continue for some time, especially if it is not feasible to use taxes and subsidies to correct for externalities in telecommunications. Externality regulation may even need to increase somewhat over its present level. Performance standards on common carrier equipment right now is concentrated on terminal equipment, not on the equipment used within the transmission paths.

Thus, the argument in favor of the use of performance standards, equipment testing and approval, and installation of circuit breaker equipment rather than monopoly rate of return regulation, is not an argument for total deregulation. Instead, it is an argument for the use of appropriate regulation and the minimum level of regulation necessary to deal with a real externality. It also says that regulation that is appropriate for dealing with an externality need not be an argument for limiting entry or competition in an industry.

The other possible kinds of telecommunications externalities, those related to ensuring universal service and emergency access to the telephone network are not convincing arguments for rate of return regulation. It is unclear that there is a significant externality from knowing everyone is connected to the system, because users may already be willing to pay the full costs of such interconnection. And in many cases, emergency access can be provided by an alternative method that is less costly than traditional telephone service. In any case, none of these externalities, if they exist, requires traditional regulation of profits, prices, entry and exit.

Income Redistribution

Finally, the paper discussed the use of regulation to redistribute income or wealth through internal cross subsidies. Just as it is true for

externalities, arguments in favor of income redistribution are not necessarily arguments in favor of entry or price controls. Since internal cross subsidies are inherently inefficient, direct taxes or cash subsidies are always preferable. Nevertheless, if internal cross subsidies are ever used, they should be carefully defined, narrowly limited, and frequently measured. Only then can the public be sure who gains and who loses from any cross subsidy that does exist.

THE IMPACT OF REGULATION OF FINANCIAL INSTITUTIONS ON COMPETITION, AND THE ALLOCATION OF RESOURCES

By Paul M. Horvitz*

CONTENTS

	Page
Summary	225
I. Introduction: The Role of Financial Institutions	226
II. Economic Theory and Regulation of Financial Institutions	229
III. Regulation, Bank Failure, and Deposit Insurance	233
IV. Impacts of Bank Regulation	237
A. Bank Structure	237
B. Activity Restrictions	242
C. Interest Rate Control	244
D. Balance Sheet Constraints	245
E. Consumer Protection Regulation	247
V. Credit Allocation	248
VI. Impacts on the Macroeconomy	250
VII. Policy Implications	254
References	255

SUMMARY

The role of financial institutions in the economy is to facilitate the flow of funds (control over real resources) from savers to those planning real investment. Because of economies of scale, expertise, and diversification, if the markets for savings and lending are perfectly competitive, and there are no governmental impediments, financial intermediation can result in both higher returns to savers and lower costs to borrowers.

For various reasons the view has been that competitive financial markets would involve an unacceptable degree of risk, and that the resulting allocation of resources, even if efficient in a Pareto sense, would not be in accord with social preferences. As a result, there are now many regulations surrounding the operations of financial institutions, many of them specifically intended to reduce competition. If competition is imperfect, or if government regulations impose costs on the institutions, then the spread between the return to savers and the cost to borowers will be increased and the role of financial intermediaries will be reduced, with a detrimental impact on the efficiency of resource allocations.

Regulations and restrictions on the operations of financial institutions generally fall into four classes: Restrictions on activities; restrictions on structure, particularly entry; regulation of prices,

* University of Houston.

particularly interest rates on deposits; and balance sheet constraints, particularly capital and liquidity requirements. The original intent of these regulations was to enhance the safety of financial transactions, usually for the benefit of the depositor or saver, though recently regulations aimed at the benefit of the borrower or another customer have grown. Some regulations are aimed at influencing the use of credit in the economy in accord with certain social priorities (savings and loan associations are limited to housing-related assets to promote home ownership and the housing goals of the Nation), or concepts of equity (usury laws, for example). While some government regulations do not bind, and hence have no effect (commercial banks would meet most State reserve requirements even in the absence of legal constraints), others do affect the costs of financial institutions, the level of competition, and the allocation of credit. Some of these are effective in promoting safety, while others are not.

Regulation of financial institutions has grown over time, and financial institutions now comprise one of the most heavily regulated sectors of the economy. Yet the U.S. economy differs substantially from that which existed when many of the regulations arose. In particular, this paper argues that the development of a credible system of Federal deposit insurance makes unnecessary much of the regulatory framework aimed at the safety of those institutions covered by such insurance. (While this includes commercial banks, mutual savings banks, savings and loan associations and credit unions, the focus in this paper is on commercial banks, largely because more empirical research has been done on commercial banking.) This conclusion is reached in Section III of the paper, following a brief introduction (Section I), and a discussion of the economic theory of the regulation of financial institutions (Section II).

Section IV consists of a survey of the empirical research that has been done on the impacts of the various forms of regulation of financial institutions. The general conclusion from that survey is that most existing regulations do reduce competition in the provision of financial services, and have adverse efficiency implications. The magnitude of these effects are relatively small, in most cases, however, and even the direction of some effects is unclear because of "second best" considerations.

Section V recognizes that much of our regulatory framework is now aimed at producing particular effects on real resource flows, and that is true of some regulations originally intended to promote bank safety. The general conclusion from empirical studies is that using regulation of financial institutions or markets to affect real resource flows is inefficient at best, and perhaps impossible.

I. Introduction: The Role of Financial Institutions

The importance of financial institutions to an economy flows from their role in facilitating saving and investment. In any economic system more advanced than the Robinson Crusoe subsistence economy, saving and investment activities are undertaken separately by different individuals and firms. Some means must be found to channel funds

from those in the economy who are accumulating surpluses (saving) to those who are spending more than current income (investing). This transfer of funds can take place by means of a direct borrowing-lending transaction, but such a transaction requires some fortuitous coincidences—borrower and lender must be interested in the same size transaction, and the same maturity. While such transactions do take place, even in the most developed economies, most flow of funds from savers to investors is funneled through financial institutions.

Financial institutions serve as intermediaries to the benefit of both the surplus and deficit sectors of the economy. Funds flow from the surplus units to financial institutions in exchange for indirect financial assets—liabilities of financial institutions. Funds flow from the financial institution to the deficit units in the economy in exchange for their primary securities. The key point is that this circuitous process of "intermediation," whereby financial institutions gather the savings of individuals and relend those funds to business firms and other investors is efficient. That is, it makes both lender and borrower better off.

Instead of lending directly to an individual or business firm needing additional funds, the saver instead acquires an indirect financial asset issue by a financial institution. These indirect assets have different names and characteristics, such as deposits (in the case of banks), shares (in the case of some savings and loan associations or credit unions), or insurance. There are several advantages to the saver in this process. First is that the indirect securities have a wider range of maturities than those in which primary securities are typically issued. Financial institutions accept funds for as short as a day, or for as long as many years. These indirect securities come in a wide range of denominations. In particular, the saver can put small amounts into a transaction with a financial institution, whereas this would not be practical in dealing directly with a business firm or other borrower. Second, the securities issued by financial institutions tend to have less risk of default than primary securities. Intermediaries are able to reduce risk by investing the funds received from money savers in a wide range of different primary securities. Diversification reduces risk, and the financial intermediary is much better able to obtain the desired diversification than the individual saver. Also, the financial institution is able to build up expertise as a result of specializing in this type of business, and is better able, therefore, to evaluate credit worthiness of borrowers.

A third advantage is that the assets acquired by dealing with a financial institution typically have greater liquidity than most primary securities; the instrument acquired from a financial institution may be converted to cash with little or no risk of loss, either on demand or on short notice. Financial institutions can afford to commit themselves to this because they know that not all their many creditors will simultaneously seek to take advantage of it. Although there may be some penalties or loss involved in early conversion to cash, these are much less than would typically be the case in dealing with primary securities.

The net effect of these advantages is that the individual saver can generally obtain a higher net return by investing in the obligations of financial intermediaries than he could obtain by lending directly

to a borrower. That is, after allowing for the differences in risk and liquidity, and saving on the costs of searching for an appropriate investment, his net return is generally higher.

There are also advantages to borrowers in the economy from this process of intermediation. Financial institutions will buy primary securities from deficit spending units in a wider range of maturities than individual buyers. Institutions make loans for as short a time as overnight or for as long as many years. They also will buy securities or make loans in larger amounts than individual lenders typically could. The net effect is that the interest cost to the borrower, after allowing for the savings on searching for funds, is lower than if we were to seek lenders willing to lend directly.

These considerations suggest the importance of financial institutions in facilitating saving and investment in the economy. They suggest further that the efficient functioning of the financial system is uniquely important to the economy. None of this indicates that regulation should be particularly necessary or desirable in assuring the smooth functioning of the financial system.

With respect to most products and services in the economy, primary reliance is on market forces to achieve efficient results. In the case of financial intermediation, however, the United States has decided not to rely completely on market forces, but has developed an intricate and comprehensive set of regulations. It is important to realize that the set of regulations of financial institutions did not result from careful analysis of the characteristics of financial intermediation that prevent competitive markets from functioning efficiently, nor from a determination of the optimum set of regulations for financial institutions. Our regulatory structure grew over time, with regulations added to meet perceived shortcomings in the system. Some of these additions were piecemeal, but others represented major changes in the system. It is correct to speak of additions, since eliminations of regulations have been rare. In particular, there has been little consideration at the time of major revisions in the regulatory system of what previously existing regulations would no longer be necessary. The creation of the Federal Reserve System in 1913 was one such major change in the system. More important was the creation of a system of Federal deposit insurance in 1933. The introduction of Federal deposit insurance made unnecessary a great many existing restrictions on operations of the banking system. Yet rather than leading to a reduction in other regulations, the timing of the establishment of the Federal Deposit Insurance Corporation in the depths of the Great Depression was accompanied by an increased set of regulations.

Most regulation is concerned with the problem of solvency. The concern is that a free market in financial intermediation will result in "too great" a number of failures. In addition to Federal deposit insurance, the financial system includes a Federal lender of last resort and other types of regulation. These other regulations consist of five types:

1. *Restrictions on entry.*—To enter the banking industry, a charter must be obtained from a government agency. The burden of proof is on the applicant to demonstrate his ability to operate a bank soundly

and to demonstrate that there is a public need for the new bank. Restrictions also apply on the ability of an existing bank to establish a branch. The severity of such restrictions varies from State to State, but no bank, regardless of need, can establish a branch in a State other than the one in which it is chartered. In many cases these restrictions on entry through branching also apply to unmanned facilities to dispense cash or accept deposits.

2. *Activity restrictions.*—Banks are severely limited in the types of activities in which they can engage. Similar restrictions apply to bank holding companies (corporations that own control of a bank). Restrictions apply not only to type of business, but also within the lending function, with limits on the size of loans to individual borrowers and on loans to insiders.

3. *Balance sheet constraints.*—Banks are subject to requirements as to their liquidity (holdings of reserves) and capital.

4. *Pricing.*—Banks are limited in the rates that can be paid on time deposits, and prohibited from payment of interest on demand deposits. In most States there are also usury laws which limit the rates banks can charge on loans.

5. *Consumer protection.*—In recent years a variety of resolutions aimed at protecting consumers in their dealings with financial institutions were introduced. As distinct from most of the preceding, these are not designed to promote solvency of the institutions.

Regulations are enforced by means of a complicated structure as well. National banks are chartered and supervised by the Comptroller of the Currency, and are also subject to regulations of the Federal Reserve System and the Federal Deposit Insurance Corporation. State chartered banks are supervised by the State banking authority and, if they choose membership in the Federal deposit insurance system or the Federal Reserve System, are subject to supervision by one or both of those Federal agencies. Supervision, in the case of banks, includes the process of bank examination, a unique form of detailed government investigation of the finances and operations of the institution. Examination has long been part of the bank regulatory framework, through its proper role has changed substantially since the introduction of deposit insurance.

II. Economic Theory and Regulation of Financial Institutions

The economic theory of regulation is rather well specified in the literature, though it has only recently been applied to financial institutions and services. Regulation is viewed as an appropriate approach when certain market failures prevent competitive forces from producing an efficient result. Such market failures include economies of scale (which preclude a stable competitive solution), information deficiencies (which prevent informed consumer choices), public goods and externalities (cases in which an equilibration of private costs and benefits is impossible or incomplete). These situations may require regulation to assure an efficient allocation of resources. Efficiency in this context means that no one can be made better off, without also making someone worse off. But society may opt for other than the efficient allocation of resources out of concern for some group of consumers. In the

banking case, we choose to protect small depositors, even if so doing violates efficiency criteria. In this section of the paper we will examine the application of these concepts of market failure to banking, to determine whether they provide justification for regulatory framework. We will also consider the justification for concern with the "equity" or "distributional" basis for regulation.

The principal reason for government intervention in the banking industry is that the distinctive product of banks is money. The liabilities of commercial banks comprise the bulk of our money supply. Government responsibility for this product is a basic part of our system, including the constitutional provision giving the Federal Government the power to "coin money and regulate the value thereof." This is a product that has some attributes of a public good; there are numerous externalities connected with its production and use, and there may be substantial information deficiencies in the unregulated production of money services.

The money supply, produced by commercial banks, has an effect on the aggregate economy. Everyone benefits from an appropriate level of aggregate economic activity, and suffers from a depressed level. Likewise, all benefit from the existence of an efficient payments system. An individual's benefit from an appropriate aggregate supply of money is not directly related to his ownership of money, and hence the money supply is in the nature of a public good. There is general agreement, even among those economists who consistently seek to minimize government involvement in the economy, that government control of the money supply is appropriate. This sets some minimum basis for government regulation of banking, though it should be noted that this does not require much of the described regulatory system. Federal Reserve ability to control the money supply does not require a structure of entry restrictions, price controls, and activity constraints. Most economists argue that even reserve requirements are not necessary for this pupose. (Carson, 1964.)

Externalities provide a more important basis for regulation. Use of a check payments system is benefited if everyone else also has a checking account. The value of a checking account is reduced if others, not having a checking account, choose to pay in currency and expect payment in currency. Regulation which encourages use of the check payment system seems appropriate.

If the banking system is to play its role in furnishing money and payments services, banks must be able to convert demand deposits into currency, even in periods of stress. Detailed regulation grows out of concern that, in the absence of regulation, bank failures will be too frequent. The key issue is whether regulation is necessary to solve the problem.

The nature of modern banking in the United States is that banks operate with limited amounts of equity capital, and that banks hold reserves equal to only a fraction of their deposit liabilities. In such a system a bank can become insolvent as a result of operating losses or losses on assets. Even in the absence of such technical insolvency, a bank may be unable to meet unexpected large demands for cash. Such a run may be caused by customer suspicion of the bank's insolvency, a reaction to the actual insolvency of other banks, or a random coincidence of demands.

Even in the absence of regulation, banks obviously have an incentive to avoid insolvency, which means a total loss to stockholders. But the high degree of leverage that prevails in banking does provide some incentive toward risk taking—if the risk is successful, the benefits accrue to owners; if unsuccessful, the losses are shared with depositors. (Jensen and Meckling, 1976.) With respect to liquidity, banks have an incentive to provide liquidity to meet even unlikely large random withdrawals. Moreover, if the bank is solvent, it is able to borrow cash to meet such demands (even without the Federal Reserve's lender of last resort responsibility). It is difficult, however, even for a conservative bank, to provide for an irrational, ill-informed run.

Externalities are part of the problem. Experience indicates that failure of one bank can lead to runs on other banks that, at least before the run, were solvent. Part of the problem is an informational deficiency. If depositors lack good information on the soundness of an individual bank, then withdrawing funds from a bank about which one has suspicions is the wise move. The depositor's incentive to participate in a run is due to the fact that if the run is unnecessary, i.e., the bank is sound, the cost is small. But if the bank is solvent, the depositor who withdraws his funds promptly will come out whole.

There are various ways of dealing with this problem. In the absence of government regulation, there will be some market pressures toward sound banking. In the early days of banking in the United States, bank capital ratios were extremely high. A lender of last resort is one means of assuring that sound banks will not be forced to close because of illiquidity. Detailed regulations can assure the public that banks are unlikely to fail because they cannot undertake substantial risks. A system of deposit insurance can prevent runs by convincing depositors not to participate in runs because they will be protected even if they are not first in line. The United States has adopted all of these approaches.

The U.S. experience with relatively free banking during the 19th century was the basis for establishment of the Federal Reserve System. Experience of the 1920's, not to mention the 1930's, was proof that existence of a lender of last resort was not sufficient to prevent an undesirable number of bank failures. The experience of the 1930's led not only to Federal deposit insurance, but also to additional regulations.

The need for stability in the money supply and to encourage use of the payments system justify some regulation on efficiency grounds. But part of the decision to regulate banks (and other financial institutions) is the desire to protect small depositors. On this basis, there is greater concern with the creditors of banks than with other business firms, not because banks are more important, but because their creditors are often small and unsophisticated.

Efficiency considerations are involved in the desire to make banking safe to protect small depositors. If small depositors lack information, or are unable to use information effectively, they will make wrong choices, resulting in inefficiency and welfare loss. If small depositors perceive an inability to make wise choices, they may abstain from using bank services, thus reducing the external benefits from widespread use of the check payment system. But the principal concern here is protection of the small depositor rather than the furtherance of eco-

nomic efficiency. The objective is to protect against the danger of losing one's life savings in a bank failure. Eliminating risky banking, eliminates the chance that a consumer will make the wrong choice by doing business with a risky bank. On the other hand, this eliminates the possibility that a risky bank is the preferred choice for some knowledgeable consumers. (Colantoni, Davis, and Swaminuthan, 1976.) Eliminating risk in banking by regulation leads to a choice between efficiency and equity considerations.

The considerations discussed above make a substantial case for some form of government intervention in banking markets. An unfettered market for banking services is likely to result in a number of banking failures that is inconsistent with economic stability, may discourage widespread use of an efficient payments system, may lead to uninformed consumer decisions because of a lack of information on the business, and may result in losses to small consumers.

There is another quite different basis for regulation of financial institutions. Society may seek to affect the allocation of resources in a particular direction, and it may be appealing to attempt this through affecting the flow of credit. This can be done directly, as in government loan programs, or through the tax system (for example, the tax exemption of interest on State and local government securities is designed to increase the availability of real resources to such government units by increasing the willingness of savers to lend to them). The most important examples of the attempt to influence the allocation of resources through an effect on the flow of credit are the variety of regulatory incentives and prescriptions aimed at influencing the flow of credit for housing. A large part of the existing regulatory structure is intended to serve this purpose.

Once the view is accepted that some regulation of banking is appropriate, it is necessary to consider whether the present framework of regulation is optimal, or whether some other set might cure the market failures that necessitate regulation with less expense and inefficiencies. Or if the purpose is to affect the allocation of resources, consideration must be given to which set of regulations can accomplish that at least cost in terms of administration and in terms of efficiency loss. This task requires an analysis of each type of regulation, and also requires an analysis of the side effects of various regulations.

Assessment of the impact of various restrictions on financial institutions is difficult in a meaningful context. It is rather easy to determine that in a world with Federal deposit insurance less regulation is better than more regulation. It is feasible, at least in principle, to evaluate the impact of a particular regulation on an otherwise unregulated world. What is difficult is the evaluation of a particular restriction in the presence of other restrictions and a particular institutional structure. Here are a few examples of the problem:

Banks are faced with restrictions on their ability to establish branches. In the absence of any other restrictions on banks, such a restriction leads to less competitive markets than would exist with unrestricted branching, and hence to higher prices and less convenient service to consumers. But consider the more realistic case in which banks are also subject to limitations on the interest rates that can be paid on

deposits. Restriction on the ability to compete on a rate basis will lead banks to compete for deposits on the basis of convenience. This will lead, if allowed, to a proliferation of branches in excess of the number that would be established in a perfectly free market. The costs of maintaining branches may reduce bank profits to no more than they would be reduced in the absence of the competition-reducing restrictions. Some bank customers are better off (those who place a high value on convenience) and some are worse of (those who place a higher value on interest income). Certainly consumer welfare is less than in the free market world in which some banks compete on a rate basis for customers and others compete on a convenience basis by establishing a branch network. But given interest rate constraints, it is not clear that restricting branching necessarily results in a loss of efficiency.

Consider the issue of generally restricting bank entry. It is easy to see that restrictions on bank entry will reduce competition in banking markets and make consumers worse off than if entry were unrestricted. But quantitatively, this loss may be insignificant if other (nonbank) firms are able to enter the market and provide services equivalent (or similar) to the services provided by banks. This seems to be a fairly widespread phenomenon in our economy, and hence the *degree* of welfare impact from such restrictions may be small even if the direction is unambiguously toward inefficiency.

An example of this situation is the restriction on interest rates that can be paid on small denomination deposits by banks and savings institutions. That restriction clearly imposes a welfare loss on small depositors. But the restriction has also stimulated the development of a new type of institution that pays a market interest rate. In such an environment, the welfare loss from the restriction imposed on banks and savings institutions is relatively modest. The new institution is the money market mutual fund, which provides an asset with credit quality and liquidity roughly equal to that of banks deposits.

Despite this problem, this paper will attempt an assessment of the impact of individual regulations. The conclusions on individual regulation constraints will have to be qualified, however, by these "second best" considerations.

III. Regulation, Bank Failure, and Deposit Insurance

Much of the structure of regulation of financial institutions is based on concern about failure. Concern about bank failures was reasonable during the 1920's and 1930's. Bank failures were common, and there were spillover effects—the failure of one bank generated runs on other nearby banks. Many of the failures of the 1920's were the result of the postwar agricultural collapse or other isolated local events that, it might be argued, the recently established Federal Reserve System was not intended to take care of. The Federal Reserve has received much more criticism for its handling of the 1930's. Massive bank failure was still possible even with a central bank.

Although the Federal Reserve turned out not to be the solution to the problem of bank failures, the depression of the 1930's led to establishment of a system of Federal deposit insurance. The Federal Deposit Insurance Corporation has turned out to represent a fundamental

change in the stability of the banking system, and has solved the problem of bank failure. Thus Friedman and Schwartz (1963, p. 11) write that:

> The enactment of Federal deposit insurance . . . probably has succeeded, where the Federal Reserve Act failed, in rendering it impossible for a loss of public confidence in some banks to produce a widespread banking panic involving severe downward pressure on the stock of money; if so, it is of the greatest importance for the subsequent monetary history of the United States.

Federal deposit insurance not only prevents the macro effects on the money supply that are the basic justification for bank regulation, but also protects the small depositor, a justification for bank supervision that is also important to some. Thus, Tussing (1967, 1970) argues that bank failure is no longer an evil to be feared, but can be dealt with easily. Mayer (1975) argues that an occasional large bank failure improves the efficiency of the banking system. In this view, the real risk of loss causes large, uninsured depositors to analyze their banks carefully, and this market surveillance reduces the need for government supervision of the banking system.

The development of deposit insurance means that the fundamental purpose of bank regulation has changed. Bank regulation is no longer necessary to protect the money supply, or to protect the small depositor—deposit insurance takes care of both needs. Regulation of banks is still necessary, however, to the extent needed to protect the deposit insurance fund. It is likely, of course, that the nature of the regulatory framework needed to protect the deposit insurance system is different from that needed to prevent bank failures and bank runs. Yet, that difference is not apparent from examination of legislation and regulations promulgated in recent years. Congress and the regulatory agencies react to real and imagined problems of bank soundness in much the same way they did before there was a successful deposit insurance system in operation.

It becomes important, therefore, to determine just which parts of the regulatory structure are necessary to protect the deposit insurance system, and which have become superfluous since the establishment of the FDIC. This requires a consideration of some aspects of the deposit insurance system.

There has been a sharp difference of opinion in the literature as to the adequacy of the deposit insurance fund. Edwards and Scott (1977, p. 8) argue that the current deposit insurance premium "has not been high enough for the Federal Deposit Insurance Corporation to accumulate a large reserve fund . . . total funds available to meet potential claims of depositors constitute about 1.5 percent of total insured bank deposits." This relatively low ratio of insurance fund to contingent liabilities has persisted throughout the life of the FDIC.

On the other hand, Gibson (1972) argued that the insurance premium is too high, as shown by the fact that losses in each year have always been a small fraction of annual income. Thus the FDIC has never had to dip into the accumulated reserve fund.

Edwards and Gibson are both right. The insurance fund is small in relation to potential liabilities. But the amount of these potential liabilities is not relevant to actual losses of the FDIC as long as bank insolvencies are detected promptly. Bank failures do not occur over-

night. Net worth of a failing bank usually declines over time. As long as the bank is closed as soon as it becomes insolvent, the FDIC stands to lose relatively little in a bank failure. This means that the deposit size of the failing bank is not a determinant of FDIC losses. The failure of Franklin National Bank in New York City in 1974, the largest in U.S. history, will probably not result in any loss to the FDIC.

Banks are examined then to detect insolvency so that the bank can be closed if it is found to be insolvent. This is particularly important in a world of Federal deposit insurance because depositors have no hesitation in making deposits in an insolvent bank. Without insurance, depositors will attempt to assess the soundness of the bank, and will not put funds in a bank with zero or inadequate capital. In such a world, even a rumor of insolvency or weakness can lead to a run on the bank. Evidence from recent failures and studies of unfavorable disclosures about bank problems suggests that even in the case of banks suspected of being insolvent, deposits may not decrease, and may even increase. (Kurtz and Sinkey, 1973.) In the absence of market pressure curtailing operations, losses can increase in the absence of official action closing the bank.[1]

It must be stressed that there is a strong tendency for a bank's losses, once they are large enough to make the bank insolvent, not only to continue, but also to accelerate. This tendency is due, in part, to the banker's attitude toward risk. Modern finance theory assumes that most participants in financial markets are risk averters. The behavior of financial markets is consistent with this assumption, and it is even more likely that bankers are personally risk averse. Decisions with respect to risk, however, are influenced by financial position. As Jensen and Meckling (1976, p. 334) point out, when the firm has little capital "the owner-manager will have a strong incentive to engage in activities which promise very high payoffs if successful even if they have a very low probability of success. If they turn out well, he captures most of the gains; if they turn out badly, the creditors bear most of the costs."

A more drastic change in attitude occurs when one's net worth is negative. Assume one has a choice between an investment opportunity that offers an expected return of 10 with a standard deviation of 0, and an alternative investment with an expected return of 8 and a standard deviation of 20. Nearly all solvent bankers and others will choose the first. The investor (banker) with a net worth of −12 is likely to choose the second, however. In fact, it would be irrational to do otherwise, as the first offers no hope of a positive outcome.

This is not merely a theoretical possibility. Bank supervisors have long noted the tendency of banks in financial difficulty to take what they view as excessive risks. During the early 1970's supervisors complained about the practice of banks buying deposits linked to out-of-area and usually risky loans. Losses on these loans found in a number of failed banks were large, and supervisors tried to stop the prac-

---

[1] The existence of uninsured depositors may provide a market pressure leading to runs on banks suspected of being insolvent. This does not seem to be the case, even in large banks with sizable uninsured deposits. It appears that even uninsured depositors expect to be protected in the case of failure. There is certainly basis for this expectation, in that all large bank failures have been handled by the FDIC through a purchase and assumption transaction that has resulted in de facto 100 percent deposit insurance. Moreover, many large depositors also have large loans from the bank, and hence through the right of offset are protected from loss of their deposit in case of failure.

56-368 O - 80 - 16

tice in the belief that it was responsible for a number of bank failures. It seems more likely that the bank managements involved were neither stupid nor dishonest in acquiring these high risk assets. They were more likely an attempt to recoup once the bank was already insolvent. They were, then, a result, rather than a cause of the insolvency.

Because of this tendency for an insolvency to involve a greater and greater loss over time, it is important to the insuring agency that insolvencies be detected promptly. The detection of insolvency before the loss becomes larger is the essential function of bank examination. A bank failure does not necessarily involve any significant loss to the FDIC (or to any depositor or general creditor), but if the insolvency goes undetected or unstopped for long, a large loss is likely.

A traditional bank examination, involving a detailed evaluation of assets, is the best means of determining insolvency, even if it cannot prevent insolvency. With deposit insurance, banks typically do not close because of runs or liquidity problems, and hence a technically insolvent bank can continue to operate almost indefinitely. It takes an evaluation of assets to determine that the bank's capital has been exhausted. It is this evaluation of assets that the bank examiner has been doing, and no other means of assessing solvency appears workable. In a world without deposit insurance, some large depositors would conduct their own examination. In the present situation, it is hard to imagine a flow of financial information which could replace the actual loan evaluation of the examination. Banks report market values of securities, and this enables the supervisor or depositor to determine the decline in net worth due to market depreciation. But actual insolvency almost always involves some loss on loans that would not show up on financial reports for a long time (it is easy to rewrite or extend maturity of loans not current to postpone writeoff.

It is important to be clear about the nature of risk to the FDIC. Loss to the FDIC is not a function of the number of bank failures or of the size of the failing banks, but rather is related to the amount of loss on failures. This is relevant to proposals for variable rate deposit insurance premiums, based on risk.[2] Correctly set insurance premiums further the efficient allocation of resources, but setting premiums requires an ability to measure risk. While it is sometimes argued that measurement of risk is the function of the bank examiner, in fact, all the examiner must do is determine whether failure has already occurred, not measure the risk of its happening in the future. In any case, it is not even clear how to define this risk. "Risk of failure" is not the appropriate measure of risk to the insurance fund, since a bank failure, recognized promptly, does not result in loss to the insurance fund. Since the FDIC's ability to recognize a failure is not a function of the riskiness of a bank's portfolio, a uniform premium structure seems appropriate.

This approach to the role of examination also means that it is not necessary for the supervisory agencies to seek lower bank risk. If bank failure is not traumatic to the economy or to the community, if depositors are protected by insurance, and if promptly recognized bank

[2] For analysis of this issue, see Mayer (1965), Shapiro and White (1965). A somewhat different approach is suggested in Barnett, Horvitz, and Silverberg (1977).

failure does not impose great costs on the insurance system, it follows that bankers can be allowed to take risks that they deem appropriate.

There is now an elaborate structure of bank regulation. The analysis above suggests that most of this regulation is superfluous, provided that there is a system of deposit insurance and that banks are operating with their own capital. A capital requirement, then, may represent sufficient bank regulation.

This argument means that many restrictions on banks are unnecessary in terms of the traditional purposes of bank regulation. Those regulations aimed at reducing the risk of bank failure by reducing competition in banking are particularly suspect—even if it were desirable to reduce the risk of bank failure, reducing the risk of failure of particular institutions by reducing the degree of competition in the industry as a whole is certainly inefficient. But even regulations designed to prevent banks engaging in risky activities seem unnecessary unless they expose the bank to losses which are sudden and large relative to the bank's capital.

It is possible to generalize in this way, and argue that many regulations can be eliminated because they are unnecessary. It is a primary purpose of this paper, however, to examine specific regulations in detail, and assess the evidence on the costs imposed by their existence. It remains true, however, as Meltzer (1967, p. 482) noted several years ago, that "the net effect of governmental laws and decisions on the volume of assets invested in financial institutions is difficult to calculate. The net effect on resource allocation of these restrictions and tax shelters is unknown also."

The following section of the paper considers numerous examples of four types of regulations aimed at promoting bank soundness: Restrictions on structure, activities of banks, interest rate contracts, and balance sheet constraints. Also considered will be the recent regulations aimed at consumer protection.

IV. Impacts of Bank Regulation

A. Bank Structure

Changes in the banking structure are restricted in several ways. The most significant are the restrictions on the chartering of new banks and on branching by existing banks. Both of these restrictions ostensibly have two purposes in reducing the risk of bank failure. First, restricting entry by a new bank or expansion of an existing bank reduces the risk that the institution itself will fail. It is possible that some organizers of new banks may misjudge the market or their capabilities, and may fail. If the agency that must approve the charter application has better information or analytical capability than the organizers who have their capital at stake, it is possible that some charter rejections may prevent failures. It is also possible that a banks may jeopardize its soundness by expanding at a time when it should not. But this is a weak basis for requiring regulatory approval of such expansion. First, there is little basis for believing that the agency's judgment is better than that of the private decisionmakers. Second, failures resulting from such misjudgments are unlikely to be

costly to the insurance fund. In fact, the purpose of requiring regulatory approval is not simply to prevent such mistaken judgments. Convincing the chartering agency that there is no risk that the new bank will fail is a necessary but not a sufficient condition for a charter to be issued. The organizers must demonstrate that there is a need for the new bank (or the branch). The agency must determine that existing institutions will not be harmed by the new competition. It has become common for economists to criticize this approach, arguing that restraining competition is not an effective means of reducing the risk of failure. There is no good evidence on this issue, but it is reasonable to believe that there are instances in which inefficiently operated banks could be endangered if exposed to free market competition. But even if that were so, there is no basis for restricting entry once we accept the view that the objective of regulation is protection of the insurance fund, rather than prevention of failure.

A more extreme restriction than the need for regulatory agency approval of branches is the existence of statutory restrictions or prohibitions on branching. Federal law defers to State law in limiting the branching powers of national banks. All States restrict branching, and some prohibit it completely. These statutory restrictions on branching have always had motives beyond the prevention of failures, though bank safety is a major basis of justification for this type of restriction of competition. Such restrictions are not necessary to the soundness of the deposit insurance fund. The relevant question is what are the costs of retaining such restrictions.

ENTRY RESTRICTION

Several studies attempted to estimate the effect of restrictions on bank entry. The simplest approach to this question is to examine the number of charter applications denied by the Comptroller of the Currency and various State banking supervisors.

Such tabulations do not provide a meaningful measure of the effect of the restrictions on bank entry. The number of applications may be higher than if entry were unrestricted because if the application of the first potential entrant is denied, other applications may follow, which would not be contemplated if the first were already in the market. Moreover, some applications represent an attempt to enter a market which will then be protected from further entry. If that protection did not exist, the application might not have been filed. On the other hand, the official number of applications received may be too low because some potential applicants may be discouraged from applying because of a correct or incorrect expectation that their applications would be denied. More important, some chartering authorities may allow an applicant to withdraw his application on the basis of an informal rejection.

A theoretically sounder basis for calculating the impact of entry restrictions was put forth by Peltzman (1965). Peltzman investigated the impact on entry into the banking industry of the legal restrictions on entry set forth by the Banking Act of 1935. Specifically, his procedure involved a comparison of the actual bank entry rates in the period not subject to entry regulation (pre-1935) with these rates for

the period experiencing strict regulation on entry (post-1935). Using annual bank data for the period 1921–62, Peltzman developed a regression model to predict the level of bank entry using bank industry characteristics which should influence the rate of entry into banking. The model hypothesized entry, in the absence of regulatory restrictions, to be a function of the expected rate of return on invested bank capital (lagged 1 year), the amount of invested capital, deposit size (net of cash assets), risk of capital loss due to bank failure, the intended change in the average level of invested capital per bank, the annual rate of bank mergers, and the annual rate of bank failures. The model included a dummy variable signaling the existence of regulations on entry. The estimated coefficient of this regulatory variable represented Peltzman's measure of regulatory impact as the average percentage point reduction in the rate of entry since 1935 (as compared with the rate from 1921–35). In discussing the resulting value of the regulatory coefficient of minus 0.579, Peltzman concluded that the net impact of entry restrictions has reduced the number of new banks entering the industry by approximately 2,200 banks from 1936 to 1962.

In discussing Peltzman's empirical investigations, Edwards and Edwards (1974), criticized Peltzman for his failure to account for the indirect effects of entry regulation on expectations about the rate of return on bank capital. Using an alternative model which included these effects, the authors concluded that Peltzman had overstated the impact of regulation on new entry by about 45 percent.

If entry restrictions have significantly reduced new entry, that raises the question of what the impacts on the banking system and the public would be if there were a substantial liberalization of entry restrictions. That question has been investigated in several important papers.

Motter (1965) focused on the operating performance of a group of new banks chartered in 1962. These 64 newly organized national banks, the first set of new banks to be created during the 1962–64 period, were distinguished by a rapid expansion of banking facilities. Several basic ratios were calculated covering such items as rates of return, growth rates for loans and deposits, loan-deposit ratios, time and savings deposits to total deposits, operating income to total assets and operating income to total capital, and effective interest rates.

Motter found that most of the new banks chartered in 1962 enjoyed substantial growth over the study period, and concluded that these banks were indeed filling a need in the communities which they serve. For the 3-year period, the various rate of return measures showed steady increases for the bank group as a whole. Using data from bank examinations, Motter found that in 1964 the assets of the 1962 class of banks were comparable in quality with a sample of older national banks representing the same size classifications.

Although he discounted the significance of impacts of entry on existing banks, Motter did present the performance results of a group of older banks operating in the same markets entered by the new 1962 banks. Attention was directed to those markets which were served by only one or two banks prior to entry. The rate of return for the "single bank" market was more effected by new bank entry as the rate

fell, while the rate of return for the "paired bank" market increased. The impact on deposit growth of new bank entry appeared to be greater for the single banks than for the paired banks. Motter found additional effects of increased competition due to entry which were beneficial from the standpoint of bank customers. After entry, there were increases in convenience factors such as extended banking hours, increased availability of installment and small business loans, and higher interest rates paid on time and savings deposits.

Motter and Carson (1964) investigated the impact of the entry of New York City banks into Nassau County which took place after legislation in 1961 removed legal barriers to such branching into this fast-growing New York county. The authors examined changes in banking structure, economic factors, and bank performance data for Nassau County for 1958–63.

In addition to presenting empirical results of the changes which occurred in the banking structure of Nassau County, Motter and Carson examined pertinent data on bank operations during the 1958–63 expansion period to determine whether entry of New York City banks created a threat to the existing Nassau County banking system. They found no evidence that an "over-banking" situation existed which would cause depressed rates of return. In fact, the rate of return, as measured by after-tax net income to total capital, for both the aggregate data and for individual banks increased over the period for Nassau-based banks and was also consistently higher than the aggregate ratio for all U.S.-insured commercial banks. Similar results on this issue were reported by Chandross (1971).

Motter and Carson also found several other changes which benefited banking customers. The expansion in banking facilities during 1958–63 led to more convenient bank locations and more convenient banking hours. Certain specialized services were made available for the first time. The increased competition in the Nassau loan market resulted in declines in installment loan rates. Interest rates on time and savings deposits increased as the expansion of banking facilities took place.

RESTRICTIONS ON BRANCHING

Restrictions on branching are only partly based on considerations of safety. It appears that historically most such restrictions were based on a concern with preventing a tendency toward concentration of financial power by large banks through widespread branching. Related to this, of course, is concern that smaller institutions would not be able to compete successfully with the large branch banks, and thus branching could lead to failure of the smaller institutions.

The social cost of such restrictions lies in part in the fact that branching provides a greater availability of banking facilities (Horvitz and Shull, 1964). This might be offset if restricting branching led to a more competitive structure. Gilbert and Longbrake (1973, 1974), in a useful survey of the literature, conclude that financial resources of commercial banks, savings and loan associations, and mutual savings banks, were relatively more concentrated in those States which permit some form of branching than in States which prohibit branching.

They found branch institutions provided more credit and other services than did unit banks and thrift institutions. Moreover, the unit institutions extended more credit when faced with branch competition.

Several studies have examined the effects of branching on productive efficiency and costs. Studies by Bell and Murphy (1968), Benston (1969), Brigham and Pettit (1969), and Longbrake (1972, 1973) lead to the conclusion that branch institutions can achieve lower costs of production for services which can be "centralized in the main office." Such services include business loans, real estate loans, and securities investment. Unit banks and savings and loans realized lower costs in producing more decentralized services such as deposits, safe deposit boxes, and installment loans. Gilbert and Longbrake, in summarizing the evidence, conclude that profitability is not significantly affected by type of organizational structure.

CONCENTRATION AND COMPETITION

Restrictions on entry and on branching attempt to protect bank solvency by preventing "excess" competition. If banks are shielded from competition by aggressive and efficient new entrants, they are less likely to fail. There is little evidence available to support the view that reductions of entry reduce bank failure. Most bank failures do not result in any way from excessive competition, but, on the other hand, because entry is restricted, such conditions have generally not been allowed to develop. The 1965 Motter paper does provide some analysis of this by dealing with the situation in which the barriers to entry were significantly reduced. Only one of the banks included in Motter's analysis failed during the period of the study and there are no indications that other banks competing with these new entrants failed. Likewise, the evidence does not suggest that failure is more common in States where branching is allowed than where it is not. In fact, when banks get into difficulty, it is considerably easier for the FDIC to arrange a takeover that preserves some of the going concern value of the failed institution in a branching State than where branching is not allowed. This is because the takeover of a failed bank in a unit banking State requires closing of the office. In a branch banking State, the failed office can be maintained as a branch of the acquiring bank.

These considerations suggest that restrictions on entry and branching designed to reduce competitive pressures on banking and hence preserve safety and soundness are not effective in accomplishing that objective. It remains to be considered whether those restrictions result in any harm to bank customers. Motter and Carson have shown that freer entry can lead to significant benefits to consumers of bank service. In some cases, eased entry means a large difference in concentration in banking markets, e.g., in many banking markets, entry may mean increasing the number of banking alternatives from one to two or two to three. In most markets, however, entry has a relatively small impact on structure in the industry.

There has been a significant body of literature over the last 20 years focused on the issue of whether structure matters. There have been

several useful surveys and summaries of this literature, beginning with Holland (1964), and followed by Smith (1966). Rhoades performed a useful survey in 1977 which considered 39 empirical studies published since 1959 of the relationship between bank market structure and bank performance. Over time, these studies have become more sophisticated both in terms of economic concept and data and statistical methodology. According to Rhoades, "This review finds that market structure clearly affects price and profit performance in commercial banking, although the effect is quantitatively small." This appears to be a fair summary of the studies that have been done. Thirty of the 39 studies reviewed by Rhoades find that in concentrated markets, prices charged to consumers of bank services are higher than in less concentrated markets. But the economic effect is small. For example, Edwards' study (1965) of interest rates on business loans found that a 10-percentage-point increase in the concentration ratio would result in a 6-basis point increase in the average rate on business loans. This result is statistically significant, but some would argue that the economic impact is small. A 20-percentage-point increase in the concentration ratio would mean only an increase of one-eighth of a percentage in the average rate on a business loan. On the other hand, some point out that the $300 billion of business loans outstanding makes this a significant economic effect. There is room for disagreement as to the implications of this research. For example, in a recent unpublished paper, Osborne (1978) argues that the studies done so far do not make a convincing case for the structure performance relationship. The weight of the evidence, however, does seem sufficient to suggest that measures which lead to more concentrated banking structures impose cost on the public. Such measures to be justified must hold promise of reducing the risk of bank failure. Significant evidence that they do so, however, is lacking.

B. Activity Restrictions

Banks are subject to substantial restrictions on the types of activities in which they can engage. For the most part, banks are limited to the business of accepting deposits, making loans, and buying certain fixed-income securities. State and National banks are allowed to engage in direct lease financing transactions, though this is a rather recent development. The economic effects of lease financing are virtually identical with secured lending. The fact that leasing has been a subject of controversy indicates how closely the activities of banks are restricted.

Historically, banks, particularly those in small towns, were allowed to engage in some transactions that were peripherally related to mainstream banking, such as maintaining an insurance or travel agency. Banks have also been allowed to engage in the underwriting of full faith in credit, State, and local government securities (but not of revenue bonds). Beyond this, banks are severely restricted in their ability to engage in the securities business. Banks are allowed to operate trust departments. This is a sizable business, not necessarily related to the accepting of deposits and the making of loans. But there are good economic reasons why banks have a comparative advantage in the performance of trust services.

The activities of banks have been limited for several reasons. The simplest and most straightforward is simply that these nonbanking activities are viewed as riskier than banking and hence the expansion of banks into those other activities would increase the risk of failure. It should be noted that this is not totally convincing since some of these activities may allow a diversification that would reduce this risk of bank failure (Blair and Heggestad, 1978). Another reason for restricting bank activities is concern that banks have excess market power and could be in a position to engage in tied sales or pursue other practices which would take unfair advantage of their unique positions as suppliers of credit. The economic logic to support this argument is weak—if banks have excess market power, they can exploit that power directly in credit markets. A third reason for restricting bank activities is concern about conflicts of interests. There may be some basis for concern about conflict of interest in banking, but the most important already exists with respect to the trust business of banks. In some areas, for example, the potential conflict between an investment banking activity and a providing of investment advice exists for other providers of such services.

A major reason why banks are restricted from entering other activities is the opposition of existing providers of those activities: Travel agents, insurance agents, computer service bureaus, securities dealers, and others who have lobbied and sued to keep banks from entering their fields. These issues became significant during the mid-1960's when James Saxon, then Comptroller of the Currency, attempted to liberalize the restrictions on activities of national banks. The issue really came to the fore, however, with the expansion of the bank holding company movement in the late 1960's and 1970's (Fischer, 1969). The same arguments concerning banking activities have been echoed with respect to the debate over allowable activities for bank holding companies. The issue is one step removed, however, when considering bank holding companies. That is, a bank engaging in a "risky" activity could increase the chance of failure of that bank. If a holding company engages in an activity that is riskier than banking, that may increase the risk of failure of the bank holding company. In general there is no particular reason to be concerned about the failure of a bank holding company. The concern really is that a bank holding company which fails, or gets into financial difficulty, may lead to failure of its subsidiary bank (Chase, 1971; Chase and Mingo, 1975). In a few cases this has occurred. The Beverly Hills National Bank and the Hamilton Bancshares cases suggest that the difficulties of a bank holding company may in fact be transmitted to the subsidiary banks.

Restricting the activities of commercial banks may adversely affect the profit potential of banks that see opportunities to move into other areas. It is unlikely, however, that the public is harmed by these restrictions on banks as long as these other activities can be reasonably competitive. This is certainly the case with respect to the provision of computer services, travel agency, or insurance agency services, for example. On the other hand, there would clearly be a public loss if banks were not allowed to engage in the trust business. In the absence of banks, competition in providing trust services would be limited. In be-

tween these examples, there may be some areas that need closer examination. One such area, and perhaps the only one on which significant empirical analysis has been done is the underwriting of municipal revenue bonds. Studies by Smith (1968) and Cagan (1978) suggest that there would be significant benefits possible to the public if commercial banks were allowed to underwrite revenue bonds. Studies by Mussa (1978) and Kaufman (1979) suggest some weaknesses in the Cagan analysis. While this is still a debatable issue, it appears that there are few other areas in which the restriction of banks from entry into particular areas adversely affect the public inteest. And in those cases, it may be that allowing bank holding companies to enter those businesses would suffice to meet the public needs.

C. Interest Rate Control

Perhaps the most significant public impact of restrictions on banking has to do with interest rate controls. Banks are prohibited from paying interest on demand deposits and are subject to ceilings set by the Federal Reserve and the FDIC on rates that may be paid on time deposits. These restrictions became law in the banking legislation of the 1930's. The presumed intent of the restrictions was to prevent banks from engaging in excessive rate competition, which, it was believed, was responsible for many bank failures during the 1920's. Subsequent historical research, including studies by Linke (1966) and Cox (1967) have demonstrated that rate competition was not a significant factor in the failures of the 1920's or a plausible basis for the restrictive legislation of the 1930's. Their historical works supported the analytical work of Benston (1964). Because interest rates were low during the 1930's and 1940's, these restrictions did not have much impact on the banking system or on the public for the first 20 years or so of their existence. For many years the time deposit ceiling was above the rate that banks were willing to pay on time deposits. The time deposit ceiling began to bind in the 1950's, but whenever it did become binding, the Federal Reserve raised the ceilings to allow banks to pay going market rates. The ceilings first became a serious problem in 1966 when, with market interest rates at record levels, commercial banks were willing to bid aggressively for time deposit business. Many savings and loan associations holding portfolios consisting of mortgages made in earlier years at relatively low rates found themselves unable to meet the rate competition of the commercial banks. The serious problems of the savings and loan industry led to attempts to protect both the savings and loans and the construction industry from the effects of tight money and high interest rates. At that time, interest rate ceilings were lowered on small deposits, but not large CD's. Since 1966, the major focus of interest rate regulation has not been on the question of commercial bank soundness, but on the need to protect savings institutions from commercial bank competition.

There is no doubt that the effectiveness of this type of regulation for this purpose has become less and less over time. Individual depositors have become more sensitive to interest rates and more aware of the ways in which they can invest directly in money market instruments. Moreover, the financial system has innovated in ways to get around

these regulations. The development of money market mutual funds is the most impressive example of the development of a new financial institution aimed primarily at providing a way for small depositors to avoid being penalized by the interest rate ceiling. Economists have generally criticized interest rate controls as being inefficient and inequitable to small depositors. There is a tendency to gloss over the fact that elimination of interest rate controls would in fact probably lead to increased failures. It is likely that in the absence of interest rate ceilings, some banks, presumably less well managed banks, would find themselves in financial difficulty as a result of making poor decisions with respect to the rates paid on time deposits. One illustration of this is the brief period during which ceilings were eliminated on 4-year certificates of deposit. One bank committed itself to high rates on an extensive volume of deposits and that decision played an important role in that bank's ultimate failure. But beyond the question of failures of commercial banks, it is likely that elimination of interest rate controls will lead to some increase in the failures of savings and loan associations. This could result from their inability to pay competitive rates leading to liquidity pressures, or from a commitment to rates higher than can be sustained. It should be noted that there is a relationship between the restrictions on interest rates and restrictions on entry. In a world of limited entry, a savings and loan may survive even if it is restricted by its portfolio to paying below market rates on time deposits. If entry were free, a savings and loan could enter the market in periods of high interest rates, such as 1979, make loans at competitive mortgage rates, and afford to pay higher than existing rates on deposits. While there have been attempts to measure the adverse effects on the public of existing interest rate ceilings, there has not been any significant effort to measure the possible impacts on the safety and soundness of financial institutions that would result from elimination of controls.

The effectiveness of such ceilings has been eroded over time as the public has adjusted to finding ways to evade the controls. Nevertheless, the impact on the public has been significant. The favorable impact of controls on bank earnings is also substantial. Silverberg (1973) has demonstrated that bank earnings benefit during those periods of time when interest rate ceilings are binding. Estimates of the impact on the public have been made by Kane (1970, 1977) and Pyle (1974).

D. Balance Sheet Constraints

Banks are subject to various restrictions on their balance sheets. The most important are liquidity requirements and capital requirements. The major liquidity requirements are reserve requirements imposed by the Federal Reserve on its members and by the various States on nonmember banks. Several recent studies have been done of the impact of State reserve requirements. These have generally found, though there are some exceptions, that most State reserve requirements are not binding, i.e., the bank would be forced to maintain assets in liquid form to meet day-to-day operating needs even in the absence of the State reserve requirements. The Federal Reserve requirements, which require the bank to hold a specified amount of

assets in the form of vault cash or nonearning deposits with the Federal Reserve bank, are binding and impose a cost burden on member banks. This has led banks to drop Federal Reserve membership and has led to proposals to solve the Federal Reserve membership problem. This is a significant issue, but it is not related to the problem of bank solvency and soundness. It is generally acknowledged now that reserve requirements are related to monetary policy considerations and not to the liquidity of individual banks. Fractional reserves to the extent that they are required cannot provide liquidity to meet deposit outflows. The legally required reserves of a bank are not available in full to meet cash needs since they must be maintained against deposits that remain with the bank.

In terms of bank soundness, capital requirements are the most significant constraint on the balance sheet of the bank. Statutory capital requirements are not a significant factor in that the minimums required for national banks, and under most State laws, are extremely low. The capital constraints that are significant are those imposed as part of the supervisory process by the banking agencies. While a great deal of work has been done on the subject of capital adequacy, there is no real agreement as to how capital adequacy can be determined or measured (Robinson, 1941; Smith and Hengren, 1947). There is agreement that capital adequacy is the most important regulatory means for protecting bank soundness. Thus, Edwards and Scott (1977) argue that "the major thrust of our analysis was to show that through a combination of equity capital and liquidity controls, regulators can successfully control bank solvency." They also argue that bank capital requirements impose the least distortion on resource allocation of any other form of regulatory controls.

In one sense, bank capital does not involve any real social cost. Funds represented by bank capital are also available to be used for other purposes. To the extent that the capital of a steel company, for example, is used to provide physical facilities, it absorbs real resources which are not available for other uses. An increase in bank capital does not diminish the availability of real resources for other uses in the economy (of course, a return must be earned on the bank capital and that has implications for the structure of interest rates through the economy). Studies have been done of the effectiveness of bank capital by analyzing the relationship between capital of failed and surviving banks (Cotter, 1966), and the relationship with market prices (Pettway, 1976). The difficulties with this type of research are involved in the measurement of bank capital. Book capital is not a good proxy for real capital of banks under stress. Even so, capital variables do turn out to be significant in the early warning studies of Sinkey (1975) and others (e.g., Korobow and Stuhr, 1975).

The analysis of the impact and effectiveness of bank solvency regulations can be summarized rather briefly. There are legitimate bases for government concern with the safety and soundness of financial institutions. These concerns are met with a credible system of Federal deposit insurance. With Federal deposit insurance, the need for regulation becomes solely a need to protect the deposit insurance system. But only modest restrictions on banks are necessary for protection of the deposit insurance system. A capital constraint and a means of col-

lecting information on banks, presumably bank examination, is sufficient to protect the deposit insurance system against significant loss. Most existing regulations, such as entry, interest rate controls, and restrictions on bank activities do have adverse efficiency implications, but the magnitude of these effects is relatively small.

There is substantial agreement in the economic and finance literature that the existing regulatory framework involves overregulation of financial institutions. Some of this is philosophical. That is, there are some who would favor reduced regulation of financial institutions in the absence of substantial evidence indicating benefits from regulation. But on the basis of evidence that exists, it is clear that reducing the extent of regulation would not impose substantial costs on the public. The risks are to the deposit insurance system, which, it can be demonstrated, can be protected with less than the existing level of regulation. The evidence then indicates rather strongly that there are net costs associated with the current regulatory structure that exceed the public benefits. Because the market is rather ingenious in finding ways around many regulatory constraints, these social costs are perhaps less extensive than some analysts have argued.

E. Consumer Protection Regulation

A relatively new development in the regulation of financial institutions is the passage of legislation and imposition of regulations designed to provide consumer protection in the credit process. The Truth in Lending Act, passed in 1968, was the first of these measures, but it was followed by the Fair Credit Reporting Act, the Fair Credit Billing Act, the Fair Collection Procedures Act, Elimination of the Holder in Due Course Doctrine, the Equal Credit Opportunity Act, and the Community Reinvestment Act, and the Home Mortgage Disclosure Act. All of this legislation was aimed at real and documented abuses in the credit process. Though there is no real evidence of the extent of such abuses, the financial institutions affected have argued that the legislation involves an overreaction to small numbers of abuses and that the costs of complying with the regulation impose costs that are greater than the benefits to consumers. Some individual institutions have made estimates of the cost of complying with Federal regulations and conclude that the costs are on the order of 10 percent of net income. These estimates, however, include the cost of reporting to the Federal Government. Because of the role of banks in the money supply process and the importance of monetary policy, it is clear that some extensive amount of statistical reporting by banks is necessary for that purpose as well as for solvency purposes. While the current reporting burden may be in excess of the optimal, it can certainly not be eliminated. Likewise, these estimates include the cost of bank examination. This is a major regulatory cost and, as we have seen, bank examination plays an important role in protecting the solvency of the insurance system.

The legitimate objections of financial institutions concern whether the cost of the recent consumer protection legislation exceeds benefits to consumers. That is a difficult area on which there has not been much research. Some studies indicate that the Truth in Lending Act, the oldest of these pieces of legislation, has resulted in consumers being

better informed as to the interest rates being paid on loans. It is not completely clear, however, whether this information has improved the competitiveness of consumer credit markets or whether consumers make use of this information in shopping for credit. The cost and benefits of consumer protection legislation is an area in which limited research has been done up to this time and in which more research is clearly needed. Attempts at quantifying these costs and benefits have been made by Shay and Sexton (1979) and Shay and Brandt (1979).

V. Credit Allocation

This paper has so far focused on the role of regulation in protecting the safety and soundness of financial institutions. That has certainly been the basis for the historical development of most of our regulatory structure. In some cases the regulation is aimed at improving the efficiency of the allocation of resources; in other cases, the regulation has unintended and undesired effects on the allocation of resources. It is not surprising that there have been many attempts to harness the regulation of financial institutions to produce desired effects on the allocation of resources. In some cases there have been changes in the stated justifications of regulation from one purpose to another, without any actual change in the regulation. Interest rate ceilings on deposits were originally supported on the basis of protecting the solvency of commercial banks against failure due to excessive competition. We still have the same regulation but the basis now is assistance to the housing market rather than protection of banks.

Credit allocation involves a variety of techniques designed to influence lenders and borrowers to discriminate among potential transactions in governmentally approved ways. As Kane indicates (1975, p. 16), "These techniques range from explicit taxes and subsidies to interest rate ceilings, reserve and collateral requirements, government loans and guarantees of repayment, quantitative limits on the size of different types of equity and debt issue, and diverse varieties of subtle and not-so-subtle government strong-arm pressures on would-be lenders and borrowers." [3]

Credit allocation programs in the United States have had a variety of intended beneficiaries—small business, consumers, the Treasury, State and local governments, and, most important, the housing industry. In recent years the housing market has been the major target of attempts to affect the flow of credit by means of regulation of financial institutions. These programs result from an acceptance of the social objective or promoting adequate housing and, more specifically, of promoting home ownership. Congress has acted in accord with the view that in the absence of government intervention, these objectives will not be achieved and, further, that other government measures to promote home ownership would be ineffective or politically unacceptable (too costly). Some support for government intervention on efficiency grounds is based on the view that housing construction, without government intervention, would suffer from such great cyclical instability that housing costs would be significantly higher than if production

[3] Descriptions of various techniques can be found in Davis (1971), Mayer (1972), and Silber (1973).

were more stable, and these higher costs (arising from the need to shift resources into and out of the industry) would contribute to inflation.

The most important approach taken to support the housing industry is the attempt to insulate housing finance from the rest of the financial market by the creation of specialized lending institutions whose access to funds is to be protected by various devices. Savings and loan associations are, for all practical purposes, limited to making residential mortgage loans. For this approach to be successful, savings and loans must be assured a continued access to funds at relatively low rates. In periods of high interest rates, a portfolio of fixed-rate mortgage loans will yield less than the shorter term assets that comprise the portfolios of commercial banks, and savings and loans will be unable to compete for funds. We have dealt with this situation by setting ceilings on the rates that commercial banks can pay on deposits, thus limiting their ability to draw funds from the savings institutions. Since large depositors had easy access to a variety of money market instruments not subject to rate ceilings, it became clear that ceilings on large deposits did not accomplish any useful purpose, and they were dropped in 1973. Over time even smaller depositors found ways around the interest ceilings, leading to further government countermeasures. For example, the minimum denomination of Treasury bills was raised from $1,000 to $10,000. During 1978 the regulatory agencies authorized depository institutions to offer a deposit instrument, with a $10,000 minimum denomination, at rates tied to the Treasury bill rate. On this instrument, as on all other deposits, the thrift institutions were allowed a one-fourth percent differential over the rate that could be offered by commercial banks (though now the differential is dropped if rates exceed 9 percent). These developments, as well as other market innovations such as the growth of money market mutual funds, limited the ability of the regulatory process to assure savings and loans a steady flow of deposits at below market interest rates.

These attempts to provide support for the housing market are financially oriented. They attempt to influence the availability and cost of mortgage credit.[4] They suffer from the fact that credit is fungible. According to Diller (1979, p. 315), "It is not possible—in theory or in practice—to isolate segments of the credit market for the purpose of changing the balance of supply and demand forces in these segments. Fungibility permits investors to look beyond the superficial differences among securities in pursuit of the basic goals of high yields and low risk * * *."

Usury laws may attempt to hold down interest rates in mortgage markets. Savings and loan associations have little choice but to put the funds they are able to obtain into the mortgage market, but other lenders are free to more of the mortgage market if yields on mortgages are lower than other comparable market instruments. Mutual savings banks, insurance companies, pension funds, and commercial banks can play a large or small role in the residential mortgage mar-

[4] This is also true of other forms of government involvement in the market, such as usury ceilings, and purchases of mortgages by Federal National Mortgage Association, Government National Mortgage Association, and the Federal Home Loan Mortgage Corporation.

ket depending upon relative yields. Thus, policies that keep mortgage rates lower than they otherwise would be almost certainly lead to a smaller supply of mortgage credit than would otherwise be the case.

Another aspect of the fungibility of credit is that it is easy to break the link between the mortgage market and housing construction. Meltzer (1975, p. 128), after describing the changes in mortgage markets that have taken place since the Great Depression, concludes that "there is no evidence that the considerable increase in the relative use of mortgages has had any effect on the distribution between housing and other assets." Meltzer found that the rise in the ratio of mortgage debt to total liabilities is largely the result of a substitution for security credit. This trend has become more pronounced during the 1960's and 1970's. That is, to the extent that the cost and availability of mortgage credit, as a result of government regulation and intervention, has been held below its equilibrium level with respect to other financial instruments, individuals have tapped the mortgage market to finance nonhousing related expenditures. This paper is not the place for consideration of whether such a development is or is not desirable; it is sufficient to note here that these considerations suggest that use of process or regulation of financial markets or financial instiutions to influence the allocation of real resources is likely to be ineffectual or at least inefficient. Thus Kane (1977, p. 67) finds that borrowers are able "to relabel debt contracts and to substitute . . . less efficient unregulated . . . forms of credit for regulated ones." An empirical analysis by Hamburger and Zwick (1979, p. 109) concludes that their evidence "supports Kane's position and casts doubt on previous tests, which suggest that credit allocation schemes could play a role in altering the composition of consumer expenditures."

VI. Impacts on the Macroeconomy

The previous sections of this paper dealt with the impact of regulation of financial institutions on competition and the allocation of resources. The emphasis was on impacts in particular markets. It is appropriate to conclude with analysis of the impact of these regulations on the economy as a whole—that is, does regulation of financial institutions have effects that are significant in terms of GNP or its major components?

Much of the discussion in previous sections concluded that even where the direction of the effect of regulation is clear, the magnitudes are not large. An important example of this is the relationship between market structure and business loan rates of commercial banks. Most of the empirical studies reviewed find a positive association between concentration in local markets and the level of interest rates on business loans. This relationship meets the usual tests of satistical significance, but the magnitude of the relationship—the economic significance—is rather small.

In some cases, the magnitude of the effects of regulation is small because of the ability of the financial system to find efficient and innovative means of avoiding the effects of regulation. Controls on the rates paid on time deposits of banks and thrift institutions have a smaller impact on consumers than might be expected because of the

development and growth of money market mutual funds. Likewise, credit unions, subject to less retrictive ceilings, have grown more rapidly in recent years than commercial banks, saving banks, and savings and loan associations. Such a development may distort the allocation of resources, as has been discussed, but the point is simply that the adverse effect on rates earned by consumers is not as great as the spread between market rates and regulation Q ceiling would suggest.

The prohibition of interest on demand deposits provides a similar illustration. The law apparently requires that bank customers receive a zero return on their transaction balances, yet the effect is not as great as that. Banks have competed by providing services free or below cost to demand deposit customers. Depositors have learned to economize on demand deposit holdings, partly because of the development of substitute transaction accounts or devices not subject to the interest prohibition. Again, these results have meant a distortion in the allocation of resources, but they have tended to minimize the magnitude of the effect of the interest prohibition on the public and the economy.

There are at least three ways in which regulation of financial institutions have an impact on the macroeconomy: First is the deadweight cost of administering and enforcing the regulation; second effects on saving in the economy; and third, effects on real investment—either the amount or its distribution. Each will be considered.

The cost of financial regulation is not trivial, as the combined budgets of the FDIC, the Federal Reserve System, the the Comptroller of the Currency totaled nearly $1 billion in 1978. In addition, States spend a substantial sum on financial regulation. Of equal or greater magnitude is the cost to the financial institutions of compliance with the regulations. But these figures exaggerate the issue. There is the possibility, under a different regulatory system, of achieving a significant reduction in cost. The opportunity for reducing costs is much less than the figures cited suggest. Much of the Federal Reserve System costs represent those associated with the conduct of monetary policy, and the same is true of some of the costs of bank compliance with regulations. Regardless of what is done about financial regulations, monetary policy activities must continue. Another $400 million of Federal Reserve costs is accounted for by services provided to member banks. If the Federal Reserve ceased to provide such services, they would be provided by the private sector, and perhaps at lower cost. The saving in social costs, however, would be at best a small fraction of that $400 million. The major element in the costs of the other regulatory agencies is due to examination of the financial institutions. It has already been noted that even under a reduced regulatory framework aimed only at protecting the deposit insurance systems, examination (or some other costly information gathering system) would still be necessary.

Any saving of an unnecessary cost is desirable, and there clearly are costs involved in financial regulation which are not necessary. But the magnitude of such saving is not likely to be significant in terms of GNP or total government expenditures. This does not mean that a reduction in the expenditures of government agencies for regulatory purposes is impossible or unnecessary—only that the effects of doing so will be relatively small in an economy with a GNP of $2 trillion.

56-368 O - 80 - 17

The principal means by which regulation of financial institutions may affect the aggregate volume of saving in the economy is through an impact on interest rates. Restrictions on rates payable on deposits are a significant form of regulation in the United States. These regulations, however, seem to be becoming less effective over time. For many years, rate ceilings applied to all commercial bank deposits. In 1973 deposits of over $100,000 were freed from rate ceilings. In the aggregate, these deposits account for a large portion of savings in the economy. In 1978 banks and savings institutions were allowed to offer approximately market rates on short-term deposits of $10,000. At this point, therefore, it is only small savers who are not allowed to receive market rates of return on savings held with regulated savings institutions. This is a serious equity problem, of course, but it is not likely that it significantly affects the total volume of saving in the economy.

There is a great deal of inertia in many financial decisions, and many depositors continue to hold deposits paying 5 percent or 5¼ percent in commercial banks when they could receive substantially more in, say, money market mutual funds. But the existence of such opportunities simply means that these depositors are not interest sensitive, and hence their decisions with respect to saving versus consumption are not likely to be affected by the existence of interest rate ceilings.

Decisions with respect to the form in which savings are held are much more likely to be sensitive to interest rates than the decision to save rather than spend on consumer goods. Theoretical considerations have long suggested, in fact, that it may be rational for saving to be a negative function of interest rates—accumulation of a fixed capital sum requires less annual saving if interest rates are high than if they are low. In any case, it has long been standard in economic analysis to assume that the interest elasticity of saving is near zero. This is the case in all current large-scale economic models.

Existing regulations affect the convenience of financial institutions. It is possible that convenience of places to save affects the decisions of potential savers. Restrictions on branching and on entering do adversely affect the convenience of financial institutions (though on the other hand, restrictions on interest rates lead existing institutions to compete more on the basis of convenience). The lack of convenient places and forms in which to save has been cited in the literature as a significant cause of low saving rates in some less-developed countries. It would be hard, however, to view this as a significant factor in the United States.

A more important factor than rate or convenience in saving decisions is likely to be the safety of the form in which savings are held. If regulation contributes to the existence of a safe financial asset, it is likely to result in a higher rate of saving.

A reasonable conclusion from this analysis is that the rate of saving in the U.S. economy is not significantly different from what it would be if there were an optimal system or regulation.

The impacts of regulation of financial institutions on investment spending in the economy are likely to be more significant than on saving. Much of our regulatory structure is designed to affect the allocation of credit. It appears that regulations do affect the allocation of

credit. But it is not clear that this effect on credit produces a similar effect on real resources. Hamburger and Zwick (1977 Ap. 1566) explored the relationship between controls of credit and resource allocation. They conclude that borrowers 'do not select particular assets according to the avaliability of particular liablities. . . . A decrease in the availability of a particular form of credit will encourge the creation of new forms of credit and new uses of existing forms of credit to finance the existing demand for assets. . . . The implication is clear: Credit controls do not affect expenditures."

These results are questioned by Thurow (1979, p. 380). Thurow argues that credit controls would be unnecessary if credit markets were perfectly competitive. In such a world, "social priorities are quite properly left to the government budget." In his view, however, "credit markets are not now neutral." The principal imperfection he points out has to do with the corporation's ability to avoid financial markets by retaining earnings.

This comes close to consideration of tax policy. There is little doubt but that tax and subsidy policies can significantly affect not only the allocation of credit but also the allocation of real resources. It is clear that the tax exemption of interest on State and local government securities results in such entities having access to a greater volume of resources than would be the case in the absence of such tax treatment. More resources are devoted to housing than would be the case if mortgage interest and real estate taxes were not deductible from personal income for tax purposes. Such considerations do not suggest that regulation of financial institutions has a significant impact on the allocation of real resources.

The existence of capital markets and a system of financial institutions facilities investment. It is conceivable that our system of financial markets would not be viable in the absence of regulation. If such were the case, it would be reasonable to conclude that regulation promotes investment. But the amount of regulation necessary to allow the development of financial institutions is less than now exists. Regulation may, on the margin, affect the interest rate on various credit instruments, though the evidence of the impact of concentration on loan rates does not provide a great deal of support for this view. Even if there is an effect, most studies of the interest elasticity of investment spending imply that the impact of regulation on investment would be small.

Because of the importance of real investment to the economy, it is important to be clear about this conclusion. Our financial system would be different if there were no government regulation of financial institutions. It would be an interesting exercise (though not one undertaken here) to speculate on what such a financial system would look like. Such an exercise is irrelevant to the real issues, however, since no one, even the most ardent supporters of free markets and minimal government intervention, advocates total elimination of government regulation of financial institutions. The relevant comparison is between investment under our existing regulatory framework and what it would be under feasible alternative systems. It is in terms of that comparison that I conclude the impact of regulation on investment is small.

VII. Policy Implications

The conclusion of the previous section that the effects of regulation of financial institutions on GNP savings and investment are small does not mean that the system is perfect, or that changes are unnecessary. The borrower in a noncompetitive banking market will pay a higher rate on a loan than he would if freer entry made that market more competitive. A change in the current situation can be supported on both equity and efficiency grounds, even if it is acknowledged that GNP will not be noticeably higher as a result. Likewise, the small depositor receiving 5¼ percent on his commercial bank savings account is being treated inequitably. This is inefficient also in that he would like to earn a higher return and the bank would like to pay a higher rate on such accounts. Elimination of the ceiling would make both parties to the transaction better off.

In terms of equity, certainly, and efficiency, probably, interest rate controls are the regulatory feature most in need of change. Virtually all existing studies indicate that consumers would gain significantly and the housing industry would not suffer substantial loss. Some banks would gain, while others would lose. Only the savings and loan industry (or portions of it) would be losers. It is hard to see what claim they have for such continued protection from competitive forces.[5]

The policy implications of this paper flow from the conclusion that some regulation of financial institutions is necessary, but with Federal deposit insurance the only appropriate regulations are those necessary to maintain the viability of the deposit insurance system. The analysis of this paper suggests that a capital requirement (and a means of seeing that it is maintained) is sufficient, and their elimination would make a significant combination to welfare.

Although the United States has a dual banking system, these issues can be dealt with solely at the Federal level. If National bank charters are available freely, restrictions on State charters become academic and would soon be eliminated or eased. Scott (1979) argued persuasively that Congress has the power to allow branching for national banks (even on an interstate basis) regardless of State law. Such action by Congress would also resolve the now troublesome issue of electronic funds transfer systems, including unmanned facilities and part-of-sale terminals.

These policy steps are easy to support because protecting the soundness of the banking system through restriction on competition is clearly inefficient. This paper has questioned, but not resolved, whether other safety and soundness regulations are necessary. Research should be encouraged to determine whether such restrictions as loan limits, fixed asset limitation, and other portfolio and activity constraints are necessary to protect the deposit insurance system. The focus on this analysis should be on whether such restrictions reduce potential issues to the insurance fund—not whether they reduce the risk of bank failure.

---

[5] The savings and loan industry recognizes this point and argues for continuation of rate controls not in terms of protection for themselves but in terms of protection to the housing market (and homebuyers).

References

Barnett, R., P. Horvitz, and S. Silverberg, "Deposit Insurance: The Present System and some Alternatives," Banking Law Journal, 1977.

Bell, F., and N. Murphy, "Costs in Commercial Banking: A Quantitative Analysis of Bank Behavior and Its Relation to Bank Regulation," Federal Reserve Bank of Boston, 1968.

Benston, G., "Interest Payments on Demand Deposits and Bank Investment Behavior," Journal of Political Economy, October 1964.

———, "Cost of Operation and Economies of Scale in Savings and Loan Associations," Study of the Savings and Loan Industry, vol. II, Federal Home Loan Bank Board, 1969.

Blair, R., and A. Heggestad, "Bank Portifolio Regulation and the Probability of Bank Failure," Journal of Money, Credit and Banking, 1978.

Brigham, E., and R. Pettit, "Effects of Structure on Performance in the Savings and Loan Industry," Study of the Savings and Loan Industry, vol. III, Federal Home Loan Bank Board, 1969.

Carson, D., "Is the Federal Reserve System Really Necessary?" Journal of Finance, 1964.

Chandress, R., "The Impact of Entry on Unit Banks in One Bank Towns," Journal of Banking Review, 1971.

Chase, S., "The Bank Holding Company as a Device for Sheltering Banks From Risk," Conference on Bank Structure and Competition, 1971.

Chase, S., and J. Mingo, "The Regulation of Bank Holding Companies," Journal of Finance, 1975.

Colantoni, C., O. Davis, and M. Swaminuthan, "Imperfect Consumers and Welfare Concerning Information and Regulation," Bell Journal of Economics and Management, 1976.

Cotter, R., "Capital Ratios and Capital Adequacy," National Banking Review, 1966.

Cox, A., "Regulation of Interest on Deposits: A Historical Review," Journal of Finance, 1967.

Davis, R., "An Analysis of Quantitative Credit Controls and Related Devices," Brookings Papers on Economic Activity, 1971.

Diller, S., "Credit Allocation and Housing," unpublished, 1978.

Edwards, F., "The Banking Competition Controversy," National Banking Review, 1965.

Edwards, F. and L., "Measuring the Effectiveness of Regulation: The Case of Bank Entry," Journal of Law and Economics, 1974.

Edwards, F., and J. Scott, "Regulating the Solvency of Depository Institutions," Research Paper No. 210, Graduate School of Business, Columbia University, 1977.

Fischer, G., "Market Extension by Bank Holding Companies: History, Economic Implications and Current Issues," Conference on Bank Structure and Competition, 1969.

Friedman, M., and A. Schwartz, "A Monetary History of the United States", Princeton, N.J.: Princeton University Press, 1963.

Gibson, W., "Deposit Insurance in the U.S.: Evaluation and Reform," Journal of Financial and Quantitative Analysis, 1972.

Gilbert, G., and W. Longbrake, "The Effects of Branching by Financial Institutions on Competition, Efficiency, and Stability: An Examination of the Evidence," Journal of Bank Research, 1973, 1974.

Hamburger, M., and B. Zwick, "Installment Credit Controls, Consumer Expeditures and the Allocation of Real Resources," Journal of Finance, 1977.

———, "The Efficacy of Selective Credit Policies," Journal of Money, Credit and Banking, 1979.

Holland, R., "Research Into Banking Structure and Competition," Federal Reserve Bulletin, 1964.

Horvitz, P., and B. Shull, "Branch Banking and the Structure of Competition," National Banking Review, 1964.

Jensen, M., and W. Meckling, "Agency Costs and the Theory of the Firm," Journal of Financial Economics, 1976.

Kane, E., "Good Intentions and Unintended Evil: The Case Against Selective Credit Allocation," Journal of Money, Credit and Banking, 1977.

———, "Short-Changing the Small Saver," Journal of Money, Credit and Banking, 1970.

———, "Deposit-Interest Ceilings and Sectoral Shortages of Credit," Government Credit Allocation, San Francisco: The Institute for Contemporary Studies, 1975.
Korobow, L., and D. Stuhr, "Toward an Early Warning of Changes in a Bank's Financial Condition," Monthly Review, Federal Reserve Bank of New York, 1975.
Kurtz, R., and J. Sinkey, "Bank Disclosure Policy, Adverse Publicity and Bank Deposit Flows," Journal of Bank Research, 1978.
Linke, C., "The Evolution of Interest Rate Regulation on Commercial Bank Deposits in the U.S.," National Banking Review, 1966.
Longbrake, W., "Productive Efficiency in Commercial Banking: The Impact of Legal Form of Organiaztion and Size," Federal Deposit Insurance Corporation, 1973.
Mayer, T., "A Graduated Deposit Insurance Plan," Review of Economics and Statistics, 1965.
———, "Financial Guidelines and Credit Control," Journal of Money, Credit and Banking, 1972.
———, "Should Large Banks Be Allowed To Fail?" Journal of Financial and Quantitative Analysis, 1975.
Meltzer, A., "Credit Availability and Economic Decisions," Government credit allocation, San Francisco: the Institute for Contemporary Studies, 1975.
———, "Major Issues of the Regulation of Financial Institutions," Journal of Political Economy, 1967.
Motter, D., "Bank Formation and the Public Interest," National Banking Review, 1965.
Motter, D., and D. Carson, "Bank Entry and the Public Interest: A Case Study," National Banking Review, 1964.
Mussa, M., "Competition and Borrowing Costs in the Municipal Revenue Bond Market," unpublished. 1979.
Peltzman, S., "Bank Entry Regulation: Its Impacts and Purpose," National Banking Review, 1965.
———, "Entry in Commercial Banking," Journal of Law and Economics, 1965.
Pettway, R., "Market Tests of Capital Adequacy of Large Commercial Banks," unpublished, 1976.
Pyle, D., "The Losses on Savings Deposits From Interest Rate Regulation," Bell Journal of Economics and Management, 1974.
Rhoades, S., "Structure-Performance Studies in Banking: A Summary and Evaluation," staff economic studies, No. 92, Federal Reserve Board, 1977.
Robinson, R., "The Capital-Deposit Ratio in Banking Supervision," Journal of Political Economy, 1941.
Shapiro, E., and W. White, "A Graduated Deposit Insurance Plan: Some Comments," Review of Economics and Statistics, 1965.
Shay, R., and D. Sexton, "Anti-Discrimination Laws in Consumer Credit Markets," unpublished, 1979.
Shay, R., and W. Brandt, "Consumer Credit Protection," unpublished, 1979.
Silber, W., "Selective Credit Policies: A Survey," Banco Nationale Del Lavore Quarterly Review, 1973.
Sinkey, J., "A Multivariate Statistical Analysis of Problem Banks," Journal of Finance, 1975.
Smith, T., "Research on Banking Structure and Performance," Federal Reserve Bulletin, 1966.
Smith, T., and R. Hengren, "Bank Capital—The Problem Restated," Journal of Political Economy, 1947.
Smith, W., "Commercial Bank Entry Into Revenue Bond Underwriting," Office of the Comptroller of the Currency, 1968.
Thurow, L. C., "Commentary," Issues in Financial Regulation, F. Edwards, ed., New York: McGraw-Hill, Inc., 1979.
Tussing, A., "The Case for Bank Failure," Journal of Law and Economics, 1967.
———, "Meaningful Bank Failure: A Proposal," Journal of Industrial Economics, 1970.

ALTERNATIVES TO REGULATION IN THE HEALTH CARE SECTOR

By Warren Greenberg*

CONTENTS

	Page
Summary	257
I. Introduction	257
II. Regulation in the Health Care Sector	258
III. Alleged Imperfections in the Health Care Sector	259
a. Physician Induced Demand for Medical Care	259
b. Widespread Insurance for Medical Care	260
IV. Injecting Competition in the Health Care Sector Given the Widespread Presence of Insurance	261
a. HMO's as Alternatives to Fee-for-Service	261
b. Use of Copayments, Deductibles, and Indemnity Provisions in Insurance Plans	261
c. Encourage Cost Control Efforts by Third Parties	262
V. Public Policy Initiatives Which Can Encourage Competition in the Health Care Sector	264
VI. Role of Competition in the Health Care Sector	268

Summary

This paper suggests that increased regulation of the health care sector has not succeeded in curbing the rapidly increasing costs of health care. Although the widespread use of insurance in health care makes consumers insensitive for the most part to relative prices, this paper argues that in three areas competition and market forces can help contain rising costs. These three areas are: (1) Competition between HMO's and fee-for-service systems; (2) use of copayments, deductibles, and indemnity provisions in insurance plans; and (3) cost control efforts by third parties. Reforming the current tax system by providing a tax supplement to those who select plans which control costs is suggested as an important way to stimulate competitive forces in health care.

I. Introduction

An important element in the Nation's current inflation is the rate of increase in health care expenditures. Fiscal year 1977's spending for health care was $163 billion, an increase of 12 percent over the previous 12 months. Expenditures for health care now account for 8.8 percent of Gross National Product compared to 8.7 percent in 1976 and 5.2 percent in 1960. Currently, Government spending for health care amounts to $68.4 billion or 42 percent of all health care expenditures, an increase of 13 percent over 1976.[1] Enactment of Government-spon-

*Senior research economist, Center for Metropolitan Planning and Research, the Johns Hopkins University.

[1] See Department of Health, Education and Welfare, "Health Care Financing Administration Notes." May 1978, p. 1.

sored National health insurance may further increase expenditures since more people will likely be more fully covered for health care expenses.

The rapid increase in health care expenditures has prompted a number of responses from Federal, State, and local governments. Nearly all responses have called for expanded regulation of medical care input and output costs, rather than reliance on competition and market forces. This paper will first fully review some of the studies which have evaluated a number of regulatory schemes. Second, it will consider arguments which suggest that competition will not work in health care. Third, the paper provides examples of where competitive forces may already be operative in health care. Finally, the paper concludes by suggesting some alternatives to increased regulation in the health care sector.

II. Regulation in the Health Care Sector

There have been a number of studies of the effects of regulation on costs and input use. Recently, Sloan and Steinwald examined the impact of regulation on hospital capital expenditure and facility expansion, revenue-cost inflation, and utilization review.[2] Using American Hospital Association data for 1,228 hospitals between 1970–75, the authors concluded that "regulatory programs did not do much to contain hospital costs and input expansion during the first half of the 1970's." [3] One reason for regulation's failure, according to the authors, is the "anticipatory" and "compensatory" response of hospitals prior to the implementation of regulations.[4] Hospitals, anticipating certificate-of-need implementation, accelerated their expansion in bed capacity prior to the regulations. Compensatory responses included increases in non-capital costs such as labor inputs. State-enacted certificate-of-need measures attempt to control rising hospital costs by mandating review of changes in an area's physical plant (hospital facilities), hospital equipment, and hospital services. In a major study of the effects of certificate-of-need, Salkever and Bice found that certificate-of-need laws did not significantly reduce total hospital investment or hospital costs. The laws did reduce the growth in the number of beds (which lessens pressure for cost increases) but at the expense of stimulating increases on facilities and equipment.[5]

Feldstein has suggested that "as long as strong financial incentives for higher spending remain, it will be politically impossible to restrain them by government controls.[6] The Administration's hospital cost containment bill which would put a lid on the rate of increase in hospital expenditures, would, according to Feldstein, result in hospital behavior that "can only be described as wasteful distortions that produce the appearance of compliance." [7]

---

[2] See Frank A. Sloan and Bruce Steinwald, "Effects of Regulation on Hospital Costs and Input Use, paper presented at the annual meeting of the American Economic Association, Chicago, Illinois, August 29, 1978.

[3] Sloan and Steinwald, p. 1.

[4] Sloan and Steinwald, p. 1.

[5] David S. Salkever and Thomas W. Bice, "Hospital Certificate-of-Need Controls," Washington, D.C., 1979, esp. pp. 75–83.

[6] Martin S. Feldstein, "Consequences of Hospital Costs." Wall Street Journal, April 12, 1979, p. 24.

[7] Ibid.

Although there has been little supportive evidence that regulatory methods have stemmed the rapid increase in health care costs,[8] there has not been a large-scale movement in public policy to inject additional competitive forces into health care. Many observers have suggested that the health care sector behaves differently from other sectors of the economy, and therefore traditional market forces do not apply. There are two imperfections which are most often cited for the lack of market forces in health care. The first suggests that the asymmetry of information between the providers and services (physicians and hospitals) and the consumers of services allows physicians to create their own demand. Consumer sovereignty does not exist in health care. Second, the widespread presence of insurance in health care causes consumer indifference to prices charged by competing providers. The next section reviews both of these imperfections.

III. Alleged Imperfections in the Health Care Sector

a. Physician Induced Demand for Medical Care

Economic theory assumes that consumers are sovereign in the purchase of goods and services. If physicians, because of the asymmetry of knowledge between the physician and patient, can create their own demand with the consumer as a passive buyer, almost all assumptions of economics are violated. In particular, increases in the supply of physicians and para-physicians may not lower prices but may raise prices.

No economist believes that physicians have unlimited capacity to create their own demand. There are costs of time and sometimes pain that all patients must bear. On the other hand, even the most skeptical of the physician-induced-demand-hypothesis school do not deny that, at times, physicians have induced demand beyond what a rational consumer would consume. The debate on this controversy centers on the extent that physicians can create their own demand.

The statistical evidence has not been unambiguously clear on this subject. Some researchers, for example, have pointed to the observed correlations between the number of physicians and the utilization of physician services, while others have focused on the observed correlations between increased physician supply and higher prices for medical services.[9] It is possible, however, for the number of physicians and the utilization of physician services to be correlated in the absence of physician creation of demand. In a competitive market an increase in the supply of a good causes prices to fall and quantity demand to rise, i.e., there is a correlation between increased supply and increased quantity of services.

---

[8] For a mixed review of prospective reimbursement rate setting, which sets a fixed amount for hospital reimbursement before expenses are incurred, see Fred J. Hellinger, "Prospective Reimbursement through Budget Review: New Jersey, Rhode Island, and Western Pennsylvania," Inquiry, September 1976, pp. 309–320. A critical review of the Professional Standards Review Organization program, which attempts to oversee physician practices in hospitals, see Clark C. Havighurst and James F. Blumstein, "Coping with Quality/Cost Trade-Offs in Medical Care: The Role of PSRO's," Northwestern University Law Review, March–April 1975, pp. 6–68. For a fairly complete review of the failures of regulation, in general, see Clark C. Havighurst, ed., Regulating Health Facilities Construction, 1974, pp. 7–25.

[9] See, for example, Robert Evans, "Supplier-Induced Demand: Some Empirical Evidence and Implications," in "The Economics of Health and Medical Care: Proceedings of a Conference Held by the International Economic Association at Tokyo," edited by Mark Perlman, 1974, and Victor R. Fuchs and Marcia A. Kramer, "Determinants of Expenditures for Physicians' Services in the United States 1948–68," New York, National Bureau of Economic Research/HEW, 1973.

It is also possible in a competitive market to observe an increased physician supply and higher prices of medical services. This would be the case if physicians spent more time with their patients (which would increase the value of the visit for the patient), or less waiting time would be incurred by the patient (which would reduce the price of the visit for the patient). Both of these patterns have been observed.[10]

In a recent paper, Victor Fuchs correlates the supply of surgeons and the demand for operations across geographical areas of the United States.[11] By using "surgeons" Fuchs avoids the problems of waiting times and lengths of visits. Fuchs finds a positive correlation between the surgeon/population ratio and per capita utilization. However, Fuchs' work does not adequately determine whether surgeons are dictating demand, or high demand is bringing forth a greater supply of surgeons.[12]

More research is needed in this area such as research on a survey based on data from particular types of physicians. It is not at all clear that increases in certain types of physicians or paramedics will not exert downward pressure on prices. Organized medical societies which have opposed the introduction of paramedics may believe that lower physician fees may result.

b. Widespread Insurance for Medical Care

The presence of widespread insurance for medical care also may weaken competitive forces in this industry. Insurance both lessens search for lower prices and increases demand for medical services. Empirical evidence has demonstrated that the increased use of insurance has resulted in more services of a more expensive variety than consumers would elect to purchase in a market based on direct payment on services.[13] It is no surprise that hospital services which are the most heavily insured are increasing most rapidly in price. Moreover, as the increased use of insurance increases the price of medical services, it increases further the demand for insurance.[14] As the Nation approaches full coverage only the supply of physicians and the time-cost of seeing physicians and being hospitalized act as restraints on usage. Currently, 70 percent of all medical services are covered by insurance; more than 60 percent of physician services and more than 92 percent of all hospital bills are insured.[15] The rapid rise of medical care expenditure has been concomitant with the rapid rise of insurance coverage. Health care expenditures grew from 5.2 percent of G.N.P. in 1960 to 8.8 percent in 1977, while third-party payments (both public and private) grew from 45 to 70 percent of total health expenditures.[16, 17]

---

[10] See Frank A. Sloan and John Lorant, "The Allocation of Physicians' Services: Evidence on Length-of-Visit," Quarterly Review of Economics and Business, Autumn 1976, pp. 85–103, and Frank A. Sloan and John Lorant, "The Role of Waiting Time: Evidence from Physicians' Practices," Journal of Business, forthcoming.

[11] Victor R. Fuchs, "The Supply of Surgeons and the Demand for Operations," Journal of Human Resources, Supplement 1978, pp. 35–56.

[12] See Victor R. Fuchs and Joseph Newhouse, "The Conference and Unresolved Problems," Journal of Human Resources, Supplement 1978, p. 10.

[13] Martin S. Feldstein, "The Rising Cost of Medical Care," Washington, D.C., 1971.

[14] See Martin S. Feldstein, "The Welfare Loss of Excess Health Insurance," Journal of Political Economy, March/April 1973, pp. 251–280.

[15] Health Education and Welfare, HCFA Health Notes, May 1978, p. 2.

[16] Ibid.

[17] Council of Economic Advisors, Economic Report of the President, January 1976, p. 118, table 35.

IV. Injecting Competition in the Health Care Sector Given the Widespread Presence of Insurance

a. HMO's as Alternatives to Fee-for-Service

One way to provide more competition in health services is to encourage (or at least not discourage) the growth of health maintenance organizations (HMO's). An HMO is basically a prepaid health insurance plan in which the organization and participating physicians accept contractual responsibility for the delivery of a stated range of health services. Sponsors of the HMO have included physician groups (Ross-Loos in Los Angeles), Blue Cross (Genesee Valley Group Health Association in Rochester, N.Y.), universities (Georgetown University Plan in Washington, D.C.), and hospitals (Michael Reese in Chicago) both for profit and not for profit. Unlike the traditional fee-for-service sector, the HMO, at risk for costs that exceed income, generally has an incentive to minimize hospital admissions, to limit the length of hospital stays, and to use non-physician personnel as substitutes for physicians. Indeed, many studies have found that HMO's provide health care at lower cost than does the fee-for-service sector.[18]

In a recent article, based on a study done for the Federal Trade Commission, Goldberg and Greenberg found that the existence of HMO's in a geographic area stimulated Blue Cross, the Nation's largest insurer to exert greater control on hospital utilization and to offer a more complete benefit package.[19] The two main features of the HMO—reduced hospital utilization and greater benefit packages—were mirrored by Blue Cross. In addition, Blue Cross responded with its own HMO's while many physician groups, in general, became more cognizant of possible excessive utilization. The major results of this study were recently substantiated by Jon Christianson in an analysis of HMO and fee-for-service competition in Minnesota and Hawaii.[20]

Unfortunately, there has been resistance to health maintenance organizations in their early beginnings, and even relatively recently physician opposition has surfaced.[21] Public policy need not necessarily endorse HMO's, but at least HMO's should be allowed to compete fairly with the fee-for-service sector without any private restraints.

b. Use of Copayments, Deductibles, and Indemnity Provisions in Insurance Plans

A second competitive alternative is to encourage insurance policies with copayments, deductibles, and indemnity provisions. The theory behind all of these is that if the consumer pays all or some of the bill he will be more cognizant of prices charged and services rendered. Under copayments and coinsurance, for example, the consumer would pay a specified amount or percentage (such as 20 percent) of the medi-

---

[18] For a review of many of these studies, see Harold S. Luft, "How Do Health Maintenance Organizations Achieve Their Savings?" New England Journal of Medicine, June 15, 1978, pp. 1336–1343.

[19] See Lawrence G. Goldberg and Warren Greenberg, "The Competitive Response of Blue Cross to the Health Maintenance Organization," forthcoming, Journal of Economic Inquiry.

[20] See Jon Christianson, "Do HMO's Stimulate Beneficial Competition?" InterStudy, April 1978.

[21] See Reuben Kessel, "Price Discrimination in Medicine," Journal of Law and Economics, October 1958, pp. 20–53, and the recent Federal Trade Commission order, Medical Service Corp. of Spokane County, et al., DOC—2858, 1976.

cal bill. Copayments and coinsurance can make the insuree somewhat, but not totally, sensitive to price since he is paying part of the bill.

Deductibles are the total amount that the consumer must pay to a set limit. To be most effective they should be set large enough so that most of the population, most of the time, will not exceed them. The higher the deductible, the more that decisions are made by consumers based on price. The consumer must, however, also bear the costs of increased risk. A proposal, first suggested by Martin Feldstein, and endorsed by many economists, emphasizes high deductibles, or catastrophic insurance to induce consumers to search for the lowest prices for most of their medical expenses.[22] Catastrophic insurance is consistent with the classic use of insurance for high risk, random, relatively expensive events, rather than for everyday predictable expenses.

Finally, there is indemnity insurance. An indemnity policy pays a certain amount of dollars per unit of services consumed, such as $100 per hospital day or $6 per office visit. Such insurance may pay a good portion of the consumer's bill, but if a consumer uses a more expensive hospital or physician he has to bear the incremental cost. Hence, there are incentives to search for less expensive providers. To the extent that this search is hampered by the lack of information on prices of providers, the benefits of indemnity insurance are reduced [23] (as would be the benefits of copayments and deductibles).

If insurers were to make greater use of copayments. deductibles, and indemnity provisions, it might create increased incentives for third parties to control costs, since consumers could realize some portion of cost savings if third parties were successful.

c. Encourage Cost Control Efforts by Third Parties

The widespread increase in health insurance coverage has been largely responsible for the rapid increases in health care costs.[24] Most studies of the effects of health insurance, however, have assumed that insurers are passive payers of health care bills. This is in contrast to the automobile insurer, for example, who is a careful payer of automobile accident claims. Yet, insurers as informed purchasers of health care services, may be in a position to control health care costs since they have greater knowledge of the medical marketplace than consumers. It is of interest to cite first an example from dentistry. Dental insurance is the fastest growing insurance in the United States. By 1980 more than 60 million people are expected to have some sort of coverage, compared to only two million people in 1965 and 30 million people in 1976.[25] Dental insurance plans vary, as do health insurance plans, but may include benefits as extensive as the United Automobile Workers benefits contract with Aetna Life and Casualty wherein diagnostic and preventive services, general services, and prosthodontics are included.[26]

---

[22] See Martin S. Feldstein, "A New Approach to National Health Insurance," The Public Interest, Spring 1971, pp. 93–105.

[23] The recent decision by a Federal Trade Commission Administrative law judge which supported the FTC's complaint against the AMA for its ban on solicitation of patients, if upheld by higher courts, should help increase the amount of information available to the public. See "AMA's Ban on Advertising by Doctors is Illegal, FTC Judge Rules," FTC News, November 29, 1978.

[24] See Martin S. Feldstein, "The Rising Cost of Hospital Care," 1971.

[25] American Dental Association, "Prepaid Dental Care," undated.

[26] "Summary of UAW Dental Benefits," furnished to author by Aetna Life and Casualty Company, December 23, 1976.

What is the behavior of dental insurance firms in terms of their actions to affect cost control? In many group dental plans, insurers review dental claims in two stages.[27] In the first stage, "Predetermination," the insurance carrier will review a "proposed treatment," which is expected to cost in excess of a certain dollar value, for "necessity and appropriateness." This review attempts to eliminate misunderstandings among the patient, the dentist, and the insurance firm before work is performed. Second, under the Alternative Course of Treatment (A.C.T.) program, dental consultants on the insurer's staff suggest alternatives to the work plan of the patient's dentist, if necessary.

Under its United Automobile Worker's contract, the largest prepaid dental contract in the United States, Aetna Life and Casualty, for example, questions not only whether various procedures are covered but also ". . . . claims where the charges are for services and supplies which appear not to be necessary for treatment, or where the charges appear to be above the prevailing charge level." [28] In addition, Aetna's investigation of questionable claims might utilize investigatory procedures such as:

(*a*) Discussion with the attending dentist.

(*b*) Examination of dental x-rays and study models.

(*c*) Case review by Aetna's dental consultant when professional judgment is required.

(*d*) Oral examination of the patient by Aetna's dental consultant.

(*e*) Referral to the local Dental Society Review Committee.[29]

Unlike health insurers, dental insurers seem willing to question professional practices. Under competition, dental insurers have incentives to control costs to keep their premiums competitive with other firms in the market. The third party, in effect, acts as an expert for the uninformed consumer by enabling the consumer or patient to make more knowledgeable choices about the various elective procedures in dentistry. The patient, however, always has the option of rejecting the insurer's advice and paying the dentist himself. In addition, dental insurers which must compete in the marketplace for employer groups (by far, the largest purchasers of dental insurance) need respect the quality of dental care in addition to price or risk consumer complaints or loss of business. For their part, with some exceptions, dentists appear willing to accept dental insurer "interference" since insured services, approved by the insurer, represent a guaranteed source of payment.[30]

Is dentistry so much different from medicine that in some instances health insurers could not perform like dental insurers? Although all insurers have similar incentives to maximize profits, and both dental

---

[27] According to the Council on Wage and Price Stability, this two-stage plan is the model for contracts with the automobile, steel, and railway workers, as well as with the United Federation of Teachers in New York City. The Council reports that this program "reduces dental claims against the U.F.T. Welfare Fund, which covers 250,000 teachers and dependents, by approximately $1 million a year." See Council on Wage and Price Stability, The Complex Puzzle of Rising Health Care Costs, December 1976, p. 116.

[28] American Dental Association, Extent of Dental Prepayment as of December 31, 1975, undated, p. 1.

[29] American Dental Association, "Prepaid Dental Care", undated, p. 2.

[30] In November 1978, the Federal Trade Commission charged the Indiana Dental Association (I.D.A.) along with its component societies with collective restriction of insurer cost containment efforts. The I.D.A. agreed, however, in a consent agreement with the Commission, to prohibit such restrictions in the future. See F.T.C. News, November 15, 1978.

and medical insurance include involvement with professionals, greater uncertainty of outcome, greater risk of certain procedures, greater ignorance of the consumer of medical procedures, and more emergency or quasi-emergency treatments, may force health insurers to behave differently in degree than dental insurers. In fact, in the early history of health insurance, in the 1930's and 1940's, health insurers in the state of Oregon actively questioned a physician's procedures, hospital length of stay and bills rendered.[31] Insurers asked for written justification from the physician prior to elective hospital admission and elective surgery. Payment was denied to physicians who were not able to justify adequately their procedures. Although physicians attempted to resist this interference in the "doctor-patient" relationship, individually the physicians preferred to have some payment by the insurer rather than boycott collectively and forego the income. The subsequent disappearance of these insurer-initiated cost controls in Oregon has been traced to the medical society's organization of a competing Blue Shield plan as a model of insurer conduct and to a simultaneous boycott by physicians of third parties as long as they persisted in questioning physicians' practices.[32] By creating a Blue Shield plan (termed the Oregon Physicians Services plan) which was receptive to physician needs, physicians were in a position to direct patients away from the more active plans, and still receive insurer revenues.

A contemporary example of a health insurer monitoring physician behavior was described in a recent study of health insurer cost control efforts.[33] In southern California, a small firm with less than one-half million enrollees, U.S. Administrators, attempts to control health care costs with a large scale computer system and the assistance of 23 physicians on its advisory board. Two major computer files are used as an initial step in monitoring physician behavior. First, the Model Treatment Profile has, for each disease in the I.C.D.A. Handbook (International Classification of Diseases, Adopted, 8th edition) a range of the number of physician visits or procedures which would be appropriate. When a physician's visits or procedures exceed this range, the computer kicks a report out for manual review. A second file, the Diagnosis File, specifies maximum lengths of stay for hospital visits. When these parameters are exceeded, a report is automatically released for manual inspection. In approximately 4 percent of its cases, U.S. Administrators have referred to pay for what it (and its 23 physicians on its advisory board) believes to be excessive utilization.[34] Thus far, there appears to be little physician resistance to the tactics of U.S. Administrators precisely because it has so few enrollees. If U.S. Administrators grows in size a greater share of potential physician income might be lost and physician resistance might be expected.

V. Public Policy Initiatives Which Can Encourage Competition in the Health Care Sector

Thus far the paper has focused on the types of competition which can occur in the health care sector even with the pervasiveness of

[31] See L. G. Goldberg and Warren Greenberg, "The Effect of Physician-Controlled Health Insurance: *U.S. v. Oregon State Medical Society*," Journal of Health Politics, Policy and Law, Spring 1977, pp. 48–78.
[32] See L. G. Goldberg and W. Greenberg, Ibid.
[33] See Robert D. Brogan and Warren Greenberg, "Health Insurer Market Structure and Cost Containment," forthcoming.
[34] See Robert D. Brogan and Warren Greenberg, Ibid.

health insurance.[35] The competitive initiatives, however, which have been discussed are not common in the economy. Health maintenance organizations enroll only 3 or 4 percent of the population. Although there are now 199 operational HMO's, overall growth has still been limited.[36, 37]

The use of copayments, deductibles, and indemnity payments has become less frequent. Third parties pay more than 94 percent of hospital expenditures. Moreover, even commercial insurers, which unlike Blue Cross encouraged cost sharing for enrollees, have no cost sharing among their largest accounts.[38] And insurers which examine physician behavior are rare, although the firm, U.S. Administrators, appears to be growing. It appears fair to ask why there are no additional forms of competitive behavior in the health care economy.

There are at least three reasons for the absence of more competition in health care.

First, there are potential threats of provider boycott similar to the tactics used by the Indiana Dental Association.[39] Second is the nature of medicine. Physicians are trained to treat patients until the marginal benefits of treatment are zero. This suggests that physicians may treat patients beyond the cost-benefit framework defined by the economist as economically efficient.[40] It becomes difficult therefore for an insurer, of any size, to monitor or contain "excessive" tests or utilization, when the physician in good faith can medically justify additional procedures. Third, the Nation's income tax laws are such that there are incentives for employers to provide untaxed fringe benefits in the form of increased health insurance coverage rather than increased potentially taxable income. Health insurance premiums paid by employers are exempt from the employee's Federal, State, and local income taxes as well as from the employee's social security taxes. Employees and unions therefore have incentives to bargain for increased health insurance benefits relative to increases in wages.[41] The result has been, in general, much more shallow coverage and much less catastrophic coverage than consumers would purchase without the subsidy. In effect, most insurance is prepayment for health care bills, not for high-risk random events, the traditional reasons for insurance. Furthermore, the need for Congress to pass legislation to help Americans who need kidney dialysis treatment, and recurrent congressional hearings high-

---

[35] There are many variants of these two types of competition, especially when one considers the various types of HMO's and pre-paid practices. See Clark C. Havighurst and Glenn M. Hackbarth. "Private Cost Containment," The New England Journal of Medicine, June 7, 1979, pp. 1298–1305.

[36] See Group Health News, December 1978, p. 1.

[37] A recent study suggests that HMO's grow most rapidly in States with the largest union membership and number of physicians practicing in group practices. See Lawrence G. Goldberg and Warren Greenberg, "The Determinants of HMO Enrollment and Growth," submitted for publication.

[38] See Robert D. Brogan and Warren Greenberg, Ibid.

[39] Recently, physicians attempted to boycott Aetna Life and Casualty Company for disallowing a portion of a physician's fee for exceeding prevailing limits. See Lawrence G. Goldberg and Warren Greenberg, "The Emergence of Physician-Controlled Health Insurance, *U.S. v. Oregon State Medical Society*," Ibid., pp 62–65.

[40] As Schwartz and Joskow point out, a large number of medical procedures engender some positive benefits (what they call medical efficacy) but one outweighed to some degree by their costs (economic efficiency). See William B. Schwartz and Paul L. Joskow, "Medical Efficacy versus Economic Efficiency: A Conflict in Values," New England Journal of Medicine, December 28, 1978, pp. 1472–74.

[41] This problem may be further exacerbated during wage and price controls or wage and price guidelines. Currently, for example, employer-paid health premiums are exempt from the President's wage and price guideline program.

lighting families with enormous uninsured medical bills suggest the absence of catastrophic insurance.

A number of plans, therefore, has evolved which concentrate on changing the tax laws. First, the most important of the proposed plans will be reviewed; second, a plan will be suggested which can help encourage competition without some of the disadvantages of the aforementioned plans.

The most drastic proposal to alter the tax subsidy for health insurance would be to tax, as any other income, all employer-paid health insurance premiums. In addition, the deduction of up to $150 in health premiums paid by the employee would be abolished. Martin Feldstein, Professor of Economics, Harvard University, has been the most forceful in promoting this view.[42] Feldstein suggests that his proposal will encourage people to purchase the catastrophic, expensive coverage rather than the shallow coverage that is currently so common. Under the Feldstein proposal, distortions in demand caused by excess insurance coverage would be eliminated. There are at least two problems with the Feldstein proposal, however. First, since people are accustomed to and have received the benefits of the tax subsidy, the abolishment of the tax subsidy would be difficult politically. Second, some health services might not be abolished at all, since employers, in attemping to provide an equivalent level of benefits, might be tempted to substitute non-taxed in-house physicians and medical services.

A more moderate approach has been embodied in a new health insurance plan proposed by Representative Al Ullman, Chairman of the House Ways and Means Committee.[43] Ullman suggests that there should be a cap placed on the Federal tax subsidy for medical insurance and that employers should pay equally to each plan which is offered. If an employee desires a more expensive plan, it would have to be paid out of his own after-tax income. Like the Feldstein proposal, this plan also, but to a lesser degree than Feldstein's plan, takes away a benefit with which people are accustomed. Second, requiring that employers pay equally to each plan offered does not necessarily eliminate the offering of more expensive plans.

Alain Enthoven's Consumer Choice Health Plan is closely related to Chairman Ullman's plan.[44] It takes as its premise that competition among plans to control costs is desirable. It envisions a number of "qualified" plans being offered by employers. Qualified plans are those which have open enrollment, community rating, comprehensive benefits, and full protection against the cost of catastrophic illness. Like the other plans, however, employer contributions to health insurer premiums would be subject to income and other taxes, but a tax credit would be given to those who selected a qualified plan. The tax credit would be the same regardless of which plan was chosen so those who joined a more expensive plan would have to pay for it with after-tax dollars. A political problem with this plan, like the Ullman plan, is

---

[42] See Martin S. Feldstein, "The Rising Cost of Hospital Care," 1971, and Martin Feldstein and Bernard Friedman, "Tax Subsidies, The Rational Demand for Insurance and the Health Care Crisis," Discussion Paper Number 382, September 1974.

[43] See speech prepared by Representative Al Ullman, Chairman of the House Ways and Means Committee, for delivery to the National Journal Conference on Health Policy in Washington, D.C., June 7, 1979.

[44] See Alain C. Enthoven, "Consumer-Choice Health Plan," New England Journal of Medicine, March 23 and March 30, 1978, pp. 650–658 and 709–720.

that by limiting the tax credit to a certain dollar amount, people will have to pay after-tax dollars for the equivalent present coverage. This might be politically unacceptable.

To circumvent many of the above problems, I would suggest a tax supplement health plan which would encourage people to enroll in plans which control costs. The essentials of the tax supplement health plans are as follows: [45]

A tax supplement would be given to any individual who enrolled in a health insurance plan which had been able to control costs at a specified level in the preceding year. That level might be equal to the increase in the consumer price index in a geographic area. The health insurer could control costs in any way the firm thought to be most efficient. The firm might examine more closely physician procedures, refuse to pay for redundant technological equipment, or offer copayment and deductible provisions in its contract. The tax supplement would be provided to the consumer who would incur the added costs of negotiating physician-insurer disputes, and the inconvenience of paying copayments. The tax subsidy would provide the individual with the explicit choice between choosing an expensive plan with no subsidy or a more modest plan with a subsidy. Health insurer plans would be required to have open enrollment and community rating so that cost control would not be based simply on the ability to enroll healthy people.

The exact dollar amount of the supplement would be based on whatever the Government determined was needed to encourage people to select cost-controlling plans. For example, one might receive a supplement of $300 at the end of the year for selecting such a plan. The tax supplement (adjusted by individual income level, if necessary) would increase the business of the more effective insurance firms and, therefore, would provide incentives for firms to control costs if they desired increased business.[46] The employee who selected such a cost-controlling plan would receive the benefits of a lower cost insurer as well as the benefits of a tax supplement. If costs are shifted by providers to the relatively passive insurer, the relative prices between those insurers who controlled costs and those who did not would increase, making the firm which controlled costs even more attractive to consumers.

The actual details of the yardstick (such as the particular regional consumer price index) for firms certified to have effectuated cost control, and the exact dollar value of the subsidy to be paid need to be worked out. It might be specified, for example, that employees would be allowed to select only one health insurance plan from an employer. (This would prevent the employee from selecting multiple, relatively incomplete plans, and receiving a tax supplement. If, however, the employee elected to purchase outside supplemental insurance, this would be consistent with consumer desires and would, of course, be allowed.) The important point here is that there is some supplement

[45] This is not meant to be a full scale national health insurance proposal. Such a proposal involves consideration of a broad range of socioeconomic factors. Rather, it is an example of the various conditions that might be needed in any national health insurance plan which would attempt to bring more competition into the health care system.

[46] Presumably, employers would offer the option of a cost-conscious insurer along with insurers who elect not to control costs, similar to the way that employers now offer the HMO option. (If, in fact, an HMO were to control costs to the specified level, employees who would choose an HMO would also receive a tax subsidy.)

which consumers will accept, and firms will engage in cost control (to a greater degree than at present) and a tax supplement should be geared to that particular level.

Revenues for the tax supplement might come from two sources. First, employer paid health insurance of upper income employees might be made subject to income taxes. This tax would also improve the distributional aspects of the plan [47] since lower income people would still receive the benefit of tax-free employer paid health premiums as well as the tax supplement. Second, lower health care costs, in general, would result (with consequent savings to the government in government sponsored health programs), since insurers would not have incentives to control costs. This might even be true for insurers who elected not to control costs, forcefully, since higher costs might compel them to be more active in this regard.

In essence, tax supplement health insurance would help control costs in the framework of consumer sovereignty with minimum distortion of current tax laws or deprivation of current benefits. There would be additional incentives created for firms in the private sector to control costs. One would expect a variety of competing insurers in the marketplace, each offering its own cost control (or lack of cost control) plan. Consumer choice could be effectuated easily without the need for the individual to commit himself to any particular group of physicians or hospitals.

VI. Role of Competition in the Health Care Sector

Regardless of the potential effectiveness of any of the health plans which have been presented, it would be unduly optimistic to suggest that the rapid increases in health care costs can be entirely curtailed. There still appears to be a number of imperfections in the medical marketplace which might be addressed by public policymakers. Some of these imperfections are:

(1) Lack of information between the providers (physician or dentist) and the patient. A.M.A. and A.D.A. restrictions on the dissemination of information by members have recently been challenged and found to be in violation of the antitrust laws by a Federal Trade Commission administrative law judge.[48]

(2) Restrictive licensing practices such as reciprocity agreements among dentists.[49]

(3) Collective actions by providers against cost control activities of third parties.[50]

Finally, even if market forces were operating more efficiently, there is a great deal of uncertainty about the effects of medicine in general (even from a physician's viewpoint) which may compel the physician

---

[47] Senator Edward Kennedy has also commented on taxing fringe benefits of upper income employees. See "Kennedy in New Attack, Says Carter Budget Approach Favors the Rich," The New York Times, January 14, 1979, p. 18.

[48] See decision of Ernest G. Barnes, Administrative Law Judge, Federal Trade Commission, D-9064, November 13, 1978.

[49] See Lawrence Shepard, "Licensing Restrictions and the Cost of Dental Care," Journal of Law and Economics, April 1978, pp. 187–201

[50] See "FTC Charges Indiana Dental Groups with Restricting Insurers' Cost-Containment Efforts," FTC News, November 15, 1978. Currently, the Federal Trade Commission is investigating possible collective action of the Michigan State Medical Society against Michigan Blue Cross/Blue Shield.

to avoid risk and opt for the latest and most expensive care.[51] Indeed, these complex imperfections make it all the more important to encourage, as much as is feasibly possible, the workings of market forces in the health care sector.

---

[51] See Mark V. Pauly, "Is Medical Care Different?" in Warren Greenberg, ed., "Competition in the Health Care Sector," 1978, pp. 11–35.

THE DIFFERENTIAL IMPACT OF GOVERNMENT REGULATION ON SMALL VERSUS LARGE BUSINESS: A REVIEW OF AN INADEQUATE INFORMATION BASE

By William M. Diamond*

CONTENTS

	Page
Summary	270
I. Introduction	273
Scope of the Study	273
Small Business	274
Regulation: The Special Concern for Small Business	275
The Impact Issue	275
Impact: An Operational Definition	276
II. Regulatory Impact—The Conceptual Framework	276
Basic Types of Regulatory Programs	276
Regulatory Impact: The Broader View	277
III. Regulatory Impact: Large Business	278
IV. The Impact of Regulation on Small Business	282
Overall Indexes of Regulatory Impact	282
Selected Industry Experience	286
Hearing: Overregulation of Small Business	287
Hearing: The Effects of OSHA on Small Business	289
Summary Evaluation of Small Business Regulatory Impact Materials	289
Benefits Accruing to Small Business From Regulation	289
The State Incremental Impact on Small Business	290
State Regulatory Burden—The California Assessment	291
The New York Experience	292
V. Recommendations	294
The Primary Methodological Requirement	294
Secondary Methodological Considerations	295
Regulatory Reform: A Plea for Moderation	295

SUMMARY

Expansion of government controls on the private sector has appeared to proceed at a near-exponential rate over the past 15 years. The dramatic and costly extension of stringent regulations over the environmental factors affecting business created must be of this impression. Another arena of major regulatory growth has encompassed the broad range of terms and conditions of employment. The impact of this second category of regulation on business, in addition to the desired changes in employment practices and job safety, can be measured in the increased paperwork burden. Legislation fixing standards of environmental contamination, by contrast, often requires substantial incremental capital investment as well as on-going operating costs to support the installations and to remain in compliance.

*School of Business, State University of New York at Albany.

The purpose of this paper is to review the 'state of the art' with respect to empirical assessment of regulatory impact on small business. To the extent that available empirical data provide an appropriate foil for the discussion of small business, the compliance experiences of large scale enterprises are also reviewed. In the context of this paper, impact refers to the costs, changes in operating conditions, and benefits accruing to business as a result of government regulation. The paper does not address the much broader question of net societal impact.

The conceptual framework of regulatory impact.—The bulk of existing material relating to regulatory evaluation is either non-empirical or non-specific to the particular segment of small business activity. Some generalizations, such as taxonomy of regulatory efforts, are relevant without regard to segmentation. A basic division may be made between traditional and non-traditional approaches. The former type tends to be proscriptive, while more recent approaches have tended to focus on incentives to elicit desired behavior patterns from the business community.

In the broad context of total regulatory activity, several measures have been attempted by researchers. The most widely quoted of these estimates assesses the total dollar burden of regulation and regulatory compliance for fiscal 1979 at $102.7 billion. This figure may, in fact, be conservative since productivity losses and opportunity costs attributable to compliance requirements are not included.

Measures of impact on large business.—Quantitative measures of regulatory impact on large scale business are limited. Some reports appearing in the business press provide an indication of the overall burden. In 1975 Dow Chemical reported total costs of $147 million for compliance with 'domestic' regulations. For the same year General Motors Corporation spent $1.258 billion on research, development, and administrative costs of regulatory compliance. The absence of any treatment of methodological approach in these reports hinders objective evaluation. This problem tends to pervade the published material. Until more detailed information is available, it will remain a near impossibility to measure the real impact of regulation on large scale business activity. A recent study by the Business Roundtable and Arthur Andersen & Company is a representative example of the results to be expected of improved methodology.

Other types of analyses, developed for a different purpose, offer some possible models for future impact studies: The reports commissioned by the Environmental Protection Agency to forecast the likely results of new regulations reflect much of the concern for objectivity that ought to pervade all impact analyses.

The impact of regulation on small business.—More attention has been focused on the small business community and its problems with regulatory compliance than is the case with large establishments. A number of organizations within both the public and private sectors has devoted a substantial amount of attention to the subject.

The quarterly survey of its membership undertaken by the National Federation of Independent Business for the past five years has repeatedly revealed a high level of concern by respondent businesses over

the volume of "government regulation and red tape." Consistently, one in 10 firms reports this area to be its most significant problem.

The Center for the Study of American Business reports a variety of general and specific impact problems faced by small business. In addition to paperwork burden and compliance costs associated with specific agencies such as OSHA, the Center has focused attention on the relationship between regulation and access to the financial markets by small firms. In this latter area, the reported data clearly demonstrate the disproportionate impact of regulation by the Securities Exchange Commission on the small business sector.

The testimony from congressional hearings clearly reflects the anger and frustration of individual organizations and industry groups. Repeatedly, business representatives cite examples of perceived unfairness, difficulty and misunderstanding. At the same time, however, these collections of testimony are not objective evidence. Certainly, the negative reports are symptomatic of problems, but further study is essential to secure a balanced evaluation.

A significant peripheral point in any discussion of impact is the economic benefit accruing to business as a result of regulation. These benefits arise throughout the industrial structure as a result of business investment for purposes of regulatory compliance. In the realm of pollution and noise control alone, one estimate places the total sales generated for the small business sector during the period 1975–80 at over $17 billion.

State regulatory requirements place a significant additional burden on the individual business. In 1976, the state of California estimated paperwork costs alone for the typical small firm at between $3,000 and $4,000 annually. For the 1978–79 fiscal year in New York it is estimated that per establishment incremental costs of compliance with state regulations equaled $5,000. This figure, added to a probable Federal burden of $19,000, contributes to a combined Federal-State compliance cost of $24,000 per establishment in New York. As an average, the $24,000 figure obscures the fact that many firms in the manufacturing sector are incurring costs many times this level annually.

Recommendations.—The primary recommendation of this paper is a call for improved impact analysis. Available data are neither sufficiently objective nor comprehensive on which to base a complete program of regulatory reform. Studies should be undertaken which are designed in full cognizance of sound empirical data collection and analysis procedures.

In the final analysis, impact studies are incomplete in the absence of benefit evaluation. More attention should be directed to the design and implementation of cost-benefit studies in the small business sector. At this stage it becomes important to consider also the secondary costs of regulation such as the delay of technological innovation that may result from stringent controls.

Finally, in view of the inadequacy of existing data it is recommended that immediate efforts at regulatory reform be limited to internal procedural improvements. Major structural changes should await objective, comprehensive impact assessment. This type of recommendation is anathema to the reformer, who, armed with recent examples of successful deregulation and convinced of the efficacy of multitiered models, is impatient to move ahead to a reconceptualized basis

of regulation. It would indeed be unfortunate if such attitudes prevail since the results could be costly experiments performed on the economic community of the whole. Fortunately, there is so much to be gained through the systematic revision of existing procedures that there is no necessity for precipitous action to accomplish major improvements in the costs and equity of regulation.

I. Introduction

Recent years have witnessed unprecedented growth in government regulation of business at both the Federal and State levels. Documentation of action-impinging government control has elevated a task that of itself is conceptually simple into a major area of profession activity. Assessment of the economic and social costs and benefits of this expanded regulatory activity has lagged, unfortunately, to the point where the effects on specific segments of the economy remain unclear.

Scope of the Study

This paper provides, first, a review of the published sources dealing with the impact of Federal and State regulations on small business; second, and to a lesser extent, the paper touches on current practices in the field of regulatory analysis; third, the paper provides some case studies of regulatory compliance experiences; and, finally, recommendations are made regarding future action. Limitations on the breadth of regulatory activity covered in the report were imposed, on the one hand, because of resource limitations. Also, the similarity of many regulations in terms of the character of their impact permits a degree of categorization. Therefore, a relatively small subset may be considered exemplary.

A simple approach to categorization is to dichotomize between those regulations where compliance is technology forcing and those where impact is measured largely in paperwork burden. In the first category is a large number of Federal laws (over 50) enacted in the decades of the 1960's and 1970's. Included and quite representative of such regulations with broad impact, *i.e.*, across industry lines are the following:

Clean Air Amendments of 1970, Public Law 604.
Federal Water Pollution Control Act Amendments of 1972, Public law 92–500.
Occupational Safety and Health Act of 1970, Public Law 91–596.
Consumer Product Safety Act of 1972, Public Law 92–573.
Emergency Petroleum Allocation Act of 1973, Public Law 93–159.
Wholesome Meat Act of 1967, Public Law 90–201.

In the second category are the reporting requirements of such regulatory bodies as:

(1) Bureau of the Census.
(2) Equal Employment Opportunity Commission.
(3) Internal Revenue Service.
(4) Office of Federal Contract Compliance.
(5) Securities and Exchange Commission.

Complementary or overlapping statutory provisions requiring separate reporting may be found in either group at the State level.

Small Business

The Federal Small Business Administration (SBA) continues to wrestle with the problem of defining small business. A report of the Office of Advocacy, SBA, recently noted the following:

> . . . Instead of focusing on competition, the one concept that ties all small businesses together, the SBA size standards system has incorporated five different units of measure into eight different definitions of small business . . . Little wonder that the SBA has had considerable difficulty in conveying to Congress and the public, just what is meant by a "small business."
>
> The contrast between the deliberations of the Congress concerning the definition of a small business, and the evolution of the SBA size standards procedures, gives support to the belief that "nothing is so simple that it cannot be made confusing." But perhaps the Congress tended to overly simplify a complex problem, while the SBA sought a degree of preciseness that exceeded reasonable bounds?
>
> Regardless of the cause, the dissatisfaction with the current method for defining a small business, shared by many within the SBA as well as those outside the Agency, requires a change. . . .[1]

Nearly two years later there is little evidence to indicate the problem has been resolved within the SBA, despite continuing efforts toward that end. The Small Business Administration convened a workshop in November 1978, to consider the problem of the definition of small business for research purposes. Noting that, at the present time, the definition selected "often depends on the data base used because there is no current and comprehensive set of data following a consistent definition . . .", the workshop reached "general agreement" on the following:

> 1. That employment should be the primary measure of business size;
> 2. Gross sales/receipts should be a second standard;
> 3. Asset size is important but difficult to measure consistently;
> 4. Age is important in understanding survival;
> 5. Enterprise is preferred over establishment;
> 6. County summaries of data would provide the best geographic base; and
> 7. Within the definition of small business, a minimum employment classification should be: 0 employees, 1–19 employees, and 20–5,000 employees; a more useful employment breakdown could be achieved by establishing a further degree of sub-classification, such as 5–19 employees. [2]

These bases for definition highlight the difficulties currently facing those who would see a common set of criteria, especially since the above qualifications are not fully represented in any existing available Federal or State data base.

Yet, an acceptable operational definition of small business is essential to effective progress in regulatory research if a basis of comparison is to exist among studies. In the past, too little attention has been paid to the qualitative element of the definition. As a result, and in recognition of different levels of the economies of scale, the SBA and others flounder about in a morass of definitions which, in reality, should be considered appropriate sub-classifications under the qualitative generalization. In such a context researchers may find it relatively easy to agree on a definition similar to the following, while reserving qualitative sub-classes for analytical purposes within a study: a small busi-

---

[1] U.S. Small Business Administration, Office of Advocacy, "The Study of Small Business," U.S. Government Printing Office, 1977, p. 17.

[2] U.S. Small Business Administration, "Report on the Small Business Data Needs Workshop," Washington, D.C., 1978, pp. 3–12.

ness is one that is independently owned and operated by an owner/manager and has a single decision-making center.[3]

Because no such agreement as yet pervads the literature, this paper temporizes and employs the concept of a prior Federal consultant: "We have used the word 'small' in a generally qualitative sense, although for our purposes, the definitions for eligibility for financial assistance and regulatory exceptions are particularly relevant." [4] The more precise definition suggested, if employed herein, would disqualify from reference most of the sources employed from an already too thin conceptual base.

Regulation: The Special Concern for Small Business

Existing studies point out the relevance of a special concern for small business in the consideration of regulatory impact. These efforts have to some extent documented that regulation places unique burdens on the small business. The organizational structure of a small business, largely devoid of staff specialization for reasons of scale, requires that the bulk of regulatory compliance effort be borne by the owner/manager. This scenario implies that the manager's creative efforts at planning and innovation are effectively reduced by the amount of time required for the routine of compliance. Here then is the crux of the matter insofar as small business is concerned: While the total costs of regulatory compliance tend to increase steadily with business size, the ability to shift the burden of work from management to staff also occurs only with large scale operations.

Therefore, small business becomes especially vulnerable to what may be the most serious negative impact of regulation, the absorption of severe opportunity losses occasioned by foregone development, innovation, and expansion. Congressional hearings have documented the highly significant contributions of small business to technology and productivity.[5] Yet this creative element is largely located in the owner/manager, the principal component of the small organization affected by the routine of regulatory compliance.

The Impact Issue

There is a substantial amount of published material dealing with the impact of regulation on small business. Some of this material is in the form of special studies, but the bulk of the information is contained in testimony before legislative committees. Something of a paradox is created by the character of the available impact documentation. On the one hand, there is sufficient factual data to establish clearly that small business is often subject to disproportionately high levels of capital, operating administrative, research and development, and product modification costs when complying with a wide range of regula-

---

[3] This definition is adapted from Roland J. Cole and Phillip D. Tegeler, "The Impacts of Government Requirements on Small Business in Washington State," Seattle, Washington: Battelle Human Affairs Research Centers, 1979, pp. 1 and 9.

[4] U.S. Small Business Administration, "The Impact on Small Business Concerns of Government Regulations That Force Technological Change," U.S. Government Printing Office, 1975, p. 13.

[5] "Future of Small Business in America," a Report of the Subcommittee on Antitrust, Consumers and Employment, Committee on Small Business, House of Representatives, 95th Congress, 2d Session, November 9, 1978.

tions. On the other hand, a principal contention of this paper is to the effect that the quality of information is at present inadequate for the purpose of generalized regulatory reform. In particular, existing documentation fails to provide cost-benefit evaluation in most instances; nor has much been done to create, strengthen, or apply a methodology of impact analysis. In general, conclusions are imputed from data collected for other purposes (such as the Economic Censuses), based upon the often subjective quality of testimony, or deduced from relatively simplistic surveys. If serious attention is to be given to regulatory modification and reform *vis a vis* small business, a prerequisite is comprehensive analysis at much higher levels of methodological sophistication than commonly employed at present.

Impact: An Operational Definition

The primary concern in this paper is with the specific effects of Government regulation on small business. These effects are the impact of interest and refer to costs imposed, benefits derived, and changes in operating behavior required of the small business when subject to regulation. This review of effects is limited to the "impact on" and does not deal with the necessity or desirability of regulation, nor is there any attempt made herein to assess the benefits to society which result from Government controls.

II. Regulatory Impact—the Conceptual Framework

Many of the thoughtful approaches to documentation and evaluation of regulatory impact deal with the business community as a whole and not only with the small business sector. Much of this more general commentary, however, is equally relevant to a consideration of the specific problem of impact on small business.

Basic Types of Regulatory Programs

The American Bar Association has noted that, traditionally, most efforts at regulation of economic activity can be grouped into one of five types: (1) Cost-of-service rate making (2) allocation in accordance with a public interest standard; (3) standard setting; (4) historically based price setting; and (5) historically based allocation.[6] Cost-of-service rate making has been applied most extensively to public utilities and common carriers. Allocation in accordance with a public interest standard refers to regulation for purposes of allocating scarce resources such as airline routes and radio and television licenses. Standard setting is employed in a great variety of situations where serious harm or injury to the public is anticipated in the absence of control. Environmental and safety standards are exemplary of such activity. Historically based price regulation is essentially price control across a broad segment of the economy. With widely varying costs and cost structures in evidence in such cases, a simple expedient is to fix the firm's price as of a past date. Finally, among the traditional

[6] American Bar Association Commission on Law and the Economy, Federal Regulation: Roads to Reform," Washington, D.C., August 1978, pp. 43–57.

forms of regulation, historically based allocation is employed by Government to ration commodities temporarily in short supply, using historical rates of consumption as the basis for control. In recent years, oil, natural gas, and water (in California) have been so allocated.

Other approaches to regulation exist and are receiving increased attention as alternatives to the more traditional methods just outlined. A relatively wide variety of possibilities exists. Illustrative of the concept, however, are various incentive-oriented approaches (taxes, subsidies) and other tracks based on the disclosure, dissemination, or persuasive presentation of information (jawboning).[7] There is a certain intuitive attractiveness to these alternative methods of achieving a desired level of control. It is hoped by proponents that their substitution for direct regulation might be both more effective and more efficient, thereby achieving a more uniform distribution of intended results with attendant lower costs of administration. The final proof of efficacy comes at the level of empirical evaluation, based on sufficient evidence for generalization. The available data on results achieved under non-direct methods of regulation are as yet inadequate for conclusive evaluation.

Regulatory Impact: The Broader View

Several studies have examined the evidence of burgeoning Government regulation. To date, the bulk of these efforts has dealt with Federal regulation of business exclusively, but it seems clear from informal observation alone that State and local controls are proceeding to grow, if not apace, at least at a significant rate.

Most recently, the total administrative costs of operating 41 Federal regulatory agencies have been placed at $4.8 billion for fiscal year 1979. This figure represents an increase of 115 percent over the $2.2 billion expended by the Federal establishment for regulatory budgets in fiscal 1974.[8] Yet these costs, which may be computed in straightforward fashion from the Federal budget, represent at most a small fraction of the total costs imposed on business by the presence of Government controls. A far larger figure, though less easily substantiated in precise quantities, is the amount expended by business in complying with Federal regulation. Murray Weidenbaum has set this compliance figure at $97.9 billion for fiscal 1979.[9] Other authors have discussed the limitations attached to the accuracy of this specific figure.[10] There is, however, little question that it represents a reasonable index of the share of regulatory costs in the economy borne by the regulatee. The total cost impact on business is not covered by the estimated $102.7 billion encompassed by the above figures for regulatory and compliance expenditures. The easiest costs to capture are the primary charges identifiable as part of the incremental cost of a compliance action. These costs have been categorized as follows:

---

[7] U.S. Department of Commerce, "Alternatives to Regulation: Possibilities; Experience; Prospects" (draft), August 1978, pp. 7–9.

[8] Murray L. Weidenbaum, "The Costs of Government Regulation of Business: A Study Prepared for the Use of the Subcommittee on Economic Growth and Stabilization of the Joint Economic Committee, Congress of the United States," U.S. Government Printing Office, 1978, p. 12.

[9] Ibid., p. 4.

[10] Julius W. Allen, "Costs and Benefits of Federal Regulation: An Overview," The Library of Congress, Congressional Research Service, 1978, pp. 17–18.

(1) Capital costs, including expenditures for portions of capital projects or for individual-fixed assets which are required to comply with a regulation.

(2) Operating and administrative costs required to operate facilities or equipment installed to comply with regulations; perform activities necessary to comply with regulations which affect the manner in which the manufacturing or other operating functions are carried out; complete paperwork; meet with government personnel; and perform other administrative activities.

(3) Product or process research and development costs, a category which includes the expenditures for determining and developing modifications needed for products or processes to meet regulatory standards, and for analyses of the impact of regulation on product costs, market share, and corporate image.

(4) Direct product costs, including material, labor, overhead and other costs incurred to conform a product to a regulatory requirement.[11]

Accurate assessment of these costs for all business would be a formidable task at best. Even given an accurate accounting, critically important categories of cost impact would remain unaccounted for: (1) productivity losses incurred where employees must devote all or part of their time to dealing with regulations (estimated by one source to be $2 billion for fiscal 1979 [12]); (2) opportunity costs, including foregone profits occasioned by delays in placing a product on the market at the opportune time and diminished managerial effectiveness owing to concentration on regulatory rather than operating problems; and (3) diminished profits and sales. It may well be that these difficult-to-assess costs are of particular relevance to small business, as noted later in this paper.

III. Regulatory Impact: Large Business

A review of the documented experience of large-scale enterprise engenders some feelings of frustration. Unlike the small business area, big business has tended to sublimate the impact of compliance requirements on the firm, substituting instead a concept of "the industry" as the focal point. There are few purposive investigations of regulatory impact on large scale business that can be accepted as objective measures of such impact. One notable exception is the recently completed study by the Business Roundtable of 48 member firms. Some existing studies attempt to assess impact on specific industries.[13] These analyses, however, tend to isolate intentionally one dimension of the regulatory scene, as in the case of the cited EPA reports, or to omit critical methodological information such as sampling procedures and data collection instruments. In the absence of investigations specifically designed to measure the impact of regulation on a cross section of industry, such as the study conducted by the Roundtable and cited previously,[14]

---

[11] Arthur Andersen & Co., "Cost of Government Regulation for the Business Roundtable: Phase I," May 1978, pp. 4–7—4–9.

[12] "Purchasing," Vol. 84, No. 11, June 7, 1978, p. 14.

[13] See, for example: Sterling Hobe Corporation, "Impacts of Government Regulations on the Iron and Steel Industry," U.S. Department of Commerce, August 1978; U.S. Environmental Protection Agency. "Economic Analysis of Effluent Guidelines: The Metal Finishing Industry," Washington, September 1974.

[14] Cost of Government Regulation Study: Phase I, op. cit.

are the only private sector evaluations available, and some conclusions can be drawn from isolated bits of information reported by business and to draw what conclusions we can from special, limited studies by and for a variety of Government agencies and legislative committees.

While inadequate for true generalization, there are some brief commentaries reported in literature that provide a bit of insight into the dimension of overall impact. Dow Chemical U.S.A. and General Motors Corporation have generated two such reports of comprehensive measures. Dow summarized a detailed internal investigation of regulatory costs for the year 1975. The overall cost impact of domestic regulative compliance was $147 million for that year.[15] Some components of these costs are interesting in themselves: $5 million in salaries and expenses for executives testifying on regulatory problems; $18.2 million of an overall research and development budget of $167 million went into regulatory compliance; $63 million of the total compliance bill was absorbed in meeting environmental standards; $45 million was "assignable" to transportation regulation; and $22 million was required to achieve required health and safety standards.

General Motors Corp., by its own account, spent $1.258 billion in 1975, and $1.3 billion in 1974 to meet the research, development, and administrative costs of regulatory compliance involving all levels of government. The Chairman of General Motors stated that it now costs the corporation $1 billion to meet each additional half mile of fuel economy required, and he estimates that similar gains by the mid-1980's could cost $3 billion to $4 billion.[16]

Industry-specific studies by non-governmental agencies are in some measure typified by two reprots on the iron industry (Sterling Hobe Corporation) and steel industry (Arthur D. Little).[17] Because of particular areas of interest or, perhaps, because of the enormous scope that would characterize a comprehensive analysis, studies by private sector contractors tend to focus on segments of the regulatory scene. The Sterling Hobe study, while rather general coverage is implied in the title, actually addresses the "effect of the plethora of regulations on innovation."[18] Also, it deals only with environmental and safety-health regulatory impacts on innovation. Citing secondary sources and interviews with industry representatives, Sterling Hobe estimates total capital expenditures for water and air pollution control equipment at $3.85 billion for the period 1975–77, and total operating and maintenance expenses for this equipment are estimated at $.4 billion over the same term.[19] Comparable figures are not provided for the assessment of Occupational Safety and Health Administration regulatory impact. The relative paucity of data is cited as the reason no quantitative analysis could be carried out.[20]

Concentrating on environmental regulations, Arthur D. Little, using industry records in a report to the Iron and Steel Institute, has provided expenditure data and estimates of future investment burden for

---

[15] Business Week, April 4, 1977, p. 50.

[16] Thomas A. Murphy, Chairman, General Motors Corporation, in remarks before National Tire Dealers and Retreaders Association, Detroit, September 16, 1978.

[17] Sterling Hobe Corporation, op. cit.; Arthur D. Little, Inc., "Steel and the Environment—Cost Impact Analysis: A Report to the American Iron and Steel Institute," Washington, May 1978.

[18] Sterling Hobe Corporation, op. cit., p. 2.

[19] Ibid., pp. 91–93.

[20] Ibid., p. 96.

environmental control through 1985. While currently running at about 20 percent of total capital investment expenditures, environmental controls investment is seen rising to 25 percent of total annual investment by 1985.[21]

The cited evidence offered by Dow and General Motors clearly offers some measure of impact. It is equally clear, however, that little generalization is possible given the absence of methodological specification. Regulatory compliance can, indeed, be costly if a firm is in an appropriate industry subject to a number of technology-forcing as well as paperwork-loading controls.

Regarding the type of evidence available from the Hobe and Arthur D. Little reports, some modest additional industry specific generalization can be gleaned because the data are industrywide on the one hand, and on the other hand. the sources are cited where data are displayed. Unfortunately, the methodological approaches are somewhat inconsistent with the desire for clearcut impact evaluation. Also, the purpose of each report being narrowly conceived, only a small segment of the overall impact of regulation on the industry is visible.

Studies performed by and for the regulatory agencies offer some of the best current sources of information on industry structure and the impact or potential impact of regulation on large scale business. In particular, the series of economic assessment studies conducted by contractors for the Environmental Protection Agency are exemplary. Unfortunately, these studies are designed to assess the potential rather than the actual impact of proposed regulations, and they are limited in each case to a specific subset of environmental controls. Therefore, they are of limited value where the objective is to capture the overall impact of regulation. At this point, too, there is limited coverage in the series of industries where large business organizations are a significant factor. Methodologically, however, these studies are suggestive of models for the future. Two examples involve forecasted impact of specific proposed EPA controls for the inorganic chemicals industry and the metal finishing industry.[22] With comparable objectives, the methodologies of the reports diverge owing in part, it would appear, to the vagaries of data availability. Both studies open with detailed factual material on the structure and characteristics of the respective industries. Analytical methodology is carefully prescribed, and the application of the analysis presented in detail.

Such *ex ante* attempts to project potential impact are fraught with error risk, of course. The task of the historical investigator should be simplified, given appropriate methodology: Historically, the impacts have occurred. The only problem is to ferret them out.

In the EPA study of the inorganic chemicals industry the empirical analysis is tied closely to the microeconomic theory of the firm.[23] Regression analysis is then employed to estimate demand functions utilizing available data on production, prices, and GNP.[24] Data problems obviated the use of a similar approach to estimate supply functions.

---

[21] Ibid., p. 87.
[22] U.S. Environmental Protection Agency, . . . The Metal Finishing Industry, op. cit.; and U.S. Environmental Agency, "Economic Assessment of Potential Hazardous Waste Control Guidelines for the Inorganic Chemicals Industry," Washington, 1976.
[23] EPA. . . . "Inorganic Chemicals Industry," op. cit., pp. VII 3–16.
[24] Ibid., pp. VII 17–22.

and an engineering cost estimation for 'model' plants was substituted.[25] Analysis of the industry produced estimates of the expected impact of hazardous waste management costs on producers in a large segment of the industry. Projected impact is developed for industry size and growth, employment, community effects, and foreign trade.[26] Data limitations and other caveats are carefully enumerated in this study, permitting a full evaluation by the user.

The study of the metal finishing industry for EPA contains similar attention to detail. In this instance the availability of sufficient industry data enabled the contractor to produce an impact assessment based largely on the analysis of empirical data. The detailed analysis of market, financial, and process data yielded estimates for nine categories of impact: (1) Industry volume; (2) operations; (3) customers and suppliers; (4) financing; (5) size of the firm; (6) plant closures; (7) employment; (8) annual costs; and (9) foreign trade and regional economy effects.[27] Again, this paper discloses data and discuss the methodology and limitations of the analysis. Metal finishing is an industry of small rather than large producers. The illustration remains valid from the standpoint of methodology considerations, however.

The most ambitious, comprehensive, and enlightening study to date of impacts on large-scale business units is that undertaken for the Business Roundtable by Arthur Andersen & Company.[28] Something of the magnitude of the general documentation task may be gleaned from the facts that this study, in two volumes, covers only 48 companies, 20 industries, and but six Federal Government agencies and programs. The report, which documents $2,621,593,000 of regulatory compliance costs by the participating companies for the year 1977, yields conservative estimates of costs for several reasons connected with the design and methodology of the study. First, the report includes only incremental costs as determined by the companies. Excluded are those associated costs which, in the judgment of individual companies, would have been incurred in the absence of regulation. For example, the study includes only the added cost of a required pollution control system beyond that which the company believed it would have installed on its own. Second, no costs of secondary effects of regulation are included because of the necessary lack of precision attending their measurement. Therefore the report omits such important impacts as opportunity costs, productivity changes, and the costs of regulation imposed delays. Third, as noted, only the effects of six agencies were assessed. No industry-specific regulations were evaluated, nor were any State or local impacts assessed.

Nevertheless, the Roundtable study is an important contribution to literature both from the standpoint of cost data provided and for the example it provides of rigorous design and methodology. The costs of such attention to objectivity and detail are staggering: Arthur Andersen & Co., alone devoted more than 400 professional months of effort to the study.

[25] Ibid., p. VII 25.
[26] Ibid., ch. 8, pp. 1–59.
[27] EPA . . . The Metal Finishing Industry, op. cit., pp. 19–26.
[28] "Cost of Government Regulation Study: Phase I," op. cit.

In summary, the available data on regulatory impact for large-scale business is generally far from clear, nor are there a sufficient number of studies extant to provide an adequate cross-sectional assessment. Regulatory compliance is a significant cost of doing business for the large firm. Even the relative dearth of reliable data enables one to draw a broad, simplistic conclusion to that effect. Such statements only serve to emphasize the real problem. If detailed quantified assessment is to be accomplished, studies will have to be performed with far greater attention to the requirements of scientific evidence than is currently the case in most instances.

IV. The Impact of Regulation on Small Business

Evidence bearing on regulatory compliance impact on small business is generally more available, comprehensive, and to the point than is the case with large-scale enterprise. In particular there has been more effort devoted to the specific issue of impact with respect to small business. Studies prepared for the Small Business Administration, hearings before legislative committees, analysis from academic institutions, and reports prepared by small business associations have all dealt directly with the question of impact. Most of this material deals with Federal regulation. Other levels of Government impose definite compliance costs on small business, however, and an initial objective of this paper involved segregating and providing measurements of the State as well as the Federal compliance burden. It has been possible to provide some insight in the incremental impact of State regulation, but in general the data base is incomplete and subject to methodological criticism. Therefore most of the following documentation relates to Federal regulation.

Overall Indexes of Regulatory Impact

Many organizations exist, public and private, which have devoted concentrated effort to isolating the specific problems prevailing in the relations of small business with government. Significant within the private sector are the Center for the Study of American Business at Washington University, St. Louis, and the National Federation of Independent Business (NFIB), San Mateo, California. Both of these organizations have conducted numerous investigations focusing on the particular issues involving small business. At the Federal level the most conspicuously involved units of Government include the Small Business Administration, the Committee on Small Business of the House, and the Select Committee on Small Business of the Senate.

Turning first to the contributions of the private groups to the assessment of regulatory impact on small business, some of the material developed by the NFIB provides an appropriate point of entry. Reporting the serial results of a membership survey conducted for some years on a quarterly basis, the Washington counsel of NFIB noted in prepared testimony:

> Over the past four and one half years, quarterly, NFIB has asked a random sample of its membership, "What is the single most important problem facing your business today?" Ten possible responses have been listed ranging from "Quality of Labor" to "Competition from Large Business." The results of this

quarterly polling show that small business problems haven't changed appreciably since the poll's inception. "Inflation" has consistently been singled out by about one in four reporting small firms as their most important problem. "Taxes" rank second, being cited by one in five. "Government regulation and red tape" is third with over one in 10 reporting it to be the single most important problem.

Somewhat surprisingly, government regulations and red tape appear to be a problem not isolated to any particular size small business or to any particular sector. The October 1977 report, for example, indicated 15 percent of the firms annually grossing $800,000 or more found government regulation and red tape to be their most important problem. Yet, 10 percent of the very smallest firms, those annually grossing $50,000 or less, arrived at the same conclusion. Viewing it on a sector by sector basis, NFIB found "Financial Services" and "Professional Services" and government regulation and red tape to be a greater burden than do others. Yet, 9 perecnt of those in "Construction" and "Manufacturing," and 10 percent of those in "Retail," "Agriculture," and "Non-Professional Services" had a similar principal problem.

These data do not mean that only slightly more than one in 10 small businesses perceives Government regulation and red tape to be a problem; they mean slightly more than one in 10 considers it to be the most important problem. How many would rank it second or third is a determination we are not able to make. But judging from membership mail, NFIB believes government regulation and red tape to be an important problem for portions of the small business community far beyond those which did not cite it first.[29]

One index of regulatory impact frequently overlooked in attempts to assess the burden imposed is the perceived level of impact by those who are regulated. The NFIB material cited sheds some light on the perception dimension. From the Center for the Study of American Business comes added evidence of the overall burden:

Government regulation tends to have a disproportionate adverse impact on small business. The limited ability of the smaller firm to pass along regulation-induced costs or to generate investment capital to finance those costs is having a debilitating effect on this segment of the economy. The special problems of the small firm in dealing with government regulation are managerial as well as financial. Government paperwork requirements command large segments of the entrepreneur's time and energy. Thus, appropriate consideration of the problems of small business should be a part of the analysis of any proposed regulation of the private sector.[30]

Expanding on these introductory points, Chilton noted the following:

The impact of the recent expansion of Federal Government regulation of business has been particularly severe on the smaller companies. Large capital expenditure requirements to meet environmental or job safety standards above those that would be followed voluntarily may represent merely an uneconomical application of resources for a large firm; but it may literally be a matter of the enterprise's life or death for the small firm. While surely there has been no deliberate intent on the part of the regulation writers to bring about the demise of the small business firms, it is accurate to say that often there has been little if any consideration given to the economic impact of regulation on the small firms...

In 1974, the National Association of Manufacturers surveyed its members and obtained estimates of the costs of meeting the OSHA standards existing at that time. The average estimate for firms with fewer than 100 employees was $35,000, a significant demand on the resources of a small firm. It is interesting to note that the only reference to small business in the whole legislative history of the Occupational Safety and Health Act is for the provision for Small Administration financial aid to improve work place safety.

[29] Prepared statement of James D. McKevitt before the Subcommittee on Special Small Business Problems of the House Committee on Small Business, March 9, 1978.

[30] Testimony of Kenneth W. Chilton, Assistant Director, Center for the Study of American Business, before the Subcommittee on Small Business Problems of the House Committee on Small Business, March 8, 1978.

Other regulatory agencies have created increased capital demands on small business. The Employee Retirement Income Security Act was designed to protect employee pensions. However, the Commissioner of Internal Revenue, Jerome Kurtz, estimates that, due to a propensity to specify uniform standards for all pension plans, large or small, as much as 30 percent of the Nation's 500,000 private pension plans may have gone out of business. An important contributor to this phenomenon is the estimated 5–10 percent increase in the costs to small business to meet pension law requirements. This increase is due to higher minimum funding and actuarial services and insurance for fiduciary liability.

Some Government regulations adversely affect the ability of small enterprises to attract investment. The "prudent man" rules included in ERISA coupled with their accent on personal liability for the "imprudent" pension fund managers have accentuated the trend to concentrate equity investments in the larger, well-established companies. Changes to these rules to allow up to 2 percent of a pension fund's portfolio in smaller, more risky firms while meant to provide some relief for small business may unfortunately merely place a ceiling on these types of investments by the pension funds.

Some of the Securities and Exchange Commission's rules also tend to work against the small firms. The SEC definition of a private offering is the solicitation of no more than 25 potential investors and sale of stock to no more than 10 investors. SEC registration for public offerings costs between $100,000 and $150,000 to prepare and takes four to six months. Table 2, from the report of the SBA Task Force on Venture and Equity Capital for Small Business, demonstrates the decline in the new issue market for small business (defined as those with net worth less than $5 million).

TABLE 2.—NEW STOCK ISSUES BY SMALL BUSINESS

	Offerings	Total amount (millions)
1969	548	$1,457.7
1970	209	383.7
1971	224	551.5
1972	418	918.2
1973	69	137.5
1974	8	13.1
1975	4	16.2

Source: U.S. Small Business Administration.

Table 3 shows the estimated cost (in 1969 dollars) of flotation of public and private debt issues by size of the issue. The difference in the two types of issue indicates, in part, the added cost of meeting requirements imposed by the Securities and Exchange Commission and ancillary accounting and legal costs associated with public offerings. It is clear that small firms (small issues) bear disproportionate costs whether private or public and that SEC requirements greatly compound the costs for small issues.[a]

TABLE 3.—COMPARISON OF COSTS OF FLOTATION AS PERCENT OF PROCEEDS PUBLICLY OFFERED AND PRIVATELY PLACED DEBT ISSUES, 1951, 1953, AND 1955

Size of issue	Publicly offered	Privately placed	Difference
$500,000 to $900,000	10.24	2.14	8.10
$1,000,000 to $1,900,000	8.00	1.52	6.48
$2,000,000 to $4,900,000	3.33	1.12	2.21
$5,000,000 to $9,900,000	1.53	.83	.70
$10,000,000 to $19,000,000	1.44	.63	.81
Over $20,000,000	1.22	.44	.78

Source: American Enterprise Institute.

[a] Ibid.

Returning to the previously cited material from the NFIB, we find more support for the above conclusions:

If a small firm can bear additional debt and if a bank can be located to supply the capital, the rates and conditions of the loan will be far more severe than for a large firm. Small business is not familiar with the prime rate; they never see it. But prime +5 points they understand very well. They are not familiar with the 20-year loan, but they understand the 5–7 year loan.

Testifying before the House Small Business Committee, Mr. Arnold F. Mazotti, Senior Research Officer for the Bank of America National Trust and Savings Association, spoke of small business' capacity to absorb nonproductive capital costs for pollution abatement. His analysis pertains equally well to other nonproductive capital costs. Said Mr. Mazotti:

"Unlike the majority of large businesses, the smaller lacked access to long-term, low-cost funds. The drain on cash flow for small business is therefore more burdensome. As a result, smaller firms experience a disproportionate impairment of operating flexibility and creditworthiness . . .

Conventional long-term financing for major corporations goes out 20 years or more for pollution, which reduces debt-service costs considerably on an annual basis. Instead of paying $5,000 a month for a facility, when you shorten it from 20 down to 5 years you are up into the $20,000 to $25,000 a month. Small businesses do not generate enough cash flow to meet that kind of obligation . . .

Perhaps the most visible effect of government regulation—and certainly one of the costliest—is the amount of paperwork that is generated by the regulatory process . . . Recently the Commission on Federal Paperwork estimated that paperwork costs the American Taxpayer $100 billion each year. While NFIB believes this is an inflated estimate, there is considerable evidence to support the Commission's contention that small business may spend between $15–20 billion a year on paperwork. (See table 5.)[32]

TABLE 5.—*Small business paperwork costs by agency*

Agency	Cost
Internal Revenue Service	$11, 310, 436, 964
Department of Labor	1, 603, 148, 404
Small Business Administration	626, 364, 165
Census Bureau	369, 748, 700
Other	1, 055, 181, 273
Total	14, 864, 879, 506

Source: Study of Federal Paperwork Impact on Small and Large Businesses. Paperwork Commission draft dated July 20, 1977, p. 22.

These quotations, from people who devote all or a major portion of their time to studying or representing small business, serve well to frame the general problem. While they contain relatively little in the way of actual data, few would contest the assertion that these commentaries reflect the frustrations and difficulties faced by small businesses attempting to cope with the perceived monolith of Federal regulation.

At the outset of this paper a distinction was drawn between technology-forcing regulations and those regulations where the primary requirement is reportorial, or paperwork generating. The comments of McKevitt and Chilton referred to compliance with both types. Wherever the reference is to regulation-induced capital investment, however, the referent regulation is generally of the technology forcing category. One major study has been designed to view the cross-sectional impact of such regulation on small business. The study, by Charleswater Associates for the Small Business Administration, is both revealing of significant impact data and exemplary in a general way of the type of assessment effort required if meaningful impact

[32] McKevitt, op. cit.

analysis is to be undertaken.[33] Charleswater identified 48 major pieces of Federal legislation which mandate technological change "through requirements having a direct and adverse effect on small businesses and involving some difficulty of implementation." [34] At this point it is sufficient to summarize the findings. Charleswater Associates, in developing estimates of capital expenditure requirements imposed by air, water, and noise pollution regulations, discovered that only four industries (primary metals, petroleum products, chemicals, and paper) accounted for two-thirds of estimated small business expenditures in these three regulatory areas.[35] In the course of the project the memberships of four small business associations were surveyed regarding impacts from environmental, worker protection, and product safety regulations. Results indicated compliance expenditures were generally under $10,000 over a three-year period, but nearly three-fourths of the respondents reported some cost increases from the compliance requirements.

Though economies of scale would indicate compliance expenditures would be inversely proportional to size, Charleswater reports the surveys showing the "smallest companies facing lesser impacts than larger small companies, no doubt a reflection of regulatory exceptions applicable to very small firms and of less strict enforcement and compliance at the lowest end of the size scale." [36]

While the Charleswater study provides valuable cross-sectional impact data, it only deals with a limited segment of the small business world. Because the paper focuses on technology-forcing regulations, it is primarily concerned with the manufacturing sector where the bulk of expenditures for environmental controls is required. The absence of similar data for wholesale and retail trends, services, finance, insurance and real estate, and for construction industries on a similar basis leaves as a residual an incomplete picture of the overall impact of regulation on small business.

Selected Industry Experience

A considerable amount of information pertaining to various dimensions of regulatory impact appears in the records of congressional committees, especially those directly concerned with the problems of small business. Aside from the general information already cited, there is a growing body of material relating to the experience of specific industries, and often as a corollary, the experience of individual firms. Testimony of this sort lacks a common methodological base; is often oriented toward a limited subset of problems; frequently contains a strong emotional vein which comes through clearly even in the printed record; accumulates over a long period; and occurs in somewhat random patterns across a variety of industries as hearings are conducted on diverse matters of interest. All these factors discount the ultimate value of the evidence as a basis for objective assessment of

[33] Charleswater Associates. "The Impact on Small Business Concerns of Government Regulations That Force Technological Change." Small Business Administration, September 1975.
[34] Ibid., p. 18.
[35] Ibid., p. 116.
[36] Ibid., p. 5.

regulatory impact. A number of these anecdotes will serve to illustrate both the problems just enumerated and the flavor of regulatory compliance assessment as seen in the hearing room.

Hearing: Overregulation of Small Business [37]

A valuable source of information to the legislator as a means of fingering the pulse of the business constituency, the testimony of small businessmen tends to be vividly descriptive of specific incidents:

> At the present time the company is engaged in providing approximately 200,000 blankets for the U.S. Government and Saudi Arabia. The company was originally prepared to manufacture 130,000 blankets, but after strong insistence by the Defense Department, the company finally agreed to increase this production by almost 50 percent. These negotiations and the final awards covering these two blanket contracts were delayed two months by arbitrary and unrealistic demands by another government agency. . . .
>
> I offer the observation that the impact of regulations enacted by Congress in the past decade has been particularly traumatic on small business. Because these enterprises are small, their administrators occasionally lack the sophistication found among the staff of large corporations . . . Because of its size, small business frequently has limited resources and staff to meet and keep pace with the expanding demands made upon it. . . .
>
> With respect to the question of "over regulation" I shall limit my remarks today to our experience with the implementation of Affirmative Action Plans for Equal Employment Opportunity. Equal opportunity is not new to Hayward-Schuster. It has been part of its personnel policies for many years. Its labor agreements with the Textile Workers Union of America follow the philosophy of equal treatment. We have yet to receive a grievance from employees or the Union as a result of discrimination because of age, sex, race, color, national origin, or handicap.
>
> But somewhere and at some time following the enactment of Equal Opportunity regulation, the 'intent' seems to have become overshadowed by overly broad interpretations and decisions of regulatory agencies. Quotas are established and goals are set based on the utilization of workers by sex and race in job categories. The plans are evaluated on a maze of statistical evidence including substantial details and additional record keeping including a Daily Log of Applicants, and lists of Job Refusals, Hirings, Trainees and Apprentice Promotions, Transfers, and Separations. All these data are to be analyzed as to sex, minorities, and job categories, and reported to the regulatory agency on a quarterly basis.
>
> In 1974, two company administrators concentrated their entire time for two months in an effort to meet the initial requirements of the regulation. Subsequently additional personnel was engaged in the Personnel Department to assist in the implementation of the Plan. In 1976 the same two administrators devoted another two months in restructuring the existing plan and validating records in keeping with guidelines and expanded reporting now prescribed by the regulatory agency. This restructuring was to accommodate the requirements under Affirmative Action which was made a part of the new Saudi Arabian contract previously mentioned.
>
> The company was subjected to what it felt to be an inflexible attitude on the part of the regulatory body particularly with regard to what seemed to be failure to recognize the untenable position of Hayward-Schuster management. As a small concern, the company could not force immediate Union action. The abrogation of long existing contract language would result in a strike affecting the entire plant. Meanwhile, the award of contract by the Defense Personnel Support Center which was being delayed represented a very small percentage of the total plant capacity. . . .[38]

---

[37] "Overregulation of Small Business," Hearing . . . "Select Committee on Small Business," U.S. Senate (94th Congress, 2d Session, April 26, 1976).

[38] Ibid., pp. 67–71. Statement of Winfield A. Schuster, President-Treasurer, Hayward-Schuster Woolen Mills, Inc.

There is a wealth of information contained in the statement. References to delays, lack of expertise, contract language, statistical evidence, records requirements, and resource application to compliance activity are all important points. But no matter how many such statements are made, as a matter of record they form a poor basis for congressional or agency action. Instead, they may be viewed as a valuable source of research objectives and hypotheses to be tested with standardized data from adequate samples.

Again, from the same hearing, a spokesman for an industry:

> Associated General Contractors of Massachusetts, an organization representing over 100 general building contractors who put into place over three-quarters of a billion dollars annually. . . .
>
> My presence here is to speak briefly, but realistically, about the mountainous outpouring of regulations emanating from both National and State regulatory bodies which confront us, trip us, bind us and, perhaps, someday in the near future, will prevent us from doing what we do best . . . Build and rebuild America. . . .
>
> We explain the regulations, answer questions, and prepare summaries of the regulations. Our members complaint is that they cannot keep up with the summaries. . . .

To be specific:

> 1. *Equal Employment Opportunity (EEO) regulations.*—Attached is the Table of Contents of an EEO Manual we recently prepared and issued to our members. This manual includes only the most basic Federal, State, and city regulations. Tab 23 in the Manual is entitled "List of EEO Documents not included in this Manual"; the list, which is attached, covers the page, single-spaced.
>
> 2. *Safety.*—Attached are the Massachusetts Construction Safety regulations, which were in force before OSHA. They were recognized as sound, adequate regulations. OSHA's Construction Regulations are 94 Federal Register pages. This does not include the general-industry standards (249 Federal Register pages), to which construction contractors are also subject. An additional burden on contractors are the referenced standards in the OSHA construction regulations. We have purchased a partial set, at a cost of $337; we know of two other such sets in Massachusetts. Yet, OSHA continues to cite and penalize contractors for violations described only in these referenced standards. To make sense of the OSHA construction regulations, we prepared in early 1973 a summary of the regulations. OSHA issued a summary in June 1974, tacitly acknowledging the burden of their regulations.
>
> 3. *Environment.*—In January 1973, AGC of Massachusetts issued to its member firms an "Environment Guide for Construction," a 100-page document containing the environmental regulations governing construction contractors. It is now outdated, and we are endeavoring to update it. As was testified at a state legislative hearing last week by a representative of the New England Construction Users Council, "The licensing process in effect today on both the State and Federal level can, at best, be described as a nightmare of complex procedures involving multiple overlapping and frequently duplicative steps." The vast majority of the licenses referred to are environmental. An unknown number—but judging from today's market in Massachusetts, it is vast—of projects never get started because the developer is asked by his lending sources "Do you have your environmental clearances" before he knows whether or not he has financing. For a typical residential building in Massachusetts, 49 permit steps are required, of which 15 are environmental.
>
> 4. *Environmental Protection Agency (EPA) Construction Grant regulations.*—The regulations governing the $18 billion construction program of the Environmental Protection Agency to clean up the Nation's waters has an incredible number of regulations bearing primarily on municipalities and design firms, many of whom are small. To monitor the EPA program and to assist in getting work out, we formed a committee. To get a grasp on some of the regulations, we put together a working manual; attached is the Table of Contents which indicates 20 pages. This omits the 67 Program Guidances (known as "PG's") issued by EPA as clarifying instruction to the state and municipal agencies and the design firms.

The above listing of "over-regulation" is but a sample . . .[39] As in the case of the previously cited testimony, numerous areas for investigation are pinpointed by the contractors' representative. Of special note here, in addition to the litany of Federal requirements, is the allusion to State and local regulations, in profusion, added to those emanating from Washington.

Hearing: The Effects of OSHA on Small Business[40]

Testimony from a state group representing a cross section of that state's industry:

> Under the present status of the law inspectors must inspect on the basis of whether an establishment meets standards, not on the basis of whether "the establishment is safe or unsafe." There is no official way for an employer of knowing as to whether or not he is in compliance with the law. This one standard package which we are presently dealing with covers all employees regardless of the type of the business or items manufactured or sold and it has, in fact, created more problems than it has resolved. And as far as I am concerned, for a law such as the Occupational Safety and Health Act to be both effective and equitable, the inspections or compliance should be based on the total overall quality of safety and health in the workplace, not merely the number of standards which are not being complied with. It seems to me that asking a small employer to comply with the thousands of pages of standards, many of which are couched in technical or scientific terms, is not unlike strapping a novice into a high-performance aircraft and asking him to take off, do sophisticated maneuvers in the air, and then land this plane safely . . .[41]

This statement and those in the preceding section are fairly representative of industry inputs to hearings on small business problems. Other types of testimony also appear, including technical papers from academic researchers and detailed responses from the administrative agencies involved.

Summary Evaluation of Small Business Regulatory Impact Materials

As with large-scale business, the overall impact of regulation on small business is significant. If the broad gauge comments of the NFIB, the Center for the Study of American Business, and the statements of those participating in legislative hearings could be accepted without serious question, it could be safely concluded that the level of regulation is excessive and injurious. Such may be the case, but, unfortunately, the evidence is not adequate to permit a comprehensive, objective conclusion to remedies. A major part of the problem lies with the methodologies employed in the collection and analysis of evidence. The methodological issue is addressed in the recommendations section of this report.

Benefits Accruing to Small Business From Regulation

Benefits accruing to small business from regulation, such as licensing and census reports, may yield benefits to small business in the form of fairer standards of competition and better information for decisions.

---

[39] Ibid., pp. 274–78. Statement of the Associated General Contractors of Massachusetts.
[40] "The Effects of the Administration of the Occupational Safety and Health Act on Small Business, Hearings . . . Permanent Select Committee on Small Business." House of Representatives (93d Congress, 2d Session, February 8 and 9, 1974).
[41] Ibid., p. 59.

Aside from these somewhat intangible results, there are direct economic benefits resulting from certain regulations, particularly in the technology-forcing realm. The data have not been studied extensively, but Charleswater Associates has produced some interesting estimates of small business sales generated as direct and indirect results of pollution and noise control compliance expenditures for equipment. Total sales accruing to the small business sector for the period 1975–80 are estimated at slightly over $17 billion. The small business segment of the industrial machinery and equipment sector (SIC 356) is projected at $7.7 billion in new sales arising directly and indirectly from pollution and noise control investment. Other industries expected to benefit by over $1 billion in added sales during the period are the small business segments of construction, maintenance, and repair; stone and clay products; primary metals; heating, plumbing, and structural metal products; and other fabricated metal products. These benefits should appear as offsets against estimates of total compliance expenditures for the affected industries.[42] Of course some care must be exercised in determining "benefits" to be offset. Sales dollars are probably an inappropriate measure in that such figures include materials and other purchased inputs which are simply passed through to the purchaser. A better figure, though posing some problems of calculation, would be value added, or sales net of purchased components and services. This approach would measure benefits to all members of the firm—owners, managers, and operatives.

The State Incremental Impact on Small Business

The pyramid of regulation includes significant elements of State and local, as well as Federal, controls. The burden is unevenly distributed geographically, reflecting state government attitudes nationally. Within states the regulations and taxes of counties and municipalities create substantial distortions in the impact pattern.

The Federal Government has consciously incorporated state controls and enforcement machinery in its own approach to regulation. This explicit recognition of state jurisdiction may occur for two principal reasons: (1) To improve or ensure compliance by industry within the state; and (2) to gain a measure of efficiency in the inspection-enforcement process through the avoidance of duplication.

This Federal-State regulatory union is most apparent in the areas covered by the more recent technology-forcing statutes. Presumably impact is not affected by these models of interlocked Federal-State jurisdiction, except that distribution effects tend to be equalized across state boundaries. This effect by itself is significantly positive to those states where there is sympathy for stringent standards. Under a system of state administration of Federal standards (as in the case with OSHA, for example) these states are at least assured of a floor below which other state's standards cannot fall. The condition mitigates a potentially destructive source of competition for economic activity.

[42] Charleswater Associates, op. cit., p. A–9.

State Regulatory Burden—The California Assessment

California has made a modest formal attempt to assess the impact on small business of State regulatory activities. The study was broad in scope of business activity covered but limited as to sample size (14) and the inquiry was further constrained to paperwork processing costs. The results indicated that "paperwork costs for the typical small firm are estimated to be between $3,000 and $4,000 per year." [43] Given that the largest firm in the sample had sales of $2.4 million annually and estimated paperwork costs of $1,885, or .07 of 1 percent of sales,[44] the impact appears significant but hardly burdensome. The detailed breakdown for this manufacturer appears in table V-1.

More significant, perhaps, is that the cost of paperwork is disproportionately higher for smaller business. A service station reported costs of $3,500; a drug manufacturer with sales of $600,000 estimated costs of $2,600 annually, or 0.43 of 1 percent of sales; a logger with annual revenues of $250,000 incurred costs of $3,840 (a substantial portion of this expense was incurred on industry-specific reporting

TABLE V-1.—HOURS SPENT BY A SMALL MANUFACTURING BUSINESS ON STATE PAPERWORK

Form No. and description	Hours	Category subtotals	Percent of paperwork processing time
A. BASIC			
1. Corporate:			
W-2: Statement of earnings	10.5		
W-3:			
W-2 summary	.3		
Franchise payment:			
Prior year	.1		
Current year	8.0		
Sales tax	18.0		
Corporation declaration	.3		
100, attach: Annual franchise tax	42.0	79.2	50.4
2. Licenses: 77R-2: Auto repair license	.3	.3	.2
3. Employees:			
596,599: Information returns	7.0		
DE3: Quarterly payroll	28.0		
DE3M: SDI, SIT deposits [1]	.5		
DE1080: Ruling	6.8		
DE1101C: Notices, new claims	10.0		
DE11906: Notices	3.3		
DE1296B: Benefit audits	.8		
DE1545: Benefit amounts	.8		
DE2503: Notice—Disability	.2		
DE3423C:			
Notices	2.0		
Requests for disability wages earned	2.0	61.4	39.1
B. REGULATORY			
CAL-OSHA report	12.0	12.0	7.6
C. MISCELLANEOUS			
FTB290M: Order to withhold wages	4.2	4.2	2.7
Total		157.1	100.0

[1] Acronyms unknown by the author. They refer to California employment reports.

Source: "The Cost of State Paperwork Requirements to a Small Business," Assembly Office of Research, California Legislature, July 1976.

[43] California Legislative, Assembly Office of Research, The Cost of State Paperwork to a Small Business, July 1976, p. 7.

[44] Ibid., pp. 7–8.

requirements such as timber harvest plans). Finally, there was some indication in the limited data that firms with "accountants" and/or well developed management information systems incurred the lowest processing costs.[45] While this result could, in fact, have stemmed from more efficient information processing capability, it may reflect only a perception brought about by the intercession of a buffer (the accountant or computer) in the system.

The New York Experience

The fiscal 1978–79 New York State budget contained, as a conservative estimate, $128 million for business-related reporting and regulatory activity. The state budget for the 11 departments, commissions, and boards involved in business regulation also contains in many cases substantial amounts for nonregulatory work. The overall budgets for the agencies involved totaled in excess of $500 million.

Weidenbaum has established a ratio of 1:20 between Federal regulatory budgets and business costs of compliance.[46] If the ratio holds also at the state level, business in New York State absorbed $2.5 billion additional for compliance with state regulations in 1978–79 beyond their expenditures to comply with Federal requirements. Employing the same assumptions, the burden of Federal compliance for New York business should be in the realm of $8 billion annually, or about $19,000 per establishment. Thus, the per facility average cost of compliance totals approximately $24,000 in the state, excluding costs of local government regulatory requirements. While this is a significant figure it is only an average and as such obscures the wide range of costs that were obtained across different industries. Real compliance costs for a small retail service establishment in New York may not exceed a few hundred dollars each year. By contrast, manufacturers in, say, primary metal processing may incur paperwork, operating, and capital costs of hundreds of thousands of dollars in a given year.

A survey conducted for this paper of a small sample of 17 New York business firms revealed that most companies at the present time are unaware of the full impact on their operations. While management typically can recount anecdotes relating to specific instances of significant capital expenditures, paperwork requirements, and procedural difficulties, none (at least in this small sample) are able to provide a record of costs or man-hours expended without recourse to special study. One manufacturing firm, however, was selected and asked to cooperate in inventorying the recurring reporting requirements for Federal and State agencies. The company involved is a high technology, relative diversified manufacturer with 850 employees. It may be considered fairly typical of corporate manufacturing operations in New York. The resulting inventory appears as follows:

[45] Ibid., p. 8.
[46] Weidenbaum, op. cit.

Agency	Form	Estimated man-hours
Securities Exchange Commission	10-K	[1] 600
Do	Proxy statement	5
Do	10-Q	32
Commerce	RD-1 survey of industrial research and development	2
Do	NC-X1A report of organization	1
Do	MA38B current industrial reports	1
Do	BE-452S plant and equipment supplement	4
Do	BE-462 plant and equipment expenditures	4
Do	BUS-80 current service trade report	1
CASB	Cost accounting standards board disclosure checklist	160
Federal Trade Commission	Form MG-4/78	1
New York	Franchise or property tax returns and quarterly estimates	10
Michigan	do	3
Delaware	do	1
New York	Job incentive tax credit application	60
Internal Revenue Service	Federal tax return and quarterly estimates	80
Do	Form 5500 pension plan disclosure	2
Wage related	Form 941—Wages subject to social security (FICA) and Federal withholdings (quarterly).	12
	NYS—WRS—Quarterly wage report to NYS	8
	Form 940 508—Federal unemployment	8
	Employees report of contributions—NYS unemployment	8
	NYS quarterly withholding statement	8
Personnel related	Employees industry and location report	4
	Federal affirmative action plan	100
	Federal EEO-1	10
	State wage survey	20
	Federal wage and salary review	20
	State labor report	2
	NYS job bank listings	1
	NYS job bank closings	1
	Pension Benefit Guaranty Corp. disclosure checklist	10
	Grand total	1,179

[1] Includes company time participating in public audit as certification is required on form 10-K.

The company reported that these man-hours translated into salary costs of approximately $14,000 yearly. The estimate noted parenthetically that the firm had spent 300 man-hours analyzing data to meet compliance with the President's Wage and Price Guidelines—another $3,500 in salary costs that may become an annual requirement.

Approximately 120 man-hours in the list are attributable to recurring State-level reporting requirements. As with Federal regulation, this State total may increase dramatically where there is substantial environmental and natural resource involvement. As might be expected, manufacturing firms in the sample were particularly sensitive to the issue of regulatory compliance burden since it is in this sector that environmental standards and occupational health and safety requirements primarily have an impact. Wholesale and retail distributors of packaged and semi-packaged products in the sample, while hardly enthusiastic about the obligations, did not feel unduly burdened. The principal complaints of this group centered about perceived redundancy in reporting requirements, on the one hand, and, nonrecurring problems involving contract negotiations with both Federal and State agencies on the other hand.

The data reported, whether from California or New York, are from samples too small to be considered adequate or fully representative. The absence of capital and nonrecurring costs implies the need for formal investigation of these elements as well as a paperwork burden across a broad cross section of economic activity. The State data, however, do reflect the fact that State requirements are a significant dimension of regulatory impact. It would seem that future evaluations

of compliance requirements must give due consideration to all levels of government if a complete and accurate impression is to be gained.

V. Recommendations

Originally, the principal intent of this paper was to summarize the existing information on regulatory impact relating to both large- and small-scale business. This summarization would provide a basis for a set of explicit measures of the differential impact of Government regulation on these two sectors. The achievement of such a specific objective was frustrated by the fundamental weakness of available data. On one side are mountains of testimony and subjective statements from Government administrators and the regulated businessman which insistently remind us that the regulatory burden is onerous, excessive and thoroughly in need of reform. One is intuitively led to agree, given the great weight of assurance that the need exists. On the other side, however—the side of objective review and evaluation—there is a veritable dearth of adequate, representative data to support the advocacy of massive reform. There are also no measures of impact that unambiguously assess the burden on the various sectors of business activity.

The Primary Methodological Requirement

Assuming that there is a perceived need by Congress and appropriate administrative agencies for impact measurement that distinguishes among size categories of business activity, the pressing requirement is for a methodology of data collection, retrieval, classification, and analysis. Or, looking at the problem from another dimension: sufficient theory exists to guide the empirical investigation. Microeconomic models clearly specify anticipated firm behavior in response to impacts on costs and revenues. Refinements in the basic model, to allow for organizational behavior changes, for example, and to account for secondary effects such as opportunity costs incurred as a result of compliance activity may be "worked in" gradually.

The timing may be especially propitious in 1980 to initiate a standardized model for impact data collection. Presumably the 1977 Economic Census data will be available in full that year. This material will provide a wealth of information that can be employed in establishing a comprehensive program at reasonable cost. Minimum requirements for the data base entail the ability to classify by line of business (appropriate SIC code) and to segment lines of business by size (employment, sales, and value added per establishment), age of facility, geographic location, and ownership characteristics. This information is available in secondary Federal sources, and the categorization indicated should be adequate for the task at hand.

Data on costs of compliance with Federal and State regulations can be secured from carefully designed samples of lines of business at reasonable expense, especially if care is exercised to reduce respondent reporting requirements to a minimum. Nonetheless, the process of collecting and analyzing empirical data for a specific purpose as envisioned here is time consuming and always entails significant costs. An administrative decision must be made as to the need. The information available seems to indicate that special and, perhaps, severe problems are created for at least some small business by the regulatory

burden imposed. That fact may be sufficient justification to implement the paper. It would be reasonable, at least, to initiate a series of pilot studies of major lines of activity (industrial, trade, service, construction, and finance) both as a means of securing initial data and as a test of survey methodology.

Secondary Methodological Considerations

There is a vast potential for specialized studies of regulatory impact. Most such inquiries would require a broad range of disciplinary skills among the investigators. The first logical extension of cost studies would be to net cost and cost-benefit analyses. While cost analysis alone may be used to eliminate obvious inequities and inefficiencies in the regulatory process, the ultimate decisions on regulatory reform must rest on public benefit. Analyses of significant secondary impacts such as opportunity costs and delays imposed on innovation are areas that could be explored fruitfully.

Nothing has been done regarding the behavioral implications of regulation. Impacts may be viewed from this dimension as well as the economic effects. Such inquiries may well include the changes imposed by regulatory requirements on organization structure, allocation of managerial time, entrepreneurial style, and the design of jobs and work.

Regulatory Reform: A Plea for Moderation

The generally non-industry-specific character of existing impact analyses suggests that reform efforts should proceed with caution. Dramatic changes in methods of regulation (e.g., removal of reporting requirements, full substitution of subsidies, taxes, and moral suasion for standards, inspection, and review), insofar as these changes are based on a size of enterprise distinction, are best based on impact evaluation.

Given the existing interest and motivation underlying pressure for regulatory change and reform, initial actions in this direction should be restricted to the streamlining of administrative processes rather than to the fundamental revision of the basic posture toward regulation. This admittedly conservative approach implies that, in the absence of impact studies, reforms should concentrate on paperwork reduction, agency staff improvement, and improvement in operating systems. Major structural changes in the existing character of regulation should await objective impact assessment.

These recommendations may appear tentative and narrow in scope. Indeed, they have been criticized as such by members of the Federal establishment. These reviewers object strenuously to what they consider to be "weak" recommendations that do not point the way to dramatic reconceptualization of the regulatory and evaluative process. Further study is seen as an unacceptable alternative; instead, we are asked to forecast the future (from inadequate data) if present conditions continue, and to state categorically the dimensions of the economic problem that exists. On reflection, such criticism seems to further illuminate the problems this paper has found to be paramount.

It is hoped still that major reform will await the support of adequate evaluative data. There is a plethora of administrative relief to be attended to in the interim.

THE ECONOMIC EFFECTS OF FEDERAL REGULATORY ACTIVITY ON STATE AND LOCAL GOVERNMENTS: CONCEPTS AND PROPOSALS

By Robert E. Firestine

CONTENTS

Page

I. Some Useful Economic Concepts — 296
- A. Federal Regulatory Requirements: Programmatic and Procedural — 296
- B. Efficiency and Equity — 297
- C. Benefits — 298
- D. Costs — 299
  - 1. Expenditures — 299
  - 2. Caveat on Factor Costs and Other Complications — 300
- E. Direct and Indirect Economic Effects of Federal Actions — 301
  - 1. Direct Effects — 301
  - 2. Indirect Effects — 302

II. Grants-in-Aid — 303
- A. Structural Considerations — 303
  - 1. Federal Requirements: Grant Related and Nongrant Related — 303
  - 2. Programmatic and Procedural Requirements — 303
- B. The Focus on Costs — 306
- C. State-Local Cost Problems by Type of Grant — 307
- D. Some Complications in Producing a Detailed Cost Analysis of Grants-in-Aid — 309
- E. Recommendations — 310

III. Regulations — 311
- A. Two Kinds of Procedural Costs — 311
- B. More Problems in the Measurement of Costs — 312
- C. Problems in the Attribution of Costs to the Federal Government — 314
- D. Improving the Cost-Effectiveness of Procedural Requirements — 315
  - 1. Intrapolicy Priority Setting, Combined With Demonstration Experiments — 316
  - 2. Federal Reimbursement of the Policy Compliance Costs of State and Local Governments — 318
  - 3. Structural Modification of Certain Federal Policy Requirements — 319
- E. Regulations Affecting the Conduct of Federal Monetary and Fiscal Policies — 320

IV. Conclusions and Recommendations — 321
- A. Conceptual Issues and Problems — 321
- B. Recommendations — 323

I. Some Useful Economic Concepts

A. Federal Regulatory Requirements: Programmatic and Procedural

Newspapers these days are filled with stories or commentaries on the burgeoning costs and magnitude of "Federal regulation" within the American economy. Under the unrelenting pressures of continu-

ing inflation, where price increases easily outrun income growth, many are asking whether the burden of well-intended government intervention in our economic lives has not perhaps become excessive, whether the real costs of some government policies have not perhaps begun to outweigh their real benefits, and whether some reasonable way might not be found to redress this seeming imbalance. Nowhere is this issue better illustrated than in the question of Federal "regulatory" impacts upon State and local governments. On the one hand, the non-Federal public sector—especially America's large municipalities—is itself often strapped for adequate operating funds and is nonetheless reluctant (or, in some cases, virtually unable) to seek higher tax revenues from its citizenry. On the other hand, State and local governments are reacting with increasing hostility to the reportedly ever-growing Federal demands for program documentation, statistical information, interagency coordination, and confirmation of compliance with Federal policy requirements. This situation is particularly ironic since a substantial proportion of these reactions seems to derive from efforts to comply with the stipulations of Federal aid programs upon which so many State and local governments have come to depend.

From the perspective of State and local governments, then, Federal regulation really encompasses that panoply of Federal policy requirements by which the presumed intent of the Federal Government is communicated to the non-Federal public sector and through which intergovernmental initiatives are enforced. These Federal policy requirements may take the form of programmatic and procedural directives for the receipt of intergovernmental aid, legislative or judicial mandates prescribing standards of State or local government performance, or other guidelines and regulations which run the gamut from programmatic strictures to detailed stipulations on the reporting of compliance with administrative procedures. Indeed, rather than categorizing these Federal requirements by the vehicles of policy (grants, mandates, or whatever), most of the rather scanty literature in this field seems to approach the problem from the standpoint of the character of the requirements themselves. In this vein, much of this paper will juxtapose the programmatic against the procedural nature of Federal policy requirements. Regardless of the specific vehicle employed, it is these essential programmatic-versus-procedural characteristics which appear to offer the most useful approach to the operational evaluation of the State-local costs of Federal policy requirements.

B. Efficiency and Equity

Preeminent among the major elements of any conceptual examination of the costs of Federal policy requirements is the inherent conflict between efficiency and equity. Efficiency, of course, encompasses (*a*) the technical efficiency objective of "least-cost" provision of public goods (or public policies) and (*b*) the allocational efficiency criterion that is directly concerned with potential distortion of State-local fiscal preferences. Distributional equity, in contrast, refers to the less easily defined objective of fairness in the treatment of individuals whose economic circumstances may differ one from the other.

While the commonly understood idea of least-cost efficiency may be generally accepted as a desirable objective, the concept of allocational

efficiency is a bit more complex. Herein, it refers primarily to the presumed distortion of State or local government fiscal priorities as a result of Federal grant incentives or other Federal policy requirements. Such community tax or expenditure distortions can occur whenever different amounts of State or local public funds are allocated to a given activity, as a result of Federal intercession, than would otherwise have occurred. In such a situation, one measure of the cost of the Federal action to the community is the net value of those public services (or private goods whose consumption was sacrificed to higher taxes) which were necessarily foregone to redirect State or local efforts toward the federally favored expenditure objective. These community budgetary distortions are associated with the so-called programmatic costs of Federal policy requirements which are referred to throughout this paper. Many allocational effects of this nature are justified by the need for national government intervention in State-local affairs to guarantee minimum standards of economic well-being for all Americans, to guarantee the minimum fiscal capacity of essential units for State or local government, or to redress certain externalities which arise among jurisdictions from the natural interstate migration of financial or human resources.

The equity objective can often conflict with these efficiency criteria, since "fair treatment" may easily imply the adoption of policies which do not yield the least-cost production of public goods and which introduce community budgetary distortions. Specifically with reference to the Federal policy requirements addressed here, the so-called national policy objectives [1] generally function as across-the-board requirements imposed upon State and local governments in the form of separate policy directives as well as comprehensive provisions of virtually all grant-in-aid programs. As identified by the ACIR, the national policy objectives fall into seven broad classifications:

Nondiscrimination,
Environmental protection,
Planning and project coordination,
Relocation and real property acquisition,
Labor and procurement standards,
Public employee standards, and
Access to Government information and decision processes.

Thus national policy objectives play a central equity role in the current debate over Federal regulations since they are incorporated into many Federal programs whose concerns are often allocative rather than distributive. As such, they seek to impose upon State and local governments equity goals which may be enforced separately as autonomously operative requirements.

C. Benefits

The economist's approach to the measurement of benefits is related to growth in economic well-being (probable net increase in personal incomes over a lifetime) as a result of implementation of the program. The key to this is the "attributability" of benefits to a govern-

[1] Advisory Commission in Intergovernmental Relations (ACIR), "Categorical Grants: Their Role and Design" (Washington, D.C.: ACIR, May 1977), ch. VII.

ment program, regardless of whomever may receive these benefits and whether some of the benefits may have been unintentional. Furthermore, it is vital to address the question of incremental benefits (net change in economic well-being) brought about by a program rather than by total benefits (in this instance, total earnings) which exist after enactment of the program.

The popular perception of public service benefit flows is much more elusive than is the general perception of government outlays for those benefits. That is, it is easy for the general public to overlook the benefit flows from many programs, but the tax costs of those same programs are always apparent to all.[2] The nonuniversal distribution and delayed occurrence of benefit flows limit their impact upon the public consciousness. Also, the program-related cost savings which accrue via the avoidance of social problems (and public expenditures to deal with them) are difficult to gauge, either because (*a*) some policies merely retard the growth of such problems rather than halting their expansion altogether, or (*b*) implementation efforts in one functional area (e.g., OSHA or EPA regulations) may well produce uncatalogued results in other areas: Reduced private medical expenses, lower public welfare and workmen's compensation outlays, improved general health levels and longer life expectancy.

D. Costs

A similar perspective should be taken in the evaluation of the cost effects of Federal policy upon State and local governments. The costs so examined should be the incremental costs of Federal action, with care being taken to distinguish those costs which would have occurred anyway, even in the absence of a Federal policy initiative. Additionally, those costs which are related to a particular program could conceivably be attributed either to the Federal Government or possibly to individual State governments which might well enforce policies similar to Federal initiative in a given policy area. Thus the correct portion of the incremental costs of, say, environmental protection policy, should be properly attributed to the Federal and State governments, as appropriate.

1. EXPENDITURES

It is in this sense that major errors of analysis are made in the easy equating of program expenditures with policy costs. In any attempt to assess the cost effects of government policies, it may seem at first that the most expeditious approach would involve a review of government expenditures by program. However, despite the titles of individual program activities, the multiplicity of actual policy goals which are incorporated into any one program must be carefully sorted out in each case. Failing this, the actual policy costs of a singular Federal objective are quite likely to be overstated by counting the costs of other policies whose objectives have been imposed upon the operation of that particular Federal program. For example,

[2] In fairness to each side of this issue. it must be pointed out that some economists believe that it is costs. rather than benefits. which may not be fully perceived by the layman. See Murray Weidenbaum, "Measuring the Costs of Regulation," Washington Post, Feb. 3, 1979, p. A–19.

56-368 O - 80 - 20

it would be tempting to allocate to "the cost of compliance with Federal regulations" the salary outlays of local school officials whose sole duty (as in large school systems) is to fill out forms required for the receipt of Federal grant-in-aid funds. This, clearly, is an inaccurate accounting of such "paperwork" or "compliance" costs, since the local district would need to report some basic data to qualify for any Federal support. Beyond that, of course, there exists a clear national interest in the compilation of certain basic statistics on the public school population in the United States—information which must originate at the local level. The local expenditures which support the provision of this information could certainly be called costs of Federal paperwork, but to ascribe them to the Federal grant-in-aid process alone would surely be to overestimate the local compliance cost of that process. The assignment of policy costs by reviewing programmatic or budgetary expenditures can easily produce an upward bias if this is in the interest of the individual producing the estimates.

2. CAVEAT ON FACTOR COSTS AND OTHER COMPLICATIONS

With regard to the concern for State-local government costs of Federal policies, an additional observation should be made about the term "costs." As used in the traditional cost-benefit evaluation of government programs, costs refers essentially to costs of output: expenditures on a given activity. An entirely different usage can emerge, however, when one turns to the economic effects of Federal policies on State and local governments. In this instance, the costs of input are important unit prices of individual factor inputs which are employed by State and local governments in producing the public services which are consumed by the local constituency. Clearly, from the State and local perspective, Federal policies can affect either of these costs. The latter, incremental input costs, are really a price effect facing State-local governments: if a Federal policy produces higher input prices facing the State and local public sector (as is widely alleged for the Davis-Bacon requirement that wages on federally supported construction projects equal the local prevailing wage, which is usually the local union wage), then an incremental price increase of these inputs may be attributed to that Federal policy. Often such an incremental price effect will be accompanied by a substitution effect, wherein the increased input price discourages the purchaser from acquiring as much of the affected input as he would otherwise have purchased, substituting instead some other combination of factor inputs or—where possibilities for such substitution are limited—simply reducing the level of planned output in accordance with the perceived increase in the price of the affected input. In our example, the hypothetical Davis-Bacon price effect would likely discourage some State-local government construction activity and lead to the reallocation of additional community public sector resources into the planned construction project at the expense of some other "competing" government activity. Finally, of course, the distinct possibility also exists that such a price effect of Davis-Bacon on the cost of labor inputs could contribute to higher State or local taxes that would have been obtained in the absence of the prevailing wage requirement.

In the more popular version of the debate on the "costs" of Federal policies, of course, the initial expenditure effect of State-local is likely of greater concern. Again, this refers simply to the additional public sector outlays which, to some degree, are imposed upon State and local governments as a result of Federal policy initiatives. While this may also be accompanied by a substitution effect,[8] it can result directly from a federally induced State-local government expenditure decision rather than indirectly from a federally induced increase in the price of a factor input. In this case, the price of individual State-local factor inputs need not change. Rather, the affected State-local governments choose to purchase a larger quantity of certain inputs (as in the example of the grantsman hired expressly to seek out additional Federal funds) or a larger quantity of all inputs (as in the case of Federal matching grants which effectively lower the price to the community of a given programmatic outcome, such as a new bridge) without affecting the unit prices which are paid for any individual factor inputs. Thus, though input prices remain unchanged, a larger quantity of input is purchased, thereby expanding overall expenditures.

In exploring the potential cost effects of Federal policies, it is important to remember that either or both of these influences—price effects or expenditure effects—can have a significant impact on the State-local public sector. Additionally, in assessing any study which offers empirical estimates of the costs of Federal policy requirements, it is incremental outlays attributable to specific policies—and not simply aggregate expenditures by program—which are the appropriate concern. Finally, as addressed above in a preliminary way, identification of the true incidence of policy costs and benefits can often prove to be an elusive goal. The following section exemplifies this problem by explaining indirect as well as direct economic effects of Federal policy initiative upon the State-local sector.

E. Direct and Indirect Economic Effects of Federal Actions

1. DIRECT EFFECTS

Especially with regard to the questions of Federal economic effects on State-local governments, it is useful to distinguish between the direct and the indirect effects of Federal action. A direct effect, the type referred to thus far, occurs when the imposition of Federal policy produces an initial response primarily from the State-local public sector itself. Normally, this would happen whenever such "subordinate" levels of government are the principal targets of Federal action, be it in the form of legislation, Executive order, or court mandate. Most Federal policy effects are generally perceived to be of this direct type. Examples include the straightforward proscription of or requirement of certain actions, such as the prohibition of racial discrimination in employment or the establishment and enforcement of Federal air quality standards, each of which has imposed incremental costs upon State and local governments. Other examples include the restrictions incorporated in many Federal grant-in-aid pro-

[8] This could occur either by distorting State-local government budgetary preferences or by substituting public expenditures for otherwise private expenditures via a tax increase to meet the heightened demands on the community public sector.

grams which link funding approval to recipient compliance with requirements perhaps unrelated to the primary functional purposes of the individual programs themselves. Such evident conflicts between efficiency and equity include the intended fair treatment of persons or businesses displaced from property as a result of federally assisted activity, A–95 comprehensive clearinghouse review of local/regional development programs, and conversion of public transportation facilities to meet the needs of handicapped persons. Thus, both with and without the threat of withheld Federal assistance, the direct effect of Federal action may impose incremental compliance costs upon State and local communities.

Most of the attempts to estimate public sector costs of Federal requirements focus on direct effects. Utilizing expenditure analyses,[4] comprehensive cost-benefit assessments of particular programs,[5] or even field interviewing [6] in some combination with these other methods, direct effects may generally be studied by fairly straightforward examination of intergovernmental (perhaps, more accurately, government-to-government) relationships. Here, in many cases, the major stumbling block is that of the nonavailability of sufficient data of suitable detail and accuracy.

2. INDIRECT EFFECTS

The indirect effect of Federal actions generally works through the private rather than the public sector.[7] In this case, Federal policy produces a two-stage reaction. First, the private sector (at which the Federal initiative nominally is aimed) responds to a new requirement (by raising prices to cover the new costs, by curtailing employment growth, by reducing stockholder dividends, or perhaps by relocating to a new plant or even shutting down operations altogether). Second, as a result of such action, the economic circumstances (the tax bases or expenditure requirements) of State and local governments are adversely affected, forcing them in turn to raise tax rates or curtail public services as an indirect effect of new Federal constraints upon the private sector.

Clearly, it would be difficult to estimate the incremental costs which may accrue upon State and local governments as a result of such indirect effects of Federal action. Beyond the complications of the analysis of direct effects, as discussed above, a major methodological prob-

---

[4] The most prominent of these is probably the controversial work of Murray Weidenbaum at the Washington University Center for the Study of American Business. See Weidenbaum, "The Cost of Regulation of Business," prepared for the U.S. Congress, Joint Economic Committee (Apr. 1978).

[5] David O'Neill, "Discrimination Against Handicapped Persons: The Costs, Benefits, and Economic Impact of Implementing Section 504 of the Rehabilitation Act of 1973 Covering Recipients of HEW Financial Assistance," revised version of an impact statement published in the Federal Register (May 17, 1976).

[6] A team of researchers under the direction of Dr. Catherine Lovell of the Graduate School of Administration of the University of California, Riverside, is currently pursuing this approach under a grant from the National Science Foundation.

[7] Although attention to this phenomenon has come largely through consideration of intergovernmental grants-in-aid, it is equally relevant to questions of the impact of Federal regulatory activity (in its broadest sense) and national macroeconomic policy.

An excellent compilation of the indirect impacts of Federal action is contained in the recent work of Steven M. Barro, Roger J. Vaughan, and Mary E. Vogel at the Rand Corp.: "The Urban Impacts of Federal Policies" series (4 volumes), especially Barro, "Fiscal Conditions," vol. 3 (Apr. 1978). An abbreviated version of the project is contained in Barro, "The Urban Impact of Federal Policies: Their Direct and Indirect Effects on the Local Public Sector," in Herrington J. Bryce, editor, "Small Cities in Transition: The Dynamics of Growth and Decline" (Cambridge, Mass.: Ballinger Publishing Co., 1977), pp. 213–243.

lem in evaluating indirect costs involves the necessary assumptions about the economic interactions between the public and the private sectors. Some behavioral model is needed, in qualitative if not quantitative form, which both describes the private sector reaction to specific Federal initiative and then predicts the attendant fiscal impact on the public sector of the affected communities. Good econometric work has been done, of course, in areas related to these concerns, and useful studies could doubtless be commissioned on individual aspects of this overall problem. Nevertheless, we are a long way from any operational capability to assess the impacts of this issue in a reliable, comprehensive, and timely way.

II. Grants-in-Aid

A. *Structural Considerations*

As an introduction to this chapter on the cost-related problems of intergovernmental grants, this section emphasizes several institutional concerns. The first is the perhaps arbitrary distinction between Federal regulations (those which are not necessarily linked to Federal grants) and grant-related requirements placed upon recipients of Federal intergovernmental aid. The second is the contrast between Federal programmatic requirements (usually associated with grants) on the one hand and Federal procedural requirements (often linked with Federal nongrant regulations) on the other. The third is the reiteration of the analytic and operational focus on costs rather than benefits.

1. FEDERAL REQUIREMENTS: GRANT RELATED AND NONGRANT RELATED

With regard to the State-local public sector, there is no clear delineation between what may be called "requirements of Federal grants" on the one hand and the "Federal nongrant regulations" on the other. This unfortunate semantic confusion has arisen in the recent publicity over the effects of Federal regulations because the term "regulations" has been loosely applied to the numerous stipulations which are often attached to individual grant-in-aid instruments. Additionally, of course, the same term is used in reference to the myriad of Federal rules and constraints which also attend many public sector activities which are not necessarily under intergovernmental grant-in-aid support from the Federal Government. In certain cases, as well, similar Federal requirements apply to grant and nongrant activities alike. The separation here of grant-in-aid requirements from nongrant regulations is largely arbitrary, and it is intended simply to make this discussion more manageable. That is, requirements imposed upon the functioning of the grant-in-aid system are examined under the present chapter on grants; requirements which are generally discussed without necessary reference to the grant-in-aid system are treated in the subsequent chapter on regulations.

2. PROGRAMMATIC AND PROCEDURAL REQUIREMENTS

Federal grant-in-aid stipulations contain requirements which may be classified as either programmatic or procedural. Programmatic requirements prescribe standards and conditions of service delivery for

the particular public good in question. They may include Federal concerns for economic (allocational) efficiency and for distributional equity by stipulating the extensiveness, quality, and (minimum) quantity of a given public service receiving Federal funding assistance. In so doing, of course, the goal of least-cost service delivery may well be sacrificed, often at the alleged financial expense of State and local jurisdictions. Procedural requirements exist for grants-in-aid as well as for nongrant-related regulations. Often derived from the national policy objectives discussed earlier, they prescribe how State-local government activities are to be carried out, particularly with regard to planning, reporting, and program management.

It would appear that Federal grant-in-aid requirements contain a strong component of programmatic as well as procedural stipulations, while procedural requirements might seem (at least to a State or local official) to play a larger role in nongrant-related Federal regulations. Nonetheless, Helen Ingram has argued that, at least at the State level, the impact of Federal programmatic requirements has been diluted to the point where the grants process is one of Federal bargaining with the States, rather than of real Federal intervention, to achieve defined programmatic compliance:

Both Federal agencies administering grants and the State agencies designated as the receptors of grant monies want the grant transaction to take place. While Federal agencies would like the transaction to bind State recipients to Federal policy, State agencies face a number of conflicting concerns. Thus the aim of State agencies in the transaction is maximum possible leeway to pursue their own separate goals with Federal money. Because each participant values settlement, neither is anxious to cause conflict if it can be avoided. For instance, Federal agencies are unlikely to embark upon enforcement practices that they anticipate will bring them little support and much criticism. Similarly, when the returns to a State are substantial and certain, and the State interest is small, the State is likely to accede to Federal requirements. The result is a process of implementation that is a complex succession of bids and counterbids between the State and Federal levels during which the initial aims of each are substantially modified.[8]

Beyond this dilution of programmatic objectives through bargaining, Ingram reemphasizes the recognized importance of including legislative "sweeteners" (particularly functional planning grants, though this would be challenged by some recent positions of the National Governors' Association to be noted shortly) in a bill solely as a means of buying acceptability for the entire package.[9] Moreover, this practice includes "a tendency to be vague about objectives and indefinite about the conditions under which grants can be denied to States," Ingram claims, leaving Federal agency administrators "little in terms of legal mandate, legislative history, or public expectations" upon which to base rigorous enforcement standards:

For instance, the Federal Water Pollution Control Act of 1956 provided only one performance requirement: the Secretary (Administrator) must approve a plan that provides for the extension of improvement of the State progam for prevention and control of water pollution. No clearly defined standard against which to measure "improvement" was provided. Similarly, the only major re-

---

[8] Helen Ingram. "Policy Implementation through Bargaining: The Case of Federal Grants-in-Aid," Public Policy, vol. 25, No. 4 (fall 1977), p. 502.
[9] Ibid., pp. 506–507.

straint placed on State educational agencies before receiving title V grants under the Elementary and Secondary Education Act was a need to show in the grant application how the agencies' leadership resources would be "strengthened." What constituted strengthening was nowhere set forth in the bill.[10]

Therefore, to the extent that such vagueness of programmatic objectives or weakness in their enforceability characterizes at least some Federal grant-in-aid efforts, State-local governments may be hampered less by program-related than by procedure-related cost effects. This is appropriately illustrated in a recent study by the National Governors' Association (formerly the National Governors' Conference) with reference to Federal support for policy planning and management assistance:

> At one time, States saw such assistance as a significant resource for rationalizing the policy process and strengthening their capacity to manage their own as well as Federal programs.
>
> It now appears to many Governors that these planning assistance programs, while intended to be part of the solution to managing the growing complexity of the intergovernmental system, have become part of the problem instead.
>
> * * * * * * *
>
> In the final analysis, these programs have contributed substantially to improving the central policy and management capacity of State and local governments. They have, however, become so fragmented and laden with red tape that many Governors question the wisdom of continuing them in their present form.[11]

Such observations as this do not, of course, belie the substantial fiscal and programmatic impacts which the Federal grants system has had on State and local government activity. They do, however, illustrate the frustrations of grant recipients in dealing with the procedural requirements of Federal grants, despite State-local recognition of the budgetarily distortive effects of programmatic requirements.[12] Indeed, even with regard to the more narrow (and predominant) categorical grants, a recent (fall 1975) survey by the Advisory Commission on Intergovernmental Relations and the International City Managers Association (ACIR/ICMA) found that

> The relatively low, "problem" rating given "the narrowness of scope and the number of program categories" is surprising in light of State and local officials' clamor for more block grants.
>
> Matters of fund allocation, performance standards, and central decisionmaking are markedly less bothersome, in the eyes of these local officials [responding chief executives of cities over 10,000 population and of counties over 50,000 population], than problems caused by the volume of paperwork, processing delay, and specific financial management requirements.[13]

---

[10] Ibid., p. 508, with citations of Sheldon M. Edner, "The Implementation of Federal Water Pollution Control Policy Through Grants-in-Aid," (a paper delivered to the 27th annual meeting of the Western Political Science Association, San Diego, Calif., 1973). and Jerome T. Murphy, "Grease the Squeaky Wheel" (Cambridge, Mass.: Harvard Research Center for Educational Policy, 1973).

[11] National Governors' Conference, "Federal Roadblocks to Efficient State Government," vol. 2, "Agenda for Intergovernmental Reform" (Washington, D.C.: National Governors' Conference, Feb. 1977), pp. 16–17.

[12] A substantial body of opinion exists among city and county officials, and particularly among those in larger jurisdictions, that Federal categorical grants have a pervasive, stimulative, and lasting effect on local decisionmaking.

From ACIR, "The Intergovernmental Grants System as Seen by Local, State, and Federal Officials," report A–54 (Washington, D.C.: ACIR, Mar. 1977), pp. 3–4. See also ACIR, "Federal Grants: Their Effects on State-Local Expenditures, Employment Levels, and Wage Rates," report A–61 (Washington, D.C.: ACIR, Feb. 1977), for a review of the literature and empirical analyses of the fiscal effects of Federal grants on State-local government activity.

[13] Ibid., p. 4.

Despite the unique constraints of the programmatic stipulations of Federal grant programs, it is the procedural requirements which evidently raise the most hackles at the State and local level. Officials there object much less these days to the distortive budgetary effects of Federal grants than to their procedural costs. Accordingly, the remainder of this paper will concentrate on the procedural, rather than the programmatic, requirements of Federal policies toward State and local governments.

B. The Focus on Costs

As in the operational evaluation of any Federal policy requirement, the proper analysis of the impact of Federal grant programs should include considerations of benefits as well as costs. Although this point was made previously in chapter I, it bears repeating here, since the purpose of any public sector program is largely the production of benefits against which the costs of that program might then be assessed. Unfortunately, as already noted, our operational (as opposed to theoretical) capacity to identify and evaluate the benefits of a given policy is even less impressive than is our capacity to quantify many of a policy's attendant costs. While the rough specification of such benefits is possible as a result of extensive field research, such an approach is too expensive and too time-consuming ever to be of practical use in the analysis of many of the diverse Federal programs. Indeed, even such large-scale analyses as the Brookings study of the Community Development Block Grant (CDBG) Program can produce only limited results in terms of benefit measurement:

> The research approach of this study involves a quite narrow concept of income-incidence, concentrating on direct and short-term benefits to lower-income persons. There are a number of strictly pragmatic reasons for doing this. Indirect benefits are, to say the least, difficult to gauge. In addition, because this report focuses on the first year of the CDBG program, longer-term benefits are necessarily beyond its scope. We intend to probe further and as fully as we can in this area; but, until we have done so, we do not plan to broaden or revise the concept of income incidence developed at the second field research conference and used in this chapter. In this respect, our approach is similar to that of the Oakland Project, started in 1966 by graduate students and faculty members of University of California, Berkeley. Major differences are that this study of the CDBG program covers a number of jurisdictions and deals with more generalized assessments of income-group impact. The Oakland Project, however, also focuses on "immediate outcomes." According to Frank S. Levy, Arnold J. Meltsner and Aaron Wildavsky, "We look to the most immediate outcomes so as to know which groups get more good outputs and less bad ones * * * we cling to close causation partly because more distant causes are difficult to disentangle.[14]

Accordingly, the benefit side of the question is necessarily given short shrift in any such examination of Federal policy impacts. Despite both the need and the rhetoric, what is produced is more akin to cost-effectiveness or cost-efficiency analysis (aimed to minimizing the costs of delivering a given programmatic outcome) than to cost-benefit analysis (wherein the value of the outcomes themselves is also assessed).

---

[14] Richard Nathan et al., "Block Grants for Community Development," first report on the Brookings Institution Monitoring Study of the Community Development Block Grant Program, prepared under contract H–2323R (Washington, D.C.: U.S. Department of Housing and Urban Development, Jan. 1977), pp. 305–306, with citation from Frank S. Levy, Arnold J. Meltsner, and Aaron Wildavsky, Urban Outcomes (Berkeley, Calif.: University of California Press, 1974), p. 4.

C. State-Local Cost Problems by Type of Grant

Federal grants fall into three basic groups: Categorical grants, block grants, and general revenue sharing.[15] Categorical grants are both the most narrowly defined and the most prominent, reported by the ACIR to consist of 442 separate grants in fiscal year (FY) 1975, amounting to approximately $37.4 billion.[16] Block grants generally go to general purpose governments in accordance with a statutory funding formula for use in a wide variety of activities within a broad functional area. General revenue sharing funds are also distributed to general purpose governments, but with virtually no restrictions on either the purposes or procedures of recipient government expenditure.

Categorical grants themselves may be classified in a number of ways, though that of the ACIR is as useful as any for our purposes: Formula grants, project grants, "formula-project" grants, and open-end reimbursement grants.[17] In formula grants, aid funds are allocated using a precise combination of factors (frequently State-local socioeconomic characteristics of recipient jurisdictions) as specified either by legislation or by administrative regulations. The most straightforward aid instrument in terms of administration, formula grants can generate extensive political maneuvering by various interests in attempts to affect the ultimate character of the legislated formula. With project grants, applications are submitted by potential recipients in accordance with a schedule and manner specified by the Federal Government. These, clearly, are the most troublesome grant vehicles for State and local governments, since the federally prescribed procedures may prove ambiguous (as noted earlier) or difficult for certain jurisdictions to meet (especially those with smaller and less-experienced "grantsmanship" staffs). With the so-called formula-project grant, a two-step process is employed: States first receive allocations based on formula provisions, and then local governments engage in intrastate project-award competitions. Finally, with the open-end reimbursement grant, the Federal Government agrees to reimburse a specified percentage of State-local program costs, thereby avoiding both the problems of project application/competition and the political difficulties of selecting the allocation formula. In practice, even the ACIR has treated open-end reimbursement grants as fomula grants, however, as may be seen in its count of categorical grants by type for fiscal year 1975:[18]

Allotted formula (including open-end reimbursement)	110
Project	296
Formula-project	36
Total	442

From the standpoint of both programmatic and procedural requirements, project grants are surely the most troublesome for State and local officials. Again, in a recent ACIR survey, the complexity and

[15] This grant typology and accompanying descriptions are drawn from ACIR, "Categorical Grants: Their Role and Design," report A-52 (Washington, D.C.: ACIR, 1977), pp. 5-6.
[16] Ibid., p. 92.
[17] Ibid., p. 5.
[18] Ibid., p. 120.

volume of grants paperwork in general and the time required for the application, review, and approval process of project grants in particular were the greatest complaints of local officials.[19] The magnitude of this problem may easily be appreciated, as project grants comprise fully two-thirds of the above total of 442 categoricals.

Beyond this, even, local officials objected to the increasing constrictions being placed on block grants which were passed through the States to localities. State officials were also concerned with the role of Congress and Federal agency personnel in the intergovernmental grants process. This was evidenced by (*a*) a growing uncertainty in year-to-year Federal grant revenue flow, (*b*) a reduced sense of the positive effects of Federal supervision and oversight of grant activities, and (*c*) a general unease with the perceived inflexibility in the Federal agency application of administrative standards. Finally, of course, "considerable local irritation" was noted with regard to the across-the-board application of national policy objectives to grants with their own diverse programmatic goals.

Similar problems were also identified by the National Governors' Conference (NGC) in its call for three basic reforms of the intergovernmental grants system: "decategorization, advance funding, and a clear congressional mandate for executive branch efforts to improve intergovernmental relations." [20] Of the six general problems which "plague intergovernmental programs," as identified in the NGC study, five relate directly to Federal procedural requirements:

1. Lack of coordination among Federal departments or agencies limits the effectiveness of programs in solving problems and increases the administrative burden on the States.
2. The Federal executive branch exceeds its proper authority in some areas, encroaching on matters which are within the proper jurisdiction of the States.
3. Federal regulations are prescriptive in methodology rather than oriented toward results.
4. Excessive reporting and paperwork requirements must be met by States participating in Federal programs.
5. Funding and program implementation are delayed by lengthy approval procedures, absence of program guidelines, and other administrative practices which cause serious dislocation and inequities at the State level.
6. Lack of Federal coordination and consistency in implementing indirect cost determination procedures creates continuing administrative confusion for States.[21]

Thus the view is reinforced that Federal procedural requirements impose an inordinate burden upon State and local jurisdictions. Among intergovernmental grant programs, project grants are surely the worst offenders in this regard, with the unique State and local costs of carrying out grantsmanship activities, coordinating grant application efforts, interpreting sometimes ambiguous Federal program and policy guidelines, marking time awaiting the sometimes delayed Federal funding announcements (a problem which extends beyond project grant activity to much intergovernmental fiscal funding), distorting community implementation activities to conform to Federal

---

[19] Ibid., pp. 279–284.
[20] National Governors' Conference, vol. 2, "Agenda for Intergovernmental Reform," p. 4.
[21] National Governors' Conference, "Federal Roadblocks to Efficient State Government," vol. 1, "A Sampling of the Effects of Red Tape" (Washington, D.C.: NGC, February 1977), pp. vii–viii.

methodological prescriptions, and providing up-dated information to numerous Federal agencies on the operation of federally funded State-local government endeavors. Again, most prominent among the State-local complaints regarding the operation of the grant-in-aid program appears to be this alleged procedural, rather than programmatic, cost. Among the major grant types, the numerically predominant project grants receive by far the harshest criticism in this regard.

D. Some Complications in Producing a Detailed Cost Analysis of Grants-in-Aid

Other than identifying Federal procedural requirements and categorical project grants as generating the greatest objections from State and local officials, what more might be done? Conceptually, one might consider constructing a table of hypothetical State-local government costs, with each type of grant program [22] listed down the left-hand column and with the various categories of State-local sector costs [23] arrayed across the top. Within each cell of this hypothetical matrix, therefore, might then be entered information on the magnitudes of each category of State-local costs for each of the major grant types.

In such a scheme, unfortunately, the relative measure of these cost burdens presents a difficult conceptual problem: to what should any such costs be related? Simply entering supposed cost estimates in dollar terms would provide no meaningful aid in assessing which costs were then too high or too low. Some standard must be proposed against which to evaluate such hypothetical data. If costs were to be specified relative to the overall outlays for each grant, the resulting indicator would surely vary unpredictably from one grant program to another. No standard of comparison could be established by which to judge the alleged degree of cost inefficiency across the many grant programs. Alternatively, if the cost burden were defined relative to some necessary minimum for each individual grant program, how could such a minimum reasonably be established? In the extreme case here, one might speculate about simply eliminating a given grant stipulation, such as some national policy objective which is functionally unrelated to the grant in question. In such an instance, the minimum necessary cost of that stipulation would, of course, be zero, with any State-local costs above it being deemed excessive. More moderately, by what benchmark could such minimum necessary State-local costs be established if even a nominal pursuit of said national policy objective were decided upon? In such quasi-cost-effectiveness situation (albeit ignoring the values of benefits associated with programs whose scope is permitted to vary), reasonable answers are not easily forthcoming. Except under a heavily constrained set of grant types (i.e., looking only at a few very similar grant programs) or of costs (i.e., specifically ignoring all indirect costs and expenditure effects due to nearly complete lack of applicable studies of these issues),

---

[22] Including, as a starter, categorical grants (formula, project, and formula-project), block grants, and general revenue sharing, with detailed breakdowns of each of these as appropriate.

[23] Programmatic costs (including price and expenditure effects, direct and indirect effects) and procedural costs (including grant-specific costs and costs of pursuing national policy objectives through the grant-in-aid mechanism).

relative comparisons of such results for policy purposes would be virtually meaningless.

Beyond this is the question of the differentiations among the grant programs themselves. With several hundred individual formula and project grants, surely it is not enough merely to identify the three or four basic types of categorical grants. At minimum, it would be necessary to distinguish among the major grant design features—varying State-local matching ratios, maintenance-of-effort requirements, for example—which are recognized for their important fiscal effects on grant impact. Such an exercise would quickly balloon into a substantial analytic enterprise, one whose costs in terms of time and financial resources might possibly outweigh its initial policy purposes. Without such a detailed analysis, inquiry into cost burdens can be misleading, as indicated in the conclusion of the excellent chapter on "Cost-Sharing Arrangements: Their Significance and Impact," from the recent ACIR study of categorical grants:

Certain general conclusions can be drawn.

> First, Federal grants are stimulative [of State-local government fiscal effort] to some degree. Particular statistics, however, are difficult to interpret because the degree to which matching requirements, as compared to more political factors, are responsible is usually unclear. In addition, consideration of the cause of apparent stimulation is important. No-match grants can be expected to elicit the most widespread participation by State and local governments, while matching grants may elicit greater non-Federal fiscal commitments.
>
> Second, situational factors appear to be crucial. These include the size and timing of the grant offerings relative to State and local activity and whether or not maintenance-of-effort requirements are in existence and effective.
>
> Third, it cannot be assumed that two similarly designed grants for two different functional areas will be equally effective, in part because political preferences that influence grant responses can be expected to vary for different types of public spending. A corollary is that the same grant for the same activity to two different recipients may elicit entirely different responses, depending on local economic conditions, political factors, and servicing tastes.
>
> Finally, the total impact of a grant cannot be fairly assessed without considering the ways in which effort-related requirements, such as matching and maintenance-of-effort provisions, interact with allocational formula characteristics.[24]

E. Recommendations

The magnitude of any comprehensive grants-cost analysis is suggested by the panoply of Federal grant programs and the variability in the characteristics and responses of State and local communities to each of these separate endeavors. As a practical matter, a comparative cost study of each of these programs, if done properly, is too massive an undertaking to be recommended as an effective policymaking tool. Short of that, however, the information presented here suggests several feasible approaches to the problem of costs of Federal grants at the State-local level: First, the procedural costs and practices of selected grant programs might be reviewed, particularly those of project grants which appear to provoke the greatest objections from State and local officials. In so doing, differentiation should be made—to the extent possible—between grant-specific procedural costs and procedural costs associated with the pursuit of overlaying national policy objectives. This is advisable because, in the potential

[24] ACIR, "Categorical Grants: Their Role and Design," pp. 182–183.

revision of a grant instrument, the former may be much easier to modify than the latter. Again, selectivity is important, for useful policy comparisons can most easily be made among the cost-impact results of similar grant programs.

Moreover, the objective here is as much to identify and eliminate wasteful and ineffective practices as it is to conduct comparative economic analyses. In many cases, the real problem seems to be the presence of cumbersome administration rather than the absence of an efficient cost-estimating technique. Unnecessary costs can still be reduced even if their true magnitudes cannot be accurately estimated.

Second, over the longer term, consideration should be given to the programmatic and political feasibility of shifting some grant funding from project grants to other aid instruments. This follows Richard Nathan's conclusion that project grants (generally the most stimulative of State-local effort, especially if accompanied by matching requirements) become less so over time as functions become established within recipient communities.[25] However, given the diverse effects of different grant aid characteristics in the several functional areas, decreased reliance on project aid ("decategorization") should be limited in nature.

Third, tentative attempts might be made to begin the practice of estimating intergovernmental cost aspects of certain bills as they pass through the legislative process. (This will also be discussed in the following chapter.) If undertaken at all, however, this endeavor should be carried out only on a most carefully defined trial basis and, at that, only for a very narrowly prescribed set of legislative proposals. If actually implemented, the experimental nature of this effort must clearly be recognized, for the results could well discourage further such attempts. On the one hand, this work could become a bottomless pit of analytic frustrations amid growing demands for legislative cost estimation, thereby defeating the timeliness of the work for policy purposes. On the other hand, the more encompassing the evaluative mandate of this endeavor, the more likely that careful analysis would necessarily give way to quick rough estimates whose potential inaccuracy could ill-serve the goal of objective policy evaluation.

Further cost-effectiveness aspects of these and similar concerns will be addressed in the following chapter on regulations and in the closing chapter with the conclusions and recommendations.

III. Regulations

A. Two Kinds of Procedural Costs

The preceding chapter concluded that procedural costs of intergovernmental grants were perceived to be a much greater burden to State and local governments than were programmatic costs of grants. If this observation is true for Federal grants, it is surely true for Federal nongrant requirements, the regulations with which this chapter is con-

[25] From U.S. Senate. Subcommittee on Revenue, Committee on Finance, hearings on general revenue sharing (Washington, D.C.: U.S. Government Printing Office, April and May 1975), p. 160, cited in ACIR, "Categorical Grants: Their Role and Design," p. 179.

cerned. Accordingly, the primary interest here will be an inquiry into the two important policy components of the procedural costs of regulation: program-specific costs and costs arising from the pursuit of overlaid national policy objectives.

Program-specific procedural costs essentially constitute those State-local government costs which are unique to a given program of public service activity but which are not directly related to the delivery of that service. For example, numerous statistical reports are required by the U.S. Department of Labor (DOL) from state employment security agencies on labor force employment trends, wage rate patterns and unemployment statistics (i.e., labor employment service program delivery areas of both Federal and State Governments).[26] Many of these reports are clearly related to labor employment service program efforts while having little to do with national policy objectives of nondiscrimination or fair labor standards. That is, their role is labor-program specific, including, of course, the simple collection of employment data deemed useful by the DOL but not oftentimes, to the State agency from which the data were derived. In contrast, a few of these reports likely would not be required were it not for the existence of national policy objectives which are imposed upon the labor employment public service activity: "Placement and referral of individuals (minority)" and "Applicants provided counseling, testing and related services (Vietnam-era veteran)" are but two illustrations. In this respect, both types of reports, program specific and national policy related, are procedural rather than programmatic in nature: They merely report information to the Federal Government and play no direct role in the actual provision of the services which they describe.

The simple separation of these types of procedural costs suggests a structural approach to reducing the State-local burden of Federal policy requirements. In individual programs, for example, it might be feasible to reduce or to consolidate the reporting requirements on program-specific activities as an administrative decision, while a more difficult to achieve legislative decision might well be required to modify a reporting requirement related to a national policy objective (nondiscrimination in employment training, for instance). Therefore, measurability of these two types of procedural costs becomes a useful economic consideration.

B. More Problems in the Measurement of Costs

Despite the analytic importance of being able to distinguish between these program-specific and national-policy-related procedural costs, the practical problem of doing so is formidable. The vagaries of such an endeavor are illustrated by a discussion on the identification of core versus noncore costs for a recent State-local paperwork impact study in California:

> Briefly, core costs were those necessary to the actual delivery of a government service to its recipient, and noncore costs were everything else. The purpose of the distinction was to permit isolation of costs that were not part of the service production function. Classifying a cost as a noncore cost did not mean that it was socially undesirable. Records kept to assure the protection of

[26] Over 100 such reports from the Ohio Bureau of Employment Services are listed in National Governors' Conference, vol. 1, "A Sampling of the Effects of Red Tape," pp. 31–36.

the 14th Amendment rights, for example, do not provide a service to school children, but through its lawmakers the country has decided it is useful to keep such records. What is or is not a core cost is subject to judgmental determination. Though the concept may be readily accepted, substantial negotiation among public administrators, officials, and members of the public will be required to develop core cost standards. It is a critical concept in determing the "burden" of Federal paperwork.[27]

Evidently, merely such a fundamental disaggregation of costs—into what approximates our programmatic and procedural categories—is much more subjective and less scientific than might initially be supposed. Even with the indicated element of "judgmental determination," a close examination of necessarily high-quality records would be required to produce an accurate result for a given existing program:

In searching for such standards, there are minimal or threshold administrative costs that must be incurred simply to operate any program. * * * Sound administrative practice requires that sufficient records be kept to assure that services are delivered. *In particular, such records should provide fiscal, personnel, and production accountability to verify funds, employee time, and actual service delivery.* [Italics added.][28]

Thus any attempt to analyze the various cost components of an existing, operating program simply at this micro-level is fraught with difficulty. Indeed, even with heavy investments of time and money on such a cost-analysis enterprise, it would seem likely that the eventual results would be seriously limited by the quality of much of the original accounting data and by the "judgmental" factor. These are vital problems to a legitimate "economic" analysis which would be most difficult to overcome, if the California paperwork study is any guide:

Reliance on budgeting or accounting data alone, therefore, will inhibit useful measure of paperwork in State government. Consequently, a "judgmental" approach is required. A line-by-line analysis was made of the California budget to determine what percentage was likely to be paperwork.

California's budget was selected for several reasons. A large State, California accounts for a significant portion of all State spending and is active in a wider variety of activities than many States. Furthermore, California separates administrative costs as well as individual program accounts within departments; consequently, costs within programs can be isolated.

* * * * * * *

As noted above, any sum of core and noncore cost at this stage of conceptual development will depend partially on the presumptions made by individual analysts.[29]

Furthermore, with a problem whose analytic applications are so open to question, there is room for bias in almost every estimate. Particularly in so political an issue as government regulation, advocacy analysis may be expected to make its way into nearly all of the quantitative studies of the situation. As Julius Allen puts it,

What has disturbed several observers of some of these cost estimates, particularly projected cost estimates for many individual regulations or regulatory programs, is that they tend to support the vested interest of the sponsor of the estimate or to fit the hypothesis of the individual making the estimate.[30]

---

[27] Academy for Contemporary Problems, "Impact of Federal Paperwork on State and Local Governments: An Assessment by the Academy for Contemporary Problems," a report to the U.S. Commission on Federal Paperwork (Columbus, Ohio: the Academy, July 1977), p. 33.

[28] Ibid., pp. 33–34.

[29] Ibid., pp. 35–36.

[30] Julius W. Allen, "Estimating the Costs of Federal Regulation: Review of Problems and Accomplishments to Date," report No. 78-205E (Washington, D.C.: Congressional Research Service, U.S. Library of Congress, Sept. 26, 1978), p. 22.

In short, having virtually abandoned hope for practically useful cost-benefit analyses of Federal policy requirements, we must now look dubiously at many estimates of costs. Despite the economic arguments for the desirability of distinguishing between procedural and programmatic costs, and, within the former, between program-specific costs and those attributable to national policy objectives, the preceding evidence—based on the pertinent studies to date—should not be encouraging in this regard.

Before turning to some suggestions which may offer some hope, not for the measurement of procedural costs but for the improvement of procedural operation, we must look at a final cost-related concern—the proper allocation of regulatory costs between Federal and State-local governments.

C. Problems in the Attribution of Costs to the Federal Government

In addition to the difficulties of reliably measuring the State-local costs of regulatory requirements, we must inquire about the degree to which those costs would be attributed to Federal, as opposed to State, Government policies. Many State laws exist in the areas of access to Government information and decisionmaking processes, worker health and safety, fair labor standards and wages, and environmental protection, whose requirements are at least as stringent as similar Federal legislation. To the extent that such State standards exceed those of the Federal Government, no such programmatic costs can be attributable to extant Federal policy. Indeed, given the irregularity of State statutes in many of these areas, even the procedural reporting requirements of the Federal Government might well be justifiable on national policy grounds (though these procedures themselves might nonetheless be technically inefficient).

Important examples of this abound. They include a plethora of State and local laws on openness in Government decisionmaking.[31] Even more topical, however, are the interactions between Federal and prevailing State wage laws in the 41 States having statutes similar to the Davis-Bacon Act aimed at maintaining the prevailing wage rate on Government-supported construction projects. As reported by the ACIR, "higher wage rates usually have been established by 31 of those States, resulting in those State laws governing rather than the Federal law." [32] Additionally, with reference to pollution abatement costs, Allen notes the cost attribution problem where State or local standards, though generally lower than Federal requirements, nevertheless impose their own burdens:

> The estimates derived from the Council of Environmental Quality, admittedly rough and subject of some controversy, are probably as good as can be found. One aspect of these estimates that may warrant examination is the proper allocation of pollution abatement costs where both Federal and State and local regulations coexist. Where Federal standards are higher than State or local, it may be difficult to determine the costs needed to meet State and local standards, and the incremental costs needed to meet Federal requirements. Where State or local standards exceed Federal standards, it could be argued either that no cost

[31] See ACIR, "Categorical Grants: Their Role and Design," p. 260.
[32] Ibid., pp. 256–257.

should be allocated for meeting the Federal requirements, or that the cost be divided between that needed to meet Federal regulations and the additional amounts needed to comply with State or local standards.[33]

Finally, perhaps the ultimate complication is suggested by those situations wherein a State or local statute would have to be enacted were it not for already-existing Federal requirements:

In many social service delivery programs—employment services, for example—the Federal Government mandates many of the information practices used in program administration. *But in the absence of such mandates the States would be required to adopt comparable provisions* to gather the names, addresses, employment history, and related applicant information simply in order to administer the program. Such paperwork is socially dysfunctional only to the extent that a State is willing to argue that the program is undesirable and is being operated within the State solely as a result of the availability of Federal funds which virtually mandate the program. Federal paperwork of this kind is too inextricably intertwined with the function, however, that it is not a likely source of reduction. Such information practices represent core costs. [Italics added.] [34]

In this circumstance, the problem is especially complex, for it would ask the analyst to pose the formidable "what-if" question of State-local actions in the hypothetical absence of Federal initiative. Here, as in each of these cases, the problem of allocating the cost burdens of regulation to the appropriate level of government is most difficult, involving as it must an evaluation of the laws of both Federal and State (and sometimes local) Governments as well as a sophisticated simulation of State-local action in an intergovernmental world which differs from the one we know. As a practical policy instrument, such analyses, indeed, have little to offer.

D. Improving the Cost-Effectiveness of Procedural Requirements

Thus the prospect does not seem bright for the acquisition of a reasonably reliable, economically feasible technique by which to measure the costs of most Federal regulations at the State-local government level. Aggregate analyses consistently appear to be too crude and inaccurate on the one hand and too philosophically biased on the other. In any case, the gross nature of aggregate results offers little to the policymaker whose decisions most frequently occur at the level of individual programs. Likewise, micro-level studies appear to be at once too costly and too time-consuming to be of regular use in the policy process. Certainly, the detailed analyses of the effects of individual programs can provide punctual and economical policy advice only on a few selected issues at a time. As evaluative tools, then, such careful inquiries can have only limited effectiveness in the evaluation of existing or proposed Federal legislation.

With empirical efforts limited at best to only a few cost analyses of carefully selected existing or proposed statutes, a more qualitative approach might prove fruitful. Several approaches may be examined which might improve the procedural cost-effectiveness of Federal regulations, especially those relating to national policy objectives. These are—

---

[33] Allen, pp. 36–37.
[34] Academy for Contemporary Problems, "Impact of Federal Paperwork on State and Local Governments," p. 35.

1. Intrapolicy priority setting, coupled with demonstration experiments;
2. Federal reimbursement of the policy compliance costs of State and local governments; and
3. Structural modification of certain Federal policy requirements.

1. INTRAPOLICY PRIORITY SETTING, COMBINED WITH DEMONSTRATION EXPERIMENTS

This is probably the most generally applicable of the three potential approaches listed above. Essentially, it implies a qualitative review of the numerous requirements imposed by a given policy with an eye toward those areas (geographic, programmatic, technical, or whatever) where diligent pursuit of Federal policy seems likely to produce the greatest success. Conversely, in other areas where achievement of the stated goal appears much less feasible (or cost effective), the current procedural requirements might be differentially reduced or demonstration experiments might be conducted to ascertain if the desired policy goals might more easily be reached through application of a different policy instrument.

A good example of this situation occurs in the pursuit of environmental quality. Along with the goal of nondiscrimination/civil rights, Federal initiatives in this field might reasonably be described as Federal "mandates" upon State and local governments.[35] Viewed as mandates, the national policy objectives of environmental quality require adherence to programmatic standards in the face of legal, administrative, or even fiscal penalties. Unfortunately, the implicit costs of achieving these mandated standards can themselves become prohibitive, as Larry Ruff notes regarding the soaring expenditure estimates for dealing with municipal stormwater treatment and control:

> When it appeared that a few dozen billions of Federal dollars would essentially eliminate municipal pollution, it was perhaps reasonable to be cavalier about the planning process, to avoid facing up to the need for institutional change in the local-State-Federal system. . . . In short, when it looked like it would be cheap to eliminate municipal pollution within a few years, there was no need to worry much about how to manage municipal pollution in the long run.
>
> But when the estimate of the cost becomes hundreds of billions of dollars, the situation is qualitatively different. Now it is clear that municipal pollution is here to stay, and the real questions concern how to use limited resources efficiently, not how to force the administration to release a few billion dollars more or to streamline the grantmaking machinery with EPA.[36]

Ruff's observation of the need for more efficient use of limited resources suggests the need for priorities in the quest of attainable

---

[35] Mandates are usually thought of as State objectives imposed upon local governments, as defined by the recent ACIR publication on the subject: "Mandates are defined as a legal requirement—constitutional provision, statutory provision, or administrative regulation—that a local government undertake a specified activity or provide a service meeting minimum state standards." See ACIR, "State Mandating of Local Expenditures," report A-67 (Washington, D.C.: ACIR, July 1978), p. 2.

A similar definition might well be applied at the Federal level. As an imposition of Federal standards, environmental protection should be so viewed in light of numerous Federal statutes (particularly the Water Pollution Control Amendments of 1972 and the Safe Drinking Water Act). Nondiscrimination/civil rights is more broadly based both in Supreme Court decisions (school desegregation and rights-of-the-accused cases, for example) and in Federal legislation (the Equal Employment Opportunity Act and the Age Discrimination Act, among others).

[36] Larry E. Ruff, "Federal Environmental Regulation," U.S. Senate, Committee on Governmental Affairs, study on Federal regulation, appendix to vol. VI, "Framework for Regulation," p. 305.

objectives of some mandated policies. That is, some prescribed performance standards are more attainable than others, given present techniques and institutional constraints. Such currently feasible objectives should therefore enjoy priority attention among the current regulatory approaches. In addition, consideration should be given to conducting demonstration experiments of new approaches in those areas where present techniques are less effective or where an especially difficult problem might respond to a new remedy. Ruff's proposed market-based system of Federal effluent fees charged to local governments and marketable "discharge rights" to municipalities[37] would surely become more complex in its practical implementation than market-system advocates are willing to admit.[38] Nonetheless, demonstration experiments in the regulation of some State-local government activity in this area might prove as interesting as have the recent experiments in the (de)regulation of the commercial airline industry.

Such demonstration experiments would be most useful in policy areas like environmental quality where priority setting is most applicable. In this instance, effluent charges might effectively control pollution from identifiable sources such as local industries but not from nonpoint sources such as farm feedlots (a pollution source which may be uncontrollable by any method).[39]

Priorities, then, would both respond to the feasibility of alternative approaches while influencing experimentation with new techniques. Thus, the possibility of implementing a programmatically distinct alternative to the administrative standards regulatory techniques might also offer a way of reducing program-specific procedural costs while not diminishing the vigor with which national environmental objectives are pursued.

In contrast, national policy mandates regarding nondiscrimination and civil rights do not easily lend themselves to the concept of priority setting for the attainment of feasible objectives. These mandates pose quite a different problem than that of environmental protection since alternative economic devices frequently cannot be as effectively applied to the issue of discrimination as they might be to that of environmental pollution. Programmatically, of course, there can be no differentiation among possible sources of discrimination, for the idea of "prioritizing" the pursuit of civil rights clearly conflicts with the basic equity goals upon which these Federal laws and court actions are based.

Even from a purely cost-effectiveness perspective, there is little promise of reducing State-local government procedural costs of compliance with Federal nondiscrimination requirements. Improved Federal interagency coordination of policy guidelines and reporting regulations could go a long way toward imposed procedural efficiency. Unfortunately, of course, nondiscrimination provisions are

---

[37] Under such a scheme, municipalities would pay a prevailing market price for the "right to pollute," with EPA options to "buy back any unneeded discharge rights in any year at market price." Ibid., p. 339.

[38] The setting of effluent charges and the ease and equity with which they could subsequently be adjusted (administered?) are difficult political and legal as well as economic concerns. For review and comment on such proposals, see Robert E. Firestine, "The Economics of Environmental Control: Practical Proposals and Philosophical Considerations," proceedings of the 1971 Du Pont environmental engineering seminar: selected papers, edited by Thomas deS. Furman, bulletin series 137, vol. XXVI, No. 1 (Gainesville, Fla.: engineering and industrial experiment station, 1972), pp. 9–16.

[39] Ruff, p. 340.

probably the most widely applied of any set of national policy objectives, and previous attempts to improve their coordination across major Federal agencies have generally not met with notable success.[40]

For environmental protection, therefore, priority setting for the achievement of feasible obectives may offer some promise, along with coordination demonstration experiments of alternative approaches to the problem. No such hope appears to exist for the much more pervasive Federal nondiscrimination/civil rights mandates.

2. FEDERAL REIMBURSEMENT OF THE POLICY COMPLIANCE COSTS OF STATE AND LOCAL GOVERNMENTS

The emergence of State mandating of some local government activities suggests yet another tool for dealing with the problem of federally imposed costs: full or partial reimbursement of those State-local costs by the Federal Government. This fiscal device is applied particularly in California, and also in the Western States of Montana, Oregon, Colorado, and Wyoming, to compensate local governments for new local activities undertaken as a result of State mandate.[41]

The details of these individually unique programs need not be explored here, although a few aspects of the general process may be reviewed with an eye toward their selective adoption at the Federal level. Probably the most interesting of these is the fiscal note—essentially a rough estimate of the cost of proposed State legislation to a State's local governments. Although carried out by rather small staffs often engaged in other duties as well, such estimates are reportedly helpful to those State legislatures which employ them, for the analysis of State legislative impacts is often not a simple task. Nevertheless, in contemplating a similar endeavor, the Federal legislative analyst faces 50 State-local systems rather than 1, with all the intergovernmental complications which this diversity implies. In particular, the relative depth of the Federal system—especially including the grant-in-aid programs—takes any prospective Federal fiscal note process into a much deeper analytic swamp than any to be encountered in a State capital. As in California, a Federal effort would require both a legislative cost-estimating agency (quite possibly to be housed within the Congressional Budget Office) and an agency for the adjudication of appeals from State and local governments regarding the size and appropriateness of their Federal fiscal reimbursements. The more extensive the scope of any Federal mandate-reimbursement policy, the larger and more cumbersome would be these technical support activities and the more difficult would be any coordinative tasks which might become necessary.

This does not mean, however, that the concept of cost reimbursement is not practical at the Federal level as an addition to the present extensive grant-in-aid efforts. It does suggest, however, that if ever such a plan were to be implemented, it should be done on a selective basis. In so doing, it might initially be directed only at a few Federal policy "mandates" where the cost-estimation and appeal-adjudication

---

[40] ACIR, "Categorical Grants: Their Role and Design," pp. 243–247.
[41] See ACIR, "State Mandating of Local Expenditures," report A–67 (Washington, D.C.: ACIR, July 1978).

burdens could be narrowly circumscribed. Even more important, activities covered by such an effort must demonstrably incur additional State-local government costs purely as a result of a Federal policy requirement not otherwise supported by a grant-in-aid.

One candidate for such reimbursement is the Uniform Relocation Assistance and Relocation Policies Act of 1970, which provides for fair treatment of those displaced from property as a result of direct Federal action or the implementation of federally assisted programs. As described by the ACIR, financing for this federally imposed equity objective may be partially covered (in an irregular way) through the Federal grants process. Moreover, the true cost of relocation can also become embedded in State or local budgetary distortions:

> The act originally allowed for additional funds to meet 100 percent of the relocation costs, but this provision was temporary. Relocation costs are not simply eligible items under the regular cost-sharing formulas of the grants that are subject to this act. When relocation costs add significantly to total project expenses, some projects have been redesigned, shifted to other sources of funding, or canceled. Of course different matching ratios and the existence of no matching requirements in a few programs result in relocation benefit costs being higher for grantees in some programs than in others.[43]

As a reimbursement, Federal support for such relocation activities might at least be more regularly distributed than under the grants mechanism. Regularized reimbursement might also involve lower State-local procedural costs than the project-grants approach described above. Finally, of course, the use of fiscal notes in a selective and experimental reimbursement process might tend to curtail "mandating" and even to cut costs at the Federal level just as its advocates believe it has at the State level.

3. STRUCTURAL MODIFICATION OF CERTAIN FEDERAL POLICY REQUIREMENTS

One way to improve perhaps both the equity and the efficiency effects of Federal policy requirements is to tailor some of the structural elements of regulations to the operating situation facing State and local governments. That is, a regulation may be well designed in principle but flawed in its technical specifics. In such an instance, the modification of a relatively minor component of the regulation might reduce the State-local cost impact without harming the functional effect of the regulation itself. Contrarily, of course, too great a change in such a parameter could completely nullify the original purpose of the legislation.

One example of this may be seen in the Davis-Bacon Act, which requires that the local prevailing wage be paid on construction projects supported by Federal funds. The equity objective of this law is to ensure that community wage standards are not undercut in the expenditure of Federal funds. Of course, this tends to increase the overall cost of public sector construction activities, especially in those areas where nonunion labor is generally available at less than the prevailing wage rate (usually union scale). Some of the community public sector costs of this provision could potentially be lowered by reducing the

[43] ACIR, "Categorical Grants: Their Role and Design," p. 254.

geographic size of the individual regions wherein prevailing wage rates are defined.[43] If done appropriately, such a revision of the Department of Labor's (DOL) area wage rate data might well reduce the Davis-Bacon construction cost burden upon local governments. The corollary to this, unfortunately, is that the cost to the Federal Government of conducting these area wage surveys would then increase to some degree, since there would presumably be a larger number of wage areas to be surveyed. To the extent that this estimating is done in conjunction with related DOL activities, coordination would suffer and overall Federal costs would likely increase.

E. Regulations Affecting the Conduct of Federal Monetary and Fiscal Policies

Finally, a few comments should be offered concerning the many regulations governing Federal monetary and fiscal policies and their potential link to State-local government costs. These activities are reviewed briefly here because they are fundamentally different in their nature and intent than are the grant requirements and regulations already examined. Moreover, much of the current interest in the cost of Federal regulation is in these latter requirements and not in the constraining effects of monetary and fiscal policies.

In general, it must be recognized that there are considerable programmatic effects of U.S. monetary and fiscal policies upon both the cost and the operation of State and local governments. Monetary policy, of course, includes a panoply of national money and credit activities which can have major stimulative or restrictive effects on interest rates and the availability of funds for industrial, commercial, residential, and personal loans. Within these are the important distributional impacts of regulation of the secondary mortgage market. Debt management is also included here as an important element in the money-and-credit picture, although the basic size of the debt is, of course, the result of fiscal decisions interacting with a changing national economy. Like the effects of monetary policy, the taxation and expenditure outcomes of Federal fiscal policy are also conditioned by the many stipulations and regulations that are integral to the functioning of these policies.

With both monetary and fiscal activity, but especially the latter, many of these regulations are explicitly legislative in nature. That is, rather precise regulatory conditions are specified in the enabling legislation, in contrast to many of the previously discussed regulatory provisions which derived from administrative interpretation of less-detailed statutes. Numerous examples may be offered which have important economic effects on State and local governments. The Federal tax exemption of interest earned on State or local government bonds clearly reduces borrowing costs for localities while arguably increasing the availability of credit to them. The deductibility of contributions to nonprofit charitable organizations in all likelihood enhances the viability of important private institutions, some of whose efforts

---

[43] The only current alternative to the "area" method of determining the prevailing wage is the "project" method. This avoids the geographic problems mentioned above, but it requires that the grantee conduct a local wage survey in accordance with Department of Labor guidelines. This, in turn, can introduce additional complications, costs, and delays.

would otherwise have to be supported directly by community public funds. More indirect effects include the availability of various credit and loan guarantees for residential borrowing and for business investment.

With these and many similar examples, the essential regulations are generally not subject to significant administrative interpretation or revision because they are specific programs or statutes enacted through the Federal legislative process. In such instances, in fact, any perceived State-local government cost impacts would be attributable not to regulations—however that term may be defined—but legislative action. This is a major distinction in light of the preceding emphasis on the administrative/procedural objections to regulation on the part of State and local governments.

Perhaps even more telling is the point that many of these monetary and fiscal stipulations were not enacted as regulatory activity at all, but rather to serve primary programmatic ends which may be only moderately concerned with State and local government. That is, while tight monetary policy undoubtedly can have severe "cost" effects on State and local governments, that outcome may be of only moderate concern to a Federal Reserve System whose macroeconomic problems far transcend those of the State-local public sector.

Accordingly, it is of little use to seek modifications in most of the Nation's monetary and fiscal policies as a means of alleviating their cost impacts upon State and local governments. More promising approaches can be made to the procedural costs and across-the-board policy requirements of the Federal Government which have been the major focus of this paper.

IV. Conclusions and Recommendations

A. *Conceptual Issues and Problems*

This paper has been a conceptual examination of the question of the impact of Federal policies on State and local government costs. Adopting a general approach to the topic, this examination has reviewed some of the analytic problems posed by this issue and suggested a few policy responses. The magnitude of the over-all question precluded any detailed exploration of particular programs and policies. Moreover, given the relative paucity of pertinent research in this area, an essay focused on recurring problems and issues seemed more likely to assist policymakers than a narrowly analytic commentary restricted to a few specific economic studies in a broad and diversified field.

Despite the equal importance of the benefits and costs of government activity, we know little about the benefits generated by public sector endeavors, and we are not likely to correct this deficiency in the future. Especially in the complicated area of intergovernmental fiscal and regulatory relations, relative ignorance of the benefit flows may be expected to remain a major impediment to the rational analysis of the impact of Federal policies on State and local governments. Beyond this, the public's perception of benefits from government activities is much weaker than its awareness of public sector tax costs. This

disparity tends to enhance political demands for greater economy in government while potentially underestimating the full contributions of some government activities to the national well-being.

With regard to costs, a number of points were made emphasizing the need for reliable measurement of the incremental or marginal costs attributable to Federal policies. First, Federal policies may produce either a price effect or an expenditure effect (or both), where the former refers to higher input prices facing the State-local sector and the latter to the more popularly perceived expansions of public sector outlays in response to increased Federal requirements.

Second, there are important direct and indirect cost effects of Federal policies. Direct effects occur when imposition of a Federal policy produces an initial response primarily from the State-local public sector itself. These are the most widely recognized responses to Federal policy requirements, but they are only part of the story. Sometimes of equal, but less apparent importance are indirect cost effects—those in which Federal action results first in a private sector response which itself then triggers higher public sector costs. This two-stage scenario frequently occurs when the private sector reaction to a new Federal policy requirement is to modify business investment or employment patterns, producing a public sector fiscal impact on the State or local tax base. Through the complexities of intergovernmental fiscal relations, direct effects are usually much easier to identify than they are to measure with any reliability. In the case of indirect effects, however, mere identification can be a challenge, with even rough estimation of their economic magnitudes perhaps becoming a truly Herculean task.

Recognizing the great practical difficulty of estimating the benefits of Federal policy requirements, the paper has concentrated on the costs of those policies to State and local governments. Within both grant-related and non-grant-related Federal policy stipulations, an important distinction was proposed between programmatic and procedural requirements. Programmatic requirements prescribe standards and conditions of service delivery for the particular public good in question, while procedural requirements indicate how State-local activities are to be conducted with regard to planning, reporting and managing of programs being supported by Federal funds or linked to Federal policy objectives. Concerns for equity are of especial importance in the pursuit of those national policy objectives. Such nonprogrammatic across-the-board requirements often conflict with the efficiency objective, defined either as least-cost technical efficiency or as the allocational efficiency which seeks to minimize the distortion of State and local budgetary priorities which can often accompany Federal grants-in-aid.

At present, the costs of procedural requirements appear to outweigh heavily those of programmatic requirements from the viewpoint of most State and local officials. That is, it is procedural costs of Federal requirements which generate the greater objections at the State-local level. Despite well-known arguments about the distributive impacts of Federal grants upon State and local economies, the current furor over the costs of Federal regulation is clearly directed at procedural

costs. Among intergovernmental grant programs, project grants bear the brunt of the criticism. The complications and costly delays of their many unique application-and-review procedures have come in for particular criticism of late, though such complaints are more widely applicable to the broader grant-in-aid system in general and to numerous nongrant Federal requirements as well.

Within procedural costs may be denoted some Federal requirements which are unique to individual program activities, whether grant supported or not. These have been called program-specific procedural requirements in this paper. In addition to their cost impacts on State and local governments, there may also be identified costs which arise from the pursuit of so-called national policy objectives. These Federal policy requirements are often applied to a wide variety of State-local government activities, and their attendant procedural regulations are a major source of State and local irritation with regard to federally imposed procedural costs of government.

As emphasized throughout the paper, the reliable estimation of the State-local cost burdens of Federal policy requirements is not an easy task. This is especially true in reference to the delineation of procedural costs into their two components. Program-specific procedural costs are, of course, often uniquely linked to the particular activity from which they arise. As such, generalizations about their measurement and control are not easily made, although approaches to individual program-specific procedural costs could be effective. Moreover, costs deriving from the general imposition of national policy objectives would be difficult to reduce on a broad general scale. This is true precisely because such economizing might unintentionally turn into an overall rollback of the national equity (and, to an extent, allocational efficiency) goals upon which many of these requirements are founded. That is, once the enforcement provisions (and costs) of such procedural requirements are reduced or eliminated, the actual implementation of national policy objectives in such areas as nondiscrimination or environmental protection can also be lost with the general reduction in administrative burden. In so comprehensive a policy arena, the potential loss of benefits from sweeping policy reforms must be considered along with the alleged reduction in costs.

B. Recommendations

Our review of these questions has led us to a cautious, rather than bold, perspective on the likelihood that the State-local cost burdens of Federal policy requirements can be substantially reduced. This suggests that comprehensive new attempts to cut these costs dramatically will likely run afoul of the resulting administrative confusion or perhaps will even manage to undermine much broader Federal objectives than ever intended by the reformers. That is not to imply, however, that no economies can be obtained through implementation of alternative policies. To the contrary, several recommendations are offered which seek prudent but deliberate changes in the very procedures by which the Federal Government pursues its own policy requirements.

From this conceptual evaluation, the following activities are recommended for further consideration or action:

1. Procedural costs and federally imposed practices of selected grant programs might be reviewed, particularly those project grants which appear to provoke the greatest objections from State and local officials. In so doing, grant-specific procedural costs should be differentiated from procedural costs associated with the pursuit of national policy objectives, since the former should prove easier to control on the level of individual grant programs. Again, selectivity is important, for useful policy comparisons can most readily be made among the cost-impact results of similar programs.

2. As a longer term effort, consideration should be given to the programmatic and political feasibility of shifting some grant funding from project grants to other aid instruments as a means of reducing their unique procedural costs. Although an increasingly popular cause at present, such an effort contains the danger of overrelaxation of Federal procedural standards in the name of economy. Therefore any further moves in the direction of "decategorization" should be accompanied by careful study of the potential quantitative and qualitative outcomes of such an initiative.

3. On a trial basis only, attempts might be made to demonstrate both the possibilities and the limitations of estimating the intergovernmental costs of a narrowly defined set of legislative bills. From the many problems of cost estimation outlined in this paper, such an effort might likely prove to be a general disappointment or, at best, a modest success in certain limited areas of analysis. In either case, a small demonstration effort should be well worth the price.

Should such a cost-estimation experiment prove analytically useful, its hypothetically more permanent successor might be institutionalized in a legislative fiscal note procedure linked to a program of partial Federal reimbursement of nationally mandated policy compliance costs of State and local governments. Again, however, this does not offer much practical promise, for any such activity could become plagued by a legalistic appeal adjudication process which would complicate—rather than simplify—the intergovernmental fiscal system.

4. With reference to such national policy objectives as environmental protection, procedural costs might be reduced through priority setting for the achievement of feasible goals. While attainable programmatic goals might then receive increased emphasis using present approaches, demonstration experiments of new techniques could be sanctioned in those areas where current practice has been less successful with an especially difficult problem. Environmental protection is well suited for such experiments. The range and sources of pollution problems are diverse, and, accordingly, success to date has been more significant in some areas than in others.

5. In a limited number of instances, structural modification of certain Federal policy requirements might reduce State-local procedural costs while doing little harm to the policy purposes upon which they are based. Redefining the area of coverage of certain programmatic parameters could yield a more analytically rational policy in practical application. Any such redefinition, however, need not guarantee lower costs, for the major purpose of any such modification should be a more allocationally appropriate standard for administrative use. From that

standpoint, State-local compliance costs might well increase in some areas while decreasing in others.

In closing, two points should be emphasized. First, in any review of Federal procedural requirements upon State and local governments, the important—though frequently conflicting—goals of rational economic allocation and distributional equity must not be forgotten even in the most compelling quest for lower costs and greater economy in government. Important and legitimate purposes of many Federal policy requirements can too easily become lost from the view even of the most well-intentioned economizer. Second, in this regard, we should not assume that economic analysis has at its disposal sufficient time, data, and technique to resolve all the problems of excessive government cost and technically inefficient government operation. In many instances, the simplification of cumbersome administration and the improvement of Federal interagency coordination could accomplish more than could an effective cost-estimation technique. Unnecessary costs can still be reduced even if their true magnitudes cannot accurately be estimated.

THE IMPACT OF SELECTED FEDERAL ACTIONS ON MUNICIPAL OUTLAYS

By Thomas Muller and Michael Fix*

CONTENTS

	Page
Summary	326
I. Introduction	328
1. Growth in Federal assistance and mandates	328
2. Response of cities to the growth of aid and mandates	329
3. Selection of jurisdictions as case studies	331
II. Federal mandates and conditions	332
1. Background	332
2. Environmental regulations—The Clean Water Act	334
3. Davis-Bacon Act	339
4. Unemployment compensation	344
5. Bilingual education	347
6. Education of the handicapped	351
7. Access to public transportation for the elderly and handicapped	356
III. Federally induced local expenditures	359
1. Background	359
2. The Comprehensive Employment and Training Act	360
3. The high impact anti-crime program	363
4. Federal Insurance Contributions Act	364
5. The community development block grant	365
IV. Findings	366
1. Magnitude of fiscal impact identified	367
2. Differences in impact and their causes	371
3. Projected fiscal effects of identified mandates	373
4. Policy recommendations	376
References	377

SUMMARY

The growing Federal presence in State and local government budgets is illustrated by both increased financial aid to the lower tiers of government and increased Federal mandates and other regulations requiring the outlay of State and local funds. While most officials view Federal aid favorably, this assistance is not without its burdensome consequences. Indeed, there appears to be a growing concern among local officials that the ability to allocate revenue in a timely and efficient manner is being slowly eroded by new Federal guidelines, regulations and other requirements.

This paper examines Federal mandates and those aspects of Federal assistance to local governments which increase local outlays. Specifically, the paper focuses on the incremental or added costs in-

*Thomas Muller is a principal research associate and Michael Fix is a research associate of The Urban Institute. This paper was completed with the assistance of Reed Hansen and Beth Humstone of the Urban Institute.

curred by local government as a result of Federal requirements. Seven jurisdictions (six cities and one county), chosen on the basis of size and geographic representation, were studied to determine the ways in which localities are affected by Federal regulations and mandates. Data were gathered primarily through discussions with government officials and others in the jurisdictions, through a review of local budgets and other financial records, and through an examination of expenditure trends before and after mandates and requirements came into effect. The selected Federal requirements examined were limited to a small group of programs whose incremental costs could be quantified and include the Clean Water Act; the Davis-Bacon Act; Unemployment Compensation; Bilingual Education; Education of the Handicapped; and Transportation for the Elderly and Handicapped. In addition, the paper examines the effect of recent changes in the Comprehensive Employment and Training Act; the High Impact Anti-Crime program; the Federal Insurance Contribution Act; and the Community Development Block Grant program.

Each program area listed above was analyzed in terms of its effects on local and/or State outlays in the case study jurisdictions. Although effects varied by locality, due to such factors as previous practices, demography, migration patterns, fiscal conditions, and the like, findings support the concern that the cost of Federal requirements to local government are substantial.

Excluding the cost of the Davis-Bacon Act, the average per capita cost of the selected mandates was $25 in 1978, with costs varying from $6.00 per capita in Burlington, Vermont, to $51.03 in Newark, N.J. The aggregate local and State annual operating costs of mandates examined in the seven jurisdictions were about $78 million, with costs expected to rise substantially during the next 3 to 4 years. Aggregate local and State capital outlays to meet the cost of the mandates were close to $150 million in the seven jurisdictions. Assuming that all capital outlay costs will be met from long-term bonds, the annual repayment cost would be about $13 million. Thus, a conservative estimate of the aggregate annual operating and capital outlays in the seven jurisdictions is over $90 million, or $33 per capita. As a result, the typical household residing in one of the jurisdictions pays close to $100 in the form of user costs and local taxes. This figure excludes the local cost of social service programs, OSHA, the Clean Air Act, data collection and reporting requirements and other administrative costs. Private sector costs, except those met from sewerage treatment user charges, and most indirect costs are also excluded.

Assessments of the cost of individual programs and projections for the future were also undertaken. Of all programs examined in the study, the Clean Water Act produced the largest financial impact on local government. While capital outlays associated with meeting the requirements of the Act are expected to be reduced by the early 1980's, increasing operating costs will become the dominant concern. The effects of the Davis-Bacon Act are likely to continue at present levels, since Federal funding levels for capital facilities are expected to remain the same. However, given the data limitation, any estimate of the Act's impact is difficult to make. The cost of unemployment compensation during fiscal year 1978 and 1979 was lower than that projected

by local officials, largely due to growth in private employment since 1976. However, a downturn in the economy would cause a sharp rise in this outlay. The costs of bilingual education are expected to increase during the next several years as a result of increasing immigration. The effects of this mandate will be most strongly felt in northeastern and southwestern cities. The provision of educational services for the handicapped is probably the most rapidly rising public expenditure in the Nation as both the costs per student and number of students requiring this service are rising. The capital cost of public transportation for the handicapped varies with the age of the transit system in each city. Operating cost data will not be available for some time.

The paper concludes that since the costs of Federal mandates to selected localities comprise a considerable share of all Federal funds received by these communities, there is a need to re-examine Federal requirements which condition the receipt of funds, particularly if the case study jurisdictions are representative of other communities nationally. Recommendations focus on the means to obtain more reliable estimates of the total costs to both the Federal and local government at the time legislation is being considered and on adapting more flexible standards to reflect the fiscal, locational, demographic, and other differences among communities.

I. Introduction

1. Growth in Federal Assistance and Mandates

The Federal-local government relationship during the 1970's has been marked by two parallel developments: the rapid growth of Federal aid to cities and a substantial expansion in the number, scope and type of Federal mandates and other regulations. To some extent, the two events are interrelated: each new Federal program providing assistance to communities creates an added set of regulations which imposes new rules on recipient jurisdictions. This, in turn, typically results in higher local expenditures.

The growing Federal presence in State and local budgets can be illustrated by examining changes in Federal aid in recent years. Between 1970 and 1978, State and local outlays on a per capita basis increased by 131 percent, or almost twice the inflation rate, while Federal aid to States and localities increased by 218 percent. Almost one-third of the total direct increase in local-State outlays between 1970 and 1978 was the result of expanded Federal aid. At the State level, Federal funds in 1977 were equal to 45 percent of all tax revenue raised by States. The most dramatic increase in Federal funding was directed at cities. Between 1970 and 1978, this aid increased from $1.3 to $8.9 billion. In 1970, this assistance was equal to less than 10 percent of the tax revenue raised locally; by 1977, Federal funds accounted for more than one-third of locally raised taxes (see table 1). As shown in table 2, the growth has been steepest in cities with fewer than 50,000 residents, although cities with 100,000 to 1 million residents are the most dependent on Federal aid. Regardless of city size, however, Federal funds represent a significant portion of local revenues.

TABLE 1.—FEDERAL FUNDS AS PERCENTAGE OF LOCAL TAXES 1970-77

[In billions of current dollars]

	Direct Federal aid	Locally raised taxes	Percent Federal aid of local taxes
Fiscal year:			
1970	$1.3	$13.6	9.6
1971	1.9	15.1	12.6
1972	2.5	17.1	14.6
1973	4.4	18.5	23.8
1974	5.5	19.4	28.4
1975	5.8	21.1	27.5
1976	7.4	23.3	31.8
1977	8.9	26.1	34.1

Source: Bureau of the Census, City Government Finances, various issues.

TABLE 2.—FEDERAL AID AS A PERCENTAGE OF LOCAL TAXES BY CITY SITE, FISCAL 1972 AND 1977

	1,000,000 and over	500 to 1,000	300 to 500	200 to 300	100 to 199	50 to 99	Under 50	Total
1972								
Federal	31.0	62.0	29.0	33.0	20.0	12.0	5.0	19.0
Local taxes	292.0	182.0	131.0	134.0	132.0	111.0	70.0	129.0
Percent Federal/local	10.6	34.1	22.1	24.6	15.2	10.8	7.1	14.7
1977								
Federal [1]	107.0	158.0	114.0	[2] 77.0		46.0	31.0	66.0
Local taxes	475.0	274.0	200.0	[2] 19.1		159.0	103.0	191.0
Percent Federal/local	22.5	57.7	57.0	40.3		28.9	30.0	34.6

[1] These data are not from the census publication noted below. The values shown represent corrected unpublished data from the Bureau of the Census.
[2] Combined for city population between 100,000 and 300,000.

Source: Bureau of the Census, City Government Finances, fiscal 1972 and fiscal 1977.

The dramatic increases in Federal aid have not been cost free to local jurisdictions. During the period of the sharp rise in Federal assistance, mandates requiring higher local outlays also proliferated. Thus, some of the increased aid was offset by higher federally induced outlays to meet growing Federal requirements which affected almost every facet of municipal functions.

2. Response of Cities to the Growth of Aid and Mandates

Despite the mandates and their associated costs, most local officials view additional Federal aid as a positive development. Indeed, the stabilization of local tax burdens between 1972 and 1977 was exclusively the result of increased Federal aid.[1] The mayor of Jacksonvile, Florida, a city with a healthy, growing economy, forcefully expressed one view on Federal funds in his budget message to city residents:

Through our Office of Intergovernmental Affairs and my own participation in national organizations (and with the assistance of many members of this Council) we have maintained a position of readiness—aggressively pursuing every single federally funded program possible. . . . Without these dollars, we would

[1] Thomas Muller, testimony to Subcommittee on Fiscal and Intergovernmental Policy of the Joint Economic Committee, March 20, 1978.

have no choice but to be taxing our citizens at the absolute maximum rate to simply maintain the basic level of service and the quality of life in Jacksonville.[2]

While the view may not be representative of all mayors, "grantsmanship" as a means of obtaining Federal funds has become a growth industry. This is not to suggest, however, that city officials are unconcerned over mandates and their implications for local outlays. Indeed, there appears to be growing concern among local officials that the ability to allocate their revenue in an efficient and timely manner is being slowly eroded by new Federal (as well as state) guidelines, mandates and requirements.[3] At the 1978 League of Cities meeting in St. Louis, a resolution introduced by Mayor Bradley of Los Angeles, part of which is quoted below, expressed the view of most mayors:

Whereas the Federal Government has historically mandated programs and activities that have caused cities to make expenditures which generally are not reimbursable; and

Whereas federally mandated programs and activities reflect a lack of sensitivity on the part of the Federal Government to understand the fiscal problems of cities; and

Whereas there are no provisions in federal mandated programs and activities to reimburse or provide adequate funds to cover any cost cities incur; be it

Resolved, That the National League of Cities reemphasizes its support and encourages the introduction of legislation in the 96th Congress that requires reimbursement to cities for any cost they incur due to federally mandated programs.[4]

This paper focuses on Federal mandates and those aspects of Federal assistance to local governments which increase local outlays. However, no attempt has been made to include all or even a small share of the hundreds of Federal mandates and regulations which can affect the level of local outlays. Specifically, the report concentrates on local expenditures induced by compliance with the following acts: the Clean Water Act, the Davis-Bacon Act, Unemployment Compensation, Bilingual Education, Education of the Handicapped, Transportation for the Elderly and Handicapped, Comprehensive Employment and Training, High Impact Anti-Crime, the Federal Insurance Contribution Act, and the Community Development Block Grant program. The data on local outlays derive primarily from a review of local budgets and other financial records.

The paper is limited to the above named Federal programs since the costs which otherwise would not have been incurred in the absence of Federal mandates could be readily estimated. The local outlays associated with meeting the requirements of the mandates may be defined as incremental costs.

While data on total outlays associated with implementing a particular requirement are typically available, jurisdictions do not maintain separate information on incremental costs. Therefore, these data had to be obtained from two sources: discussions with municipal personnel—those with professional expertise and others with managerial or fiscal responsibility—and through an examination of expenditure trends before and after mandates were enacted.

---

[2] City of Jacksonville, Florida, 1977–1978 budget.

[3] While outside the scope of this paper, state mandates, many resulting from State-Federal agreements, are at least as costly as Federal mandates.

[4] National League of Cities, "Proposed Resolution—EG No. 14; Continued Support for Federal Funding of Federal Mandates," submitted by Tom Bradley, mayor of Los Angeles, Nov. 1, 1978, approved Nov. 29, 1978.

In most instances, personal interviews were conducted, frequently followed by telephone conversations and correspondence. This approach restricted to a small number the pool of mandates and regulations selected for study. The investigation itself was limited to an examination of added local and State government outlays. No attempt was made to identify and quantify benefits which may have accrued as a result of implementing the mandates discussed in this report. Therefore, the report should not be considered a definitive examination of a very complex issue, but rather an initial step which provides insight into the magnitude of the problem as illustrated by the data from a small group of cities. As noted in a recently released background paper by the Congressional Budget Office, aggregate costs provide little information regarding the burden imposed by specific requirements on levels of government. This paper, rather than estimating aggregate outlays, is aimed at assessing the incremental costs of specific mandates at the local level. In most instances, however, only direct costs are considered, although the authors recognize that a comprehensive analysis would have to take into account indirect costs to assess the full impact of mandates on the national economy.

3. Selection of Jurisdictions as Case Studies

Jurisdictions were selected to represent each of the Nation's regions, with differing tax rates, responsibility for public services and personal income. Very large and small cities were excluded from the sample. As shown in table 3, jurisdictions range in population from 39,000 (Burlington) to 848,000 (Dallas). Three of the cities are in northern States: the three other cities and one urban county are in the Southern or Western States. Property taxes vary from $0.88 per $100 market value in Seattle to $4.30 per $100 in Newark, with income, and thus ability to pay for services, also differing substantially.

TABLE 3.—JURISDICTIONS SELECTED AS CASE STUDIES

City: Region	Population (1976) (In thousands)	Effective property tax rate [1] (1977)	School responsibility	Per capita income (1976)
				1
Alexandria: South	108	$1.38	Yes	$7,785
Burlington: Northeast	39	NA	Yes	4,343
Cincinnati: North-Central	410	1.22	No	4,845
Dallas: South	848	NA	No	5,712
Fairfax County: South	531	1.43	Yes	7,47
Newark: Northeast	331	4.30	Yes	3,58[6]
Seattle: West	491	.88	No	6,44[1]
Average	394	1.84		5,740

[1] Per $100 market value.

Source: 1977 Census of Government, vol. II, Population Estimates and Projections, p. 25.

A comparison of growth in Federal aid in the six cities with the national pattern, as shown in table 4, demonstrates that Federal and State aid to these jurisdictions is above the average received by all cities nationally. This is attributable to two factors: (a) the average city size (400,000) is considerably above the national city average, and larger cities typically receive more Federal revenue compared to

smaller cities; and (b) half the cities include schools in their budgets, compared to only about 10 to 20 percent nationally.

While the information collected may be roughly representative of medium sized cities, the inclusion of only one smaller jurisdiction—Burlington—means that there is very limited data on cities with a population of fewer than 100,000 residents, although smaller cities are the residences of the majority of the Nation's urban population. Thus, caution needs to be exercised in relating results from the selected jurisdictions to others in the Nation.

TABLE 4.—GROWTH OF FEDERAL OUTLAYS PER CAPITA, FISCAL 1972-77[1]

[In millions of dollars]

Revenue source	Alexandria[2]		Burlington[2]		Cincinnati		Dallas		Newark[2]		Seattle	
	1972	1977	1972	1977	1972	1977	1972	1977	1972	1977	1972	1977
Federal	$16	$279	$31	$133	$49	$208	$5	$45	$6	$137	$26	$112
State	72	160	51	56	23	68	1	7	245	513	32	49
Local	380	631	705	872	234	306	174	282	341	432	195	337
Federal as percent of local	4	44	4	15	21	68	3	16	2	32	13	33

	6-city average per capita		Share by source (peicent)		All U.S. cities per capita		Share by source (percent)		Cities under 50,000 per capita	
	1972	1977	1972	1977	1972	1977	1972	1977	1972	1977
Federal	$22	$152	5	17	$19	$65	7	15	$5	$55
State	71	142	16	16	63	104	24	24	23	44
Local	338	582	79	67	178	269	69	62	107	162

[1] Fairfax County excluded.
[2] School revenue included.

Source: Bureau of the Census, City Government Finances in 1971–72 and 1976–77, and financial data from city of Burlington.

II. Federal Mandates and Conditions

1. *Background*

Local expenditures affected by Federal requirements and programs can be grouped into several broad categories. This paper will consider three of the categories: (1) Outlays which result from Federal mandates; (2) expenditures caused by meeting conditions attached to the use of Federal revenue; and (3) costs which arise from replacing Federal funds with local revenue following the reduction, retargeting or elimination of Federal grants. The first two categories are discussed in this section. In addition to these categories, Federal assistance is provided as part of various categorical programs which require minimal or no local funds. The effects of such programs on local expenditures are beyond the scope of this paper.

The cost of government mandates has been the subject of several reports undertaken by State and local agencies. Both Colorado and Oregon, for example, examined the effects of mandates on local outlays in the 1973–74 period (2, 3). As the focus of these studies was State mandates, less effort was made to examine the effects of Federal requirements. A more comprehensive examination of both Federal and

State mandates was undertaken by Fairfax County, Va., in 1978 (4). Elements of this effort, updated to include more recent information, have been incorporated into this paper.

FEDERALLY MANDATED LOCAL EXPENDITURES

Mandated costs to local governments arise as a result of Federal laws which compel jurisdictions either to undertake cost-producing actions or to refrain from engaging in cost-saving activities. The regulations affect all jurisdictions to which they apply, regardless of the extent to which jurisdictions use, or refrain from using, Federal funds. These laws regulate municipalities' activities in much the same way that civil and criminal laws impose affirmative obligations on the behavior of private individuals.

Federally mandated local expenditures can be further separated into two categories. The first subset includes those Federal mandates which require local outlays while simultaneously establishing a Federal-funding mechanism to offset these expenditures. An example is the construction grant programs established under title II of the Clean Water Act. These mandates can be distinguished from those which require the local jurisdiction to bear most or all of the costs which arise as a result of the regulation. This category is illustrated by the Safe Drinking Water Act which sets certain mandatory standards for the quality of local drinking water, but which provides no support for local jurisdictions which are compelled to update or construct new treatment facilities to comply with the Act.

Federal mandates can also be divided into two groups on the basis of legislative intent: (*a*) Those which are enacted with the express legislative objective of curbing or encouraging certain activities at the local level; and (*b*) those which are the indirect result of legislation motivated by another objective. The mandate of Public Law 94–142 which requires that a "free and appropriate education be offered to all handicapped children between the ages of 3 and 21" is an example of the former. The Privacy Act, which prohibits the Internal Revenue Service from disclosing private taxpayers' earnings to local officials on an unsolicited basis, is an example of a cost to local governments which results in lower levels of compliance and the added expense associated with obtaining information on earnings.

CONDITIONS ON THE USE OF FEDERAL FUNDS RESULTING IN ADDED LOCAL OUTLAYS

Costs in this category are incurred only if Federal funds are used at the local level. Thus, in theory (and perhaps only in theory), local governments have the option of foregoing these costs by refusing to accept Federal assistance. For example, if a local government should refuse all Federal support for a municipal construction project, that government would no longer be obliged to pay Davis-Bacon wages to workers on the project. In reality, few local governments reject Federal aid because of requirements associated with such aid, since typically (but with notable exceptions) the cost of compliance is less than the level of funding.

There are two types of conditions which result in added local outlays when Federal funds are utilized. The first can be termed "substantive conditions"—regulations which in virtually all instances require the alteration of local operations or policies to comply with Federal guidelines. An example would be the proposed regulations issued by the Department of Transportation which require that public transportation agencies receiving Federal funds equip buses with lifts in order to make them accessible to the handicapped. The Davis-Bacon wage requirements for all construction projects undertaken with Federal assistance or funding would also be illustrative of substantive conditions which must be met prior to the receipt of Federal assistance.

The second type of condition resulting in added local outlays imposed by the Federal Government is the requirement that local governments provide a certain designated percentage of total program cost—a local "match"—as a precondition to the use of Federal funds. For example, the local and/or State share which must be contributed toward the construction of secondary waste treatment facilities under title II of the Clean Water Act represents 25 percent of total project cost.

Federal grants and aid also impose other, less direct costs on local governments—costs which for several reasons can be said to derive from local government reliance upon Federal funds. First, these costs result from delay in the exercise of Federal Government responsibility and from inefficiencies which can arise from the skewing of local priorities to obtain "cheap" Federal dollars. Second, as Federal aid increases, the incentive for private assistance diminishes, while at the same time there is also a transfer of costs from the private to the public sector. This phenomenon is best illustrated by developments in the area of education for the handicapped. Finally, where the share of Federal matching aid is high—75 or 80 percent—there is a tendency to overdesign facilities beyond the level of reasonable demand and thus produce inefficiencies which result in high operating costs.

Any comprehensive examination of mandates and requirements should take into account these and other aspects of Federal activities, such as the degree to which Federal funds stimulate additional outlays or substitute Federal for local dollars. This paper, however, is limited to estimating the fiscal effects of six mandates and several other programs in seven case study jurisdictions.

2. *Environmental Regulations—The Clean Water Act*

The Clean Water Act requires the development and implementation of wastewater treatment management plans which meet pollution discharge standards promulgated by the U.S. Environmental Protection Agency.[5] The Act also authorizes a large Federal grants program for local and regional governments to finance the construction or conversion of waste-water treatment facilities to comply with Federal effluent limitation levels. Most Federal grants are made available to local governments at a Federal–non-Federal matching ratio of 75 percent to 25 percent.

[5] The Clean Water Act PL. 92–500, as amended by PL. 95–217 (1977).

Contributions from local and regional governments to cover construction costs of compliant waste-water treatment facilities comprise one of the largest federally-mandated local expenditures. Increased local government expenditures arising from the secondary treatment mandate of the Clean Water Act also derive from the rapidly escalating operation and maintenance costs which users and local governments must bear. No Federal reimbursement is available to mitigate local operation and maintenance costs.[6] Operating and capital outlays for the seven jurisdictions are shown in table 5.

TABLE 5.—LOCAL AND STATE OUTLAYS ATTRIBUTABLE TO THE CLEAN WATER ACT

[In millions of 1978 dollars]

	Total operating, local	Total capital		Per capita operating, local	Per capita capital	
		local	State		local	State
Alexandria	$1.6	$6.8	$2.3	$14.81	$62.96	$21.30
Burlington	0	0	0	0	0	0
Cincinnati	[1] 5.0	13.0	0	12.20	31.70	0
Dallas	4.1	12.3	0	4.83	14.50	0
Fairfax County	[2] 4.1	30.2	3.8	[2] 7.72	56.87	7.20
Newark	10.4	20.7	31.1	31.42	62.54	93.96
Seattle	2.3	[3] 12.1	18.1	4.68	24.64	38.96
Total	27.5	95.1	55.3			

[1] For 1980.
[2] Projected to be $13,500,000, or $25.23 per capita when facilities are operational.
[3] Currently projected.

ALEXANDRIA, VA.

The city of Alexandria, currently constructing an advanced treatment plant, would not have considered such an undertaking in the absence of Public Law 92–500. However, it would have expanded its secondary sewage treatment facilities from 18 to 36 million gallons per day capacity regardless of any Federal legislation. The cost of plant expansion, which incorporates secondary treatment capabilities, would have been $18 to $22 million in current dollars. By contrast, the cost of the tertiary treatment plant, when completed in 1983, will be about $100 million, of which $80 million is already under contract. In current (1979) dollars, the total costs will be approximately $95 million. The differential in outlay, or $75 million, is attributable to the Federal mandate.

The total State contribution toward the tertiary treatment plant is $2.3 million, the local share is $21.4 million, with the balance paid by the federal government.[7] The incremental local capital cost attributable to Public Law 92–500 is $16.9 million. Of this total, Alexandria residents will pay $6.8 million, with the balance of the treatment capacity used and paid for by Fairfax County. Incremental operating outlays associated with tertiary treatment are estimated at $4.1 million, with the added operating cost for Alexandria residents of $1.6 million, or $14.81 per capita. Higher operating outlays are a

[6] These costs are met by user charges imposed on householders and other users.
[7] The lowlevel of State aid is linked to the 75 percent Federal funding. Prior to Public Law 92–500, the state share was higher, but state funds would have been allocated for less expensive facilities in the absence of the public law.

combination of three factors: additional labor, chemical products, and the rapidly rising cost of energy required to operate tertiary treatment plants.

BURLINGTON, VT.

Burlington was already in the process of upgrading its primary treatment plants to meet the secondary treatment requirements of state law when Public Law 92–500 was passed in 1972. Therefore, no additional local expenditures were imposed on the city as a result of the 1972 legislation. The Burlington Director of Water Pollution Control did note, however, that Federal and state reimbursement for 90 percent of capital construction costs was available both before and after Public Law 92–500 was passed.[8] Despite the existence of state standards prior to the availability of Federal funds, the city and state might have resisted upgrading the sewerage treatment plant if Federal cost sharing were not available.[9]

Unlike the situation in rural towns where the availability of Federal funds frequently results in "overdesign" of sewage treatment plants, the expansion of Burlington's facilities was necessitated by industrial growth.[10]

CINCINNATI, OHIO

The Metropolitan Sewer District of Cincinnati serves the city, 32 of 37 incorporated municipalities, and the unincorporated parts of Hamilton County. The cost of electricity to operate one of the system's two largest treatment plants, the Mill Creek Plant, increased from $500,000 a year for primary treatment to $3 million under secondary treatment. The Metropolitan Sewer District staff estimates that the annual operating budget of the Operations Division for the district will increase approximately 55 percent as a result of Federal secondary treatment requirements—from $9 million in 1978 to $14 million in 1980 when secondary treatment facilities will be fully operational. The increase will represent an added per capita operating cost of $32.

The conversion cost to secondary treatment will require the city of Cincinnati to contribute over $25 million toward the construction of waste-water treatment facilities. Unlike Burlington which had undertaken plant improvements in advance of Public Law 92–500, Cincinnati officials estimate that less than half this amount would have been expended had it not been for Federal mandates and the availability of Federal construction grants. There is no State contribution for the construction of sewage treatment facilities.

---

[8] The problem of isolating the predominating influence in the conversion from primary to secondary—whether it has been the "stick" of the Federal mandate of the Clean Water Act standards, or the "carrot" of the construction grants program, is difficult to solve in any definitive manner. For the purpose of this rather limited inquiry, we have been compelled to accept the opinion of local officials in determining the degree to which either Federal dollars or Federal penalties have proved to be more persuasive.

[9] Under a previous Federal act (Public Law 89–177), Burlington was eligible for 60 percent Federal funding, owing to an existing waste-water management plan.

[10] To date the issue of tertiary treatment at the Burlington sewage treatment plants is unresolved. According to State and local officials, EPA is giving grants for tertiary treatment a lower priority than grants for secondary treatments. Therefore, funding for tertiary treatment is not assured. An alternative to upgrading the plants is to lower the state water quality standards, an action which may not be approved by EPA.

DALLAS, TEX.

To comply with state requirements, Dallas has relied upon secondary treatment since the mid-1950's. Despite the city's early conversion to secondary treatment, state stream quality standards, promulgated in response to Public Law 92-500, have compelled the city to implement advanced secondary treatment.

The Director of the Dallas Water Utilities Department states that the simultaneous expansion and conversion of the city's 150 million gallons/day (MGD) Central Plant resulted in approximately $19 million in increased costs which could be attributed to Federal mandates. (Costs attributable to the expansion of capacity have been excluded.) As the expansion was begun before the passage of Public Law 92-500, the city's share of the cost of construction was 45 percent of the total or $8,500,000. However, plant design has been dictated by standards promulgated in accordance with Public Law 92-500.

Dallas is also engaged in the expansion and conversion of its second, smaller Southside plant. The project, which is funded under the 75-25 Federal/non-Federal ratio established by the Clean Water Act, will require outlays of $15 million to meet Federal requirements. The city's share of capital costs imposed by Federal mandates will be $3,750,000.

Operation and maintenance costs: Dallas water utilities department

Fiscal year:	*Cost (in thousands)*
1975	$1,881
1976	3,332
1977	5,226
1978	8,718
Percent change 1975-78	363

The increased operating and maintenance costs of conversion to advance secondary treatment have been dramatic in Dallas, as elsewhere. As shown above, these costs have risen by 363 percent in 3 years. The increase in operating costs for each of the years is principally attributable to changes in the technological sophistication of treatment methods needed to meet Federal standards. Approximately 75 percent of the $5.4 million increase in operating and maintenance costs between fiscal 1976 and 1978, or $4.1 million, can be imputed to the shift from secondary treatment to advanced secondary treatment. The cost increases have resulted primarily from the increased costs of chemicals and power required for advanced treatment. To illustrate the effects of higher costs, during 1977, city sewer rates rose 62 percent while water rates rose only 10.5 percent.

FAIRFAX COUNTY, VA.

As a result of Public Law 92-500 and subsequent state legislation, Fairfax County undertook a massive construction program to upgrade its treatment facilities. In addition, Alexandria, as noted earlier, is upgrading its treatment plants, with part of its capacity to be used by

Fairfax County. The estimated incremental outlays are based on the premise that the county would have provided secondary treatment even in the absence of Public Law 92–500. Estimated local incremental capital outlays, as shown in table 6, are $30.2 million. However, added operating outlays are estimated to be $13.5 million each year by the early 1980's when facilities currently under construction will be fully operational. By comparison, incremental operating outlays were only $1.7 million in fiscal 1978.

TABLE 6.—SEWERAGE TREATMENT FACILITY EXPANSION FOR FAIRFAX COUNTY USERS: CAPITAL OUTLAYS

[In millions of dollars]

Location	Total project cost	Federal grants	State grants	Local outlays	Estimated incremental outlays
Fairfax County [1]	115.0	[2] 66.3	3.8	24.9	19.7
Alexandria [3]	57.0	[4] 42.3	1.4	13.3	10.5

[1] Includes pumpover stations and Lower Potomac advanced water treatment (AWT).
[2] Grants approved as of March 1979. Projected grants should increase Federal share to about $86,300,000.
[3] Share of Alexandria facilities to be used by Fairfax County, with costs to be recovered from user charges.
[4] Projected Federal grants.

Annual incremental operating outlays

	Millions
Current Fairfax County—fiscal 1979	$4.1
Projected when facilities completed:	
Fairfax County facilities	11.0
Alexandria facilities	2.5

Source: Fairfax County Wastewater Treatment Division, Fairfax County Budget Office, City of Alexandria.

As noted by the Director of the Fairfax County Wastewater Treatment Division, Federal and State discharge standards have a "devastating effect on capital construction costs." When an advanced wastewater treatment facility has to be expanded to accommodate growth, the probability of Federal or State grant funds to pay for part of the expansion under current policies is extremely low. This leaves a jurisdiction with the responsibility for totally funding construction at a cost of 4 to 5 times greater than the cost of expanding an improved secondary treatment plant.

NEWARK, N.J.

Sewage treatment for the City of Newark, N.J., is performed by two regional authorities, the Passaic Valley and the Joint Union Meeting. Newark contributes about 40 percent of the sewage flow to the Passaic Valley plant and about 15 percent of the Joint Union Meeting plant. Capital costs arising from the conversion from primary to secondary treatment at the Passaic Valley plant are expected to exceed $500 million, while capital costs of conversion at the Joint Union Meeting plant have already reached $50 million.

Newark is scheduled to contribute 10 percent of its proportionate share of capital costs to construct each plant.[11] With construction costs

[11] The other 15 percent of the state and local match required under the Capital Construction Grants program of the Clean Water Act is contributed by the State.

related to the city's usage of each sewage plant, Newark's share of the Joint Union facility cost will be $750,000 (1.5 percent of the total outlay), while its contribution to the conversion of the Passaic Valley plant will be 4 percent, or $20 million. Annual operating costs, which will have to be paid by Newark users, will rise from the current primary treatment cost of $121,000 at the Joint Union plant and $14 million at the Passaic Valley facility. Upon conversion of the two plants, the added annual costs will rise by over $10.4 million, or over $30 per capita, from current levels.

SEATTLE, WASH.

The Municipality of Metropolitan Seattle (Metro) operates four waste-water treatment plants along Puget Sound at West Point, Alki, Corkeek Park, and Richmond Beach; these plants provide primary treatment only. A fifth plant, located at Renton on the Green-Duwamish River, provides secondary treatment before discharging waste-water into the river. Metro-supported studies indicate that secondary treatment of waste-water discharge from the Puget Sound plants would be of limited usefulness as Metro's present discharge is not substantially degrading the Sound's water quality.

While the Environmental Protection Agency is currently preparing waiver regulations regarding secondary treatment requirements, it is not possible to determine whether the waiver regulations will substantially affect Metro. In the meantime, Metro has yet to comply with EPA's regulations for secondary waste-water treatment. A comparison between the capital cost of the baseline alternative and the least costly alternative which would comply with the Public Law 92–500 requirements indicates a difference of $182 million (1978 dollars). The Federal and State matching contributions to capital construction are 75 percent and 15 percent, respectively.[12] Hence, the difference in locally funded capital costs between the alternative plans is $18.2 million. Of this total, the Seattle share is about 67 percent, or $12.1 million. Annual operating and maintenance cost of the "no action" alternative is $10.1 million, while that of the secondary treatment compliance alternative is $13.6 million, a difference of $3.5 million. Assuming the 90 percent construction grant, customer water rates would be $5.45 per month under the no action alternative and approximately $7.00 per month under the secondary treatment compliance alternative, a difference of $19 per annum.

3. Davis-Bacon Act

The Davis-Bacon Act requires that contractors working on federally assisted projects pay their employees wages and fringe benefits which are no less than those prescribed by the U.S. Department of Labor (DOL). The prescribed wages are those found by DOL to be prevailing for similar types of building in the locality where the construction is taking place.

In contrast to the other mandates discussed in this paper, the Davis-Bacon Act has been in effect for almost 40 years and has been fre-

[12] The State may not have the necessary funds to meet its 15 percent share in future years.

quently amended. In view of the act's controversial nature, it has been studied and debated for years by legislators, researchers, and various groups within the construction industry. Within the context of this paper, discussion will be limited to the act's impact on the seven selected jurisdictions, without reference to related issues.

The major focus of current discussions concerns the inflationary impact of the act. Where local prevailing wages fall below the level determined by the Department of Labor and projects are funded jointly by Federal and local agencies, the Davis-Bacon Act can have the effect of increasing construction costs above what they would otherwise be. While the act also has a potential inflationary effect on both private construction taking place in a community and projects paid for and financed by local dollars, these issues are beyond the scope of this paper.

By 1974, all States but nine had enacted "little Davis-Bacon Acts." These acts, which generally apply to municipal and State sponsored construction, vary widely from State to State. Only two of the jurisdictions—Alexandria and Burlington—examined in this paper were located in States which had not passed Little Davis-Bacon Acts. Each of the other four States has its own act, but applies varying criteria for determining prevailing wages.

ALEXANDRIA, VA.

The Davis-Bacon Act has been quite controversial in Alexandria because of the city's long involvement in the construction of the Washington Metropolitan Area Transit Authority Metrorail system. For Alexandria, as for other Washington area governments, the local share of the capital construction costs of the Metro system has apparently increased as a consequence of the Davis-Bacon wage mandates.[13] The computation of a local government's share of the construction costs of the system is determined partly by the cost of building Metrorail lines and facilities within a jurisdiction's political boundaries. Thus, the impact of the Davis-Bacon wage scales on local government expenditures can be determined by the extent to which local wage scales differ from those rates set by the Department of Labor.

The construction industry in Alexandria and northern Virginia is non union to a greater degree than is the case in other parts of the Washington metropolitan area. The open shop nature of the industry explains at least part of the difference in local heavy construction wages between northern Virginia and the other jurisdictions in the Washington metropolitan area.[14] Since Davis-Bacon wage rates for heavy construction in Alexandria are based on a scale which applies to the entire metropolitan region, their impact on Alexandria and other northern Virginia governments is probably greater than on other nearby jurisdictions.

While Davis-Bacon heavy construction wage rates for Metro remain the same throughout the Washington Metropolitan area, the U.S. Department of Labor recognized 2 years ago that the prevailing wage levels of the northern Virginia region in building trades were

[13] For a discussion of the Metro wage issue, see Armand J. Thiebolt, Jr. The Davis-Bacon Act, University of Pennsylvania, 1975, pp. 109–117.

[14] Due to lower State and local taxes, the cost of living is probably somewhat lower in northern Virginia compared to the District of Columbia, Montgomery, and Prince George's Counties, Md., which could also affect wage rates.

lower than those found elsewhere in the metropolitan area. As a result, DOL began to issue local wage standards targeted specifically to the area. Thus, in northern Virginia, the differential between local and Davis-Bacon wages in building trades construction has been substantially reduced. As data for Alexandria illustrate, there are substantial differences between city employee and Davis-Bacon Act wages for similar occupations (see table 7). However, city wages are no doubt lower than those paid to private non-unionized workers in many of the occupations. In addition, skill levels of city employees are not comparable to those of private sector workers.

TABLE 7.—COMPARISON OF WAGES BETWEEN ALEXANDRIA CITY BASIC HOURLY RATES AND DAVIS-BACON ACT RATES

Worker category	City of Alexandria, Va., basic hourly rates	Basic hourly rates to be paid under Davis-Bacon	Percent difference
Building construction:			
Bricklayers	$5.02	$8.82	75.7
Carpenters	5.02	7.29	45.2
Carpet layers	5.02	6.85	36.5
Cement masons	5.02	7.57	50.8
Drywall finishers	4.52	7.96	76.1
Drywall hangers	4.52	7.50	65.9
Laborers, unskilled	3.96	4.91	24.0
Painters, brush	5.02	5.59	11.4
Pipefitters	5.02	11.48	128.7
Plumbers	5.02	6.00	19.5
Roofers	5.02	6.25	24.5
Tile setters	5.02	9.23	83.9
Power equipment operators:			
Backhoe operators	4.52	7.09	56.9
Bulldozer operators	4.52	8.64	95.6
Mechanics	5.40	9.19	70.2
Mean	4.84	7.64	57.8

Source: Memorandum from Alexandria assistant city manager.

BURLINGTON, VT.

Since Burlington is a relatively small city, the level of federally funded construction in any given year tends to be small. Nevertheless the act has been responsible for increasing local outlays. The city recently built an ice skating rink, financing the construction, in part, with $800,000 in general revenue sharing funds. As the city initially neglected the Davis-Bacon wage requirements, it was forced to pay an additional $30,000 to cover the costs of meeting federally mandated wage rates.

Local officials report that Federal mandates, one of them being the Davis-Bacon Act, increased the costs of a $707,000 EDA-financed school project by 25 to 30 percent.[15] The Federal requirements to which most of the added costs could be attributed were Federal wage mandates and minority contracting requirements.[16] Table 8 shows that Davis-Bacon wages during 1974 in Chittenden County, which includes Burlington, were about 27 percent above prevailing wages determined by a survey conducted by the contractor's association for the Department of Labor. The differences shown in surveys led to a downward adjustment in Davis-Bacon wages.

[15] The labor share of contracts is usually about 35 to 40 percent.
[16] Since minorities comprise only 0.6 percent of the State's population, local administrators had to obtain minority firms from out of State, many of which were later disqualified, adding to the cost of the projects.

Comparable data for 1978 are available for only the two occupation categories for Chittenden County, as shown below:

	Davis-Bacon	Survey[1]	Percent difference 1973	Percent difference 1978
Carpenter (1978)	$7.95	$6.83	15.4	16.4
Laborer (1978)	4.32	3.96	43.0	9.1

[1] Survey undertaken by State of Vermont.

As the data show, the differential increased slightly for carpenters, but decreased substantially for laborers.

CINCINNATI, OHIO

Cincinnati poses a striking contrast to Alexandria and Burlington, presumably because union strength is much greater in Ohio than in Virginia and Vermont—predominantly open shop States. In 1931, a year before the passage of the Federal act, Ohio enacted its own "little Davis-Bacon Act" which mandated that the State Department of Industrial Relations determine the prevailing wage for each political subdivision of the State which contracted for the construction of public improvements.[17] As required by State law, State-established prevailing wage scales have in the past been equal to or higher than those set by the U.S. Department of Labor (see table 9).

TABLE 8.—CHITTENDEN COUNTY, VERMONT: DAVIS-BACON—SURVEY COMPARISON OF HOURLY RATE, 1973–74

	Davis-Bacon Act	Survey	Percent difference
Bricklayer	$8.41	$7.76	8.4
Carpenter	7.40	6.41	15.4
Cement mason	8.41	7.76	8.4
Electrician	8.12	5.96	36.2
Laborer	6.25	4.37	43.0
Painter	5.40	3.50	54.3
Plumber	8.02	6.60	21.5
Roofer	8.70	3.82	127.8
Sheet metal worker	7.64	7.64	0
Mean	7.59	5.98	26.9

Source: Armand J. Thieblot Jr., Davis-Bacon Act, University of Pennsylvania, 1975, table 16.

TABLE 9.—DAVIS-BACON AND STATE PREVAILING HOURLY WAGE LEVELS, CINCINNATI, OHIO, 1978

Type of worker	Federal building construction	Federal heavy highway construction	State prevailing wage
Carpenters	$12.20	$12.55	$12.20
Plumbers	12.82	12.82	12.82
Electricians	12.50	12.50	13.10
Lathers	13.32	NA	13.32
Pipefitters	13.00	13.00	13.10
Iron workers	12.33	12.33	12.33
Asbestos workers	13.16	NA	13.16
Brick layers	12.845	12.845	12.845
Cement masons	12.08	12.03	12.08
Elevator constructors	11.78	NA	12.36

Source: Ohio Department of Industrial Relations.

[17] Ohio Rev. Code Ann. Sec. 4115.04, Supp. 1975.

One reason for this relatively rare phenomenon is that State law prescribes that local prevailing wages are not to be less than those payable "under collective bargaining agreements or understandings between employees and bona fide organizations in the labor force."[18] Although union labor is used for all city construction, the fact that Davis-Bacon and State prevailing wage scales are roughly comparable means that no additional labor costs are imposed on the city by this Act.

DALLAS, TEX.

Texas law requires that all political subdivisions of the State which undertake public works projects must determine the prevailing wage rates for each category of employee engaged in construction projects within the locality.[19] The classification of construction employees which Dallas has developed is generally the same as that drawn by the U.S. Department of Labor. The locally established wage scales set the minimum wage for all construction undertaken by municipalities without Federal funds.

Two categories of Federal wage scales are found in Dallas—building construction and heavy construction rates. The impact of Davis-Bacon is primarily felt by the city government in the building construction field, as the building trades industry is more heavily unionized than the heavy construction industry. Within the building construction industry, Davis-Bacon wages approximate union rates—which are significantly higher than locally established prevailing wages. Within the heavy construction field, Davis-Bacon wage levels reflect the lower prevailing wage rate established by the city government (see table 10).

TABLE 10.—DIFFERENCE BETWEEN DALLAS CITY AND FEDERAL DAVIS-BACON WAGE RATES FOR BUILDING CONSTRUCTION

Trade	Base Davis-Bacon wage rate,[1] 1978	Base city wage		Percent Difference Davis-Bacon recommended, 1979
		Prevailing, 1978	Recommended, 1979	
Bricklayers	$10.20	$8.71	$9.91	3.0
Carpenters	10.20	8.01	8.37	12.4
		8.14	9.17	12.8
Cement mason	9.49	7.82	7.82	2.14
Electricians	11.03	8.56	9.69	13.8
Laborers	6.32	5.54	4.48	41.1
Operating engineers (heavy)	10.00	7.93	6.04	65.6
Painters	9.52	8.58	8.30	14.7
Plumbers	10.23	9.36	9.59	6.7
Roofers	9.21	7.42	7.59	21.3
Sheet metal	10.42	9.00	7.58	37.5
Mean	9.68	8.10	8.05	20.2

[1] Source: Federal Register, vol. 43, No. 242, Dec. 15, 1978.

The relationship between Davis-Bacon and union wages is interesting since, in 1977, 95 percent of all building projects in the city with a value of over $50,000 were apparently undertaken with non-union labor. Projects of this size accounted for 81 percent of the total value of all construction within the Dallas-Fort Worth region in that year.[20]

[18] *Ibid.*
[19] Vernon's Ann. Civ. St. Art. 5159a (1969).
[20] The North Texas Contractors Association, 1977.

FAIRFAX COUNTY, VA.

Fairfax County solicited bids from contractors for one sewage pump and one water main project during 1978. Each bidder was asked to submit two cost estimates for the same project based upon EPA and non-EPA funding options. As a result, each contractor's bids reflected both the presence and absence of Davis-Bacon wage requirements.

Of the six bids received for the first project, five estimated slightly lower costs for the non-EPA funding option. However, the lowest bid showed no difference between the two alternatives. Six bids were likewise received for the second project. Four were the same for the two alternatives, whereas the other two were higher if EPA and thus Davis-Bacon provisions were followed. Again, the lowest bid showed no differences between EPA and non-Davis-Bacon costs. The average difference for the first project was 1.6 percent or $101,000, while the average difference for the second project was 0.5 percent, or $10,000. However, since the lowest bids for both projects were accepted, the county incurred no added cost as a result of the Davis-Bacon Act. In the view of local officials, the high demand for skilled labor in the Washington, D.C. market caused a convergence between Davis-Bacon and other wages since the mid 1970's.

NEWARK. N.J.

Discussions with city officials indicated that the Davis-Bacon wage requirements currently impose no incremental cost to the city.

SEATTLE, WASH.

As in Cincinnati, the Davis-Bacon Act produces no cost-inducing impact in Seattle. Union wage scales, typically close to those certified for Federal construction, dominate the construction labor market. Further, State law closely parallels Federal legisation.

4. *Unemployment Compensation*

With the enactment of the Federal Special Unemployment Assistance Act (SUA) in 1974, most local and State employees became eligible, for the first time, to receive unemployment compensation. Under the provisions of this legislation, unemployment compensation payments were financed in full by the Federal Government, since the fiscal problems resulting from the severe recession were assumed to make it difficult for local governments to absorb the cost of widespread lay-offs of public employees. In 1976, Congress passed legislation which permanently entitled non-elected State and local public employees to the same unemployment compensation coverage as other workers. However, under the new legislation, most of the burden of claims payments was shifted to local and State governments. There was, and continues to be, some concern over the cost of these payments which vary considerably as indicated in table 11.

TABLE 11.—LOCAL COST OF UNEMPLOYMENT COMPENSATION

[In thousands]

	Outlays in 1978	Per capita outlays, 1978
Jurisdiction:		
Alexandria (schools only)	$97	$0.90
Burlington	35	.90
Cincinnati	[1] 350	.85
Dallas	[2] 40	.05
Fairfax County (including schools)	[3] 120	.23
Newark	60	.18
Seattle	225	.46
Total/average	927	[4] .51

[1] Payments were $500,000 in 1977 and are expected to be $400,000 in fiscal 1979.
[2] Payments projected to be $80,000 in 1979–80.
[3] Payments expected to be about $200,000 in 1979–80.
[4] Mean not weighted by population.

ALEXANDRIA, VA.

General government in Alexandria which has carefully avoided increasing the number of its municipal employees, made no unemployment compensation payments in 1978. While the Alexandria school system appropriated $97,000 for such payments during fiscal 1979, actual payments may be lower.

BURLINGTON, VT.

In Burlington, the mandate increased the city's operating budget by about $35,000 annually. Most of the increase—$32,000—is attributable to claims payments, and $3,000 to $4,000 to administrative expenses in the school department. The law has also resulted in altered staffing patterns in the school department which do not necessarily meet Burlington's needs. In some cases the work week of part time staff has been limited to 18 hours so that employees do not qualify for the act's coverage. In addition, proposals have been advanced to start a summer school which would employ part time aides who would not be eligible for unemployment compensation.

CINCINNATI, OHIO

In Cincinnati, federally mandated unemployment compensation payments cost the city $500,000 in 1977 and $350,000 in 1978. These expenditures were the result of city budget cuts which led to the layoff of more than 450 city employees in 1976. The Director of the Cincinnati Office of Research Evaluation and Budget states that the most dramatic effect of this Federal mandate is that two city employees must now be laid off to obtain the cost savings which were formerly achieved by laying off a single employee.[21] Thus, the size of personnel lay-offs and service cutbacks has been exaggerated by the requirement that local governments pay unemployment compensation.

[21] For a limited time period, until benefits are exhausted.

DALLAS, TEX.

The city of Dallas, with a municipal work force of 13,762, paid out a total of $9,405 during the final quarter of 1978. The low figure is largely attributable to the strong economy of the Dallas-Ft. Worth metropolitan area and to the relative prosperity of the Dallas city government. (Early 1979 statistics show that the Dallas SMSA's unemployment rate is only 3.7 percent.) Unemployment compensation payments made by Dallas, unlike those made by Newark and Cincinnati, have not been occasioned by layoffs of city personnel. Rather, they derive primarily from claims made by workers who have been fired but who remain eligible for unemployment compensation payments under Texas law.[22] Dallas expects the current quarterly payment levels of approximately $10,000 to double over the upcoming year. However, the projected $80,000 annual outlay represents only a small share of the total city budget.

FAIRFAX COUNTY, VA.

Fairfax County, a growing suburban area, has been expanding its general government employment base. While the County allocated $358,000 in its 1978 budget for unemployment benefits, its actual compensation outlays, based on current experience, will be closer to $60,000 annually. The Fairfax County school system budgeted $400,000 for unemployment compensation in 1978 and $166,000 in 1979. Current compensation payments and employment trends indicate that total 1979 outlays will be $60,000, rising to $100,000 in fiscal 1980.

NEWARK, N.J.

Newark, like Cincinnati, has incurred some costs as a result of the 1976 unemployment compensation legislation. In 1978 the city paid $1.5 million into a trust fund to be used to reimburse the State for the amount of unemployment compensation paid to eligible employees of Newark. The Director of Finance for the city estimates that approximately $30,000 was paid in claims during the second half of 1978—the first period during which the city and state assumed responsibility for payment.

SEATTLE, WASH.

Unlike other jurisdictions in the sample, Seattle had instituted a self-insured compensation program prior to 1974 which provided limited unemployment payments to laid-off city employees. It is estimated that in 1978, the difference in cost between currently federally mandated compensation payments and the previous, city-initiated assistance was $225,000 a year. Further, since the city does not participate in the Federal Unemployment Assistance Act, Seattle is required to pay 100 percent of extended unemployment benefits.

---

[22] Vernon's Ann. Civ. St. Art. 5221b–2(b) (1977).

5. Bilingual Education

The Federal legislative mandate for bilingual education derives from three statutory sources—the Civil Rights Act of 1964, the Bilingual Education Act of 1968 and the 1974 Bilingual Education Act. The last elaborates the Federal mandate with the greatest specificity. It states that special provision must be made for the education of persons of limited English speaking ability. The act further declares it to be national policy to: (*a*) encourage bilingual education methods and techniques; and (*b*) encourage State and local education agencies to carry out such programs at the elementary and secondary levels. Current requirements direct school districts to develop a formal bilingual education program if 20 or more students in a school district speak the same foreign language. Local enrollments and outlays for the case study jurisdictions are presented in table 12.

TABLE 12.—INCREASED COST TO LOCAL GOVERNMENT OF BILINGUAL PROGRAMS, 1978

	Students	Percent of enrollment	Total local outlays (in thousands)	Per capita outlay
Alexandria [1]	[1] NA	NA	0	0
Burlington	0	0	0	0
Cincinnati	33	0.1	0	0
Dallas	NA	NA	$392	$.46
Fairfax County [2]	[3] 2,200	1.7	500	1.07
Newark (1979)	8,000	13.3	[3] 2,300	6.95
Seattle	1,475	3.0	1,000	2.05
Total/mean			4,192	1.50

[1] NA=Not available.
[2] Not strictly bilingual program.
[3] Fiscal 1979–80. State share $1,800,000.

To some degree, the 1974 Act codified judicially decreed bilingual education requirements which derived from litigation brought under title VI of the 1964 Civil Rights Act. Perhaps the most significant impetus to the creation of bilingual education programs was a 1974 U.S. Supreme Court decision, *Lau* v. *Nichols*.[23] In *Lau*, the Supreme Court held that the failure of the San Francisco schools to provide appropriate language instruction to 1,800 Chinese-American students denied them meaningful participation in the school program and thus violated title VI of the Civil Rights Act. The Court ordered the school district to rectify the situation but prescribed no specific remedy.

In an October 20, 1978, letter to the Chairman of the House Committee on Education and Labor, the Director of the Office of Civil Rights detailed the basic requirements imposed on local school systems by Federal bilingual education mandates. The letter states:

> A school district is responsible for assuring that national origin minority students receive equal educational opportunity. The first step in this process requires the district to identify all students with limited English language pro-

[23] *Lau* v. *Nichols*, 414 U.S. 563 (1974).

ficiency. The next step is the district's assessment of the ability of children in each of the languages. After undertaking these steps, the district then becomes responsible for providing necessary instructional services for students identified as limited English speaking.[24]

The letter further notes that the overriding legal obligation of local school systems is "to assure students are not kept from progressing in required subject areas because of their limited English speaking ability."

ALEXANDRIA

Alexandria does not offer bilingual education and has therefore incurred no additional costs as a result of the Federal mandate, despite the fact that the children in the system come from as many as 53 foreign language backgrounds. The predominant foreign groups in the schools are Hispanics and Asians who represent 3 and 5 percent of the overall population, respectively. School officials claim that the language needs of children in the system are fully met by the schools' English as a Second Language (ESL) program rather than by HEW's bilingual education options. The ESL program, which was instituted prior to the enactment of the relevant Federal mandates, emphasizes remedial language training for students who come from foreign language backgrounds.

ENROLLMENT BY RACE: ALEXANDRIA PUBLIC SCHOOLS, MARCH 1979

Race	Students	Percent of total
American Indian	8	0
Hispanic	396	3
Asian	520	5
Black	5,558	48
White	4,146	44

BURLINGTON, VT.

Burlington's school system enrolls virtually no students who would qualify for bilingual training, offers no such program and, as a result, expends no funds in this area. In fact, Burlington refused to fill out required Federal reporting forms on the number of non-English speaking students enrolled because the number of bilingual students in the system did not justify the administrative costs of meeting the reporting requirements.

CINCINNATI, OHIO

Out of a total enrollment of 62,996 in 1978, Cincinnati has very few foreign language speaking students enrolled in the public schools. In 1978, 18 Indo-Chinese students received supplementary language instruction supported by Federal funds under the Indo-Chinese Refugee Act. In addition, 15 students were identified as potentially bilingual as a result of a Civil Rights survey undertaken in 1978. However, only eight students are currently receiving tutorial language instruction. Unlike other cities which have large bilingual enrollments,

[24] cf. National Advisory Council on Bilingual Education, "Second Annual Report." Washington, D.C., 1976, pp. 14–16.

all funds used for this purpose in Cincinnati are provided by the Federal Government.

DALLAS, TEX.

Of the seven case study jurisdictions, Dallas has both the highest absolute number and highest proportion of students whose primary language is not English (see table 13). Reflecting a statewide pattern, a survey found that for 30.4 percent of all Texas families, English was not the household language. School enrollment data for the Dallas Independent School District (DISD) show that students of Hispanic origin made up 14 percent of the total school enrollment and 17 percent of the total elementary school population in 1976. By 1978, Hispanic students comprised 16 percent of all students and 19 percent of all elementary school students.

The degree to which bilingual education expenditures in Dallas are attributable to Federal mandates is difficult to assess as local programs have been developed in response to both State and Federal statutes and Federal judicial decrees. Nonetheless, the cost of providing bilingual education for the Dallas Independent School District for fiscal year 1979 will be approximately $1,475,000, or $69 for each Hispanic student enrolled in the system.[25] Federal funds make up approximately $500,000 of the total, while local funds represent approximately $635,000; the balance is funded by the State. It can be assumed that in the absence of Federal mandates, about 40 percent of current outlays would take place.

TABLE 13.—ENROLLMENT BY RACE: DALLAS INDEPENDENT SCHOOL DISTRICT, 1978

Race	Students	Percent of total
American Indian	475	0.4
Hispanic	21,392	16.0
Asian	794	.6
Black	64,281	49.0
White	45,119	34.0
Total	132,061	100.0

Source: Dallas Independent School District.

FAIRFAX COUNTY, VA.

Prior to 1974, Fairfax County offered only a small, unorganized language program for non-English speaking students. In academic 1974–75, owing to increasing numbers of foreign students (and concern no doubt over new Federal mandates), the county identified 1,100 foreign language students. The program has grown rapidly as more students from foreign countries have entered the Washington metropolitan area. Fairfax County currently enrolls 2,188 students in its language program, which represents languages from A (Amharic, official language of Ethiopia) to Z (Zulu) and includes Marathi (spoken in eastern Bombay state), Tagalog (the Philippines), and Gujurati (Northwestern India). Of the 50 or so languages, 15 are spoken by 20 or more students, thus requiring the development of a formal bilingual program (see table 14).

[25] Not all Hispanic students, however, require bilingual education.

TABLE 14.—*Languages in Fairfax County Schools, February 1979*

	Number		*Number*
20 or more students:[1]		10 to 19 students:	
Korean	532	Dutch	19
Vietnamese	396	French	16
Spanish	419	Italian	12
Chinese	118	Punjabi	10
Persian/Farsi	98	Tagalog	16
Arabic	98	German	12
Laotian	49	Filipino	10
Urdu	65	Turkish	16
Cambodian	25	All other	102
Greek	32		
Japanese	41	Total	2,188
Hindi	30		
Thai	26		
Swahili	25		
Portuguese	21		

[1] If 20 or more students speak the same foreign language, the school district, following HEW guidelines, has to develop a detailed formal education program.

Source: Fairfax County public schools, division of curriculum services.

As in Alexandria, the Fairfax program is based on the ESL approach rather than on any one of the three options required by HEW. In comparison to the HEW options, the county considers the ESL approach both preferable educationally and more cost efficient. The HEW mandated options, however, require that the programs rely upon the student's native language and cultural factors in instruction. Given the array of languages and lack of skilled instructors, it may not be feasible to meet most bilingual education requirements.

The direct cost of the current program is over $1 million. An additional $360,000 in indirect costs produces a total outlay of $1.4 million, an amount substantially less than would be needed to meet HEW guidelines. The incremental costs to the county are made up of legal fees, and administrative and recordkeeping costs arising from the county's efforts to document the adequacy of its ESL program. The documentation has been undertaken in response to threats by HEW to find the county in noncompliance for failure to implement a bilingual education program. Should HEW find the county to be in noncompliance all Federal aid to education in Fairfax could be cut off. While the actual incremental cost is difficult to isolate, it is approximately $500,000 in direct and indirect costs.

NEWARK, N.J.

Newark's bilingual education program is mandated by both Federal and State law. While 22 language categories have been identified for inclusion in the program, bilingual services are currently being provided in only three languages: Spanish, Portuguese and Italian. In the academic year 1979–80, the local share of bilingual education was $2.3 million; State share, $1.8 million; and Federal grants, $508,-000. The per pupil cost for the approximately 8,000 students in the program is $575, with most funds allocated for 215 teachers, 121 of whom teach in the native language.

SEATTLE, WASH.

The Seattle School District is currently providing bilingual programs to 1,475 students with backgrounds in 30 languages. The annual local cost of the program (1979) is $1 million, with the State government providing an additional $0.5 million. In addition to State and local funds, the Federal Government has provided Seattle with grants which account for only a small percentage of total bilingual education outlays.

6. Education of the Handicapped

Public school enrollment in special education for the handicapped increased nationally from 1.5 million in 1963 to 2.7 million in 1970–1971. However, a survey undertaken by HEW in 1970 estimated that there were almost 4.8 million handicapped in public schools, indicating that less than 60 percent of potential students was receiving special education.

One of the primary effects of recent Federal legislation mandating standards for the education of the handicapped has been a large increase in the number of handicapped identified and served over the last several years.[26] It is therefore likely that the percentages referred to above are currently significantly understated.

Several requirements form the core mandates for the Nation's public school systems. Among other provisions they require that:

(1) Handicapped persons, regardless of the nature or severity of their handicap, are to be provided a free, appropriate public education. Public school systems which receive Federal aid must either educate handicapped children in their regular programs or provide an appropriate alternative education at the public expense;

(2) Handicapped students are to be educated with nonhandicapped students to the maximum extent appropriate to their needs;

(3) School districts are to identify all unserved handicapped children;

(4) Evaluation procedures are to be improved in order to avoid the educational problems resulting from the misclassification of students; and

(5) Procedural safeguards are to be established to enable parents and guardians to influence decisions regarding the evaluation and placement of their children.

A Federal assistance program has been created to aid the financing of education for the handicapped. To be eligible for Federal funds, States must submit plans indicating that a free and appropriate education will be available for all handicapped children within the State. A second important Federal requirement—that educational facilities be made accessible to the handicapped—is addressed most directly in the Rehabilitation Act of 1973. The act dictates that programs ad-

[26] The Education of the Handicapped Act, Public Law 93–380 (1974) as amended, and section 504 of the Rehabilitation Act of 1973. In addition, several judicial opinions since 1971 have shaped educational requirements for the handicapped.

ministered by recipients of Federal funds "must be readily assessible to and usable by handicapped persons."

Local operating and capital outlays as well as enrollment data are shown in table 15.

TABLE 15.—LOCAL COST OF EDUCATING THE HANDICAPPED ATTRIBUTABLE TO FEDERAL MANDATES

	Percent handicapped of total enrollment	Total outlays (in thousands)		Per capita outlays	
		Operating	Capital	Operating	Capita
Alexandria	12.0	$1,000	$25	$9.26	$0.24
Burlington	9.7	65	[1] 700	1.67	17.95
Cincinnati	5.1	0	0	0	0
Dallas	7.0	1,628	[1] 2,500	1.92	2.95
Farfax County	10.0	12,080	[2] 1,197	22.58	2.24
Newark	10.0	2,150	[1] 1,700	6.32	5.00
Seattle	5.5	1,800	[1] 200	3.67	.41
Total/average	8.3	18,723	6,322	6.49	4.11

[1] Cumulative.
[2] Debt service only.

ALEXANDRIA, VA.

The proportion of Alexandria students defined as handicapped increased from 5 percent in fiscal 1972 to 12 percent in fiscal 1979. The absolute number of handicapped students increased by 32 percent. While an indeterminate share of the increased enrollment is attributable to Federal mandates and improvements in screening, the steepest cost increases derive specifically from Federal requirements that, as of academic year 1978–79, direct school systems to cover fully the tuition costs for students enrolled in specialized facilities. Prior to that year, Virginia schools, in accordance with State law, were obligated to pay no more than $1,200 per child for students enrolled in specialized facilities. The balance of the cost was assumed by parents or social service agencies. As a result of Federal mandates, however, costs increased from $790,000 in fiscal 1978 to $1.25 million in fiscal 1979 and to a projected $1.7 million in fiscal 1980. The number of special contract students increased from 46 in 1973 to 209 in 1979. The current average annual cost per student is $6,000.[27] The combination of more students, partially attributable to mandates, and Alexandria's full absorption of tuition costs resulted in a net incremental increase of about $1 million. Capital outlays totalled $25,000.

BURLINGTON, VT.

In response to the handicapped education mandates, the Burlington School District recently hired 12 teachers and 10 aides at a total cost of $200,000. Twenty-five percent of the costs, or $50,000, was paid for by the city, while the remainder was paid for by the State. In addition, school officials estimate that another $5,000 in administrative services and $10,000 in teacher in-service time are attributable to the statutory mandates.

[27] The average direct and indirect cost per student in centers/schools within Alexandria was $4,581 in 1977 and $6,635 in 1978, in self-contained classes, $2,677 in 1977 and $3,702 in 1978.

To provide accessibility to school buildings for the handicapped, capital costs of $50,000 were incurred for improvements to three existing facilities. In two instances, the costs were reimbursed by the Federal Government. More recently, a committee of both State and local officials identified a set of improvements which must be made by the Burlington School District by 1980 to meet Federal accessibility requirements. The estimated cost of the improvements was $1 million—an amount equal to the entire State allocation from the Federal Government to help subsidize such construction. City officials claim that Burlington cannot afford the funds necessary to satisfy Federal requirements under the current 30 percent State—70 percent local match formula. Consequently, the city expects to make improvements only as demanded by handicapped pupils or to send students to other schools with appropriate facilities. Handicapped students in Burlington represent almost 10 percent of the total enrollment.

CINCINNATI, OHIO

Federal handicapped education mandates have produced an interesting effect on special education expenditures in Cincinnati. Rather than compelling increased expenditures, they have impeded local spending for the construction of several new facilities for the education of the handicapped. In 1975, the Cincinnati Board of Education passed a resolution to issue $12 million in bonds for the exclusive purpose of providing centralized facilities for the mentally and physically handicapped and, in 1978, $6 million in school bond anticipation notes were sold. However, construction of the facilities has not yet commenced. Federal standards that require handicapped persons to be educated with nonhandicapped persons "to the maximum extent appropriate to the needs of the handicapped person in question" appear to prohibit plans previously contemplated by the school board. As a result, earlier designs for the centralized facilities have had to be redrawn with a greater emphasis placed on the modification of existing schools. School officials state that the impetus for the original projects cannot be traced to Federal handicapped education mandates.

Since Cincinnati has been a center for education of the handicapped since the 1940s, Federal mandates do not appear to have resulted in any net cost increases in instruction and administration. Instead, Federal legislation has produced a net savings for the system as Federal and State aid for special education has increased significantly over the past several years and the controlling statutes have grown increasingly stringent.

The handicapped student enrollment in the Cincinnati public schools has grown only slightly in recent years—from 3,180 in fiscal 1974 to 3,217 in fiscal 1978—an increase of only 37 pupils. Handicapped children accounted for 4.3 percent of the total school enrollment in fiscal 1974 and 5.1 percent in fiscal 1978. Given the higher levels of Federal and State aid now available, the traditionally high level of commitment to handicapped education on the part of the school board, and the minimal increase in handicapped enrollment, Cincinnati has been able to contain operating costs for education of the handicapped.

DALLAS, TEX.

The current handicapped population of the Dallas schools is approximately 10,000 students, representing about 7 percent of the total school district population. Given the current rise in the number of handicapped students and the overall decline in school enrollment, official project that within 2 years handicapped pupils could represent 10 to 12 percent of the total school population.

Unlike Cincinnati, Dallas has experienced a sudden rise in handicapped enrollment. This increase is attributable to three factors: the increased publicity surrounding the availability of free public education for all children; an intensive program undertaken by the Dallas Independent School District to identify handicapped children not previously enrolled in school; and Federal mainstreaming requirements which have resulted in the transfer for handicapped children from full-time residential institutions to public schools.

In fiscal 1979, the school district will spend approximately $3.4 million for institutional and support services for the handicapped as a direct result of Federal mandates. Of this total, State funds constitute approximately 40 percent of the total, or about $1.4 million, with direct Federal grants representing an additional $371,000. As of April 1, 1979, the district had spent almost $2 million on capital construction to satisfy the decentralization and accessibility requirements of Federal laws. The total cost of building modifications is estimated to be $2.5 million.

FAIRFAX COUNTY, VA.

The Fairfax County school system is one of the largest in the Nation, with an enrollment of about 130,000 students. The number of handicapped in all categories has grown substantially since 1972. However, the number of teachers providing services to the handicapped has grown even more rapidly. Expenditures reflect this growth: direct outlays for special education in fiscal 1972 were $4.3 million, or 3.1 percent of the county's operating budget, and increased to $23.1 million, or 8.9 percent of the budget in fiscal 1979 (see table 16). The number of personnel assigned to aid handicapped students will have risen from 3.5 percent of all county teachers in 1973 to 10.1 percent in 1979. By contrast, outlays for handicapped as a percentage of total operating outlays have remained stable at about 2.6 percent between 1965 and 1971.

TABLE 16.—SPECIAL EDUCATION IN FAIRFAX COUNTY—SELECTED CHARACTERISTICS, 1973–79

Item	1973	1979	Percent change, 1973–79
Number of special education students [1]	1,875	3,324	77.3
Special education as percent of total enrollment	1.38	2.57	86.2
Total direct special education outlays (in thousands)	$4,349	$23,097	432.1
Special education as percent of total operating outlays	3.1	8.9	187.1
Number of special education teachers	370	1,207	226.2
Special education teachers as percent of all teachers	3.5	10.1	188.6

[1] Enrollment in county special schools only. Excludes self-contained classes and students mainstreamed. The total number of handicapped students was 12,800 in 1978 and is projected to rise to 14,900 or 11.8 percent of total enrollment in 1980.

Based on the rate of change in direct outlays for handicapped students as a share of the total operating budget prior to 1973, outlays should equal 4.3 percent of the total in 1979. The differente between outlays based on this percentage and the actual outlay of 8.9 percent is $11.9 million, an amount which has been attributed to Federal mandates by State officials. The local share of these added costs was about $8.9 million in 1978. In addition, the local share of indirect outlays (employee benefits, administration, plant operation) added another $3.2 million, for a total incremental cost of $12.1 million.

NEWARK, N.J.

Federal handicapped education standards will cost Newark an additional $2.15 million in operating revenues during the 1978–79 school year. Total handicapped/special education expenditures for the year will be $14 million, with most funds provided by the State. Despite the high costs, school officials state that only two-thirds of all eligible children within the school district are being served. The failure to locate and identify the other eligible students is attributable to three factors: (1) Overcrowded facilities for the handicapped; (2) insufficient numbers of child guidance personnel; and (3) the difficulties in isolating the handicapped child within a school population which is generally underachieving. Newark currently enrolls 6,000 students in handicapped programs and estimates that an additional 3,000 students have not yet been identified.

The high additional costs imposed on the Newark system are somewhat surprising given the fact that State legislation has set relatively high educational standards for the handicapped over the past 10 years. Components of the added $2.15 million are as follows:

35 additional psychologists and learning consultants	$770,000
30 new special education teachers	495,000
2 new administrators	60,000
Materials	30,000
Additional costs of contracting for residential facilities	550,000
Additional transportation costs	250,000
Total	2,155,000

In addition, the capital costs of making schools accessible to the handicapped have resulted in $1.7 million in city outlays to date.

SEATTLE, WASH.

As in many large cities, Seattle restricted its provision of facilities for the education of the handicapped to several designated schools. New Federal mandates, requiring that support services for the handicapped be available at all facilities, have required $200,000 in capital outlays. However, little public support for any capital improvements can be found as overall enrollment in the schools has declined by half in recent years—from 100,000 students in 1968 to 50,000 in 1979.

Federal handicapped education mandates have also resulted in the expenditure of $1.8 million in city funds, primarily for teacher compensation. Because special education teachers in Seattle are only allowed to instruct handicapped pupils, one of the problematic results

of Federal mandates has been a sharp increase in the number of specialized teachers who cannot be used in regular classrooms—despite the obligation to mainstream the handicapped. Thus, the increase in teachers in Seattle has not effectively reduced the pupil-teacher ratio which is established by contract at 29–1.

7. *Access to Public Transportation for the Elderly and Handicapped*

The current requirement that public mass transportation facilities be made accessible to the elderly and the handicapped derives primarily from the Rehabilitation Act of 1973. Recently promulgated regulations implementing the Act as it pertains to public transportation require that governments which receive Federal aid for transportation must, within 10 years, make at least half of their buses used in peak service wheelchair accessible. The rule also requires that "these buses must be utilized before inaccessible buses during offpeak hours so as to maximize the number of accessible buses in service." [28]

This inquiry focuses on costs local governments are presently incurring as a result of Federal requirements for urban bus services. In several cities, however, cost estimates associated with fixed facilities and commuter rail transit have been included.[29]

TABLE 17.—ADDITIONAL LOCAL COST OF PROVIDING ACCESS TO PUBLIC TRANSPORTATION FOR THE HANDICAPPED

	Total (in thousands)		Per capita	
	Operating	Capital	Operating	Capital
Alexandria	NA	$102	NA	$0.97
Burlington	$43	0	$1.10	0
Cincinnati	400	[1] 7,140	.91	16.17
Dallas	0	14	0	.02
Fairfax County	NA	1,700	NA	3.09
Newark	NA	[2] 2,960	NA	[1] 8.84
Seattle [3]	100	150	.31	.21
Total	543	12,066		

[1] Fiscal 1979. Cumulative 1979–83, projected cost for Cincinnati (85 percent of total) is $17,900,000.
[2] Potential cost to Newark will depend on Essex County decision. The above data assume the city will pay 5 percent of total cost, the State 15 percent.
[3] Includes only bus fleet cost.

ALEXANDRIA, VA.

As of April 1, 1979 Alexandria provided no public door to door transportation for the elderly and handicapped. However, 130 lift-equipped buses have recently been purchased by the Washington Metropolitan Area Transit Authority (WMATA); $9,500 of the overall cost of each bus is attributable to the added design features which respond to the needs of the elderly and handicapped. As Alexandria's percentage of total bus miles of the WMATA system is 8.3 percent, the city's share of the increased costs of providing these features can be estimated as approximately $102,000. Service in Alexandria commenced in June 1979.

[28] 44 Federal Register 31442 (May 31, 1979).
[29] The costs documented in this paper are attributable to regulations promulgated by the Department of Transportation in April of 1976. Those regulations required that local governments "make special efforts" to, among other things, purchase or equip buses in service "so that one-half of a jurisdiction's fleet is wheelchair accessible." (41 Federal Register 18234 (Apr. 30, 1976).

BURLINGTON, VT.

In April 1979, responsibility for federally funded transportation of the elderly and handicapped was assumed by the regional transportation authority. No local or regional capital outlays are projected. However, the cost of operating the program with local funds, based on the first month of operation, was somewhat higher than projected. Outlays in April were $4,270, or $1.14 per pessenger mile, offset by $675 in revenue from fares. On an annual basis, the net local cost will be about $43,000.

CINCINNATI, OHIO

The public transportation system for Cincinnati and surrounding Hamilton County is administered by the Southwest Ohio Regional Transit Authority (SORTA).[30] To date, the system's efforts to meet the needs of its handicapped population have concentrated on providing door to door service for wheelchair users and the semiambulatory. All current activities and costs are wholly and attributable to the impact of Federal regulations. The cost of operating door to door service was $400,000 for fiscal 1978, with the same amount budgeted for service in fiscal 1979.

The Research Department of SORTA has projected costs for the purchase of two types of buses which incorporate differing standards designed to accommodate the elderly and the handicapped. The advanced design bus simply includes a "kneeling" feature which lowers the chassis to ease entry for the elderly and handicapped and several other minor design changes. The federally designated Transbus features a lower floor, tandem axles and a lift. The additional costs to be incurred by the purchase of new buses (costs which exceed those of buses currently available) are estimated to be as follows:

[In millions]

	1979	1980-83
Advanced design	$1.1	$1.6
Transbus	7.3	11.1
Transbus operating	.4	2.3
Total	8.8	15.0

The costs shown above would typically be met by a Federal appropriation to cover 80 percent of capital costs, while State government would contribute an additional 10 percent. However, all available Federal funds in Cincinnati are currently allocated to build garages, leaving no funds for the purchase of buses. The State contribution toward operating outlays is offset by the cost of meeting State requirements which exceed Federal standards.

DALLAS, TEX.

The Dallas transit system has responded to Federal mandates by instituting door to door service in small, lift-equipped buses. The service, which went into operation on December 18, 1978, is only available

[30] About 85 percent of the system's services is provided within the boundaries of the city of Cincinnati.

to those city residents who use wheelchairs, are semiambulatory, or are mentally retarded. Twelve buses were purchased by Dallas Transit at a total cost of $204,000. Eighty percent of the cost was borne by the Federal Government; 13 percent by the State; and 7 percent or $14,000 by the city. Five additional buses had been ordered as of April 1, 1979. At this time, city transit officials do not expect to have to pay operating costs as the buses have been leased to a private firm which will attempt to make a profit from the return on fares. Dallas Transit has also purchased 120 advanced design buses which are equipped with "kneeling" devices to use on its line haul routes.

FAIRFAX COUNTY, VA.

Public transportation in Fairfax County is provided by the Washington Metropolitan Transit Authority (WMATA). WMATA has purchased 130 lift-equipped buses with service for the elderly and handicapped which began in June 1979. The cost of this service to Fairfax County will be based on the county's share of total weekday revenue miles of the WMATA system. Since the service is new, the total cost to the county is not yet known. However, the maintenance costs, based on preliminary data, are anticipated to be high. Fairfax County's share of buses purchased by WMATA is about $200,000 and the cost of Metrorail access is $1.5 million, with an annual debt service of $135,000.

NEWARK, N.J.

Public transportation in Newark is part of a tri-State regional system which includes New York and Connecticut. Data shown below for the city are based on average unit costs derived for the total regional system in meeting section 504 interim regulations.

Facility	Facilities—Tristate region [1]		Unit cost (in thousands)	Facilities—Newark only [1]	
	Units	Total cost (millions of 1978 dollars)		Units	Total cost (millions of 1978 dollars)
Bus retrofit	1,108 buses	$11.6	$10	650 buses	$6.5
Commuter rail	149 stations	418.5	2,808	4 stations	11.2
Lifts	512 units [2]	63.2	176	14 lifts	2.4
Newark subway system	11 stations-elevators	39.0		11 stations	39.1
Total					59.2

[1] Based on correspondence between the New Jersey State Department of Transportation and the Urban Institute.
[2] For commuter rail.

In New Jersey, the state is the recipient of all UMTA funds and requires, for some projects, a 25 percent local match. While either cities or counties may provide the required matching funds, it is not known what course Newark will follow. If Newark (rather than Essex County) were allocated a pro-rata share of the $59.2 million total, its contribution would be 25 percent of the non-Federal share, or $2.96 million.

SEATTLE, WASH.

Seattle's Metro Transit is currently testing a lift, designed by its own engineers, which should cost approximately $7,000 per bus to in-

stall. Early cost forecasts indicate that outfitting half the system's 600 bus fleet with the lift would cost approximately $250,000 per year, including operating and amortized capital costs. In addition, officials state that 50 percent of the total bus stops will eventually be made accessible. As an adjunct to the accessibility program, Metro Transit has arranged a subsidized taxi service for the elderly and the handicapped. Specially designed vans have been purchased to insure that the taxi system is accessible to the handicapped. Cab fares are discounted 40 percent for the target population.

Seattle officials state that a neighboring community has recently purchased five specially equipped buses for its system without Federal money. The jurisdiction was able to purchase buses in Canada which met Federal guidelines at a cost one-third below the outlay necessary to install the required equipment.

III. Federally Induced Local Expenditures

1. *Background*

Despite the growth of Federal aid during the last 5 years, a number of Federal programs have been reduced or eliminated. Explanations for these changes in Federal assistance programs vary. Some grants of limited duration were given on the premise that they would support innovative pilot projects which would typically not be undertaken by local governments because of the uncertainty of success. Thus, aid was given as "seed money" to initiate innovative efforts which, if successful, would then be funded from local revenues. Other programs were initiated to respond to a particularly critical or politically visible problem. The High Impact Anti-Crime program, for example, which granted substantial Federal funds to eight large cities, was enacted in response to high crime rates and the political visibility of law enforcement issues at the time.

Adverse economic conditions spurred the development of Federal assistance programs designed specifically to be countercyclical. During a national economic downturn, certain areas and population segments are particularly hard hit. To offset these conditions, three major programs were initiated in the mid-1970's—the Comprehensive Employment and Training Act (CETA), Anti-Recession Fiscal Assistance, and the Emergency Local Public Works Program (LPW). As economic conditions improved, assistance provided under these programs was reduced or eliminated.

In some instances, Federal programs are not eliminated but are, rather, retargeted or redefined. By redrawing entitlement guidelines, amended regulations may prohibit particular uses of funds previously permitted. The recently modified Community Development Block Grant Act, for example, now requires that cities focus all their CD-supported social service funds in specific neighborhoods. As a consequence, social agencies not located in these designated neighborhoods are no longer eligible for HUD funds. Thus, if local governments wish to continue programs in other neighborhoods, they must replace CD funds with State or locally collected dollars.

Another category of induced local expenditures are Federal actions which require rising local contributions. Added Federal Insurance

Contributions Act costs are an example of such an activity, as briefly discussed in this section.

2. The Comprehensive Employment and Training Act (CETA)

The Comprehensive Employment and Training Act (CETA) provides public service jobs for the unemployed. Several grant programs are authorized under the act, the largest of which are the manpower block grant program, authorized under title I, and the public services employment programs, authorized under titles II and IV.

Under regulations recently ordered by Congress, a large proportion of CETA employees currently enrolled in the program will be terminated on or before October 1, 1979. The new CETA regulations reflect the decline in the national unemployment level, evidence that some local governments improperly used the program to provide essential city services, and concern that the program has failed to provide jobs for the hard core unemployed. The new regulations are intended to insure that CETA positions, particularly those within local government, are to be entry level positions and are not to replace regular city jobs.

The regulations mandate several new policy directions for the program. The eligibility requirements have been amended to limit an individual's participation in the program to 1½ years. Local governments' previous unlimited ability to supplement Federal wages has been restricted, and the percentage of the local Public Service Employee (PSE) work force which can receive supplemental salaries has been reduced. These changes in some instances reduce the incentive of local governments to hire and train employees under CETA.

The new regulations have had differing impacts on the case study jurisdictions. Variation can be ascribed to a number of factors, among them:

Strength of the local economy;
Number of CETA employees in essential municipal positions;
Wage levels of the municipal government; and
Fiscal strength of the municipal government.

ALEXANDRIA, VA.

Restrictions placed on local government's supplements to Federal wages mean that more PSE employees will be drawn from the hardcore unemployed. The shift in regulations, coupled with the abbreviated term of employment, will limit the city's participation in the program. Fourteen CETA workers of the 91 currently employed by Alexandria must necessarily be laid off before October 1, 1979. City officials believe that 13 of the 14 to be terminated will be placed on the city payroll. However, none of the positions into which these CETA employees will be placed was created as a result of modifications in the program.

BURLINGTON, VT.

Burlington, in contrast to other cities included in the sample, currently employs only three CETA workers, considerably below the 40 to 50 employed several years ago. The Burlington Personnel Director

stated that the agency designated as the prime sponsor has been difficult to work with and the enforcement of CETA guidelines has been too strict. As a result, the city has rejected most of its CETA employee entitlement and will, therefore, be basically unaffected by the amended regulations.

CINCINNATI, OHIO

As of February 1, 1979, the CETA program administered by the city of Cincinnati enrolled 1,765 people. Of this total, 562 were employed within city agencies, which account for 10 percent of the total city work force. An additional 1,203 were employed in outstation positions in nonprofit organizations. Under regulations recently enacted by the Congress, up to 90 percent of the city's CETA employees will be terminated as of October 1, 1979.

Despite efforts by city officials to the contrary, budgetary constraints have led several city departments to become dependent on CETA employees in recent years. The departments which have become most reliant on these workers are: waste collection, police communications and records, parks and recreation maintenance, and highway maintenance. Officials in the city's Public Works Department estimate that the termination of all CETA employees would reduce the number of laborers in highway maintenance by 60 percent.

One of the largest costs borne by Cincinnati as a result of the new regulations will be the assumption of the salaries of employees previously paid by the Federal Government. As of February 1, 1979, city officials estimated that 100–200 CETA employees would be transferred to the city payroll by October 1, 1979. Among the less direct costs which the rewritten regulations will impose on the city are the cost of training new CETA enrollees every 18 months. In addition, the limits set on CETA salaries reduce the number of PSE's who can be assigned to positions other than low skill entry level jobs. Accordingly, the duration of PSE employment will be reduced, the frequency of training will be increased, and the number of positions in which PSE's can be employed will be limited.

The new regulations are not the only source of the city's woes with regard to the administration of CETA. Cincinnati is currently faced with a large cutback—from $22.5 million to $16 million—in CETA funding which will force further PSE layoffs. One city official expressed his frustration regarding the political and fiscal costs of administering the CETA program:

> The new regulations have been drafted in a way which doesn't permit public service employees to address the needs of the city. Why should we run the program for them when we are the ones who have to take the heat from the PSE's when we lay them off in 18 months? The people think that they work for the city while in fact the city has almost no say over the rules which govern their employment. My own feeling is to let the feds run the program.[n]

DALLAS, TEX.

The new CETA regulations will have virtually no effect on the operation of the Dallas city government. Of the 1,151 CETA slots authorized for Dallas under titles II and VI, 92 are within city govern-

---

[n] Interview with James Buckalew, Director of Management Service Operation Group, the Office of Research Evaluation and Budget, the city of Cincinnati, Jan. 31, 1979.

ment. The remainder are assigned to private nonprofit organizations which deliver a wide range of services. Further, in October 1978, prior to the issuance of the redrafted Federal eligibility guidelines, the Dallas City Council voted to limit enrollment in CETA to 1 year—6 months less than the Federal maximum. Thus, new Federal limits on an employee's tenure within CETA will have no impact on city policy.

For two reasons, the termination of CETA employees does not pose the same political problems in Dallas as it does in cities such as Cincinnati and Newark. First, the city has 300 to 500 job vacancies at all times, according to the Assistant City Manager. Thus, it can find jobs for almost any employee simply by taking advantage of attrition within the city work force. Second, the economy of the Dallas region is so healthy, with current unemployment rates estimated to be 3.7 percent, that almost all employees can find work in the private sector.

Recent restrictions on the amount local governments can subsidize CETA wages will produce no effect on the Dallas government as city wage scales are far lower than is the case in the other case study jurisdictions. Less than two percent of the city's CETA enrollees currently receive any supplement.

While the city government may be immune from the effects of the new CETA regulations, the city's private non-profit organizations, which are called upon to deliver many essential social services, will feel the impact more severely. These organizations, which employ over two-thirds of Dallas' authorized CETA personnel, have relied on public service workers to perform day to day administrative and support tasks. Because these agencies have depended on CETA employees to provide program continuity, they have supplemented enrollees' salaries more substantially than the city and will be unable to replace outgoing employees with new PSE's at the same professional level.

FAIRFAX COUNTY, VA.

As of June 1979, Fairfax County was allotted 245 CETA positions of which 190 have been filled. The differential is attributable to the recently imposed wage supplement limitations and the length of administrative time required to determine client eligibility for CETA employment. In order to meet its quota of PSE's, the county is actively soliciting new CETA positions. In addition, the county, in response to the shortened term of employment currently permitted, has increased the level of training available to PSE's and, in so doing, is providing clients with skills that will give them access to the private sector.

NEWARK, N.J.

The recently enacted changes in CETA regulations will produce a dramatic effect on Newark. Up to 25 percent of the city's non-uniformed work force could be affected by the regulations which set an 18-month limit on the tenure of a CETA employee.[32] City officials state that the most serious problem arising from the layoffs will be the financial burden of training substitute employees for short-term employment. Restrictions on CETA employees' salaries will make it

---

[32] The total city work force is made up of 5,000 employees, 2,000 of whom are in the police or fire department. The number of CETA employees working for the city on Feb. 1, 1979, was 1,300, 70 percent of whom will have to be laid off on Sept. 30, 1979.

virtually impossible to find people who can quickly develop the skills of the outgoing employees, many of whom have been in their current position for 3 or 4 years. Officials are concerned that the new regulations will lead to a substantial decline in the quality of the workforce in the city.

The city's administration of the CETA program has resulted in two other types of cost. The first resulted from the 100-percent payback provision of the regulations which requires that improperly spent CETA funds be returned to the Federal Government from city funds. To date, the Department of Labor has required the city to pay penalties amounting to almost $1.4 million, including close to a $1 million repayment for the city's failure to comply with CETA's "maintenance of effort" provisions. These provisions prohibit CETA employees from filling positions that would otherwise be filled by a municipal employee.

The other costs which CETA imposes on the city derives from the requirement that CETA staff be released before municipal employees in comparable positions can be laid off. The Director of the Budget Office for Newark stated that the selection of personnel to be laid off has in the past been partially determined by the number of CETA employees who would have to be released. Thus, the calculation of costs associated with lay-offs of city personnel takes into consideration the relative numbers of CETA employees who would be lost—skewing local personnel decisions and imposing efficiency costs on city governments.

SEATTLE, WASH.

Seattle's 500 CETA employees currently constitute about 5 percent of the total municipal work force. In the past, Seattle emphasized useful public service as its main program objective in administering CETA, particularly since a high unemployment rate enabled the city to hire highly skilled workers. The recent growth of private sector employment, however, has produced a shift in program emphasis to less skilled workers. General economic conditions and the 18 month maximum term of employment have reduced the city's incentive to train its CETA employees. In addition, the problem is complicated by low level CETA wages which are only slightly higher than the lowest paying city positions. Nonetheless, the city contributes about $500 thousand per year to the city's retirement fund for CETA workers.

3. The High Impact Anti-Crime Program

The High Impact Anti-Crime program, enacted in January 1972, provided approximately $20 million over a 4-year period to each of eight American cities selected on the basis of crime rate, city size and geographic region. The impact program's goal was to reduce serious crime by 5 percent in 2 years and by 20 percent in 5 years through the introduction of innovative crime control strategies. Two of the six jurisdictions examined in our study, Newark and Dallas, were selected to be among the eight cities to receive High Impact Anti-Crime Aid.

NEWARK, N.J.

At the time of the first grant award in 1972, Newark reported the highest crime rate of any major city in the country. Over the course of the next 4 years, the city received approximately $17.8 million in Impact Aid which was used to set up 27 separate criminal justice

programs. Because of the serious fiscal problems which plagued Newark both during and after the period of Impact Aid, relatively few of the 27 programs which were begun with the $17.8 million survived. Of the 12 programs that are still operating, none is currently funded at a level comparable to that enjoyed during the years of generous Federal subsidies. Figures obtained from the Newark Office of Criminal Justice Planning reveal that $250,000 of the current city operating budget is used to support programs initiated under Impact Aid.

It is clear Newark could not have been expected to allocate city funds to replace any significant percentage of the $4–$5 million received annually under the Impact program. City officials state that the goal of subsidizing the start-up costs for innovative programs, eventually intended to be institutionalized with support from city funds, is wholly unrealistic when applied to cities with fiscal conditions similar to those in Newark.

DALLAS, TEX.

The fate of programs begun with the $18.7 million in Impact Aid received by Dallas offers a dramatic contrast to Newark's experience. The city of Dallas [33] used most of its Impact Aid allocation to hire new personnel in a number of areas of criminal justice administration: 100 new officers for the Tactical Division of its police force, six attorneys to provide legal assistance to the department, 33 new officers for use in investigations of more serious "impact" crimes, and 12 counselors for a youth services bureau. All positions were retained even after the termination of Federal assistance. The annual cost of each of Dallas's programs noted above was approximately $3.1 million.

4. Federal Insurance Contributions Act (FICA)

The programs which constitute the Nation's Social Security system include the old age, survivors. and disability insurance program and hospital insurance benefits programs which the Federal Government offers to the aged and disabled. Both programs are financed largely from taxes paid by employees and employers into the Social Security system under provisions of the Federal Insurance Contributions Act. With the exception of a few nonprofit organizations, payment of FICA taxes for private sector employers and employees is mandatory.

Unlike private sector employees and employers, State and Federal Governments and their workers have the option of participating in the Federal insurance program. As such, FICA cannot be strictly classified as a Federal mandate. Indeed, a State voluntarily may enter into an agreement with the Secretary of the Department of Health, Education, and Welfare to make benefits available to State employees as well as to employees of the political subdivisions of the state. Further, states entering into agreements do not have to extend coverage to all such employees if workers are already enrolled in a private or locally administered retirement plan. The only State and local employees now legally required to participate in the Federal social security program are those employed by a transportation system acquired for public ownership after 1950. Concern that the Federal Government will extend required participation in the social security

---

[33] Federal Impact Aid funds were divided between the governments of the city of Dallas and Dallas County.

system to all public employees dictates that FICA be termed a "potential mandate."

Because of steep increases in social security taxes (FICA) over the past two decades (from 2.25 percent in 1957 to 6.13 percent in 1979), some local governments which first participated when rates were far lower than their current and prospective levels have recently considered opting out of the Federal retirement system and establishing their own local pension plans. However, for any State to terminate coverage for an individual group, 2 years advance notice of intention to terminate must be given and coverage must have been in effect for at least 5 years at the time notice is given.

Officials in San Diego County, Calif., for example, claim that converting from participation in the Federal plan to a roughly comparable locally administered pension plan could save the city and its employees between $6.1 and $7.9 million per year. The County Board of Supervisors voted to notify the state Public Employee Retirement System of its intention to drop out of the program effective December 31, 1980. The final decision will be made at an election among city employees planned to take place in the next few months. Fairfax County has also been considering terminating FICA. Two of the seven case study cities—Dallas and Cincinnati—chose not to participate in the Federal plan, while the other five opted for participation (see table 18).

TABLE 18.—FICA TAXES PAID BY CASE STUDY CITIES

City	Year	FICA total (in thousands)
Alexandria	Fiscal year 1978	$1,472.
Burlington	Fiscal year 1978	$307.
Cincinnati		Exempt.
Dallas		Exempt.
Fairfax County	Fiscal year 1978	$3,774.
Newark		Not received.
Seattle		Not received.

5. *The Community Development Block Grant (CDBG)*

Federal regulations [34] issued after the enactment of the Housing and Community Development Act of 1977 [35] imposed new restrictions on the extent to which Community Development Block Grant funds could be used to support public services. The regulations required that public services supported by CDBG funds be provided exclusively to residents of Neighborhood Strategy Areas (NSA), that is, neighborhoods where CDBG financed construction and rehabilitation activities were being conducted in a concentrated manner. Thus, CDBG regulations mandated for the first time that social service programs be tied to physical improvement efforts.

NEWARK, N.J.

The new regulations have already played a critical role in forcing the closing of seven drug treatment centers in Newark. The centers, which had previously depended heavily on CDBG funds, were neither located in nor drew their largely low-income clients from the city's

[34] 43 FR. 8434 (Mar. 1, 1979).
[35] The Housing and Community Development Act of 1977, Public Law 96–128, amending the Housing and Community Development Act of 1974.

designated NSA's. As the centers were no longer eligible for CDBG funding, they were forced to turn to the city's general fund for support—support that has not been forthcoming due to Newark's continuing fiscal problems.

CINCINNATI, OHIO

In Cincinnati, the requirement that social services funded with CDBG grant moneys only be provided to NSA recipients has raised serious political problems about the future of a number of the city's youth centers and day care centers. The city's youth centers, located throughout the city, have relied heavily on the availability of CDBG funding. City officials, responding to the new regulations and finding no general fund support available, announced that centers located outside of NSA's could no longer receive city funds. However, the intense political pressure which developed in reaction to the announcement led to a redrawing of the boundaries of the city's NSA's so that all youth centers would remain eligible for CDBG funding.

A problem which has arisen in both Newark and Cincinnati as a result of the new requirements is the inability of the cities to provide day care subsidies to low income working parents or to supplement the subsidies received by such parents under other Federal aid programs. Contributions in each city are largely made up of CDBG funds. As eligibility is determined by income level rather than residence, many current recipients could be declared ineligible under the new regulations. Despite the fact that city officials in both Cincinnati and Newark are aware of the new requirements, they are continuing to grant day care subsidies made up of CDBG funds as they have in the past.

IV. Findings

> I see that, faithful to its popular origin, the government * * * looks after the poor, distributes annually millions to the schools, pays for all services, and rewards its humblest agents liberally. Though such a way of government seems useful and reasonable to me, I am bound to admit that it is expensive.
>
> De Toqueville, *Democracy in America* [38]

Financing the myriad activities associated with the three layers of American government remains the same expensive propostion that it was at the time of De Toqueville's expeditions to the United States in the 19th century. Indeed, the major political movements of the 1960's and 1970's in the areas of civil rights and the environment resulted in a further expansion of governmental authority which, in turn, led to the assumption of major new costs. The growth of the public sector is best illustrated by the fact that between 1955 and 1976 the local and State share of the GNP nearly doubled.

One insufficiently explored aspect of this expansion is the added cost to local government of meeting various Federal requirements. These several requirements derive from two forms of Federal legislation examined in this paper: (a) Direct orders to local government, that is, Federal mandates; and (b) congressionally established standards which serve as conditions to the receipt of Federal assistance.

As illustrated by the experience of the seven case study jurisdictions, Federal legislation frequently exacts compliance costs from the very

[38] A. De Toqueville, *Democracy in America*, trans. by G. Lawrence; ed. J. P. Mayer, Anchor Books, New York, 1969, p. 211.

jurisdictions which are the beneficiaries of Federal assistance. Concern with the cost of compliance of local government has received, at least implicitly, judicial scrutiny. The Supreme Court, in its somewhat puzzling decision in *National League of Cities* v. *Usery*[37] has stated that Federal mandates directed to State and local governments pose serious constitutional questions. In *National League of Cities*, the Court invalidated the 1974 amendments to the Fair Labor Standards Act (FLSA) which had extended Federal minimum wage and maximum hour provisions to almost all State and municipal employees.[38] The Court held that the law in question exceeded Federal authority over matters which were "essential to the separate and independent existence" of the States and was thus beyond the reach of congressional power under the commerce clause. In addition to the federalism issue which the decision raises, the opinion calls attention to the costs imposed by the FLSA as well as to the potential reductions in Government services which could result from its implementation. Although the court appears to have carefully avoided resting its decision on any factual conclusions about the cost of the regulation for State and local government, its discussion of the issues acknowledges that compliance costs were considered to be a pertinent factor in the opinion.

The Court has yet to issue any subsequent opinions which would clarify the scope of the decision. Hence it is unclear whether federally mandated local expenditures are unconstitutional per se or whether a balancing test is to be undertaken. Such a test could weigh the size of the costs against other considerations, including the relative societal need for the achievement of a federally mandated goal, or the extent to which traditionally independent local functions would be usurped by Federal orders. The findings of this paper, despite its limitations, lend support to the concern that the costs of mandates to local government and its residents are substantial and thus should be legitimately included in any such review of Federal requirements. The paper also supports the linkage between the growth of Federal aid and mandates—as all but one of the six came into effect during the 1970's, when Federal assistance to localities accelerated.

1. Magnitude of Fiscal Impact Identified

The fiscal impacts identified in the preceding sections are limited to those induced by local involvement in six federally mandated programs. Because of the difficulty and level of effort required to isolate the incremental costs of complying with and administering such Federal mandates as the Occupational Safety and Health Act, the Clean Air Act, and various welfare programs, an examination of the fiscal effects of these and other programs is not included. Indeed, our investigation revealed only one attempt by local government to develop a comprehensive accounting of all capital and operating costs induced by compliance with the entire spectrum of relevant Federal mandates.

---

[37] 426 U.S. 833 (1976).

[38] However, a number of jurisdictions, including Fairfax County, are challenging the amendments to FLSA in the courts based on the *Usery* decision. In addition, a number of others have questioned the Security and Exchange Commission's power to issue stringent bond disclosure regulations. They claim, *inter alia*, that the increased information collection and distribution costs would reduce the amount of borrowed funds that could be used to provide services.

LOCAL FISCAL IMPACT

Total local incremental operating costs during 1978 and cumulative incremental capital outlays totaled $165 million in the seven jurisdictions. Applying the conservative assumption that all capital costs are met from long-term borrowing at modest interest rates, the average incremental annual cost in 1978 for the six mandates was about $62 million, or $22.50 per capita.[89] Effects of the Davis-Bacon Act are excluded (except for one project in Burlington) due to data limitations, although it is likely that incremental costs are incurred by three of the seven case study jurisdictions. The costs of the mandates identified vary from $6 per capita in Burlington—where the Clean Water Act produced no impact—to $51.51 in Newark—where the highest costs were identified (see table 19).

TABLE 19.—LOCAL ANNUAL COST OF MEETING SELECTED MANDATES

[Per capita]

	Total	Alexandria	Burlington	Cincinnati	Dallas	Fairfax County	Newark	Seattle
Clean Water Act:								
Operating	$27.5	[1] $14.81	0	$12.20	$4.83	$7.72	$31.42	$4.68
Capital [1]	8.3	5.47	0	2.95	1.26	4.95	5.44	2.16
Educating handicapped:								
Operating	18.7	9.26	1.67	0	1.92	22.58	6.32	3.67
Capital [1]	.5	.02	1.56	0	.27	2.24	.43	.04
Access for handicapped:								
Operating	.5	[2] NA	1.10	.91	0	[2] NA	[2] NA	.31
Capital [1]	1.1	.95	0	1.41	.03	[1] .27	.77	.02
Bilingual education:								
Operating	4.1	0	0	0	.46	1.07	6.95	2.05
Capital	0							
Unemployment compensation:								
Operating	.9	.90	.90	.85	.05	.23	.18	.46
Capital	0							
Davis-Bacon Act		[3] Yes	.77	0	[3] Yes	[3,4] Yes	0	0
Total:								
Operating	51.7	31.41	6.00	18.32	8.79	39.06	51.51	13.39
Capital	9.9							

[1] Cumulative capital outlays amortized over 20 yr based on 6-percent rate of borrowing to facilitate comparisons among jurisdictions.
[2] Not available.
[3] The Davis-Bacon Act has created an effect in this jurisdiction but the effect could not be quantified.
[4] None in 1978 except for Metro.

Local costs as a percentage of total Federal assistance for the six mandates equaled 8.8 percent in Cincinnati, rising to 46.0 percent in Fairfax County (see table 20).

TABLE 20.—LOCAL PER CAPITA COST OF MEETING MANDATES AS PERCENTAGE OF REVENUE SHARING AND ALL FEDERAL FUNDS [1]

	Alexandria	Burlington	Cincinnati	Dallas	Fairfax County	Newark	Seattle	Mean
Total local cost of meeting mandates	$31.41	$6.00	$18.32	$8.79	$39.06	$51.51	$13.39	$24.07
Total Federal aid	[2] $279.00	[3] $67.00	[2] $208.00	[2] [4] $57.00	[2] $85.00	[3] $160.00	[3] $88.00	[3] $135.00
Percent mandate cost of Federal aid	11.3	9.0	8.8	15.4	46.0	32.2	15.2	18.9

[1] Based on assumption all local capital outlays are met by long-term debt.
[2] 1977.
[3] 1978.
[4] Includes $7 per capita in assistance to Dallas schools in fiscal 1979 and $5 per capita in EPA grants in fiscal 1977.

Note: Because of the difficulty in obtaining consistent data on Federal aid for the same fiscal year, the percentages are only approximate.

[89] Because of differences in fiscal years and, in some instances, projected capital outlays, some of the costs identified were not incurred in calendar 1978.

A more inclusive analysis of local costs induced by Federal mandates was conducted by Fairfax County in 1978. The results of the analysis as shown in table 21, indicate that the per capita local and State cost to meet requirements for all relevant Federal mandates identified by the county in fiscal year 1979 was over $62. The relatively small differential between the per capita cost for all mandates and the per capita cost for the six selected mandates would seem to suggest that the costs associated with the six selected mandates constitute a major proportion of all federally induced local per capita costs. In fact, the differential in most communities is probably considerably larger than that identified by Fairfax. Social service costs, in particular, probably represent local governments' largest single federally mandated outlay among those not included in this paper. As a relatively affluent community, Fairfax, however, is atypical of most jurisdictions in its low demand and outlay for such social service programs as Aid to Dependent Children (ADC), the Work Incentive Program and title XX of the Social Security Act.[40]

TABLE 21.—FAIRFAX COUNTY INCREMENTAL OPERATING AND CAPITAL COSTS TO MEET FEDERAL MANDATES FISCAL 1979

[In thousands]

	Local share	State share
Noncapital costs:		
Education of handicapped [1]	[2] $12,080	[2] $4,123
Bilingual education [1]	900	0
Clean Air Act [3]	241	0
OSHA	207	0
Public health	50	----------
Unemployment compensation [1]	120	0
Social services (title XX)	653	288
Aid to families with dependent children (title IVa, XVI)	936	2,319
Mental health/drug abuse	270	111
Administrative costs, reporting requirements	277	0
Subtotal	15,734	6,814
Operation and maintenance of sewerage treatment:		
Lower Potomac [1]	2,600	0
Blue Plains [1]	1,700	0
Debt service:		
Sewerage treatment [1]	[3] 4,345	0
Metrorail access [1]	135	0
Facilities for handicapped [1]	1,197	0
Capital outlays:		
Sex-bias access	216	0
Integrated sewer system (1978 and 1979) [1]		618
Total	25,927	7,432
Per capita	48.46	13.89

[1] Items considered as part of this report.
[2] Includes indirect cost.
[3] Differs from methodology applied in the paper.

Source: Basic data source for information based on memo from Fairfax County Office of Management and Budget to deputy county executive dated Aug. 4, 1978 (memo). Data have been adjusted to reflect newer information and modified to reflect definition of mandates and incremental cost as applied in this report. For example, FICA payments have been excluded.

STATE FISCAL IMPACT

While this paper has focused on the impact of mandates on local governments, there is a substantial State cost associated with the Federal actions identified. Many States share in the cost of meeting Clean Water Act standards as well as in meeting bilingual and handicapped

[40] However, some States fully assume responsibility for all social services; thus, the cost of Federal mandates is shifted to the State.

education requirements. For example, the State contribution to meet handicapped education requirements in Vermont is 75 percent and in Texas 40 percent of total local costs. Davis-Bacon Act provisions apply to all State projects funded with Federal money and thus to a good share of all major highway projects.

New Jersey and Washington State contribute close to $50 million in capital funds in Newark and Seattle, respectively, to meet the provisions of the Clean Water Act. The State capital contribution is the highest in New Jersey, with Newark obtaining about $53 million in State assistance to meet Federal mandates. Total State operating and capital outlays to meet Federal mandates in the seven jurisdictions identified totalled $90 million for the six mandates (see table 22). For every $1 spent by localities, States spend an additional $0.56.

TABLE 22.—STATE COSTS ASSOCIATED WITH MANDATES AT LOCAL LEVEL—1978

[In millions of dollars]

	Alexandria	Burlington	Cincinnati	Dallas	Fairfax County	Newark	Seattle	Total
Clean Water Act:								
Capital	$2.3	0	0	0	$3.8	$31.1	$18.1	$55.3
Operating	0	0	0	0	0	0	0	0
Educating handicapped:								
Capital		$0.3	0	NA	0	NA	0	.3
Operating		.2	0	$1.6	3.9	11.5	5.8	23.0
Access for handicapped:								
Capital		0	0	.03	0	8.9	0	8.9
Operating		0	0		0	NA	0	
Bilingual eduction	0	0	0	.4	0	1.8	.5	2.7
Unemployment compensation	0	0	0	0	0	0	0	0
Total	2.3	.5		2.0	7.7	53.3	24.4	90.2

TOTAL IMPACT

The aggregate annual State and local operating costs derived from the selected mandates identified for the seven jurisdictions in the 1978 and 1979 time periods was $78 million, or about $28 per capita. Aggregate State and local capital outlays to meet these mandates amounted to about $178 million. These costs were met from both long-term bonds and current revenue. Based on the conservative assumption that all capital outlays (except where debt service was identified) were met from long-term bonds at 6-percent interest over 20 years, the annual repayment cost would be about $15.5 million. Thus, a conservative estimate of total annual State and local costs associated with the mandates identified is about $93 million, or $34 per capita (see table 23).[41]

The discussion of local and State, as well as aggregate impact of federally induced outlays should be considered in the context of a few important caveats. First, the incremental costs do not include those associated with the provision of social services and numerous other mandates. In other instances, such as operating outlays associ-

[41] This means that each household residing in these communities paid about $100 annually in 1970–1978 in local and State taxes and fees to meet Federal requirements.

TABLE 23.—TOTAL OPERATING AND CAPITAL OUTLAYS GROUPED BY MANDATE

[In millions]

Mandate	Operating		Capital [1]		Capital [2]		Total [3]	
	Local	State	Local	State	Local	State	Local	State
Clean Water Act	$27.5	0	$95.1	$55.3	$8.3	$4.8	$122.6	$55.3
Unemployment compensation	.9	0	0	0	0	0	.9	0
Bilingual	4.1	$2.7	0	0	0	0	4.6	2.7
Educating handicapped	18.7	23.0	6.2	.3	.5	0	24.9	23.3
Handicapped access	.5	(4)	12.1	8.9	1.1	.8	8.4	8.9
Total	51.7	25.7	113.4	64.5	9.9	5.6	165.1	90.2
Per capita	18.75	9.32	41.12	23.40	3.58	2.03	59.87	32.72

[1] Capital outlays not amortized.
[2] Capital outlays amortized over 20 yr at 6-percent annual interest.
[3] Does not include amortized outlays.
[4] Not available.

ated with transportation of the handicapped, data will not be available for several months. Difficulties in determining differentials in service levels both with and without Federal assistance precluded the development of an accurate estimate of federally induced outlays. Second, administrative costs associated with data collection reporting requirements and the like have also been excluded from the scope of work as have all private sector costs excepting user costs.

Finally, because of the small sample size and variations found among cities identified, it would be hazardous to argue that costs identified represented all cities nationally. However, it may be feasible to approximate crudely the cost of implementing the selected mandates nationally with a larger sample.

2. *Differences in Impact and Their Causes*

While all communities within the small sample realized identifiable incremental costs associated with Federal regulations, the magnitude of the costs imposed by the six selected Federal requirements varied substantially among the seven case study jurisdictions. Given the wide variations found, it is important that causes for these differences be identified. Factors which appear critical in determining the magnitude of the cost effect associated with each of the several mandates are discussed as follows:

PRIOR ACTIVITY

In a few instances, jurisdictions acting independently or under the dictates of State law were already providing services subsequently mandated by the Federal Government. Cincinnati's long term involvement in meeting the needs of the handicapped children is an example. The provision of secondary treatment in advance of Federal requirements by Burlington and Dallas is illustrative of instances where overall compliance costs have been partially mitigated by a jurisdiction's earlier efforts. In Burlington, the availability of Federal funds spurred local action while in Dallas a combination of State regulation and the availability of Federal money contributed to the city's decision to provide secondary treatment.

One area in which there has been a considerable amount of State-level activity in advance of Federal requirements is in the establishment of prevailing wage rates. By 1976, 44 States had enacted "Little Davis-Bacon Acts" which apply to the construction of State and municipal improvements. Additional local cost, if any, therefore, is likely to depend, in part, on the prior existence of State "Little Davis-Bacon Acts" and their provisions. In Ohio, for example, State law requires the local prevailing wage rate to be tied to collective bargaining agreements or to understandings between employees and bona fide labor organizations. Texas law, on the other hand, simply requires that local prevailing wages reflect the general prevailing wages in the locality in which the work is to be performed.

DEMOGRAPHIC CHARACTERISTICS, MIGRATION PATTERNS AND SCREENING PROCEDURES

The number and proportion of persons eligible for special education and bilingual education varies greatly among cities. Differences found in each instance are attributable to a number of factors, including identification standards and the socioeconomic and ethnic composition of the population. Coastal cities and cities near the Mexican border have a high percentage of non-English speaking students in their school systems, and, thus, larger bilingual programs. By contrast, cities further inland have more homogeneous population and require fewer programs. Cincinnati, for example, incurs practically no bilingual education costs. Differences in both the screening standards and composition of the local population resulted in Burlington's identifying 10 percent of its students as handicapped, while the percentage in Cincinnati is only 5.7 percent of the student population.

FISCAL CONDITIONS

Unemployment compensation for municipal employees is illustrative of a mandate for which differences in impact are, to a large degree, related to the relative fiscal strength of local jurisdictions. Cities in poor fiscal condition, such as New York, Newark, Cleveland, and Detroit have been forced to lay off a large number of municipal employees since 1976. By contrast, growing cities in the South and West have typically expanded their public employment base in areas other than education. Thus, the burden of paying unemployment compensation tends to fall on cities which can least afford such added costs. While these costs are not currently large, another economic downturn could cause severe problems, including an increase in compensation payments. On the other hand, growing jurisdictions can absorb certain capital costs, such as those associated with improving access for the handicapped, more easily than older jurisdictions since structural modifications can be instituted at the planning stages rather than after the fact. Nevertheless, expansion pressures in growing cities are such that added regulatory costs are difficult to absorb.

LOCAL AND STATE GOVERNMENT ATTITUDE

The attitude of local government officials can, to some degree, condition their response to Federal mandates. Burlington, for example,

refused to fill out the forms necessary to determine whether the jurisdiction needed to initiate a bilingual education program. In Alexandria, school officials were skeptical of the merit of teaching basic academic skills in a foreign language. In most cases, however. local governments appear to cooperate closely with Federal agencies, probably as a result of pressure exerted by their own constituencies. Expansion of bilingual education opportunities in Dallas came about both as a result of Federal law and the efforts of the city's large Hispanic community. In regard to compliance, it is evident that a city can usually postpone, with relative impunity, the implementation of a mandate by relying upon a number of delaying tactics, including costly legal challenges.

3. Projected Fiscal Effects of Identified Mandates

CLEAN WATER ACT

The fiscal impact of the Clean Water Act on local governments and residents of affected communities far surpasses the fiscal effects of the other mandates considered in this report. While capital outlays associated with meeting the Clean Water Act provisions should stabilize in the early 1980's, the price of constructing the required facilities has been rising more rapidly than other components of the economy. Thus, above average inflation, attributable to sharp increases in demand for treatment facilities, will continue to result in larger dollar outlays. Concurrently. State contributions are likely to be reduced. Among the six states for which data were collected, two—Virginia and Washington—have practically no matching funds remaining for assistance to localities. In Virginia, local governments have had to assume the full non-Federal share of the cost. Vermont, however, is an exception to this pattern. In view of rising operating and capital costs which have produced higher user fees, many localities are refusing to pay their share for secondary treatment and for proposed tertiary treatment for phosphates. In these instances, the State of Vermont is anticipated to pay the local share of capital costs which exceed $150 per user.

The most serious economic effect for cities, however, will derive from rising operating outlays typically met by user charges. As data in this paper illustrate, operating outlays have risen sharply as a result of mandates for more sophisticated treatment processes which tend to be both labor and energy intensive. None of the technical experts in the subject area anticipates that operating costs will stabilize during the next 5 or 10 years, while most elected officials expect vigorous opposition to rising rates.

One projected cost associated with this mandate is found in a recent report issued by the Council on Environmental Quality which estimates that the aggregate incremental capital investment (excluding interest) for water pollution abatement will be $26.4 billion between 1977 and 1986, or about $122 per capita (see table 24). This implies a local and State incremental outlay of about $32 per capita in 1978 dollars. Based on the data collected for seven jurisdictions, this estimate is low, although given the small sample size, caution must be exercised. More important to localities are CEQ estimates of cumula-

tive incremental operating and maintenance costs of $31.4 billion for the 1977–1986 period, or $3.5 billion in each year ($17 per capita in 1978 dollars).[42]

TABLE 24.—ESTIMATED TOTAL AND INCREMENTAL PUBLIC COST OF WATER POLLUTION ABATEMENT (PRIMARILY SEWERAGE TREATMENT)

[In billions of 1977 dollars]

	1977			1986			Cumulative (1977–86)			
	Operating and maintenance costs	Capital costs[1]	Total annual	Operating and maintenance costs	Capital costs[1]	Total annual	Operating and maintenance costs	Capital costs[1]	Total annual	Capital investment
Base[2]	$1.7	$8.5	$10.2	$2.4	$15.2	$17.6	$21.5	$127.8	$149.3	$38.9
Incremental	1.3	.4	1.7	5.0	2.4	7.4	31.4	13.8	45.2	26.4
Total	3.0	8.9	11.9	7.4	17.6	25.0	52.9	141.6	194.5	65.3

[1] Interest and depreciation.
[2] Outlays in the absence of Federal environmental legislation.

Source: Environmental Quality Ninth Annual Report of the Council on Environmental Quality, Washington, D.C., December 1978.

THE DAVIS-BACON ACT

The effects of the act can be expected to continue at a level similar to the current pattern as Federal aid for capital facilities at the local and State levels is likely to remain stable. In view of severe data restrictions, it is difficult to estimate even crudely the level of impact. It is likely, however, that the Act has and will continue to have its greatest impact on highway and sewer system construction projects, owing to the size of these programs. In 1976, the value of federally aided highway projects was $5.1 billion, while independent State projects amounted to only $1.0 billion. Sewer system construction in 1976 amounted to $5.0 billion, with a substantial share funded by Federal grants.

Current differences between prevailing and Davis-Bacon wages in areas identified in this paper are likely to remain. While there has been some convergence in construction wages at the regional level. it is not known whether intra-area differences are being reduced. The data for Vermont suggest some convergence, but the State may not be representative of other areas. Given the complexity of determining prevailing wages and/or holding productivity constant, the data collected are insufficient to estimate the magnitude of the legislation's local fiscal effects. None of the local governments included in this paper maintains systematic records of federally funded projects which could be used to compare labor and other costs to nonfederally funded projects.

UNEMPLOYMENT COMPENSATION

The cost of this mandate in fiscal 1978 and 1979 is substantially lower than was projected by most local officials, due primarily to the

[42] The per capita values discussed are determined by dividing total incremental cost by the total national population. Since no advanced treatment facilities are required in many rural areas, the incremental annual cost for areas with advanced sewerage treatment is considerably higher. Adjusting for rural areas, the CEQ incremental operating costs estimates appear reasonable for the late 1970's. However, rising energy costs are likely to make these estimates too conservative.

expansion of private employment since 1976. Therefore, jurisdictions which were forced to lay off a number of municipal employees found that many of these workers were able to find jobs in the private sector. The improved fiscal conditions of most cities resulted in added revenue which, in turn, enabled some of the cities to expand employment, particularly in areas other than education. However, if the recession anticipated by many economists does occur, the cost of meeting potential unemployment claims could prove to be a serious problem, particularly in older northern cities.

BILINGUAL EDUCATION

As a result of increasing legal and illegal immigration from a large number of nations, the demand for and cost of bilingual education can be expected to increase in the next several years. With total enrollment in most school districts declining, outlays for this service will most likely require an increasingly large share of local educational budgets. However, future costs will depend, in part, on the outcome of various legal challenges to current bilingual education guidelines. If current requirements remain in force and are complied with strictly, costs can be expected to escalate, especially in cities along the Atlantic Coast and in southwestern States. Midwestern States will continue to enroll only small numbers of students requiring bilingual education.

EDUCATION OF THE HANDICAPPED

The provision of education services for the handicapped is probably the most rapidly rising public expenditure in the Nation. Outlays for this service in Fairfax County increased three times more rapidly than the other school services since the passage of relevant Federal mandates. Most other jurisdictions are reporting similar rates of cost increases. Both the number of handicapped in the public school system and the cost of providing services to these students on a per capita basis are rising rapidly. While one could anticipate that declining enrollments will reduce the adverse fiscal effects of these outlays in most school districts, reduced enrollments do not result in parallel reduced outlays. In fact, it is likely that handicapped students will represent an increasingly larger percent of school enrollments for the next several years.

ACCESS TO PUBLIC TRANSPORTATION FOR THE HANDICAPPED

In view of recent developments, such as unavailability of buses which meet current guidelines and proposed modifications to regulations, the potential cost of this mandate is difficult to estimate. It is evident, however, that in cities with older mass transit systems, as in the New York region, the costs could be close to $1 billion. In cities such as Washington, D.C., where Metro incorporated facilities for the handicapped into the design of its system, the added costs would be relatively low. Data on operating costs to implement current requirements will not be available for several months in most case study jurisdictions as these communities only began service in April 1979.

4. Policy Recommendations

Our policy recommendations are grouped into two categories which respond to the need for: (1) more complete analysis of the total costs which State and local governments are forced to bear in response to constraints imposed by the Federal Government and (2) change in requirements to allow general flexibility in program administration and cost effective implementation.

In the past, regulators and administrators' deliberations regarding new regulations have varied in the attention paid to costs imposed on State and local governments. Recently, however, legislative and executive branches have begun to acknowledge the need for a more complete analysis of economic impact in advance of the enactment of new regulations.[43]

We suggest that the following issues be included in any assessment of the cost of Federal regulations.

(1) *Operating Costs.*

Regulations which require major capital outlays, such as the wastewater treatment requirements of the Clean Water Act, should consider the full magnitude of operating and maintenance costs. Data recently assembled by the Council on Environmental Quality show that the total incremental operating and maintenance costs associated with sewage treatment facilities will exceed total capital outlays (both Federal and local) in about 7 years. While some higher operating costs were anticipated at the time the legislation was drafted, the scale of current increases (particularly those that have been energy related) could not have been contemplated. There is some question if the act would have been passed in its current form had the increase in operating costs been anticipated at that time.

(2) *Expansion in Demand.*

The availability of sizable amounts of Federal grants moneys to finance the construction of capital facilities may result in costly and inefficient overinvestment in "cheap" capital goods by local governments. Where local governments' construction of waste water treatment plants have been subsidized up to 90 percent by Federal and State government, many plants have been built to capacities which exceed foreseeable need. These oversized facilities have led to higher operating costs which must be borne by local consumers.

An additional area of demand-related analysis would assess whether cost estimates for Federal regulations accurately reflect total demand for programs and services once they are made available. In the case of handicapped education, the number of handicapped students in schools increased dramatically following the passage of Federal legislation. The overall cost of the federally mandated program may

---

[43] A report issued by the Congressional Budget Office notes that the House Committee on the Budget has recommended that the CBO include estimates of costs to State and local government which would arise from public bills which are reported out of all communities except Appropriations.

The CBO report points out, in addition, that Executive Order 12074, requiring the preparation of "urban and community impact analyses" and Executive Order 12044, requiring an assessment of the economic consequences for geographical regions of regulations which would have an annual economic effect of $100 million or more, acknowledge the need for more complete analysis of the impact of Federal regulations on State and local government.

have been substantially understated as estimates of total demand were based on "known" or "anticipated" cases.

The price effects of increased demand for a product or service should also be considered. When additional funds become available for goods and services prices rise as the market attempt to adjust to increased demand. Application of this concept is as valid in the health care field, (where the number of physicians, nurses and other specialized health workers has not grown in proportion to the increased demand which is attributable in part to Federal actions) as it is in the production of sewage treatment equipment. Aside from direct cost increases, one problem which arises from increases in utility equipment prices is the rapid exhaustion of funds set aside for State matching grants. As a result localities are forced to absorb a higher share of total cost.

(3) *Differential Impact.*

Finally, the differential effects of Federal mandates on communities with varying fiscal capacity, economic viability and State-local services responsibility must be considered. It is evident, even from the small sample of cities considered in this paper, that substantial variation exists. Failure to take fiscal capacity into account can aggravate urban fiscal problems—particularly when older central cities are forced to shift scarce local funds away from pressing economic needs.

FLEXIBILITY IN MEETING REQUIREMENTS

Many mandates impose rigid standards which are uniformly applied. In Burlington, for example, buses were to be lift equipped, despite the fact that they would have been useless during Vermont winters. Expensive bilingual education requirements remain in effect despite substantial disagreement among educators as to their effectiveness and appropriateness. Incremental gains which can be achieved by compliance with tertiary and advanced secondary treatment standards continue to be questioned by engineers. While the authors make no pretense to be technical experts in bus design, language education or environmental engineering, it seems likely that flexibility in meeting regulatory objectives can frequently produce cost savings to both localities and states as well as to the Federal treasury without diluting legislative intent.

REFERENCES

1. Congressional Budget Office. *Federal Constraints on State and Local Government Actions*, April 1979.
2. Behrns, Lynn P., *Mandated Costs for Colorado Local Governments*, Colorado Division of Local Governments, Denver, Colo., 1975.
3. League of Oregon Cities, *The City Budget Squeeze*, Salem, Oreg., 1974.
4. Fairfax County, Office of Management and Budget, "The Cost of Local and State Mandates," Memo to Deputy County Executive, August 4, 1978.
5. Muller, Thomas, "Service Costs in the Declining City," in *How Cities Can Grow Old Gracefully*, U.S. House of Representatives, Committee on Banking, Finance, and Urban Affairs, Subcommittee on the City, December 1977.

FEDERAL REGULATIONS AND HIGHER EDUCATION

By Carl Kaysen* and Crystal C. Lloyd-Campbell**

CONTENTS

		Page
I.	Introduction	378
II.	Some general observations about regulation	381
III.	Classification system for regulatory analysis	384
IV.	Equal opportunity, academic standards, and employment	387
V.	Conclusion	403
Bibliography		405
Appendixes		409

I. Introduction

Relations between the Federal Government and higher education have become increasingly troubled in recent years. From the academic side, there has been a well-articulated and widely heard stream of complaints, much of it directed at "overregulation" and "government interference." The government side has felt disappointment at the failures of the academy to obey the rules and has found its complaints self-serving. It is important to note that these complaints and counterclaims—government overregulation and alleged nonresponsive conduct—are not unique to higher education. This unhappy litany is a common theme between government and organizations (business as well as non-profit) subject to regulation. This should not be surprising since in many instances the problems of regulation (enforcement and compliance) which give rise to the complaints are common to all organizations subject to regulatory command.

Universities and colleges, however, are relative latecomers to this complaining chorus. Fifteen to 20 years ago, they saw Washington as a munificient patron. Federal support for scientific research was growing rapidly; a drive to encourage and assist new "centers of excellence" through institutional grants and student fellowships led to new or enlarged graduate programs in a broad spectrum of fields; and grants and loan guarantees were made available for the construction of dormitories and, later, academic buildings.

*David W. Skinner, Professor of Political Economy, Massachusetts Institute of Technology, Cambridge, Mass.
**Associate professor of law, Suffolk University Law School, Boston, Mass.

NOTE.—The authors are currently on leave from their respective institutions, and are serving as vice chairman and director of research and senior research staff member on the Sloan Commission on Government and Higher Education. The material here presented was gathered in the course of the work for the commission. Any policy views expressed are those of the authors, not those of the commission. The commission's analysis and recommendations will be presented in its report. The junior author is primarily responsible for the legal analysis; the senior author, for the discussion of the context of regulation. Both share the responsibility for the paper exposition.

In the same period, public institutions' experience of their state governments was, in general, equally agreeable. Existing institutions expanded in size, scope, and program level. New institutions and branch campuses were created in large numbers. Growth engendered a buoyant and optimistic atmosphere, in which the notes of "master-planning" and "coordination" rang with quite different overtones in both campus and capital than they do today.

The growth of Federal regulation, applied to colleges and universities, is not the only reason for the change in their attitude toward government. The end of the enrollment boom and the worsening cost situation of the institution are central elements of the context in which the growth of regulation is experienced.

The unfavorable demographic and economic contexts affect all of higher education. Fifteen years ago, colleges and universities were at the peak of their steady postwar growth. In the period 1950–65, enrollments grew at an average annual rate of over 9 percent: nearly 400,000 students were added each year. This growth derived from a combination of the increasing size of the 18-year-old cohort and a sustained increase in the fraction of the college-age population enrolling.

The number of 18-year-olds peaks in 1980 and will decline almost steadily for 15 years to a level only 75 percent of its peak value. Further, the current ratios of high school graduates to 18-year-olds (75 percent) and first-time enrollments in higher education to high school graduates (62 percent) are both high by historical standards. The former ratio increased fairly steadily from 56 percent in 1950 to about 75 percent in 1966 and has fluctuated within a small compass since. The latter ratio has moved fairly steadily from 43 percent in 1950 to 62 percent in 1970 and, again, has shown only small fluctuations since. Whether there is enough possibility for future growth in either or both of these ratios to counter the favorable demographic prospect remains to be seen, but the recent record provides little basis to expect it. Unlike the underlying demography, however, both high school completion rates and, to an even greater extent, college enrollment rates can be influenced by public policy.

The unfavorable costs situation of higher education, especially (but not only) of its private sector, has deeper roots. In the first place, there is a sense in which colleges and universities are always operating at the margin and under great pressure to spend all the money they can get and a little more. As nonprofit institutions, they do not aim at always achieving a surplus of revenues over costs but, at best, a balance. Publicly supported institutions usually present their budgets and receive their support in a way which achieves just this result. Then, each institution has a large inventory of desirable activities to undertake that will improve the quality of its performance: such as increasing the range and variety of course offerings, reducing class size, expanding the libraries, adding new laboratory space and equipment, and raising faculty salaries. There are always enough desirable items in the inventory to consume any emerging surplus.

Second, the nature of higher education is such that there is a tendency for its costs to grow more rapidly than its revenues at constant price levels. Colleges and universities operate on what is essentially a

handicraft technology, with high utilization of costly skilled labor and little or no opportunity to benefit either from scale economies (above a rather small size) or technological advance. Indeed, some important education resources, such as libraries, seem to suffer from significant technical diseconomies of scale. Yet, large and growing—and thus more costly—libraries are essential for research and advanced training in many disciplines. Compensation for both faculty and staff employees is broadly determined by competitive forces, both within and outside the academy. The rise of incomes in the whole economy reflects technological progress and other sources of increasing productivity per man. Many forces press academia to increase its pay scales in a roughly parallel way. Thus it experiences a longrun upward trend in unit costs.

The two decades of rapid expansion in the whole system added a further impulse to higher costs. Increased enrollments demanded more teachers, and supply expanded slowly and tardily. Competition for faculty, especially for leading people with established reputations and promising younger faculty, was keen. The competitive effects were particularly felt in the sciences, where rivalry between public universities with more recently established and newly expanded research programs and the private universities which had dominated the much smaller research scene before World War II, was fueled with Federal money. The result was a rapid increase in academic salaries, both absolutely and relative to other occupations. It was not confined to natural scientists, though greatest for them. Considerations of equity within institutions helped the faculty in the humanities and some of the social sciences achieve larger gains than rising demand alone might have produced. Extension of the coverage of minimum wage laws, social security, and unemployment insurance, increasing unionization, as well as the general forces of economic growth led to increases in the costs of administrative, clerical, maintenance, and housekeeping staff as well.

Until well past the middle of the 1960's, increasing enrollments and the continued growth in Federal support of research generally contributed enough on the revenue side to cover these increasing costs. In the late 1960's the situation began to change however. The growth of Federal research support in real terms ceased. Private institutions began to feel the pressure to raise tuitions. As inflation, generated by the Vietnam War and the later oil crisis, gathered momentum, these pressures increased and tuitions started to climb. In the period 1966–67 through 1976–77 average tuition at all private four-year institutions rose by an average of more than 9 percent per year, compared to a 6 percent average annual increase in the Consumer Price Index (CPI).

Continuing inflation also had unfavorable effects on publicly supported institutions, though in somewhat different ways. Increases in tuition fees at public institutions began much later than in the private sector; up to 1974 they were below the rate of increase in the CPI. More intense demands on state budgets, reflecting both an expansion of services (including higher education) and rapidly rising employment costs in the public service, left governors and legislatures less willing to provide large annual increases in the budgets for higher education. The slowing of enrollment growth and the prospect of future decline added to this reluctance to fund increased budgets. The same pressures led to some rises in tuition and wider differentials for out-of-state tuitions over those for state residents.

The underlying demographic forces hold no cheerier prospect for the public than for the private institutions. The falling off of the cohorts of college age is paralleled by a rise in those at older ages. Their demands for public services focus on health care and welfare services of various kinds, rather than on education. With fewer children and, soon, grandchildren for whom higher education is an important public good, the potential weight of the older cohort will be shifted to support increases in the services they require for themselves. Governors and legislatures can be expected to respond, and the relative weight that higher education can claim in total state expenditures will diminish.

All organizations find adaptation to declining demands for their services difficult. But it is even harder for colleges and universities, because the institution of tenure makes a large part of their labor costs rigid in the short and medium run.

In the same recent period, the depth of reliance on Federal monies by universities and colleges increased. In 1978 the dollar value of all kinds of Federal support for all higher education, including that which came to institutions indirectly through the channel of aid-receiving students, was between 12 percent and 16 percent of their aggregate current fund revenues. This compares with the 30 percent share of the states, through their direct budgetary support of public institutions. Nearly 40 percent of all students receive some form of Federal student aid, and for a large fraction of these, that aid is crucial to their decisions to attend. This means that some part of the state appropriations (which are usually geared to enrollments) depends, in a way, on the availability of Federal money.

A decade-and-half ago, Federal research support was twice as large as Federal student aid. Today, the ratios are just about reversed. Nearly every one of the 3,000-odd institutions in the collegiate sector benefits from the student aid program. Research support, on the other hand, is highly concentrated: about 100 universities get almost all of it. So the shift in the nature of Federal funding has greatly extended the Federal presence throughout all of higher education.

II. Some General Observations About Regulation

While there are regulatory issues that are unique to the system of higher education, the academy is not alone in questioning the present environment of regulatory oversight. Businessmen, economists, lawyers, political scientists, and philosophers, among others, fear that the present "regulatory state" undermines fundamental American concepts of constitutional democracy, market economy, and the rule of law.[1]

The increasing amount of regulation affecting higher education is a reflection of the increasing amount of government regulation every-

---

[1] American Bar Association Commission on Law and the Economy, "Federal Regulation: Roads to Reform" (Exposure Draft) (Washington, D.C.: American Bar Association, 1978); Freedman, James O., "Crisis and Legitimacy: The Administrative Process and American Government." (Cambridge: Cambridge University Press, 1978); Theodore J. Lowi, "The End of Liberalism," (New York: W. W. Norton & Co., Inc., 1969); Roger G. Noll, "Reforming Regulation" (Washington, D.C.: The Brookings Institution, (1971); John Rawls, "A Theory of Justice" (Cambridge: Harvard University Press/Belknap Press, 1971); Charles L. Schultze, "The Public Use of Private Interest" (Washington, D.C.: The Brookings Institution, 1977); and Roberto Mangabeira Unger, "Law in Modern Society" (New York: The Free Press, 1976).

where in society. Regulation came to the campuses because colleges and universities, like other institutional groups in this country, were slow to respond voluntarily to increased social pressure for reform.

In one area after another—housing; product, environmental, and employee safety; education; energy; and employment—Americans asked their government to relieve them from the burden of private enforcement of their rights in favor of public regulation of these activities to benefit the whole society. The call for public regulation had two broad consequences. One was a tendency to look to the Federal government for solutions, and the second was to give Federal agencies broad, open-ended and multifaceted problems for resolution without also giving them a sharply defined concept of the social goals to be achieved.

This is a change from the early history of government regulation of business. Initially, regulatory agencies were created by legislation aimed at eliminating specific problems in specific industries. This problem-oriented regulatory response was based on a theory that administrative and technical expertise would be respected by the regulated industry, the courts, and the public, alike.[2]

In contrast to the prior approach, the problems addressed by recent regulation are broader than those of the past. The new public issues, in many instances, are related to changing market behavior. The identification of a social problem does not mean, however, that there is a uniform view regarding the underlying causes of the problem or the solution. As a consequence, there has been a tendency to utilize a legislative "shotgun" approach in which lawmakers have responded to public pressure over a wide spectrum of social concerns (education, housing, health, employment) by enacting a variety of general laws calling for reform. Rarely, however, has Congress recognized that in this proliferating regulatory law there are inherent conflicts in goals and policies. In the absence of a clear consensus concerning specific content, many laws have become increasingly abstract in their legislative command.[3]

The abstraction and generality have created other tensions. Concepts of a democratic government of separated powers blur, and the initial regulatory theory of administrative expertise bows to the politics of interest group bargaining rather than to an exercise of agency expertise.[4] As a result, procedural concepts and liberalized interpretations of "standing" which grant broad access to the courts have become all important.[5] This increased use of judicial process intertwined with administrative process has intensified the adversary nature of the Government's regulatory role. In the final analysis, ambiguous policy

[2] See generally, L. Jaffe, "Law Making by Private Groups," 51 Harv. L. Rev., 201 (1937).

[3] The decline and fall of the concept of law as a series of absolute principles is ultimately related to this increasing generality and abstraction of legal command. What is "right" has become relative, and "relativity" is more easily accomplished within an abstraction. See Edward A. Purcell, Jr., "The Crisis of Democratic Theory" (Lexington : The University Press of Kentucky, 1973)

[4] Theodore J. Lowi, "The End of Liberalism," supra, n.1.

[5] Abram Chayes, "The Role of the Judge in Public Litigation," 89 Harv. L. Rev. 1281 (1976) ; Owen M. Fiss, "The Supreme Court 1978 Term, Foreward : The Forms of Justice," 93 Harv. L. Rev., 1 (1979) ; Richard Stewart, "The Reformation of American Administrative Law," 88 Harv. L. Rev., 1667 (1975) ; L. H. Tribe, "Structural Due Process," 10 Harv. CR–CL L.R. 269 (1975).

fosters conflict. The courts are increasingly called upon to balance conflicting policy interests and develop the substantive content of regulatory law.[6]

Review of the self-studies of government regulation made by 21 institutions for the Sloan Foundation in connection with the Sloan Commission on Government and Higher Education indicates that many of the specific comments about regulation of higher education are intimately related to the systemic problems of regulatory law in general.

From the perspective of a university administrator, understanding how to respond to regulatory command is made more difficult when the desired goal is entangled in myriad laws and regulations. Among the more specific complaints of academic institutions are:

Duplication of regulatory efforts, necessitating duplicative response by the institution;

Changing, conflicting, vague, and at times nonexistent standards for measuring institutional compliance;

The heightened cost burden in light of this regulatory uncertainty; and

The adversarial nature of the relationship between government agency and institution, caused by regulatory emphasis on new procedures, broad access to multiple forums and agencies for grievants, and, in many instances, the overwhelming magnitude of regulatory sanctions for noncompliance.

These complaints are all related to the systemic problems noted earlier: multiple laws, many of which are too general and too abstract and many of which appear to present internal and external conflicts of ends and means without any coordination.

There is no one congressional committee or administrative agency responsible for overseeing the burden of Federal Government regulation on any one institution. The 35 congressional committees (16 in the Senate, 18 in the House, and one joint committee) and their 70-plus subcommittees oversee 439 Federal programs affecting higher education administered by over 35 Federal agencies. This is not the complete picture since some nonprogrammatic agencies also regulate the activities of academic institutions. Then too, a growing number of courts is involved in issues directly related to academic administration. Policies and practices in connection with student admissions,[7] student government,[8] academic credit,[9] employment and tenure,[10] student retention,[11] and system-wide reorganization and merger [12] are just a few of the academic issues recently decided in the courts.

---

[6] See, e.g., *Califano* v. *Westcott*, 99 S. Ct. 2655 (1979); "The Supreme Court, 1978 Term," 93 Harv. L. Rev., 60, 133 n.21; *Dunlop* v. *Bachowski*, 421 U.S. 560 (1975).

[7] *Cannon* v. *University of Chicago*, 99 S. Ct. 1946 (1979); *Davis* v. *Southeastern Community College*, 99 S. Ct. 2361 (1979); *Regents of the University of California* v. *Bakke*, 438 U.S. 265 (1978).

[8] *Uzell* v. *Friday*, 401 F. Supp. 775, 547 F.2d 801 (CA 4, 1977); cert. granted, vacated 46 L.W. 3803 (1979); 591 F.2d 997 (CA 4, 1979).

[9] *Wayne State University* v. *Cleland*, 440 F. Supp. 806 (E.D. Mich., 1977), aff'd. in part, rev'd. in part, 590 F.2d 627 (CA 6, 1978).

[10] *Board of Trustees of Keene State College* v. *Sweeney*, 99 S. Ct. 295 (1979), remand, 604 F.2d 106 (CA 1, 1979); *Powell* v. *Syracuse University*, 580 F.2d 1150 (CA 2, 1978) cert. denied, 417 L. W. 3369 (1978); *Kunda* v. *Muhlenberg College*, 463 F. Supp. 294 (E.D. Pa. 1978).

[11] *Board of Curators of the University of Missouri* v. *Horowitz*, 435 U.S. 78 (1978).

[12] *Geier* v. *Blanton*, 427 F. Supp. 644 (D.C. M.D. Tenn. 1977), aff'd., 597 F.2d 1056 (CA 6, 1979), cert. denied, 48 L.W. 3222 (1979).

It is the combined effect of all these factors—the growth of governmental institutions (Judiciary, Congress, and Executive) with overlapping powers of command and control, the universities' increasing reliance on government support, and the changing economic environment—which generates the intensity of frustration and complaint concerning government regulation and the academy.

III. Classification System for Regulatory Analysis

Federal regulations affecting higher education can usefully be divided for analytic purposes as follows: (*i*) Rules and standards generally applicable to all organized activity in society (with or without an associated receipt of government funds); (*ii*) rules affecting activities specific to higher education; and (*iii*) rules related to financial accountability in connection with support for activities specific to academic institutions, such as student aid and research support.

Examples of the first category—regulation which generally applies to all organizations—cover a broad range of government regulations. Included in this category are laws as diverse in nature as: Social security taxes and benefits, workmen's compensation. Occupational Safety and Health administration (OSHA), environmental protection, Employee Retirement Income Security Administration (ERISA), Fair Labor Standards, public broadcasting, and a wide range of equal opportunity laws affecting housing, employment, and educational programs.

Some of these laws apply only to organizations which directly or indirectly receive Federal funds—such as Executive Order 11246 (as amended by 11375),[13] which requires affirmative action in employment by all Federal contractors—but most of them are not linked to the receipt of Federal funds.

These across-the-board laws share certain common characteristics. First is that they generally originated in the context of problems arising in business organizations. Only later were these applied to nonprofit organizations such as hospitals, museums, and colleges and universities.[14] Second, these laws generally relate to the financial security, health, and safety of employees. Third, these laws have an unusually heavy financial impact on colleges and universities because higher education is so labor intensive. Indeed, because these costs are taxes on employment, some writers have contended that they constitute an erosion of the value of the tax-exempt status of the nonprofit organizations.[15]

These laws and regulations and others of similarly broad scope serve social purposes that have commanded the degree of acceptance neces-

---

[13] Exec. Ord. 11,246 (30 Fed. Reg. 12, 319) (1965) as amended, 32 Fed. Reg. 14, 303 (1967).

[14] The Fair Labor Standards Act of 1938, was amended in 1966 to apply to hospitals and institutions of higher education; Fair Labor Standards Amendments of 1966, Pub. L. No. 89–601, 80 Stat. 836. The Old Age Assistance Program, instituted pursuant to the Social Security Act of 1935, was extended to include nonprofit educational institutions, two-thirds of whose employees voted for coverage in a referendum. Social Security Act Amendments of 1950, § 204(e), 64 Stat. 477, 535. Because the votes were so overwhelmingly for coverage, Congress eliminated the referendum requirement in 1960. Social Security Amendments of 1960, § 105(a), Public Law No. 86–778, 74 Stat. 924, 942.

[15] C. Van Alstyne and S. Coldren, "The Costs of Implementing Federally Mandated Social Programs at Colleges and Universities," Policy Analysis Service—Special Report (Washington, D.C.: American Council on Education, June 1976), p. 15.

sary for the enactment of legislation in a democratic and pluralist society. Colleges and universities are not alone in finding them burdensome and costly. Yet, legislatures which have approved them must have made the calculation, implicitly if not explicitly, that the social benefits justify their costs.

In contrast to the across-the-board health and safety regulations are regulations which apply only to organizations that are recipients of Federal funds. They are primarily equal opportunity or nondiscrimination regulations.[16] Most of them were enacted relatively recently and they reflect a broad social movement against discriminatory practices that have denied equal opportunity to many groups.

Examples of regulations of the second category (affecting activities specific to higher education) include the Family Education Rights and Privacy Act (Buckley Amendment) and Department of Health, Education, and Welfare (DHEW) regulations governing the use of human subjects and the care of animals used in research. An example of a different kind of regulation in the second category is the Health Professions Educational Assistance Act (1976) which, among other things, requires that medical schools seek to alter the distribution of medical students among areas of specialization as a condition of receiving Federal capitation funds. Unlike other examples cited, this Act reaches directly and deeply into the heart of the educational enterprise.

The nature of the third type of regulation is almost self-evident: rules prescribing the purposes for and the terms on which Federal grants, loans, and contracts are available and persons or organizations entitled to receive them; and reporting and auditing requirements placed on the recipients.

The number of regulations of all three types has grown rapidly since the mid-1960's, with the greatest growth in the first and second categories. The third has also grown, as new student aid programs have been legislated, with their corresponding eligibility requirements, application procedures, and reporting rules. The volume and scope of rules governing research grants have changed relatively little, but their burdensomeness appears to have increased with the slackening growth in research support.

Most laws and regulations that have stimulated the anxious concern of academia so far have fallen into the first and third categories. Those in the second category are relatively few; and almost all are concerned with medical schools, their associated hospitals, and biomedical research, rather than with higher education generally. The Buckley Amendment protecting the privacy of student records is exceptional in this respect.

The greatest volume of complaint appears to arise in response to the first category of regulation, which is the one that does not focus specifically on higher education. There are several reasons for this. Rules in the third category, those connected with the receipt of Federal funds (in the form of research support or student aid) impose admin-

---

[16] There are three major equal employment laws which are linked to the commerce clause rather than to the receipt of Federal funds: Title VII, the Age Discrimination in Employment Act, and the Equal Pay Act. Even though these statutes are not linked with receipt of Federal funds, because they deal with the regulation of equal employment, they are being discussed with this group of regulations.

istrative burdens and costs of the same sort—and frequently of greater magnitude—as do those of the first category. They also give rise to similar dissatisfactions: unnecessary red tape, duplicative or irrelevant paperwork, and incompetent government administration. However, they accompany tangible benefits of great value to the institutions. Almost no institution gives up the benefits to avoid the costs, although every one has the choice. No direct tangible institutional benefits correspond to the costs and burdens of compliance with regulations of the first category; hence the complaints are, so to speak, free. Many administrators and professors who express strong support for the goals of these enactments and regulations nonetheless view the whole process of regulation in this area as a transaction in which "they" impose costs on "us" in pursuit of "their" social goals. The further possibility must be allowed that, for some, complaints about the burdens of compliance and the incompetence of the administration are means of expressing disapproval of the social goals of the regulations themselves.

It is difficult to estimate the cost burdens that the regulatory process imposes on academic institutions.[17] The Sloan Foundation sponsored self-studies of a sample of 21 institutions drawn from all parts of the higher education spectrum in the summer of 1977. These suggested a representative figure for compliance cost of, at most, something like 2 percent added to the operation budget through increases in clerical and administrative staffs.[18] This number, while it clearly varies widely for specific institutions, is on the average neither overwhelming nor insignificant. The figure alone cannot answer several fundamental questions: do the results justify the costs? could equally good results be achieved more economically? who should bear the cost? And, of course, such an estimate is unable to assign a dollar value to the effects of increased regulation on institutional operations. This is widely experienced as an increase in formality and bureaucracy at the cost of ease, informality, and collegiality. The cost estimate also does not cover the diversion of time and energy of senior administrators and faculty members to the problems of compliance, finite resources which might otherwise be spent on more centrally educational problems.

Some regulations within this first category of across-the-board regulation, particularly the equal opportunity regulations, are seen as threatening to impair the capacity of academic institutions to perform their functions effectively by transferring from the academic community to government bureaucracies[19] crucial decisions on who should

EDITOR'S NOTE.—Following the discussion, the authors present their conclusion in which it is stated: "Measuring the impact of federal regulatory policies in higher education is a difficult business. One man's burden is another's benefit." The authors also state that regulatory laws achieve outcomes for minorities, women, the handicapped, aged and veterans that might not otherwise have occurred. Therefore, the following analysis should be read with the knowledge that benefits are addressed in the conclusion.

---

[17] See "The Costs of Implementing Federally Mandated Social Programs at Colleges and University," supra, note 15.

[18] See J. A. Kershaw, "Government and Higher Education, A Survey of 21 Constitutions," Sloan Commission on Government and Higher Education, October 1977.

[19] Richard H. Lester, Antibias Regulation of Universities, (New York: McGraw-Hill, 1974).

teach, who should be taught, and even what should be taught. This discussion will therefore focus on these laws—government regulation affecting all organizations—which may unintentionally be altering the function and structure of academia.

IV. Equal Opportunity, Academic Standards and Employment

In the 21 self-studies prepared for the Sloan Foundation there appeared to be "universal agreement" that Federal equal opportunity laws covering educational programs, policies, and employment have "more of an impact at the campus level than other types of regulations." [20] Judged by the number of books and articles written on the topic recently, it is fair to conclude that the effects of these laws rank as the most salient issue in any discussion of government regulation and higher education.[21]

At least in part, the increased degree of regulation of the academic enterprise was provoked by the failure of academics to critically view their own institutions and practices. In the wake of this failure, regulations applicable to business organizations were extended to academic organizations, and new regulations did not exempt them.[22] As a result, many agencies and procedures are involved in overseeing the activities of academic institutions. The functions of many of these agencies overlap, creating unnecessary duplication of effort and enervating repetitive response.

This problem of duplication and concurrent jurisdiction is well illustrated in the area of equal opportunity laws. Thirteen different laws or executive orders concerning this objective recently have been enacted or issued. Including the post-Civil War constitutional amendments and civil rights statutes, there are 17 Federal laws generally related to equal opportunity. A minimum of eight different government agencies have direct responsibility for enforcement of these provisions in higher education, and several more have similar responsibilities by virtue of Federal monies granted by them to academic institutions. In addition to this Federal administrative jurisdiction, the Federal courts have jurisdiction over private suits which may be brought under these statutes. (State constitutions and nondiscrimination laws also address these issues; the paper does not discuss them.) The functions of many of the administrative agencies overlap, creating duplication of effort and demanding enervatingly repetitive

[20] Irene K. Spero, "Government and Higher Education, A Summary of 21 Institutional Self-Studies," paper for the Sloan Commission on Government and Higher Education, January 1978, p. 35, citing the Rice University self-study, p. 23.

[21] See generally. Walter C. Hobbs, "Government Regulation of Higher Education" (Cambridge: Ballinger Publishing Co., 1978); Carnegie Council on Policy Studies in Higher Education, "Making Affirmative Action Work in Higher Education" (San Francisco: Jossey-Bass Publishers, 1975); Richard A. Lester, "Antibias Regulation of Universities" (New York: McGraw-Hill, 1974); Bernice Sandler, "Backlash in Academe: A Critique of the Lester Report," vol. 76, No. 3; Robert A. Scott, "The Hidden Costs of Government Regulation," Change, vol. 10, No. 4, April 1978, pp. 16–23; McGeorge Bundy, "The Issue Before the Court: Who Gets Ahead in America?" Atlantic Monthly, vol. 240, No. 5, November 1977, pp. 41, 46; Chester E. Finn, Jr., "Federal Patronage of Universities in the United States: A Rose by Many Other Names" Minerva, vol. XIV, No. 4, winter 1976–77, pp. 496–529.

[22] One recent exception is the Age Discrimination in Employment Act Amendments of 1978, Public Law No. 95–256, 92 Stat. 189, in which the ban on mandatory retirement before age 65 was amended to extend to age 70. The Amendments, which generally took effect Jan. 1, 1979, will not be effective for tenured employees in higher education until July 1, 1982, however.

responses from the institutions.[23] Procedures for enforcement and standards for compliance vary among agencies.

The complexity and uncertainty thereby created perhaps can best be illustrated by considering a pair of hypothetical situations:

(i) The Case of John Doe, Student

Chem 220 is one of the most famous courses offered at Oxbridge, a leading (private) research university. The course is famed for its rigor and for the fact that any student who gets a grade of "B" or above in the course is practically assured admission to medical school. Only 250 students can be admitted to the course, and because of its reputation it is always oversubscribed. The class is given in the Holsworthy Lecture Room, the only one large enough to hold a class that size. Holsworthy is located on the top floor (third) of the Old Science Building, a magnificant late-Victorian building.

John Doe, a senior at Oxbridge, is a 37-year-old black veteran. He lost one of his legs in the Vietnam War and uses crutches to get around campus. John wants to go to medical school. Last year John was one of the few minority students enrolled in Chem 220. It was difficult for him to get to Holsworthy, and he generally arrived late. He received a "C" in the course, which was the highest grade received by any of the minority students in the course. This year John applied to the state medical school but was not admitted there or to any other medical school. He ranked below the fiftieth percentile in the National Medical College Admission Test (MCAT).

To understand the issues raised in this situation, some understanding of the legal concept of discrimination is necessary. Discrimination is a denial of equal opportunity because of some irrelevant condition such as race, sex, national origin, religion, age, or handicap. Discrimination can occur in two ways: Disparate treatment of similarly situated persons explicitly on the basis of the irrelevant condition, or the disparate impact of a facially neutral practice on a previously disadvantaged group.

There are three different standards for determining unlawful discrimination: Constitutional, statutory, and contractual. Under the constitution (Amendments V and XIV), unlawful discrimination requires evidence of discriminatory intent accompanying the allegedly discriminatory action.[24] Hence the person charged with a constitutional violation must show an absence of an intent to discriminate.

In some instances, however, statistical disparities in outcomes may be evidence of discriminatory intent.[25] In contrast, the legal standard

[23] In May 1978 President Carter's Reorganization Plan (No. 1 of 1978, 43 Fed. Reg. 19807 (1978)) for consolidating some of the responsibility for enforcement of the equal employment opportunity programs was approved. This reorganization plan was phased-in over the course of two years. By July 1979 the *three* agencies chiefly responsible for enforcement of the employment opportunity laws were: EEOC, Department of Labor/OFCCP (Executive Order Program); and the Department of Justice. This consolidation of enforcement responsibility enumerates some of the confusion in the area of employment, but it does not affect enforcement of the educational program equal opportunity laws, such as Titles VI, IX and § 504.

[24] *Village of Arlington Heights* v. *Metropolitan Housing Development Corp.*, 429 U.S. 252 (1977); *Washington* v. *Davis*, 26 U.S. 229 (1976).

[25] *Village of Arlington Heights* v. *Metropolitan Housing Development Corp.*, supra note 24. "Determining whether invidious discriminatory purpose was a motivating factor demands a sensitive inquiry into such circumstantial and direct evidence of intent as may be available. The impact of the official action—'whether it bears more heavily on one race than another.' *Washington* v. *Davis*, 426 U.S. at 242—may provide an important starting point. Sometimes a clear pattern, unexplainable on grounds other than race, emerges from the effect of the state action even when the governing action appears neutral on its face." (Citations omitted) 429 U.S. at 266. See also, *Washington* v. *Davis*, supra, note 20, J. Stevens, concurring, 426 U.S. at 254: "I agree, of course, that a constitutional issue does not arise every time some disproportional impact is shown. On the other hand, when the disproportion is as dramatic as in *Gomillion* v. *Lightfoot*, 364 U.S. 339, or *Wick* [*Wo* v. *Hopkins*, 118 U.S. 356 (1886)], it really does not matter whether the standard is phrased in terms of purpose or effect." See also, *Board of Education of City School District of New York* v. *Harris*, 100 S. Ct. 363 (1979). But compare, *Personnel Administrator* v. *Feeney*, 99 S. Ct. 2282 (1979) and "The Supreme Court, 1978 Term," 93 *Harv. L. Rev.*, 60, 137, 138 (1979)

for statutory discrimination is one set by the legislature, or implicitly by the administrative agency or court applying the particular statute. Generally, statutory standards for unlawful discrimination do not require evidence of discriminatory intent.

For instance, under Title VII (employment), statistical evidence that a neutral employment practice has disparate impact is sufficient to support a finding of unlawful discrimination.[26] Thus the employer charged with statutory discrimination has a more difficult test to meet: He must demonstrate that he did not discriminate even inadvertently. Finally, contractual standards of discrimination further complicate the picture. The standard for Executive Order 11246 is based on an affirmative obligation to remedy underrepresentation of certain groups of the employer's work force.[27] This obligation arises out of the contractual relationship between government and contractor. This contractual standard is thus more demanding than either the constitutional or statutory standard.

The hypothetical situation could lead to various charges of discrimination or unequal treatment. Of course, a claim does not mean the charge is proved. The period between an initial claim and a final decision, however, can be long and the process costly (in energy as well as money) for all of the parties concerned.

Doe, a black student in a private institution, may raise claims against Oxbridge under the following statutes requiring nondiscrimination in educational programs: Title VI (race),[28] Age Discrimination Act,[29] and Section 504 of the Rehabilitation Act of 1973 (handicap).[30] Doe may also claim a violation of the XIV Amendment. In addition to these programmatic claims, Doe may use the link between the course, Chem 220, and admission to medical school to assert that both Oxbridge and the state medical school violated several nondiscrimination-in-employment statutes, specifically Section 503 of the Rehabilitation Act of 1973 (handicap),[31] the Vietnam Era Veterans Readjustment Assistance Act of 1974,[32] and the Age Discrimination in Employment

---

[26] Compare, *Griggs* v. *Duke Power Co.*, 401 U.S. 424 (1971) (under the statute, Title VII, disparate treatment on the basis of some suspect classification (such as race or sex) is presumptive evidence of discriminatory intention) with *Washington* v. *Davis*, supra note 20 (discriminatory intention is a necessary element of proof of constitutional violation under the 5th and 14th amendments.) Even under the statutory standards, however, intention may become a relevant issue. When an alleged discriminatory practice is "facially neutral", evidence of disparate impact presents a prima facie case of discrimintion.

A . . . prima facie showing is not the equivalent of a factual finding of discrimination however. Rather, it is simply proof of actions taken by the employer from which we infer discriminatory animus because experience has proved that in the absence of any other explanation it is more likely than not that those actions were bottomed on impermissible considerations. When the prima facie showing is understood in this manner, the employer must be allowed some latitude to introduce evidence which bears on his motive. * * * . . . [T]he District Court was entitled to consider the racial mix of the work force when trying to make the determination as to motivation. *Furnco Construction Corp.* v. *Water*, 438 U.S. 567 (1978) at 579–580.

See also, *Board of Trustees of Keene State College* v. *Sweeney*, U.S. , 99 S. Ct. 295 (1978), where the Court stated: While words such as "articulate," "show," and "prove," may have more or less similar meanings depending upon the context in which they are used, we think that there is a significant distinction between merely "articulat[ing] some legitimate nondiscriminatory reason" and prov[ing] absence of discriminatory motive." . . . [W]e made it clear that the former will suffice to meet the employee's prima facie case of discrimination. *Id.* at 295 and 296, fn. 2.

[27] See Office of Federal Contract Compliance Program Task Force, U.S. Department of Labor, A Preliminary Report on the Revitalization of the Federal Contract Compliance Program, Sept. 16, 1977, p. xix. The recent case of *United Steelworkers of America* v. *Weber*, .U.S. , 99 S. Ct. 2721 (1979) held that voluntary affirmative action undertaken by a government order program did not violate the discrimination standards of Title VII.

[28] 42 U.S.C. § 2000(d) (1976).

[29] 42 U.S.C. § 6101ff. (1975).

[30] 29 U.S.C. § 794 (1976).

[31] 28 U.S.C. § 793 (1973).

[32] 38 U.S.C. § 2011 (1974).

Act of 1967.[33] Although these employment-based claims are certainly more attenuated, two courts have said that the term "employment agency" in similar statutes should be liberally construed to include law and medical schools.[34]

Doe has a private right of action under some of these laws and therefore could immediately bring suit against Oxbridge.[35] He might prefer, however, to bring his complaints to the agencies responsible for enforcing the relevant laws and ask them to investigate his charges.[36] For the programmatic statutes (Title VI, Title IX, Section 504, ADA), the relevant agency would be the Office for Civil Rights (OCR) in the Department of Health, Education, and Welfare.[37] The relevant agency for the employment-based claims would be the Office of Federal Contract Compliance Programs (OFCCP) in the Department of Labor.

On the basis of these facts, the regulatory agencies might determine that Oxbridge has violated the law. The third-floor classroom presents an access problem for a handicapped student. Oxbridge may have to move the class, build ramps, or install an elevator to be in compliance with Section 504.[38]

In addition, the statistical evidence on the number of minority students admitted to chem 220 and the lack of any minority students receiving a "good" grade may lead to a finding of discrimination based on disparate impact under Title VI.[39] If any of the programmatic

---

[33] 29 U.S.C. § 621–636, as amended (1978).

[34] See, *Kaplowitz* v. *University of Chicago*, 387 F. Supp. 42 (N.D. Ill., 1974); *Cannon* v. *University of Chicago*, 559 F. 2d 1063 (1977), . U.S. , 99 S. Ct. 1946 (1979) (reversed and remanded on different issue).

[35] See, *Cannon* v. *University of Chicago*, supra, note 34 (Title IX) . Since Title IX is modeled after Title VI, the decision has some impact on the question of whether there is a private right of action under Title VI as well. But see *Bakke*, supra, note 7 (four Justices assumed a private right of action under Title VI in this case, four Justices declined to decide the issue and one Justice denied a private right of action under Title VI); see also *Lau* v. *Nichols*, 414 U.S. 563 (1974). And see *Southeastern Community College* v. *Davis*. U.S. 99 S. Ct. 2361 (1979) (§ 504 of the Rehabilitation Act of 1973).

[36] But see Age Discrimination Act, supra, note 22, c. 45 C.F.R. § 90.43(d), 90.50. 44 Fed. Reg. 33768 (1979) which requires exhaustion of administrative remedies prior to bringing a civil suit in the courts. But see *Cannon* v. *University of Chicago*, supra, note 34.

[37] The enforcement responsibilities of OCR for education will be shifted to the new Department of Education. This will occur 180 days after the Secretary of Education takes office, or sooner, at the President's discretion. See Public Law 96–88, 96th Congress 2d session (Oct. 17, 1979).

[38] The recent case, *Southeastern Community College* v. *Davis*, supra note 32, does not appear to reach this type of *physical* access problem. The Court held that "an otherwise qualified person [who may not be excluded from participation in an educational program] is one who is able to meet all of a program's requirements in spite of his handicap." (Id. at 2367). However the basis for the Court's decision in Davis was that Section 504 did not require "fundamental alterations" in *curriculum* to meet the needs of handicapped students, "Section 504 imposes no requirement upon an educational institution to lower or to effect substantial modifications of standards to accommodate a handicapped person." Id. at 2369, 2370. Furthermore the Court felt that an interpretation of the regulations which required extensive curriculum modifications—"beyond those necessary to eliminate discrimination against otherwise qualified individuals . . . would do more than clarify the meaning of § 504. Instead they would constitute an unauthorized extension of the obligations imposed by the statute." (Id. at 2369). However modification of physical facilities to meet the needs of otherwise qualified handicapped students does appear to be an instance "where an insistence on continuing past requirements and practices might arbitrarily deprive genuinely qualified handicapped persons of the opportunity to participate in a covered program." Id. at 2370.

[39] This may still be a "hypothetical" situation in the context of higher education. This type of issue has been raised in other educational situations, however. See, McClung, "Competency Testing Programs: Legal and Educational Issues," 47 Ford L. Rev. 651 (1979); also, *Larry P.* v. *Riles*, 343 F.2d 963 (CA9 1974); permanent injunction issued, No. C–71–2270 RFP, (U.S. Dist. Ct. N.D. Calif., Oct. 16, 1979) (enjoining use of I.Q. Tests as a basis for placement in "stigmatizing" educable mentally retarded category by school authorities). See also, *Board of Education City Sch. Dist. of N.Y. v. Harris*, supra fn. 25.

claims ultimately are resolved in Doe's favor, Oxbridge may lose its Federal funding including student financial aid as well as government contracts.[40]

Doe is just one student; multiply him by the number of potential student claimants on that campus and one will understand why academic administrators are concerned about the impact of these laws. Another example, this one involving faculty members, will further illustrate these problems.

The Case of Associate Professors Robert Roe and Mary Moe

Old State is the "star" institution in the public system of higher education in the state of Newconsin. It is a highly respected research university with several outstanding graduate degree programs. In the past, Old State's sociology department did not have a strong reputation; but in the last six or seven years the department has grown from 8 to 21 fulltime positions, and it has increased its emphasis on research as well as teaching excellence. Throughout this period, Old State was also actively seeking to increase the size of its black faculty. Old State has several major contracts of more than $1 million with the Federal government.

Robert Roe, a black male, and Mary Moe, a white female, have teaching appointments in the sociology department at Old State. Roe had taught for several years at another college and had completed all but his dissertation (ABD) when he was recruited and hired as an associate professor by the department three years ago. He asked for and received a commitment to pay him $15,000 during the first year of his three-year contract if he received his Ph.D. (which he did) by the time he arrived on campus to begin teaching in the fall. This was $3,000 above the rate which the sociology department was paying for Ph.D.'s.

In the spring of his second year, Roe was evaluated by the department and the decision was made not to renew his contract. The reasons given for nonrenewal were that Roe's performance in the areas of teaching and scholarship were below the standard the department wished to maintain. Roe has not produced any scholarly publication since coming to Old State, and student enrollment in his "Black America" course fell dramatically during the two years he has been at Old State. Roe has been involved in community activities, and he is chairman of the local branch of the NAACP.

Mary Moe is an older woman who returned to graduate school to complete her graduate work after her youngest child entered high school. She received her doctorate from one of this country's most prestigious institutions. She was hired as an instructor (at a salary of $9,500) in the sociology department seven years ago. Her contract was renewed and she has received regular promotions. Moe is now as associate professor.

Mary Moe's research work concentrated on studies of the blue-collar work force. In the past two years she has increasingly focused her research on the role of the workingclass woman in American society. Moe has published extensively in this area and has several women graduate students who came to Old State to study with her specifically. She has been active in organizing a women's group on campus. This year, her final year on contract, she submitted her dossier for consideration by the faculty tenure committee. The committee acknowledged her teaching excellence but felt her recent publications were "insufficiently scholarly" in view of the objectives of the department and therefore denied her request for tenure. Old State has an "up or out" rule and, as a result, Moe must seek a teaching position elsewhere.

---

[40] See *Bob Jones University* v. *Johnson*, 396 F. Supp. 597 (D.C. S.C. 1974) which held institutional eligibility for continued receipt of Federal student financial aid is conditional on compliance with Federal equal opportunity laws and regulations (Title VI). See, also, In the Matter of Hillsdale College, ALJ decision, HEW 8/23/78 (refusal to sign assurance of compliance form under Title XI not evidence of noncompliance until specific violation of statute or regulations shown); reversed, HEW Civil Rights Reviewing Authority, XIX Chronicle of Higher Education, No. 10, p. 16 (Nov. 5, 1979).

The Roe and Moe problem also presents a multiple statute/multiple forum situation. On the basis of these facts, they may be able to make claims under several statutes with the following agencies:

Roe	
XIV Amendment	Individual or class action private suit in Federal or State Court.
Title VII (nondiscrimination in employment).	EEOC (Equal Employment Opportunity Commission), individual complaint investigation, may lead to a "right to sue" letter or EEOC may refer the matter to the DOJ (Department of Justice) for filing of a "pattern and practice" suit against a public institution.
Executive Order 11246, as amended by Executive Order 11375.	OFCCP/DOL (Office Federal Contract Compliance Programs, Department of Labor), individual complaints referred to EEOC. Systemic discrimination investigation and compliance enforcement OFCCP/DOL; may be referred to DOJ for further action.
Moe	
Title IX (discrimination on the basis of sex in educational programs).	Complaint filed with OCR/HEW. Moe may have a private right of action under Title IX. Title IX.
Title VII	EEOC/DOJ, individual complaint investigation, may lead to a "right to sue" letter or EEOC may refer the matter to the Department of Justice for filing of a "pattern and practice suit" against a public institution.
Age Discrimination in Employment Act.	EEOC.
Equal Pay Act	EEOC.
Executive Order 11246, as amended by Executive Order 11375.	OFCCO/DOL, individual complaints referred to EEOC. Systemic discriminaton investigation and correction enforcement OFCCP/DOL; may be referred to DOJ for further action.

Assuming Roe and Moe ultimately prevail on any of these claims, Old State may have to reinstate Roe and grant Moe tenure.[41] They could be subject to an order requiring equalization of salaries and an award for back pay. These complaints also may cause problems for Old State in connection with the Executive order contract compliance program, and its contracts with the Federal Government may be delayed or terminated. One additional note: Both Moe and Roe may be able to raise these issues before a state antidiscrimination commission under the relevant state employment nondiscrimination statute. In some instances, the state law does not follow the Federal antidiscrimination standards.[42]

The examples give some idea of the bewildering array of equal opportunity/nondiscrimination laws, regulations, and enforcement agencies affecting higher education. These situations are drawn from actual cases.[43] An identical factual situation, such as an apparent un-

---

[41] See, e.g., *Scott* v. *The University of Delaware,* 455 F. Supp. 1102 (D.C. Del. 1978); *Kunda* v. *Muhlenberg College,* 463 F. Supp. 294 (E.D. Pa. 1978).

[42] See, e.g., Mass Gen. Laws, c.151B. Compare, *Mass. Electric Co.* v. *Mass. Commission Against Discrimination,* 375 N.E. 2d 1192 (1978) with *General Electric Co.* v. *Gilbert,* 429 U.S. 125 (1976). See, also, *Smith College* v. *Mass. Commission Against Discrimination,* 380 N.E. 2d 121 (1978), *Wheelock College* v. *Mass. Commission Against Discrimination,* 355 N.E. 2d 309 (1976).

[43] See, generally, *Scott* v. *University of Delaware,* supra, n. 41; *Cannon* v. *the University of Chicago,* supra, n. 7; *Sweeney* v. *Board of Trustees of Keene State College,* supra, n. 10; and *Powell* v. *Syracuse University,* 580 F.2d 1150 (CA 2, 1978), cert denied, Nov. 28, 1978, 417 L.W. 3369; *Seattle University* v. *Department of Health, Education, and Welfare,* W.D. Washington, Jan. 3, 1978, C77–631S; *Lamphere* v. *Brown University,* 71 F.R.D. 641 (D.C. R.I. 1976).

equal pay structure, can trigger multiple compliance reviews by several different agencies. Since each agency has its own perspective, each agency's particular "lens" will color its perception of the facts. Therefore, identical facts may result in different outcomes in the several agencies and perhaps among different branch offices of the same agency.

The complexity and confusion of one enforcement procedure is well illustrated by a case now called *Adams v. Califano.*[44] This 10-year-long Title VI action involves a privately initiated suit against HEW requesting the agency to enforce the nondiscrimination policies in educational institutions and programs in 10 formerly *de jure* segregated state systems. In this discussion it is impossible to give the details of the impact of the action on the various state systems involved. At the Federal court level the matter has involved six consecutive Secretaries of HEW, six General Counsels, and six Directors of the Office for Civil Rights (OCR), the HEW agency directly responsible for enforcement. By 1975, of the 10 states originally involved, two states refused to respond or negotiate with OCR and their cases were transferred to the Department of Justice for prosecution (and, thereafter no further action was taken in one state), one negotiated a settlement, and one state countered by bringing suit against HEW.[45] The remaining six states attempted to negotiate with OCR to develop state plans to promote racial integration in their formerly dual systems.

The initial state plans, approved by OCR/HEW in 1974, later were repudiated by the agency in 1976 and found unacceptable by the Federal district court in Washington, D.C. in 1977. Thereafter each of these six states worked to develop new plans acceptable to the agency. By 1979, five state plans had been given provisional and, later, final approval by the agency. The state plan submitted by North Carolina was provisionally approved in May 1978 but denied final aproval in March 1979. ORC announced it was going to institute formal administrative enforcement proceedings possibly leading to termination of all Federal funds to the state. The state brought suit to enjoin funding delays and the court granted a temporary injunction. Administrative hearings in this matter were to begin in January 1980.[46]

This recital does not do justice to the feelings of frustration and uncertainty of academic administrators responsible for running these systems for the past 10 years nor to the complex politics involved in assuring state legislative support for the increased academic budgets promised in these plans.[47] The economic cost and emotional cost have been high, and yet at no point has the Federal agency made clear what its standard for measuring compliance is.

In the *Adams* case, both the plaintiffs and OCR have applied legal and regulatory standards developed in the context of elementary and

---

[44] *Adams* v. *Califano,* 433 F. Supp. 118 (D.C.D.C. 1977), the full history of the case appears in: *Adams* v. *Richardson* 351 F. Supp. 636 (D.C.D.C. 1972, as amended, 1973); *Adams* v. *Richardson,* 356 F. Supp. 92 (D.C.D.C. 1973); *Adams* v. *Richardson,* 480 F.2d 1159 (C.A.D.C. 1973); *Adams* v. *Califano,* 430 F. Supp. 118 (D.C.D.C. 1977). The four *Adams* cases will be referred to as "the *Adams* cases."

[45] *Mayor and City Council of Baltimore* v. *Mathews and Mandel* v. *HEW,* 571 F.2d 1276 (C.A. 4 1978) (per curiam, withdrawing opinions issued in 562 F.2d 914 (C.A. 4 1977) because of death of one of the circuit judges prior to issuance of opinion), cert den'd, 99 S. Ct. 184 (1978).

[46] *State of North Carolina* v. *Department of Health, Education, and Welfare,* No. 79–217–Civ–5 (E.D.N.C. June 8, 1979).

[47] For a detailed case study of this case, with particular emphasis on higher education systems in North Carolina, see Crystal Lloyd-Campbell, "Adams v. Califano: A Case Study in the Politics of Regulation," working paper for the Sloan Commission on Government and Higher Education. January 1978. Postscript June 1979.

secondary public school education, with little examination of the appropriateness of these standards for higher education. The substantive content of standards developed for elementary and secondary schools defines equal education opportunity in terms of an integrated unitary model of education.[48] The social science methodology which supported the development of that model, however, has been questioned by subsequent sociologists, particularly those working in related areas.[49]

The importance of this division of opinion is not the relative merits of one view or another, but the acknowledgment that there is a substantial difference of opinion on these issues, a difference which is even more apparent when one moves from compulsory elementary and secondary education to a situation of voluntary choice of educational institution. This element of voluntary choice is, of course, one of the primary characteristics of higher education. The agency, however, has taken little notice of these considerations. Furthermore, these standards may have an unintended (and harmful) effect on the black colleges in the *Adams* states. Almost all of these institutions have played an important role in providing educational opportunity to minority students and many believe there is a continuing need for their services.[50]

The problems noted here are problems of competing values. Professor Lindblom has noted that public policy often is made by making choices among policies rather than by a rational means-end analysis:

> The idea that values should be clarified, and in advance of the examination of alternative policies, is appealing. But what happens when we attempt it for complex social problems? The first difficulty is that on many critical values or objectives, citizens disagree, congressmen disagree, and public administrators disagree.
>
> Administrators cannot escape these conflicts [among objectives] by ascertaining the majority's preference, for preferences have not been registered on most issues. . . . By the impossibility of doing otherwise, administrators often are reduced to deciding policy without clarifying objectives first.[51]

---

[48] See *Columbus Board of Education* v. *Pennick*, 99 S. Ct. 2941 (1979) ; *Keyes* v. *School Dist. No. 1*, 413 U.S. 189 (1973) ; *Swann* v. *Charlotte-Mecklanburg*, 402 U.S. 1 (1971) ; *Green* v. *County School Board*, 391 U.S. 430 (1968).

[49] See, James S. Coleman, "Equality of Educational Opportunity," a report for the U.S. Department of Health, Education, and Welfare, Office of Education (Washington, D.C.: National Center for Education Statistics, 1966). Compare, Frederick Mosteller and Daniel Patrick Moynihan, "On Equality of Educational Opportunity" (New York: Random House, 1972) ; Christopher Jencks, "Inequality" (New York: Harper & Roe, 1972). Coleman himself appears to have had second thoughts about his earlier conclusions in 1975. Professor Coleman reassessed his original conclusions and concluded that mandatory desegregation had more negative than positive effects. As he stated recently : "This belief in the inherent inferiority of an all-black school has a curiously racist flavor." See Lorenzo Middleton, "The Effects of School Desegregation: The Debate Goes On," The Chronicle of Higher Education, vol. XVII, No. 10, Nov. 6, 1978; and James S. Coleman, letter to the editor, The Chronicle of Higher Education, Vol. XVII, No. 15, Dec. 11, 1978.

[50] In 1975, 5 years after the *Adams* suit was started, an organization of Black Educators (NAFEO) filed an amicus brief in support of HEW, in opposition to the plaintiffs' (NAACP-LDF) "immediate integration" request. NAFEO said the black colleges had historically fulfilled a "crucial need" and should continue to play an important role in black education. The court acknowledged their concerns by noting that desegregation might "place a greater burden" on black colleges, or eliminate existing opportunities for black students. He ordered HEW to set guidelines for higher education desegregation which took "into account the importance of Black colleges *and at* the *same time*, comply with the congressional mandate." *Adams* v. *Califano*, 430 F. Supp. at 120 (emphasis added). In practice, of course, this is a difficult tight-rope to cross. Achievement of goals for increased access by black students into traditionally white colleges and universities, may shift black students at the expense of black colleges. The ability of the traditionally black institutions to attract new sudents (white or black) is hampered by this past history and the charges by civil rights leaders and others that "black schools are educationally bankrupt" and deny equal educational opportunity. See, D. Bell, Jr., "Serving Two Masters: Integration Ideals and Client Interests in School Desegregation Litigation," 85 Yale L. Rev. 470, 479 (1976) ; See also J. H. Wilkson III, "From Brown to Bakke: The Supreme Court and School Integregation," 1954–1978 at 187–189 (1979). See also, Randolph, "Academic Irony: Black College Seeking to Stay Black Undergo Pressure to Integrate, the Wall Street Journal, Mar. 19, 1979, at 1.

[51] Charles E. Lindblom, "The Science of 'Muddling Through'," 19 Pub Ad Rev 79, 86 (1959).

But in many areas Federal administrators have not accomplished even this incremental evaluation of competing policies with respect to higher education. The programmatic equal opportunity laws (Title VI, Title IX, Section 504, and the Age Discrimination Act) present many difficult policy problems for regulatory agencies and higher education. In many instances, however, policy has not been developed, and regulatory standards range from inappropriate to nonexistent. For instance, notwithstanding HEW's long-term involvement with the application of Title VI to systems of higher education, no separate regulations concerning this subject have been issued.

In 1977, HEW issued desegregation criteria to assist the six *Adams*-affected states. This approximated regulatory policy, but court-ordered criteria are not regulations. They were not exposed to even the loose hearing and comment protection accorded under the informal rule-making procedures of the Administrative Procedures Act. Thus, the appearance of impartiality and expert evaluation which accompany rulemaking procedures is absent from the fundamental issue of the meaning of equal educational opportunity within higher education.[52]

Along with Title VI, other programmatic equal opportunity laws (Title IX, Section 504, and the Age Discrimination Act) also present regulatory difficulties. Here, too, regulatory standards are often inappropriate or nonexistent. There are critical areas of regulatory enforcement which raise major issues concerning academic standards. Some of the issues raised in connection with the application of Title VI to higher education have already been mentioned. The problems of developing standards for Title VI are compounded when considering Title IX and Section 504 regulations. The concepts of race discrimination developed out of a history of judicial, executive, and administrative experience; such is not the case for theories of discrimination on the basis of sex or handicap. What is the practical meaning of equal educational opportunity on the basis of sex? The troubled development of regulatory policy in connection with providing equal opportunity in college and university athletic programs spotlights this critical problem. The practical meaning and content of the recent regulations governing access for the handicapped are even more worrisome. Uniform standards for access amid the particularized problems of the handicapped are subject to wide variations of interpretation which are just now being tested in the courts. This is one example where the extraordinary financial impact of the access requirements on the institution is interrelated with serious issues of the internal academic decisions of whom to admit and how to teach them.

And what are the further implications of nondiscriminatory access? Is access enough? Will the concepts of statistical discrimination which have developed in connection with equal opportunity in employment be extended to the area of student admission and retention?[53] The *Adams* criteria (which relate to remedial efforts to correct past discriminatory policies) require the public institutions involved to analyze admission and retention patterns and correct deficiencies in retention between minority and nonminority groups. The indications

---

[52] There are problems with administrative proceedings too: see R. Stewart, "The Reformation of American Administrative Law," 88 Harv. L. Rev. 1667 (1975); D. Horowitz, "The Courts and Social Policy" (1977); James O. Freedman, "Crisis and Legitimacy: The Administrative Process and American Government," supra note 1.

[53] See Fn. 25 and 39 supra.

from Washington are contradictory, and these standards may have implications extending beyond the remedial concepts of *Adams*. For instance, the U.S. Commissioner of Education recently stated, in connection with student aid regulations, that "The content of the institution's standards of satisfactory progress is strictly an institutional matter."[54] Recently proposed student aid regulations, however, state that if an institution has a dropout rate exceeding 33 percent a year, it will be ineligible to participate in student aid programs.

The absence of clear standards creates other problems. Varying standards are applied in essentially similar situations, or compliance standards shift erratically. Both factors created problems for the university administrators in the *Adams*-affected states. Many state administrators questioned why their states were subject to these Title VI enforcement proceedings when other states with similar histories and similar dual systems remained untouched by the Federal agency.[55] Furthermore, what compliance standard there was shifted as agency personnel changed. For example, in the period 1970–76, the agency emphasis was on entry of white students into traditionally black institutions. Now under the *Adams* criteria the emphasis is on the number of black students entering traditionally white institutions. The policies and budget commitments made by the states in the earlier period may not be appropriate for these new goals. Shifting priorities and objectives in public systems of higher education is a long-time process not amenable to these sudden shifts in regulatory directions.

Notwithstanding the shared purpose of equal opportunity legislation, each agency develops its own procedural style. For instance, the orientation of EEOC is focused on voluntary settlement of individual complaints. In large measure, the agency has relied on private litigants to develop the substantive standards of Title VII.[56] In contrast, enforcement of the Executive Order program is seldom triggered by complaints. The procedural orientation of the enforcement agency (OFCCP) focuses on elimination of systemic problems by contract. It emphasizes result-orientated procedures and specific numerical standards (affirmative action plans) irrespective of an employer's actual history of discrimination.[57]

The different procedural styles may account for differential enforcement efforts by the regulatory agencies when enforcing similar

---

[54] 43 Fed. Reg. 37898 (Aug. 24, 1978).

[55] J. H. Wilkinson, "From *Brown* to *Bakke*," supra note 50 at 217; see Lloyd, fn. 46 at 55–56.

[56] See R. Belton, "A Comparative Review of Public and Private Enforcement of Title VII of the Civil Rights Act of 1964," 31 Vand. L. Rev. 905, 914, 917 (1978); EEOC does have authority to bring "pattern and practice" suits, but it had not used this authority effectively. In the past, the agency focus on individual complaints was a product of its statutory duty to make an investigation of every complaint. See 42 U.S.C. 2000(e)–5. See also, David Copus, "Long-Term Problems in Title VII Enforcement," Internal Memorandum, Special Investigation and Conciliation Division, EEOC. See also, R. W. Walker, "Title VII: Complaint and Enforcement Procedures and Relief and Remedies," 7 B.C. Ind. and Comm. L. Rev. 495 (1966).

[57] See, "A Preliminary Report on the Revitalization of the Federal Contract Compliance Program," Office of Federal Contract Compliance Programs Task Force, U.S. Department of Labor, pp. 5, 7, 13–16, 30 (1977).

The Department of Justice also illustrates this problem of differing procedural styles resulting in differentiated enforcement of the laws. Department of Justice has exclusive authority to litigate public sector Title VII employment cases and suits against Federal contractors under the Executive Order program. The Department has concentrated on race discrimination cases and has not "adequately addressed" the problems of sex discrimination in employment, The Federal Civil Rights Enforcement Effort—1977: To Eliminate Employment Discrimination: A Sequel Report of the U.S. Commission on Civil Rights, pp. 246–247, 272–273 (1977).

regulatory provisions. Agencies with similar responsibilities and overlapping jurisdiction under the laws apply different standards and approach enforcement differently. For instance, the Department of Justice, which has exclusive authority to litigate public sector Title VII employment cases and suits against Federal contractors under the Executive order program,[58] has concentrated on race discrimination cases and has not adequately addressed the problems of sex discrimination in employment.[59] Similarly, HEW has taken action against certain state systems of higher education under Title VI but not against *private* institutions in these same states, although they may have the same history of *de jure* race discrimination.

Other problems arise in the enforcement of regulatory policy within higher education. Those problems reflect the lack of experience and training in higher education of all the compliance coordinators or equal opportunity specialists—the agency personnel (by whatever name)—who investigate equal opportunity questions on the campus. For ORC the grade range of personnel is from GS 9 to GS 13; for the Wage and Hour Division of the Department of Labor it is from GS 5 to GS 13.[60] A recent vacancy announcement for such positions in OCR listed experience as a condition of employment, but not a college degree or even attendance. In some instances, education will be substituted for experience, but it is obvious from the job description that experience is the crucial factor. What type of experience? Experience with the institutions being regulated is not sought. Rather, experience in various "community programs designed to promote equality," "special experience" demonstrating "knowledge of the causes and effects of discriminatory practices against minorities and women," and experience "working with or promoting equal opportunity for handicapped persons" constitute the desiderata.[61]

Another problem is decentralized control within the agencies. Decentralized agency structure frustrates the development of new approaches as well as the development of uniform standards and procedures. Moreover, decentralization widens the gap between regulatory theory and practice in both the development and implementation of policy. Policymakers may express respect for the concept of institutional autonomy as they develop equal opportunity policies. For instance, regulations which are directed to achievement of goals and outcomes without mandating specific process and procedures (such as the *Adams* criteria) are examples of such autonomy-oriented policy objectives. The pressures on administrators of enforcement agencies, however, are different from the pressures on policymakers. When enforcement of new policy becomes the responsibility of others, the procedure may become result orientated.

---

[58] 42 U.S.C. 2000(e)(5). Prior to the 1972 amendments, the Department of Justice had exclusive authority to initiate "pattern and practice" suits against private employers. See U.S. Commission on Civil Rights, supra note 57, pp. 246–247; Office of Federal Contract Compliance Programs Task Force, U.S. Department of Labor, "A Preliminary Report on the Revitalization of the Federal Contract Compliance Program, Sept. 16, 1977."

[59] See U.S. Commission on Civil Rights, supra note 57, pp. 272–273.

[60] U.S. Commission on Civil Rights, supra note 57, p. 151; and conversation with OCR personnel officer 2/9/79.

[61] Office of the Secretary of HEW, Vacancy Announcement #79–533–B, Opportunity for Equal Opportunity Program Specialist, opening date 12/27/78, closing date 1/11/79.

As reported by an official in Region 1, at one pre-award compliance investigation conducted by OFCCP at a prestigious private university, none of the agency's on-site personnel had had any prior experience with higher educational institutions.

The distinction between theory and practice which may be noticed at headquarters in Washington is even more pronounced when comparing actions taken in the field with policies defined in Washington. For instance, pre-award compliance reviews under the Executive Order program can have different outcomes depending on where the institution is located and what regional office is performing that review. Enforcement outcomes which differ because of territorial distinctions, rather than factual distinctions, underscore the subjective involved in equal opportunity enforcement efforts and further discourage institutional initiative toward voluntary compliance.

In contrast to the variety of procedures and forums available to initiate complaints, there is relatively little flexibility in connection with the remedy to be applied once a violation of an equal opportunity regulatory standard is established. In the past 15 years, 13 laws have been enacted focusing on nondiscrimination and equal opportunity programs or employment which affect higher education. Nine of these statutes provide that the sanction for violation of statutory or regulatory standards is termination of Federal funds and Federal contracts to the noncomplying institution or an entire system of higher education. This is a powerful sword, perhaps too powerful and overwhelming to be useful.

The threat posed by fund termination has a divisive impact on the collegial structure of academic institutions and magnifies the specter of Government coercion. Both research funds and student-based funds are subject to these statutory sanctions.[62] As a consequence, this sanction has an unfair, coercive quality and at the same time a mythic quality. Though to date no institution has had its Federal funds terminated under the equal opportunity laws, many institutions have been threatened with loss of Federal funds.[63]

For academic administrators, fund termination may represent the difference between institutional life or death. Thus administrators take the threats seriously. Commitments in the name of equal opportunity may be made without realistic promise of performance. These sanctions have an aura of unreality for the faculty since no institution has actually lost its Federal funds. Yet the achievement of the goals of nondiscrimination and equal opportunity depends primarily on changes in attitudes and behavior at the faculty level. Thus, the utility of this sanction as a means of inducing real change in social behavior in the faculty can be questioned.

The threat—not the reality—divides the academy, administrators against faculty and even faculty against itself. A traditional departmental autonomy is seen as threatened. Faculties in the sciences and humanities may have different perspectives on recruiting or hiring procedures, but it is the science contract money which is at stake if the humanities faculty does not meet the regulatory standards. Thus, in

[62] See *Adams* v. *Califano,* 430 F. Supp. 118 (1977) ; *Bob Jones University* v. *Johnson,* 396 F. Supp. 597 (1974) ; In the Matter of Hillsdale College, ALJ Dec. 8/23/78.

[63] Currently (1979) the public university system of North Carolina faces possible termination of Federal funds in connection with enforcement of Title VI, and a private college in Texas faces a cut-off of student aid funds in connection with alleged fiscal mismanagement. The University of California, the University of Michigan, and others, have faced fund termination threats in connection with the Federal contract compliance program.

terms of impact upon higher education, this regulatory process may not achieve its intended object, assuring nondiscriminatory behavior. The unintended consequence is a breakdown of the collegial structure.

Furthermore, the significance of these equal opportunity regulations and the perception of regulatory "burden" must be seen in terms of the context within which colleges and universities experience it—specifically, the unfavorable economic and demographic prospects facing higher education. The dean who would find it irritating in any case to appear in a proceeding to answer an allegation of sex discrimination will find it doubly so if he is preoccupied with trying to justify the budget for the maintenance of his present faculty, 80 percent of whom have tenure, when the enrollment in his program already has begun to fall. An equally important aspect of this context is the large and still increasing role of the Federal Government as patron of higher education. In this context, real, as well as threatened, burdens have a major impact. The fund termination sanction is illustrative of Federal regulatory policies which have both a financial impact as well as a structural impact on higher education.

In contrast to the potential impact of fund termination, the variety of procedures—forums for complaint and enforcement agencies presently associated with equal opportunity regulation—has a direct impact on institutions of higher education. This impact can be measured in terms of increased financial, emotional, and structural costs. The financial impact is measured in terms of increased cost of operation, shifting scarce resources from academic research, and creating a larger administrative staff at the expense of the faculty.

But the real impact cannot be measured in economic terms. The real impact is measured in the gradual erosion of the academic collegial structure in a ceaseless round of complaints and rebuttals. This is not to say that academia is not in need of internal reform. Like other organizations it was slow to respond voluntarily to the social and judicial pressure for reform enunciated in the late 1950's and early 1960's. The equal opportunity/nondiscrimination regulations are a public response to the frustrations and failures of organizational self-reformation. However, regulations which apply across-the-board to all organizations—without recognition of their different structure, function or environmental context—may in the long run frustrate the real goals of equal opportunity.

There are real differences between the operation of academic organizations and industrial and business organizations. The differences are often ignored by the agencies that enforce the laws and regulations prescribing equal opportunity. These differences are rooted in the distinctive function of higher education—to preserve, transmit, and develop knowledge—and the unique organizational structure which has evolved in universities and colleges to perform this function. The paradigmatic academic organization is a relatively nonhierarchial body based on the collective leadership of colleagues rather than on a bureaucratic administration with a hierarchial structure of authority. Its organizational structure is designed to foster relationships of equal intellectual interchange between academic peers and discipleship between students and teacher.

In the recent *Yeshiva*[64] case, the appellate court distinguished the industrial model of administration (managerial hierarchy) from the "shared authority" concepts of collegial administration. It noted that "The interests of the faculty and of the University were almost always coextensive." Faculty and management "substantially and pervasively" operated the enterprise.[65]

The court stated:

> The university's unique set of goals (education, research, and service) is achieved only by a series of specialist communities working together through their common concern for enlarging and applying their own spheres of knowledge. Thus, there is no sharp dividing line between the community of administrators and the community of faculty, for both have the common goal of striving to further the institution as a house of learning.[66]

In several instances the theory supporting the collegial structure differs from practice. The collegial organization can provide an excuse to avoid responsibility for making hard decisions, or an arena in which discriminatory decisions may be made in the name of collegiality. But some errors in practice do not necessarily prove the theory wrong. Indeed, rather than discarding the theory, the discrepancy between theory and practice may warrant efforts to reform the practice. At the very least, the discrepancy should not justify irrevocably changing the present system as an *inadvertent* consequence of regulation. In two major areas the already tense relations between Government and higher education are likely to become more difficult in the next 10 to 15 years. The first is faculty employment, and the second—a more general category—involves issues of academic standards including student access and retention, and educational content. The issue of quality—how it is defined, produced and tested—lies at the heart of both of these categories.

In contrast to the declining pool of college-age students, the pool of potential women faculty members is increasing in total size and distribution across academic disciplines. The demographic changes which have occurred in doctoral programs since 1970 are startling. In 1970, women received 3,970, or 13.5 percent, of the doctorals awarded in the United States; in 1977, women received 7,845, or 24.8 percent, of the doctorates awarded.[67] Thus, while men have been enrolling in Ph.D. programs in fewer numbers since 1976, the number of women enrolled in doctoral programs continues to increase. In addition, women have increased their share of the doctorates awarded in each academic field.[68]

The outlook for academic employment is poor. Three factors combine to create this situation. One is, of course, the contraction caused by declining enrollments. Second, the present tenured faculty is relatively young. Over one-half of the American professoriate has been hired in the past 15 years. The faculty members hired in the 1960's,

---

[64] *NLRB* v. *Yeshiva University*, 582 F.2d 686 (CA 2, 1978), cert. granted, 99 S. Ct. 1212 (1979).

[65] Id, at p. 700.

[66] Id. at p. 701, quoting Kahn, "The NLRB and Higher Education: The Failure of Policy-making Through Adjudication," 21 UCLA L. Rev. 63, 63 (1973). Compare *NLRB* v. *Wentworth Institute*, 515 F.2d 550 (CA 1, 1975).

[67] Data from the National Research Council Doctorate File. The information in this paragraph is discussed in greater detail in David O. Levine, "The Condition of Women in Higher Education," paper for the Sloan Commission on Government and Higher Education, January 1979.

[68] See table 6 in Levine, supra note 67.

the expansion period, are now tenured.[69] On many campuses, the proportion of tenured faculty exceeds 50 percent. Third, as noted earlier, Congress has recently raised the permissible mandatory retirement age from 65 to 70.[70] Despite the three year exemption period, the best predictions indicate there will be few openings for faculty employment in the next 15 years. Thus, there will be more women qualified to teach and few positions available. Every teaching slot will be subject to fierce competition, and the competitive tensions will grow as the new faculty moves up the promotion ladder to tenure.

The situation is different for minorities. There are job opportunities but too few qualified candidates. After a decade of effort, blacks still constitute an extremely small proportion of the doctorates produced annually in this country. In 1977, blacks received less than 4 percent of the total of all Ph.D.'s awarded, and 60 percent of these minority doctorates were in the field of education. Nonetheless, even in this area, black scholars constituted less than 9 percent of the total.[71]

Thus, there is every indication that academic policies and decisions relating to students (admission, retention, graduation) and faculty employment (hiring, promotion and tenure) increasingly will be scrutinized and challenged by those dissatisfied by the results. Many women and minorities will not be hired, promoted, or given tenure. Some will perceive themselves as victims of discrimination sometimes rightly so, and complain. As noted, there are multiple procedures and forms available for potential grievants—individuals, groups, and regulatory agencies—to challenge the outcome of academic policies. Here, of course, is where the problems created by vague, inappropriate or nonexistent regulatory standards, have the most impact.

For instance, the uniform application of Federal equal employment opportunity laws raise problems. Consider the problem of a discrimination standard based on statistical comparisons. Statistical data is all-important in determining whether there is a "pattern and practice" of discrimination under Title VII [72] or work force "underutilization" for purposes of the Executive order.[73] The regulatory goal is to achieve statistical parity between the employer's work force and the relevant labor pool. Selection of the appropriate relevant labor pool is therefore crucial, particularly in academia: yet, here is where the law and regulations are vague and inappropriate.[74]

The academic labor force is small (537,000) when compared to the entire professional and technical work force (13,329,000) or the clerical (15,558,000) or craft blue-collar (11,278,000) work force.[75] The statis-

---

[69] *Roy Radner* and *Charlotte Kuh*, "Preserving a Lost Generation," Carnegie Council on Policy Studies in Higher Education, 1978.

[70] National Center for Education Statistics, "Salaries, Tenure, and Fringe Benefits of Full-Time Instructional Faculty in Institutions of Higher Education," 1977–78 Report.

[71] National Center for Education Statistics, Projections of Education Statistics to 1968–87, 1978.

[72] See *Hazelwood School District* v. *United States*, 433 U.S. 299 (1977) ; *International Brotherhood of Teamsters* v. *United States*, 431 U.S. 324 (1977) ; *Albemarle Paper Co.* v. *Moody*, 422 U.S. 405 (1975) ; *Franks* v. *Bowman Transportation Co.*, 424 U.S. 747 (1974) ; *Griggs* v. *Duke Power Co.*, 401 U.S. 424 (1971).

[73] 41 CFR §§ 60–2.11 to .12 (1977).

[74] See also, Thomas Reed Hunt and George Pazuniak, "Special Problems in Litigating Upper Level Employment Discrimination Cases," 4 Delaware J. of Corporate Law 114 (1978), at pps. 147, 149–152; "A proposal of Reconciling Affirmative Action with Nondiscrimination under the Contractor Antidiscrimination Program," 30 Stan. L. Rev., 803, 815 (1978).

[75] These figures can be found in Department of Labor, Employment and Training Report of the President, 1977, pp. 186–187.

tical standards have been developed in the context of blue-collar or entry-level industrial management situations. They do not realistically describe the available academic labor pool.[76]

Furthermore, the regulations relate neither to a problem of "impacted" employment nor to a situation where members of a preferred group have greater employment opportunities across the industry than white males. This is the problem facing academia today, where blacks now command a premium in the academic marketplace.[77]

Standards developed in business and industry also may be inappropriate in the context of academic employment where specialized training and experience and relative qualifications and quality are such an important part of the hiring process.[78] These "relative quality" considerations are even more important in connection with promotion and tenure decisions in academia, since the commitments may span the entire professional career of a professor. The Supreme Court has acknowledged that:

> When special qualifications are required to fill particular jobs, comparisons to the general population (rather than to the smaller group of individuals who possess the necessary qualifications) may have little probative value.[79]

No regulations have been issued, however, to address the problem of what would have probative value in the context of academic employment. A serious collateral problem arising from the regulation of academic employment is the need to balance confidentiality of faculty dossiers, regulatory oversight, and public disclosure law.[80] The prob-

---

[76] Draft memorandum of William C. Bowen, "Utilization Analyses and Goals and Timetables," July 20, 1976, pp. 4–5, submitted to the Federal Advisory Committee on Affirmative Action in Employment at Institutions of Higher Education, Robben Flemming, Chairman; Making Affirmative Action Work in Higher Education, "Excerpts from University of North Carolina Memorandum," Jan. 17, 1975, pp. 240–242 and *Id.* at pp. 56–59 (1975). One commentator has suggested that the Griggs statistical disparity standard is appropriate where the discrimination issue involves promotion of incumbent employees, but not where the group is "a large shifting, ill-defined, potentially limitless and endlessly manipulable group of would-be employees, as in *Davis.*" See B. Lerner *"Washington* v. *Davis:* Quantity, Quality and Equality in Employment Testing," The Supreme Court Rev. pp. 263, 270 (1976).

[77] Recent research indicates that black faculty now command a premium in the academic marketplace. See, R. B. Freeman, "The New Job Market for Flack Academicians," 30 Ind. and Lab. Rel. Rev. 161, 168–169 (1977); W. Williams, "Higher Education and Minority Opportunities," 7 Howard L.J., 545, 552 (1978); see, also, "A proposal for Reconciling Affirmative Action with Nondiscrimination under the Contractor Antidiscrimination Program," supra note 74 at 814–815 (1978).

One may cheer this development but for academic administrator it presents a paradox. The premium salary paid to black male professors may alter the statistics of academic salaries at an institution to the point that nonminority female professors would appear to have a facially valid Equal Pay Act complaint. See Ester Greenfield, "From Equal Pay to Equivalent Pay: Salary Discrimination in Academia," 6 J. of Law and Education, 41, 42 (1977).

[78] See Carnegie Council on Policy Studies in Higher Education, "Making Affirmative Action Work," supra note 76, pp. 56–59. See also, *Scott* v. *University of Delaware,* supra note 43 where the plaintiff challenged the Ph.D. requirement as discriminatory because it "has a disparate impact on blacks which is not justified by the legitimate needs of the University," at p. 1123.

Another aspect of this "qualification" problem is contained in the OFCCP affirmative action guidelines relating to contractor workforce analysis. "Employee specifications cannot set higher qualification standards than those of the lowest qualified incumbent," 41 C.F.R. § 60–2.24(f)(5)(1974) (emphasis added). This lowest common denominator rule could mean that an institution that wished to "upgrade" its faculty would be held to its present standards until the "lowest qualified incumbent" retired!

The courts have not required a showing by the plaintiff of superiority, but only minimal competence to meet the competency standard. See *Powell* v. *Syracuse University,* 580 F.2d 1150, 1155 (CA 2 1978): "[P]roof of competence sufficient to make out a *prima facie* case of discrimination under Title VII was never intended to encompass proof of superiority or flawless performance. * * * In the context of this case, Ms. Powell has demonstrated that she possesses the *basics* necessary for the performance of her job, and has thereby made out a *prima facie* showing of competence." (emphasis added).

[79] *Hazelwood School District* v. *United States,* supra note 72 at p. 308 fn. 13.

[80] See, Freedom of Information Act, 88 Stat. § 6, 5 U.S.C. § 552, Privacy Act of 1974, 88 Stat. 1095, 5 U.S.C. § 552(a); see also, Family Educational Rights and Privacy Act, 20 U.S.C. §§ 1232 g–i.

lem here is less a question of inappropriate regulation than it is a lack of any standards to guide administrators in a crucial area. The faculty dossier has unique qualities: "pound for pound" they are files that "count most" in the development of an academic carrer.[81]

The academic progression through the ranks is accompanied by an ever-growing dossier; once in the files, nothing is discarded. Access to the files is generally restricted. The assessments, evaluations, and comments of academic review committees (such as promotion and tenure committees) are made and given in confidence. These confidential assessments, evaluations, and comments are not available for direct review even by the person under consideration.

Historically, the rationale which supports strict confidentiality has been the need to protect the integrity of the peer review process and the institutional desire to make offers and commitments only to the most qualified person available in any particular field. Confidentiality protects the process in two ways: It promotes open, honest, and (in theory) unbiased evaluations by the assessor concerning the candidate's qualifications while protecting his or her reputation from injury.

In this age of regulation and agency enforcement, the boundary between rights of privacy and public access has become blurred. "Statutory law basically gives the agency wide latitude on use within the scope of its legitimate function and does not provide very direct protection against indiscriminate use."[82] For instance, the Executive order regulations provide that the contractor must allow the Government investigators complete access to all records including copying and taking material that may be relevant to assuring compliance with the order.[83] At the same time, however, the Freedom of Information Act grants individuals broad access to agency records and materials on matters which affect the interests of those individuals. One of the aims of the public disclosure laws is to protect the "consumer" and provide public oversight of agency actions. If the university must relinquish confidential material, however, is there any way the public agency can protect the privacy of these academic files while in its hands? How shall the privacy interest of academia be balanced against the public interest? There is no guidance in case law or regulations on this crucial question. To date, these issues have been left to ad hoc, case-by-case determination, without public discussion of the policy issues or development of regulatory standards to guide agencies and institutions.[84]

V. Conclusion

Measuring the impact of Federal regulatory policies in higher education is a difficult business. One man's burden is another's benefit. Regulations aimed at changing entrenched patterns of social behavior are particularly difficult to analyze according to traditional cost-bene-

---

[81] See Clark "The Dossier in Colleges and Universities" in Wheeler, "On Record: Files and Dossiers in American Life." (1969).

[82] David W. Leslie, "Legal Protection for Privacy," in Protecting Individual Rights in Higher Education, 1977. See. also, "Draft Presidential Decision Memorandum for the Privacy Poicy Coordinating Committee, National Telecommunications and Information Administration." U.S. Dept. of Commerce. pp. 2–3 (1978).

[83] See, e.g., 41 CFR § 60–1, 43, §60–60.3 (1974).

[84] See *Department of Health Education and Welfare, and Department of Labor* v. *University of California,* ALJ Opinion September 1979, OFCCP No. 5239475.

fit methods. From the perspective of the minorities, women, handicapped, aged and veterans—those for whom these laws are designed—the laws do achieve outcomes that might not otherwise have occurred.

On the one hand, there is disagreement among various groups about the efficacy of the regulatory approach to affect changes in organizations and institutionalized behavior patterns.

On the other hand, notwithstanding the turmoil and the strife, the regulatory effort has made a difference.[85] The numbers of minority students and women students who have come into the higher education system within the past decade speak for themselves. Black high school graduates now are attending four-year colleges at a rate higher than their proportion in the population. Also, enrollments of women continue to increase. The aim of these laws and regulations is to achieve equality of opportunity in educational organizations as well as in other sectors. The real issue and the hard question is: Is the social cost in terms of the time, money, and effort expended by all parties—regulatory, regulated, and client group—worth the result? Many will agree that it is. Times are changing, social values are changing, and academia is changing. Many others will agree that some governmental regulatory efforts are necessary but will question whether the present methods and procedures are necessary or appropriate for institutions of higher education. And, of course, there are a few at either end of the spectrum who will argue that the present procedures are not enough or, in opposition, that any regulation of higher education is inappropriate and unnecessary.

Not too surprisingly, many members of the client group question the appropriateness of some regulatory efforts within the academic setting. Notwithstanding occasional intimations to the contrary, in terms of academic employment, evidence does not support the notion that members of the client groups want to enter the academic ranks at the expense of academic quality. Although some may, and do, question how that quality is defined, minorities, women, and others do not want to discard the concept of merit.[86] Rather, they and the public want to be assured that academic badges are honorably awarded, and that the academic processes are not tainted by illegal discrimination.

The present regulatory machinery does not serve either dissatisfied academics or academia well.[87] Whatever the outcome, the scars left by the present process may be as onerous as the condemned discriminatory behavior. Faculty status is not won in the courtroom. Remedies which compel employment or tenure ignore the sociology of the collegial departmental structure. Remedies short of compelling employment, such

---

[85] See Making Affirmative Action Work in Higher Education, supra note 76, pp. 53–54; "Climbing the Academic Ladder: Doctoral Women Scientists in Academia," Comm. on the Education and Employment of Women in Science and Engineering. Comm. on Human Resources, National Research Council, National Academy of Science 1979, Chapters 3 and 4; Patricia A. Graham, "Expansion and Exclusion: A History of Women in American Higher Education," 3 Signs 759 (Summer 1978), pp. 722–723; Special Subcommittee on Education of the Committee on Education and Labor, U.S. House of Representatives, Hearings on Federal Higher Education Programs Institutional Eligibility, Civil Rights Obligations Hearings, Part 2B, 93rd Congress 2nd Sess. (1974), pp. 1271–1418.

[86] See Special Subcommittee on Education of the Committee on Education and Labor, Hearings on Federal Higher Education Programs, Institutional Eligibility, supra note 85 at 1244, statement of Lilli S. Horneg, Executive Director, Higher Education Resource Services, Brown University.

[87] In fact, few academic employment discrimination cases are actually "won" by the plaintiffs. Of 44 higher education cases reviewed, only seven resulted in court decisions that awarded plaintiff relief. See charts 1 and 2 following this chapter.

as back pay, may recoup financial loss, but litigiousness is not an attribute which enhances one's reputation in an academic discipline. An individual litigant may win the immediate battle at great emotional and economic cost, only to lose the academic war.[88]

It is the composite effect of all these factors—economic costs, emotional costs, political costs, organizational costs—which makes the impact of regulation upon higher education. Impact measured in these terms has a subjective quality which runs counter to the spirit of objective scientific inquiry. In the long run however, it will be these subjective evaluations of social cost and social benefit which will determine the direction of future policy. The policy goals of nondiscrimination and equal opportunity served by these regulations have great social importance, as evidenced by their repeated ratification in new legislation by Congress and the continued reinforcement of their purposes by the courts.

Many argue that it is the methods, not the goals, that are generating the rising tone of complaint. On the other hand, disagreement concerning the proper means to achieve ends may mask disagreement concerning the ends themselves. These conflicts reflect a more basic American dilemma: The tension between the ideal of equality of opportunity and the reality of inequality of condition; the tension between traditional individualism and a new egalitarianism.[89] In the current controversy surrounding the regulation of higher education, the basic dilemmas arising from these dual values are thrown into sharp focus.

The problems created by this underlying dilemma defy easy resolution. They are an essential component of our political system. It is not unreasonable, however, to expect that public policies concerning common goals achieve a higher level of administrative and procedural rationality than they exhibit. Regulatory reform efforts directed to these ends can be useful.

Much of the current literature concerning regulatory reform advocates greater reliance on tax incentives and other market-oriented devices as regulatory mechanisms, in lieu of administrative setting of mandatory standards.[90] Here, too, the higher education sector raises unusual issues. These regulatory alternatives are not as relevant to nonprofit and tax-exempt organizations.

Regulation to achieve public goals will not fade away. It is a fact of our political existence. Critical analysis of the current regulatory system can assist administrators, public and private, to achieve the regulatory goals.

Bibliography

Books and Papers

Abramson, Joan. *The Invisible Woman.* San Francisco: Jossey-Bass Publishers, 1975.

Allison, G. *Essence of Decision: Explaining the Cuban Missile Crisis.* Boston, Little, Brown and Company, 1971.

[88] For a good picture of the tensions created for grievants in academia see Joan Abramson, "The Invisible Woman." (San Francisco: Jossey-Bass Publishers, 1975).

[89] See T. R. Pole, "The Pursuit of Equality in American History," (Berkeley, University of California 1978), pp. 119, 348–494, 352–358.

[90] See Charles L. Schultze, "The Public Use of Private Interest," supra note 1.

American Bar Association Commission on Law and the Economy. *Federal Regulation: Roads to Reform* (Exposure Draft). Washington, D.C.: American Bar Association, 1978.

Berger, R. *Government by Judiciary.* The Transformation of the Fourteenth Amendment. Cambridge; Harvard University 1977.

Bickel, A. *The Morality of Consent.* New Haven: Yale Univ. 1975.

Carnegie Council on Policy Studies in Higher Education. *Making Affirmative Action Work in Higher Education.* San Francisco: Jossey-Bass Publishers, 1975.

Clark, . "The Dossier on Colleges and Universities," from Wheeler, *On Record: Files and Dossiers on American Life.* 1969.

Cohen, M. D. and March, J. G. *Leadership and Ambiguity.* Prepared for the Carnegie Commission on Higher Education. San Francisco, 1974.

Coleman, James S. *Equality of Educational Opportunity,* a report for the U.S. Department of Health, Education, and Welfare, Office of Education. Washington, D.C.: National Center for Education Statistics, 1966.

Copus, David. "Long-Term Problems in Title VII Enforcement" Internal Memorandum, Special Investigation and Conciliation Division, EEOC.

Davis, Kenneth Culp. *Administrative Law of the Seventies.* Rochester: The Lawyers Cooperative Publishing Co., 1976.

———, *1978 Supplement to Administrative Law Treatise.* San Diego: K. C. Davis Publishing Co., 1978.

Davis, K. C. and Borders-Patterson. . . . "Black Students in Predominantly White North Carolina Colleges and Universities." ETS Report — 2, 1973.

Deitch, Kenneth M. "Federal Financial Aid: General Issues, Specific Background Recommendations." A Working Paper for the Sloan Commission on Government and Higher Education. Cambridge, November 1978.

———. "Financial Aid: A Resource for Improving Educational Opportunities." A Working Paper for the Sloan Commisison on Government and Higher Education. Cambridge, March 1978.

Doermann, Humphrey. *Toward Equal Access.* New York: College Entrance Examination Board, 1978.

Finkin, Matthew W. *Federal Reliance on Educational Accreditation: The Scope of Administrative Discretion.* Washington, D.C.: The Council on Postsecondary Education, 1978.

Freedman, James O. *Crisis and Legitimacy: The Administrative Process and American Government.* Cambridge: Cambridge University Press, 1978.

Fuller, L. *The Morality of Law.* New Haven: Yale Univ., 1964.

Gruson, Edward S. "The National Politics of Higher Education." A Working Paper for the Sloan Commission on Government and Higher Education. Cambridge, November 1977.

Heclo, H. *A Government of Strangers: Executive Politics in Washington,* Washington, D.C. Brookings Institution, 1977.

Hobbs, Walter C. *Government Regulation of Higher Education.* Cambridge: Ballinger Publishers Co., 1978.

Hofstadter, Richard. *The Paranoid Style in American Politics.* New York: Vintage Books, 1967.

Horowitz, D. *The Courts and Social Policy.* Washington, D.C.: Brookings Institution, 1977.

Horwitz, Morton J. *The Transformation of American Law, 1780–1860.* Cambridge: Harvard University Press, 1977.

Jencks, Christopher. *Inequality.* New York: Harper and Row, 1972.

Jencks, Christopher and Riesman, David. *The Academic Revolution.* Anchor Books; Garden City, New Jersey 1969.

Kaysen, Carl. "Government and Higher Education: The Significant Policy Issues in Their Context." Draft Paper for the Sloan Commission on Government and Higher Education. Cambridge, August 1978.

Kluger, S. *Simple Justice.* New York, Alfred A. Knoph, 1975.

Lee, E. C. and Bowen, F. M. *Managing Multicampus Systems.* Report for the Carnegie Commission on Policy Studies in Higher Education. San Francisco, 1975.

———. *The Multicampus University.* Report for the Carnegie Commission on Policy Studies in Higher Education. San Francisco, 1971.

Leslie, David W. "Legal Protection for Privacy" in *Protecting Individual Rights in Higher Education.* 1977.

Lester, Richard A. *Antibias Regulation of Universities.* New York: McGraw-Hill, 1974.
Levine, David O. "The Condition of Women in Higher Education." Paper for the Sloan Commission on Government and Higher Education. Cambridge, January 1979.
Lipsett, S. M. and Riesman, D. *Education and Politics at Harvard.* Prepared for the Carnegie Commission on Higher Education. San Francisco, 1975.
Lloyd, C. "Adams v. Califano: A Case Study in the Politics of Regulation." A Working Paper for the Sloan Commission on Government and Higher Education. Cambridge, 1978.
Lowi, Theodore J. *The End of Liberalism.* New York: W. W. Norton & Co. Inc., 1969.
Maine, H. *Ancient Law.* 1861. Dent Edition, 1954.
Mingle, James R. *Black Enrollment in Higher Education: Trends in the Nation and the South.* Atlanta: Southern Regional Education Board, 1978.
Mosteller, Frederick and Moynihan, Daniel Patrick. *On Equality of Educational Opportunity.* New York: Random House, 1972.
National Center for Education Statistics. *The Condition of Education.* Washington, D.C., 1978.
National Center for Education Statistics. *Projections of Education Statistics to 1968–1987.* Washington, D.C., 1978.
———. *Salaries, Tenure and Fringe Benefits of Full-Time Instructional Faculty in Institutions of Higher Education, 1977–78.* Washington, D.C.
Neustadt, R. *Presidential Power.* New York, John Wiley and Sons, 1960.
Noll, Roger G. *Reforming Regulation.* Washington, D.C.: The Brookings Institution, 1971.
Peters, B. Guy. *The Politics of Bureaucracy: A Comparative Perspective.* New York, Longman, Inc., 1978.
Purcell, Edward A., Jr. *The Crisis of Democratic Theory.* Lexington: The University Press of Kentucky, 1973.
Radner, Roy and Kuh, Charlotte, *Preserving a Lost Generation.* San Francisco: Carnegie Council on Policy Studies in Higher Education, 1978.
Rawls, John. *A Theory of Justice.* Cambridge: Harvard University Press/Belknap Press, 1971.
Schultze, Charles L. *The Public Use of Private Interest.* Washington, D.C.: The Brookings Institution, 1977.
Sennett, R. and Cobb, . *The Hidden Injuries of Class.* New York: Vintage, 1973.
Shulman, Carol Herrnstadt. *Compliance with Federal Regulations: At What Cost?* Washington, D.C.: American Association for Higher Education, 1978.
Spero, Irene K. "Government and Higher Education, A Summary of 21 Institutional Self-Studies." Paper for the Sloan Commission on Government and Higher Education. Cambridge, January 1978.
Task Force on Civil Rights Reorganization, President's Reorganization Project. "Reorganization of Equal Employment Opportunity Programs." Background and Option Papers, 1977.
Unger, Roberto M. *Law in Modern Society.* New York: The Free Press, 1976.
U.S. Department of Labor. *Employment and Training Report of the President.* Washington, D.C., 1977.
U.S. Department of Labor, Office of Federal Contract Compliance Programs Task Force. *A Preliminary Report on the Revitalization of the Federal Contract Compliance Program.* Washington, D.C., September 16, 1977.
U.S. Government, *The Budget of the United States Government,* Fiscal Year 1979.
U.S. House of Representatives, Special Subcommittee on Education and the Committee on Education and Labor. *Hearings on Federal Higher Education Programs Institutional Eligibility.* Civil Rights Obligations Hearings, Part 2B, 93rd Congress, 2nd session, 1974.
Van Alstyne, C. and Coldren S. *The Costs of Implementing Federally Mandated Social Programs at Colleges and Universities.* Policy Analysis Service—Special Report. Washington, D.C.: American Council on Education, June 1976.
Vetter, Jan. "Affirmative Action in Faculty Employment Under Executive Order 11246." A Draft Report for the Committee on Grant and Benefit Programs of the Administrative Conference of the United States, September 1974.
Yale University. *The Report of the President, Yale University: 1974–75.*

Articles

Belton, Robert. *A Comparative Review of Public and Private Enforcement of Title VII of the Civil Rights Act of 1964.* 31 Vanderbilt L. R. 905.

Bundy, McGeorge. "The Issue Before the Court: Who Gets Ahead in America?" *Atlantic Monthly,* Vol. 240, No. 5., November 1977.

Chayes, Abram. *The Role of the Judge in Public Law Litigation.* 89 Har. L. R. 1281 (1976).

Clark, "Discrimination Suits: A Unique Settlement." *Educational Record,* Summer 1977.

Coleman, James S. Letter to the editor. *The Chronicle of Higher Education.* Vol. XVII, No. 15. December 11, 1978.

Comment. *Implication of a Private Right of Action Under Title IX of the Education Amendments of 1972.* 73 Nw. U. L.R. 772 (1978).

——— *Private Rights of Action Under Title IX.* 13 Har. CR–CL L.R. 425 (1978).

Cuddy, Robert W. *Age Discrimination Amendments and Their Impact on Personnel.* 4 Emp. Rel. L.J. 339 (1978).

Davis, Kenneth Culp. *Informal Administrative Action: Another View.* 26 Am. U. L.R. 836 (1977).

Dentler, R. "Improving Public Education: The Boston School Desegregation Case." *The Advocate,* Suffolk University Law School. 1975.

Evans and Novak. *Washington Post,* August 29, 1977.

Finkin, Matthew W. *The Supervisory Status of Professional Employees.* 54 Fordham. L.R. 805 (1977).

Finn, Chester E., Jr. "Federalism and the Universities: The Balance Shifts." *Change.* Vol. 7, No. 10. December–January, 1975–76.

——— "Federal Patronage of Universities in the United States: A Rose by Many Other Names?" *Minerva.* Vol. XIV, No. 4. Winter 1976–77.

Frankel, Marvin E. *The Search for Truth: An Empirical View.* 123 U. Penn. L. R. 1031 (1975).

Freeman, Richard B. *The New Job Market for Black Academicians.* 30 Ind. Lab. R. 161, 168–169 (1977).

Friendly, Henry J. *Some Kind of Hearing.* 123 U. Penn. L.R. 1267 (1975).

Frug, Gerald E. *The Judicial Power of the Purse.* 126 U.Pn. L.R. 715 (1978).

Gardner, Warner W. *The Informal Actions of the Federal Government.* 26 Am. U. L.R. 799 (1977).

Greenfield, Ester. *From Equal Pay to Equivalent Pay: Salary Discrimination in Academia,* 6 J. of Law and Education 41, 42 (1977).

Haslam, C. L. *Age Discrimination in Campus Employment.* 2 J. Coll. and Univ. Law 326 (1975).

Hunt, Thomas Reed and Pazuniak, George. *Special Problems in Litigating Upper Level Employment Discrimination Cases.* 4 Del. J. of Cor. L. 114 (1978).

Jaffe, L. *Law Making by Private Groups.* 51 Har. L.R. 201 (1937).

Johnson, George E. and Stafford, Frank P. *The Earnings and Promotion of Women Faculty.* 64 Am. Econ. Rev. 888 (1974).

Kahn, Kenneth. *The NLRB and Higher Education: The Failure of Policymaking Through Adjudication.* 21 U.C.L.A. 63 (1973).

Kennedy, Duncan. *Form and Content.* 2 J. of Law and Contemporary Society. 1973.

Lamber, Julia. *Private Causes of Action Under Federal Agency Nondiscrimination Statutes.* 10 Conn. L.R. 859 (1978).

Larner, Barbara. *Washington* v. *Davis: Quantity, Quality and Equality in Employment Testing.* The Supreme Court Review. 1976.

Levinson, L. Harold. *Elements of the Administrative Process: Formal, Semiformal and Free-Form Models.* 26 Am. U. L.R. 872 (1977).

Lindblum, Charles E. *The Science of Muddling Through.* 19 Pub. Ad. Rev. 79. 86 (1959).

Lindsey, Robert. "U.S. Hispanic Populace Growing Faster Than Any Other Minority." *The New York Times.* February 18, 1979.

McGowen, Carl. *Congress, Court and Control of Delegated Power.* 77 Colum. L.R. 1119 (1977).

Middleton, Lorenzo. "The Effects of School Desegregation: The Debate Goes On." *The Chronicle of Higher Education.* Vol. XVII, No. 10. November 6, 1978.

Moynihan, Daniel Patrick. "Social Science and the Courts." *Public Interest.* 12, 21. Winter 1978.

Oaks, Dallin H. *A Private University Looks at Government Regulation.* 4 J. Coll. and Univ. Law 1 (1977).
Rabin, Robert L. *Job Security and Due Process: Monitoring, Administrative Discretion Through a Reasons Requirement.* 44 U. Chi. L.R. 60 (1976).
Reich, D. *The New Property.* 83 Yale L.J. 733 (1964).
Rosenblatt, Rand E. *Health Care Reform and Administrative Law: A Structural Approach.* 88 Yale L.J. 243 (1978).
Scott, Robert A. "The Hidden Costs of Government Regulation." *Change.* Vol. 10, No. 4. April 1978.
Steward, Richard. *The Reformation of American Administrative Law.* 88 Har. L.R. 1667 (1975).
Tolchin, Martin. "Battle Intensifies Over Authority of President to Control Agencies." *The New York Times.* January 17, 1979.
Tribe, Laurence H. *Structural Due Process.* 10 Har. CR-CL L.R. 269 (1975).
Van Alstyne, William. *Cracks in "The New Property": Adjudicative Due Process in the Administrative State.* 62 Cornell L.R. 445 (1977).
Verkvil, Paul R. *A Study of Informal Adjudication Procedures.* 43 U. Chi. L.R. 739 (1976).
——— *The Emerging Concept of Administrative Procedure.* 78 Colum. L.R. 258 (1978).
——— *The Ombudsman and the Limits of the Adversary System.* 75 Colum. L.R. 845 (1975).
William, W. *Higher Education and Minority Opportunities.* Howard L.J. 545, 552 (1978).

Appendix A.—A Lay Guide to Equal Opportunity-Nondiscrimination Provisions

The equal opportunity/nondiscrimination laws mentioned at various points in this paper and analyzed in the charts of Appendix A are briefly described below in the least technical jargon a lawyer can muster:

Title VII.—(42 U.S.C. § 2000[e] et seq.) Part of the Civil Rights Act of 1964, this title deals with discrimination in employment. It prohibits discrimination in nearly every aspect of employment, from the decision of whom to hire to the salary paid, fringe benefits granted, and working conditions provided. Originally, educational institutions were exempt from the requirements of Title VII; but an amendment in 1972 eliminated the exemption, and now educational institutions are subject to the provisions of the Act.

Under Title VII, discrimination based on race, color, religion, national origin, and sex is declared illegal.

Title VI.—(42 U.S.C. § 2000[d] et seq.) Another part of the Civil Rights Act of 1964, this title focuses on discrimination in programs supported by Federal funds. In contrast to the employment provisions of Title VII, Title VI excludes most employment issues and trains the regulatory eye on the benefits provided to the recipients of the program. Either segregation (or other separate treatment) or unequal distribution of benefits (ranging from the opportunity to receive benefits at all to the level and character of those benefits) may constitute a violation of Title VI.

Discrimination on the basis of race, color, and national origin is prohibited by Title VI.

Title IX.—(20 U.S.C. § 1681 et seq.) This title was part of the Education Amendments of 1972 and is limited to sex discrimination[1] in *educational* programs receiving Federal funds. As a programmatic statute modeled after Title VI of the Civil Rights Act of 1964, Title IX is generally considered to apply to discrimination in educational programs, not employment. Like Title VI, Title IX prohibits denial of or differential in benefits, but in this instance based on sex. Admissions decisions at all undergraduate private institutions and single-sex public institutions, however, are exempted from this ban on discrimination.

Title IX applies only to sex discrimination and only to educational institutions.

Executive Order 11246.—(30 Fed. Reg. 12319 [1965]) A different provision from the three outlined above, the Executive Order was promulgated by the President in 1965 and requires that, in addition to not discriminating, all employers who contract with the government take "affirmative action" to ensure proper representation of minorities and women (the latter added in 1967) in the

[1] Title IX also prohibits discrimination on the basis of blindness, but this form of discrimination is now subject to the more comprehensive provisions of Section 504.

employer's work force. The Executive Order, then, is an employment (as opposed to a programmatic) directive which arises out of the contract relationship between government and contractor.

The Executive Order applies to race, color, religion, national origin, and sex.

Equal Pay Act.—(29 U.S.C. § 206(d)) A relatively specific employment provision, this 1963 legislation does not speak in terms of discrimination; rather, it requires equal pay for equal work and bans any salary distinction for performance of similar work duties based on sex. Although it applies to all employers covered by the Fair Labor Standards Act, professional employees were not covered by the Act until 1972.

Again, the Act makes sex-based salary distinctions illegal.

Section 503.—(29 U.S.C. § 793) Part of the Rehabilitation Act of 1973, this section parallels Executive Order 11246 in requiring affirmative action by employers with government contracts to hire and promote handicapped persons.

Section 504.—(29 U.S.C. § 794) Another section of the Rehabilitation Act of 1973, this parallels the programmatic nature of Title VI. Section 504 prohibits discrimination on the basis of handicap in all federally assisted educational programs.

Architectural Barriers Act.—(42 U.S.C. § 4151 et seq.) Together with Section 502 of the Rehabilitation Act of 1973 (29 U.S.C. § 792), this 1968 Act requires that buildings financed by Federal grants or loans be accessible to the handicapped.

Age Discrimination in Employment Act (ADEA).—29 U.S.C. § 621 et seq.) This 1967 Act prohibits employment discrimination against persons 40–70 years old. including mandatory retirement programs. It was not until 1974 that the Act was extended to public sector employees and only in 1978 that the Act's coverage extended to employees 65–70 years olds.

Age Discrimination Act (ADA).—(42 U.S.C. § 6101 et seq.) Another programmatic provision modeled after Title VI, this 1975 Act prohibits discrimination on the basis of age (not limited to persons 40–70 years old) in any educational program receiving Federal funds or assistance.

Vietnam Era Veterans Readjustment Assistance Act.—(38 U.S.C. § 2012 et seq.) This 1974 employment provision requires affirmative action to employ and promote Vietnam veterans by government contractors. Modeled after Executive Order 11246, the affirmative action requirement is an element of the employer's government contract.

Fair Housing Act.—(42 U.S.C. § 3601 et seq.) Passed in 1968, this act prohibits housing discrimination based on race, national origin, religion, and sex in housing financed with Federal aid or by state and local agencies receiving Federal assistance.

Sections 799A and 845.—(42 U.S.C. § 16) These sections of the Public Health Service Act prohibit sex discrimination in admissions by schools of medicine, schools of nursing, and schools in other health-related fields receiving Federal assistance under the Act.

In addition to the laws noted above, there are several other laws which have been subject to a century of judicial interpretation pertaining to present equal opportunity/nondiscrimination efforts. It is best simply to quote them in pertinent part:

14th Amendment.—(U.S. Const. Amend XIV § 2) "No State shall . . . deny to any person within its jurisdiction the equal protection of the laws."

42 U.S.C. § 1981.—"All persons . . . have the same right . . . to make and enforce contracts . . . and to the full and equal benefit of all laws and proceedings for the security of persons and property as is enjoyed by white citizens, and are subject to like punishment, pains, penalties, taxes, licenses and exactions of every kind and to no other."

42 U.S.C. § 1983.—"Every person who, under color of any statute, ordinance, regulation, custom or usage, of any State or Territory, subjects, or causes to be subjected, any citizen of the United States . . . to the deprivation of any rights, privileges, or immunities secured by the Constitution and laws, shall be liable to the party injured in action at law, suit in equity, or other proper proceeding for redress."

Appendix B

Procedural and Jurisdictional Chart of the Equal Opportunity-Nondiscrimination Laws

	Age	Color	Handicap	National origin	Race	Religion	Sex	Veteran	Employment
Title VII		X		X	X	X	X		X
Title VI		X		X	X				X
Title IX			(1)				X		(7)
11246		X		X	X	X	X		X
Equal Pay Act							X		X
§ 503			X						X
§ 504			X						X
Architectural barriers and § 502			X						
Age discrimination in employment	X								X
Age Discrimination Act	X								X
Vietnam Era Veterans Act								X	X
§ 1981		X			X				X
§ 1983	(7)	X	X	X	X	X	X		X
5th Amendment and 14th Amendment		(2)	(2)	(2)	(2)	(2)			X
Fair Housing				X	X	X	X		
§§ 799A and 845 of the Public Health Service Act							X		X

[1] Blindness only.

[2] Could be any with varying levels of judicial scrutiny.

Procedural Regulations of the Antidiscrimination Statutes

Statute [1]	Administration agency	Prehearing conciliation	Agency remedies after hearing	Agency action on behalf of complainant	Private cause of action?	Exhaustion of administration remedies required?	Court remedies	Special regulations for higher education
Title VII 42 U.S.C. § 2000(e)-(o)17 1972.	EEOC	Yes, if EE-OC finds reasonable cause after investigation.	None. (No final hearing authority). Referral to DOJ for public institutions.	EEOC may sue on behalf of complainant. EEOC may initiate pattern and practice suit.	Yes	Yes [2]	General equitable remedies. [3] (hereinafter G.E.R.).	Recordkeeping regulations, 29 C.F.R. §§ 1602.47-1605.55.
Equal Pay Act 29 U.S.C. § 206 (d) 1966/1972. [4]	EEOC [5]	Not required; most disputes settled voluntarily.	None	EEOC may sue on behalf of complainant.	Yes, if EEOC does not file suit.	No	G.E.R./fines and imprisonment after criminal prosecution for willful violations.	None.
Age Discrimination in Employment Act § 9 U.S.C. 621-634 1967/1974. [6]	EEOC [5]	Required before EEOC files suit.	do	do	do	No, however individual must give EEOC 60 days notice before filing suit.	G.E.R	Do.
Executive Order 11246, as amended 1965.	DOC/OFCCP	Encouraged, but not mandatory. OFCCP refers all *individual* complaints to EEOC for processing in accordance with EEOC, Title VII procedures.	Cancel, terminate, or suspend contracts; debarment of contract from further contracts.	May refer to Department of Justice for suit. The Department of Iustice may also initiate suit after an independent investigation	No, but an individual may sue to compel agency enforcement proceeding.	NA	Department of Justice may sue for G.E.R. Also, contracts may have judicial review of agency decisions.	No, however, 3 memoranda by OCR in 1974 and 1975 are still in effect.
§ 503 of the Rehabilitation Act of 1973. 29 U.S.C. § 793, 1973.	DOL/OFCCP	Yes	Termination or suspension of the contract; debarment from future contracts.	Director of OFCCP may seek judicial action.	No, but an individual may sue to compel agency to initiate enforcement proceeding.	NA	GER, judicial review of agency decisions to impose sanctions on contractors.	None.
Vietnam Era Veterans Readjustment Assistance Act 38 U.S.C. § 2012, 1974.	DOL/OFCCP	Yes	Termination of contract; withholding of progress payments to contractor; debarment from future contracts.	May refer to Department of Justice for suit.	No, but an individual may sue to compel agency to initiate enforcement proceedings.	NA	G.E.R., contractors may have judicial review of unfavorable agency decisions.	Do.

Title VI 42 U.S.C. § 2000(d)-(d)6 1964.	HEW/OCR [7]	Yes	Termination of funds	do	Yes [8]	No	G.E.R., judicial review of final agency decisions.	Do.
Title IX 20 U.S.C. §§ 1681–1686, 1972.	HEW/OCR [7]	Yes	do	do	Yes [8]	No	do	45 C.F.R. §§ 86.21–86.23.
§ 504 of the Rehabilitation Act of 1973. 29 U.S.C. § 794, 1973.	HEW/OCR [7]	Yes	do	do	Yes [8]	No	do	45 C.F.R. §§ 86.41–86.47.
Age Discrimination Act 42 U.S.C. § 6101–6107, 1975.	HEW/OCR [7]	Yes, the regulations also require mediation.	do	do	Yes	Yes [8]	do	None.
§§ 799A–845 of the Public Health Service Act 42 U.S.C. § 216, 1971.	HEW/OCR [7]	Yes	do	do	Yes [8]	No [8]	do	Applies only to medical and nursing schools. 45 CFR Part 83.
Fair Housing Act 42 U.S.C. §§ 3601–3619, 1968.	HUD	Yes	NA	Yes. The Attorney General may institute a pattern and practice suit for equitable relief.	Yes	No	G.E.R.	None.
Architectural Barriers Act, 42 U.S.C. 4151–4156 (1968) and 502 of the Rehabilitation Act of 1973.	GSA. The Architectural and Transportation Barriers Compliance Board was created in 1973 to insure compliance with the act.	Yes	Withhold or suspend funds with respect to buildings not in compliance.	No	No	No	Judicial review of final board orders.	Do.

[1] Dates refer to year the statute was made applicable to higher educational organizations.

[2] Administrative remedies are exhausted if, after the expiration of 180 days from the date the complaint was filed, the agency has reached no final decision or has made a "no action" determination.

[3] General equitable remedies include backpay, and injunctive relief, i.e., reinstatement, hiring, altered seniority systems, training, or other remedial programs, recruitment, wage increases, award of attorney's fees, and other relief including continuing court oversight.

[4] The Equal Pay Act was applied to "hourly" or "nonexempt" employees of colleges and universities in 1967. In 1972, coverage under the act was extended to include administrative and faculty personnel of such organizations.

[5] The authority to enforce the Equal Pay Act and the Age Discrimination in Employment Act was transferred from the Department of Labor to the EEOC July 1, 1979. EEOC has said that it will follow the investigation and enforcement procedures previously established by the Wage and Hour Division of the Department of Labor. The interpretations of the act, however, which were previously issued by by the DOL are currently being reviewed by the EEOC.

[6] The ADEA has applied to private institutions since its passage in 1967. In 1974 the coverage of the act was extended to include public institutions.

[7] There will be an Office for Civil Rights within the recently created Department of Education that will take over the responsibility for enforcing the statute formerly exercised by OCR within the Department of Health, Education, and Welfare.

[8] The Supreme Court held in Cannon v. University of Chicago, 99 S. Ct. 1946 (1979) that individuals had a private right of action under title IX. In the light of Cannon, Regents of the University of California v. Bakke, 98 S. Ct. 2733 (1978), Lau v. Nichols, 414 U.S. 563 (1974), and Southeastern Community College v. Davis, 99 S. Ct. 2361 (1979), it seems likely that title IX and § 504 will be held to provide similar private rights of action.

Appendix C

Chart I.—Analysis of higher education employment discrimination complaints: Administrative agency and judicial action.

[This analysis is based on a review of 44 cases that involve allegations of race or sex discrimination in faculty or professional staff employment in colleges and universities. An alphabetical list of the actions reviewed in this analysis follows chart II.]

INDIVIDUAL OR CLASS ACTION

FORM OF DISCRIMINATION ALLEGED IN COMPLAINT

[4 complaints alleged both race and sex discrimination]

PRACTICE CHALLENGED

[Many plaintiffs challenged more than one practice]

Promotion (23)

Salary (15)

Tenure (15)

Non Reappointment (23)

Hiring (7)

Practices: 0 5 10 15 20 25 30

NUMBER OF STATUTES CLAIMED VIOLATED IN COMPLAINT

1 (17)

2 (10)

3 (9)

4 (4)

5 (2)

6 (2)

Statutes: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 17 18

NUMBER OF AGENCIES TO WHICH COMPLAINTS WERE MADE

Number of agencies	Count
0	(13)
1	(21)
2	(9)
3	(1)

Agencies: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

ADMINISTRATIVE AGENCIES

EEOC (23)

State or Local Agency (11)

HEW (2)

OFCCP (2)

DOL (1)

DOJ (4)

NUMBER OF COURTS AND AGENCIES WITH WHICH PLAINTIFF FILED A COMPLAINT

1 (10)

2 (18)

3 (7)

4 (8)

5 (1)

Number of Courts and Agencies: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

LATEST RULINGS

For Defendant Institution (24)

For Plaintiff (5)

Final Decision Pending (15)

CHART II.—Lengths of time elapsed between various stages of proceedings in cases involving allegations of race and sex discrimination by faculty members and professional staff (41 cases).

FIRST COMPLAINT WITHIN INSTITUTION TO FIRST COURT COMPLAINT

0 - 1

1 - 2

2 - 3

3 - 4

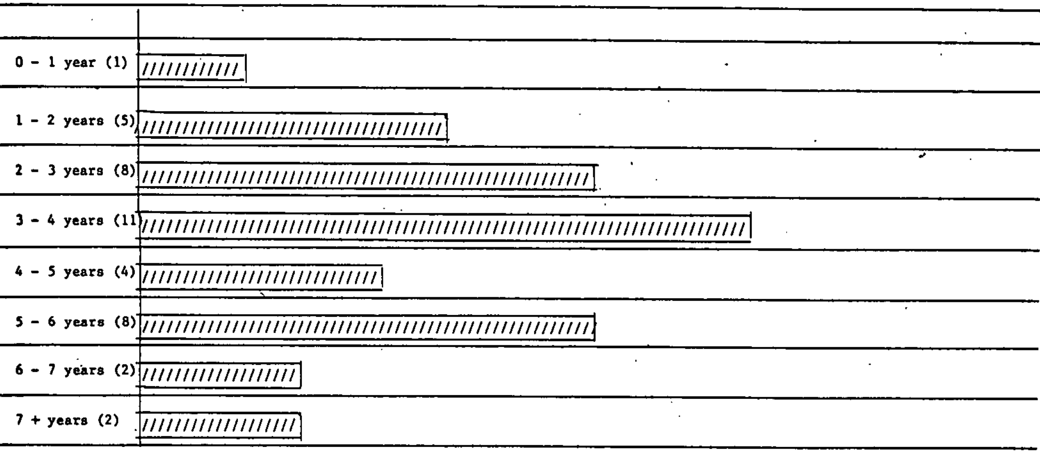
4 - 5

Years: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Mean length of time: 2 years, 1.24 months

INITIAL NOTICE OF GRIEVANCE WITHIN INSTITUTION TO FIRST COURT DECISION ON THE MERITS

[Mean length=3 years, 8.5 months]



INITIAL NOTICE OF GRIEVANCE WITHIN INSTITUTION TO LATEST COURT DECISION ON THE MERITS

[Mean length=4 years, 4.2 months]

Length	Number
0 - 1 year	(0)
1 - 2 years	(4)
2 - 3 years	(3)
3 - 4 years	(10)
4 - 5 years	(8)
5-6 years	(10)
6 - 7 years	(4)
7 + years	(2)

Years: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

INITIAL ADMINISTRATIVE COMPLAINT TO FIRST COURT COMPLAINT

0 - 1 year
(6)

1 - 2 years
(7)

2 - 3 years
(3)

3 - 4 years
(5)

Years: 0 1 2 3 4 5 6 7 8 9 10

Mean length of time: 1 year, 8.57 months

INITIAL AGENCY COMPLAINT TO FIRST COURT DECISION ON THE MERITS [1]

[Mean length=3 years, 6.9 months]

0–1 year (2)

1–2 years (2)

2–3 years (8)

3–4 years (5)

4–5 years (3)

5–6 years (4)

6–7 years (1)

7+ years (1)

Years: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

[1] Of the 15 cases not included in this graph and the following graph, in 5 there was no agency involvement. In the other 10, the date of the agency complaint is unknown.

INITIAL AGENCY COMPLAINT TO LATEST COURT DECISION ON THE MERITS

[Mean length=4 years, 2.4 months]

0-1 year (1)

1-2 years (2)

2-3 years (4)

3-4 years (4)

4-5 years (6)

5-6 years (6)

6-7 years (2)

7+ years (1)

Years: 0 1 2 3 4 5 6 7 8 9 10

56-368 O - 80 - 28

INITIAL COURT COMPLAINT TO FIRST COURT DECISION ON THE MERITS

[Mean length=1 year, 7.2 months]

Length	Cases
0-1 year	(7)
1-2 years	(14)
2-3 years	(4)
3-4 years	(3)
4-5 years	(1)
5-6 years	
6-7 years	
7+ years	

Years: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

INITIAL COURT COMPLAINT TO LATEST COURT DECISION ON THE MERITS

[Mean length=2 years, 3.4 months]

0-1 year (3)

1-2 years (11)

2-3 years (10)

3-4 years (1)

4-5 years (4)

5-6 years

6-7 years

7-8 years

Years: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

ALPHABETICAL LIST OF ALL FACULTY AND PROFESSIONAL EMPLOYMENT DISCRIMINATION CASES INVOLVING COLLEGES AND UNIVERSITIES

1. *Al-Hamdarii v. State University of New York*, 438 F. Supp. 299 (1977).
2. *Berry v. University of Texas*, HEW Pocket No. CC–10. (1971, 1979).
3. *Carrion v. Yeshiva University*, 535 F. 3d 722 (1976).
4. *Citron v. Jackson State University*, 456 F. Supp. 3 (1977).
5. *Clap v. Lehigh University*, 450 F. Supp. 460 (1978).
6. *Clark v. Atlanta University*, 15 FEP 1139 (1976).
7. *Craig v. Alabama State University*, 451 F. Supp. 1207 (1978).
8. *Cramer v. Virginia Commonwealth University*, 586 F. 2d 297 (1978).
9. *Cussler v. University of Maryland*, 430 F. Supp. 602 (1977).
10. *Davis v. Weidner*, 421 F. Supp. 594 (1976).
11. *Dyson v. Lavery*, 417 F. Supp. 103 (1976).
12. *Egelston v. State University College at Genesco*, (1976).
13. *EEOC v. Tufts Institution of Learning*, 421 F. Supp. 152 (1975).
14. *Faro v. New York University*, 502 F. 2d 1229 (1974).
15. *Green v. Board of Texas Tech University*, 474 F. 2d 594 (1973).
16. *Hanshaw v. Delaware Technical and Community College*, 405 F. Supp. 292 (1975).
17. *Hill v. Nettleton*, 455 F. Supp. 514 (1978).
18. *Huang v. College of the Holy Cross*, 436, F. Supp. 639 (1977).
19. *Jawa v. Fayetteville State University*, 426 F. Supp. 218 (1978).
20. *Johnson v. University of Pittsburg*, 435, F. Supp. 1328 (1977).
21. *Keyes v. Lenoir Rhyne College*, 552 F. 2d 579 (1977).
22. *Kunda v. Muhlenburg College*, 463 F. Supp. 294 (1978).
23. *Labat v. Board of Higher Education of the City of New York*, 401 F. Supp. 753 (1975).
24. *Lewis v. Chicago State College*, 299 F. Supp. 1357 (1969).
25. *Mecklenburg v. Montana State Board of Regents*, 13 FEP 462 (1976).
26. *Moore v. Kibbie*, 400 F. Supp. 1367 (1975).
27. *Morpurgo v. Board of Higher Education of the City of New York*, 400 F. Supp. 1135 (1977).
28. *Mosby v. Webster College*, 563 F. 2d 901 (1977).
29. *Pace College v. Commission on Human Rights*, 38 NY 2d 28 (1975).
30. *Pendrell v. Chatham College*, 370 F. Supp. 494 (1974).
31. *Perham v. Ladd*, 436 F. Supp. 1101 (1977).
32. *Peters v. Middlebury College*, 409 F. Supp. 857 (1976).
33. *Power v. Syracuse University*, 580 F. 2d 1150 (1978).
34. *Presseisen v. Swarthmore College*, 442 F. Supp. 593 (1977).
35. *Rackin v. University of Pennsylvania*, 386 F. Supp. 992 (1974).
36. *Scott v. University of Delaware*, 455 F. Supp 1102 (1978).
37. *Sime v. Trustees of California State University and Colleges*, 526 F. 2d 1112 (1975).
38. *Smith College v. MCAD*, 380 NE 2d 121 (1978).
39. *Stevens v. Junior College District of St. Louis, St. Louis County*, 548 F. 2d 779 (1977).
40. *Sweeney v. Board of Trustees of Keene State College*, 569 F. Supp 169, (further proceedings December 1979).
41. *United States v. University of Maryland*, 438 F. Supp 742 (1977).
42. *Wagner v. Long Island University*, 419 F. Supp 618 (1976).
43. *Wheelock College v. MCAD*, 355 NE 2d 309 (1976).
44. *Van de Vate v. Bolling*, 379 F. Supp. 925 (1974).

SOCIALLY IMPOSED COSTS OF HIGHER EDUCATION [1a]

By Howard R. Bowen*

CONTENTS

	Page
Conclusions	432
Socially Imposed Costs as a Widespread Phenomenon	433
Some Conceptual Issues	434
The Special Position of Higher Education	438
Specific Sources of Socially Imposed Costs in Higher Education	439
Estimated Dollar Amounts of Socially Imposed Costs Affecting Higher Education	450
Appendix	454
Bibliography	456

During the past quarter century, institutional expenditures for higher education in the United States have increased at a prodigious rate. (See Table 1.) Total current expenditures grew from $2.2 billion in 1949–50 to an estimated $41.9 billion in 1975–76—a 19-fold increase. Part of this remarkable growth was due to an explosion of enrollment and part was due to price inflation.[1b] But even after adjustment for these two factors, expenditures per student in constant dollars nearly doubled over the 26-year period, increasing at the average rate of about 2½ percent a year. Though the rate of growth has been slower in recent years than in the 1950's and 1960's, it has still been substantial.

There are various reasons for the rapid growth of expenditures. Some critics allege that the higher educational community expanded its functions and its work load unnecessarily, that it has not been adequately attentive to efficiency and cost control, and that it has been able to indulge these follies through access to funds beyond its essential needs. Others, more friendly toward higher education, argue that the increased expenditures were necessary to bring about much-needed improvements in quality of instruction and research; or they argue that the increase expenditures were thrust upon higher education by unavoidable increases in wages, salaries, and other costs that could not possibly have been offset by improvements in productivity. Doubtless there is substance to all of these explanations.

---

*Sloan Commission on Government and Higher Education.

[1a] Paper given in shortened form as a Henry Lecture at the University of Illinois, 1979.
[1b] Later in this paper, the question of the meaning and measurement of the rate of inflation will be considered. Some of what is ordinarily called inflation may have been due to the very socially imposed costs that are under review in this paper.

TABLE 1.—GROWTH OF CURRENT EXPENDITURES IN ALL COLLEGES AND UNIVERSITIES, UNITED STATES, 1950 TO 1976

	Dollar amounts				Average annual percentage increases compounded (approximate)			
	1949–50	1959–60	1969–70	1975–76	1949–50 to 1959–60	1959–60 to 1969–70	1969–70 to 1975–76	1949–50 to 1975–76
Total current expenditures:[1]								
Currentdo llars (billions)[2]	2.2	5.6	21.1	41.9	10.0	14.0	12.0	12.0
Constant 1975–76 dollars (billions)[3]	5.1	10.6	31.0	41.9	7.5	11.0	5.25	8.5
Constant 1975–76 dollars per student	2,333	3,820	4,906	5,173	2.5	2.75	1.6	2.5
As percent of GNP	0.8	1.1	2.2	2.6				
Total educational and general expenditures:[4]								
Current dollars (billions)[2]	1.7	4.5	15.8	32.7	10.25	13.0	13.0	12.5
Constant 1975–76 dollars (billions)[3]	3.9	8.5	23.2	32.7	7.75	10.5	6.0	8.5
Constant 1975–76 dollars per student	1,773	3,063	3,671	4,037	3.25	2.4	2.25	2.5
As percent of GNP	0.6	0.9	1.6	2.0				
Total enrollment (millions)[5]	2.2	2.8	6.2	8.1	2.5	8.0	2.5	3.5
Consumer price index (1975–76=100)[6]	43.2	53.0	68.1	100.0	2.0	2.5	4.0	3.0
Gross National Product (billions of current dollars)[7]	272.1	496.3	958.9	1,607.0	6.0	7.0	5.25	7.0
Gross National Product (billions of 1975–76 dollars)	629.9	936.4	1,408.1	1,607.0	4.0	4.0	1.4	3.5

[1] Total current expenditures include outlays for institutional operations plus those for student aid and auxiliary enterprises.

[2] Sources: Bureau of the Census, U.S. Department of Commerce, Historical Statistics of the United States, Washington, D.C.: U.S. Government Printing Office, 1975, Vol. I, p. 384. National Center for Educational Statistics, U.S. Department of Health, Education, and Welfare, Projections of Education Statistics to 1984–85, Washington, D.C.: U.S. Government Printing Office, 1976, p. 97.

[3] Deflated on the basis of the U.S. Department of Labor Consumer Price Index.

[4] Excludes outlays for student aid and auxiliary enterprises.

[5] Refers to full-time equivalent enrollment in the fall of each year.

Source: National Center for Education Statistics, U.S. Department of Health, Education, and Welfare, 1967, p. 23; 1976, p. 29. Figure for 1949–50 estimated by the author on the basis of total enrollment.

[6] U.S. Department of Labor.

[7] U.S. Department of Commerce.

In addition, another explanation of the growth in expenditures is that they were induced—in part at least—by social demands for new services, new activities, and new standards of operation. Such demands were an outgrowth of changing social values and expectations often leading to reenforcement through government laws and regulations. It is widely asserted that they imposed heavy costs for substantive compliance and also for onerous reports and redtape connected with government intervention. In this paper, the concern is primarily with the growth of higher educational expenditures attributable to the new social demands. The focus will be on socially imposed costs. The main objective will be to identify the sources of these costs, to estimate their amounts, and to judge their impacts on higher education. These socially imposed costs are of special concern to educators who fear that social pressures and government intervention will erode institutional autonomy and thus impair academic freedom. A spirited discussion of socially imposed costs occurred in higher educational circles in recent years, a considerable though fragmented literature appeared, and some partial and preliminary cost estimates were made. The Federal Government has responded with some initial efforts to assess the situation and to seek solutions.

CONCLUSIONS

All higher educational costs are socially imposed in the sense that all are in response to the needs or demands of society. The question of which ones are to be selected out and labeled socially imposed or governmentally mandated is wholly arbitrary. It is asking only: What is society expecting higher education to do today that it was not doing last year, 10 years ago, or 25 years ago? Or what is higher education voluntarily doing today, in the public interest as perceived by educators, that it was not doing before? With this all-inclusive definition of socially imposed cost, the dollar amount over a given time could be computed merely by comparing total expenditures at the beginning of the selected period with those at the end of the period. Even the rising costs due to inflation could be said to be socially imposed. Thus, efforts to calculate the amount of expenditures associated with socially imposed costs is not a hugely rewarding activity. The basic issues relating to socially imposed costs are philosophical or political.

One issue is whether some of the socially imposed programs or activities are directed toward improper objectives. Is higher education being asked to engage in socially harmful or socially useless activities? The literature suggests that there is little complaint among educators about the objectives of most of the socially imposed activities. Hardly anyone opposes efforts to improve personal security, equality of opportunity and access, or environmental conditions.

A second issue pertains to the appropriateness of the procedures and the skill and efficiency with which government programs are administered. In this connection, there is much criticism of clumsy administration, lack of understanding of the academic community, arbitrariness, tactlessness, redundancy, and inefficiency. There is also complaint about the large number of new programs being imposed over so short a period that cannot be readily assimilated.

A third issue pertains to finances. The plea is often made that the Government should finance the additional costs it foists upon higher education. This plea, however, is simply part of the ongoing debate over the question: How much money should higher education receive? If society (represented by legislative bodies and donors) believes that higher education gets more money than it needs, then the imposition of new costs does not necessarily call for increased appropriations and gifts, but only for rearranged priorities within given levels of expenditures. On the other hand, if it is believed that higher education is impoverished, then the new costs will call for comparable increases in revenues. The question of financing socially imposed costs is thus simply a subtopic under the general questions of the proper level of financing for higher education and of the distribution of the burden among the Federal Government, State governments, donors, and students. Included within this issue is the question of the relative difficulty that many institutions of higher education face in shifting socially imposed costs forward.

A fourth issue relates to governmental control and academic freedom. Many educators believe that the proliferation of social imposed activities and governmentally mandated programs represents a threat to the kind of institutional autonomy which is the foundation of aca-

demic freedom. This view leads to the recommendation that desirable ends sought by society should be achieved through means that would bear down less heavily on institutional autonomy. In the end, however, this view is reduced to the question of appropriateness of the procedures and the skill of administration; that is, to the question of efficiency broadly conceived.

The objectives being sought are, on the whole, laudable; the execution is often clumsy and wasteful; more new demands are being piled on the higher educational community in a short span than can be readily assimilated; in the financing of higher education, inadequate attention is being given to financial needs created by socially imposed costs; and as Earl Cheit has pointed out (1975, p. 33), educational leadership may be losing a sense of "the larger vision" because of its preoccupation with social and governmental demands. Even more important, the proliferation of new social imposed programs and government laws and regulations is a serious threat to academic freedom. Perhaps the greatest need in connection with socially imposed costs is to explore carefully the question of how the legitimate needs of society for security, equality and environmental improvement may be reconciled with the needs of the academy for intellectual freedom. These matters have been discussed with great wisdom and clarity by Presidents Bok (Harvard, 1976) and Brewster (Yale, 1975).

Socially Imposed Costs as a Widespread Phenomenon

Higher education is, of course, not the only sector of the economy that is subject to socially imposed costs. The operations of virtually all businesses, nonprofit organizations, municipal and State agencies, and even agencies of the Federal Government are subject to social pressures and governmental programs resulting in significant increases in expenditures. For example, most are faced with demands relating to:

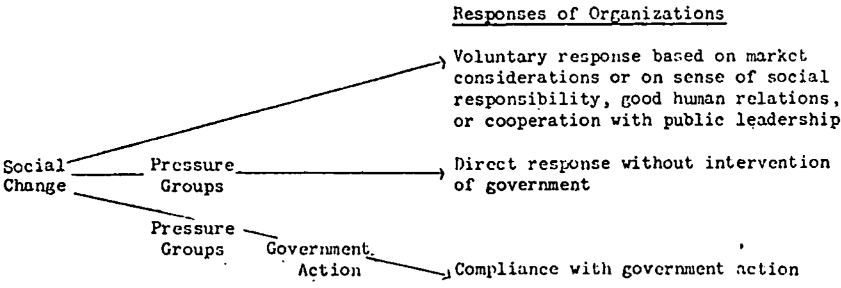
- Health, safety, security, and welfare of workers;
- Prescribed standards of wages, hours, and working conditions;
- Collective bargaining;
- Equality and openness of personal opportunity and of access to economic and cultural facilities;
- Participation of constituent groups in decisionmaking;
- Protection of personal privacy;
- Public disclosure of information;
- Consumer protection;
- Environmental protection; and
- Development and conservation of natural resources.

In addition, most industries are subject to particular socially imposed costs related to their special circumstances. For example, railroads are subject to special rate regulation and labor standards, oil companies are subject to special environmental demands, multinational companies are subject to the requirements of U.S. foreign policy, colleges and universities are subject to the shared-cost concept in connection with Federal grants and contracts.

The socially imposed costs, to which organizations of all kinds have become subject in recent decades, were due mostly to profound social changes. These changes were worldwide though they appeared in

unique forms and combinations in different nations. They were part of the evolution of industrial society. Asserting that many socially imposed costs are due to basic social change does not imply that social pressures or governmental action leading to increased cost were inevitable or well-advised any more than it implies that organizational responses are always appropriate.

Organizations of all kinds, including higher education, have responded to these social changes in three ways: (1) By voluntarily adjusting to the new conditions in ways that will improve their market positions;[2] (2) by voluntarily modifying their behavior according to their own concepts of social responsibility, of good relations with workers, consumers, suppliers, and the general public, or of cooperation with public leadership; (3) by altering their behavior in ways forced upon them by collective bargaining, community opinion, mass protest, and threat of government action; and (4) by changing their behavior in conformity with specific government sanctions, Governmental sanctions in turn were responses to these same social changes. Thus, basic social change—not merely the arbitrary or whimsical decisions of government—lie at the root of much of the behavioral modification of organizations. The path of causation is illustrated in the following diagram:



The changed behavior of organizations, and any associated increases in costs, are ultimately due to social change—sometimes mediated through government and sometimes not. Government programs and regulations are thus seen as part of a more fundamental social process. In this paper, all socially imposed costs, not merely with those in which government happens to be involved, are considered.

Some Conceptual Issues

When contemplating assigning dollar amounts to the socially imposed costs, several basic conceptual issues emerge, quite apart from technical problems of estimation.

First, social change of the kinds under consideration may raise cost, lower cost, raise income, lower income, or have no effect on either cost or income. The frequent assumption that such social change will always raise cost is surely false. For example, when an organization responds to social demands regarding the conditions of work, the

[2] Even nonprofit organizations such as colleges and universities are mindful of the effects of their actions upon their markets.

morale or the health or the job satisfaction of workers may be enhanced and, as a result, worker productivity may be raised and cost lowered. Even when this effect is not immediately noticeable within the particular organization, improved working conditions may in the long run bring about a rise in productivity for the whole society (eventually including the particular organization). Comparably, a change in tax laws intended to enhance the fairness of the tax systems may reduce philanthropic contributions and thus lower the income of nonprofit organizations, an effect which would be tantamount to an increase in cost. In view of these possibilities socially imposed costs should be defined to include positive and negative effects on both costs and income in the long run.[3]

Second, socially imposed costs may be usefully divided into two groups (1) Costs for actual program operations, and (2) costs associated with compliance or information. For example, one might distinguish between social security taxes paid versus filling out the returns and recordkeeping; operating a black studies program versus providing statistical information on minority enrollments; improving animal care facilities versus preparing annual compliance reports; increasing the faculty-student ratio as suggested by an accreditation team versus preparing an institutional self-study. These distinctions may break down when pressed too far, but they are important because much of the furor about social costs pertains to costs of compliance or of gathering and supplying information rather than to program costs which in the aggregate are much greater than costs of compliance.

Third, the time over which socially imposed costs are to be measured must be specified. Among the possibilities are: (1) To measure increases in these costs between two selected dates with the results presented as total increases over the period or as average annual increases, or (2) to measure the total expenditure for socially imposed costs in a given year. The second possibility would be tantamount to measuring the increase in social costs since the first such costs were incurred, probably since the inception of higher education in this country. Ideally, it would be desirable to know the total social cost in the latest year and the rate of growth in social cost over recent years. In connection with specifying the time, a distinction should be made between capital costs and recurrent operating costs and between startup costs and continuing recurrent costs. Contemporary discussion tends to be fuzzy concerning these distinctions. For example, some of the costs of the Occupational Safety and Health Act (OSHA) may be in the form of one-time capital costs which loom large in the future but which would seem less overpowering if amortized over the life of the capital improvements. Whereas another program involving mainly operating costs—for example, establishing women's intercollegiate athletics—may be seemingly less awesome in

[3] When judging the effect of a social change on cost, one must ask: Compared to what? For example, if the workers of the Nation are demanding universal health insurance, their morale and productivity might fall if the demand is rejected. Meeting the demand, then, may be said to increase productivity and lower cost, compared with what it would have been with rejection but not compared with past performance. On the other hand, in a different psychological climate, adopting universal health insurance might raise morale and increase productivity compared with past performance. The effect on cost, then, may be relative to the psychological climate.

the future but may in the long run be more costly than OSHA. Similarly, in many cases, costs of compliance or of supplying statistical data are relatively high in the first year when new procedures and new computer programs must be adopted and employees must be trained, but become much lower in later years when the operations can be routinized.

Fourth, a distinction must be made between the increasing of costs and the changing of priorities. When social or government intervention impose new programs and activities upon organizations, it does not necessarily follow that aggregate costs will be increased. Such intervention may merely change the priorities of the organizations, causing a substitution of new programs or activities for old ones, with total costs unchanged. For example, if the specifications of the Employment Retirement Income Security Act (ERISA) didn't require higher education to introduce ethnic studies and to follow affirmative action, then the same money might have been spent to raise the faculty-student ratio, to expand the library, or build a new sports arena. The amount of money available to spend would not necessarily have been greater because of the imposed social costs. The priorities would simply have been altered. Similarly, when administrators and faculty members are called upon to assemble statistics, prepare reports, entertain site visitors, negotiate, lobby, and read governmental regulations, these activities may not increase aggregate expenditures but may merely divert staff time and energy from accustomed work to new tasks. These new demands may change the priorities but not necessarily increase aggregate costs. Do priorities involve an improvement or a deterioration in the efficiency of the organization viewed from the standpoint of the broad public interest?

Consideration of these four conceptual issues suggests caution in rushing to the conclusion that every social pressure or government regulation raises cost or reduces efficiency. The probabilities are high, however, that the combined net effect of these social changes has indeed raised costs of traditional services or has changed priorities in ways that impair efficiency of conventional operations. When these social changes are viewed in a broader context, however, they may mean that in recent decades society chose to take an increasing share of the national product in the form of security, safety, health, participation, due process, equality, privacy, consumer protection, agreeable working conditions, environmental improvement, conservation of resources and information, and a smaller share in the form of conventional goods and services. Society apparently chose a more secure, a more humane, and a more informed world in place of a more affluent world. This is not necessarily a bad choice—though people may disagree on the details—but the consequence is a significant change of the composition of the national product. It happens, however, that our conventional social accounting does not include these new goods relating to security, humaneness, and information as part of the national product and so there are rising costs of production without corresponding increases in the measured quantity of goods and services. Rather these costs, regardless of their merit, are assigned to the production of ordinary goods and services and show up as reduced efficiency or impaired productivity (in the production of these ordinary goods and

services) rather than as legitimate costs that yield distinctive and useful, albeit intangible, products.

To finance these costs (which are akin to taxes) organizations must either increase their revenues or make offsetting reductions in regular costs. Groups may increase revenues by raising the prices of the ordinary goods and services they produce. Or they may reduce their regular costs by lowering the range or quality of goods and services produced. In either of these cases, costs are shifted forward to consumers. Nonprofit organizations may shift some or all of the cost to philanthropists and taxpayers by obtaining increased amounts of gifts or appropriations.[4] On the other hand, organizations may shift the new costs backward to workers by paying wages and salaries lower than would otherwise prevail. It is often argued that payroll taxes and other fringe benefits paid by employers are shifted in the long run to workers in the form of lower wages. This conclusion is based on the assumption that real labor costs (including wages, salaries, and fringe benefits) are limited by the net productivity of workers. This conclusion, however, may be questioned for a labor market that is dominated by collective bargaining, minimum wages, and regular annual wage increases as the norm for nonunionized as well as unionized workers. On the other hand, it is often asserted that costs for purposes such as environmental improvement, consumer protection, conservation of resources, and affirmative action, are usually shifted to consumers in the form of higher prices. The question of the precise shifting and incidence of taxes and related charges is far from settled, especially in the case of taxes and charges that apply across the entire economy or across large parts of it. Since most members of the population are both workers and consumers, it may not make much practical difference whether the shifting is to workers or to consumers. But to the extent that the shifting is to consumers (a large part of it does go in this direction), it is likely that some of the steady, relentless inflation experienced year after year is due to socially imposed costs which are passed through to consumers in the form of higher prices or reduced quality of product, or to taxpayers and philanthropists through increased appropriations and more generous gifts.

To the extent that the new costs are shifted to consumers in the form of higher prices, they affect the general price indexes, accelerating the rate of increase in these indexes. Some inflation, then, is almost certainly due to cost increases generated by social change. Strictly, these costs should not be registered in the price indexes because they are incurred for the purpose of producing important social values such as security, job satisfaction, participation, and environmental improvement. But since these values are not counted statistically in the national product, they show up as higher prices for ordinary goods and services with which they are jointly produced. It is not known to what extent the price indexes were biased upward because of these social costs. But it is certain that the effect was substantial, and that the true rate of inflation derived from monetary and fiscal factors was less than usually supposed.

---

[4] In the case of higher education, the possibility of shifting these costs forward may be affected by the discriminatory price policy charactersitic of colleges or universities.

Because some socially imposed costs become incorporated in the general price indexes, it is necessary in estimating social costs to avoid the error of deflating expenditures on the basis of a general price index and then to consider the social costs as part of the deflated expenditures. Rather, at least some social costs must be considered as part of the expenditures reflected in the rising price indexes.

THE SPECIAL POSITION OF HIGHER EDUCATION

Up to this point, the discussion has been concerned with socially imposed costs in general as applied to all sectors of the economy. Colleges and universities are only one class of organizations affected. As indicated, organizations of different types and in different industries are subject to special costs or are affected in special ways. This section describes the special position of colleges and universities.

Higher education became subject to some of the socially imposed costs later than other industries. For example, higher education was exempted in the early years of social security, minimum wage laws, collective bargaining legislation, and other social programs and was brought in only after many years. More recently, however, higher education is regarded for purposes of social legislation as no different from other industries, and in the newer programs, it has been included from the start. As a result, higher education has had to adjust to the many socially imposed costs over a shorter period than other industries. This problem has been exacerbated because the cost and inconvenience of social programs is greater at the time of startup than over the longer pull when they become routinized.

Higher education (and other nonprofit organizations) may have more difficulty than profitmaking enterprises in shifting social costs. The problem may be especially severe in the case of institutions that are near the margin of survival. Whereas profitmaking enterprises are financed almost wholly from the sale of their products in the market, colleges and universities receive only part of their income from tuition and fees and the rest from appropriations, gifts, and investment income. For higher educational institutions to recover increased social costs, they must usually enlist the aid of legislators and donors as well as raise tuitions and fees. If they are unsuccessful in this effort, then their only recourse is to absorb at least some of the costs within their budgets at the expense of the range of quality of their programs. In this case, the social costs are indeed shifted to consumers but through deterioration of programs, a way that is seldom acceptable to educators. This is not to say that profitmaking enterprises can always shift new social costs instantly through higher prices to consumers, but only that their power to do so may be greater than that of higher education [Van Alstyne & Coldren, 1976, pp. 18–19].

Colleges and universities may have more difficulty in coping with new social costs than business firms because they may be less adequately staffed and, in view of their decentralized organization, less capable of dealing with government.

It is sometimes argued that higher education is harder hit by new social costs than other organizations because colleges and universities

are relatively labor intensive. It may be that the payroll of higher education makes up a somewhat larger part of total expenditures than that of other industries on the average. Wages and salaries in the entire economy, however, are about four-fifths of total national income, and it is doubtful if higher educational payrolls are much above this share.[5]

Higher education differs from business in that it is largely tax exempt. This special status cuts two ways. On the one hand, higher education receives a hidden subsidy in the form of general governmental services for which it does not pay. This may justify government in expecting colleges and universities to bear some social costs or to serve in partnership with government in some costly research and training program. On the other hand, tax exemption means that higher education is not able to deduct socially imposed costs from taxable income, as private business can do, and that colleges and universities must bear the whole cost rather than about half the cost as in the case of private companies of comparable size.

Though the special conditions within higher education may put it at some disadvantage in comparison with a profitmaking business, the differences are not pronounced. On the basis of the differences mentioned, it is hard to make a case that higher education should have special treatment except the obvious one that its funding source should take socially imposed costs into account in setting the amount of appropriations and gifts. There is, however, one additional difference that stands alone because of its central importance. This is the special need of colleges and universities for freedom from government controls and social pressures. Profitmaking enterprises are, of course, as concerned as any other organizations to protect their freedom of decision and action. Colleges and universities as centers of teaching and research, however, have a special interest in autonomy because of their responsibility to protect freedom of thought from encroachments either of interest groups or of government. For this reason, the sensitivity and the resistance of the higher educational community to some forms of external pressure and control are likely to be greater than the resistance of profitmaking enterprises—and properly so.

SPECIFIC SOURCES OF SOCIALLY IMPOSED COSTS IN HIGHER EDUCATION

This section presents an inventory of the sources of socially imposed costs to which colleges and universities are subject. In a later section, the question of the magnitude of these costs will be considered.

Personal security.—The increasing concern in our society for the protection of individuals against hazards of life such as unemployment, illness, accident, old age, and premature death result in many informal influences on organizations to manage their affairs in ways that will enhance personal security, and also result in a flood of legislation—Federal and State—requiring organizations to provide for the security of their workers. Among the specific laws, regulations, and influences are the following:

[5] Some rough calculations suggest that payrolls make up 80 to 85 percent of total educational and general expenditures.

Federal Legislation or Regulations [6]

Social Security Act of 1935 as amended:
- Old-age pensions.
- Survivors' insurance.
- Disability insurance.
- Unemployment compensation.
- Health insurance.
- Occupational Safety and Health Act of 1970 (OSHA).
- Poison Prevention Packaging Act.
- Radiation safety.
- Protection of human and animal subjects used in research.
- Health Maintenance Organization Act of 1973.
- Employment Retirement Income Security Act of 1974 (ERISA).
- Intermittent wage and price controls designed for broad economic stabilization.

State and Local Legislation or Regulations

- Workmen's compensation.
- Building codes for protection against fire, flood, and earthquakes.
- Rules and inspections relating to public health and health-care services.
- State laws and programs relating to retirement, occupational health and safety, and unemployment compensation, some of which overlap with Federal programs.

Institutional Practices Resulting From Informal Social Influences

- Fringe benefits in the form of pensions, health and disability insurance, life insurance, and severance pay.
- Tenure and other long-term employment contracts.
- Procedures making termination of employment legally and practically difficult.

Most of these programs or practices involve substantial costs for both the benefits afforded employees (and in some cases students) and for expenses involved in compliance.

Personal opportunity.—Another important source of rising costs to higher education is the increasing concern in our society for open access of all persons to education, work, public facilities, and cultural amenities without discrimination on the basis of race, religion, sex, physical and mental handicaps, age, personal appearance, and life style. This concern led to far-reaching voluntary decisions on the part of colleges and universities, to heavy informal pressures on them, and to a great deal of legislation and court-made law. Some of the specific laws, regulations, and influences are as follows:

[6] Some colleges and universities are exempt from some Federal programs. For example, they participate voluntarily in social security old-age pensions, survivors' insurance, and disability insurance. In the case of unemployment compensation, ERISA, and OSHA, many are subject to comparable State laws or programs.

Federal Legislation or Regulations

Affirmative action: Executive Order 11246 of 1965 as amended in 1967 to include discrimination on the basis of sex.
Employment Act of 1967 (relates to discrimination on the basis of age).
Table VII of the Civil Rights Act of 1964 as amended by the Equal Employment Opportunity Act of 1972 (prohibits discrimination in employment practices on the basis of sex, race, creed, or national origin).
Title IX of the Educational Amendments of 1972 (bars sex discrimination in educational policies, programs, and employment practices).
Financial aid to students (a variety of grants, loans, work-study programs, and special provisions for minorities).
Internal Revenue Service regulations concerning discrimination in employment and in student admissions.
Many judicial decisions affecting employment, student recruitment, and educational practices.

State and Local Legislation or Regulations

Some States and municipalities enacted legislation or regulations pertaining to discrimination in student admissions or in employment. In particular, open admissions or special provisions for minorities were mandated in some jurisdictions.
Programs of financial aid to students and of assistance to minorities are available in many States, and have tended to increase the student load. These programs, however, also expanded the opportunities of the institutions to obtain revenue from both appropriations and tuition.

Institutional Practices Resulting From Informal Social Influence

In carrying out newly perceived social responsibilities and in reacting to pressures from minority groups and women, many institutions actively recruited minority students and women and provided special remedial programs, counseling, curricula, facilities, and financial aid for these groups.

Because higher education is a vital ingredient of social mobility for disadvantaged groups and a strategic point of access to the mainstream of American society, colleges and universities are focal points in the struggle for personal opportunity and human equality. Their role is more important in the areas of student recruitment and educational programs than in the area of employment practices, though the latter is essential to nondiscriminatory education. An institution that discriminated in employment practices would not be a fitting environment for nondiscriminatory education. To some degree, the involvement of colleges and universities in equal opportunity antedated both the intense social pressure and government programs of the 1950's

and later. Without the pressure—informal and governmental—it is doubtful if colleges and universities would have extended themselves as much as they have, however.

The efforts of colleges and universities to respond to social demands for equal personal opportunity involved substantial costs—especially in recruiting and admitting students,[7] supplying student aid, and providing special programs and facilities. Some of these costs are for one-time capital investments, but most tend to be recurrent and increasing. Complying with various related governmental programs involves costs of data collection, report preparation and legal services. Also, dealing with campus unrest associated with minority issues involved significant time and effort on the part of faculty and administrative officers.

Public programs of student aid have undoubtedly relieved institutional student aid budgets. On the other hand, social pressures on institutions required them to stretch their resources for student aid and also to bear significant administrative or matching costs connected with public student aid.

Two new areas of Federal intervention with substantial cost implications are pending. One is the extension of affirmative action to student admissions [8] and the other is providing facilities to accommodate students with physical and mental handicaps.[9] The cost of these two initiatives is likely to be high and it remains to be seen how these costs will be financed.

Participation, openness, due process, and privacy.—A prominent characteristic of contemporary society is the rejection of paternalism, arbitrary decisions, and secrecy in the conduct of affairs. As a corollary, there is a strong demand for participation of individuals in matters affecting them. Employing organizations, educational institutions, churches, political parties, and government have all been persuaded or forced to adopt more open and democratic decisionmaking processes. Closely related to the demand for participation is a new insistence on due process in matters affecting individuals including formal grievance and appeal procedures. In the present climate of opinion, most organizations find it expedient or mandatory to reach decisions openly, with the participation of those affected, and with formal provisions for due process and review. This is especially so because decisions often lead to litigation and courts have leaned increasingly toward reviewing organizational decisions not only on the basis of substance but also of procedure. Indeed, the mandate for participation, openness, and due process comes largely from judicial decisions rather than legislation.

The principal legislation on these matters is the first amendment to the Constitution (as continuously reinterpreted); the Wagner Act of 1935 (as frequently amended and interpreted) providing the right of workers to organize, to bargain collectively, and to strike, and formulating rules governing collective bargaining; various Federal and State laws prohibiting secrecy in the conduct of public business.

---

[7] The costs to particular institutions of recruiting minority students may level off as institutions gain more experience and as they develop networks of minority alumni who can assist in the recruitment process.

[8] Speech of Joseph A. Califano, Jr., at City College of New York, June 5, 1977. Quoted in Higher Education and National Affairs, American Council on Education, June 10, 1977, p. 1.

[9] "Califano Signs Guidelines on Handicapped; Costs of Compliance Worries Colleges," Chronicle of Higher Education, May 9, 1977, pp. 3, 12–14.

Otherwise, the sanctions derive primarily from informal group pressure and court-made law.

The demands for participation, openness, and due process deeply affected the governance and decisionmaking within higher education. They led to increasing influence of students, faculty, and nonacademic staff in the affairs of colleges and universities. They led to greater attention to the codification of rules, to consultation, and to procedures. They gave rise to pressure groups among faculty, students and employees; to new tactics for exerting pressure including demonstrations and strikes; to the establishment of the office of ombudsman; to formal grievance procedures; and to frequent litigation. Collective bargaining was also a manifestation of the demand for participation, openness and due process. Though relatively new in higher education, collective bargaining is commonplace among nonacademic employees and is increasingly prevalent among faculty members, especially in public community colleges and State colleges.[10]

The new demands for participation, openness, and due process require much time, effort, and money in establishing formal procedures, in settling specific complaints, in obtaining legal advice, and in defending law suits. They also greatly increase the time and effort in ordinary decisionmaking and impair the flexibility and innovativeness of institutions. They diverted people—administrators, faculty, staff, and students—from their ordinary pursuits. Whether they resulted in better decisions, more amicable human relationships, higher morale, greater productivity, and better education is not known, but it is doubtful that they have had these effects. They almost certainly strengthened the position of faculty, staff, and students in the decisionmaking process and protected them from arbitrary decisions, but in doing so have injected new costs and constraints into the operation of colleges and universities.

Closely related to the concern for due process was a new interest in the protection of individual privacy. This was manifested in many ways, beginning with the traditional limitation on the search of private premises, and extending to more recent restrictions on wiretapping and other covert surveillance, and to limitations on the use of educational histories, credit records, financial transactions, medical histories, criminal records, and letters of recommendation. The concern for privacy also raises issues about the use of identity cards and identity numbers.

These developments have profound significance for higher education. Colleges and universities have long accumulated records on individual students regarding academic performance, campus work, personal behavior, and health; faculty members and administrators added to these records by conferring awards and prizes and by writing letters of recommendation; and these records were shared with other educational institutions and with prospective employers including the Armed Services. On occasion, these records figured in criminal proceedings. These records certainly had a bearing on crucial educational decisions such as admissions, promotion, retention, award of

---

[10] The effect of collective bargaining and other forms of participation in the decisionmaking process on wages and salaries will be considered in another section of this discussion.

56-368 O - 80 - 29

degrees and transfer to other institutions. In the present climate of opinion, institutions of higher education are cautious about the collection, storing, and conveyance of such information, and systematic rules and procedures were adopted. One factor in their caution was the threat of law suits.

Concerns about privacy for students culminated in the Family Educational Rights and Privacy Act of 1974. Known as the Buckley amendment, it gives students the right to inspect their files, to seek changes in records they feel are inaccurate, to demand formal hearings of complaints, and to prevent disclosure of information of a private nature without the student's permission. The act also requires that institutions obtain permission from past informants before the information they have supplied is released to students. The Buckley amendment may have reduced the capacity of colleges and universities to supply pertinent information about students to prospective employers and to other educational institutions, but it may have protected students from the dissemination of information about them.

Emancipation of youth.—In recent decades, a revolution occurred in the degree of freedom of young people with respect to manners, mores, dress, general mode of life, and personal life decisions. Along with this came a rejection of paternalism both within the family and within institutions. The chief legal manifestation of this social change was the constitutional amendment lowering the age of majority to 18.

Higher education was deeply influenced by the emancipation of youth. It has largely abandoned the concept of in loco parentis and deals with its students as adults. It has greatly reduced specific curricular requirements and social rules, and has narrowed its responsibilities for supervision and guidance. The percentage of students living in institution residences declined. Also, in view of the reduced formality of life style and manners, emphasis on gracious surroundings, served meals, and formal parties and receptions, were reduced. These changes may have resulted in lowered costs. On the other hand, changing conditions in the job market led to expanding services in vocational guidance and placement and to the introduction of new vocational curricula. The new sexual mores led to new demands for medical services. The increasing use of cars, TV sets, hi-fit sets, and other gadgets by an affluent generation of students increased energy consumption and parking costs. And new behavior patterns increased the need for campus security. On balance, it is unlikely that the emancipation of students reduced costs.

Perhaps the most important aspect of the emancipation of youth is its potential effect on student finances. The basic concept underlying student financial aid is that the amount of funds to be granted any student is largely determined by the financial circumstances of the student's family—the family defined to include the student, his parents, and other contributing members. But if a student becomes emancipated from his family, as increasing numbers do, the parental family is no longer relevant to student aid allotments. The family of a single student then becomes simply the student himself and the family of a married student becomes the student, a spouse, and children if any. Thus the amount of aid due an emancipated student is usually more than that due one who is still part of a parental family. This situation is

aggravated by the increasing average age of students. Another aspect of emanicpation affects in-State versus out-of-State tuitions in public colleges and universities. Historically, the residence of a student was deemed to be the residence of his parents. But emancipated students, and others as well, may establish residence in the State where they go to college (regardless of where their parents reside) and thus qualify for in-State tuition. The impact of emancipation upon expenditures for student aid and on income from tuitions is substantial. If the trend toward emancipation continues, the financial effect could be enormous.

Consumer protection.—A social concern that has a long history but that has been gathering steam in recent decades is consumer protection. Its purpose is to promote health and safety and is therefore akin to the concern for security already discussed. But it has another purpose; namely, to prevent deception and exploitation of consumers. The legislation in this area centers on Federal laws establishing the Federal Trade Commission, the Food and Drug Administration, the Securities and Exchange Commission, and a host of agencies scattered throughout the Federal Government and State and local governments that inspect, license, accredit, cajole, and otherwise supervise and regulate organizations for the protection of consumers.

In the case of higher education, society depended primarily on voluntary accreditation for the protection of consumers. In some States, however, departments of education or boards of regents also exercised surveillance and control. Recently, abuses in the proprietary sector, the appearance of degree mills, and also impaired confidence in colleges and universities led to demands for consumer protection. Also there were a few court cases in which plaintiffs alleged that they were misled by the claims of institutions attended or that institutions did not produce the educational results they had promised. Generally, however, the rising demand for consumer protection did not produce a significant increase in higher educational costs, though potential future increase of modest amounts are likely.

Work standards.—Over the years, standards for wages, hours, and other working conditions have steadily risen. The rising standards grew out of, and were often backed up by, legislative mandate. The principal legislation is the following:

- The Wagner Act of 1935 (as frequently amended) providing the right of workers to organize, to bargain collectively, and to strike, it also formulates rules governing collective bargaining.
- The Fair Labor Standards Act of 1938 (as frequently amended) providing minimum wages, maximum hours, and time-and-a-half for overtime work.
- The Equal Pay Act of 1963 providing that employees doing similar work should receive the same pay regardless of the employee's sex.
- The economic Stabilization Act of 1970 instituting temporary wage and price controls.
- Many State laws regarding working conditions other than wages and hours.

Higher education has become subject to legislation of this type relatively recently but it has already had considerable impact.

Though there is a question whether collective bargaining in the long run raises the level of real wages in the economy as a whole, it probably raises the wages of particular disadvantaged groups. It happens that colleges and universities have a long record of paying less than prevailing wages to their nonacademic employees. They have been able to do this in part by offering steady employment, by providing pleasant and humane working conditions, by exploiting captive workers such as students and wives of students and professors, by placing workers under less pressure than is customary in private industry, and by providing special fringe benefits (for example, reduced tuition or tickets to public events). The unionization of nonacademic employees has probably raised their wages and thus increased higher educational costs. In the case of faculty members, who on the whole were well paid over recent decades, there is some doubt whether collective bargaining raised compensation. Two recent studies suggest that its effect may have been weak.[11]

Under conditions that can easily be visualized, however, the impact of collective bargaining might be considerable over the long run—especially in preventing the decline in real earnings that might otherwise accompany a weakening of the demand for academic labor.

The minimum wage law and the requirement of time-and-a-half for overtime both undoubtedly raised higher educational costs. Their effect was to raise money wage payments without necessarily increasing productivity or taking away fringe benefits. Also, the secular trend toward higher standards in working conditions other than wages and hours may also have raised costs in higher education.

Environmental protection and natural resource use.—A major social change is concern for protection and improvement of the environment and for conservation and development of natural resources. Numerous environmental laws and programs (Federal, State, and local) were enacted or initiated. These relate to improvement of air and water quality, sanitation, noise levels, natural beauty, architectural design, and land use. Similarly, numerous laws and programs were enacted or initiated relative to conservation of land, water, forests, energy, minerals, fisheries, and other natural resources. With a growing population, increasing demands on the environment, and steady depletion of natural resources, concern for the environment and for resources is likely to become more intense.

Higher education is neither a heavy polluter of the environment nor a heavy user of natural materials. Its effect on the environment tends on the whole to be favorable or at worst neutral. Colleges and universities sometimes emit pollutants from their powerplants, or discharge liquid wastes into streams, or dispose of solid wastes inadequately; they sometimes create minor neighborhood nuisances related to traffic, parking, and student behavior; they sometimes use (or wish to use) land in ways contrary to overall city plans; and they sometimes engage in research that may conceivably endanger the environment, for example, research involving radiation or recombinant DNA. Environmental concerns as expressed in government programs and policies, however, are on the whole a minor source of increased expenditures for colleges and universities—though in some

[11] "Do Unionized Faculty Members Get Bigger Pay Increases?" Chronicle of Higher Education, Dec. 6, 1976.

instances institutions were subjected to major costs or major inconveniences because of municipal zoning regulations or demands for smoke abatement.

The use of energy in higher educational institutions is chiefly for heating, air conditioning, lighting, and operation of small equipment and machinery. The use of water is chiefly for drinking, sanitation, cleaning, and air conditioning. The uses of energy and water correspond to that of homes and office buildings—not to that of factories or transportation. And colleges and universities are minor users of other natural materials. Rising relative costs of energy, water, and other natural materials raised the expenditures of colleges and universities. But the increase was probably due predominantly to the policies of the oil cartel and to market forces rather than to domestic governmental decisions and actions. The cause of the rising costs of energy and water were not so different from the causes of the sharp increases in the prices of legal services, books, paper, building construction, postage, and many other items educational institutions purchased.

Need for knowledge and information.—A complex industrial society has an almost insatiable appetite for knowledge and information. They are needed as an important ingredient of broad cultural and technological advancement, they are needed for private policymaking, and they are needed for government policymaking, policy, evaluation, and program enforcement. These needs result in a host of grants and contracts from government in support of research to be conducted by colleges and universities. They also result in a surfeit of questionnaires and reports that are requested or demanded of colleges and universities by government agencies and also in numerous inspections, site visits, and negotiations.

A major issue in connection with governmental grants and contracts to support research in colleges and universities concerns the provision for indirect or overhead costs. The Federal Government discriminates between universities and private business in allowances for overhead and other indirect costs. The theory underlining the less generous treatment of higher education is in two parts: (1) because one of the established functions of colleges and universities is to conduct research, their regular funding sources should share in the costs of governmentally sponsored research, and (2) because universities are research enterprises, the overhead costs related to governmentally sponsored research are not increased in proportion to the volume of research undertaken and reimbursement of marginal rather than average overhead costs is adequate. At any rate, it is debatable whether research grants and contracts from the Federal Government raise university and college costs in ways that place increased burdens on the regular funding sources of the institutions. Many educators believe that institutional costs are probably increased by reasons of Federal research grants and contracts and that what the Government euphemistically calls "cost-sharing" probably places a considerable burden on higher educational institutions. In any case, the negotiations about reimbursement of overhead costs are exceedingly time-consuming.

The need for knowledge produced another result. Colleges and universities—along with companies and private individuals—were faced

with increasing governmental demands for data. Some of the requests are seemingly useless, some are ambiguous, some are duplicative, and most request data in unique forms that necessitate special analyses. The time, effort, and frustration involved are substantial. Nevertheless, that large quantities of information are needed can hardly be disputed. Colleges and universities are in an embarrassing position when they object to the load of questionnaires and reports because they are important research organizations themselves and are among the leading perpetrators of requests for data. Literally hordes of graduate students and professors are seekers of information in the form or responses to questionnaires or personal interviews. Moreover, much of the data collected by government is of greater use to academic researchers than to any other group. Colleges and universities clearly place a heavier burden on society for the supply of information than society places on higher education. Compliance with government in the supply of data does in fact raise higher educational costs, however.

In recent years, adverse public attitudes about higher education and increased political interest in colleges and universities have led to the "accountability" movement involving new demands for information, self-studies, reports, justifications of fund requests, inspections, investigations, and new layers of accreditation. Most of this activity is at the State level and most—but not all—is directed toward public institutions. Legislative committees, Governors, State finance officers, State personnel offices, coordinating boards, and other agencies request information, often in detail and on short notice. Attacks upon higher education by political candidates and public officials are frequent and must be answered. The effect of these developments on institutions was not only additional demands for information but loss of autonomy as decisionmaking has been shifted upward to statewide boards, legislative committees, and State finance officers. The result was reduced administrative initiative and flexibility and increasing management by formulae imposed by absentee administration. These developments undoubtedly raised costs but, more important, probably impaired efficiency and academic freedom.

Ironically, the growing involvement of higher education with Federal and State governments has increased the need in the institutions for information about existing laws, regulations, and programs and about proposals for new ones. It also required increasing attention to formulating positions on future legislation and to advocating these positions. Some of the increased effort takes the form of membership and activity in organizations representing higher education and some takes the form of direct lobbying. All of these activities take time, effort, and money.

Tax reform.—In recent years, a slow erosion of the traditional tax-exempt privileges of higher education has occurred. In some respects, this tended to raise costs and in others to lower income. The Federal Government modified taxes in ways that may inhibit philanthropy while it constantly threatens to take additional drastic action. The Federal Internal Revenue Service increasingly scrutinized the operations of nonprofit organizations—partly to discover taxable income and partly to enforce affirmative action. As a result, significant new demands were made on the time and effort of college and university administrative staffs. In many areas of the Nation, local and State governments were less lenient with respect to tax exemption on the

property and sales of nonprofit institutions. Colleges and universities are often persuaded—in some cases pressured by threats of adverse rulings on zoning—to make contributions in lieu of property taxes. And some local and State governments have stiffened their policies on taxation of property not directly used for educational purposes, taxation of "unrelated" income, and tax treatment of charitable contributions.

Neighborhood deterioration.—In recent decades there has been a visible increase in urban decay, crime, vandalism, and militant protest in the communities within which higher education operates. These condtions and the costs they generate may be attributed in part to failures of government, but they may be also attributed to failures of the family, the church, the media, and education itself. But whatever the cause, they are among the social changes that have increased the costs of higher education. Neighborhood deterioration resulted in increased provision for security, higher insurance premiums, losses from theft and vandalism, and unwelcome changes in mode of operation. Also some institutions found it expedient to incur large expenditures for property acquisition, urban renewal, and neighborhood improvement in areas adjacent to their campuses.

Changes in technology and proliferation of knowledge.—Developments in technology over the past several decades have affected instructional methods, but there is little evidence that they have improved productivity of higher education. The introduction of TV, computer-assisted instruction, advanced audio-visual systems, and miniaturization of printed materials may have improved teaching-learning slightly, but have not reduced expenditures significantly. At the same time, new technologies laid responsibilities on the institutions to teach new subjects such as computer science, TV journalism, labor relations, modern genetics, and new techniques in medical and allied health fields. They also required institutions to secure new equipment for both teaching and research purposes, equipment such as advanced computers, electron miscroscopes, linear accelerators, and advanced optical and radio telescopes. Meanwhile, the proliferation of knowledge brought about the introduction of new specialties into both teaching and research and required a great acceleration of library acquisitions with associated problems of storage and retrieval of information. Government had little to do with creating these problems, and generously assisted higher education in meeting them. Indeed, that our colleges and universities were able to keep up-to-date in equipment and books is due in no small measure to the assistance of government. In assessing the effect of government on higher educational costs, it is necessary to consider the assistance given as well as the costs imposed.

Special problems of teaching hospitals and clinics.—Government imposed special requirements on teaching hospitals and clinics that tend to add cost or reduce income. These requirements relate partly to the time and effort involved in health planning procedures, patient care review accreditation and licensure, accounting requirements, drug and blood regulations, radiation protection, use of human and animal subjects for research and liaison with the Veterans' Administration. The impacts of government also relate to the reimbursement of hospitals and clinics for medical services performed for medicare, medicaid, and patients in other programs whose bills are paid partly or

wholly by public agencies. The issue here is similar to that of indirect cost reimbursement in the case of research contracts and grants.

Special fiscal problems.—The efficiency of programs involving the Government is often said to be impaired because of annual budgeting combined with erratic variations in level of support for particular programs. In some States, the fiscal situation is further complicated by State governments' taking Federal funds (and other in-State funds as well) into the State treasury and reappropriating them to the universities and colleges.

Estimated Dollar Amounts of Socially Imposed Costs Affecting Higher Education

In the preceding sections, we reviewed a wide range of social changes that affected higher educational costs. In most cases, the impacts of these changes were mediated to some degree through government. It is clear from this review that a vast array of new or accentuated social influences have come to bear on higher education and that most of them tend to raise institutional costs, a few tend to lower institutional income, and a few have a negligble financial effect. In this section, we shall consider the impact of these social changes on institutional costs. Accounting records of colleges and universities are not well adapted to isolating socially imposed costs and there are serious conceptual problems. Nevertheless, it may be possible to gain some idea of general orders of magnitude.

Expenditures for general administration and general expense.—A first step in reaching a rough approximation of the dollar amount of costs socially imposed upon higher education is to observe changes in the percentage of educational and general (E and G) expenditures devoted to general administration and general expense. These administrative expenditures include many, though not all, of the socially imposed costs. They include most payroll taxes and other employee fringe benefits and much of the legal work, accounting, and other paper work involved in compliance. As shown in table 2, the percentage of E and G devoted to general administration and general expense rose sharply from 13.1 percent in 1959–60 to 18.6 percent in 1974–75, an increase of 5.5 percentage points. One might conclude from this that the socially imposed costs added since 1959–60 were approximately $1.5 billion (5.5. percent of 1974–75 E and G expenditures of $26.3 billion). This figure leaves out substantial amounts carried in the departmental budgets, plan operation and maintenance, carried research, and other accounts. It might also include some voluntary increases in costs related to improvement of educational quality. On balance, the figure of $1.5 billion is almost surely less than the total of socially imposed costs.

Most of the increase in general administration and general expense was probably due to increases in payroll taxes and other empolyee fringe benefits. But some was probably due to the rise in administrative work associated with government programs. This hypothesis is supported by data in table 3 showing the distribution of academic employment among instructional faculty, other professionals, and nonprofessional employees. As shown in table 3, the percentage of other professionals (many of whom are executive, administrative, and managerial persons) increased substantially in recent years.

Fringe benefits.—Employee fringe benefits increased rapidly in the past decade as shown in table 4. This table refers only to faculty and not to other professional and administrative persons or to nonprofessional persons. If the same rate of increase had occurred for all employees,[12] the rise in total cost between 1964–65 and 1975–76 would have been about $1.5 to $2 billion. This is a rough estimate, again designed to identify an order of magnitude rather than a precise figure.

TABLE 2.—GENERAL ADMINISTRATIVE AND GENERAL EXPENSE AS PERCENTAGES OF TOTAL EDUCATIONAL AND GENERAL EXPENDITURES, ALL U.S. INSTITUTIONS OF HIGHER EDUCATION, 1949–50 TO 1974–75

Year	Total educational and general expenditures (millions of dollars)	General administrative and general expense (millions of dollars)	Percentages
1949 to 1950	1,706	213	12.5
1959 to 1960	4,511	592	13.1
1969 to 1970	15,789	2,628	16.6
1972 to 1973	21,078	3,713	17.6
1973 to 1974	23,435	4,243	18.1
1974 to 1975	26,347	4,900	18.6

Sources: National Center for Education Statistics: Digest of Education Statistics 1975 Edition (1976, p. 130); Financial Statistics of Institutions of Higher Education: Current Fund Revenues and Expenditures (1976, p. 8); Projections of Education Statistics to 1984–85 (1976, pp. 97–98).

TABLE 3.—NUMBERS AND PERCENTAGES OF FULL-TIME-EQUIVALENT EMPLOYEES IN U.S. INSTITUTIONS OF HIGHER EDUCATION, 1966–67 TO 1972–73

[In thousands]

Year	Instructional faculty	Other professionals	Nonprofessional employees	Total employees
Numbers:				
1966 to 1967	324	196	599	1,120
1967 to 1968	348	215	609	1,172
1970 to 1971	417	284	603	1,304
1972 to 1973	428	290	611	1,329
Percentages:				
1966 to 1967	29	18	53	100
1967 to 1968	30	18	52	100
1970 to 1971	32	22	46	100
1972 to 1973	32	22	46	100

Source: National Center for Education Statistics, Number of Employees in Institutions of Higher Education (1976, p. 7)

TABLE 4.—FRINGE BENEFITS AS PERCENTAGES OF FACULTY COMPENSATION, ALL U.S. INSTITUTIONS OF HIGHER EDUCATION, 1964–65 TO 1975–76

Year	Average compensation	Average fringe benefits	Percentage
1964 to 1965	10,050	709	7.1
1966 to 1967	11,289	902	8.0
1968 to 1969	12,951	1,191	9.2
1970 to 1971	14,792	1,508	10.2
1972 to 1973	16,413	1,861	11.3
1974 to 1975	18,709	2,306	12.3
1975 to 1976	20,015	2,565	12.8

Source: American Council on Education, A Fact Book on Higher Education, third issue 1976, p. 170.

[12] The rate of increase was probably about the same for other professional and administrative persons, and possibly higher for nonprofessional employees. Fringe benefits for the latter were traditionally relatively low and increases in recent years were probably more rapid than increases for the faculty and other professional groups.

The Van Alstyne-Coldren study.—In 1976, Carol Van Alstyne and Sharon L. Coldren, under the sponsorship of the American Council on Education, published an illuminating study, "The Costs of Implementing Federally Mandated Social Programs at Colleges and Universities." They obtained data from six varied institutions: University of Illinois, Miami-Dade Junior College, Duke University, Georgetown University, Hampton Institute, and College of Wooster.

They limited their study to 12 Federal programs:

Equal employment opportunity: Title VII of the Civil Rights Act of 1964, as amended by the Equal Employment Opportunity Act of 1963.

Equal pay: Equal Pay Act of 1963.

Affirmative action: Executive Order 11246, issued in 1965, as amended by Executive Order 11375 to include discrimination on basis of sex, 1967.

Age discrimination: Employment Act of 1967, as amended.

Wage and hour standards: Fair Labor Standards Act (FLSA), as amended.

Unemployment compensation: Social Security Act of 1935; Employment Security Amendments, 1970.

Social security tax increases: Social Security Act of 1935; Employment Security Amendments, 1970.

Health maintenance organizations (HMO's): Health Maintenance Organization Act of 1973.

Retirement benefits: Employment Retirement Income Security Act (ERISA) of 1974. (Note: Public institutions excluded.)

Wage and salary controls: Economic Stabilization Act of 1970. (Note: Public institutions excluded; nonprofit institutions exempted January 25, 1974.)

Occupational safety and health: Occupational Safety and Health Act (OSHA) of 1970.

Environmental protection: Regulations implemented under several laws by the Environmental Protection Agency.

Subject to numerous caveats about limitations of data and difficulty of interpretation, they reached the following conclusions:

The combined cost to the six institutions increased from about $500,000 in 1965–66 to about $4 million in 1970–71 and to about $10 million in 1974–75 (pp. 26, 31).

Cost trends over time for different Federal programs varied, but the aggregates tended to increase steadily (p. 27).

The 1974–75 costs due to the 12 programs amounted to 1 to 4 percent of total operating budgets (p. 14).

Because of the small number of the institutions represented in their data base, they did not report the costs of each of the 12 programs separately. They did, however, indicate that one-half the increase was accounted for by social security taxes (p. 29) and that the other major sources of cost (in descending order of importance) were equal employment opportunity programs, occupational safety and health, minimum wage, environmental protection, and unemployment compensation.

If their figure of 1 to 4 percent could be generalized to all of higher education, the dollar amount would range from $0.3 billion to $1.1 billion (1 to 4 percent of total E and G expenditures of $26.3 billion).

Some guesses about overall cost.—On the basis of the Van Alstyne-Coldren study, the anecdotal information from particular institutions (as presented in the Appendix) and other indications, one might guess that the additional to aggregate higher educational expenditures resulting from socially imposed costs over the period since 1949–50 were in the neighborhood of 8 to 10 percent of total educational and general expenditures. E and G expenditures in 1973–75 were $26 billion; 8 to 10 percent of this amount would be $2.1 to $2.6 billion.

Of course, guesses of this kind lack reliability not only because appropriate numbers are not available but also because they depend on what is included within the elastic concept of "socially imposed costs." The 8 to 10 percent would include the following:

- Payroll taxes and fringe benefits;
- Impact of collective bargaining and wages and hours legislation;
- Equal opportunity laws, affirmative action, and other programs for minorities and women;
- Shared costs in connection with government grants and contracts;
- Mandatory changes in buildings resulting from new building codes, fire marshal directives, OSHA, provision for the handicapped, new needs for security precautions, and [13]
- Costs of general compliance, statistical reports, and other paperwork.

Most of this 8 to 10 percent would be connected with employee compensation, shared costs of grants and contracts, and changes in physical plant. All the rest would be of the order of 1 or 2 percent. Had such items as student financial aid, adult education, and public service programs been included, the percentage would have been larger. There is no way of knowing to what extent the included socially imposed costs have been shifted to workers in the form of reduced wages and salaries (as compared with what otherwise might have been paid), or shifted to government, donors, and students in the form of greater institutional income (than would otherwise have been received).

In the recent and present psychological and financial climate in which higher education is operating, it is likely that a substantial part of these costs was financed by retrenchment of program, impairment of educational quality, or reduction in operating efficiency. To the extent that funds must be spent on higher wages and fringe benefits, on minority programs, on shared costs, on building safety, and on paperwork and negotiation, expenditures must be cut back for upgrading faculty, for program and facility improvement, for long-range planning, for building maintenance, and for accumulation of endowment. To the extent that faculty and administrative staff must spend time filling out forms, reading government regulations, negotiating with public officials, lobbying in Washington or in State capitals, and mollifying pressure groups, their effectiveness as educational leaders may be impaired. It may be argued, on the contrary, that socially imposed costs have higher priority than some kinds of expenditures educators would make if the decisions were left wholly to them, and it can be argued that it is good for educators to be brought into the processes of democratic government rather than to be isolated in

---

[13] Conceptually, I have considered modifications of physical plant to be amortized over a period of years.

their ivory towers. There may be honest differences of opinion on these matters. Because of the crucial importance of academic freedom, however, and because of the precarious financial position of higher education, it is likely that the heavy dose of socially imposed costs in recent years impaired educational excellence and threatened academic freedom.

It should be observed that some of the socially imposed costs may be reflected in the price indexes used to deflate higher educational expenditures. To the extent that business firms and other organizations can shift socially imposed costs to buyers, these costs will tend to raise the indexes of both wholesale and consumer prices. Thus, the indexes used to deflate higher educational expenditures contain unknown amounts of socially imposed costs. To count the rise in higher educational expenditures from so-called inflation and then to count the rise in expenditures from socially imposed costs separately would probably involve some improper double counting.

The figures on the amount of socially imposed costs are sheer guesses based upon limited, fragmentary, and ambiguous data. If one were to seek definitive statistics on the amount of socially imposed costs, the following steps would be necessary: (1) to formulate precise definitions of the expenditure items to be included; (2) to select a sizable sample of institutions; (3) to examine their expenditures over at least 10 years in great detail and thus to identify increases in expenditure associated with particular socially imposed costs; (4) to classify the resulting data by types of institutions. Even with considerable care, the estimates would still be ambiguous. The outcome would depend on more or less arbitrary judgments of what expenditure items to include. Also, many of the relevant costs do not show up in accounting records and depend on arbitrary assumptions (e.g., faculty time used in complying with government regulations, amount of shared costs, or extent of backward shifting of fringe benefits). Indeed, such studies may not be worth the effort. They would impose upon institutions still another costly data-collection effort laid on by outside agencies would produce results having spurious accuracy and would achieve little more than the satisfaction of curiosity.

APPENDIX

ANECDOTAL INFORMATION

Some insight into the impact of socially imposed costs can be gleaned from miscellaneous reports contained in various public statements of college and university executives. The most complete and detailed of these comes from the University of Iowa where a survey was made throughout the institution on "The Impact of Federal Regulations" (1976). Other useful documents were provided by Harvard University (Bok, 1976), Ohio State University (Enarson, 1976), and the University of Wisconsin (undated, probably 1976). The Southern Association of Colleges and Schools (1976) conducted a survey of member institutions and assembled the results in the form of anecdotal reports. Also, various bits of information on the amount of socially imposed costs appeared from time to time in the daily press, news magazines, and other publications. In most cases, the information is merely illustrative. Often, the definitions of categories of cost are ambiguous and the amounts are unclear. Nevertheless, one does get from these scattered sources a sense of the magnitudes involved.

Some gleanings from these sources are contained in the following three subsections:

(a) *Proliferation of Federal Laws and Regulations Affecting Higher Education*

"In 1972 the Office of Education published 32 documents in the Federal Register. In 1976, it expects to publish 270 official notices and regulations." [Cheit, 1975, p. 32.]

"Nearly 400 Federal programs now directly affect higher education." [U.S. News and World Report, July 5, 1976, p. 91.]

". . . there are over 400 different laws which affect colleges directly, producing many thousands of separate regulations." [Mezvinsky, 1976, p. H9662.]

The Library of Congress identified 439 separate statutory authorities affecting postsecondary education. In 1973, the National Commission on the Financing of Post-Secondary Education identified some 375 separate programs lodged in more than 35 agencies besides the U.S. Office of Education. The programs are placed "among 18 or 22 standing committees of the House and 16 of 18 Senate committees." [Andringa, 1976, p. 28.]

"A study at the University of California found that 229 'unique reports' are regularly sent to 32 federal agencies" [Cheit, 1977, p. 94.]

. . . the Federal Register, where agencies publish their regulations, have grown from 3,450 pages in 1937 to 35,591 pages in 1973, 45,422 pages in 1974 and 60,221 pages in 1975 . . . during the past year Congress enacted 402 laws, whereas in the same period 7,496 new federal regulations appeared. [McGill, 1976, p. 11. See also Lilley and Miller, 1977, p. 50.]

(b) *Illustrative Indications of Costs of Particular Federal Programs*

Program	Institution	Cost
Unemployment compensation	University of Iowa	Annual benefits $120,000; clerical and administrative costs, $17,000.
Workman's compensation	do	Annual compliance, $10,000.
Employment Retirement Income Security Act (ERISA).	do	Annual compliance, $2,000.
Occupational health and safety	do	Capital, $25,000,000 over a period of years. Annual administrative cost, $102,000.
	Ohio State University	Capital, $9,100,000 over a period of years. Preparatory activity, $885,000 over the last 2 years.
	Unidentified southern university.	Capital, $50,000,000 in a single hospital complex.
Radiation safety	University of Iowa	Annual compliance, $1,000.
Affirmative action	University of California, Berkeley.	Planning document required 70,000 separate statistical calculations.
	University of Wisconsin	Planning document ran to 6,000 pages in 16 volumes and required over 100,000 analyses of employee actions. Personnel office addition of more than $300,000 a year plus expenses at department and college levels.
	University of Iowa	Addition of 5 new administrative positions plus $10,000 in computer costs and costs in departments and colleges, e.g., $8,000 in College of Dentistry.
Student financial aid	Ohio State University	Last year cost was $55,000 for administration of Basic Educational Opportunity Grants (BEOG's). This year another full-time person added.
	University of Iowa	Cost of financial needs analysis increased from $53,000 in 1967 to $186,000 in 1975. Staff increased from 5 to 29 persons from 1967 to 1977.
	University of Wisconsin	Administration of Guaranteed Student Loans will cost additional $50,000. Administration of BEOG's will require 2½ professional persons plus support staff.
Buckley amendments	Ohio State University	$250,000 in staff time and computer charges to establish the program.
	University of Iowa	1 clerical person added.
Wages and hours	do	Compliance, $6,000 a year.
Minority student recruitment and counseling.		New procedures and 2 new staff members.
Environmental protection	Ohio State University	$50,000 a year in new costs to haul solid waste to land fill.
	University of Iowa	Waste disposal, $64,000. Routine compliance with environmental quality standards, $2,500 annually.
Research on human subjects	Ohio State University	Annual compliance, $25,000 and probably as much again in staff time.
	University of Iowa	Part time of several persons.

(c) *Estimates of Overall Costs of Federal Regulation to Particular Universities*

"At Harvard, the total cost of administering five government programs—equal employment opportunity laws, the Buckley amendment, occupational health and safety rules, environmental protection, and pension reform—has been running from $4.6 to $8.3 million." [Bok, 1976, p. 20.]

"At Harvard, for example, compliance with Federal regulations consumed over 60,000 hours of faculty time in 1974–75 alone." [Bok, 1976, p. 18.]

"Recently a survey was conducted at the University of Iowa to determine the impact of Federal regulations. It was estimated that the direct dollar costs alone amount to approximately $6 million annually, which represents 4 percent of the university budget." [Mezvinsky, 1976, p. H9662.]

"At the University of Rochester, government redtape is estimated to add $1 million annually to operating costs. The University of Kansas, which gets about $24 million in Federal funds annually, spends more than $750,000 a year on compliance reports." [U.S. News and World Report, 1976, p. 92.]

"* * * Columbia University spends easily in excess of 1 million each year in meeting its various Federal reporting obligations." [McGill, 1976, p. 13.]

Compliance regulations of the Federal Government "represents 2 percent of our budget." [Southern Association of Colleges and Schools, 1976, p. 12.]

"* * * we spend better than 13 percent of the total effort output of the college campus in involvement in Federal reporting." [Ibid., p. 13.]

"* * * we are using the equivalent of approximately 13 full-time staff members to handle all of the paper work, reporting, etc., required by the Federal Government." [Ibid., p. 13.]

"Indeed, if the $2 billion figure [for the cost to all of higher education] mentioned in a national magazine is accurate, the Southern Association of Colleges and Schools (SACS) estimates are very, very low." [Ibid., p. 13.]

"We have made actual comparisons in the report requirements imposed at the Federal level over the past 5 years and find that in our instance the volume of material has grown 1600 percent." [Ibid., p. 6.]

"* * * current annual cost resulting from the Federal Government regulatory impact upon the operations of ——— University is approximately 8.41 full-time equivalent (F.T.E.) positions at an expense of about $126,745." [Ibid., p. 10.]

"The total cost of this activity and the attendant supplies and equipment consumed total at least $439,000 for the fiscal year." [Ibid., p. 14.]

"Of a total annual budget of approximately $6 million, between 14 and 20 percent is expended for the purposes as above stated." [Ibid., p. 16.]

Bibliography

Andringa, Robert C. "New Demands by Government for More Information From Postsecondary Education." Address delivered at the Second Forum on New Planning and Management Practices in Postsecondary Education, Chicago. November 16, 1973 (mimeo).

Andringa, Robert C. "The View From the Hill." Change, April 1976, pp. 26–30.

Bender, Louis W. Federal Regulation and Higher Education. Washington: American Association for Higher Education, 1977.

Birnbaum, Robert. "Unionization and Faculty Compensation: Part II." Educational Record, spring 1976, pp. 116–18.

Bok, Derek C. "Harvard University: the President's Report, 1974–75." Cambridge: Harvard University, 1976.

Bork, Robert H. "Can Universities Escape the Regulation They Prescribe for All Other Institutions?"

Bowen, Howard R. "Faculty Salaries: Past and Future," Educational Record, winter 1968, pp. 9–21.

Brewster, Kingman. "Higher Education and the Federal Government," Congressional Record, March 10, 1975, pp. S3515–S3517.

Brown, William W., and Stone, Courtenay C., "Academic Unions in Higher Education." Economic Inquiry, July 1977, pp. 385–396.

Chambers, M. M. "Should State Legislatures Insist on Appropriating Federal Funds?" Normal, Ill.: 1977.

Cheit, Earl F. "What Price Accountability?" Change, November 1975, pp. 30–34. 60.

———. "The Benefits and Burdens of Federal Financial Assistance to Higher Education," American Economic Review, February 1977, vol. 67, No. 1, pp. 90–95.

———. "Do Unionized Faculty Members Get Bigger Pay Increases?" Chronicle of Higher Education, December 6, 1976.

Enarson, Harold L. "The Common Good: Foundation for Partnership," Educational Record, spring 1977, pp. 123–31.

———. Restoring the Partnership. Columbus: Ohio State University, 1976 (mimeo).

Finn, Chester E., Jr. "Federalism and the Universities: The Balance Shifts," Change, winter 1975–76, pp. 24–29, 63.

Lilley, William, III, and Miller, James C., III. "The New 'Social Regulation,'" The Public Interest, spring 1977, pp. 49–61.

McGill, William J. "The University and the State," Educational Record, spring 1977, pp. 132–145.

Mezvinsky, Edward. "Higher Education Strangling in Red Tape," Congressional Record, September 9, 1976, pp. H9662–H9663.

Oaks, Dallin H. "A University President Looks at Government Regulation," address to the Association of Governing Boards of Universities and Colleges, St. Louis, Missouri, October 18, 1976. Excerpted in AGB Report, January–February 1977, pp. 41–46.

Saunders, Charles B. "Association View of Federal Impact on Education," Educational Record, spring 1975, pp. 89–95.

Saunders, Charles B., Jr. "Easing the Burden of Federal Regulation: The Next Move Is Ours," Educational Record, vol. 57, No. 4, pp. 217–224.

Southern Association of Colleges and Schools. "Colleges and Universities Speak Out on Cost of Federal Compliance in Higher Education." Atlanta: Commission on Colleges, Southern Association of Schools and Colleges, 1976.

University of Iowa. "Impact of Federal Regulations on the University of Iowa: A Compilation of Campus-Wide Responses," Iowa City: Office of the Vice President for Educational Development and Research, University of Iowa, 1976 (mimeo).

University of Wisconsin. "Report on Questionable Federal Regulatory Activities," undated (mimeo).

U.S. News and World Report. "Another Campus Revolt—This Time Against Washington," July 5, 1976, pp. 91–94.

Van Alstyne, Carol, and Coldren, Sharon L., "The Costs of Implementing Federally Mandated Programs at Colleges and Universities." Washington: American Council on Education, 1976.

Wile, Howard P. "What's So Special About Higher Education?" NACUBO, Studies in Management, vol. 3, No. 8, 1974.

REGULATION IN THE FEDERAL SYSTEM: WHAT DO WE WANT FROM THE INTERGOVERNMENTAL SYSTEM?

By Robert B. Hawkins, Jr.

CONTENTS

	Page
What do we want from the intergovernmental system?	459
What is the intergovernmental system?	461
The Dow siting controversy	463
The interests of key actors	469
Evaluating the intergovernmental system	471
Reform and intergovernmental system	472
Summary	473

In the fall of 1974 executives of Dow Chemical Company began discussions with local, State and Federal officials over the feasibility of building a new petro-chemical plant in the eastern reaches of the San Francisco Bay. Two years later Dow terminated its plans for the new plant, deciding instead to expand existing facilities in Texas and Louisiana. In part, this decision was fostered by a complex intergovernmental-regulatory system. Dow was required to deal with four planes of government, with 28 different departments or agencies, that required 65 permits for construction. Proponents and opponents use this case as an example of how the intergovernmental system either works or doesn't work. For our purposes it provides an excellent opportunity to explore a number of important issues regarding the workings of the intergovernmental system.

The Dow case, and more recently the decision of Standard of Ohio to terminate its five year plans to sell Alaskan oil overland from Los Angeles, highlight trends in the American intergovernmental system: Each year all planes of government call upon the intergovernmental system to do more. Each year it takes on new responsibilities and becomes more complex. Each year it has a greater impact, whether intended or not, upon governmental and private sector decision making. The key policy question in light of this growing complexity is whether policy makers in this system can make decisions that are timely, efficient and sound, or whether institutional changes need to be made in order to realize these goals.

In this case study we examine the intergovernmental system with the purpose of assessing its effectiveness in regulating public and private behavior. In the first section we begin by asking what we want from this system, describe recent reforms and end with a description of the key operational characteristics of the intergovernmental system. We integrate this description with a set of policy questions that need to be asked regarding the capabilities and limitations of the intergovernmental system. In the second section we provide a description

of the decision making system, decisional processes, key actors and interest groups that were involved in the Dow case. We treat not only the development of this issue but also how the intergovernmental system, with its various positions of authority, was used by proponents and opponents to affect the final outcome. In the last section we draw policy conclusions from the case study and suggest possible remedies.

What Do We Want From the Intergovernmental System?

Or: How do we evaluate the Dow experience? To provide an understandable answer we must briefly trace the efforts to reform this system over the last 50 years. Beginning in the late 1890's a fundamental critique of the American Federal system began under the leaderships of such prominent figures as Woodrow Wilson and supported by the emerging populist movement. The dominant assumption of this critique was that a modern industrial society with its increasing interdependence needed unified political and administrative leadership.[1] From such a perspective it was not hard to see that the Federal system with its multiplicity of semi-autonomous units of government was a pathological form of organization. While the Depression and the administration of Franklin Roosevelt propelled Federal dominance over State and local communities, it is also fair to say that constitutional protections, a political system built on a compounded base and strong tradition, have frustrated efforts at reform. Because structural reforms have been hard to come by, reformers have changed their tactic to one of streamlining the intergovernmental system: simplification and programatic rationalization have been the key goals. Two recent reforms illustrate this trend.

In the early 1960's regionalization of Federal programs into 10 Federal regions, with State and local officials working with their regional counterparts, was thought to be an answer to the growing inefficiencies in program development and delivery. In the early 1970's the President, by executive order, created 10 regional councils in a further attempt to strengthen regions and to integrate Federal programmatic efforts. During this same period of time, through Federal stimulus, local-regional organizations, known as council of governments, were formed in most of the local regions of the United States. The goal of these organizations was to integrate local decision making on regional issues and to interface with Federal regional agencies. In neither case have the results been impressive. Administrative integration is not a substitute for political integration. Neither State, local nor Federal officials, all with legal authority and political standing, have been willing to delegate real authority or political power to these organizations.[2]

The New Federalism of the Nixon administration sought the same ends as previous reforms through simplification of the grants process

[1] Vincent Ostrom, "The Intellectual Crisis of American Public Administration," University of Alabama Press. 1973, see chapters one and two. Daniel J. Elazar, "Is Federalism Compatible with Prefectorial Administration?", a paper presented at the 1978 annual meeting of the American Political Science Association. The assumptions of reform fit nicely with the administrative theories of Max Weber and Talcott Parsons.

[2] Robert B. Hawkins, Jr., "Regionalism: A Salable Commodity in Today's Political Market." Paper presented at the National Conference on American Federalism in Action, ACIR, February 21, 1974.

and decentralization. Revenue sharing and block grants were the first thrusts of this program, with States and regional agencies being given the first right of refusal to implement Federal programs in such areas as environmental protection, health and safety, and areawide planning.[3] While the goal was to simplify and streamline the intergovernmental system, revenue sharing and block grants are already starting to look more like categorical programs. Thus, while simplification remains the mainstay in the rhetoric of reform, its realization in practice is far from certain.

However, while the reforms of New Federalism may have failed to simplify and rationalize, these forms are still in place and are directly related to the Dow case. States and regional agencies, especially in the environmental area, are integral parts of the decision making process.[4] In fact, the advent of this system of decision making is relatively new to the intergovernmental process. States and regional agencies are asked to enforce Federal standards and rules, where in the past the Federal Government enforced its own rules and regulations or States carried out broad national programs under their own rules that were negotiated with Federal agencies.[5] These new systems of regulation are predicated on several assumptions regarding performance.

(1) It is assumed that there are very real limitations to national policy making and implementation. Decentralized systems are predicated on the assumption that they are more efficient.

(2) It is also assumed that State and regional structures will realize the following goals:

Will provide a rational planning system;

Will provide a system of conflict resolution;

Will internalize national goals;

Will take into account significant external costs;

Will allow diverse groups of interest to participate in decision making; and

Will allow for more efficient trade-offs between competing interests and goals.

Whether these performance standards are met is both an empirical and normative question. To understand one, requires the other. Just as importantly, simplification, rationalization, and decentralization have demonstrated serious institutional weaknesses.[6] They have not worked as predicted. After many reforms the system is neither simplified nor decentralized. Why these results? A closer look at the fundamental tenets of the American Federal system aids in answering this question.

---

[3] Robert D. Thomas and Ralph Luken, "Balancing Incentives and Conditions in the Evolution of a Federal Program: A Perspective on Construction Grants for Waste Water Treatment Plants," Publius, Volume 4, No. 3, Summer 1974, pp. 43–64. For a fuller treatment of these programs and their foundations see James L. Sundquist, "Making Federalism Work," The Brookings Institution, Washington, D.C., 1969.

[4] Charles O. Jones, "Federal-State-Local Sharing in Pollution," Publius, Volume 4, No. 1, Winter, 1974, pp. 69–86. Also see Carl W. Stenberg and David B. Walker, "The Block Grant: Lessons from Two Early Experiments," Publius, Volume 7, No. 2, Spring 1977, pp. 31–61, for an excellent discussion of the logic of intergovernmental reform.

[5] For an excellent treatment of how this change has occurred in the area of transportation see, Arthur E. Bauer, "Solving Transportation Problems in the Federal System: Is There a Role for State and Local Governments." Publius, Volume 8, No. 2, Spring 1978, pp. 77–98.

[6] See Vincent Ostrom, "The Third Century: Some Anticipated Consequences of Government Reorganization," Publius, Volume 8, No. 2, Spring 1978, pp. 121–140. Also see in this same volume an excellent debate between national leaders on just what government reorganization means.

WHAT IS THE INTERGOVERNMENTAL SYSTEM?

It is not a neutral system of organization. Whether intended or not it is the direct result of the American constitutional system. Its fundamental foundations are not fiscal flows or decentralized structures; they are constitutional, legal and political. To count and enumerate the numbers of governmental units or to list fiscal transactions between plans of government is to look at symptoms rather than causes. In fact, if one concentrates on these symptoms, the system becomes so complex as to defy comprehension, leading one to call for simplification. This is particularly true in the last 20 years when the velocity of government has increased so dramatically.[7]

If the intergovernment system is the product of activities between units of government, we can draw out one of its most important design characteristics. The intergovernmental system does not operate from principles of hierarchy. It is not a common structure, with State and local governments being the operational units. It is precisely the noncentralized nature of the intergovernmental system that must be understood in order to make any sense out of the Dow experience and similar cases. By way of a shorthand method we will outline the key operational presumptions of the intergovernmental system:

(1) *Structural considerations.*—Based on a noncentralized system of governance, where power is divided and then shared, the federal system unites separate entities under an overarching political system. Where State constitutions and/or political cultures support home rule the noncentralized nature of the system is increased.[8] Even today, State and local governments are relatively immune from Federal interference into matters regarding their own structure.[9] Although legal analysis of our constitutional system is no longer in vogue, two critical dimensions must be studied. First, our constitutional system conveys or grants acting authority to a wide variety of actors, through political-legal institutions. The very definition of a political institution in our system is one that has the authority, or decision making capability, to make authoritative decisions. Second, this system of dispersed authorizations is also an opportunity structure; it creates multiple points of access through which diverse groups enter the political system.[10] The importance of these two points to the intergovernmental system is obvious: Participants have compound opportunities to pursue goals through political action.

(2) *Politics and the intergovernment system.*—Traditional interpretations of the Constitution held that it created a politics of separation—dual federalism. Recent research indicates that this doctrine was more of a theoretical fiction and a staple of political rhetoric than

---

[7] Using any set of fiscal or regulatory indicators only one conclusion can be reached: In the last 20 years the velocity of governmental activity in and through the intergovernmental system has increased dramatically. For an excellent summary of these trends see David B. Walker, "A New Intergovernmental System in 1977," Publius, Volume 8, No. 1, Winter 1978, pp. 79–100.

[8] Vincent Ostrom, "The Political Theory of a Compound Republic," Center for Public Choice, Blacksburg, Va., 1971.

[9] The most recent and perhaps trend setting event in this area of the Supreme Court's decision in *Usery* v. *The National League of Cities.* The Court held that the Congress had exceeded its powers under the Commerce clause when it attempted to set employment standards for States and their subdivisions.

[10] The integration of legal, economic and behavioral factors of collective action is to be found in the work of John R. Commons, "The Legal Foundation of Capitalism," University of Wisconsin Press, 1950. Daniel J. Elazar, "American Federalism: A View from the States," New York, Thomas A. Crowell Company, New York, 1972, Chapter Two.

an operational principle of political practice. Historical research into political practice during the 18th century indicates that conflict, negotiation, compromise and the joint production of services best describe the intergovernmental system.[11] Daniel Elazar finds the roots of this cooperative federalism in three dominant elements of the American experience: The existence of a frontier where the Federal Government dominated until States were formed, the decentralized building blocks for political organizations, and the American constitutional structure. In analyzing this structure he found that there were 17 provisions for either concurrent or cooperative behavior between the Federal and State governments.[12] This structure, plus a supporting political structure, is the crucible from which the dominant traits of intergovernmental politics emerged.[13] States and local governments became the institutional structures which groups of interest used to make authoritative decisions. Since no State or groups held a monopoly on decision making rights and since the Constitution provides authority for each plane of government to undertake collective action, a system of conflict and conflict resolution emerged. With the notable failure of this system during the Civil War, it has worked. How well is of course an empirical question.

(3) *The courts.*—Since the ability to make authoritative decisions is a necessary ingredient to political standing, in the intergovernmental system, the courts have played a critical role. To maintain such cherished notions as partnership or cooperative federalism implies that all parties have some standing to make claims. Or if the decision making rights of States and local governments can be abridged by the national government through mere legislation, then legal and political standing are of less importance. The Supreme Court throughout our history has played a critical role in defining the rights of parties to the intergovernmental process. Over the last 40 years the rights of States and local governments to act have been greatly limited through the expansion of the Commerce Clause. Yet, recent court decisions indicate that the pendulum may be swinging back. The right of States and local governments to determine their own structure, to set standards higher than Federal standards and the integrity of State courts have all been affirmed in recent decisions.[14] These decisions all have the effect of increasing the ability of these units of government to sustain their political identity and to pursue political agendas with the Federal Government on a more equitable basis.

(4) *Service provision through the intergovernmental system has been a product of the aforementioned incentives and structures.*—Beginning with the Federal assumption of State Revolutionary War debts, continuing with the development of the joint-stock company to undertake internal improvements in the early 1800's, and leading to

---

[11] Daniel J. Elazar, "The American Partnership," University of Chicago Press, Chicago, 1962.

[12] Ibid., p. 309.

[13] Op. cit. Elazar's view of political culture, its variation and underpinning of American politics has found empirical substantiation in recent work. See Ira Sharkansky. "The Utility of Elazar's Political Culture: A Research Note," Polity, 2 Fall 1968, pp. 68–83 and Ira Sharkansky, "Regionalism, Economic Status and the Public Policies of American States," Social Science Quarterly, Volume 49, June 1968, pp. 9–26.

[14] For example see the following cases: *California* v. *U.S.*, 1978, where the Supreme Court reaffirmed a significant role for States in the management of all water projects. *Usery* v. *The National League of Cities*, and *Stone* v. *Powell*, 1976, where the court reaffirmed the integrity of State court processes.

the Morrell Act, there has been a strong tendency for joint or cooperative provision. Even with the quantum jump in Federal activity during the Depression these tendencies held. It wasn't until the 1960's that the Federal Government began to exercise greater authority over the intergovernmental system, the topic of this paper.[15]

It is within this framework that the intergovernmental system is created and operates. We can best describe the intergovernmental system as one that is founded upon the authority of various institutions to take and sustain decisions. Intergovernmental relationships, conflict and conflict resolution arise because joint authority exists and requires that competing institutions and actors be brought into some form of action. From a design perspective our task is to understand how authority can be distributed to this complex system so that decisions are reached responsibly and rationally. The following criteria outline some of the standards that should be used in judging outcome from this system:

Political accountability.—That any decision making structure has clear lines of authority and responsibility, so that decision making areas are clearly defined as to who has standing to enter and by what means decisions will be reached.

Decisional predictability.—That existing procedures insure that decisions will be made upon predictable grounds and within time frames that are tenable and known to all participants. The degree of predictability will vary according to the complexity and newness of a function.

Rational and institutional arrangements.—Do existing arrangements facilitate policy making, taking into account diverse interests and effected parties, while still being capable of making economic, political and social trade-offs? Furthermore, do these arrangements increase the probability that decisions will be made upon rational groups, based on a strategy of knowledge?

The goals inherent in these criteria are political, they are predicated on the importance of access, conflict and resolution through decision. They are also predicated on the assumption that political structures can be rational; that they can weigh multiple ends with multiple means. While laudable, their realization within the intergovernmental system is far from certain; in a word, they are problematic. This assessment becomes clear when we review the Dow case.

THE DOW SITING CONTROVERSY

Four facts go a long way in explaining why the Dow Chemical Company decided to expand its facilities in the San Francisco Bay Area.[16] First, in 1973 Solano County, the location for the new facilities, updated its general plan. Called the Solano County General Plan—1990, it outlined the general policy of the county towards growth. Within the Southeastern County Planning Area, the policy outlined by the county was as follows:

---

[15] Elazar, op. cit.

[16] In this descriptive account of the Dow siting case we have not attempted to deal with the motives or intentions of actors within the intergovernmental system. Nor has it been possible or feasible to interview all of the key actors in the controversy. Instead we have attempted to merely trace the action route by which the decision progressed to final decisions. We have used news files, working memorandum and some personal interviews.

To provide lands for industrial growth and development with a minimum of adverse environmental effect and maximum accessibility.[17]

Second, EPA regulations requiring the production of unleaded gasoline and the availability of Alaskan oil created a steady supply of Naptha from Bay Area refineries, a basic feedstock for many of Dow's products. Third, Dow's existing plant in Contra Costa County did not have sufficient vacant lands on which to build the needed facilities. Finally, increased transportation costs for finished products from Dow's Gulf Coast and eastern plants made the California location attractive.

In 1974 Dow Chemical purchased 2,700 acres of land in Solano County that was adjacent to their Pittsburg plant but separated by the Sacramento River. Dow was not the only industrial concern interested in this area. Several years earlier Anheuser-Busch had purchased industrial land to the north for a large brewery. Likewise, the Atlantic-Richfield Company (ARCO) had purchased land adjacent to Dow's for a new petro-chemical plant, while National Steel, the Southern Transportation Company and Pacific Gas and Electric Company all owned land for industrial expansion.

Dow's proposed plant was to be located in the northeastern reaches of the San Francisco Bay on the Sacramento River. The river, a primary transportation route, separated their existing plant from the one proposed for Solano County. Dow's plans called for expansion of their existing facilities, as well as the construction of the new facilities across the river and tied together by four underwater pipelines. The existing land use for the proposed site was agricultural. On a three year cycle, the land was dry farmed one year, used for sheep grazing the next and allowed to lay fallow during the third year. Located in an inactive earthquake zone, the Southeastern Solano County Planning Area is sparsely populated, representing only a small portion of the county's population. In terms of hydraulic regions the proposed site was influenced and could influence two basins: the San Francisco and Central Valley basins. Ongoing studies by the U.S. Bureau of Reclamation indicated that while the basins do not meet certain standards, the overall quality of the basins was good.

Two environmental facts dominate the region. First, the Susan Marshes, adjacent to the proposed new facilities is a major sanctuary for water fowl, a major recreational area and an environmental-scientific testing ground. Second, is the problem of air pollution. The West Coast is highly affected by temperature inversions. These inversions have the effect of blocking the flow of pollutants into the higher atmosphere. In the San Francisco Bay Area evidence indicates that inversions will occur in two out of every three days. The fact of these inversions when coupled with high population densities explain why the State of California and its local entities have promulgated such stringent pollution standards.

It is within this context that Dow, in the fall of 1974, held meetings with local, regional, State, and Federal agencies in an attempt to determine whether they could meet all of the environmental require-

[17] Solano County General Plan.

ments for their proposed facilities. After these meetings, the leadership of Dow was convinced that they could in fact meet environmental standards and decided to move ahead with the project. In February of 1975 Dow publicly announced its plans to proceed with the project; and, the Governor's Office of Planning and Research designated Solano County as the lead agency for preparing the Environmental Impact Report, with Contra Costa County being designated as a party of special interest. Under the California Environmental Quality Act (CEQA), the two counties were legally responsible for determining all significant environmental impacts that would be associated with the proposed facilities. The CEQA also requires that an open set of hearings be held to discuss the possible impacts of the project. The controlling document for these hearings is a draft Environmental Impact Report (EIR).

After Dow submitted its EIR to both counties in August 1975, nine public hearings were held, plus acceptance of written comments from interested parties. During the review of Dow's EIR, three Federal agencies, 13 State agencies, 4 regional agencies, 11 local governments, 7 proponents, 13 environmental organizations, and 8 individuals took part or submitted written comments. In late October 1975 Dow submitted its final EIR to Solano County for decision. In December 1975 the County Board of Supervisors certified that the EIR was an adequate assessment of all significant environmental impacts and that no adverse environmental impacts existed. In addition the supervisors rezoned 824 acres of Dow's land from agricultural to industrial use, taking this land out of the Williamson Act, a State system for the preservation of agricultural land. For this last action Dow was required to pay $230,000 in deferred property taxes.

Soon after Solano County approved the EIR the Sierra Club, Friends of the Earth and People for Open Space filed a petition in the California Superior Court asking for a writ of mandate naming Solano County as the defendant, with Dow being named the real party. Amongst other things, the petition claimed that Solano County had acted improperly in making its findings in the EIR process, that it had not taken all significant environmental impacts into account and that it had acted improperly in canceling the agricultural preserve contract when it rezoned the needed land for the new plants.

Early in 1975 Dow had begun similar environmental review processes with the State and Federal governments. Federally, the Corps of Engineers, from whom Dow needed four permits, had been designated as the lead agency. Dow also needed permits from the U.S. Department of Fish and Wildlife, the Department of Commerce and the Environmental Protection Agency (EPA). Through the State of California and its subdivisions, Dow was required to obtain permits from the following agencies: Regional Water Quality Control, Water Resources Control Board, Department of Water Resources, Department of Fish and Game, the Reclamation Board, the Seismic Commission and the State Lands Commission. At the local level Dow was required to obtain permits from the Bay Area Air Pollution Control District (hereafter BAAPCD) and from the counties of Solano, Contra Costa, and Sacramento.

TABLE 1.—*Summary of required regulations for new Dow facilities*

Agency	*Action required*
Federal:	
U.S. Army Corps of Engineers	Issue permits for construction in navigable waters (dock, turning basin, water intake structure, outfall, pipelines); certify environmental impact statement.
U.S. Fish & Wildlife Service	Approval of USCE permits.
U.S. Department of Commerce	Approval of USCE permits.
U.S. Environmental Protection Agency.	Approval of USCE permits and NPDES permit.
State:	
California Regional Water Quality Control Board, Central Valley Region.	Issues NPDES permit, issue certificate of conformance for USCE permits.
Department of Fish and Game	Issue stream bed alteration permit.
State Water Resources Control Board.	Issue appropriate water right.
Department of Water Resources	Issue permit to construct dam.
State Reclamation Board	Issue permit to alter levee.
State Lands Commission	Issue lease for easements across State lands.
Bay Area Air Pollution Control District.	Issue permits to construct and operate.
Local:	
Contra Costa County	Certify environmental impact report and issue building, grading, and drainage correction permits.
Solano County	Certify environmental impact report and issue building, grading, and sanitary permits, plus alter zoning and remove land from the Williamson Act, V.
Sacramento County	Issue Use Permit for pipeline construction.

In May 1976 Dow applied to BAAPCD for the first of 13 required construction and operation permits for its first facility. Through delegated Federal and State authority, the BAAPCD had the responsibility of setting ambient air standards for the San Francisco air basin. In the Dow case the BAAPCD had to determine whether the stationary sources of pollutions from the new facilities would exceed the standards set for the basin. In May 1976, the District turned down Dow's first set of permits for construction. While the District engineers had found that the emissions from Dow's first plant were well within the standards for new facilities, they also found that such emissions would cause total basin wide emissions to exceed the authorized standard. Officials of the District indicated that such emissions would violate both Federal and State standards as set down in the Federal Clean Air Act of 1970 and the regulations of the California Air Resources Board.

Several matters of interpretation and discretion caused conflict. First, the District refused to see the new facilities as an extension of the old plant, while Dow argued that the two facilities were an integrated production unit connected by four underwater pipelines. If the District had seen these new facilities as additions to existing ones, Dow would have had to meet less stringent air standards. Second, the District measured total emissions from sources within each facility,

while Dow argued that it should follow EPA regulations of measuring emissions at the boundary of the plant. By adopting such a rule, Dow argued, their facilities would not exceed the ambient air standards of the basin.

Dow appealed the District's decision to an administrative hearing board and the Sierra Club filed a suit to require the District to disqualify the whole project rather than passing on each plant of the new facility. Through informal discussion with the State's Air Resources Board and the EPA, Dow was informed that, while BAAPCD's regulations were more stringent than either State or Federal standards, the regulations were well within the BAAPCD's statutory authority. Thus, while the EPA was willing to accept pollution measurements at the boundaries of a plant, it was bound to accept the BAAPCD's rule of measuring pollutions at the site.

During the summer and fall of 1976 the focus of decision making shifted to the State. In July the Secretary of the State's Resources Agency, the primary environmental agency of the State, requested the Corps of Engineers to delay their permit process until the State of California had satisfied all of its concerns. During the fall, mounting pressures from labor and business groups aroused State legislative leaders to express concern about the siting decision and urged the Governor to speed up the permit process. During this same time, the president of Dow Chemical met with Governor Brown to discuss the issue. One effect of these pressures was a decision of the Governor to hold an integrated public hearing, comprised of the State agencies required to grant permits, as a means of isolating outstanding issues. The meeting also had the purpose of speeding up the State's permit process. Another reason for the meeting was an Attorney General's Advisory Opinion that was to unravel most of the decision reached through the local EIR process. In this opinion, the Attorney General ruled that the EIR, which had been found to be adequate by Solano and Contra Costa counties, was not binding on State agencies. The result of this opinion was that each State agency in the permit process began its own environmental assessment of the proposed project. For example, while Solano County had found that seismic activity from existing faults near the proposed plant proposed no significant environmental impact, the State's Commission on Seismic Safety contended that the Antioch Fault was active, posed a possible environmental threat, and would require further State consideration before permits could be granted.

Also, during the fall, Dow continued its discussion with BAACPD, exploring possible ways in which it could comply with the District's regulations. One proposal offered by Dow was a set of trade-offs between its new plants and its existing plants. According to the EIR and Dow its new 13-plant facility would emit a total of 4,800 pounds of emissions per day. Dow was willing to accept conditional construction and operational permits from BAACPD, in the form of improved emissions requirements on existing facilities. Dow stated publicly that they were willing to reduce emissions by 6,800 pounds per day by 1980 on their old facilities, meaning that there would be a net reduction in emissions, from both plants, of 2,000 pounds a day. In other words, 2,000 pounds less of emissions would be going into

the atmosphere with both plants than is presently the case without the new facilities. This compromise was refused by the districts because they held that existing law and regulations do not permit trade-offs.

On January 20, 1977, Dow Chemical announced that it was withdrawing its proposed plans for an expanded chemical facility in Solano County. Reaction for all interested parties was immediate. Labor, business and local government officials were quick to lay blame on State government and the Governor. The chairman of the Assembly Ways and Means Committee introduced legislation to override existing regulations blocking the Dow project. Environmental groups responded positively to the decision arguing that Dow was attempting to force State and local officials to violate environmental regulations and that Dow's decision to withdraw was more of an internal-company decision over scarce capital resources. The Governor who publicly had adopted a neutral position on the Dow question, drew criticisms from all quarters, especially from environmentalists who thought he was attempting to rig the State permit game in favor of Dow. The Governor's director of the Office of Planning and Research expressed dismay, claiming that the State permit process would have been completed within 70 to 90 days.

Dow's reasons for withdrawal were multiple. There is substance to the argument that scarce capital resources played a part of their decisions. Dow publicly admitted that it was losing $5 million per month because of the delay. Other Dow projects were competing for scarce capital reserves and the uncertainty as to when they could be spent on the Solano project worked against the new facilities. The most salient reason for Dow's withdrawal was that they knew they could not modify the rulings of the BAAPCD. Dow's decision was also tempered by the fact that within three months of their decision they had received permits to expand two of their plants in Texas and Louisiana to produce the products originally planned for the Solano plants.

Calls for regulatory reform in the permit process were immediate. The Speaker of the State Assembly introduced legislation, AB 884, setting a maximum 18 month limit for complex permit approvals or disapprovals, once local EIR's had been completed. Also legislation was passed that increased the State's authority to short circuit local decisions that were not seen as being in the State's general interest. At this time of writing, several other interesting facts emerge as a result of the Dow experience. ARCO, National Steel and the Southern Pacific Transportation Company have all given up their plans to develop lands in the southern part of Solano County. The only company still considering location in the area is the Pacific Gas and Electric Company. With the apparent blessing of the State's Energy Commission, Pacific Gas and Electric is developing its plans to put a major coal burning plant on property directly adjacent to Dow's proposed site. If regulatory reform has taken place, one can hardly notice by the permit process that Pacific Gas and Electric is going through.

The Interests of Key Actors

At first sight, one is tempted to conclude that the existing regulatory and permit process for new industrial plants is absolutely incoherent. A casual reading of any case study leads one to see an archaic system, with fragmented authority, where the rules of the game are either unintentionally rigged in favor of environmental concerns or the rules of the game are in a state of constant change. Before evaluating the intergovernmental system it is of value to draw out the purposes of the major actors in this case study.

While Solano County met the legal requirements of the California Environmental Quality Act, held open meetings, posed serious environmental questions to Dow for answer and attempted to take into account a significant number of environmental and social factors in their decision process, they were not without significant interest. First, it was predicted that the County would realize a 14 percent net benefit in revenues from the new facilities. Second, the new facilities would open the door for the further development of the southern part of the county. There is a third reason that is even more compelling: Solano County's imbalanced economy. Forty-eight percent of its work force is employment by government, 3 percent in agriculture, 7 percent in manufacturing, with the rest being distributed in construction, trade and service. Furthermore, the county has the lowest per capita income of any of the nine Bay area counties and has one of the highest unemployment rates. These facts had lead the county in the late 1960's to develop their general plans with an eye towards developing a diversified economy that was not dependent upon one industry. Thus, the county and its local governments were most interested in seeing Dow locate its facilities in the county.

This interest also affected Dow. It fit nicely with Dow's interest in finding a West Coast facility that would allow them to compete more effectively with their competitors. The key strategy and political question that Dow had to answer was whether to start the permit process through the EIR-local government process or whether to gain the most difficult permits first and then proceed with the EIR. Dow obviously saw merit in starting with the EIR as a means of building support among local government officials, business leaders, the press and labor leaders. Judging by the support they received, this part of the strategy was a success. It is less clear whether Dow had a strategy with State, Federal, and regional governments. If there was one, it was to move aggressively with all governmental units. At the end, it is clear that Dow attempted to focus attention on State decision makers as a means of forcing permit authorizations and of putting pressure on the administrative hearing of BAAPCD to grant the necessary permit to begin construction of the first plant.

Assessing the State and regional governments' role is more difficult. BAAPCD's role on the technical aspects of the issue appears to be straightforward and professional. Dow never questioned their calculations, only their interpretations of the regulations. At no time during the process were the regulations changed. However, interpretations of

regulations regarding where to measure pollutions and what constitutes a new facility are inherently political and are the very stuff of American politics. The unwillingness of BAAPCD to negotiate or modify existing regulations suggests a strong commitment to environmental standards. While there is no way to determine what kinds of relationships existed between BAAPCD and State officials, this much is clear: BAAPCD regulations had been approved by the State's Air Resources Board and a broad range of professional contacts existed between the two agencies. While it is true that BAAPCD had the legal authority to implement guidelines and regulations that are more stringent than the State's, it is also true that pressure from the State could have significantly influenced the District's interpretation of those rules in the Dow decision.

At the State level it is more difficult to determine what, if any, strategy was adopted towards Dow. The general response of the Legislature was positive towards the Dow facility. The critical factor was the executive branch. From press reports and statements by State officials, one can glean at best a neutral strategy. In the winter of 1976, the Governor and his environmental chief came out against the Standard Oil Company of Ohio (SOHIO) proposal for a pipeline to transport Alaskan crude to the Midwest. In the fall of that year the Governor would only say that he would streamline the permit process but that the Dow case would have to stand on its own merits. Likewise, the Governor's director of the Office of Planning and Research, past executive officer of the Planning and Conservation League, publicly took a neutral stand.

If one analyzes strategic actions over a period of time, a different picture emerges. Critical to this analysis is the way in which the decision making area was structured and controlled. The Attorney General's Office of Environment Protection's (long a pro-environmental department) issuance of an advisory opinion, at a strategic time, had the effect of delaying a timely decision by opening up settled issues for study and debate. The State's consolidation permit hearing process in December, by all accounts, was not a fact finding process but rather a political forum used by those against the Dow project. The Air Resources Board also came out openly against the project and used such delaying tactics as asking Dow to study the likely impacts of total development of the area, to openly criticizing Dow's estimates of probable air pollution. Yet the telling tale is the longrun record of the State with regard to industrial development in California. The ARCO facility, planned first for Solano County and then upon encouragement from the administration to move to Southern California was never approved. Dow was disapproved. Standard of Ohio withdrew. National Steel has given up plans to build in the area. Although Pacific Gas and Electric is still considering building a coal powered plant in Solano County, observers predict that it will never be built. Such results should not be surprising given the environmental perspective of most of the appointees in the present administration. Nor should this fact concern us in the present analysis. Again, our interest is in how effectively the system operates.

Finally, the position of the environmentalists is clear: No Dow plant. Dow was never really the key concern of environmentalists. The real issue to the environmentalists was development and violation of the rural areas of Solano County. If Dow were approved, it would open the Susan Marsh road for further development. In a very real sense, environmentalists were playing a zero-sum game. Fortunately for them, they were able to use the rules of the game to make playing a zero game highly profitable.

EVALUATING THE INTERGOVERNMENTAL SYSTEM

With this background let us evaluate the performance of the intergovernmental system. True to form, one finds traditional patterns of politics within this system. Proponents and opponents chose points of access and strategies that they thought would maximize their chances of obtaining favorable hearings and decisions. In terms of access, concerned interests had multiple points of access. Local, regional, State, and Federal agencies held public meetings, took testimony and attempted to integrate their efforts. Furthermore, environmental groups used the courts on two occasions in attempts to overturn decisions taken by political bodies.

Using our general evaluative criteria outlined above, we can draw the following conclusions. In terms of political accountability the following points can be made:

While the system is complex, lines of authority are in most instances clearly specified.

To merely state numbers of permits required, as Dow did, is misleading. Even in an integrated, one-step permit process, the system of permit granting will be complex.

The Attorney General's Advisory Opinion clearly demonstrated a weakness in the system. If State agencies can override local decisions, made through a State mandated process with State participation, then the local EIR takes on the nature of an advisory process rather than a legally binding procedure.

Thus, the Dow case raises questions about who has the means to make decisions regarding industrial development. If State and Federal goals, to be implemented by local governments, are overriden by higher authorities because they disagree with the decisions, then those local structures have very little importance.

If similar industrial siting cases form a pattern, it suggests that predictability is gained from unpredictability. One of the clear tactics used by opponents in the Dow case was to draw the controversy out and make it complex, thus costing Dow not only time and money to present its case, but also to increase the opportunity costs for using scarce resources in more attractive ways. The following points can be made regarding decisional predictability:

There is a predictable bias in California against heavy industrial development.

In the Dow case there were no predictable grounds upon which to expect a decision, since a significant environmental impact is an undefined term.

Time frames are unpredictable and amenable to utilization by opponents as a resource.

It is in the area of rational organization arrangements that the Dow case highlights very real problems. The following points demonstrate these problems:

At no point in the process was there any decision making structure that attempted to integrate the costs and benefits of social, economic, and environmental factors.

The Clean Air Act of 1970 and its 1976 amendments, guarantee that single value decision making will continue.

There are not agreed upon standards of how one would determine the trade-offs between competing values.

There is no limit to the complexities that can be generated by the EIR process. In fact, since there is very little definition to the EIR and what it should entail, opponents can always claim that it is inadequate.

The present process guarantees that decision making will be fragmented.

Present ambient air standards, even with recent allowances for the trading of pollution rights and the bubble concept, are time specific and do not allow for responsible industrial development or at a minimum do not allow for complex trade-offs.

Finally, it is fair to say that the present system imposes unreasonable costs on citizens, corporations and governments. While Dow estimates that it spent over 10 million dollars on the planning of its plant, the costs of environmental concern is probably much higher. In 1974 the California Assembly Committee on Local Government evaluated the EIR process. In terms of costs they found that to implement the Act local governments spent between $50 million and $75 million in 1974. They also found that the Act probably added about $150 to each new housing unit in 1974. In 1974 over 4,000 EIR were written, with local governments processing over 90 percent. In 53 percent of the cases approval was granted, while 31 percent were required to make changes to avoid environmental damage. This study also found that many time local governments approved projects that had serious environmental impacts but were outweighed by social, economical or political justification.

Reform and Intergovernmental System

Calls for one-stop shopping, consolidation of governmental entities, and State land use plans are all popular remedies but probably of little practical value. The following steps appear to hold the most promise for meaningful reform:

Through cooperative Federal-State-regional-local action develop an acceptable sequence for environmental reports and permit processing.

Develop specific time frames, by industry or complexity of issues, for the reaching of a decision.

Decisional processes must have integrity. To require individuals or corporations to go through processes that are time consuming and costly only to have these decisions overturned by other

units of government is inefficient and erodes the legitimacy of the system.

Develop an EIR process that State, Federal, and local officials must participate in and reach decisions through. There is little value in having State and Federal officials participate in local or regional process if their decisions do not hinge on such participation.

Develop a decision-making structure, including concerned governmental agencies, that has the responsibility for determining the costs and benefits of alternative strategies.

Develop a unified State and national statement on what constitutes critical information for the environmental impact.

Give serious study on how environmental standards can be integrated with economic and social standards. There must be trade-offs. High levels of unemployment are not cost free tradeoff for the attainment of pure air.

SUMMARY

From one case study it is impossible to draw definitive conclusions. However, we can draw conclusions regarding process and questions that need further study. Whether the ultimate decision in the Dow case was good or bad is irrelevant. Likewise, to side with one plane of government against others is not only indefensible but obscures both the nature of the problems faced and possible solutions. If local governments are to play a critical role in the EIR and industrial siting process they must have clear instructions from States and the national government. Their authority and range of decisions that they can make must be outlined. Likewise, where local governments clearly are limited in the capacity to make complex decisions regarding State and national forces, these factors should be removed for decisions at the appropriate planes of government. What is clear is that processes must be started that allow for this delineation to take place.

O