OUTLINE AND SELECTED READINGS FOR AAEA LEARNING WORKSHOP

"MODERN POLITICAL ECONOMICS AND ITS APPLICATION TO AGRICULTURAL ECONOMICS"

Wednesday, August 7, 1991

Annual Meeting of the American Agricultural Economics Association

Kansas State University, Manhattan, Kansas

#20212 - 20217

132

Learning Workshop--1991
August 7, Manhattan, Kansas

analytics, p. 292, 324, 372, 395, 449

"MODERN POLITICAL ECONOMICS AND ITS APPLICATION TO AGRICULTURAL ECONOMICS"

MORNING SESSION, 10:15 A.M., Tanya Roberts, presiding

1. "The Historical Underpinnings of Modern Political Economics," Paul T. Heyne, University of Washington

2. "The Economic Analysis of Government Behavior: Choices in the Public Arena," Bruce Yandle, Clemson University

3. "Political Economics and Other Disciplines," Roger Noll, Stanford University

A Period of Open Discussion

Luncheon, K-State Union
Speaker: Mancur Olson, University of Maryland
"The Political Economy of Agriculture"

AFTERNOON SESSION, 1:30 P.M., Paul W. Barkley, presiding

4. "Agriculture Policy and the Political Process," Bruce Gardner, USDA, on leave from University of Maryland

5. "The Political Economics of Environmental Issues," Katherine Reichelderfer, Resources for the Future

6. "Water and the Political Economy," B. Delworth Gardner, Brigham Young University

A Period of Open Discussion

7. "Summary and Integration," Paul W. Barkley, Washington State University

Readings on

THE HISTORICAL UNDERPINNINGS OF POLITICAL ECONOMICS

Policy analysis in the broad manner of Adam Smith

Adam Smith. "Digression concerning the Corn Trade and Corn Laws," from Book IV, Chapter v of *The Wealth of Nations*; the roots and results of "colonialism" from Part Third in Chapter vii of Book IV of *The Wealth of Nations* (1776).

Policy analysis in the narrow manner of David Ricardo

David Ricardo. "On Wages" (1817). Chapter V from *On the Principles of Political Economy and Taxation.*

Policy analysis by John Stuart Mill

John Stuart Mill. "The Corn Laws" (1825). Reprinted in *The Collected Works of John Stuart Mill,* Volume IV: *Essays on Economics and Society.*

Policy analysis by Alfred Marshall

Alfred Marshall. "Where to House the London Poor" (1884). Reprinted in *Memorials of Alfred Marshall*, edited by A.C. Pigou.

Policy Analysis by John Maynard Keynes

John Maynard Keynes. "Concluding Notes on the Social Philosophy towards which the General Theory Might Lead" (1936). Final chapter of *The General Theory of Employment, Interest and Money* .

Joseph Schumpeter's seminal chapters on the "economics of democracy"

Joseph A. Schumpeter. "Socialism and Democracy" (1942). Part IV (Chapters XX-XXIII) in *Capitalism, Socialism and Democracy.*

James Buchanan relates his work to "what has gone before"

James M. Buchanan. "Marginal Notes on Reading Political Philosophy" (1962). Appendix 1 in James M. Buchanan and Gordon Tullock, *The Calculus of Consent.*

Gordon Tullock reviews theoretical antecedents

Gordon Tullock. "Theoretical Forerunners" (1962). Appendix 2 in James M. Buchanan and Gordon Tullock, *The Calculus of Consent.*

As an introduction to "The Historical Underpinnings of Modern Political Economics," contrast the policy analysis characteristic of contemporary economics with Adam Smith's approach in the following two excerpts from *The Wealth of Nations*. --PH

[a]*Digression concerning the Corn Trade and Corn Laws*[a]

1 I cannot conclude this chapter concerning bounties, without observing that the praises which have been bestowed upon the law which establishes the bounty upon the exportation of corn, and upon that system of regulations which is connected with it, are altogether unmerited. A particular examination of the nature of the corn trade, and of the principal British laws which relate to it, will sufficiently demonstrate [291] the truth of this assertion. The great importance of this subject must justify the length of the digression.

2 The trade of the corn merchant is composed of four different branches, which, though they may sometimes be all carried on by the same person, are in their own nature four separate and distinct trades. These are, first, the trade of the inland dealer; secondly, that of the merchant importer for home consumption; thirdly, that of the merchant exporter of home produce for foreign consumption; and, fourthly, that of the merchant carrier, or of the importer of corn in order to export it again.

3 I. The interest of the inland dealer, and that of the great body of the people, how opposite soever they may at first sight appear, are, even in years of the greatest scarcity, exactly the same. It is his interest to raise the price of his corn as high as the real scarcity of the season requires, and it can never be his interest to raise it higher. By raising the price he discourages the consumption, and puts every body more or less, but particularly the inferior ranks of people, upon thrift and good management. If, by raising it too high, he discourages the consumption so much that the supply of the season is likely to go beyond the consumption of the season, and to last for some time after the next crop begins to come in, he runs the hazard, not only of losing a considerable part of his corn by natural causes, but of being obliged to sell what remains of it for much less than what he might have had [292] for it several months before. If by not raising the price high enough he discourages the consumption so little, that the supply of the season is likely to fall short of the consumption of the season, he not only loses a part of the profit which he might otherwise have made, but he exposes the people to suffer before the end of the season, instead of the hardships of a dearth, the dreadful horrors of a famine. It is the interest of

[a-a] *2–6*

following reference to bounties on exportation: 'These in Scotland and England together amount to about £300,000 a year; exclusive of the Bounty upon Corn which in some years has amounted to a sum equal to all the other bounties.' See above, I.xi.g.18.

the people that their daily, weekly, and monthly consumption, should be proportioned as exactly as possible to the supply of the season. The interest of the inland corn dealer is the same. By supplying them, as nearly as he can judge, in this proportion, he is likely to sell all his corn for the highest price, and with the greatest profit; and his knowledge of the state of the crop, and of his daily, weekly, and monthly sales, enable him to judge, with more or less accuracy, how far they really are supplied in this manner. Without intending the interest of the people, he is necessarily led, by a regard to his own interest, to treat them, even in years of scarcity, pretty much in the same manner as the prudent master of a vessel is sometimes obliged to treat his crew. When he foresees that provisions are likely to run short, he puts them upon short allowance. Though from excess of caution he should sometimes do this without any real necessity, yet all the inconveniencies which his crew can thereby suffer are inconsiderable in comparison of the danger, misery, and ruin, to which they might sometimes be exposed by a less [293] provident conduct. Though from excess of avarice, in the same manner, the inland corn merchant should sometimes raise the price of his corn somewhat higher than the scarcity of the season requires, yet all the inconveniencies which the people can suffer from this conduct, which effectually secures them from a famine in the end of the season, are inconsiderable in comparison of what they might have been exposed to by a more liberal way of dealing in the beginning of it. The corn merchant himself is likely to suffer the most by this excess of avarice; not only from the indignation which it generally excites against him, but, though he should escape the effects of this indignation, from the quantity of corn which it necessarily leaves upon his hands in the end of the season, and which, if the next season happens to prove favourable, he must always sell for a much lower price than he might otherwise have had.

4 Were it possible, indeed, for one great company of merchants to possess themselves of the whole crop of an extensive country, it might, perhaps, be their interest to deal with it as the Dutch are said to do with the spiceries of the Molluccas,[1] to destroy or throw away a considerable part of it, in order to keep up the price of the rest. But it is scarce possible, even by the violence of law, to establish such an extensive monopoly with regard to corn; and, wherever the law leaves the trade free, it is of all commodities the least liable to be engrossed or monopolized by the force of a few large capitals, [294] which buy up the greater part of it. Not only its value far exceeds what the capitals of a few private men are capable of purchasing, but, supposing they were capable of purchasing it, the manner in which it is produced renders this purchase altogether impracticable. As in every civilized country it is the commodity of which the annual consumption is the greatest, so a greater quantity of industry is annually employed in producing

[1] The same point is made at I.xi.b.33 and IV.vii.c.101.

corn than in producing any other commodity. When it first comes from the ground too, it is necessarily divided among a greater number of owners than any other commodity; and these owners can never be collected into one place like a number of independent manufacturers, but are necessarily scattered through all the different corners of the country. These first owners either immediately supply the consumers in their own neighbourhood, or they supply other inland dealers who supply those consumers. The inland dealers in corn, therefore, including both the farmer and the baker, are necessarily more numerous than the dealers in any other commodity, and their dispersed situation renders it altogether impossible for them to enter into any general combination.[2] If in a year of scarcity therefore, any of them should find that he had a good deal more corn upon hand than, at the current price, he could hope to dispose of before the end of the season, he would never think of keeping up this price to his own loss, and to the sole benefit of his rivals and competitors, but would immediately lower [295] it, in order to get rid of his corn before the new crop began to come in. The same motives, the same interests, which would thus regulate the conduct of any one dealer, would regulate that of every other, and oblige them all in general to sell their corn at the price which, according to the best of their judgment, was most suitable to the scarcity or plenty of the season.

5 Whoever examines, with attention, the history of the dearths and famines which have afflicted any part of Europe, during either the course of the present or that of the two preceding centuries, of several of which we have pretty exact accounts, will find, I believe, that a dearth never has arisen from any combination among the inland dealers in corn, nor from any other cause but a real scarcity, occasioned sometimes, perhaps, and in some particular places, by the waste of war, but in by far the greatest number of cases, by the fault of the seasons; and that a famine has never arisen from any other cause but the violence of government attempting, by improper means, to remedy the inconveniencies of a dearth.

6 In an extensive corn country, between all the different parts of which there is a free commerce and communication, the scarcity occasioned by the most unfavourable seasons can never be so great as to produce a famine; and the scantiest crop, if managed with frugality and œconomy, will maintain, through the year, the same number of people that are commonly fed in a more affluent manner by one of moderate plenty. [296] The seasons most unfavourable to the crop are those of excessive drought or excessive rain. But, as corn grows equally upon high and low lands, upon grounds that are disposed to be too wet, and upon those that are disposed to be too dry, either the drought or the rain which is hurtful to one part of the

[2] The problem of dispersed situation is frequently mentioned in the discussion of economic power. See, for example, IV.ii.21, IV.vii.b.24, and IV.viii.4, 34.

country is favourable to another; and though both in the wet and in the dry season the crop is a good deal less than in one more properly tempered, yet in both what is lost in one part of the country is in some measure compensated by what is gained in the other. In rice countries, where the crop not only requires a very moist soil, but where in a certain period of its growing it must be laid under water, the effects of a drought are much more dismal. Even in such countries, however, the drought is, perhaps, scarce ever so universal as necessarily to occasion a famine, if the government would allow a free trade. The drought in Bengal, a few years ago, might probably have occasioned a very great dearth. Some improper regulations, some injudicious restraints imposed by the servants of the East India Company upon the rice trade, contributed, perhaps, to turn that dearth into a famine.[3]

7 When the government, in order to remedy the inconveniencies of a dearth, orders all the dealers to sell their corn at what it supposes a reasonable price, it either hinders them from bringing it to market, which may sometimes produce a famine even in the beginning of the season; or if they bring it thither, it enables [297] the people, and thereby encourages them to consume it so fast, as must necessarily produce a famine before the end of the season. The unlimited, unrestrained freedom of the corn trade, as it is the only effectual preventative of the miseries of a famine, so it is the best palliative of the inconveniencies of a dearth; for the inconveniencies of a real scarcity cannot be remedied; they can only be palliated. No trade deserves more the full protection of the law, and no trade requires it so much; because no trade is so much exposed to popular odium.

8 In years of scarcity the inferior ranks of people impute their distress to the avarice of the corn merchant, who becomes the object of their hatred and indignation. Instead of making profit upon such occasions, therefore, he is often in danger of being utterly ruined, and of having his magazines plundered and destroyed by their violence. It is in years of scarcity, however, when prices are high, that the corn merchant expects to make his principal profit. He is generally in contract with some farmers to furnish him for a certain number of years with a certain quantity of corn at a certain price. This contract price is settled according to what is supposed to be the moderate and reasonable, that is, the ordinary or average price, which, before the late years of scarcity, was commonly about eight-and-twenty-shillings for the quarter of wheat, and for that of other grain in proportion. In years of scarcity, therefore, the corn merchant buys a great part of his corn for the ordinary [298] price, and sells it for a much higher. That this extraordinary profit, however, is no more than sufficient to put his trade upon a fair level with other trades, and to compensate the many losses which he sustains upon other occasions, both from the perishable

[3] See below, IV.vii.c.101.

nature of the commodity itself, and from the frequent and unforeseen fluctuations of its price, seems evident enough, from this single circumstance, that great fortunes are as seldom made in this as in any other trade. The popular odium, however, which attends it in years of scarcity, the only years in which it can be very profitable, renders people of character and fortune averse to enter into it.[4] It is abandoned to an inferior set of dealers; and millers, bakers, mealmen, and meal factors, together with a number of wretched hucksters, are almost the only middle people that, in the home market, come between the grower and the consumer.

9 The ancient policy of Europe, instead of discountenancing this popular odium against a trade so beneficial to the publick, seems, on the contrary, to have authorised and encouraged it.

10 By the 5th and 6th of Edward VI. cap. 14.[5] it was enacted, That whoever should buy any corn or grain with intent to sell it again, should be reputed an unlawful engrosser, and should, for the first fault, suffer two months imprisonment, and forfeit the value of the corn; for the second, suffer six months imprisonment, and forfeit double the value; and for the third, be set in the pillory, suffer imprisonment during the [299] king's pleasure, and forfeit all his goods and chattels.[6] The ancient policy of most other parts of Europe was no better than that of England.

11 Our ancestors seem to have imagined that the people would buy their corn cheaper of the farmer than of the corn merchant, who, they were afraid, would require, over and above the price which he paid to the farmer, an exorbitant profit to himself. They endeavoured, therefore, to annihilate his trade altogether. They even endeavoured to hinder as much as possible any middle man of any kind from coming in between the grower and the consumer; and this was the meaning of the many restraints which they imposed upon the trade of those whom they called kidders or carriers of corn, a trade which nobody was allowed to exercise without a licence ascertaining his qualifications as a man of probity and fair dealing. The authority of three justices of the peace was, by the statute of Edward VI. necessary, in order to grant this licence.[7] But even this restraint was afterwards thought insufficient, and by a statute of Elizabeth, the privilege of granting it was confined to the quarter-sessions.[8]

[4] It is remarked above, I.x.b.34, in the discussion of 'net advantages', that of the five circumstances which affect wages, only two are relevant in the determination of profits, the agreeableness of the business and the risk involved.

[5] 5 and 6 Edward VI, c. 14 (1551).

[6] See above, III.ii.21, and below, IV.v.b.26, where Smith compares the popular fear of engrossing and forestalling with the terrors of witchcraft.

[7] 5 and 6 Edward VI, c. 14, s. 5 (1551) allowed for licensing and so made the act less rather than more stringent. But in general the Act was aimed against regrators, forestallers, and engrossers.

[8] 5 Elizabeth I, c. 12 (1562), by transferring the power of licensing to quarter sessions, confirmed the impression of stringency in the provision for concessions.

12 The antient policy of Europe endeavoured in this manner to regulate agriculture, the great trade of the country, by maxims quite different from those which it established with regard to manufactures, the great trade of the towns. By leaving the farmer no other customers but either the [b]consumers[b] or [c]their[c] immediate factors, the kidders and carriers of corn, it endeavoured to [300] force him to exercise the trade, not only of a farmer, but of a corn merchant or corn retailer. On the contrary, it in many cases prohibited the manufacturer from exercising the trade of a shop-keeper, or from selling his own goods by retail. It meant by the one law to promote the general interest of the country, or to render corn cheap, without, perhaps, its being well understood how this was to be done. By the other it meant to promote that of a particular order of men, the shopkeepers, who would be so much undersold by the manufacturer, it was supposed, that their trade would be ruined if he was allowed to retail at all.

13 The manufacturer, however, though he had been allowed to keep a shop, and to sell his own goods by retail, could not have undersold the common shopkeeper. Whatever part of his capital he might have placed in his shop, he must have withdrawn it from his manufacture. In order to carry on his business on a level with that of other people, as he must have had the profit of a manufacturer on the one part, so he must have had that of a shopkeeper upon the other. Let us suppose, for example, that in the particular town where he lived, ten per cent. was the ordinary profit both of manufacturing and shopkeeping stock; he must in this case have charged upon every piece of his own goods which he sold in his shop, a profit of twenty per cent. When he carried them from his workhouse to his shop, he must have valued them at the price for which he could have sold them to a [301] dealer or shopkeeper, who would have bought them by wholesale. If he valued them lower, he lost a part of the profit of his manufacturing capital. When again he sold them from his shop, unless he got the same price at which a shopkeeper would have sold them, he lost a part of the profit of his shopkeeping capital. Though he might appear, therefore, to make a double profit upon the same piece of goods, yet as these goods made successively a part of two distinct capitals, he made but a single profit upon the whole capital employed about them; and if he made less than this profit, he was a loser, or did not employ his whole capital with the same advantage as the greater part of his neighbours.

14 What the manufacturer was prohibited to do, the farmer was in some measure enjoined to do; to divide his capital between two different employments; to keep one part of it in his granaries and stack yard, for supplying the occasional demands of the market; and to employ the other in the cultivation of his land. But as he could not afford to employ the latter for less than the ordinary profits of farming stock, so he could as little afford to

[b–b] consumer *1* [c–c] his *1*

employ the former for less than the ordinary profits of mercantile stock. Whether the stock which really carried on the business of the corn merchant belonged to the person who was called a farmer, or to the person who was called a corn merchant, an equal profit was in both cases requisite, in order to indemnify its owner for employing it in this manner; in order to put his business upon a level with other [302] trades, and in order to hinder him from having an interest to change it as soon as possible for some other. The farmer, therefore, who was thus forced to exercise the trade of a corn merchant, could not afford to sell his corn cheaper than any other corn merchant would have been obliged to do in the case of a free competition.

15 The dealer who can employ his whole stock in one single branch of business, has an advantage of the same kind with the workman who can employ his whole labour in one single operation.[9] As the latter acquires a dexterity which enables him, with the same two hands, to perform a much greater quantity of work; so the former acquires so easy and ready a method of transacting his business, of buying and disposing of his goods, that with the same capital he can transact a much greater quantity of business. As the one can commonly afford his work a good deal cheaper, so the other can commonly afford his goods somewhat cheaper than if his stock and attention were both employed about a greater variety of objects. The greater part of manufacturers could not afford to retail their own goods so cheap as a vigilant and active shop-keeper, whose sole business it was to buy them by wholesale, and to retail them again. The greater part of farmers could still less afford to retail their own corn, [d]or[d] to supply the inhabitants of a town, at perhaps four or five miles distance from the greater part of them, so cheap as a vigilant and active corn merchant, whose [303] sole business it was to purchase corn by wholesale, to collect it into a great magazine, and to retail it again.

16 The law which prohibited the manufacturer from exercising the trade of a shopkeeper, endeavoured to force this division in the employment of stock to go on faster than it might otherwise have done. The law which obliged the farmer to exercise the trade of a corn merchant, endeavoured to hinder it from going on so fast. Both laws were evident violations of natural liberty, and therefore unjust; and they were both too as impolitick as they were unjust. It is the interest of every society, that things of this kind should never either be forced or obstructed. The man who employs either his labour or his stock in a greater variety of ways than his situation renders necessary, can never hurt his neighbour by underselling him. He may hurt himself, and he generally does so. Jack of all trades will never be rich, says

[d-d] *om. 4–6*

[9] Smith comments on the advantages accruing to the London merchant dealing in a single type of linen, as compared to his counterpart in Glasgow or Aberdeen who might handle goods from Scotland, Ireland and Hamburg. See above, 32 n. 6.

the proverb. But the law ought always to trust people with the care of their own interest, as in their local situations they must generally be able to judge better of it than the legislator can do.[10] The law, however, which obliged the farmer to exercise the trade of a corn merchant, was by far the most pernicious of the two.

17 It obstructed, not only that division in the employment of stock which is so advantageous to every society, but it obstructed likewise the improvement and cultivation of the land. By obliging the farmer to carry on two trades in-[304]stead of one, it forced him to divide his capital into two parts, of which one only could be employed in cultivation. But if he had been at liberty to sell his whole crop to a corn merchant as fast as he could thresh it out, his whole capital might have returned immediately to the land, and have been employed in buying more cattle, and hiring more servants, in order to improve and cultivate it better. But by being obliged to sell his corn by retail, he was obliged to keep a great part of his capital in his granaries and stack yard through the year, and could not, therefore, cultivate so well as with the same capital he might otherwise have done. This law, therefore, necessarily obstructed the improvement of the land, and, instead of tending to render corn cheaper, must have tended to render it scarcer, and therefore dearer, than it would otherwise have been.

18 After the business of the farmer, that of the corn merchant is in reality the trade which, if properly protected and encouraged, would contribute the most to the raising of corn. It would support the trade of the farmer in the same manner as the trade of the wholesale dealer supports that of the manufacturer.

19 The wholesale dealer, by affording a ready market to the manufacturer, by taking his goods off his hand as fast as he can make them, and by sometimes even advancing their price to him before he has made them, enables him to keep his whole capital, and sometimes even more than his whole capital, constantly employed in manu-[305]facturing, and consequently to manufacture a much greater quantity of goods than if he was obliged to dispose of them himself to the immediate consumers, or even to the retailers. As the capital of the wholesale merchant too is generally sufficient to replace that of many manufacturers, this intercourse between him and them interests the owner of a large capital to support the owners of a great number of small ones, and to assist them in those losses and misfortunes which might otherwise prove ruinous to them.

20 An intercourse of the same kind universally established between the farmers and the corn merchants, would be attended with effects equally beneficial to the farmer. They would be enabled to keep their whole capitals, and even more than their whole capitals, constantly employed in cultivation. In case of any of those accidents, to which no trade is more

[10] Similar sentiments are expressed, for example, at IV.ii.10, IV.v.b.43, and IV.ix.51.

liable than theirs, they would find in their ordinary customer, the wealthy corn merchant, a person who had both an interest to support them, and the ability to do it, and they would not, as at present, be entirely dependent upon the forbearance of their landlord, or the mercy of his steward. Were it possible, as perhaps it is not, to establish this intercourse universally, and all at once, were it possible to turn all at once the whole farming stock of the kingdom to its proper business, the cultivation of land, withdrawing it from every other employment into which any part of it may be at present diverted,[11] and were it possible, in order to support and assist upon occasion the [306] operations of this great stock, to provide all at once another stock almost equally great, it is not perhaps very easy to imagine how great, how extensive, and how sudden would be the improvement which this change of circumstances would alone produce upon the whole face of the country.

21 The statute of Edward VI.,[12] therefore, by prohibiting as much as possible any middle man from coming in between the grower and the consumer, endeavoured to annihilate a trade, of which the free exercise is not only the best palliative of the inconveniencies of a dearth, but the best preventative of that calamity: after the trade of the farmer, no trade contributing so much to the growing of corn as that of the corn merchant.

22 The rigour of this law was afterwards softened by several subsequent statutes, which successively permitted the engrossing of corn when the price of wheat should not exceed twenty, twenty-four, thirty-two, and forty shillings the quarter. At last, by the 15th of Charles II. c. 7. the engrossing or buying of corn in order to sell it again, as long as the price of wheat did not exceed forty-eight shillings the quarter, and that of other grain in proportion, was declared lawful to all persons not being forestallers, that is, not selling again in the same market within three months.[13] All the freedom which the trade of the inland corn dealer has ever yet enjoyed, was bestowed upon it by this statute.[14] The statute of the twelfth

[11] It is pointed out at II.v.37 that agriculture was 'almost every where capable of absorbing a much greater capital than has ever yet been employed in it'.

[12] 5 and 6 Edward VI, c. 14 (1551).

[13] See above, II.v.10, where the productive role of the merchant is explained.

[14] Smith's use of statutes in support of his argument in this paragraph is confusing. 5 and 6 Edward VI, c. 14 (1551) held 'it shall be lawful to every person or persons not forestalling, to buy engross and keep in his or their garners or houses such corn of the kind aforesaid': wheat at 6*s.* 8*d.* a quarter and other grain at related prices. The distinction was thus made clear, and was confirmed by 5 Elizabeth I, c. 12 (1562); between a forestaller as someone who bought or tried to influence the price of commodities on their way to market, a regrator who bought and sold grain in a market within a radius of four miles, and engrossers who bought growing corn. In 15 Charles II, c. 7 (1663) the provisions of 5 and 6 Edward VI, c. 14 were repeated but in a way which led to the confusion in the text between forestallers and regrators: 'It shall be lawfull for all and every person and persons (not forestalling nor selling the same in the same Market within three Months after the buying thereof) to buy in open Market, and to lay up and keep in his and their Graineries or Houses.' [*continues*]

of the present king, which repeals almost [307] all the other ancient laws against engrossers and forestallers, does not repeal the restrictions of this particular statute, which therefore still continue in force.[15]

23 This statute, however, authorises in some measure two very absurd popular prejudices.

24 First, it supposes that when the price of wheat has risen so high as forty-eight shillings the quarter, and that of other grain in proportion, corn is likely to be so engrossed as to hurt the people. But from what has been already said, it seems evident enough that corn can at no price be so engrossed by the inland dealers as to hurt the people: and forty-eight shillings the quarter besides, though it may be considered as a very high price, yet in years of scarcity it is a price which frequently takes place immediately after harvest, when scarce any part of the new crop can be sold off, and when it is impossible even for ignorance to suppose that any part of it can be so engrossed as to hurt the people.

25 Secondly, it supposes that there is a certain price at which corn is likely to be forestalled, that is, bought up in order to be sold again soon after in the same market, so as to hurt the people. But if a merchant ever buys up corn, either going to a particular market or in a particular market, in order to sell it again soon after in the same market, it must be because he judges that the market cannot be so liberally supplied through the whole season as upon that particular occasion, and that the price, therefore, must [308] soon rise. If he judges wrong in this, and if the price does not rise, he not only loses the whole profit of the stock which he employs in this manner, but a part of the stock itself, by the expence and loss which necessarily ᵉattendᵉ the storing and keeping of corn. He hurts himself, therefore, much more essentially than he can hurt even the particular people whom he may hinder from supplying themselves upon that particular market day, because they may afterwards supply themselves just as cheap upon any other market day. If he judges right, instead of hurting the great body of the people, he renders them a most important service. By making them feel the inconveniencies of a dearth somewhat earlier than they otherwise might do, he prevents their feeling them afterwards so severely as they certainly would do, if the cheapness of price encouraged them to consume faster than suited the real scarcity of the season. When the scarcity is real,

ᵉ⁻ᵉ attends *1–2*

The various statutes determining the prices at which engrossing was permitted are difficult to trace. Smith may have been thinking of statutes which permitted exportation at certain prices. He refers to these statutes at IV.v.b.37, 38.

[15] It is doubtful if Smith's interpretation of 12 George III, c. 71 (1772) is wholly valid. The Act was a general statute repealing several laws against engrossers. 15 Charles II, c. 7 (1663) was not repealed until 10 Edward 7 and 1 George V, c. 8, s. 96 (1910), but the effectiveness of its restrictions on forestallers after the enactment of 12 George III, c. 71 is difficult to see.

the best thing that can be done for the people is to divide the inconveniencies of it as equally as possible through all the different months, and weeks, and days of the year. The interest of the corn merchant makes him study to do this as exactly as he can; and as no other person can have either the same interest, or the same knowledge, or the same abilities to do it so exactly as he, this most important operation of commerce ought to be trusted entirely to him; or, in other words, the corn trade, so far at least as concerns the supply of the home-market, ought to be left perfectly free.

26 [309] The popular fear of engrossing and forestalling may be compared to the popular terrors and suspicions of witchcraft.[16] The unfortunate wretches accused of this latter crime were not more innocent of the misfortunes imputed to them, than those who have been accused of the former. The law which put an end to all prosecutions against witchcraft, which put it out of any man's power to gratify his own malice by accusing his neighbour of that imaginary crime, seems effectually to have put an end to those fears and suspicions, by taking away the great cause which encouraged and supported them. The law which should restore entire freedom to the inland trade of corn, would probably prove as effectual to put an end to the popular fears of engrossing and forestalling.

27 The 15th of Charles II. c. 7. however, with all its imperfections, has perhaps contributed more both to the plentiful supply of the home market, and to the increase of tillage, than any other law in the statute book. It is from this law that the inland corn trade has derived all the liberty and protection which it has ever yet enjoyed; and both the supply of the home market, and the interest of tillage, are much more effectually promoted by the inland, than either by the importation or exportation trade.

28 The proportion of the average quantity of all sorts of grain imported into Great Britain to that of all sorts of grain consumed, it has been computed by the author of the tracts upon the corn trade, does not exceed that of one to five hun-[310]dred and seventy.[17] For supplying the home market, therefore, the importance of the inland trade must be to that of the importation trade as five hundred and seventy to one.

29 The average quantity of all sorts of grain exported from Great Britain does not, according to the same author, exceed the one-and-thirtieth part of the annual produce.[18] For the encouragement of tillage, therefore, by providing a market for the home produce, the importance of the inland trade must be to that of the exportation trade as thirty to one.

30 I have no great faith in political arithmetick, and I mean not to warrant

[16] See above, § 10, and also III.ii.21, where the laws affecting engrossing are described as 'absurd'.

[17] Charles Smith, *Three Tracts on the Corn Trade and Corn Laws* (1766), 145. See above, IV.ii.20, and IV.v.a.8.

[18] Ibid. 144. See above, IV.ii.20 and IV.v.a.8.

the exactness of either of these computations[19]. I mention them only in order to show of how much less consequence, in the opinion of the most judicious and experienced persons, the foreign trade of corn is than the home trade. The great cheapness of corn in the years immediately preceding the establishment of the bounty, may perhaps, with reason, be ascribed in some measure to the operation of this statute of Charles II., which had been enacted about five-and-twenty years before, and which had therefore full time to produce its effect.

31 A very few words will sufficiently explain all that I have to say concerning the other three branches of the corn trade.

32 II. The trade of the merchant importer of foreign corn for home consumption, evidently contributes to the immediate supply of the home market, and must so far be immediately bene-[311]ficial to the great body of the people. It tends, indeed, to lower somewhat the average money price of corn, but not to diminish its real value, or the quantity of labour which it is capable of maintaining. If importation was at all times free, our farmers and country gentlemen would, probably, one year with another, get less money for their corn than they do at present, when importation is at most times in effect prohibited; but the money which they got would be of more value, would buy more goods of all other kinds, and would employ more labour. Their real wealth, their real revenue, therefore, would be the same as at present, though it might be expressed by a smaller quantity of silver; and they would neither be disabled nor discouraged from cultivating corn as much as they do at present. On the contrary, as the rise in the real value of silver, in consequence of lowering the money price of corn, lowers somewhat the money price of all other commodities, it gives the industry of the country, where it takes place, some advantage in all foreign markets, and thereby tends to encourage and increase that industry. But the extent of the home market for corn must be in proportion to the general industry of the country where it grows, or to the number of those who produce something else, and therefore have something else, or what comes to the same thing, the price of something else, to give in exchange for corn. But in

[19] Cantillon remarked: 'There is no branch of knowledge in which one is more subject to error than Statistics when they are left to imagination, and none more demonstrable when they are based upon detailed facts.' (*Essai*, 175, ed. Higgs 133.) In Letter 249 addressed to George Chalmers, dated 10 November 1785, Smith commented that he had 'little faith in Political Arithmetic' and cited as an example the difficulties which had encumbered Alexander Webster's attempt to offer an accurate account of the population of Scotland. Webster (1707–84) had prepared *An Account of the Numbers of People in Scotland in the year 1755* (1755; reprinted in J. G. Kyd, *Scottish Population Statistics*, Scottish Historical Society Publication, 3rd series, xliii (Edinburgh, 1952)). In the same letter, Smith referred to Webster as 'of all the men I have ever known, the most skilful in Politic Arithmetic'. Despite his reservations about political arithmetic Smith was able to refer to the 'ever honoured' Sir William Petty, in Letter 30 addressed to Lord Shelburne, dated 4 April 1759.

every country the home market, as it is the nearest and most convenient, so is it likewise the greatest and most important market for corn. That rise in the [312] real value of silver, therefore, which is the effect of lowering the average money price of corn, tends to enlarge the greatest and most important market for corn, and thereby to encourage, instead of discouraging, its growth.

33 By the 22d of Charles II. c. 13. the importation of wheat, whenever the price in the home market did not exceed fifty-three shillings and four pence the quarter, was subjected to a duty of sixteen shillings the quarter; and to a duty of eight shillings whenever the price did not exceed four pounds.[20] The former of these two prices has, for more than a century past, taken place only in times of very great scarcity; and the latter has, so far as I know, not taken place at all. Yet, till wheat had risen above this latter price, it was by this statute subjected to a very high duty; and, till it had risen above the former, to a duty which amounted to a prohibition. The importation of other sorts of grain was restrained [f]at rates, and[f] by duties[g], in proportion to the value of the grain, almost equally[g] high*. [313] [t]Subsequent laws still further increased those duties.[t]

34 The distress which, in years of scarcity, the strict execution of [j]those laws[j] might have brought upon the people, would probably have been very great. But, upon such occasions, its execution was generally suspended by temporary statutes, which permitted, for a limited time, the importation of foreign corn.[21] The necessity of these temporary statutes sufficiently demonstrates the impropriety of this general one.

[h]* Before the 13th of the present king, the following were the duties payable upon the importation of the different sorts of grain:

Grain.	*Duties.*		*Duties.*	*Duties.*
Beans to 28s. per qr.	19s. 10d. after till	40s.	— 16s. 8d.	then 12d.
Barley to 28s.	19s. 10d.	32s.	— 16s.	12d.
Malt is prohibited by the annual Malt-tax Bill.				
Oats to 16s.	5s. 10d. after			9½d.
Pease to 40s.	16s. od. after			9½d.
Rye to 36s.	19s. 10d. till	40s.	— 16s. 8d.	then 12d.
Wheat to 44s.	21s. 9d. till	53s. 4d.	— 17s.	then 8s.

till 4 *l.* and after that about 1s. 4d.

Buck wheat to 32s. per qr. to pay 16s.

These different duties were imposed, partly by the 22d of Charles II. in place of the Old Subsidy, partly by the New Subsidy, by the One-third and Two-thirds Subsidy, and by the Subsidy 1747.[h] [Smith has apparently taken his table from Charles Smith, *Three Tracts on the Corn Trade and Corn Laws*, 83. Charles Smith claims to have taken his from H. Saxby, *The British Customs*, but, apart from some inconsistencies in the rounding off of some very unwieldy fractions, Charles Smith miscopied some items from Saxby. The table is dervied from Saxby, 111–14.]

[f–f] 2–6 [g–g] proportionably 1 [h–h] 2–6 [i–i] 3–6 [j–j] this statute 1–2

[20] See above, III.iv.20, IV.ii.1, IV.ii.16, IV.v.a.23, and below, IV.v.b.37–8, IV.vii.b.33, V.ii.k.13.

[21] See below, § 38.

35 These restraints upon importation, though prior to the establishment of the bounty, were dictated by the same spirit, by the same principles, which afterwards enacted that regulation. How hurtful soever in themselves, these or some other restraints upon importation became necessary in consequence of that regulation. If, when wheat was either below forty-eight shillings the quarter, or not much above it, foreign corn could have been imported either duty free, or upon paying only a small duty, it might have been exported again, with the benefit of the bounty, to the great loss of the publick revenue, and to the entire perversion of the institution, of which the object was to extend the market for the home growth, not that for the growth of foreign countries.

36 III. The trade of the merchant exporter of corn for foreign consumption, certainly does not contribute directly to the plentiful supply of the home market. It does so, however, indirectly. From whatever source this supply may be usually [314] drawn, whether from home growth or from foreign importation, unless more corn is either usually grown, or usually imported into the country, than what is usually consumed in it, the supply of the home market can never be very plentiful. But, unless the surplus can, in all ordinary cases, be exported, the growers will be careful never to grow more, and the importers never to import more, than what the bare consumption of the home market requires. That market will very seldom be overstocked; but it will generally be understocked, the people, whose business it is to supply it, being generally afraid lest their goods should be left upon their hands. The prohibition of exportation limits the improvement and cultivation of the country to what the supply of its own inhabitants requires. The freedom of exportation enables it to extend [k] cultivation for the supply of foreign nations.

37 By the 12th of Charles II. c. 4.[22] the exportation of corn was permitted whenever the price of wheat did not exceed forty shillings the quarter, and that of other grain in proportion. By the 15th of the same prince,[23] this liberty was extended till the price of wheat exceeded forty-eight shillings the quarter; and by the 22d,[24] to all higher prices. A poundage, indeed, was to be paid to the king upon such exportation. But all grain was rated so low in the book of rates, that this poundage amounted only upon wheat to a shilling, upon oats to four-pence, and upon all other grain to six-pence the quarter. By the 1st of William and Mary,[25] the act which established [315] the bounty, this small duty was virtually taken off whenever the price of

[k] its 1–2

[22] See above, IV.iv.3, and below, IV.viii.41 and V.ii.k.23–4.

[23] 15 Charles II, c. 7 (1663). See above, IV.v.b.22.,

[24] 22 Charles II, c. 13 (1670). See above, III.iv.20, IV.ii.1, IV.ii.16, IV.v.a.23, IV.v.b.33; and below, IV.vii.b.33 and V.ii.k.13.

[25] 1 William and Mary, c. 12 (1688). See also I.xi.g.4, III.iv.20, IV.v.a.8, V.ii.k.13.

wheat did not exceed forty-eight shillings the quarter; and by the 11th and 12th of William III. c. 20. it was expressly taken off at all higher prices.[26]

38 The trade of the merchant exporter was, in this manner, not only encouraged by a bounty, but rendered much more free than that of the inland dealer. By the last of these statutes, corn could be engrossed at any price for exportation; but it could not be engrossed for inland sale, except when the price did not exceed forty-eight shillings the quarter.[27] The interest of the inland dealer, however, it has already been shown, can never be opposite to that of the great body of the people. That of the merchant exporter may, and in fact sometimes is. If, while his own country labours under a dearth, a neighbouring country should be afflicted with a famine, it might be his interest to carry corn to the latter country in such quantities as might very much aggravate the calamities of the dearth. The plentiful supply of the home market was not the direct object of those statutes;[28] but, under the pretence of encouraging agriculture, to raise the money price of corn as high as possible, and thereby to occasion, as much as possible, a constant dearth in the home market. By the discouragement of importation, the supply of that market, even in times of great scarcity, was confined to the home growth; and by the encouragement of exportation, when the price was so high as forty-eight shillings the quarter, that [316] market was not, even in times of considerable scarcity, allowed to enjoy the whole of that growth. The temporary laws, prohibiting for a limited time the exportation of corn, and taking off for a limited time the duties upon its importation, expedients to which Great Britain has been obliged so frequently to have recourse,[29] sufficiently demonstrate the impropriety of her general system. Had that system been good, she would not so frequently have been reduced to the necessity of departing from it.

39 Were all nations to follow the liberal system of free exportation and free importation, the different states into which a great continent was divided would so far resemble the different provinces of a great empire. As among the different provinces of a great empire the freedom of the inland trade appears, both from reason and experience, not only the best palliative of a dearth, but the most effectual preventative of a famine; so would the freedom of the exportation and importation trade be among the different states into which a great continent was divided. The larger the continent,

[26] 11 William III, c. 20 (1698) in *Statutes of the Realm*, vii.610–11; 11 and 12 William III, c. 20 in Ruffhead's edition.

[27] Because of 15 Charles II, c. 7 (1663). See above, IV.v.b.22 and 37.

[28] See above, III.iv.20, IV.ii.1,16, IV.v.a.23, IV.v.b.33; and below, IV.vii.b.33 and V.ii.k.13.

[29] In his *Three Tracts on the Corn Trade*, 44–5, C. Smith lists the major statutes about corn from 1534 to 1766 and then comments that 'although the Bounty hath been before suspended, and the Exportation prohibited, yet, till 1757, the Importation was never allowed duty free' (46). The statute to which he refers is 30 George II, c. 7 (1757), which allowed imports duty free until 25 August 1757.

the easier the communication through all the different parts of it, both by land and by water, the less would any one particular part of it ever be exposed to either of these calamities, the scarcity of any one country being more likely to be relieved by the plenty of some other. But very few countries have entirely adopted this liberal system. The freedom of the corn trade is almost every where more or less restrained, and, [317] in many countries, is confined by such absurd regulations, as frequently aggravate the unavoidable misfortune of a dearth, into the dreadful calamity of a famine. The demand of such countries for corn may frequently become so great and so urgent, that a small state in their neighbourhood, which happened at the same time to be labouring under some degree of dearth, could not venture to supply them without exposing itself to the like dreadful calamity. The very bad policy of one country may thus render it in some measure dangerous and imprudent to establish what would otherwise be the best policy in another. The unlimited freedom of exportation, however, would be much less dangerous in great states, in which the growth being much greater, the supply could seldom be much affected by any quantity of corn that was likely to be exported. In a Swiss canton, or in some of the little states of Italy, it may, perhaps, sometimes be necessary to restrain the exportation of corn. In such great countries as France or England it scarce ever can. To hinder, besides, the farmer from sending his goods at all times to the best market, is evidently to sacrifice the ordinary laws of justice to an idea of publick utility, to a sort of reasons of state; an act of legislative authority which ought to be exercised only, which can be pardoned only in cases of the most urgent necessity. The price at which the exportation of corn is prohibited, if it is ever to be prohibited, ought always to be a very high price.

40 [318] The laws concerning corn may every where be compared to the laws concerning religion. The people feel themselves so much interested in what relates either to their subsistence in this life, or to their happiness in a life to come, that government must yield to their prejudices, and, in order to preserve the publick tranquillity, establish that system which they approve of. It is upon this account, perhaps, that we so seldom find a reasonable system established with regard to either of those two capital objects.[30]

41 IV. The trade of the merchant carrier, or of the importer of foreign corn in order to export it again, contributes to the plentiful supply of the home market. It is not indeed the direct purpose of his trade to sell his corn there. But he will generally be willing to do so, and even for a good deal less money than he might expect in a foreign market; because he saves in this manner the expence of loading and unloading, of freight and insurance. The inhabitants of the country which, by means of the carrying trade,

[30] It is pointed out below, V.i.g.8, that positive law with regard to religion will always be 'more or less influenced by popular superstition and enthusiasm'.

becomes the magazine and storehouse for the supply of other countries, can very seldom be in want themselves. Though the carrying trade [t]might[t] thus contribute to reduce the average money price of corn in the home market, it would not thereby lower its real value. It would only raise somewhat the real value of silver.

42 The carrying trade was in effect prohibited in Great Britain, upon all ordinary occasions, by the high duties upon the importation of foreign [319] corn[m], of the greater part of which there was no drawback[m]; and upon extraordinary occasions, when a scarcity made it necessary to suspend those duties by temporary statutes, exportation was always prohibited. By this system of laws, therefore, the carrying trade was in effect prohibited upon all occasions.

43 That system of laws, therefore, which is connected with the establishment of the bounty, seems to deserve no part of the praise which has been bestowed upon it. The improvement and prosperity of Great Britain, which has been so often ascribed to those laws, may very easily be accounted for by other causes. That security which the laws in Great Britain give to every man that he shall enjoy the fruits of his own labour, is alone sufficient to make any country flourish, notwithstanding these and twenty other absurd regulations of commerce; and this security was perfected by the revolution, much about the same time that the bounty was established.[31] The natural effort of every individual to better his own condition,[32] when suffered to exert itself with freedom and security, is so powerful a principle, that it is alone, and without any assistance, not only capable of carrying on the society to wealth and prosperity, but of surmounting a hundred impertinent obstructions with which the folly of human laws too often incumbers its operations; though the effect of these obstructions is always more or less either to encroach upon its freedom, or to diminish its security. In Great Britain industry is perfectly secure; and though [320] it is far from being perfectly free, it is as free or freer than in any other part of Europe.

[t–t] must *6* [m–m] *2–6*

[31] The link between personal security and economic growth is mentioned at II.i.30, and applied in explaining the rapid rate of growth attained in England, for example, at II.iii.36 and IV.vii.c.54. The same point is made with reference to the English colonies at IV.vii.b.51ff.

[32] The term 'bettering our condition' occurs frequently, for example, at II.iii.28, III.iii.12, and IV.ix.28. Hume in his essay, 'Of Commerce', provides a rather interesting contrast with this passage: 'The poverty of the common people is a natural, if not an infallible effect of absolute monarchy; though I doubt, whether it be always true, on the other hand, that their riches are an infallible result of liberty. Liberty must be attended with particular accidents, and a certain turn of thinking, in order to produce that effect.' He continues: 'Where the labourers and artisans are accustomed to work for low wages, and to retain but a small part of the fruits of their labour, it is difficult for them, even in a free government, to better their condition . . .' (*Essays Moral, Political, and Literary*, ed. Green and Grose, i.297.)

44 Though the period of the greatest prosperity and improvement of Great Britain, has been posterior to that system of laws which is connected with the bounty, we must not upon that account impute it to those laws. It has been posterior likewise to the national debt. But the national debt has most assuredly not been the cause of it.[33]

45 Though the system of laws which is connected with the bounty, has exactly the same tendency with the police of Spain and Portugal; to lower somewhat the value of the precious metals in the country where it takes place; yet Great Britain is certainly one of the richest countries in Europe, while Spain and Portugal are perhaps among the most beggarly. This difference of situation, however, may easily be accounted for from two different causes. First, the tax in Spain, the prohibition in Portugal of exporting gold and silver,[34] and the vigilant police which watches over the execution of those laws, must, in two very poor countries, which between them import annually upwards of six millions sterling,[35] operate, not only more directly, but much more forcibly in reducing the value of those metals there, than the corn laws can do in Great Britain. And, secondly, this bad policy is not in those countries counter-balanced by the general liberty and security of the people. Industry is there neither free nor secure, and the civil and ecclesiastical governments of both Spain [321] and Portugal, are such as would alone be sufficient to perpetuate their present state of poverty, even though their regulations of commerce were as wise as the greater part of them are absurd and foolish.

46 The 13th of the present king, c. 43.[36] seems to have established a new system with regard to the corn laws, in many respects better than the ancient one, but in one [n]or two respects[n] perhaps not quite so good.

47 By this statute the high duties upon importation for home consumption are taken off [o]so[o] soon as the price of [p]middling wheat rises to[p] forty-eight shillings the quarter; [q]that of middling rye, pease or beans, to thirty-two shillings; that of barley to twenty-four shillings; and that of oats to sixteen shillings;[q] and instead of them a small duty is imposed of only six-pence upon the quarter of wheat, and upon that of other grain in proportion. [r]With regard to all these different sorts of grain, but particularly with regard to wheat, the home market is thus opened to foreign supplies at prices considerably lower than[r] before.

48 By the same statute the old bounty of five shillings upon the [s]exportation[s] of wheat ceases [t]so soon as the price rises to forty-four shillings the quarter,

[n–n] respect *1* [o–o] as *1* [p–p] wheat is so high as *1* [q–q] *2–6*

[r–r] The home market is in this manner not so totally excluded from foreign supplies as it was *1* [s–s] quarter *1*

[t–t] when the price rises so high as forty-four shillings, and upon that of other grain in

[33] Smith discusses the impact of a large and growing national debt on economic growth in V.iii.

[34] See above, IV.v.a.19. [35] See above, I.xi.g.33. [36] 13 George III, c. 43 (1772).

instead of forty-eight, the price at which it ceased before; that of two shillings and six-pence upon the exportation of barley ceases so soon as the price rises to twenty-two shillings, instead of twenty-four, the price at which it [322] ceased before; that of two shillings and sixpence upon the exportation of oatmeal ceases so soon as the price rises to fourteen shillings, instead of fifteen, the price at which it ceased before. The bounty upon rye is reduced from three shillings and sixpence to three shillings, and it ceases so soon as the price rises to twenty-eight shillings, instead of thirty-two, the price at which it ceased before.[t] If bounties are as improper as I have endeavoured to prove them to be, the sooner they cease, and the lower they are, so much the better.

49 The same statute permits, at [u]the lowest[u] prices, the importation of corn, in order to be exported again, duty free, provided it is in the mean time lodged in [v]a warehouse under the joint locks of the king and the importer[v]. This liberty, indeed, extends to no more than twenty-five of the different ports of Great Britain. They are, however, the principal ones, and there may not, perhaps, be warehouses proper for this purpose in the greater part of the others.[w]

50 So far this law seems evidently an improvement upon the antient system.

51 [x]But by the same law a bounty of two shillings the quarter is given for the exportation of oats whenever the price does not exceed fourteen shillings. No bounty had ever been given before for the exportation of this grain, no more than for that of peas or beans.[x]

52 [y]By the same law too, the exportation of wheat is prohibited so soon as the price rises to forty-[323]four shillings the quarter; that of rye so soon as it rises to twenty-eight shillings; that of barley so soon as it rises to twenty-two shillings; and that of oats so soon as they rise to fourteen shillings. Those several prices seem all of them a good deal too low, and there seems to be an impropriety, besides, in prohibiting exportation altogether at those precise prices[y] at which that bounty, which was given in order to force it, is withdrawn. The bounty ought certainly either to have been withdrawn at a much lower price, or exportation ought to have been allowed at a much higher.

53 So far, therefore, this law seems to be inferior to the antient system.

proportion. The bounties too upon the coarser sorts of grain are reduced somewhat lower than they were before, even at the prices at which they take place *1*

[u–u] all *1* [v–v] king's warehouse *1*

[w] Some provision is thus made for the establishment of the carrying trade. *1*

[x–x] *2–6*

[y–y] But by the same law exportation is prohibited as soon as the price of wheat rises to forty-four shillings the quarter, and that of other grain in proportion. The price seems to be a good deal too low, and there seems to be an impropriety besides in stopping exportation altogether, at the very same price *1*

[z]With all its imperfections, however, we may perhaps say of it what was said of the laws of Solon, that, though not the best in itself, it is the best which the interests, prejudices, and temper of the times would admit of. It may perhaps in due time prepare the way for a better.[z 37]

[z–z] *2–6*

[37] TMS VI.ii.2.18 makes an interesting point: 'Some general, and even systematical, idea of the perfection of policy and law, may no doubt be necessary for directing the views of the statesman. But to insist upon establishing, and upon establishing all at once, and in spite of all opposition, every thing which that idea may seem to require, must often be the highest degree of arrogance. It is to erect his own judgment into the supreme standard of right and wrong.' The example of Solon is cited in § 16.

PART THIRD

Of the Advantages which Europe has derived from the Discovery of America, and from that of a Passage to the East Indies by the Cape of Good Hope.

1 SUCH are the advantages which the colonies of America have derived from the policy of Europe.

2 What are those which Europe has derived from the discovery and colonization of America?

3 Those advantages may be divided, first, into the general advantages which Europe, considered as one great country, has derived from those great events; and, secondly, into the particular advantages which each colonizing country has derived from the colonies which particularly be-[401]long to it, in consequence of the authority or dominion which it exercises over them.

4 The general advantages which Europe, considered as one great country, has derived from the discovery and colonization of America, consist, first, in the increase of its enjoyments; and, secondly, in the augmentation of its industry.[1]

5 The surplus produce of America, imported into Europe, furnishes the inhabitants of this great continent with a variety of commodities which they could not otherwise have possessed, some for conveniency and use, some for pleasure, and some for ornament, and thereby contributes to increase their enjoyments.

6 The discovery and colonization of America, it will readily be allowed, have contributed to augment the industry, first, of all the countries which trade to it directly; such as Spain, Portugal, France, and England; and, secondly, of all those which, without trading to it directly, send, through the medium of other countries, goods to it of their own produce; such as Austrian Flanders, and some provinces of Germany, which, through the medium of the countries before mentioned, send to it a considerable quantity of linen and other goods. All such countries have evidently gained a more extensive market for their surplus produce, and must consequently have been encouraged to increase its quantity.[2]

7 But, that those great events should likewise have contributed to encourage the industry of countries, such as Hungary and Poland, which [402] may never, perhaps, have sent a single commodity of their own produce to America, is not, perhaps, altogether so evident. That those events have done so, however, cannot be doubted. Some part of the produce of America is consumed in Hungary and Poland, and there is some demand there for the sugar, chocolate, and tobacco, of that new quarter of the world.

[1] See below, § 81.

[2] The vent for surplus as a gain from trade is cited, for example, at IV.i.31.

But those commodities must be purchased with something which is either the produce of the industry of Hungary and Poland, or with something which had been purchased with some part of that produce. Those commodities of America are new values, new equivalents, introduced into Hungary and Poland to be exchanged there for the surplus produce of those countries. By being carried thither they create a new and more extensive market for that surplus produce. They raise its value, and thereby contribute to encourage its increase. Though no part of it may ever be carried to America, it may be carried to other countries which purchase it with a part of their share of the surplus produce of America; and it may find a market by means of the circulation of that trade which was originally put into motion by the surplus produce of America.

8 Those great events may even have contributed to increase the enjoyments, and to augment the industry of countries which, not only never sent any commodities to America, but never received any from it. Even such countries may have received a greater abundance of other commodities from countries of which the surplus [403] produce had been augmented by means of the American trade. This greater abundance, as it must necessarily have increased their enjoyments, so it must likewise have augmented their industry. A greater number of new equivalents of some kind or other must have been presented to them to be exchanged for the surplus produce of that industry. A more extensive market must have been created for that surplus produce, so as to raise its value, and thereby encourage its increase. The mass of commodities annually thrown into the great circle of European commerce, and by its various revolutions annually distributed among all the different nations comprehended within it, must have been augmented by the whole surplus produce of America. A greater share of this greater mass, therefore, is likely to have fallen to each of those nations, to have increased their enjoyments, and augmented their industry.

9 The exclusive trade of the mother countries tends to diminish, or, at least, to keep down below what they would otherwise rise to, both the enjoyments and industry of all those nations in general, and of the American colonies in particular. It is a dead weight upon the action of one of the great springs which puts into motion a great part of the business of mankind. By rendering the colony produce dearer in all other countries, it lessens its consumption, and thereby cramps the industry of the colonies, and both the enjoyments and the industry of all other countries, which both enjoy less when they pay more [404] for what they enjoy, and produce less when they get less for what they produce. By rendering the produce of all other countries dearer in the colonies, it cramps, in the same manner, the industry of all other countries, and both the enjoyments and the industry of the colonies. It is a clog which, for the supposed benefit of some particular countries, embarrasses the pleasures, and encumbers the industry

of all other countries; but of the colonies more than of any other. It [a]not[a] only excludes, as much as possible, all other countries from one particular market; but it confines, as much as possible, the colonies to one particular market: and the difference is very great between being excluded from one particular market, when all others are open, and being confined to one particular market, when all others are shut up. The surplus produce of the colonies, however, is the original source of all that increase of enjoyments and industry which Europe derives from the discovery and colonization of America; and the exclusive trade of the mother countries tends to render this source much less abundant than it otherwise would be.

10 The particular advantages which each colonizing country derives from the colonies which particularly belong to it, are of two different kinds; first, those common advantages which every empire derives from the provinces subject to its dominion; and, secondly, those peculiar advantages which are supposed to result from provinces of so very peculiar a nature as the European colonies of America.

11 [405] The common advantages which every empire derives from the provinces subject to its dominion, consist, first, in the military force which they furnish for its defence; and, secondly, in the revenue which they furnish for the support of its civil government. The Roman colonies furnished occasionally both the one and the other. The Greek colonies, sometimes, furnished a military force; but seldom any revenue. They seldom acknowledged themselves subject to the dominion of the mother city. They were generally her allies in war, but very seldom her subjects in peace.[3]

12 The European colonies of America have never yet furnished any military force for the defence of the mother country. Their military force has never yet been sufficient for their own defence; and in the different wars in which the mother countries have been engaged, the defence of their colonies has generally occasioned a very considerable distraction of the military force of those countries. In this respect, therefore, all the European colonies have, without exception, been a cause rather of weakness than of strength to their respective mother countries.[4]

13 The colonies of Spain and Portugal only have contributed any revenue towards the defence of the mother country, or the support of her civil government.[5] The taxes which have been levied upon those of other European

[a-a] *3-6*

[3] Sometimes revenue was paid: 'These Cotyorites are our colonists, and it was we who gave over to them this land, after we had taken it away from barbarians; therefore they pay us a stated tribute, as do the people of Cerasus and Trapezus.' (Xenophon, *Anabasis*, V.v.10, translated by C. L. Brownson in Loeb Classical Library (1921), 134–5.) The relationship between the Greek colonies and the mother country is described at IV.vii.a.2.

[4] The military costs of the colonies to Great Britain are examined at § 64. See also IV.vii.b.20 and V.iii.92.

[5] See above, IV.vii.b.20.

nations, upon those of England in particular, have seldom been equal to the expence laid out upon them in time of peace, and never sufficient to defray that [406] which they occasioned in time of war. Such colonies, therefore, have been a source of expence and not of revenue to their respective mother countries.

14 The advantages of such colonies to their respective mother countries, consist altogether in those peculiar advantages which are supposed to result from provinces of so very peculiar a nature as the European colonies of America; and the exclusive trade, it is acknowledged, is the sole source of all those peculiar advantages.

15 In consequence of this exclusive trade, all that part of the surplus produce of the English colonies, for example, which consists in what are called enumerated commodities,[6] can be sent to no other country but England. Other countries must afterwards buy it of her. It must be cheaper therefore in England than it can be in any other country, and must contribute more to increase the enjoyments of England, than those of any other country. It must likewise contribute more to encourage her industry. For all those parts of her own surplus produce which England exchanges for those enumerated commodities, she must get a better price than any other [b]countries[b] can get for the like parts of theirs, when they exchange them for the same commodities. The manufactures of England, for example, will purchase a greater quantity of the sugar and tobacco of her own colonies, than the like manufactures of other countries can purchase of that sugar and tobacco. So far, therefore, as the manufactures of England and those [407] of other countries are both to be exchanged for the sugar and tobacco of the English colonies, this superiority of price gives an encouragement to the former, beyond what the latter can in these circumstances enjoy. The exclusive trade of the colonies, therefore, as it diminishes, or, at least, keeps down below what they would otherwise rise to, both the enjoyments and the industry of the countries which do not possess it; so it gives an evident advantage to the countries which do possess it over those other countries.[7]

16 This advantage, however, will, perhaps, be found to be rather what may be called a relative than an absolute advantage; and to give a superiority to the country which enjoys it, rather by depressing the industry and

[b-b] country *1*

[6] The enumerated goods are described at IV.vii.b.25,35.

[7] Pownall commented on this part of Smith's argument, *Letter*, 40: 'You in words advance upon the ground of *probable reasons for believing* only, you prove by probable suppositions only; yet most people who read your book, will think you mean to set up an absolute proof, and your conclusion is drawn as though you had.' See above, II.v, where Smith advances his thesis with regard to the different employments of capital: a thesis on which much of the argument of this section would seem to depend.

produce of other countries, than by raising those of that particular country above what they would naturally rise to in the case of a free trade.

17 The tobacco of Maryland and Virginia, for example, by means of the monopoly which England enjoys of it, certainly comes cheaper to England than it can do to France, to whom England commonly sells a considerable part of it. But had France, and all other European countries been, at all times, allowed a free trade to Maryland and Virginia, the tobacco of those colonies might, by this time, have come cheaper than it actually does, not only to all those other countries, but likewise to England. The produce of tobacco, in consequence of a market so much more extensive than any which it has hitherto enjoyed, might, and probably would, by [408] this time, have been so much increased as to reduce the profits of a tobacco plantation to their natural level with those of a corn plantation, which, it is supposed, they are still somewhat above.[8] The price of tobacco might, and probably would, by this time, have fallen somewhat lower than it is at present. An equal quantity of the commodities either of England, or of those other countries, might have purchased in Maryland and Virginia a greater quantity of tobacco than it can do at present; and, consequently, have been sold there for so much a better price.[9] So far as that weed, therefore, can, by its cheapness and abundance, increase the enjoyments or augment the industry either of England or of any other country, it would, probably, in the case of a free trade, have produced both these effects in somewhat a greater degree than it can do at present. England, indeed, would not in this case have had any advantage over other countries. She might have bought the tobacco of her colonies somewhat cheaper, and, consequently, have sold some of her own commodities somewhat dearer than she actually does. But she could neither have bought the one cheaper nor sold the other dearer than any other country might have done. She might, perhaps, have gained an absolute, but she would certainly have lost a relative advantage.

18 In order, however, to obtain this relative advantage in the colony trade, in order to execute the invidious and malignant project of excluding as much as possible other nations from any share [409] in it, England, there are very probable reasons for believing, has not only sacrificed a part of the absolute advantage which she, as well as every other nation, might have derived from that trade, but has subjected herself both to an absolute and to a relative disadvantage in almost every other branch of trade.

41 The monopoly of the colony trade too has forced some part of the capital of Great Britain from all foreign trade of consumption to a carrying trade; and, consequently, from supporting more or less the industry of Great Britain, to be employed altogether in supporting partly that of the colonies, and partly that of some other countries.

42 The goods, for example, which are annually purchased with the great surplus of eighty-two thousand hogsheads of tobacco annually re-exported from Great Britain, are not all consumed in Great Britain. Part of them, linen from Germany and Holland, for example, is returned to the colonies for their particular consumption. But, that part of the capital of Great Britain which buys the tobacco with which this linen is afterwards bought, is necessarily withdrawn from supporting the industry of Great [424] Britain, to be employed altogether in supporting, partly that of the colonies, and partly that of the particular countries who pay for this tobacco with the produce of their own industry.

43 The monopoly of the colony trade besides, by forcing towards it a much greater proportion of the capital of Great Britain than what would naturally have gone to it, seems to have broken altogether that natural balance which would otherwise have taken place among all the different branches of British industry.[23] The industry of Great Britain, instead of being accommodated to a great number of small markets, has been principally suited to one great market. Her commerce, instead of running in a great number of small channels, has been taught to run principally in one great channel. But the whole system of her industry and commerce has thereby been rendered less secure; the whole state of her body politick less healthful, than it otherwise would have been. In her present condition, Great Britain resembles one of those unwholesome bodies in which some of the vital parts are overgrown, and which, upon that account, are liable to many dangerous disorders scarce incident to those in which all the parts are more properly proportioned. A small stop in that great blood-vessel, which has been artificially swelled beyond its natural dimensions, and through which an unnatural proportion of the industry and commerce of the country has

[g-g] possesses *1*

[23] The concept of a natural balance of industry is developed below, § 97. See also IV.ii.3, IV.ii.12,31, IV.iv.14, and IV.v.a.39.

been forced to circulate, is very likely to bring on [425] the most dangerous disorders upon the whole body politick.[24] The expectation of a rupture with the colonies, accordingly, has struck the people of Great Britain with more terror than they ever felt for a Spanish armada, or a French invasion.[25] It was this terror, whether well or ill grounded, which rendered the repeal of the stamp act[h],[26] among the merchants at least, a popular measure[h]. In the total exclusion from the colony market, was it to last only for a few years, the greater part of our merchants used to fancy that they foresaw an entire stop to their trade; the greater part of our master manufacturers, the entire ruin of their business; and the greater part of our workmen, an end of their employment.[27] A rupture with any of our neighbours upon the continent, though likely too to occasion some stop or interruption in the employments of some of all these different orders of people, is foreseen, however, without any such general emotion. The blood, of which the circulation is stopt in some of the smaller vessels, easily disgorges itself into the greater, without occasioning any dangerous disorder; but, when it is stopt in any of the greater vessels, convulsions, apoplexy, or death, are the immediate and unavoidable consequences. If but one of those overgrown manufactures, which by means either of bounties, or of the monopoly of the home and colony markets, have been artificially raised up to an unnatural height, finds some small stop or interruption in its employment, it

[h-h] a popular measure, among the merchants at least *1*

[24] Pownall rejected the analogy of the blood vessel and argued that the fact that trade had felt no such 'convulsions or apoplexy' on the obstruction of our 'American artery' proved that America was 'not our principal, much less our sole channel of commerce'. He rejected Smith's explanation for this fact, developed below, which relied on the impact of five unforeseen and unthought of events. *Letter*, 45.

[25] Smith was evidently very interested in the current difficulties with America, and prepared a memorandum for Alexander Wedderburn, a former pupil and latterly Solicitor-General in Lord North's administration. The document forms a part of the Rosslyn MSS. (Ann Arbor, Michigan); it is dated February 1778 and endorsed 'Smith's Thoughts on the State of the Contest with America'. It is hereinafter referred to as 'Thoughts on America'. The text is included in the volume of Correspondence which forms a part of this edition of Smith's *Works*. It was first published by G. H. Guttridge in the *American Historical Review*, 38 (1932–3), hereinafter cited as AHR.

[26] The stamp act is mentioned at I.viii.50. Stamp duties are also discussed at V.ii.h.12 where they are stated to be of 'very modern invention'.

[27] In Letter 149 addressed to Smith, dated 8 February 1776, Hume complained about the delay in publication of the WN and reminded his friend that 'If you wait till the Fate of America be decided, you may wait long.' He went on: 'The Duke of Buccleugh tells me, that you are very zealous In American Affairs. My Notion is, that the Matter is not so important as is commonly imagind. If I be mistaken, I shall probably correct my Error, when I see or read you. Our Navigation and general Commerce may suffer more than our Manufactures.' In Letter 233 addressed to William Eden, dated 15 December 1783, Smith wrote: 'I have little anxiety about what becomes of the American commerce. By an equality of treatment to all nations, we might soon open a commerce with the neighbouring nations of Europe infinitely more advantageous than that of so distant a country as America.'

frequently occasions a mutiny and disorder alarming to go-[426]vernment, and embarrassing even to the deliberations of the legislature. How great, therefore, would be the disorder and confusion, it was thought, which must necessarily be occasioned by a sudden and entire stop in the employment of so great a proportion of our principal manufacturers?

44 Some moderate and gradual relaxation of the laws which give to Great Britain the exclusive trade to the colonies, till it is rendered in a great measure free, seems to be the only expedient which can[1], in all future times,[1] deliver her from this danger, which can enable her or even force her to withdraw some part of her capital from this overgrown employment, and to turn it, though with less profit, towards other employments; and which, by gradually diminishing one branch of her industry and gradually increasing all the rest, can by degrees restore all the different branches of it to that natural, healthful, and proper proportion which perfect liberty necessarily establishes, and which perfect liberty can alone preserve. To open the colony trade all at once to all nations, might not only occasion some transitory inconveniency, but a great permanent loss to the greater part of those whose industry or capital is at present engaged in it. The sudden loss of the employment even of the ships which import the eighty-two thousand hogsheads of tobacco, which are over and above the consumption of Great Britain, might alone be felt very sensibly.[28] Such are the unfortunate effects of all the regulations of the mercantile [427] system! They not only introduce very dangerous disorders into the state of the body politick, but disorders which it is often difficult to remedy, without occasioning, for a time at least, still greater disorders. In what manner, therefore, the colony trade ought gradually to be opened; what are the restraints which ought first, and what are those which ought last to be taken away; or in what manner the natural system of perfect liberty and justice ought gradually to be restored, we must leave to the wisdom of future statesmen and legislators to determine.[29]

45 Five different events, unforeseen and unthought of, have very fortunately concurred to hinder Great Britain from feeling, so sensibly as it was generally expected she would, the total exclusion which has now taken place for more than a year (from the first of December, 1774) from a very important branch of the colony trade, that of the twelve associated provinces of North America. First, those colonies, in preparing themselves for their non-importation agreement, drained Great Britain completely of all the commodities which were fit for their market: secondly, the extra-

[1–1] *2–6*

[28] Cf. IV.ii.42 where Smith comments favourably on the ability of the economy to absorb dramatic changes.

[29] See below, IV.ix.51, where Smith describes the system of 'perfect liberty', and cf. I.x.a.1, I.vii.6,30, and IV.ix.17.

ordinary demand of the Spanish Flota has, this year, drained Germany and the North of many commodities, linen in particular, which used to come into competition, even in the British market, with the manufactures of Great Britain: thirdly, the peace between Russia and Turkey has occasioned an extraordinary demand from the Turkey market, which, during the distress of the country, and [428] while a Russian fleet was cruizing in the Archipelago, had been very poorly supplied: fourthly, the demand of the north of Europe for the manufactures of Great Britain, has been increasing from year to year for some time past: and, fifthly, the late partition and consequential pacification of Poland, by opening the market of that great country, have this year added an extraordinary demand from thence to the increasing demand of the North. These events are all, except the fourth, in their nature transitory and accidental,[30] and the exclusion from so important a branch of the colony trade, if unfortunately it should continue much longer, may still occasion some degree of distress. This distress, however, as it will come on gradually, will be felt much less severely than if it had come on all at once; and, in the mean time, the industry and capital of the country may find a new employment and direction, so as to prevent [1]this distress[1] from ever rising to any considerable height.

46 The monopoly of the colony trade, therefore, so far as it has turned towards that trade a greater proportion of the capital of Great Britain than what would otherwise have gone to it, has in all cases turned it, from a foreign trade of consumption with a neighbouring, into one with a more distant country; in many cases, from a direct foreign trade of consumption, into a round-about one; and in some cases, from all foreign trade of consumption, into a carrying trade. It has in all cases, therefore, turned it, from a direction in which it would have main-[429]tained a greater quantity of productive labour, into one, in which it can maintain a much smaller quantity. By suiting, besides, to one particular market only, so great a part of the industry and commerce of Great Britain, it has rendered the whole state of that industry and commerce more precarious and less secure, than if their produce had been accommodated to a greater variety of markets.[31]

47 We must carefully distinguish between the effects of the colony trade and those of the monopoly of that trade. The former are always and necessarily beneficial; the latter always and necessarily hurtful. But the former are so beneficial, that the colony trade, though subject to a monopoly, and notwithstanding the hurtful effects of that monopoly, is still upon the

[1–1] it *1*

[30] In 1779 the official value of English domestic exports was the lowest since 1747.

[31] See below, § 97, where Smith comments further on the disadvantages involved in artificially constraining the use of stock.

whole beneficial, and greatly beneficial; though a good deal less so than it otherwise would be.

48 The effect of the colony trade in its natural and free state, is to open a great, though distant market for such parts of the produce of British industry as may exceed the demand of the markets nearer home, of those of Europe, and of the countries which lie round the Mediterranean sea. In its natural and free state, the colony trade, without drawing from those markets any part of the produce which had ever been sent to them, encourages Great Britain to increase the surplus continually, by continually presenting new equivalents to be exchanged for it. In its natural and free state, the colony trade tends to increase the quantity of productive labour in Great [430] Britain, but without altering in any respect the direction of that which had been employed there before. In the natural and free state of the colony trade, the competition of all other nations would hinder the rate of profit from rising above the common level either in the new market, or in the new employment. The new market, without drawing any thing from the old one, would create, if one may say so, a new produce for its own supply; and that new produce would constitute a new capital for carrying on the new employment, which in the same manner would draw nothing from the old one.

49 The monopoly of the colony trade, on the contrary, by excluding the competition of other nations, and thereby raising the rate of profit both in the new market and in the new employment, draws produce from the old market and capital from the old employment.[32] To augment our share of the colony trade beyond what it otherwise would be, is the avowed purpose of the monopoly. If our share of that trade were to be no greater with, than it would have been without the monopoly, there could have been no reason for establishing the monopoly. But whatever forces into a branch of trade of which the returns are slower and more distant than those of the greater part of other trades, a greater proportion of the capital of any country, than what of its own accord would go to that branch, necessarily renders the whole quantity of productive labour annually maintained there, the whole annual produce of the land and labour of [431] that country, less than they otherwise would be. It keeps down the revenue of the inhabitants of that country, below what it would naturally rise to, and thereby diminishes their power of accumulation. It not only hinders, at all times, their capital from maintaining so great a quantity of productive labour as it would otherwise maintain, but it hinders it from increasing so fast as it would otherwise increase, and consequently from maintaining a still greater quantity of productive labour.

50 The natural good effects of the colony trade, however, more than counter-

[32] See above, I.ix.11.

balance to Great Britain the bad effects of the monopoly, so that, monopoly and all together, that trade, even as it is carried on at present, is not only advantageous, but greatly advantageous. The new market and [k]the[k] new employment which are opened by the colony trade, are of much greater extent than that portion of the old market and of the old employment which is lost by the monopoly. The new produce and the new capital which has been created, if one may say so, by the colony trade, maintain in Great Britain a greater quantity of productive labour, than what can have been thrown out of employment by the revulsion of capital from other trades of which the returns are more frequent. If the colony trade, however, even as it is carried on at present is advantageous to Great Britain, it is not by means of the monopoly, but in spite of the monopoly.

51 It is rather for the manufactured than for the rude produce of Europe, that the colony trade [432] opens a new market.[33] Agriculture is the proper business of all new colonies; a business which the cheapness of land renders more advantageous than any other. They abound, therefore, in the rude produce of land, and instead of importing it from other countries, they have generally a large surplus to export. In new colonies, agriculture either draws hands from all other employments, or keeps them from going to any other employment. There are few hands to spare for the necessary, and none for the ornamental manufactures. The greater part of the manufactures of both kinds, they find it cheaper to purchase of other countries than to make for themselves.[34] It is chiefly by encouraging the manufactures of Europe, that the colony trade indirectly encourages its agriculture. The manufacturers of Europe, to whom that trade gives employment, constitute a new market for the produce of the land; and the most advantageous of all markets; the home market for the corn and cattle, for the bread and butchers-meat of Europe; is thus greatly extended by means of the trade to America.

52 But that the monopoly of the trade of populous and thriving colonies is not alone sufficient to establish, or even to maintain manufactures in any country, the examples of Spain and Portugal sufficiently demonstrate. Spain and Portugal were manufacturing countries before they had any considerable colonies. Since they had the richest and most fertile in the world, they have both ceased to be so.

53 [433] In Spain and Portugal, the bad effects of the monopoly, aggravated by other causes, have[l], perhaps, nearly overbalanced[l] the natural good effects of the colony trade. These causes seem to be, other monopolies of different kinds; the degradation of the value of gold and silver below what it is in most other countries;[35] the exclusion from foreign markets by

[k–k] *3–6* [l–l] entirely conquered *1*

[33] See above, IV.vii.b.40. [34] See above, II.v.21 and IV.vii.b.44.
[35] See above, IV.i.13.

improper taxes upon exportation, and the narrowing of the home market, by still more improper taxes upon the transportation of goods from one part of the country to another; but above all, that irregular and partial administration of justice, which often protects the rich and powerful debtor from the pursuit of his injured creditor, and which makes the industrious part of the nation afraid to prepare goods for the consumption of those haughty and great men, to whom they dare not refuse to sell upon credit, and from whom they are altogether uncertain of repayment.[36]

54 In England, on the contrary, the natural good effects of the colony trade, assisted by other causes, have in a great measure conquered the bad effects of the monopoly. These causes seem to be, the general liberty of trade, which, notwithstanding some restraints, is at least equal, perhaps superior, to what it is in any other country; the liberty of exporting, duty free, almost all sorts of goods which are the produce of domestick industry, to almost any foreign country; and what, perhaps, is of still greater importance, the unbounded liberty of transporting them from any one part of our own country [434] to any other, without being obliged to give any account to any publick office, without being liable to question or examination of any kind;[37] but above all, that equal and impartial administration of justice which renders the rights of the meanest British subject respectable to the greatest, and which, by securing to every man the fruits of his own industry, gives the greatest and most effectual encouragement to every sort of industry.[38]

55 If the manufactures of Great Britain, however, have been advanced, as they certainly have, by the colony trade, it has not been by means of the monopoly of that trade, but in spite of the monopoly. The effect of the monopoly has been, not to augment the quantity, but to alter the quality and shape of a part of the manufactures of Great Britain, and to accommodate to a market, from which the returns are slow and distant, what would otherwise have been accommodated to one from which the returns are frequent and near. Its effect has consequently been to turn a part of the capital of Great Britain from an employment in which it would have maintained a greater quantity of manufacturing industry, to one in which it maintains a much smaller, and thereby to diminish, instead of increasing, the whole quantity of manufacturing industry maintained in Great Britain.

56 The monopoly of the colony trade, therefore, like all the other mean and malignant expedients of the mercantile system, depresses the industry [435] of all other countries, but chiefly that of the colonies, without in the

[36] See above, I.xi.n.1, where it is remarked that although the feudal government had been eliminated in Spain and Portugal, it had not been succeeded by a much better.

[37] The wool trade being the exception. It is stated at IV.viii.21 that the restrictions imposed upon it were 'very burdensome and oppressive'.

[38] Cf. II.iii.36 and IV.v.c.43, where Smith comments on the experience of England.

least increasing, but on the contrary diminishing, that of the country in whose favour it is established.

57 The monopoly hinders the capital of that country, whatever may at any particular time be the extent of that capital, from maintaining so great a quantity of productive labour as it would otherwise maintain, and from affording so great a revenue to the industrious inhabitants as it would otherwise afford. But as capital can be increased only by savings from revenue,[39] the monopoly, by hindering it from affording so great a revenue as it would otherwise afford, necessarily hinders it from increasing so fast as it would otherwise increase, and consequently from maintaining a still greater quantity of productive labour, and affording a still greater revenue to the industrious inhabitants of that country. One great original source of revenue, therefore, the wages of labour, the monopoly must necessarily have rendered at all times less abundant than it otherwise would have been.

58 By raising the rate of mercantile profit, the monopoly discourages the improvement of land.[40] The profit of improvement depends upon the difference between what the land actually produces, and what, by the application of a certain capital, it can be made to produce. If this difference affords a greater profit than what can be drawn from an equal capital in any mercantile employment, the improvement of land will draw capital from all mercantile employments. If [436] the profit is less, mercantile employments will draw capital from the improvement of land. Whatever therefore raises the rate of mercantile profit, either lessens the superiority or increases the inferiority of the profit of improvement; and in the one case hinders capital from going to improvement, and in the other draws capital from it. But by discouraging improvement, the monopoly necessarily retards the natural increase of another great original source of revenue, the rent of land. By raising the rate of profit too the monopoly necessarily keeps up the market rate of interest higher than it otherwise would be. But the price of land in proportion to the rent which it affords, the number of years purchase which is commonly paid for it, necessarily falls as the rate of interest rises, and rises as the rate of interest falls.[41] The monopoly, therefore, hurts the interest of the landlord two different ways, by retarding the natural increase, first, of his rent, and secondly, of the price which he would get for his land in proportion to the rent which it affords.

59 The monopoly indeed, raises the rate of mercantile profit, and thereby augments somewhat the gain of our merchants. But as it obstructs the natural increase of capital, it tends rather to diminish than to increase the sum total of the revenue which the inhabitants of the country derive

[39] Above, II.iii.15.

[40] Though Smith has already recognized the beneficial effects when mercantile profits are subsequently invested in land. See above, III.iv.3.

[41] Above, II.iv.17.

from the profits of stock; a small profit upon a great capital generally affording a greater revenue than a great profit upon a small one. The monopoly raises the rate of profit, but it [437] hinders the sum of profit from rising so high as it otherwise would do.

60 All the original sources of revenue, the wages of labour, the rent of land, and the profits of stock,[42] the monopoly renders much less abundant than they otherwise would be. To promote the little interest of one little order of men in one country, it hurts the interest of all other orders of men in that country, and of all [m] men in all other countries.

61 It is solely by raising the ordinary rate of profit that the monopoly either has proved or could prove advantageous to any one particular order of men. But besides all the bad effects to the country in general, which have already been mentioned as necessarily resulting from a high rate of profit; there is one more fatal, perhaps, than all these put together, but which, if we may judge from experience, is inseparably connected with it. The high rate of profit seems every where to destroy that parsimony which in other circumstances is natural to the character of the merchant.[43] When profits are high, that sober virtue seems to be superfluous, and expensive luxury to suit better the affluence of his situation. But the owners of the great mercantile capitals are necessarily the leaders and conductors of the whole industry of every nation, and their example has a much greater influence upon the manners of the whole industrious part of it than that of any other order of men. If his employer is attentive and parsimonious, the workman is very likely to be so too; but if the master is dissolute [438] and disorderly, the servant who shapes his work according to the pattern which his master prescribes to him, will shape his life too according to the example which he sets him. Accumulation is thus prevented in the hands of all those who are naturally the most disposed to accumulate; and the funds destined for the maintenance of productive labour receive no augmentation from the revenue of those who ought naturally to augment them the most. The capital of the country, instead of increasing, gradually dwindles away, and the quantity of productive labour maintained in it grows every day less and less. Have the exorbitant profits of the merchants of Cadiz and Lisbon augmented the capital of Spain and Portugal?[44] Have they alleviated the poverty, have they promoted the industry of those two beggarly countries?

[m] the *6*

[42] The original sources of revenue are discussed at I.vi.17.

[43] Smith comments on the relationship between manners and frugality at II.iii.12 and IV.vii.b.20. Sir James Steuart also noted that where high profits are sustained 'for a long time, they insensibly become *consolidated*, or, as it were, transformed into the intrinsic value of the goods', thus affecting the competitive position of the country or industry concerned. See especially, *Principles*, II.x.

[44] See below, § 82, where Smith also refers to the 'sumptuous profusion' of the merchants of Cadiz and Lisbon.

Such has been the tone of mercantile expence in those two trading cities, that those exorbitant profits, far from augmenting the general capital of the country, seem scarce to have been sufficient to keep up the capitals upon which they were made. Foreign capitals are every day intruding themselves, if I may say so, more and more into the trade of Cadiz and Lisbon. It is to expel those foreign capitals from a trade which their own [n] grows every day more and more insufficient for carrying on, that the Spaniards and Portugueze endeavour every day to straiten more and more the galling bands of their absurd monopoly. Compare the mercantile manners of Cadiz and Lisbon with those of Amsterdam, and you will be sen-[439]sible how differently the conduct and character of merchants are affected by the high and by the low profits of stock.[45] The merchants of London, indeed, have not yet generally become such magnificent lords as those of Cadiz and Lisbon; but neither are they in general such attentive and parsimonious burghers as those of Amsterdam. They are supposed, however, many of them, to be a good deal richer than the greater part of the former, and not quite so rich as many of the latter. But the rate of their profit is commonly much lower than that of the former, and a good deal higher than that of the latter. Light come light go, says the proverb; and the ordinary tone of expence seems every where to be regulated, not so much according to the real ability of spending, as to the supposed facility of getting money to spend.

62 It is thus that the single advantage which the monopoly procures to a single order of men is in many different ways hurtful to the general interest of the country.

63 To found a great empire for the sole purpose of raising up a people of customers, may at first sight appear a project fit only for a nation of shopkeepers.[46] It is, however, a project altogether unfit for a nation of shopkeepers; but extremely fit for a nation [o]whose government is influenced[o] by shopkeepers. Such [p]statesmen[p], and such [q]statesmen[q] only, are capable of fancying that they will find some advantage in employing the blood and treasure of their [r]fellow citizens[r], to found and [s]to[s] maintain such an empire. Say to a shopkeeper, Buy me a good estate, and I shall always buy my cloaths at your shop, even [440] though I should pay somewhat dearer than what I can have them for at other shops; and you will not find him

[n] capital *1* [o–o] that is governed *1* [p–p] soverigns *1* [q–q] soverigns *1*
[r–r] subjects *1* [s–s] *om. 4–6*

[45] Smith refers to the low rate of return on capital in Amsterdam at V.ii.k.80, and above, I.ix.10.

[46] See below, IV.viii.53. Pownall objected to the tone of this passage, since what he called 'creating and securing' an 'encreasing nation of appropriated customers' was the only idea which he could find 'precisely to define the relation which a commercial country bears to its colonies' (*Letter*, 44,n.). Cf. the same author's *Administration of the Colonies* (4th ed., London, 1768), vol. i, chapter viii.

very forward to embrace your proposal. But should any other person buy you such an estate, the shopkeeper would be much obliged to your benefactor if he would enjoin you to buy all your cloaths at his shop. England purchased for some of her subjects, who found themselves uneasy at home, a great estate in a distant country. The price, indeed, was very small, and instead of thirty years purchase, the ordinary price of land in the present times, it amounted to little more than the expence of the different equipments which made the first discovery, reconnoitred the coast, and took a fictitious possession of the country. The land was good and of great extent, and the cultivators having plenty of good ground to work upon, and being for some time at liberty to sell their produce where they pleased, became in the course of little more than thirty or forty years (between 1620 and 1660) so numerous and thriving a people, that the shopkeepers and other traders of England wished to secure to themselves the monopoly of their custom.[47] Without pretending, therefore, that they had paid any part, either of the original purchase-money, or of the subsequent expence of improvement, they petitioned the parliament that the cultivators of America might for the future be confined to their shop; first, for buying all the goods which they wanted from Europe; and, secondly, for selling all such parts of their own produce as those traders might find it convenient to buy. For [441] they did not find it convenient to buy every part of it. Some parts of it imported into England might have interfered with some of the trades which they themselves carried on at home. Those particular parts of it, therefore, they were willing that the colonists should sell where they could; the farther off the better; and upon that account proposed that their market should be confined to the countries south of Cape Finisterre. A clause in the famous act of navigation established this truly shopkeeper proposal into a law.[48]

64 The maintenance of this monopoly has hitherto been the principal, or more properly perhaps the sole end and purpose of the dominion which Great Britain assumes over her colonies. In the exclusive trade, it is supposed, consists the great advantage of provinces, which have never yet afforded either revenue or military force for the support of the civil government, or the defence of the mother country.[49] The monopoly is the

[47] See above, IV.vii.b.15–17, where Smith comments on the prosperity of the American colonies, and § 63 of the same section where it is stated that the interest taken in them by the mother country was consequent on their success.

[48] See above, IV.ii.24–31, where the main provisions of the act are reviewed.

[49] Smith also mentions the lack of an American contribution to the costs of defence in Letter 221 addressed to John Sinclair, dated 14 October 1782. Again, Pownall objected: 'I will beg leave to suggest to you some facts that induce me, and may perhaps you also, to be of a very different opinion. That very naval force, which by their armed vessels they are now so destructively exerting against our West-India trade and transports, they did very effectively in the two late wars, especially in the last, exert to the ruin of the West India commerce of France and Spain . . .' He added, with respect to the 'point of revenue'

principal badge of their dependency,[50] and it is the sole fruit which has hitherto been gathered from that dependency. Whatever expence Great Britain has hitherto laid out in maintaining this dependency, has really been laid out in order to support this monopoly. The expence of the ordinary peace establishment of the colonies amounted, before the commencement of the present disturbances, to the pay of twenty regiments of foot; to the expence of the artillery, stores, and extraordinary provisions with which it 'was' necessary to supply them; and to the ex-[442]pence of a very considerable naval force which "was" constantly kept up, in order to guard, from the smuggling vessels of other nations, the immense coast of North America, and that of our West Indian islands. The whole expence of this peace establishment was a charge upon the revenue of Great Britain, and was, at the same time, the smallest part of what the dominion of the colonies has cost the mother country.[51] If we would know the amount of the whole, we must add to the annual expence of this peace establishment the interest of the sums which, in consequence of her considering her colonies as provinces subject to her dominion, Great Britain has upon different occasions laid out upon their defence. We must add to it, in particular, the whole expence of the late war, and a great part of that 'of the war' which preceded it. The late war was altogether a colony quarrel, and the whole expence of it, in whatever part of the world it may have been laid out, whether in Germany or 'the East Indies, ought justly to be stated to the account of the colonies. It amounted to more than ninety millions sterling, including not only the new debt which was contracted, but the two shillings in the pound additional land tax, and the sums which were every year borrowed from the sinking fund.[52] The Spanish war

t–t is *1* u–u is *1* v–v *2–6* w in *1*

that 'before we went to decided war, a revenue might have been had upon compact, on terms which would have established the constitutional sovereignty of this country, regulating at the same time the trade and naval powers of the colonies, if those terms might have gone, at the same time, to securing the rights of those colonies as granted by the government of that mother country.' (*Letter*, 38.)

[50] Smith uses the term 'badge' in a similar context above, IV.vii.b.44.

[51] See above, IV.vii.b.20 and IV.vii.c.12.

[52] Smith reviews the costs of wars at II.iii.35, IV.i.26, IV.viii.53, and V.iii.92; cf. V.iii.88, where he states that the colonies should contribute to costs incurred on their behalf, and IV.vii.c.13, where he remarks that only the Spanish and Portuguese colonies had so contributed. In commenting on the saving of costs to be expected from the emancipation of America, Smith pointed out that Britain's two most expensive wars, the Spanish War of 1739 and the French War of 1755, 'were undertaken, the one chiefly, the other altogether on account of the colonies'. He went on to point out that the British had at one time complained of involvement in the affairs of Hanover 'with which we should, otherwise, have had nothing to do. But we, surely, have had much more reason to complain, upon the same account, of our connexion with America.' ('Thoughts on America', § 12, AHR 717–18.) Smith restated this theme in Letter 221 addressed to Sir John Sinclair, dated 14 October 1782: 'The real futility of all distant dominions, of which the defence is necessarily most expensive, and which contribute nothing, either by revenue or military

which began in 1739, was principally a colony quarrel. Its principal object was to prevent the search of the colony ships which carried on a contraband trade with the Spanish main. This whole expence is, in reality, a bounty which has [443] been given in order to support a monopoly. The pretended purpose of it was to encourage the manufactures, and to increase the commerce of Great Britain. But its real effect has been to raise the rate of mercantile profit, and to enable our merchants to turn into a branch of trade, of which the returns are more slow and distant than those of the greater part of other trades, a greater proportion of their capital than they otherwise would have done; two events which, if a bounty could have prevented, it might perhaps have been very well worth while to give such a bounty.

65 Under the present system of management, therefore, Great Britain derives nothing but loss from the dominion which she assumes over her colonies.

66 To propose that Great Britain should voluntarily give up all authority over her colonies,[53] and leave them to elect their own magistrates, to enact their own laws, and to make peace and war as they might think proper, would be to propose such a measure as never was, and never will be adopted, by any nation in the world.[54] No nation ever voluntarily gave up the dominion of any province, how troublesome soever it might be to govern it, and how small soever the revenue which it afforded might be in propor-

force, to the general defence of the empire, and very little even to their own particular defence, is, I think, the subject upon which the public prejudices of Europe require most to be set right.' Sinclair had apparently commented to Smith on the bleak prospect of the American War, and that 'if we go on at this rate, the nation must be ruined', to which Smith replied: 'Be assured, my young friend, that there is a great deal of *ruin* in a nation.' Sinclair Corr., i.390-1. In Letter 158 addressed to Strahan, dated 3 June 1776, Smith wrote: 'The American Campaign had begun awkwardly. I hope, I cannot say that I expect, it will end better. England, tho' in the present times it breeds men of great professional abilities in all different ways, great Lawyers, great watch makers & Clockmakers, &c. &c., seems to breed neither Statesmen nor Generals.'

[53] Pownall objected to Smith's conclusion that the colonies should be given up at least in so far as that conclusion was based on the general analysis of the natural progress of opulence and Smith's views as to the different employments of capital: 'If we lose our colonies, we must submit to our fate; but the idea of parting with them on the ground of system, is much like the system which an ironical proverb recommends, *'of dying to save charges.'* (37.) Pownall's criticism of Smith's views on the colony trade appears mainly at pp. 37–48 of the *Letter.*

[54] Cf. Smith's 'Thoughts on America', with regard to the emancipation of the colonies: 'tho' this termination of the war might be really advantageous, it would not, in the eyes of Europe appear honourable to Great Britain; and when her empire was so much curtailled, her power and dignity would be supposed to be proportionably diminished. What is of still greater importance, it could scarce fail to discredit the Government in the eyes of our own people, who would probably impute to mal-administration what might, perhaps, be no more than the unavoidable effect of the natural and necessary course of things. (it) . . . would have everything to fear from their rage and indignation at the public disgrace and calamity, for such they would suppose it to be, of thus dismembering the empire.' (§ 13, AHR 718.)

tion to the expence which it occasioned.[55] Such sacrifices, though they might frequently be agreeable to the interest, are always mortifying to the pride of every nation, and what is perhaps of still greater consequence, they are always contrary to the private interest of the governing part of it, [444] who would thereby be deprived of the disposal of many places of trust and profit, of many opportunities of acquiring wealth and distinction, which the possession of the most turbulent, and, to the great body of the people, the most unprofitable province seldom fails to afford. The most visionary enthusiast would scarce be capable of proposing such a measure, with any serious hopes at least of its ever being adopted. If it was adopted, however, Great Britain would not only be immediately freed from the whole annual expence of the peace establishment of the colonies, but might settle with them such a treaty of commerce as would effectually secure to her a free trade, more advantageous to the great body of the people, though less so to the merchants, than the monopoly which she at present enjoys. By thus parting good friends, the natural affection of the colonies to the mother country, which, perhaps, our late dissentions have well nigh extinguished, would quickly revive.[56] It might dispose them not only to respect, for whole centuries together, that treaty of commerce which they had concluded with us at parting, but to favour us in war as well as in trade, and, instead of turbulent and factious subjects, to become our most faithful, affectionate, and generous allies; and the same sort of parental affection on the one side, and filial respect on the other, might revive between Great Britain and her colonies, which used to subsist between those of ancient Greece and the mother city from which they descended.[57]

67 [445] In order to render any province advantageous to the empire to which it belongs, it ought to afford, in time of peace, a revenue to the publick sufficient not only for defraying the whole expence of its own peace establishment, but for contributing its proportion to the support of the general government of the empire. Every province necessarily contributes, more

[55] But cf. V.iii.92, where Smith recommends that Britain should give up her imperial pretensions.

[56] Cf. 'Thoughts on America': 'tho' Canada, Nova Scotia, and the Floridas were all given up to our rebellious colonies, or were all conquered by them, yet the similarity of language and manners would in most cases dispose the Americans to prefer our alliance to that of any other nation. Their antient affection for the people of this country might revive, if they were once assured that we meant to claim no dominion over them, . . . By a federal union with America we should certainly incur much less expense, and might, at the same time, gain as real advantages, as any we have hitherto derived from all the nominal dominion we have ever exercised over them.' (§ 12, AHR 718.) It is worth observing that Smith meant by 'federal union' a set of links established by the executive (federal) power rather than the type of constitutional arrangement which was later to be adopted in America. His own preference was for an incorporating union, of the kind which was later applied to Ireland, which would have given the colonies representation at Westminster. See below, § 77–9, and V.iii.90.

[57] See above, IV.vii.a.2.

or less, to increase the expence of that general government. If any particular province, therefore, does not contribute its share towards defraying this expence, an unequal burden must be thrown upon some other part of the empire. The extraordinary revenue too which every province affords to the publick in time of war, ought, from parity of reason, to bear the same proportion to the extraordinary revenue of the whole empire which its ordinary revenue does in time of peace.[58] That neither the ordinary nor extraordinary revenue which Great Britain derives from her colonies, bears this proportion to the whole revenue of the British empire, will readily be allowed. The monopoly, it has been supposed, indeed, by increasing the private revenue of the people of Great Britain, and thereby enabling them to pay greater taxes, compensates the deficiency of the publick revenue of the colonies. But this monopoly, I have endeavoured to show, though a very grievous tax upon the colonies, and though it may increase the revenue of a particular order of men in Great Britain, diminishes instead of increasing that of the great body of the people; and consequently diminishes instead of [446] increasing the ability of the great body of the people to pay taxes.[59] The men too whose revenue the monopoly increases, constitute a particular order, which it is both absolutely impossible to tax beyond the proportion of other orders, and extremely impolitick even to attempt to tax beyond that proportion, as I shall endeavour to shew in the following book.[60] No particular resource, therefore, can be drawn from this particular order.

[58] '. . . there never was an Idea of exempting the Colonies: on the Contrary, Restraints upon their Trade, and Taxes on their Consumption, have always gone together: And together compose the System, by which they have been constantly and happily governed . . . Our Taxes have been since encreased many-fold: Their Abilities have been enlarged still faster . . . The Proportion between the publick Burthens on the Mother-country and the Colonies, as divided when they were in their Infancy, is entirely lost: And to restore that Proportion, and again to make something like a Partition of those Burthens, is no more than maintaining the System, upon which we have always acted, and to which I own I am partial, because the Colonies have flourished under it beyond all Example in History, and I cannot prefer visionary speculations and novel Doctrines to such an Experience.' (*Considerations on the Trade and Finance of the Kingdom* (London, 1766), 81, attributed to T. Whateley and often ascribed to George Grenville.)

[59] above, § 59.

[60] Below, V.ii.f.6. Cf. Hume, *History of England* (1754), i.243–4: 'To complain of the parliament's employing the power of taxation, as the means of extorting concessions from their sovereign, were to expect, that they would intirely disarm themselves, and renounce the sole expedient, provided by the constitution, for ensuring to the kingdom a just and legal administration. In all periods of English story, there occur instances of their remonstrating with their princes in the freest manner, and of their refusing supply, when disgusted with any circumstance of public conduct. 'Tis, however, certain, that this power, tho' essential to parliaments, may easily be abused, as well by the frequency and the minuteness of their remonstrances, as by their intrusion into every part of the king's councils and determinations. Under color of advice, they may give disguised orders; and in complaining of grievances, they may draw to themselves every power of government. Whatever measure is embraced, without consulting them, may be pronounced an oppression of the people; and till corrected, they may refuse the most necessary supplies to their indigent sovereign.'

68 The colonies may be taxed by their own assemblies, or by the parliament of Great Britain.

69 That the colony assemblies can ever be so managed as to levy upon their constituents a publick revenue sufficient, not only to maintain at all times their own civil and military establishment, but to pay their proper proportion of the expence of the general government of the British empire, seems not very probable. It was a long time before even the parliament of England, though placed immediately under the eye of the sovereign, could be brought under such a system of management, or could be rendered sufficiently liberal in their grants for supporting the civil and military establishments even of their own country. It was only by distributing among the particular members of parliament, a great part either of the offices, or of the disposal of the offices arising from this civil and military establishment, that such a system of management could be established even with regard to the parliament of England. But the distance of the colony assemblies from the eye of the sovereign, [447] their number, their dispersed situation, and their various constitutions, would render it very difficult to manage them in the same manner, even though the sovereign had the same means of doing it; and those means are wanting. It would be absolutely impossible to distribute among all the leading members of all the colony assemblies such a share, either of the offices or of the disposal of the offices arising from the general government of the British empire, as to dispose them to give up their popularity at home and to tax their constituents for the support of that general government, of which almost the whole emoluments were to be divided among people who were strangers to them. The unavoidable ignorance of administration, besides, concerning the relative importance of the different members of those different assemblies, the offences which must frequently be given, the blunders which must constantly be committed in attempting to manage them in this manner, seems to render such a system of management altogether impracticable with regard to them.

70 The colony assemblies, besides, cannot be supposed the proper judges of what is necessary for the defence and support of the whole empire. The care of that defence and support is not entrusted to them. It is not their business, and they have no regular means of information concerning it. The assembly of a province, like the vestry of a parish, may judge very properly concerning the affairs of its own particular district; but can have no proper means of judging [448] concerning those of the whole empire. It cannot even judge properly concerning the proportion which its own province bears to the whole empire; or concerning the relative degree of its wealth and importance, compared with the other provinces; because those other provinces are not under the inspection and superintendency of the assembly of a particular province. What is necessary for the defence and

support of the whole empire, and in what proportion each part ought to contribute, can be judged of only by that assembly which inspects and superintends the affairs of the whole empire.

71 It has been proposed, accordingly, that the colonies should be taxed by requisition, the parliament of Great Britain determining the sum which each colony ought to pay, and the provincial assembly assessing and levying it in the way that suited best the circumstances of the province. What concerned the whole empire would in this way be determined by the assembly which inspects and superintends the affairs of the whole empire; and the provincial affairs of each colony might still be regulated by its own assembly. Though the colonies should in this case have no representatives in the British parliament, yet, if we may judge by experience, there is no probability that the parliamentary requisition would be unreasonable. The parliament of England has not upon any occasion shown the smallest disposition to overburden those parts of the empire which are not represented in parliament. The islands of Guernsey and Jersey, [449] without any means of resisting the authority of parliament, are more lightly taxed than any part of Great Britain. Parliament in attempting to exercise its supposed right, whether well or ill grounded, of taxing the colonies, has never hitherto demanded of them any thing which even approached to a just proportion to what was paid by their fellow subjects at home. If the contribution of the colonies, besides, was to rise or fall in proportion to the rise or fall of the land tax, parliament could not tax them without taxing at the same time its own constituents, and the colonies might in this case be considered as virtually represented in parliament.

72 Examples are not wanting of empires in which all the different provinces are not taxed, if I may be allowed the expression, in one mass; but in which the sovereign regulates the sum which each province ought to pay, and in some provinces assesses and levies it as he thinks proper; while in others, he leaves it to be assessed and levied as the respective states of each province shall determine. In some provinces of France, the king not only imposes what taxes he thinks proper, but assesses and levies them in the way he thinks proper.[61] From others he demands a certain sum, but leaves it to the states of each province to assess and levy that sum as they think proper. According to the scheme of taxing by requisition, the parliament of Great Britain would stand nearly in the same situation towards the colony assemblies, as the king of France does towards the states of those provinces [450] which still enjoy the privilege of having states of their own, the provinces of France which are supposed to be the best governed.

73 But though, according to this scheme, the colonies could have no just reason to fear that their share of the publick burdens should ever exceed

[61] See below, V.ii.k.70.

the proper proportion to that of their fellow-citizens at home; Great Britain might have just reason to fear that it never would amount to that proper proportion. The parliament of Great Britain has not for some time past had the same established authority in the colonies, which the French king has in those provinces of France which still enjoy the privilege of having states of their own. The colony assemblies, if they were not very favourably disposed (and unless more skilfully managed than they ever have been hitherto, they are not very likely to be so) might still find many pretences for evading or rejecting the most reasonable requisitions of parliament. A French war breaks out, we shall suppose; ten millions must immediately be raised, in order to defend the seat of the empire. This sum must be borrowed upon the credit of some parliamentary fund mortgaged for paying the interest. Part of this fund parliament proposes to raise by a tax to be levied in Great Britain, and part of it by a requisition to all the different colony assemblies of America and the West Indies. Would people readily advance their money upon the credit of a fund, which partly depended upon the good-humour of all those assemblies, far distant from the seat of the [451] war, and sometimes, perhaps, thinking themselves not much concerned in the event of it? Upon such a fund no more money would probably be advanced than what the tax to be levied in Great Britain might be supposed to answer for. The whole burden of the debt contracted on account of the war would in this manner fall, as it always has done hitherto, upon Great Britain; upon a part of the empire, and not upon the whole empire. Great Britain is, perhaps, since the world began, the only state which, as it has extended its empire, has only increased its expence without once augmenting its resources. Other states have generally disburdened themselves upon their subject and subordinate provinces of the most considerable part of the expence of defending the empire. Great Britain has hitherto suffered her subject and subordinate provinces to disburden themselves upon her of almost this whole expence. In order to put Great Britain upon a footing of equality with her own colonies, which the law has hitherto supposed to be subject and subordinate, it seems necessary, upon the scheme of taxing them by parliamentary requisition, that parliament should have some means of rendering its requisitions immediately effectual, in case the colony assemblies[62] should attempt to evade or reject them; and what those means are, it is not very easy to conceive, and it has not yet been explained.

74 Should the parliament of Great Britain, at the same time, be ever fully established in the right of taxing the colonies, even independent of [452] the consent of their own assemblies, the importance of those assemblies would from that moment be at an end, and with it, that of all the leading

[62] These assemblies are described above, IV.vii.b.51.

men of British America. Men desire to have some share in the management of publick affairs chiefly on account of the importance which it gives them.[63] Upon the power which the greater part of the leading men, the natural aristocracy of every country, have of preserving or defending their respective importance, depends the stability and duration of every system of free government.[64] In the attacks which those leading men are continually making upon the importance of one another, and in the defence of their own, consists the whole play of domestick faction and ambition. The leading men of America, like those of all other countries, desire to preserve their own importance. They feel, or imagine, that if their assemblies, which they are fond of calling parliaments, and of considering as equal in authority to the parliament of Great Britain, should be so far degraded as to become the humble ministers and executive officers of that parliament, the greater part of their own importance would be at an end. They have rejected, therefore, the proposal of being taxed by parliamentary requisition, and like other ambitious and high-spirited men, have rather chosen to draw the sword in defence of their own importance.

75 Towards the declension of the Roman republick, the allies of Rome, who had borne the principal burden of defending the state and ex-[453]tending the empire, demanded to be admitted to all the privileges of Roman citizens. Upon being refused, the social war broke out. During the course of that war Rome granted those privileges to the greater part of them, one by one, and in proportion as they detached themselves from the general confederacy. The parliament of Great Britain insists upon taxing the colonies; and they refuse to be taxed by a parliament in which they are not represented. If to each colony, which should detach itself from the general confederacy, Great Britain should allow such a number of representatives as suited the proportion of what it contributed to the publick revenue of the empire,[65] in consequence of its being subjected to the same taxes, and in compensation admitted to the same freedom of trade with its fellow-subjects at home; the number of its representatives to be augmented as the proportion of its contribution might afterwards augment; a new method of acquiring importance, a new and more dazzling object of ambition would be presented to the leading men of each colony.[66] Instead of piddling for

[63] Smith provides another example below, V.ii.k.80, drawn from Holland.

[64] Cf. 'Thoughts on America': 'The principal security of every government arises always from the support of those whose dignity, authority and interest, depend upon its being supported.' (§ 10, AHR 716.)

[65] See below, V.iii.68, where it is stated that taxation with representation conforms to British constitutional practice.

[66] Smith points out in the 'Thoughts on America' that a form of union with America *might* be possible: 'The leading men of America, we may believe, wish to continue to be the principal people in their own country. After a union with Great Britain, they might expect to continue to be so; in the same manner as the leading men of Scotland continued to be the principal people in their own country after the union with England.' (§ 14, AHR

the little prizes which are to found in what may be called the paltry raffle of colony faction; they might then hope, from the presumption which men naturally have in their own ability and good fortune, to draw some of the great prizes which sometimes come from the wheel of the great state lottery of British politicks. Unless this or some other method is fallen upon, and there seems to be none more obvious than this, of [454] preserving the importance and of gratifying the ambition of the leading men of America, it is not very probable that they will ever voluntarily submit to us; and we ought to consider that the blood which must be shed in forcing them to do so, is, every drop of it, the blood either of those who are, or of those whom we wish to have for our fellow-citizens. They are very weak who flatter themselves that, in the state to which things have come, our colonies will be easily conquered by force alone.[67] The persons who now govern the resolutions of that they call their continental congress, feel in themselves at this moment a degree of importance which, perhaps, the greatest subjects in Europe scarce feel. From shopkeepers, tradesmen, and attornies, they are become statesmen and legislators, and are employed in contriving a new form of government for an extensive empire, which, they flatter themselves, will become, and which, indeed, seems very likely to become, one of the greatest and most formidable that ever was in the world. Five hundred different people, perhaps, who in different ways act immediately under the continental congress; and five hundred thousand, perhaps, who act under those five hundred, all feel in the same manner a proportionable rise in their own importance. Almost every individual of the governing party in America, fills, at present in his own fancy, a station superior, not only to what he had ever filled before, but to what he had ever expected to fill; and unless some new object of ambition is presented either to him or [455] to his leaders, if he has the ordinary spirit of a man, he will die in defence of that station.

76 It is a remark of the president Henaut that we now read with pleasure the account of many little transactions of the Ligue, which when they happened were not perhaps considered as very important pieces of news.

719.) It was in this context that Smith made the ingenious suggestion that should the idea of union fail, the solution might be 'An apparent restoration of the old system, so contrived as to lead necessarily, but insensibly to the total dismemberment of America, might, perhaps, satisfy both the people of Great Britain and the leading men of America: the former mistaking, and the latter understanding, the meaning of the scheme.' (§ 16.) Another ingenious suggestion, to be applied in the case of the complete emancipation of America, was that we should restore Canada to France and the Floridas to Spain, thereby rendering our own colonies the 'natural enemies of those two monarchies and consequently the natural allies of Great Britain'. In this way, Smith hoped that 'old enmities, and probably old friendships' might be revived. § 12, AHR 718.

[67] It is pointed out in V.i.a.27 that while militias are generally inferior to standing armies, this need not be the case where the former are long in the field, and that another campaign would place the American militia on a par with the British army.

But every man then, says he, fancied himself of some importance; and the innumerable memoirs which have come down to us from those times, were, the greater part of them, written by people who took pleasure in recording and magnifying events in which, they flattered themselves, they had been considerable actors.[68] How obstinately the city of Paris upon that occasion defended itself, what a dreadful famine it supported rather than submit to the best and afterwards the most beloved of all the French kings, is well known. The greater part of the citizens, or those who governed the greater part of them, fought in defence of their own importance, which they foresaw was to be at an end whenever the ancient government should be re-established. Our colonies, unless they can be induced to consent to a union, are very likely to defend themselves against the best of all mother countries, as obstinately as the city of Paris did against one of the best of kings.

77 The idea of representation was unknown in ancient times. When the people of one state were admitted to the right of citizenship in another, they had no other means of exercising that right but by coming in a body to vote and deli-[456]berate with the people of that other state. The admission of the greater part of the inhabitants of Italy to the privileges of Roman citizens, completely ruined the Roman republick. It was no longer possible to distinguish between who was and who was not a Roman citizen. No tribe could know its own members. A rabble of any kind could be introduced into the assemblies of the people, could drive out the real citizens, and decide upon the affairs of the republick as if they themselves had been such.[69] But though America ˣwasˣ to send fifty of sixty new representatives to parliament, the door-keeper of the house of commons could not find any great difficulty in distinguishing between who was and who was not a member. Though the Roman constitution, therefore, was necessarily ruined by the union of Rome with the allied states of Italy, there is not the least probability that the British constitution would be hurt by the union of Great Britain with her colonies. That constitution, on the contrary, would be completed by it, and seems to be imperfect without it.[70] The assembly which deliberates and decides concerning the affairs of every part of the empire, in order to be properly informed, ought certainly to have representatives from every part of it. That this union, however, could be easily effectuated, or that difficulties and great difficulties might not occur in the execution, I do not pretend. I have yet heard of none, however, which

ˣ–ˣ were *4–6*

[68] C. J. F. Hénault, *Nouvel Abrégé chronologique de l'histoire de France* (Paris, 1768), 581.

[69] Cf. Montesquieu, *Considerations*, 93: 'Once the peoples of Italy became its citizens, each city brought to Rome its genius, its particular interests, and its dependence on some great protector. The distracted city no longer formed a complete whole.'

[70] See below, V.iii.89–90, where Smith elaborates on the economic and political benefits of union with regard to the colonies and Ireland.

appear insurmountable.[71] The principal perhaps arise, not from the nature of things, but from the prejudices and opinions [457] of the people both on this and ʸonʸ the other side of the Atlantic.

78 We, on this side the water, are afraid lest the multitude of American representatives should overturn the balance of the constitution, and increase too much either the influence of the crown on the one hand, or the force of the democracy on the other. But if the number of American representatives ᶻwasᶻ to be in proportion to the produce of American taxation, the number of people to be managed would increase exactly in proportion to the means of managing them; and the means of managing, to the number of people to be managed. The monarchical and democratical parts of the constitution would, after the union, stand exactly in the same degree of relative force with regard to one another as they had done before.

79 The people on the other side of the water are afraid lest their distance from the seat of government might expose them to many oppressions. But their representatives in parliament, of which the number ought from the first to be considerable, would easily be able to protect them from all oppression. The distance could not much weaken the dependency of the representative upon the constituent, and the former would still feel that he owed his seat in parliament, and all the consequence which he derived from it, to the good-will of the latter. It would be the interest of the former, therefore, to cultivate that good-will by complaining, with all the authority of a member of the legislature, of every outrage which [458] any civil or military officer might be guilty of in those remote parts of the empire. The distance of America from the seat of government, besides, the ᵃnativesᵃ of that country might flatter themselves, with some appearance of reason too, would not be of very long continuance. Such has hitherto been the rapid progress of that country in wealth, population and improvement, that in the course of little more than a century, perhaps, the produce of American might exceed that of British taxation. The seat of the empire

ʸ–ʸ *2–6* ᶻ–ᶻ were *4–6* ᵃ–ᵃ nations *1*

[71] Smith considered the desirability of an incorporating union in the 'Thoughts on America', somewhat along the lines of the existing union between Scotland and England, but added that such a plan as would 'certainly tend most to the prosperity, to the splendor, and to the duration of the empire, if you except here and there a solitary philosopher like myself, seems scarce to have a single advocate' (§ 11, AHR 717). He added that the Americans especially, in their 'present elevation of spirits' were unlikely to agree, and as to British opinion, he believed the most popular solution was military victory.

Both Lord Kames and Benjamin Franklin supported the idea of consolidating union, although the latter clearly recognized that delay would make it increasingly unlikely as a solution. None the less, it is interesting to recall that the First Continental Congress of 1774 debated and narrowly defeated Joseph Galloway's plan for a 'grand legislative council' which was to be responsible for controlling the affairs of the union. The whole issue of union is extensively discussed in Richard Koebner's *Empire* (Cambridge, 1961), chapter 4.

would then naturally remove itself to that part of the empire which contributed most to the general defence and support of the whole.[72]

80 The discovery of America, and that of a passage to the East Indies by the Cape of Good Hope, are the two greatest and most important events recorded in the history of mankind.[73] Their consequences have already been very great: but, in the short period of between two and three centuries which has elapsed since these discoveries were made, it is impossible that the whole extent of their consequences can have been seen. What benefits, or what misfortunes to mankind may hereafter result from those great events no human wisdom can foresee. By uniting, in some measure, the most distant parts of the world, by enabling them to relieve one another's wants, to increase one another's enjoyments, and to encourage one another's industry, their general tendency would seem to be beneficial. To the natives, however, both of the East and West Indies, all the commercial bene-[459] fits which can have resulted from those events have been sunk and lost in the dreadful misfortunes which they have occasioned.[74] These misfortunes, however, seem to have arisen rather from accident than from any thing in the nature of those events themselves. At the particular time when these discoveries were made, the superiority of force happened to be so great on the side of the Europeans, that they were enabled to commit with impunity every sort of injustice in those remote countries. Hereafter, perhaps, the natives of those countries may grow stronger, or those of Europe may grow weaker, and the inhabitants of all the different quarters of the world may arrive at that equality of courage and force which, by inspiring mutual fear, can alone overawe the injustice of independent nations into some sort of respect for the rights of one another.[75] But

[72] It is interesting to note that in reading this section of Smith's work, Hugh Blair expressed regret that he had given the colonial affair 'a representation &c. which I wish had been omitted, because it is too much like a publication for the present moment. In Subsequent editions when public Measures come to be Settled, these pages will fall to be omitted or Altered.' (Letter 151 addressed to Smith, dated 3 April 1776.) In Letter 147 addressed to Smith, dated 1 November 1775, John Roebuck, friend and former partner of James Watt, stated the opposite opinion: 'I hoped by this time to have seen your Name in the Papers. The meeting of Parlt. is the proper time for the Publication of such a work as yours. It might also have been of general use in influencing the Opinion of many in this American contest.' In Letter 153 addressed to Smith, dated 8 April 1776, William Robertson commented that: 'Many of your observations concerning the Colonies are of capital importance to me. I shall often follow you as my Guide and Instructor. I am happy to find my own ideas concerning the absurdity of the limitations upon the Colony trade established much better than I could have done myself.'

[73] See above, IV.i.33. Cf. G. T. F. Raynal, *Histoire philosophique*, i.1, trans. J. Justamond, i.1: 'The discovery of the new world, and the passages to the East Indies by the Cape of Good Hope, is one of the most important events in the history of the human species.'

[74] Smith comments on the savage injustice inflicted by the Europeans on the native populations in IV.i.32 and below, § 100.

[75] It is pointed out in LJ (B) 339, ed. Cannan 265, in the course of considering the laws of nations, that: 'where there is no supreme legislative power nor judge to settle differences, we may always expect uncertainty and irregularity.'

nothing seems more likely to establish this equality of force than that mutual communication of knowledge and of all sorts of improvements which an extensive commerce from all countries to all countries naturally, or rather necessarily, carries along with it.

81 In the mean time one of the principal effects of those discoveries has been to raise the mercantile system to a degree of splendor and glory which it could never otherwise have attained to. It is the object of that system to enrich a great nation rather by trade and manufactures than by the improvement and cultivation of land, rather by the industry of the towns than by that of the country. But, in consequence of those dis-[460] coveries, the commercial towns of Europe, instead of being the manufacturers and carriers for but a very small part of the world (that part of Europe which is washed by the Atlantic ocean, and the countries which lie round the Baltick and Mediterranean seas), have now become the manufacturers for the numerous and thriving cultivators of America, and the carriers, and in some respects the manufacturers too, for almost all the different nations of Asia, Africa, and America. Two new worlds have been opened to their industry, each of them much greater and more extensive than the old one, and the market of one of them growing still greater and greater every day.

82 The countries which possess the colonies of America, and which trade directly to the East Indies, enjoy, indeed, the whole shew and splendor of this great commerce. Other countries, however, notwithstanding all the invidious restraints by which it is meant to exclude them, frequently enjoy a greater share of the real benefit of it.[76] The colonies of Spain and Portugal, for example, give more real encouragement to the industry of other countries than to that of Spain and Portugal. In the single article of linen alone the consumption of those colonies amounts, it is said, but I do not pretend to warrant the quantity, to more than three millions sterling a year. But this great consumption is almost entirely supplied by France, Flanders, Holland, and Germany. Spain and Portugal furnish but a small part of it. The capital [461] which supplies the colonies with this great quantity of linen is annually distributed among, and furnishes a revenue to the inhabitants of those other countries. The profits of it only are spent in Spain and Portugal, where they help to support the sumptuous profusion of the merchants of Cadiz and Lisbon.[77]

83 Even the regulations by which each nation endeavours to secure to itself the exclusive trade of its own colonies, are frequently more hurtful to the countries in favour of which they are established than to those against which they are established. The unjust oppression of the industry of other countries falls back, if I may say so, upon the heads of the oppres-

[76] See above, IV.vii.c.6.

[77] See above, § 61.

sors, and crushes their industry more than it does that of those other countries. By those regulations, for example, the merchant of Hamburgh must send the linen which he destines for the American market to London, and he must bring back from thence the tobacco which he destines for the German market; because he can neither send the one directly to America, nor bring back the other directly from thence. By this restraint he is probably obliged to sell the one somewhat cheaper, and to buy the other somewhat dearer than he otherwise might have done; and his profits are probably somewhat abridged by means of it. In this trade, however, between Hamburgh and London, he certainly receives the returns of his capital much more quickly than he could possibly have done in the direct trade to America, even though we should suppose, [462] what is by no means the case, that the payments of America were as punctual as those of London. In the trade, therefore, to which those regulations confine the merchant of Hamburgh, his capital can keep in constant employment a much greater quantity of German industry than it possibly could have done in the trade from which he is excluded. Though the one employment, therefore, may to him perhaps be less profitable than the other, it cannot be less advantageous to his country. It is quite otherwise with the employment into which the monopoly naturally attracts, if I may say so, the capital of the London merchant. That employment may, perhaps, be more profitable to him than the greater part of other employments, but, on account of the slowness of the returns, it cannot be more advantageous to his country.

84 After all the unjust attempts, therefore, of every country in Europe to engross to itself the whole advantage of the trade of its own colonies, no country has yet been able to engross to itself any thing but the expence of supporting in time of peace and of defending in time of war the oppressive authority which it assumes over them. The inconveniencies resulting from the possession of its colonies, every country has engrossed to itself completely. The advantages resulting from their trade it has been obliged to share with many other countries.

85 At first sight, no doubt, the monopoly of the great commerce of America, naturally seems to be an acquisition of the highest value. To the [463] undiscerning eye of giddy ambition, it naturally presents itself amidst the confused scramble of politicks and war, as a very dazzling object to fight for. The dazzling splendor of the object, however, the immense greatness of the commerce, is the very quality which renders the monopoly of it hurtful, or which makes one employment, in its own nature necessarily less advantageous to the country than the greater part of other employments, absorb a much greater proportion of the capital of the country than what would otherwise have gone to it.

86 The mercantile stock of every country, it has been shewn in the second

book,[78] naturally seeks, if one may say so, the employment most advantageous to that country. If it is employed in the carrying trade, the country to which it belongs becomes the emporium of the goods of all the countries whose trade that stock carries on. But the owner of that stock necessarily wishes to dispose of as great a part of those goods as he can at home. He thereby saves himself the trouble, risk, and expence, of exportation, and he will upon that account be glad to sell them at home, not only for a much smaller price, but with somewhat a smaller profit than he might expect to make by sending them abroad. He naturally, therefore, endeavours as much as he can to turn his carrying trade into a foreign trade of consumption. If his stock again is employed in a foreign trade of consumption, he will, for the same reason, be glad to dispose of at home as great a part as he can of the home goods, which [464] he collects in order to export to some foreign market, and he will thus endeavour, as much as he can, to turn his foreign trade of consumption into a home trade. The mercantile stock of every country naturally courts in this manner the near, and shuns the distant employment; naturally courts the employment in which the returns are frequent, and shuns that in which they are distant and slow; naturally courts the employment in which it can maintain the greatest quantity of productive labour in the country to which it belongs, or in which its owner resides, and shuns that in which it can maintain there the smallest quantity. It naturally courts the employment which in ordinary cases is most advantageous, and shuns that which in ordinary cases is least advantageous to that country.

87 But if any of those distant employments, which in ordinary cases are less advantageous to the country, the profit should happen to rise somewhat higher than what is sufficient to balance the natural preference which is given to nearer employments, this superiority of profit will draw stock from those nearer employments, till the profits of all return to their proper level. This superiority of profit, however, is a proof that in the actual circumstances of the society, those distant employments are somewhat understocked in proportion to other employments, and that the stock of the society is not distributed in the properest manner among all the different employments carried on in it. It is a proof that something is either bought cheaper or sold dearer [465] than it ought to be, and that some particular class of citizens is more or less oppressed either by paying more or by getting less than what is suitable to that equality, which ought to take place, and which naturally does take place among all the different classes of them. Though the same capital never will maintain the same quantity of productive labour in a distant as in a near employment, yet a distant employment may be as necessary for the welfare of the society as a

[78] See above, II.v.

near one;[79] the goods which the distant employment deals in being necessary, perhaps, for carrying on many of the nearer employments. But if the profits of those who deal in such goods are above their proper level, those goods will be sold dearer than they ought to be, or somewhat above their natural price, and all those engaged in the nearer employments will be more or less oppressed by this high price. Their interest, therefore, in this case requires that some stock should be withdrawn from those nearer employments, and turned towards that distant [b]one[b], in order to reduce its profits to their proper level, and the price of the goods which it deals in to their natural price. In this extraordinary case, the publick interest requires that some stock should be withdrawn from those employments which in ordinary cases are more advantageous, and turned towards one which in ordinary cases is less advantageous to the publick: and in this extraordinary case, the natural interests and inclinations of men coincide as exactly with the publick interest as in all other ordinary cases, [466] and lead them to withdraw stock from the near, and to turn it towards the distant employment.

88 It is thus that the private interests and passions of individuals naturally dispose them to turn their stock towards the employments which in ordinary cases are most advantageous to the society.[80] But if from this natural preference they should turn too much of it towards those employments, the fall of profit in them and the rise of it in all others immediately dispose them to alter this faulty distribution. Without any intervention of law, therefore, the private interests and passions of men naturally lead them to divide and distribute the stock of every society, among all the different employments carried on in it, as nearly as possible in the proportion which is most agreeable to the interest of the whole society.

89 All the different regulations of the mercantile system, necessarily derange more or less this natural and most advantageous distribution of stock. But those which concern the trade to America and the East Indies derange it perhaps more than any other;[81] because the trade to those two great continents absorbs a greater quantity of stock than any two other branches of trade. The regulations, however, by which this derangement is effected in those two different branches of trade are not altogether the same. Monopoly is the great engine of both; but it is a different sort of monopoly. Monopoly of one kind or another, indeed, seems to be the sole engine of the mercantile system.[82]

90 [467] In the trade to America every nation endeavours to engross as

[b-b] employment *1*

[79] A related point is made at II.v.34.
[80] See above, IV.ii.9.
[81] See above, § 46.
[82] Smith refers to restraints on imports and the encouragement of exports as the 'two great engines' of the mercantile system, at IV.i.35 and IV.viii.1.

much as possible the whole market of its own colonies, by fairly excluding all other nations from any direct trade to them. During the greater part of the sixteenth century, the Portugueze endeavoured to manage the trade to the East Indies in the same manner, by claiming the sole right of sailing in the Indian seas, on account of the merit of having first found out the road to them. The Dutch still continue to exclude all other European nations from any direct trade to their spice islands. Monopolies of this kind are evidently established against all other European nations, who are thereby not only excluded from a trade to which it might be convenient for them to turn some part of their stock, but are obliged to buy the goods which that trade deals in somewhat dearer, than if they could import them themselves directly from the countries which produce them.

91 But since the fall of the power of Portugal, no European nation has claimed the exclusive right of sailing in the Indian seas, of which the principal ports are now open to the ships of all European nations. Except in Portugal,[83] however, and within these few years in France, the trade to the East Indies has in every European country been subjected to an exclusive company. Monopolies of this kind are properly established against the very nation which erects them.[84] The greater part of that nation are thereby not only excluded from a trade to which it might be con-[468] venient for them to turn some part of their stock, but are obliged to buy the goods which that trade deals in, somewhat dearer than if it was open and free to all their countrymen. Since the establishment of the English East India company, for example, the other inhabitants of England, over and above being excluded from the trade, must have paid in the price of the East India goods which they have consumed, not only for all the extraordinary profits which the company may have made upon those goods in consequence of their monopoly, but for all the extraordinary waste which the fraud and abuse, inseparable from the management of the affairs of so great a company, must necessarily have occasioned. The absurdity of this second kind of monopoly, therefore, is much more manifest than that of the first.

92 Both these kinds of monopolies derange more or less the natural distribution of the stock of the society: but they do not always derange it in the same way.

93 Monopolies of the first kind always attract to the particular trade in which they are established, a greater proportion of the stock of the society than what would go to that trade of its own accord.

94 Monopolies of the second kind, may sometimes attract stock towards the particular trade in which they are established, and sometimes repel it

[83] See below, IV.vii.c.100.
[84] See above, I.viii.26. Smith considers the disadvantages of exclusive companies with regard to the colonies at IV.vii.b.22 and offers an extensive account of their record in V.i.e.

N 80

from that trade according to different circumstances. In poor countries they naturally [469] attract towards that trade more stock than would otherwise go to it. In rich countries they naturally repel from it a good deal of stock which would otherwise go to it.

95 Such poor countries as Sweden and Denmark, for example, would probably have never sent a single ship to the East Indies, had not the trade been subjected to an exclusive company. The establishment of such a company necessarily encourages adventurers. Their monopoly secures them against all competitors in the home market, and they have the same chance for foreign markets with the traders of other nations. Their monopoly shows them the certainty of a great profit upon a considerable quantity of goods, and the chance of a considerable profit upon a great quantity. Without such extraordinary encouragement, the poor traders of such poor countries would probably never have thought of hazarding their small capitals in so very distant and uncertain an adventure as the trade to the East Indies must naturally have appeared to them.[85]

96 Such a rich country as Holland, on the contrary, would probably, in the case of a free trade, send many more ships to the East Indies than it actually does. The limited stock of the Dutch East India company probably repels from that trade many great mercantile capitals which would otherwise go to it. The mercantile capital of Holland is so great that it is, as it were, continually overflowing, sometimes into the publick funds of foreign countries, sometimes into loans [470] to private traders and adventurers of foreign countries, sometimes into the most round-about foreign trades of consumption, and sometimes into the carrying trade. All near employments being completely filled up, all the capital which can be placed in them with any tolerable profit being already placed in them, the capital of Holland necessarily flows towards the most distant employments.[86] The trade to the East Indies, if it [c]was[c] altogether free, would probably absorb the greater part of this redundant capital. The East Indies offer a market both for the manufactures of Europe and for the gold and silver as well as for several other productions of America, greater and more extensive than both Europe and America put together.

97 Every derangement of the natural distribution of stock is necessarily hurtful to the society in which it takes place; whether it be by repelling from a particular trade the stock which would otherwise go to it, or by attracting towards a particular trade that which would not otherwise come

[c-c] were *4–6*

[85] See below, V.i.e.30, where Smith defends temporary monopolies, including those granted to merchants who first establish a hazardous trade. See also § 2 of the same section.

[86] See above, II.v.35.

to it.[87] If, without any exclusive company, the trade of Holland to the East Indies would be greater than it actually is, that country must suffer a considerable loss by part of its capital being excluded from the employment most convenient for that part. And in the same manner, if, without an exclusive company, the trade of Sweden and Denmark to the East Indies would be less than it actually is, or, what perhaps is more probable, would not exist at all, those two countries must likewise suffer a con-[471]siderable loss by part of their capital being drawn into an employment which must be more or less unsuitable to their present circumstances. Better for them, perhaps, in their present circumstances, to buy East India goods of other nations, even though they should pay somewhat dearer, than to turn so great a part of their small capital to so very distant a trade, in which the returns are so very slow, in which that capital can maintain so small a quantity of productive labour at home, where productive labour is so much wanted, where so little is done, and where so much is to do.

98 Though without an exclusive company, therefore, a particular country should not be able to carry on any direct trade to the East Indies, it will not from thence follow that such a company ought to be established there, but only that such a country ought not in these circumstances to trade directly to the East Indies. That such companies are not in general necessary for carrying on the East India trade, is sufficiently demonstrated by the experience of the Portugueze, who enjoyed almost the whole of it for more than a century together without any exclusive company.

99 No private merchant, it has been said, could well have capital sufficient to maintain factors and agents in the different ports of the East Indies, in order to provide goods for the ships which he might occasionally send thither; and yet, unless he was able to do this, the difficulty of finding a cargo might frequently make his ships lose the season for returning, and the ex-[472]pence of so long a delay would not only eat up the whole profit of the adventure, but frequently occasion a very considerable loss. This argument, however, if it proved any thing at all, would prove that no one great branch of trade could be carried on without an exclusive company, which is contrary to the experience of all nations. There is no great branch of trade in which the capital of any one private merchant is sufficient, for carrying on all the subordinate branches which must be carried on, in order to carry on the principal [d]one[d]. But when a nation is ripe for any great branch of trade, some merchants naturally turn their capitals towards the principal, and some towards the subordinate branches of it; and though all the different branches of it are in this manner carried on, yet it very seldom happens that they are all carried on by the capital of one private merchant.

[d-d] branch *1*

[87] See above, § 43, and IV.ii.3 for an elaboration of this point.

If a nation, therefore, is ripe for the East India trade, a certain portion of its capital will naturally divide itself among all the different branches of that trade. Some of its merchants will find it for their interest to reside in the East Indies, and to employ their capitals there in providing goods for the ships which are to be sent out by other merchants who reside in Europe. The settlements which different European nations have obtained in the East Indies, if they were taken from the exclusive companies to which they at present belong and put under the immediate protection of the sovereign, would render this residence both safe and easy, at least to the mer-[473]chants of the particular nations to whom those settlements belong. If at any particular time that part of the capital of any country which of its own accord tended and inclined, if I may say so, towards the East India trade, was not sufficient for carrying on all those different branches of it, it would be a proof that, at that particular time, that country was not ripe for that trade, and that it would do better to buy for some time, even at a higher price, from other European nations, the East India goods it had occasion for, than to import them itself directly from the East Indies. What it might lose by the high price of those goods could seldom be equal to the loss which it would sustain by the distraction of a large portion of its capital from other employments more necessary, or more useful, or more suitable to its circumstances and situation, than a direct trade to the East Indies.

100 Though the Europeans possess many considerable settlements both upon the coast of Africa and in the East Indies, they have not yet established in either of those countries such numerous and thriving colonies as those in the islands and continent of America. Africa, however, as well as several of the countries comprehended under the general name of the East Indies, are inhabited by barbarous nations. But those nations were by no means so weak and defenceless as the miserable and helpless Americans; and in proportion to the natural fertility of the countries which they inhabited, they were besides much more populous. The most barba-[474]rous nations either of Africa or of the East Indies were shepherds; even the Hottentots were so.[88] But the natives of every part of America, except Mexico and Peru, were only hunters; and the difference is very great between the number of shepherds and that of hunters whom the same extent of equally fertile territory can maintain.[89] In Africa and the East Indies, therefore, it was more difficult to displace the natives, and to extend the European plantations over the greater part of the lands of the original inhabitants. The genius of exclusive companies, besides, is unfavourable, it has already been observed, to the growth of new colonies, and has pro-

[88] Smith remarked in FA that the Hottentots 'are the most barbarous nation of shepherds that is known in the world'.

[89] This point is elaborated below, V.i.a.5.

bably been the principal cause of the little progress which they have made in the East Indies.[90] The Portugueze carried on the trade both to Africa and the East Indies without any exclusive companies,[91] and their settlements at Congo, Angola, and Benguela on the coast of Africa, and at Goa in the East Indies, though much depressed by superstition and every sort of bad government, yet bear some faint resemblance to the colonies of America, and are partly inhabited by Portugueze who have been established there for several generations. The Dutch settlements at the Cape of Good Hope and at Batavia, are at present the most considerable colonies which the Europeans have established either in Africa or in the East Indies, and both ᵉtheseᵉ settlements are peculiarly fortunate in their situation. The Cape of Good Hope was inhabited by a race of people almost as barbarous and quite as inca-[475]pable of defending themselves as the natives of America. It is besides the half-way house, if one may say so, between Europe and the East Indies, at which almost every European ship makes some stay both in going and returning. The supplying of those ships with every sort of fresh provisions, with fruit and sometimes with wine, affords alone a very extensive market for the surplus produce of the colonists. What the Cape of Good Hope is between Europe and every part of the East Indies, Batavia is between the principal countries of the East Indies. It lies upon the most frequented road from Indostan to China and Japan, and is nearly about mid-way upon that road. Almost all the ships too that sail between Europe and China touch at Batavia; and it is, over and above all this, the center and principal mart of what is called the country trade of the East Indies; not only of that part of it which is carried on by Europeans, but of that which is carried on by the native Indians; and vessels navigated by the inhabitants of China[92] and Japan, of Tonquin, Malacca, Cochin-China, and the island of Celebes, are frequently to be seen in its port. Such advantageous situations have enabled those two colonies to surmount all the obstacles which the oppressive genius of an exclusive company may have occasionally opposed to their growth. They have enabled Batavia to surmount the additional disadvantage of perhaps the most unwholesome climate in the world.

101 [476] The English and Dutch companies, though they have established no considerable colonies, except the two above mentioned, have both made considerable conquests in the East Indies. But in the manner in which they both govern their new subjects, the natural genius of an exclusive company

ᵉ-ᵉ those *1*

[90] Above, IV.vii.b.22, and cf. IV.vii.c.103.

[91] See above, IV.vii.c.91.

[92] See above, I.ix.15, where Smith comments on the discouragement to foreign trade in China.

has shown itself most distinctly.[93] In the spice islands the Dutch ʃare said toʃ burn all the spiceries which a fertile season produces beyond what they expect to dispose of in Europe with such a profit as they think sufficient.[94] In the islands where they have no settlements, they give a premium to those who collect the young blossoms and green leaves of the clove and nutmeg trees which naturally grow there, but which this ᵍsavageᵍ policy has now, it is said, almost compleatly extirpated. Even in the islands where they have settlements they have very much reduced, it is said, the number of those trees. If the produce even of their own islands was much greater than what suited their market, the natives, they suspect, might find means to convey some part of it to other nations; and the best way, they imagine, to secure their own monopoly, is to take care that no more shall grow than what they themselves carry to market. By different arts of oppression they have reduced the population of several of the Moluccas nearly to the number which is sufficient to supply with fresh provisions and other necessaries of life their own insignificant garrisons, and such of their ships as occasionally come [477] there for a cargo of spices. Under the government even of the Portugueze, however, those islands are said to have been tolerably well inhabited. The English company have not yet had time to establish in Bengal so perfectly destructive a system. The plan of their government, however, has had exactly the same tendency. It has not been uncommon, I am well assured, for the chief, that is, the first clerk of a factory, to order a peasant to plough up a rich field of poppies, and sow it with rice or some other grain. The pretence was, to prevent a scarcity of provisions; but the real reason, to give the chief an opportunity of selling at a better price a large quantity of opium, which he happened then to have upon hand. Upon other occasions the order has been reversed; and a rich field of rice or other grain has been ploughed up, in order to make room for a plantation of poppies; when the chief foresaw that extraordinary profit was likely to be made by opium.[95] The servants of the company have upon several occasions attempted to establish in their own favour the monopoly of some of the most important branches, not only of the foreign, but of the inland trade of the country. Had they been allowed to go on, it is impossible that they should not at some time or another have attempted to restrain the production of the particular articles of which they had thus usurped the monopoly, not only to the quantity which they themselves could purchase, but to that which they could expect to sell with such a profit as they might think sufficient. In [478] the course

ʃ-ʃ *2–6* ᵍ-ᵍ barbarous *1*

[93] The English colonies in the East Indies are cited as an example of the decaying economy in I.viii.26. The history of the East India Company is reviewed at V.i.e.26–30.

[94] A similar point is made above, IV.v.b.4 and I.xi.b.33.

[95] See above, IV.v.b.6.

of a century or two, the policy of the English company would in this manner have probably proved as compleatly destructive as that of the Dutch.

102 Nothing, however, can be more directly contrary to the real interest of those companies, considered as the sovereigns of the countries which they have conquered, than this destructive plan. In almost all countries the revenue of the sovereign is drawn from that of the people. The greater the revenue of the people, therefore, the greater the annual produce of their land and labour, the more they can afford to the sovereign. It is his interest, therefore, to increase as much as possible that annual produce. But if this is the interest of every sovereign, it is peculiarly so of one whose revenue, like that of the sovereign of Bengal, arises chiefly from a land-rent.[96] That rent must necessarily be in proportion to the quantity and value of the produce, and both the one and the other must depend upon the extent of the market. The quantity will always be suited with more or less exactness to the consumption of those who can afford to pay for it, and the price which they will pay will always be in proportion to the eagerness of their competition. It is the interest of such a sovereign, therefore, to open the most extensive market for the produce of his country, to allow the most perfect freedom of commerce, in order to increase as much as possible the number and the competition of buyers; and upon this account to abolish, not only all monopolies, but [479] all restraints upon the transportation of the home produce from one part of the country to another, upon its exportation to foreign countries, or upon the importation of goods of any kind for which it can be exchanged. He is in this manner most likely to increase both the quantity and value of that produce, and consequently of his own share of it, or of his own revenue.

103 But a company of merchants are, it seems, incapable of considering themselves as sovereigns, even after they have become such.[97] Trade, or buying in order to sell again, they still consider as ʰtheirʰ principal business, and by a strange absurdity, regard the character of the sovereign as but an appendix to that of the merchant, as something which ought to be made subservient to it, or by means of which they may be enabled to buy cheaper in India, and thereby to sell with a better profit in Europe. They endeavour for this purpose to keep out as much as possible all competitors from the market of the countries which are subject to their government, and consequently to reduce, at least, some part of the surplus produce

ʰ-ʰ the *1*

[96] See below, V.ii.a.13, where it is pointed out that land rents have provided the funds for many a state which has escaped from the shepherd stage. Details of the land tax in Bengal are given in William Bolts, *Considerations on India Affairs, particularly respecting the Present State of Bengal and its Dependencies* (London, 1772), i, chapter XII.

[97] See above, IV.vii.b.11, where the government of merchants is described as the 'worst of all'.

of those countries to what is barely sufficient for supplying their own demand, or to what they can expect to sell in Europe with such a profit as they may think reasonable. Their mercantile habits draw them in this manner, almost necessarily, though perhaps insensibly, to prefer upon all ordinary occasions the little and transitory profit of the monopolist to the great and permanent re-[480]venue of the sovereign, and would gradually lead them to treat the countries subject to their government nearly as the Dutch treat the Moluccas. 'It is the interest of the East India company, considered as sovereigns, that the European goods which are carried to their Indian dominions, should be sold there as cheap as possible; and that the Indian goods which are brought from thence should bring there as good a price, or should be sold there as dear as possible. But the reverse of this is their interest as merchants. As sovereigns, their interest is exactly the same with that of the country which they govern. As merchants their interest is directly opposite to that interest.'

104 But if the genius of such a government, even as to what concerns its direction in Europe, is in this manner essentially and perhaps incurably faulty, that of its administration in India is still more so.[98] That administration is necessarily composed of a council of merchants, a profession no doubt extremely respectable, but which in no country in the world carries along with it that sort of authority which naturally over-awes the people, and without force commands their willing obedience. Such a council can command obedience only by the military force with which they are accompanied, and their government is therefore necessarily military and despotical. Their proper business, however, is that of merchants. It is to sell, upon their masters account, the European goods consigned to them, and to buy in return Indian goods for the [481] European market. It is to sell the one as dear and to buy the other as cheap as possible, and consequently to exclude as much as possible all rivals from the particular market where they keep their shop. The genius of the administration, therefore, so far as concerns the trade of the company, is the same as that of the direction. It tends to make government subservient to the interest of monopoly, and consequently to stunt the natural growth of some parts at least of the surplus produce of the country to what is barely sufficient for answering the demand of the company.

105 All the members of the administration, besides, trade more or less upon their own account, and it is in vain to prohibit them from doing so. Nothing can be more compleatly foolish than to expect that the clerks of a great counting-house at ten thousand miles distance, and consequently almost quite out of sight, should, upon a simple order from their masters, give up

'–' 2–6

[98] Smith considers the constitutional reforms of 1773 at V.i.c.26, and see also V.ii.a.7.

at once doing any sort of business upon their own account, abandon for ever all hopes of making a fortune, of which they have the means in their hands, and content themselves with the moderate salaries which those masters allow them, and which, moderate as they are, can seldom be augmented, being commonly as large as the real profits of the company trade can afford.[99] In such circumstances, to prohibit the servants of the company from trading upon their own account, can have scarce any other effect than to enable the superior servants, under pretence of executing their masters order, to oppress such of the inferior ones [482] as have had the misfortune to fall under their displeasure. The servants naturally endeavour to establish the same monopoly in favour of their own private trade as of the publick trade of the company. If they are suffered to act as they could wish, they will establish this monopoly openly and directly, by fairly prohibiting all other people from trading in the articles in which they chuse to deal; and this, perhaps, is the best and least oppressive way of establishing it. But if by an order from Europe they are prohibited from doing this, they will, notwithstanding, endeavour to establish a monopoly of the same kind, secretly and indirectly, in a way that is much more destructive to the country. They will employ the whole authority of government, and pervert the administration of justice, in order to harass and ruin those who interfere with them in any branch of commerce which, by means of agents, either concealed, or at least not publickly avowed, they may chuse to carry on. But the private trade of the servants will naturally extend to a much greater variety of articles than the publick trade of the company. The publick trade of the company extends no further than the trade with Europe, and comprehends a part only of the foreign trade of the country. But the private trade of the servants may extend to all the different branches both of its inland and foreign trade. The monopoly of the company can tend only to stunt the natural growth of that part of the surplus produce which, in the case of a free trade, would be exported to Europe. That of the servants tends to stunt the natural growth of [483] every part of the produce in which they chuse to deal, of what is destined for home consumption, as well as of what is destined for exportation; and consequently to degrade the cultivation of the whole country, and to reduce the number of its inhabitants. It tends to reduce the quantity of every sort of produce, even that of the necessaries of life, whenever the servants of the company chuse to deal in them, to what those servants can both afford to buy and expect to sell with such a profit as pleases them.[100]

[99] See above, I.ix.21.

[100] In his *Considerations on India Affairs*, William Bolts made the same points several times, e.g. i.206–7: 'We have seen all merchants from the interior parts of Asia effectually prevented from having any mercantile intercourse with Bengal, while, at the same time, the natives in general are in fact deprived of all trade within those provinces, it being

32

106 From the nature of their situation too the servants must be more disposed to support with rigorous severity their own interest against that of the country which they govern, than their masters can be to support theirs. The country belongs to their masters, who cannot avoid having some regard for the interest of what belongs to them. But it does not belong to the servants. The real interest of their masters, if they were capable of understanding it, is the same with that of the country*, and it is from ignorance [k]chiefly,[k] and the meanness of mercantile prejudice, that they ever oppress it. But the real interest of the servants is by no means the same with that of the country, and the most perfect information would not necessarily put an end to their oppressions. The regulations accordingly which have been sent out from Europe, though they have been frequently weak, have [l]upon most occasions[l] been [484] well-meaning. More intelligence and perhaps less good-meaning has sometimes appeared in those established by the servants in India. It is a very singular government in which every member of the administration wishes to get out of the country, and consequently to have done with the government, as soon as he can, and to whose interest, the day after he has left it and carried his whole fortune with him, it is perfectly indifferent [m]though[m] the whole country was swallowed up by an earthquake.[101]

[l]* The interest of every proprietor of India Stock, however, is by no means the same with that of the country in the government of which his vote gives him some influence. See Book V. Chap. i. Part 3d.[l]

[l–l] *3–6* In ed. 2 the note reads: This would be exactly true if those masters never had any other interest but that which belongs to them as Proprietors of India stock. But they frequently have another of much greater importance. Frequently a man of great, sometimes even a man to moderate fortune, is willing to give thirteen or fourteen hundred pounds (the present price of a thousand pounds share in India stock) merely for the influence which he expects to acquire by a vote in the Court of Proprietors. It gives him a share, though not in the plunder, yet in the appointment of the plunderers of India; the Directors, although they make those appointments, being necessarily more or less under the influence of the Court of Proprietors, which not only elects them, but sometimes over-rules their appointments. A man of great or even a man of moderate fortune, provided he can enjoy this influence for a few years, and thereby get a certain number of his friends appointed to employments in India, frequently cares little about the dividend which he can expect from so small a capital, or even about the improvement or loss of the capital itself upon which his vote is founded. About the prosperity or ruin of the great empire, in the government of which that vote gives him a share, he seldom cares at all. No other sovereigns ever were, or from the nature of things ever could be, so perfectly indifferent about the happiness or misery of their subjects, the improvement or waste of their dominions, the glory or disgrace of their administration; as, from irresistible moral causes, the greater part of the Proprietors of such a mercantile Company are, and necessarily must be. [The matter of this note is incorporated in V.i.e.25, i.e. in a section of the WN which first appeared in 2A and ed. 3.]

[k–k] only *1* [l–l] commonly *1* [m–m] if *1*

wholly monopolized by a few Company's servants and their dependents: In such a situation, what commercial country can flourish.'

[101] Smith makes a related point at V.ii.k.74 in discussing the activities of tax-farmers who are indifferent to the fate of the people they abuse, unlike the sovereign whose proper interest it is to protect them. The example of China 'suddenly swallowed up by an earthquake' is cited in TMS III.3.4.

107 I mean not, however, by any thing which I have here said, to throw any odious imputation upon the general character of the servants of the East India company, and much less upon that of any particular persons. It is the system of government, the situation in which they [n]are[n] placed, that I mean to censure; not the character of those who have acted in it. They acted as their situation naturally directed, and they who have clamoured the loudest against them would, probably, not have acted better themselves.[102] In war and negociation, the councils of Madras and Calcutta have upon several occasions conducted themselves with a resolution and decisive wisdom which would have done honour to the senate of Rome in the best days of that republick. The members of those councils, however, had been bred to professions very different from war and politicks. But their situation alone, without education, experience, or even example, seems to have formed in them all at once the great qualities which it required, and to have inspired them both with abilities and virtues which they [485] themselves could not well know that they possessed. If upon some occasions, therefore, it has animated them to actions of magnanimity which could not well have been expected from them, we should not wonder if upon others it has prompted them to exploits of somewhat a different nature.

108 Such exclusive companies, therefore, are nuisances in every respect; always more or less inconvenient to the countries in which they are established, and destructive to those which have the misfortune to fall under their government.

[n–n] were *1–2*

[102] See below, V.i.e.26, where it is pointed out that abuses of this kind often reflect the circumstances prevailing; what Smith calls 'irresistible moral causes'.

RECOMMENDED READINGS

The Economic Analysis of Government Behavior: Choices in the Public Arena.

Becker, Gary S. "A Theory of Competition among Pressure Groups for Political Influence." Quarterly Journal of Economics 98(August 1983): 371-400.

Buchanan, James M. "Rent Seeking and Profit Seeking." In Toward a Theory of the Rent-Seeking Society, edited by J.M. Buchanan, R.D. Tollison, and G. Tullock. College Station: Texas A&M Press, 1980.

Buchanan, James M., R.D. Tollison and Gordon Tullock, eds. Toward a Theory of the Rent-Seeking Society. College Station: Texas A&M Press, 1980.

Buchanan, James M. and Gordon Tullock. The Calculus of Consent. Ann Arbor: University of Michigan Press, 1965.

Krueger.Anne O. "The Political Economy of the Rent-Seeking Society," American Economic Review 64(June 1974): 291-303.

McChesney, Fred. "Rent Extraction and Rent Creation in the Economic Theory of Regulation." Journal of Legal Studies 16(January 1987): 101-18.

McCormick, Robert E. "A Review of the Economics of Regulation: The Political Process." In Roger E. Meiners and Bruce Yandle, Regulation and the Reagan Era. San Francisco: The Independent Institute, 1989, 16-37.

Noll, Roger G. and Bruce M. Owen. The Political Economy of Deregulation. Washington: American Enterprise Institute, 1983.

Peltzman, Sam. "Toward a More General Theory of Regulation." Journal of Law and Economics 19(August 1976): 211-40.

Posner, Richard A. "Regulation by Taxation." Bell Journal 2(Spring 1971):22-50.

Stigler, George J. "The Theory of Economic Regulation." Bell Journal 2(Spring 1971):3-21.

Tollison, Robert D. "Rent Seeking: A Review." Kyklos 35(1982): 575-602.

Tullock, Gordon. "The Welfare Costs of Tariffs, Monopolies, and Theft." Western Economic Journal 5(June 1967): 224-32.

Yandle, Bruce. "Bootleggers and Baptists." Regulation 7(May,June 1983):12-16.

European Journal of Political Economy 6 (1990) 165–179. North-Holland

THE DECLINE AND RISE OF POLITICAL ECONOMY

Bruce YANDLE*

Clemson University, Clemson, SC 29634, USA

Accepted for publication April 25, 1990

Political economy and Public Choice emerged in the mid-1960s as sub-disciplines involving economists, political scientists, and legal scholars. Interestingly, political economy was present at the birth of these fields of inquiry but was displaced by new specialized disciplines. This essay offers an economic explanation of the decline and rise of political economy. The theory presented suggests that political economy experiences a rebound following periods of increased government intervention. But instead of inducing further specialization, the impulse leads to unification within and among disciplines.

1. Introduction

The application of economic logic to political behavior for the purpose of explaining the way the world works is now the basis of a burgeoning academic enterprise for economists, political scientists, philosophers, and legal scholars. Interestingly enough, those who labor in the political economy vineyard, which includes, inter alia, Public Choice, law and economics, property rights and constitutional economics, have returned to the roots of their disciplines in their search for logical models that explain human behavior.

Viewing political actors as economic agents, the political economists and political scientists who share common models make falsifiable predictions about political behavior. With empirical and rich institutional data in hand, the underlying theory can be tested and the results compared with competing notions. In many cases, the economist's narrowly construed efficiency model and political scientist's public interest theory are found wanting. Ideological and political variables do not dominate explanations of public sector decision making. The richer institutional theories of the past are more powerful.

What explains the post-1960s rise of the new political economy? And better yet, what explains the decline of the old political economy, what might

*The author expresses appreciation for helpful comments and criticisms to Hugh H. Macaulay, Clark Nardinelli, and the editor of this journal.

be termed classical economics, that evolved in the 19th century? Do cycles of government intervention explain the pronounced cyclical life of political economy? Does the twentieth-century growth of government explain the resurgence of political economy?

This essay examines these questions and puts forward elements of a theoretical explanation of the decline of the old and rise of the new political economy. The theory is founded on notions of markets for ideas along with the entry of economic agents to the political process. While the essay focuses decidedly on the evolution of economic thought, discussions of related changes in other disciplines are also included.

The essay's next section scans the development of economics and describes the evolution of a dichotomy that divided the world into private and public sectors. While elegant theoretical models were being constructed to use in examining the two social spheres, the older notions of political economy were snipped away from their philosophical roots and pure economic models were transplanted in their place. The section offers an explanation of how that happened and illustrates dramatic changes in the economist's way of looking at the political world.

The section that follows examines the decline and rise of political economy using a theory of markets. The discussion there assumes that the market for ideas determines partly the success and failure of idea producers, which is to say that the demand for intellectual products is as important as supply. The value of ideas is based on their usefulness in predicting human behavior. Ideas with higher predictive content displace those with a lesser amount and that occurs with a vengeance when demand increases.

The last major section of the essay focuses on the U.S. regulatory experience and applies some of the concepts discussed in the earlier section. Certain characteristics of the U.S. experience that invigorate the rise of political economy are identified and discussed. The section argues that the expansion of special interest rules of the 1970s provided the necessary laboratory where competing theories of government action were tested. The new political economy became the stronger contender in the intellectual market place that emerged. Some final thoughts conclude the essay.

2. From political economy to economics

Discussions of economic agents working in the political arena accompanied the birth of economics as a discipline and gave the new field its distinguishing name: political economy. One needs only consider the titles of early treatises in economics to see this: Adam Smith: *An Inqury into the Nature and Causes of the Wealth of Nations* (1776); J.B. Say: *A Treatise on Political Economy* (1803); David Ricardo: *Principles of Political Economy* (1817); Nassau Senior: *An Outline of the Science of Political Economy* (1836);

and John Stuart Mill: *Principles of Political Economy* (1848). However, by Ricardo's time, the focus of political economy had narrowed considerably in a search for laws that explained economic behavior.

Unlike later writers, the authors of these earlier treatises saw one world without a public sector/private sector demarcation. Though disintegrating rapidly, there was a certain unification of thought. Fields of inquiry such as moral philosophy and jurisprudence became more specialized disciplines of philosophy, politics, legal studies, and economics. Each of these was partly subsumed in the early field of political economy.

Another important feature of the classical writers is seen in their openness in writing about freedom and liberty. They apparently viewed those conditions of man as legitimate subjects for discussion and inquiry. Now viewed as being somehow inappropriate concepts for scholarly inquiry by economists, these scarce institutional features of social life often formed the basis of models that had to do with human well-being.

The cumulative effects of the age of rationalism, the enlightenement, and the attractiveness and power of newly emerging sciences, especially physics, contributed to a reshaping of political economy, to one that projected models devoid of institutional content to be used for theoretical analysis. By 1874 when Walras published his *Elements of Pure Economics*, the discipline had been reshaped. Just before the publication of Walras' tour de force, Jevons had published his *Political Economy* (1871). Though Jevons indicated his next major treatise would be named *Economics*, he did not live to complete it.[1]

The new science of economics contained two analytical engines. One engine was used to examine the behavior of economic agents in private sector settings. Another set of models took shape for the analysis of the behavior of political agents who were generally viewed by economists as seeking efficient policy, though systematically missing the mark. There were no analytical bridges connecting the two analytical worlds.

Though useful to be sure, the stripping away of institutional clothing that occurred left a major unanswered question: Why do politicians systematically fail to achieve efficiency goals? The unanswered question raised serious doubts about the power of economic models to predict or explain real world outcomes. Undeterred by that, analysts focused on welfare economics, which is to say the implicit answer to the question had to do with informing politicians about efficiency. Instead of claiming to explain the way the world worked, welfare economists resorted to saying how it should work, again assuming that politicians were supremely interested in efficiency as narrowly construed in economic models. All along, the politicians seemed uninterested in, if not unconvinced by, the analyses.

A comparison of the thoughts of Adam Smith (1776) and J.B. Say (1803)

[1]I am indebted to Clark Nardinelli for these insights.

with those of A.C. Pigou (1920) illustrates the transition and the frustration. Using the analysis of political economy that viewed economic motivation as a driving force in all walks of life, which of course included politics, Smith had this to say about businessmen and politics [Smith (1937, p. 150)]:

> The interest of the dealers ... in particular branch of trade or manufactures, is always in some respects different from, and even opposite to, that of the public. To widen the market and narrow the competition is always the interest if the dealers. To widen the market may be agreeable enough to the interests of the public; but to narrow the competition must always be against it, and can serve only to enable the dealers, by raising their profits above what they naturally would be, to levy, for their own benefit, an absurd tax upon the rest of their fellow citizens. *The proposal of a new law which comes from this order ought never to be adopted till after having been long and carefully examined.* (emphasis added)

Smith implies special interest influence, but implicitly described an insulated legislature that independently ponders proposed laws. He clearly sees one world of economic agents competing directly for customer patronage and then jointly for rules to reduce the force of the ensuing competition. A profit maximizing dealer will presumably spend resources across all margins of influence, attaining equilibrium, let us say, where the return from buying political favors is equal to that obtained from improving service and product for consumers. Along with his effort to explain the way the world works, Smith offered a few words of caution to the legislature.

J.B. Say points directly to an activist legislature in his political economy and speaks of its sharing in the political gains created for special interest groups. He says [Say (1966, pp. 146–147)]:

> If one individual, or one class, can call in the aid of authority to ward off the effects of competition, it acquires a privilage to the prejudice and at the cost of the whole community, it can then make sure of profits not altogether due to the productive services rendered, but composed in part of an actual tax upon consumers for its private profit; which tax it commonly shares with the authority that thus unjustly lends its support.

By 1920, the world of economics and the view of politicians had changed dramatically. To illustrate, A.C. Pigou offers pages of guidance to legislators with few words for those who seek simply to understand how politicians interact with economic agents to form the world's legal institutions. Pigou has a benign view of politics. His analysis sees a clear separation between public and private sectors. As if adopting Freud's behavioral model where the ego and super ego struggle to control the id, Pigou leaves the reader with a vision of government as the controller/perfecter of errant market forces.

Finding numerous problems where private sector incentives seem to fail, the collective superhuman mind of government can fine-tune the private sector, taxing a little here, subsidizing there until all rejoice with a grand salute to Pareto's ghost.

Elaborating on problems where the private cost of decision makers fails to include elements of cost imposed on unrelated parties, Pigou (1932, p. 192) sees the situation this way:

> It is plain that divergences between private and social net product of the kinds we have so far been considering cannot ... be mitigated by modifications of the contractual relation between any two contracting parties, because the divergence arises out of a service or disservice rendered to persons other than the contracting parties. It is, however, possible for the State, if it so chooses, to remove the divergence in any field by 'extraordinary encouragements' or 'extraordinary restraints' upon investments in that field. The most obvious forms which these encouragements and restraints may assume are, of course, those of bounties and taxes.

'It is, however, possible for the State, if it so chooses.' With that casual statement, Pigou brushed against political choice briefly as he continued the development of neat arguments in an institutional vacuum. While the economic vacuum is analyzed, politicians wait quietly with their law-making engines in neutral, considering whether or not to accept Pigou's advice.[2] Unlike the legislators in Smith's world, Pigou's are a feature of nature, not a part of the world struggle described by Smith and J.B. Say. Explaining how they choose to act was not Pigou's purpose. Explaining how they should act was more important, which is simply another way of saying that is what welfare economics is about.

Though of key importance in the development of economic theory, Pigou's strictures are minor when compared with the vast outpouring of politically sterile models that flowed from the institutional vacuum. Eventually, however, some economists became increasingly aware that politicians were (1) not particularly influenced by their writing about efficiency, and (2) greatly influenced by economic agents who seemed unaware of the theoretical barriers that separated private from public sectors. Some economists began to think about the political scientists' notion of capture. The conditional comment 'If they choose' became a point of departure for the study of collective choice. The new question was 'Why do they make certain choices?'

The rise of Public Choice in the 1960s with its unrelenting focus on the

[2]As the distinguished Professor of economics at Cambridge, Pigou was indeed teaching future government leaders. Perhaps, later welfare economists failed to realize that unlike Pigou they seldom had the future political elite in their lecture halls. (I am indebted to Clark Nardinelli for calling this to my attention.)

political choices Pigou touched in passing marked a return of the old political economy. Indeed, when the name of the new Public Choice Society was being debated in 1967, the political scientists and economists at the meeting struggled between 'Public Choice' and the 'New Political Economy' and finally settled on the former.[3] Once again, a group of economists, political scientists, and other scholars were seriously attempting to explain the way the world worked. And once again, scholars unabashedly inquired about freedom and the challenges faced by free people when organizing publicly supported activities.

3. The cyclical demand for political economy

What might explain the decline and rise of political economy? Does it reflect something about the world? Or something about economists, political scientists and the incentives they face? Surely the intelligence of generations of economists has not varied significantly. The explanation must be found in the intellectual marketplace where demand and supply affect the allocation of intellectual inputs and the determination of product characteristics. For that to be the case, one might look for variations in the value of government to special interest agents that correspond with variations in the relative attention paid by economists and others to political economy. When government expands its supply of special interest benefits, which is to say government becomes valuable to private parties, political economy would seem to be more in vogue, if not more valuable. When government reduces its special interest activities, political economy wanes.

By way of illustration, classical writers like Smith and Say devoted much of their attention to Mercantilism, the intervention of government into the working of international markets that had evolved over several centuries.[4] In the salad days of Mercantilism, the strictures were extensive and valuable. World forces changed, Mercantilism subsided relatively, and by the end of the 19th century the Western world experienced a long wave of economic prosperity. The collapse of industrial economies in the 1930s ushered in a new wave of state intervention that sowed the seeds of the next crop of political economists and provided a stimulus for broadening the scope of social science scholarship. All this suggests that political economy becomes productive after long periods of government intervention. But there is still a

[3]See note (1) in William C. Mitchell (1988, p. 117). Mitchell's article contains an excellent survey of early as well as seminal contributions to the hybrid discipline. Also, see Mitchell (1989). It is interesting that the new research program in constitutional political economy is sometimes referred to as the 'new political economy'. See Jack Wiseman (1990).

[4]For a good survey of thought on Mercantilism as well as historical background, see Robert B. Ekelund, Jr. and Robert F. Hebert (1983, pp. 32-60). Also, see Robert B. Ekelund, Jr. and Robert D. Tollison (1981).

puzzle to resolve: Why do diverse disciplines reunite and return to classical roots in the process of reforming political economic thought?

In a famous quotation, Adam Smith reminded his readers that the division of labor – specialization among factors of production – is limited by the extent of the market. In a famous article named by that quotation, Stigler (1968) developed an explanation of the organization of industries. As the market for a newly introduced product expands, specialized functions within the innovating firms can be spun off to form new firms. The resulting disintegration leads to lower production costs for components through specialization and the exploitation of ultimate economies of scale. Eventually, every possible function that can become independent is a free-standing firm. Transactions costs that enter among the various firms become a limiting factor in the ensuing pageant of competitively determined lower costs.

The reverse process indicates that when the product market shrinks, the children of the original firm tend to return home, becoming absorbed in the original firm, which is to say vertical mergers and consolidations occur. According to the theory, integration or unification is an indicator of declining markets, all else being equal. Disintegration is an indicator of expanding markets.

The cyclical story of political economy just outlined suggests the unification theory does not hold for intellectual products. We see just the reverse. Strong state intervention, a proxy for the market demand for political economy analysis, appears to be associated with integration of diverse thought; the thoughts of political scientists, philosophers, legal scholars and economists converge in an investigation of political economy. In a sense, members of the scholarly disciplines return home and re-examine the notions of Adam Smith, Hume, Locke and other founders.

If the division of labor story is generally valid, periods of extensive government intervention and regulatory growth should be followed by further specialization within disciplines, as opposed to the integration of thought represented by the joining of economics and political science to form Public Choice – the new political economy.

This conundrum requires some explanation, and Stigler offers some useful insights. The division of labor in the social sciences that split the world and then focused on political behavior apparently did not produce value-enhancing results along certain margins. For example, the institutional vacuums in economics that addressed efficiency did not adequately explain political action. Furthermore, the older Public Interest theory of political science – the counterpart to economic's efficiency theory – could not alone explain special interest motivation.

Other offspring of the parent disciplines performed better. Some spin-off disciplines appear to have flourished in their new specialized markets. For example, management, management science, and financial economics have

staked out territories and expanded. Public administration and policy sciences have as well. But the specialists that focused on the behavior of political agents became intellectually bankrupt. Unification of knowledge, the regeneration of political economy, is the counterpart of mergers involving failing firms. Theories of unification and specialization imply that value-enhancing sub-disciplines that offer predicative power about the world will survive and perhaps flourish. Those that cannot explain and predict will not.

The purveyors of predictions regarding the behavior of political agents in a political economy fell into the latter group, as did those who pushed the pure public interest theorist of political science. The return to the roots of economics that has occurred and the extraordinary growth of Public Choice as a hybrid discipline are evidence that an intellectual market test has been passed. Predictions that emerge from Stigler's division of labor theory imply that new subdisciplines within Public Choice and political economy will soon be observed.[5]

For the theory of cyclical political economy to hold water, there must be some explanation of the timing of the unification of politics and economics. Why does political economy re-emerge following periods of extreme government intervention? Consider the social scientist's laboratory and the number of natural experiments occurring in it. If there are few experiments and little attention focused on them, weak theories are apt to go unpunished. Upon observing behavior that defies his explanation, a public interest theorist can say there are always some bad apples. What are needed are better politicians. If wasteful and inefficient regulation is imposed by the legislature, the efficiency analyst can call on his stupidity theory. The politicians are just beyond hope. They must learn some basic welfare economics.

But if the number and frequency of natural experiments increase markedly and massive quantities of regulation emerge, the theorists are faced with a severe problem. Given the size of their sample, they must either explain outcomes or be discredited. The problem faced is comparable to that found when limited anecdotal evidence is all that can be mustered in an argument. The failure of a few anecdotes to jibe with respected theories is not quite so serious. However, when the theories are tested with a large statistically useful data set in a multiple regression, a lack of significant variables and low explanatory power can be very disturbing. It is equally disturbing to observe an almost uninterrupted series of anecdotes that appear to violate theoretical logic.

[5]As this was being written, the first issue of *Constitutional Political Economy* was coming off the press. James M. Buchanan (1990) describes carefully the new research program captured by the journal's title and explains how it relates to Public Choice.

4. The U.S. regulatory experience

The U.S. experience with federal regulation may illustrate the point. Since the late 1800s, there have been three distinct periods of regulatory growth.[6] The first period, sometimes called the Progressive Era, marks the origin of federal regulation. That was when the Interstate Commerce Commission was formed (1887), which was the first federal regulatory agency. The nation's first antitrust statute (1890) was also passed in that period. In response to the regulatory impulse, political scientists hawked the public interest theory and economists addressed equity and efficiency. Meanwhile evidence was accumulating. The second period, the 1930s, reflects the Great Depression when a large number of major regulatory agencies were formed. Again, old theories were proffered and time passed while evidence was accumulating. Finally, the last period reflects the environmental revolution of the 1970s. By then, almost a century had passed since the first regulatory phase and the evidence was in hand that would either refute or support competing hypotheses of political behavior. Political economy and Public Choice were already appearing on the scene.

On reviewing the actions of the Interstate Commerce Commission, some scholars were perplexed.[7] Neither the public interest nor efficiency theory would hold water. The agency seemed to serve special interests, and those interests varied across time. Examination of the behavior of the Food and Drug Administration, the Federal Trade Commission, the Occupational Safety and Health Administration, and countless other agencies revealed the same thing. Eventually, the vacuum-produced theories were discredited. There was integration in the intellectual market place. The political scientists' capture theory became more interesting. Political economy and Public Choice emerged as dominant disciplines for predicting the behavior of political agents in a political economy.

The resurgence of political economy, marked by a return to intellectual roots of the discipline, presented richer theories of political behavior based on old logic. Yes, political agents were assumed to be pursuing their self interests while serving their special interest principals – the people who put them in office. But no, the broader interests were not completely neglected. Although the substance of the public interest and efficiency theories are no longer the dominant theories, their clothing has survived.

The richer theories of political economy recognize that political actions require justification to broader audiences than those represented by special interests. And the membership composition of special interest groups varies depending on the issues to be debated and decided. At any given moment,

[6]More detail is provided in Yandle (1986).

[7]For a review of literature on the agencies mentioned here as well as others, see McCormick (1984, 1989).

some who were active members of a lobby group are now part of an unorganized mass of consumers. The variation in membership requires that sophisticated politicians offer acceptable and credible justification for their actions. To justify safety regulations that impose differential costs on old and new firms, giving advantage to the former, the politician must speak about human health and make a public interest appeal.[8]

To support costly command and control environmental regulations that have little to do with improving air quality but much to do with protecting employment of coal miners and preserving rents for particular firms, the political agent must speak as an environmentalist. With success, the politician serves competing special interest groups – those that presume to speak for consumers and others who quietly but effectively obtain restrictions on competitors' output. We are reminded of Adam Smith's advice about the dealers: 'The proposal of a new law (from the dealers) ought never to be adopted till after having been carefully considered.' The new theories argue that producer interests will be served, but not totally at the expense of consumer interests. The theories also suggest that politicians will position themselves to articulate demand for political services in ways that cause the quantity demanded always to exceed the quantity supplied.

Excess demand for political favors became part and parcel to the increased rate of production of special interest legislation and regulation in the 1970s. The resulting increase in contracts between politicians and economic agents generated more data for analysis by those who created the new political economy. How excess demand was managed was important to the emerging body of research findings.

Today's extensive literature on rent-seeking or the special interest theory of government includes substantial empirical support for the revived theories of political behavior.[9] When coupled with historical data that reaches across centuries the combined evidence tells us that market forces recognize no social boundaries. Just as Say told us, the arena for votes and political favors are traded in a market that facilitates political entry and exit. Political actions that create wealth for favored constituents are the services sold in the market, but how the actions are packaged and the certainty introduced to the contracts has great bearing on the process and on the resurgence of political economy that began in the 1960s.

Unlike unfettered markets for goods and services provided by firms where prices tend to balance the quantities demanded and supplied, political markets are characterized by perpetual excess demand. Of course, there is a tendency for the process to equilibrate, but not on the basis of some apparent market price. Whereas traders can enter and exit most political

[8]On this, see Yandle (1989).
[9]For a substantial survey of this literature, see McCormick (1984).

markets, entry is constitutionally barred for new political units. There are a fixed number of elected representatives.

The legal barriers provide potential monopoly power to the producers of favors, especially those who organize dominant parties in the legislature. As price setters, they can determine which deals are struck and how much is paid for each transaction. As a result, politicians tend to write laws that promise more than can ever be delivered universally. Phrases such as 'Every person has a right to a decent home, a safe workplace, clean air, the absence of hazardous waste in his community, and basic education' are common legislative language. It is known at the outset that no one has a duty to make good on the announced rights. Demand is inevitably greater than supply.

This feature of political demand and supply allows political brokers to discriminate across constituent groups. Those who pay the highest price (render the greatest political support) are among the first to obtain the desired service. In a multi-tiered political system – one with local, state, and national governments organized along party lines – where party support must be built at each level, valuable units of the service can be designated for discriminatory allocation at each tier. Five hundred units of subsidized housing, for example, may be provided annually in a regime that claims every citizen has a right to decent housing. Far short of the amount needed to satisfy the political promise, the 500 units are predictably allocated in ways that maximize the political rents they generate, not on the basis of some measure of need. Because of this, the efficiency oriented economist can always point to failed government programs; the outcomes never match the official goals. However, the political economist more often sees what theory predicts; the official goals are important items that disguise the private interest objectives of the process.

The desirability of excess demand predicts that politicians will not generally support the use of economic incentives in allocating governmental provided goods and services. The exception is found during periods of fiscal stringency when additional revenues for redistribution are more valuable at the margin than politically created rents.[10]

Command and control is the order of the day. For example, an approach for providing lower income families access to higher valued housing could be made with government provided vouchers, instead of government provided housing. A limited number of vouchers allocated to families on the basis of income and other characteristics that can be traded in ordinary housing markets will equilibrate demand and supply.

[10]The current expansion of user fees, including some that could be crudely termed Pigouvian, relates to this point. The growing U.S. deficit has apparently caused politicians to become more interested in using economic incentives. [For discussion, see Yandle (1989b, ch. 6)].

Although vouchers might be good for technical efficiency, they are not good for politicians. The politician supporting such schemes gives away a valuable rent-creating opportunity. He cannot predict with precision who will obtain the improved housing. The recipients of the benefits may be people from outside his political region. They will obviously not be obligated to him in the future.

In a similar fashion, pollution permits for specified amounts of allowable emissions that can be traded among firms that discharge common pollutants will reduce the overall cost of achieving a set pollution reduction goal. The permits are exceedingly valuable to firms that have high control costs. They are of less value to firms with lower control costs, which of course makes the market function. But the firms initially endowed with the permits receive the rents, not the politicians. Even if the politician chose to auction permits to the highest bidder in an open market, there would be problems. Payment in kind is preferred to payment in cash in most political economy markets. It is a matter of control and support. The successful politician must be able to deliver on implicit contracts. In exchange for support, he will arrange certain privileges. An invitation to a public auction is no privilege. The proceeds of the auction accrue to the treasury, not to the politician's campaign chest.

The political marketplace shares a common characteristic with all other markets, including the intellectual market discussed earlier. If the market for political services is small – there being limited scope for political redistribution, the politician will tend to handle transactions directly. In any case, we would expect the most highly valued transactions to be handled this way. Direct face-to-face contact with each recipient of political favors has a general advantage. Agency cost is reduced to zero when the elected official writes his own contracts, as opposed to delegating those negotiations to hired staff. However, like all producers of goods and services who rely on the use of specialized assets, the politician confronts the problem of positioning himself to maximize his lifetime wealth.

By passing laws that address pressing social problems, the politician increases demand for his services. At some point, the excess demand so generated is more than can be managed effectively on a direct basis. A division of labor is predicted to be associated with the growth of government, all else equal.

The growth of administative law and the expansion of administrative agencies provide the elected politicians with agents to handle less complex political deals. With the gains from specialization come agency costs. The appointed head of an administrative agency cannot be expected to perform exactly the way the politician would perform. To minimize that cost, the elected politicians grooms staff members for future administrative posts and makes the administrative agency subject to his oversight. Even so, important subtleties associated with political contracts are easily lost. The politician

must weigh the gains from specialization against the losses that spring from agency cost.

The bureacratic hierarchies that evolve with excess demand management and growth of government leave the theorist in a precarious position. Public interest theories that might explain the behavior of an abolsute monarch may prove robust, if the king is benevolent. The same theories are less likely to explain the behavior of a parliament, since division of labor abounds and competing special interest groups operate across the membership. Members of parliament have to be re-elected. The king has other problems.

Theories are further tested when additional layers of bureaucracy are added to the picture. While some parliamentary groups may be shown to serve the interests of large unorganized groups of consumers, it is unlikely that administrative agencies will have the same appearance. As specialized arms of the legislature, these units typically work to serve a well-defined group. They are special interest at the outset, though their rationale is based on a public interest theory. Transportation agencies naturally serve the interests of the transportation industry. International trade commissions naturally serve the interests of domestic firms that seek protection. Environmental protection agencies understandably balance the interests of producers who must operate within environmental rules against the demands of organized groups who seek to alter environmental use.

As the number of political contacts available to special interest groups grows, the observations to include in models that test public interest and private interest theories expand. In addition to there being more observations to consider, along with a larger outpouring of rules designed specifically for special groups, linkages between the expanding bureaucracy and elected politicians are stretched. Agency cost increases at every margin – between voter and elected official, elected official and government bureaucrat, and between bureaucrat and those subject to regulations.

The elected politicians write legislative blueprints to be followed by administrative agencies. Administrative law provides a vehicle for fine-tuning the final rules promulgated by the agencies. Meanwhile the legislator moves to newer territories and negotiates with other special interest groups. All along public interest arguments are used as wrappers for private interest benefits, but at some point in the hierarchy, the wrapper becomes thin. Theories based on efficiency, equity, and public interest fail to explain the way the complex world works.

5. Some final thoughts

This essay has presented a theoretical explanation for the decline and rise of political economy. At the outset, it is curious that what was standard fare 200 years ago would re-emerge and dominate explanations of political

behavior. It is even more curious that the new political economy represents a reunification of thought, as opposed to further intellectual specialization.

There is a competing theory that might be mustered to explain the cycle of thought discussed here. Economic imperialism, the take over of other disciplines by economic logic, is the notion. However, the fact that modern analysts appeal to classical logic, which included a general vision of the world that included politics, philosophy, law, and ethics, suggests a return to roots, not a takeover.

The theoretical explanation for this long cyclical pattern indicates what social scientists know from their training: It is difficult to displace theories in the absence of data. The rapid growth of government intervention in the 1970s provided a large amount of data that could be focused on the study of special interest theories of government. Viewing the political arena as just another dimension of the world where economic agents strive to better themselves, political economy argues simply that political entry will occur whenever the return is high enough. Going even further, the theory argues that politicians will act in ways that expand opportunities for transactions with the economic agents that support their political careers.

The resulting market for political favors is competitive, but it is controlled. There are political constraints that make it possible for politicians to articulate the distribution of favors across agents who always seek more than can be provided. Expansion of the political enterprise follows as government grows and a limited number of elected officials work to redistribute wealth.

If the division of labor argument outlined here is a valid explanation, the theory predicts that new subdisciplines within political economy will emerge. The theory also predicts that any future period of relative government decline will be followed by a contraction in the demand for political economy. However, unlike previous periods, the loss of demand for political economy will be related to the contraction of government, not the inability of the theory to provide insights that explain the behavior of economic agents who seek political control.

References

Ekelund, R.B. and R.F. Hebert, 1983, A history of economic theory and method, 2nd. ed. (McGraw-Hill, New York).

Ekelund, R.B. and R.D. Tollison, 1981, Mercantilism as a rent-seeking society (Texas A&M Press, College Station).

McCormick, R.E., 1984, The strategic use of regulation: A review of the literature, in: R.A. Rogowsky and B. Yandle, eds., The political economy of regulation: Private interests in the regulatory process (Federal Trade Commission, Washington) 13–32.

McCormick, R.E., 1989, A review of the economics of regulation, in: R.E. Meiners and B. Yandle, eds., Regulation in the Reagan era (The Independent Institute, San Francisco) 16–37.

Mitchell, W.C., 1988, Virginia, Rochester, and Bloomington: Twenty-five years of public choice and political science, Public Choice 56, 101–119.

Mitchell, W.C., 1989, The calculus of consent: Enduring contributions to public choice and political science, Public Choice 60, 201–210.

Pigou, A.C., [1920] 1932, Economics of welfare (Macmillan, London).

Say, J.B., [1803] 1966, A treatise on political economy, translated from: the 4th ed. by C.R. Prinsep (Claxton, Remsen and Haffelfinger, Philadelphia, 1880), reprint ed. A.M. Kelley, New York.

Smith, A., [1776] 1937, An inquiry into the nature and causes of the wealth of nations, in: E. Cannan, ed. (Random House, Inc., New York).

Stigler, G.J., 1968, The division of labor is limited to the extent of the market, reprinted in G.J. Stigler, The organization of industry (The University of Chicago Press, Chicago) 129–141.

Wiseman, J., 1990, Principles of political economy: An outlined proposal, illustrated by application to fiscal federalism, Constitutional Political Economy 1, 101–124.

Yandle, B., 1986, The evolution of regulatory activities in the 1970s and 1980s, in: P. Cagen, ed., Contemporary economic problems: 1986 (American Enterprise Institute, Washington) 105–145.

Yandle, B., 1989a, Bootleggers and baptists in the market for regulation, in: J.F. Shogren, ed., The political economy of government regulation (Kluwer, Boston) 29–53.

Yandle, B., 1989b, The political limits of environmental regulation (Quorum Books, Westport, CT).

50

Toward a Theory
OF THE
Rent-Seeking Society

Edited by

JAMES M. BUCHANAN, ROBERT D. TOLLISON,

and GORDON TULLOCK

Texas A&M University Press
COLLEGE STATION

3

The Welfare Costs of Tariffs, Monopolies, and Theft

by

GORDON TULLOCK

IN recent years a considerable number of studies have been published that purport to measure the welfare costs of monopolies and tariffs.[1] The results have uniformly shown very small costs for practices that economists normally deplore. This led R. A. Mundell to comment in 1962 that "unless there is a thorough theoretical re-examination of the validity of the tools upon which these studies are founded . . . someone will inevitably draw the conclusion that economics has ceased to be important."[2] If one can judge from conversations with graduate students, a number of younger economists are in fact drawing the conclusion that tariffs and monopolies are not of much importance. This view is now beginning to appear in the literature. On the basis of these measurements Professor Harvey Leibenstein has argued, "Microeconomic theory focuses on allocative efficiency to the exclusion of other types of efficiencies that, in fact, are much more significant in many instances."[3]

It is my purpose to take the other route suggested by Mundell and to demonstrate that the "tools on which these studies are founded" produce an underestimation of the welfare costs of tariffs and monopo-

[1] These studies are conveniently listed with a useful table of the welfare losses computed in each in Harvey Leibenstein, "Allocative Efficiency vs. 'X-Efficiency'," *American Economic Review* 56 (June, 1966): 392–415.

[2] R. A. Mundell, "Review of L. H. Janssen, *Free Trade, Protection and Customs Union*," *American Economic Review* 52 (June, 1962): 622.

[3] Leibenstein, "Allocative Efficiency," p. 392. In this article Leibenstein consistently uses the phrase *allocative efficiency* to refer solely to the absence of tariffs and monopolies.

lies. The classical economists were not concerning themselves with trifles when they argued against tariffs, and the Department of Justice is not dealing with a miniscule problem in its attacks on monopoly.

Statics

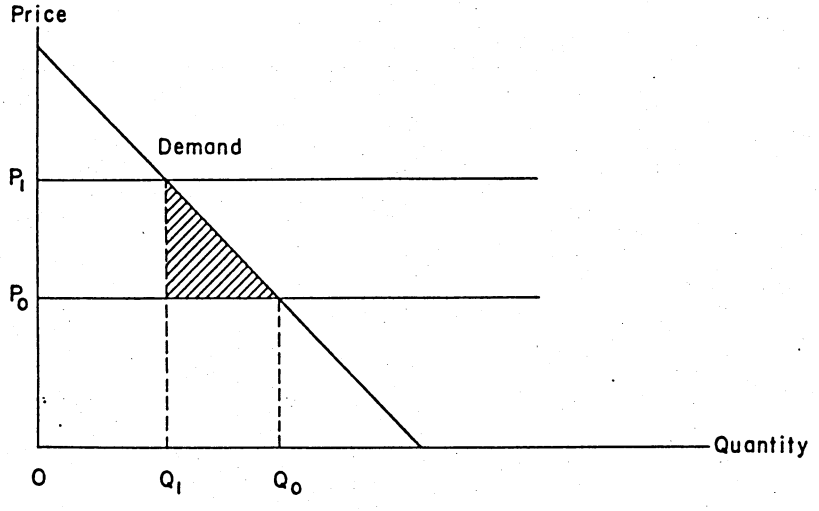
The present method for measuring these costs was pioneered by Professor A. C. Harberger.[4] Let us, therefore, begin with a very simple use of his diagram to analyze a tariff. Figure 3.1 shows a commodity that can be produced domestically at the constant cost of P_1 and imported at P_0. With the given demand and no tariff, Q_0 units will be purchased at a price of P_0. If a prohibitive tariff is imposed, Q_1 units will be bought at a price of P_1. The increase in price, it is argued, is merely a transfer from some members of the community to others, and the only welfare loss is consequently the shaded triangle. The studies purporting to measure the welfare costs of tariffs have simply computed the value of this triangle. From the geometry it is fairly obvious that the amount would normally be small.

There are a considerable number of costs that are ignored by this procedure. As a starter, collection of a tariff involves expenditure on customs inspectors and others who do the actual collection and on coast guards who prevent smuggling. Further, customs brokers are normally hired by the shipper to expedite the movement of their goods through customs.[5] Normally we pay little attention to collection costs because they are small, but in this case they may well be larger than the welfare triangle, which is also small. Thus, by simply adding in collection costs, we significantly increase the "social cost" of the tariff.

[4] A. C. Harberger, "Using the Resources at Hand More Effectively," *American Economic Review* 49 (May, 1959): 134–146. It should be noted that Harberger suggested the method for the measurement of the welfare costs of monopoly, but its extension to cover tariffs was the work of other scholars. The more careful scholars who have measured the welfare costs of tariffs have not all used this very simple application of Harberger's method, but a method such as illustrated in figure 3.2. I have chosen to begin with this method of measurement partly because it simplifies the exposition and partly because this procedure is the conventional wisdom on the matter. (Also see Leibenstein, "Allocative Efficiency.")

[5] Strictly speaking, the customs brokerage should be added on to the tax, thus producing a larger welfare triangle.

FIGURE 3.1



For a more significant criticism of this method of measuring the welfare cost, let us apply the procedure to a standard excise tax instead of a tariff. Assume that figure 3.1 shows a constant supply cost and a declining demand for some commodity in some country. Q_0 units are bought at a price, P_0. Now suppose that a tax is imposed, raising the price to P_1, and reducing sales to Q_1. The welfare cost of this tax is measured by the shaded triangle. But suppose further that the revenues raised by this tax are completely wasted—on building tunnels, for example, that go nowhere. Now the social cost of the total package of tax and wasteful expenditure is the welfare triangle plus the total tax revenue, or the trapezoid bounded by the lines showing cost, the cost-plus-tax, and the demand function. The people buying the product pay more than the cost, but no one benefits from the expenditure.[6] The funds are not transferred because no one benefits from the existence of the tax. The whole economy is poorer not just by the triangle, but by the whole amount of wasted resources.

The tariff involves a similar waste of resources, and consequently

[6]The government action might slightly increase the rents on the resources used to build the tunnel, and thus the owners of specialized resources might benefit slightly, but clearly this is a very trivial effect.

its social cost cannot be measured simply by the welfare triangle. Figure 3.1 can also be used to show the foreign and domestic costs of some type of good and the national demand for it. Since domestic cost is higher than the (delivered) cost of the foreign good, none would be produced domestically in the absence of a tariff. Q_0 units would be imported and consumed at a price shown by P_0. The country now puts on a prohibitive tariff, and the higher-cost domestic production takes over the complete market. Q_1 units are sold at P_1. The welfare triangle has been used to measure the welfare cost of this operation.[7] The argument for this procedure is, essentially, that the higher prices paid by the consumers represent a transfer payment, not a real loss to the economy. But who receives this transfer? The owners of the resources now engaged in inefficiently producing the commodity receive no more than they would have received had the tariff never been introduced and had they been employed in other industries.[8] These resources, however, are being inefficiently utilized, and the rectangle between P_1 and P_0 and bounded by the vertical axis and Q_1 measures the social cost of this waste. Thus the total welfare cost of the tariff is the triangle plus the much larger rectangle to its left.

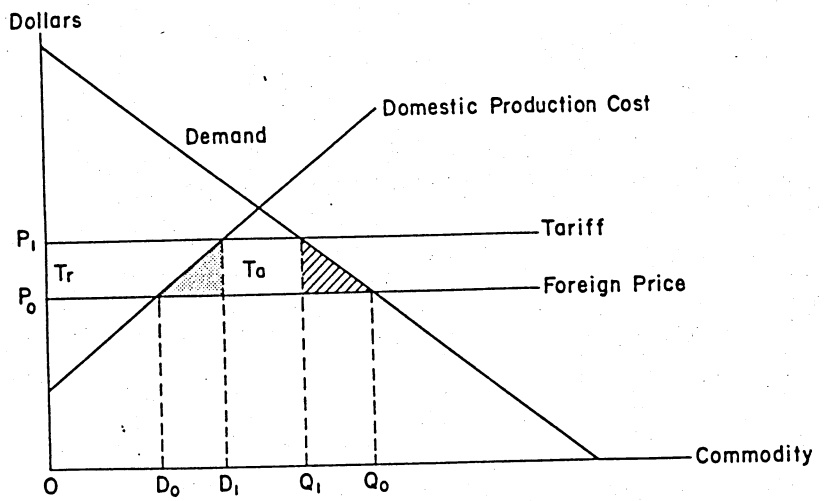
The situation is identical to that which would arise if the government required an established domestic industry to abandon an efficient method of production and adopt an inefficient one. This could be graphed on the same diagram, and it would be generally agreed that the welfare loss would not be just the welfare triangle, but would also include the inefficient use of resources required by the governmental regulation shown in the rectangle to the left of the triangle. Since a tariff shifting production from export goods to import-replacement goods where the country has a comparative disadvantage is, in fact, a governmental requirement that the goods be obtained in an inefficient manner, the cases are identical. The cost of a protective tariff is the triangle plus the difference between domestic cost of production and the price at which the goods could be purchased abroad.

Let us, however, consider the situation in which there is some domestic production before the imposition of a tariff. Figure 3.2 shows

[7] Tibor Scitovsky, *Economic Theory and Western European Integration* (Palo Alto: Stanford University Press, 1958).

[8] There might be sizable but temporary rents to the firstcomers when the industry was first established.

FIGURE 3.2



a commodity, part of which is imported and part produced domestically. The supply elasticity of the commodity from foreign sources is assumed infinite, but domestic production is carried on in conditions of increasing costs. Without the tariff, the price is P_0, domestic producers turn out D_0 units and $Q_0 - D_0$ units are imported to make up the total consumption of Q_0. Suppose, now, that Mr. Gladstone is prime minister and imposes a tariff on imports and an excise tax of the same amount on domestic production. With the new price, P_1, consumers will want only Q_1 units, and the shaded triangle measures the excess burden. Domestic production will remain D_0, but imports will shrink from $Q_0 - D_0$ to $Q_1 - D_0$. The government will receive a tax revenue equivalent to the entire rectangle bounded by the two price lines, the vertical axis, and Q_1.

Let us now change our example by assuming that the domestic excise tax is repealed, so that we have only a protective tariff. Domestic consumption and price would remain the same, but domestic production would expand to D_1 and imports would shrink accordingly. There would be an inefficient use of resources in producing things better imported, represented by the dotted triangle. Governmental revenues would shrink to the rectangle marked T_a and the owners of the resources in the domestic industry would receive an amount of re-

sources equal to the area of the trapezoid T_r.[9] Clearly the social cost of the tariff is not just the shaded triangle, but also the dotted triangle, which shows a net waste of resources in inefficient production.

Dynamics: The Cost of Transfers

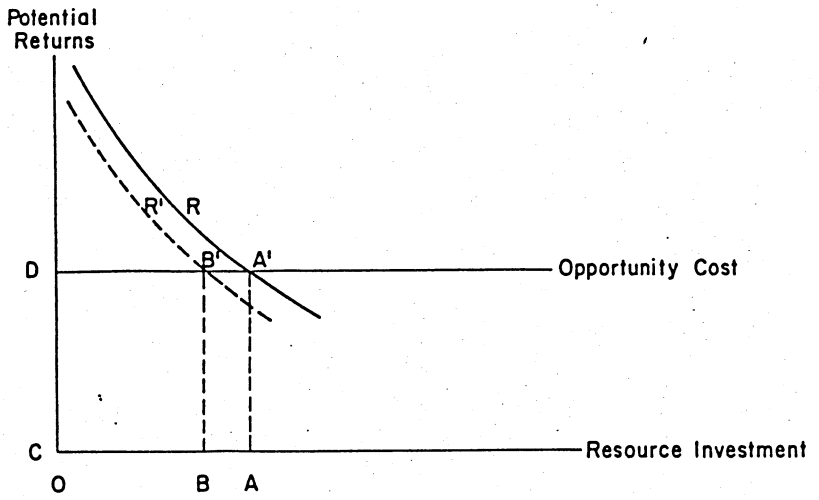
The trapezoid T_r, however, would appear to be a pure transfer and hence not to be included in the computation of the cost of the tariff. Strictly speaking this is so, but if we look at the matter dynamically, there is another social cost involved, and its magnitude is a function of the size of this transfer trapezoid. Generally governments do not impose protective tariffs on their own. They have to be lobbied or pressured into doing so by the expenditure of resources in political activity. One would anticipate that the domestic producers would invest resources in lobbying for the tariff until the marginal return on the last dollar so spent was equal to its likely return, which would produce the transfer. There might also be other interests trying to prevent the transfer and putting resources into influencing the government in the other direction. These expenditures, which may simply offset each other to some extent, are purely wasteful from the standpoint of society as a whole; they are spent not in increasing wealth, but in attempting to transfer or resist transfer of wealth. I can suggest no way of measuring these expenditures, but the potential returns are large, and it would be quite surprising if the investment was not also sizable.

Monopolies involve costs of a somewhat similar nature, and it follows that I will not be able to produce a method to measure their social costs. I will, however, be able to demonstrate that the welfare triangle method greatly underestimates these costs. The argument is customarily explained with the aid of a figure like figure 3.1. The monopolist charges the monopoly price P_1 instead of the cost P_0 for the commodity, and consumption is reduced from Q_0 to Q_1. The welfare triangle is a clear loss to the community, but the rectangle to its left is merely a transfer from the consumers to the owners of the monopoly. We may object to the monopolist's getting rich at the expense of the rest of us, but this is not a reduction in the national product.

In order to demonstrate that this line of reasoning ignores impor-

[9] See J. Wemelsfelder, "The Short-Term Effect of the Lowering of Import Duties in Germany," *Economic Journal* 70 (March, 1960): 94–104.

FIGURE 3.3



tant costs, I should like to take a detour through the economics of theft.[10] Theft, of course, is a pure transfer and therefore might be assumed to have no welfare effects at all. Like a lump-sum tax, it produces no welfare triangle and hence would show a zero social cost if measured by the Harberger method. This would, of course, be incorrect. In spite of the fact that it involves only transfers, the existence of theft has very substantial welfare costs. Our laws against theft do not deal with a trivial problem any more than do our laws against monopoly.

Figure 3.3 shows the situation confronting the potential thief. On the horizontal axis is shown the quantity of effort and capital (burglars' tools, etc.) he might invest in a career of crime. On the vertical axis are shown potential returns. The "opportunity cost" line shows the returns he could get for the same investment of work and material in

[10] The economics of illegal activities is an underdeveloped area, but Harold Demsetz discusses the subject briefly in "The Exchange and Enforcement of Property Rights," *Journal of Law and Economics* 7 (October, 1964) 11–26. J. Randolph Norsworthy's doctoral dissertation, "A Theory of Tax Evasion and Collection" (University of Virginia, 1966), is a more comprehensive examination of one type of illegal activity. Two unpublished items have been circulated among a few scholars: Gary Becker's "A Theory of Government Punishments and Rewards" and my own "Law and Morals," the unfinished manuscript of a book I began four years ago that has languished in draft form for almost all of those four years.

other occupations. It is assumed to be constant. Let us begin by assuming that taking another's property is not illegal. Under these circumstances the returns on various amounts of investment in the activity are shown by line *R*. The potential thief would invest the quantity of resources shown at *A* in theft, the cost to him would be the rectangle *AA'DC*, and his net return on the investment would be the triangular area above *A'D*.

The situation of a person who wished to guard his own assets, who might, of course, be the thief hoping to hold onto his loot, may also be shown on figure 3.3. On the horizontal axis are shown the resources invested in loss-minimizing activities.[11] The cost of each unit of resources put to this use is shown by the horizontal opportunity line, and the savings are on the vertical axis. The line *R* now shows the returns in the form of savings for each unit of "theft prevention." The total amount of resources invested would again be *A*.

The two situations are interrelated by more than the fact that they can be shown on the same diagram. The height of the *R* curve for the thief would depend upon the amount of resources invested by other members of the community in locks and other protections. Similarly, the individual in considering how many locks to buy would find that his *R* curve depended upon the resources being invested in attempts at theft by the rest of the population. When a potential thief invests money, say, in an improved lock pick, the *R* curve for people trying to protect their property moves downward. Similarly, hiring an armed guard to watch one's valuables moves the *R* curve for potential thieves down. Putting a new lock on my door reduces the chance that I will be robbed, but whether the gain will be worth the cost will depend upon the effort the thieves are willing to put into getting in. Over time the interaction between the investment in locks, the payoff on lock picks, and the investment in nitroglycerine and safes would come to equilibrium.

This equilibrium, however, would be extremely costly to the society in spite of the fact that the activity of theft involves only transfers. The cost to society would be the investments of capital and labor in the activity of theft and in protection against theft. If we consider

[11] The word *activities* may be misleading. One way of minimizing loss by theft is to have little or nothing to steal. In a world in which theft was legal, we could expect this fact to lead to a reduction in productive activities and a great expansion in leisure.

figure 3.3 as representing the entire society instead of individuals, then the social costs would be the area covered by the rectangle *AA'DC*. Transfers themselves cost society nothing, but for the people engaging in them they are just like any other activity, and this means that large resources may be invested in attempting to make or prevent transfers. These largely offsetting commitments of resources are totally wasted from the standpoint of society as a whole.

This lesson has been learned by almost all societies that have adopted a collective method of reducing this sort of income transfer. This collective procedure, laws against theft and police and courts to enforce them, can also be shown on figure 3.3. On the horizontal axis we now have resources invested by police and courts, with their opportunity cost shown as a horizontal line. The "protection" given by each unit of resources invested in these activities is shown by the *R* line. The society would purchase *A* amount of protective services, and the total cost would be the usual rectangle. The effect of this would be to reduce the expected returns on theft and the savings to be made by private investment in locks and safes. The new returns are shown by *R'* on figure 3.3, and there is a corresponding reduction in the resources invested in each of these fields to *B'*. Whether the establishment of a police force is wise or not depends upon an essentially technological question. If police activities are, for a range, more efficient than private provision of protection, then the *R* line will have the shape shown, and the police and court rectangle will have an area smaller than the sum of the two "savings" rectangles for theft and locks.[12] This is, of course, what we normally find in the real world.

Note, however, that we do not carry investment in police protection to the extent that it totally replaces private protection expenditures. Clearly it is more efficient to have some protective expenditures by the owners of property. Automobiles are equipped with locks and keys, presumably because the expansion of the police force that could be paid for from the cost of leaving them off would be less effective in preventing theft than locks and keys are.[13] The total social cost of theft

[12] It may be suggested that society should not be interested in saving the resources of thieves and, hence, that the value of the protection afforded by the police should be measured by the lock rectangle only. This, however, would be correct only to the extent that the resources would not be reallocated to socially acceptable production.

[13] James Buchanan and Gordon Tullock, "Public and Private Interaction under Re-

is the sum of the efforts invested in the activity of theft, private protection against theft, and the public investment in police protection. The theft itself is a pure transfer and has no welfare cost, but the existence of theft as a potential activity results in very substantial diversion of resources to fields where they essentially offset each other and produce no positive product. The problem with income transfers is not that they directly inflict welfare losses, but that they lead people to employ resources in attempting to obtain or prevent such transfers. A successful bank robbery will inspire potential thieves to greater efforts, lead to the installation of improved protective equipment in other banks, and perhaps result in the hiring of additional policemen. These are its social costs, and they can be very sizable.

But this has been a detour through the criminal law, and our major subject is monopoly. To return to figure 3.1, the rectangle to the left of the welfare triangle is the income transfer that a successful monopolist can extort from the customers. Surely we should expect that with so large a prize dangling before us, potential monopolists would be willing to invest large resources in the activity of monopolizing. In fact the investment that could be profitably made in forming a monopoly would be larger than this rectangle, since it represents merely the income transfer. The capital value, properly discounted for risk, would be worth much more. Entrepreneurs should be willing to invest resources in attempts to form a monopoly until the marginal cost equals the properly discounted return.[14] The potential customers would also be interested in preventing the transfer and should be willing to make large investments to that end. Once the monopoly is formed, continual efforts either to break the monopoly or to muscle into it would be predictable. Here again considerable resources might be invested. The holders of the monopoly, on the other hand, would be willing to put quite sizable sums into the defense of their power to receive these transfers.

As a successful theft will stimulate other thieves to greater indus-

---

ciprocal Externality," in *The Public Economy of Urban Communities*, ed. Julius Margolis (Baltimore: Johns Hopkins University Press, 1964), pp. 52–73.

[14]The margin here is a rather unusual one. Additional units of resources invested in attempting to get a monopoly do not increase the value of the potential monopoly, but only the likelihood of getting it. Thus they change the discount rate rather than the payoff.

try and require greater investment in protective measures, so each successful establishment of a monopoly or creation of a tariff will stimulate greater diversion of resources to attempts to organize further transfers of income. In Gladstone's England few resources were put into attempts to get favorable tariff treatment. In the United States today large and well-financed lobbies exist for this purpose. The welfare cost in the first case was very low; in the second it must be quite sizable. An efficient police force reduces the resources put into the activity of theft, and free trade or an active antitrust policy will reduce the resources invested in lobbying or attempting to organize monopolies.

The problem of identifying and measuring these resources is a difficult one, partly because the activity of monopolizing is illegal. The budget of the antitrust division and the large legal staffs maintained by companies in danger of prosecution would be clear examples of the social cost of monopoly, but presumably they are only a small part of the total. That very scarce resource, skilled management, may be invested to considerable extent in attempting to build, break, or muscle into a monopoly. Lengthy negotiations may be in real terms very expensive, but we have no measure of their cost. Similarly, a physical plant may be designed not for maximum efficiency in direct production, but for its threat potential. Again, no measure is possible. As a further problem, probably much of the cost of monopoly is spread through companies that do not have a monopoly but have gambled resources on the hopes of one. The cost of a football pool is not measured by the cost of the winner's ticket, but by the cost of all tickets.[15] Similarly the total costs of monopoly should be measured in terms of the efforts to get a monopoly by the unsuccessful as well as the successful. Surely more American businessmen know that the odds are against their establishing a paying monopoly, and they therefore discount the potential gain when investing resources in attempting to get one. The successful monopolist finds that his gamble has paid off, and the unsuccessful "bettor" in this particular lottery will lose, but the resources put into the "pool" would be hard to find by economic techniques. But regardless of the measurement problem, it is clear that the resources put into monopolization and defense against monopolization would be a func-

[15] This helpful analogy was suggested to me by Dr. William Niskanen.

tion of the size of the prospective transfer. Since this would be normally large, we can expect that this particular socially wasteful type of "investment" would also be large. The welfare triangle method of measurement ignores this important cost, and hence greatly understates the welfare loss of monopoly.

63

The Political Economy of Deregulation

Interest Groups in the Regulatory Process

Roger G. Noll and Bruce M. Owen

American Enterprise Institute for Public Policy Research
Washington and London

64

1

Introduction: The Agenda for Deregulation

Regulation is a peculiarly American institution, though all nations use political and legal processes to constrain the economic activities of their citizens. The most common method of implementing such policies in other countries is to give government officials great direct authority. Many governments nationalize important industries or set up a controlling bureaucracy that has far more power than the typical American regulatory agency.

American regulation is a reflection of the democratic and egalitarian principles held by the Founding Fathers, especially their fear of centralized government power. Its organizing principle is that decisions should be based upon objective analysis in a process that allows people who are likely to be affected by the decision to have their views heard and considered. Elaborate rules regarding rights of participation, the evidence pertaining to a decision, and the statutory basis for a policy action have developed to serve this principle.

Like many legal processes, regulation in America seeks to base decisions on objective facts, principles of equity, and the public interest, but does so in a decision-making environment that is populated primarily by advocates of particular economic interests. For the most part, participants in the regulatory process are motivated by their economic stakes in the decision, and as a result their behavior in the process—the kinds of evidence and arguments that they will produce—is quite predictable.[1]

During the decade of the 1970s, numerous federal regulatory policies were reexamined with an eye toward major reform and often complete deregulation (see table 1–1). The passage late in 1982 of legislation to deregulate intercity bus service is but the most recent in a series of policy moves to free competitive forces from federal supervision. Stock brokers' fees, railroads, trucks, airlines, petroleum, cable television, radio stations, air cargo service, savings and loan institutions, banks, securities issuers, and other industries have, to varying

degrees, been deregulated since 1974. These industries obviously have different structures and have had different sorts of regulation

TABLE 1–1
MAJOR DEREGULATORY INITIATIVES, 1971–1982

Year	*Deregulatory Initiative*
1971	Transportation deregulation (proposed)
	FCC: Specialized common carrier decision
1972	FCC: Domestic satellite open skies policy
1975	SEC: Abolition of fixed brokerage fees
	Trucking deregulation (proposed)
	Banking deregulation (proposed)
	Airline deregulation (proposed)
1976	Railroad Revitalization and Reform Act
1977	Air Cargo Deregulation Act
1978	Airline Deregulation Act
	Natural Gas Policy Act
	OSHA: Standards revocation
	EPA: Emissions trading policy
1979	FCC: Deregulation of satellite earth stations
1980	Motor Carrier Reform Act
	Household Goods Transportation Act
	Staggers Rail Act
	Depository Institutions Deregulation and Monetary Control Act
	International Air Transportation Competition Act
	FCC: Deregulation of cable television
	FCC: Deregulation of customer premises equipment and enhanced services
1981	Decontrol of Crude Oil and Refined Petroleum Products (Executive Order)
	Federal Reserve Board: Truth in lending simplification
	NHTSA: Auto industry regulatory relief package
	FCC: Deregulation of radio
1982	Bus Regulatory Reform Act
	Garn–St Germain Depository Institution Act
	FCC: Deregulation of resale and transponders

NOTE: FCC = Federal Communications Commission; SEC = Securities and Exchange Commission; OSHA = Occupational Safety and Health Administration; EPA = Environmental Protection Agency; NHTSA = National Highway Traffic Safety Administration.

applied to them, and so required separate analyses to sustain the case for reform. Nevertheless, the deregulation debate in each case followed predictable lines.

The deregulation debates continue. One major issue is whether to repeal certain regulations of the Federal Communications Commission (FCC) that constrain the business operations of the three national television broadcast networks.[2] It was this issue that provided the original motivation for this study. The FCC rules were adopted a decade ago on the basis of a record compiled largely in the late 1950s and early 1960s. The rules restrict the ability of the networks to acquire certain valuable rights in the television programs that they air. The networks are on one side of this debate. On the other side are the major Hollywood studios that produce and license television programs to the networks and distribute programs in syndication.

Because deregulation of the television networks involves the media, the political debate has been bathed in publicity. Even though the interests and issues may be no less parochial, network deregulation has received more media attention than many of the other important regulatory policy debates, such as those involving the Clean Air Act or natural gas prices. Yet the debate is likely to be carried out in a vacuum, as if this were the first rather than merely the latest proposal to deregulate an important industry.

It is our thesis that useful insights can be gained from a general understanding of the political economy of deregulation and from taking a broad view of the history of regulation itself. Regulatory regimes, whether in airlines, banking, or broadcasting, have much in common. Each tends to create (and to destroy) groups with special economic interests. The views and arguments of these groups on the question of deregulation are rooted in their own interests. The debate about network deregulation, and other future deregulation debates, will be more enlightened if the positions of the parties and their arguments are not viewed in isolation, but are instead seen as part of a long history of regulatory policy, broadly defined.

The discussion of the regulatory process in the following seven chapters is based largely on this interest group "model." An understanding of the model and of its value in predicting behavior is useful, perhaps most useful to regulators in finding ways to rise above its predictions. The last chapter of this book introduces these more normative concerns and shows that while the interest group model is descriptive of forces at work in the policy process, it need not—indeed, should not—determine the outcome of the regulatory process.

The problem of regulators is to identify a general public interest in a process that is populated largely by interest groups pursuing narrow

aims. We recognize that the public interest is an elusive concept, and we do not propose to offer a comprehensive definition of it. Nevertheless, an important aspect of the public interest is to advance the interests of members of society acting in their roles as consumers, and to do so in a manner that promotes economic efficiency. Every citizen has numerous interests, according to occupation, industry of employment, residency, nature of principal investments, political orientation, important social relationships, and pattern of consumption expenditures. A principal insight of the interest group theory of regulatory processes is that some aspects of a citizen's interests are more likely to be effectively represented than others. The task of the regulator is to work out, for each case, the biases that are likely to emerge from the patterns of participation in the regulatory process.

Organization of This Study

This first chapter provides the reader with a brief survey of some current deregulation controversies. We include a summary description of the issues and the stakes the various interest groups have in them. The idea is not in this brief space to analyze in any depth the particular pros and cons of each deregulatory proposal. Our intent is instead to motivate the discussion that follows, in chapter 2, of the interactive role of interest groups and regulators and, in chapter 3, of the types of arguments used by interest groups to defend their positions.

In chapter 2 we explore the political and economic origins of interest groups, their recognition in the federal system by the Founding Fathers, and the relationship between economic regulation and the representation of interest groups in the policy process. Numerous examples from a variety of regulatory arenas are used to illustrate our general analysis of the role of interest groups in regulatory policy.

Chapter 3 provides a survey of the arguments that have been used by opponents of deregulation. Regardless of the regulatory policy in question, groups that have a vested interest in continued regulation make predictable arguments about cross-subsidization and predictions of destructive competition, excessive risk, and harm to consumers. To the extent possible, we comment on the evidence that has become available in these industries since deregulation that bears on these predictions. Such arguments and predictions will inevitably be made in future deregulation debates. Arguments arising from self-interest are not necessarily invalid, but an assessment of the validity of these arguments in other industries can shed light on how to assess their validity in current deregulation debates.

Chapters 4 through 8 contain detailed case studies of particular deregulation episodes, both completed and ongoing. In chapter 4, Andrew Carron examines the effort to reform or to dismantle the constraints that federal regulators have placed on the banking system. Robert Crandall reviews, in chapter 5, the fascinating convergence of environmental and coal-producing interests that led to the 1977 Clean Air Act Amendments. In chapter 6, Joseph Kalt explores the special interest groups that opposed deregulation of the energy industry in the aftermath of the 1973 oil embargo. Finally, in chapters 7 and 8, two highly distinguished economists and former regulators review the actual experience with deregulation in two traditional fields: Marcus Alexis, former commissioner of the Interstate Commerce Commission (ICC), examines deregulation of surface transportation, and Alfred Kahn, former chairman of the Civil Aeronautics Board (CAB), writes about deregulation of the airlines.

Chapter 9 sums up the lessons learned and attempts to generalize beyond the interest group model. This is necessary because it would otherwise be difficult to say why, in at least some cases, the interest group model does not fully explain events—why, for example, airline or trucking deregulation eventually took place despite interest group pressures. We do not understand this political process as well as we understand the essentially economic model of interest group formation. What we do know is that any policy maker considering deregulation must understand the economic basis of interest groups, or else risk giving undue weight to a number of illegitimate arguments.

The balance of this first chapter is devoted to a survey of some major current regulatory reform controversies. There are several reasons for doing this. The first is simply to motivate the discussion in chapter 2 of the interest group model of regulatory policy making. The second is to show how this analysis can be relevant to current policy issues, most of which remain undecided. The third is to begin to demonstrate that apparently unrelated deregulation debates, often conducted in isolation from each other, in fact have much in common. We begin with the FCC's network financial interest and syndication rules.

Network Television Deregulation

The network syndication and financial interest rules impose restrictions on the ability of the three broadcast networks to acquire certain financial interests in the programs that they buy for network exhibition and on their ability to participate in the syndication business.[3] The rules apply only to broadcast networks (that is, ABC, CBS, and NBC).

They do not apply to their competitors, such as the major studios in their roles either as financiers of television program production or as syndicators, or to cable networks, or to potential competitors, such as COMSAT's recently authorized direct broadcast satellite system. The rules are opposed by the broadcast networks, for they restrict the networks' freedom of action. The rules are generally supported by the major movie studios, who supply most of the series programming purchased by the networks and who are given a favored market position by the restrictions these rules place on important former competitors. Although not prominent among the original proponents of the rules, the major Hollywood studios soon became the principal opponents of deregulation. The rules are also supported by program syndicators who are sheltered from the potential competition of the networks as syndicators.

While the studios and the networks are the central antagonists, there are other important groups with less well-defined, homogeneous, or articulated interests, such as independent producers of television shows, television stations of various kinds, and advertisers. Finally, and most important, there are viewers. The effect on the viewing public presumably ought to constitute the basis for any decision to deregulate. The FCC, as the guardian of the public's interest, must weigh the arguments of interest groups as to the positive effects—if any—of regulation on viewers, while remaining indifferent to arguments concerning the effects on the interest groups themselves.

We cannot in this brief space attempt to analyze in any depth the economic effects of the rules that the networks want repealed. Such an analysis has been provided already by the special network inquiry staff of the FCC.[4] But in order to connect the issues in network regulation to the interest group analysis that forms the heart of this book, we will provide a summary of the economic analysis of the effects of the FCC rules.

The production of television programs is risky because the popularity of a program, and hence its effectiveness in generating advertising revenues, is highly unpredictable. One way to cope with risks is to pool them—that is, to hold several risky assets in the expectation that good fortune on some will balance bad luck on others. This is the theory behind investing in a portfolio of different investments rather than putting all of one's wealth in a single asset.

Networks are in a position to pool risks by investing in a portfolio of programs. Before the syndication and financial interest rules, any producer with a creative idea could seek financing from the networks or the movie studios and in exchange sell syndication rights and finan-

cial interests. Producers, especially small ones, typically want to sell these rights. With one or only a few programs over which to spread the risks of failure, small producers are unable efficiently to bear such risks. Therefore, one would expect a mutually advantageous sale of the rights having the most uncertain value to larger entities (such as networks) that are in a better position to bear risks. Because of the rules, small producers cannot engage in such agreements with the networks.

The imposition of the rules left small program producers unable to use the networks as a means of reducing their risks. As a consequence, motion picture studios soon became the primary source of funds for financing the production of television series. The studios began to negotiate financial arrangements with program producers that paralleled the old program producer contracts with networks.

The increased role of motion picture studios in spreading the risks of program production has several important economic ramifications. First, it illustrates that the basic form of the old agreements with networks to provide financing in return for a partial sale of rights has continued to play an important role in the industry, but now with different sources of financing. Obviously these arrangements were not purely a manifestation of some special market power that was possessed by networks.

Second, the rules that required the elimination of the networks from their risk-spreading role may remove one of the cheapest methods of spreading risks, by introducing an economically unnecessary middleman. To the extent that networks, because they acquire many programs or are in possession of useful information about potential program success, are able to bear the risks of program failure more efficiently than the movie studios, one would expect the costs of program production to be increased by the rule. Moreover, three-way negotiations about program production and distribution are likely to be more expensive than bilateral negotiations.

Third, the elimination of the networks as sources of risk capital for independent producers may reduce program financing competition for the established motion picture studios. Other things equal, this could produce lower revenues to program talent for financial interest and syndication rights because three important bidders for those rights have been eliminated from competition.

Fourth, by forcing new entrants to the programming industry to obtain financing through a shorter list of established program producers, the rules could be expected to raise entry barriers into program production. By eliminating more efficient bearers of risk (networks),

the rules will raise the cost of capital to new or potential entrants in program production. All current producers could be expected to like this aspect of the rules.

Fifth, because network entertainment series programming is a riskier undertaking than many other types of programming, the rules will tend to increase concentration among suppliers of prime-time series programs. The major movie studios may well control a larger fraction of this business than they did before the rules, because they are better able to bear risks than their smaller competitors.

Sixth, the rules reduce the number of competitors, or potential competitors, in syndication markets. Barring the networks from this activity has the effect of making the market potentially less efficient.

The rules have additional effects, unrelated to risk bearing, on the positions of networks and their competitors. The rules effectively partition the entertainment business into segments. Competition across segments is hampered by the rules. The rules restrain the three major networks from acquiring certain valuable rights. Therefore, cable networks, such as Home Box Office (HBO), and other potential entrants into over-the-air network broadcasting, such as syndicators, perceive that the barriers to entry into network broadcasting have been increased by the rules. If they were to enter over-the-air network broadcasting, they could reasonably expect the FCC to expand the rules to cover them. Thus, entry into this kind of broadcasting would force them to give up valuable rights that they currently find desirable to acquire.

The rules clearly place the networks at a competitive disadvantage to other forms of program distribution in today's rapidly changing home video marketplace. The rules reduce the ability of networks to compete for programming with other program distributors. A considerable amount of jockeying among movie houses, networks, and other programmers and distributors is currently occurring as new media develop and new distribution patterns for programming are formed. It would not be surprising to find that the firms competing with networks to distribute programs to new media are comfortable with the constraints the rules place on the networks' ability to compete.

The movie studios and the networks are taking fairly predictable positions on elimination of the financial interest and syndication rules. In many different areas, the ability of networks to compete with movie studios, cable networks, and syndicators is severely hampered by the rules, naturally leading these groups to favor the rules and the networks to oppose them.

Two general themes of this book are that regulation itself can

create or reinforce interest groups and that it is possible to make predictions about which groups affected by regulation are likely or unlikely to coalesce into effective interest groups. Both these themes are illustrated by the role of interest groups in the debate over the network restrictions. They are equally well illustrated by the fate of proposals to repeal the fifty-five mile-per-hour speed limit. While repeal of this regulation has often been proposed in Congress, the issue has not attracted much support, has been strongly opposed by some, and has never progressed far enough to be voted on.

The Fifty-five Mile-per-Hour Speed Limit

The major benefits claimed for the lower speed limit are energy savings and reduced highway fatalities. The major costs imposed by the rule are increases in travel time and truck freight transportation costs. Lower speeds mean that more drivers and trucks are needed to deliver a given amount of material. A small academic literature tends to indicate that the costs of the rule outweigh the benefits.[5] Whatever the net benefits, the costs are substantial. There is, therefore, an interesting question: why have those who bear the costs of the lower limit not organized a campaign for repeal? Why have those who would be hurt by repeal been so effective in promoting their views?

To answer these questions it is necessary first to ascertain who benefits and who loses from the lower speed limit. The costs of the lower speed limit are spread widely. All consumers bear some costs, in the form of inconvenience and higher prices for final commodities because of higher transportation costs. Companies that use trucks to ship their products are also presumably hurt by the lower speed limit. Their costs are increased. But for most consumers and companies the effect will be small, since transportation costs are usually a small percentage of total costs. In any case, the cost increases apply more or less equally to all competitors in an industry and thereby significantly disadvantage only a few marginal firms.

Some individuals may feel that they benefit from the law because of the slight reduction in the chance of being involved in an accident. We do not know whether this view is generally shared by the citizenry. Many drivers routinely choose to violate the speed limit, however, and are obviously and noisily upset if their progress is delayed by a law-abiding citizen ahead of them on the highway. Such citizens presumably favor repeal of the rule at least as it applies to themselves.

But the most influential interest groups are those who have direct and perhaps substantial economic interests in the perpetuation of the

lower speed limit. One strong source of organized support for the fifty-five mile-per-hour limit is the trucking industry. Trucking companies and drivers would certainly suffer economic losses if the limit were repealed, for repeal would immediately increase the productivity of the trucking industry. Each truck and driver could make more trips in a given period of time, so that fewer trucks and drivers could transport the same amount of freight. Price reductions could presumably restore full capacity utilization, but the industry might rationally fear that the net effect would be lower incomes for truckers and trucking companies. The specters of increased unemployment among truck drivers, more excess capacity among operators, and subsequent reductions in prices, wages, and profits provide each group with a strong economic interest for opposing repeal.

The interest of truck drivers and trucking company operators in the fifty-five mile-per-hour limit was actually created by the imposition of the limit. Prior to enactment the drivers and operators might have supported the lower limit because they could expect transitional gains (higher prices and wages) before new capacity was added to the industry. They might also have reasoned, prior to passage, that such gains would be transitory and would lead to higher costs. Moreover, because trucking was then subject to price regulation, truckers might expect that some time would pass before prices were allowed to rise to reflect the higher costs. All of these reasons led to neutrality or opposition to the lower limit prior to its passage. Now that the limit is in place and the industry has adjusted to it there can be no question about the economic effect of repeal on the industry. It will be hurt. This is another excellent illustration of the general point that regulation can create interests in the perpetuation of regulation even where none existed before. Thus, a rule once in place may have far stronger support than it originally commanded, even though the rule might impose greater costs than benefits on the public. A corollary of this point is that one regulatory constraint can create a constituency that clamors for further regulation. This is quite vividly illustrated by the battle taking place within the real estate services industries over repeal of section 8 of the Real Estate Settlement Practices Act of 1974, which outlaws certain payments made by title insurance companies to real estate brokers.

Rebates in Real Estate Services

President Reagan proposed in 1982 that Congress repeal section 8 of the Real Estate Settlement Practices Act of 1974, which prohibits the

payment of kickbacks and rebates by title insurance companies and other providers of ancillary real estate services. The debate over this proposal has been long and heated, with many different factions within the real estate services industry.

In many states insurance regulation (together with the McCarran-Ferguson federal antitrust exemption) allows title insurance companies to fix prices. Title companies tended, however, to compete away the gains from their price fixing in payments to the real estate agents who referred business to them. Section 8 of the act makes such payments a federal crime, thus helping the title companies to maintain their price-fixing profits. If section 8 is repealed, at least some of the excess profits will be passed on to consumers as a result of competition among real estate brokers.[6]

The rules against rebates have created an interest group of title companies owned by real estate brokers that would not exist but for the rebate prohibition. In order to avoid the prohibition, brokers have opened title insurance agencies in their offices. This vertical integration allows the illegal rebates to be paid in the form of legal commissions and dividends. That they did not own such agencies before the Real Estate Settlement Practices Act suggests that brokers would prefer not to be title insurance agents. Indeed, the interest of real estate brokers in the whole subject would be minimal were it not for state insurance regulations that give rise to the excess profits and hence to the incentive to compete through rebates. Many title insurance firms have pressed for regulations or legislation that would outlaw these vertical arrangements. Thus, we find one interest group created by the system of regulation and another pressing to retain the benefits that it obtains from regulation by seeking additional regulation. Regulation itself can give rise to demands for more regulation, either by adversely affected groups or by beneficiaries of regulation who argue that the original rules did not go far enough.

The analysis of the interest groups created by state insurance regulation is relatively simple. But there are some deregulation debates in which the problem of identifying interests is complex. We turn now to the area of new drug licensing, where these analytic problems are more difficult. The debate over drug patent life extension is another example of new regulations being proposed in response to the distortions caused by old ones.

Regulatory Reform in New Drug Applications

New drugs to be marketed in the United States are subjected to a

period of rigorous testing, subject to Food and Drug Administration (FDA) supervision, followed by up to two years of official government review of the test results. Consequently, the time between inventing and marketing a new drug in America can easily exceed a decade. While even drug manufacturers would not want to cut significantly the testing of new drugs, they argue that delays in the FDA approval process raise the cost of new drugs, threaten business predictability, and cause many needed drugs to be available overseas long before they are approved for use here.

Two reforms have been proposed to answer this concern. Patent extension bills in the House and Senate would extend patent life for new drugs to compensate for the delays introduced by the FDA approval process. Meanwhile, both the Carter and the Reagan administrations have undertaken extensive efforts to streamline that process. This has already resulted in an increase in the number of new drugs approved, as well as a decline in the number of enforcement actions taken by the FDA in the marketplace.[7]

An interesting feature of the debate over patent extension is that both sides claim to be placing greater reliance on market incentives. Because the FDA approval process amounts to a regulatory constraint, lengthening the patent period would permit drug companies to profit from inventions to the same extent enjoyed by firms in nonregulatory environments. The arguments of the drug companies here are analogous to those of the broadcast networks. Both would like to have an opportunity to exploit their business skills under the same rules that other companies enjoy. The pharmaceutical firms claim, moreover, that the de facto shortening of the patent period by regulatory review creates a disincentive for research in an industry where greater investment in research is in the public interest.

The interests of the pioneer pharmaceutical firms in the debate over patent life extension are fairly clear-cut. Similarly clear-cut are the interests of drug manufacturers that specialize in marketing low-priced "generic" drugs after patents have expired. They oppose extension of patent life. Both groups, of course, favor decreased delays in the FDA approval process.

The interests of drug retailers, health professionals, and consumers are decidedly more complex. In fact, both interests and views vary substantially within these groups, making them difficult to organize effectively.

Consumers, for example, have interests that vary according to their age and state of health. Older consumers with health problems have relatively little to gain from future long-term research, and more

to gain from lower prices for existing drugs. They also have less to fear from harmful side effects of drugs that have long latency periods (as with most carcinogens). Younger or healthier consumers have relatively more to gain from encouraging future research. They also are more concerned about being safe from long-term carcinogens and mutagens. Similarly, drug retailers' stands on patent extension vary according to the mix of generics and brand-name pharmaceutials that each distributes. Small druggists, concerned with liability problems, tend to support the pioneer drug companies. The position of health professionals is also diverse. Different physicians and pharmacists may evaluate the trade-off between innovation and retail price competition differently. Both the AMA and the American Pharmacists' Association have nevertheless come out quietly in favor of patent extension.

Among the interest groups not represented at all in the debate are pioneer companies that would exist if the FDA delays did not reduce the economic return from new research.

Consumer advocate groups have been active in opposing patent life extension. They argue that an important interest of consumers lies in lower current drug prices. They also point out that there are other industries in which patents are not regarded as effective for the full term of a patent, such as the semiconductor and chemical industries, where greater competition has, if anything, stimulated innovation. Finally, they supported the recently enacted bill to promote research on "orphan drugs" as a more effective, targeted strategy to enhance the inventory of drugs. This program provides tax subsidies for developing new drugs to treat illnesses that are not widespread enough to provide sufficient profit incentives for drug company research.

Patent life extension is but one possible response to the problem of FDA red tape. The FDA's review of procedures for the approval of new drug applications (NDAs) and the regulatory reforms proposed by former Commissioner of Food and Drugs Donald Kennedy attack the problem more directly.

Recently both the pioneer and the generic drug manufacturers have supported steps to speed all NDA proceedings. There is a broad industry-wide consensus on reducing the number of points along the path to introduction of a drug at which regulatory review is required, and on imposing deadlines on the FDA, which, when reached without FDA action, give automatic approval to proceed to the next stage of the testing and marketing process. But some consumer groups oppose speeding up all NDA proceedings. They applaud the record of safety in America and question whether safeguards will continue to be ade-

quate. They have, however, favored speedier processes for generics and breakthrough drugs.

The drug debate shows that there are groups whose interests are difficult to reflect in the policy debate because of their heterogeneity, because of the difficulty of identifying their members, or because of the remote stakes of each member in the outcome. In the discussion of auto emissions standards that follows, we will see another example of groups created by regulation itself, and more important, the nonrepresentation of one of the most important groups affected by the standards—automobile owners.

Automobile Emissions Standards

Reform of the standards and procedures for regulating pollution from automobiles continues to be a highly controversial area of public policy. Several issues are currently being hotly contested, encouraged in part by the debate over revisions in the nation's flagship legislation in environmental policy, the Clean Air Act.

The desirability of a uniform national emissions standard for automobiles is in dispute. Meteorological conditions, geography, altitude, and traffic densities all affect the relationship between auto emissions and air pollution from region to region. One proposal is a "two-car" strategy: to have one automobile emissions standard in regions with the most severe air pollution problems and another for the rest of the country.

The argument in favor of allowing auto emissions standards to vary from region to region is that it would save costs in areas with little or no air pollution problem. The argument against it is that it would raise the cost of enforcement. Each region that had tough emissions requirements would have to implement a system to prevent people from importing automobiles from regions with relatively lax standards.

Another controversy is whether the standard should be the same for all models or whether it should be set as a fleet-wide average. Emissions are more cheaply controlled from less powerful engines, so the total costs of achieving any given emissions ceiling would be lower if emissions reductions were made mostly from small cars.

A third issue is the appropriate balance between auto emissions controls and controls on stationary sources of pollution, such as power plants. The current practice is to divide responsibility for standards among federal, state, and local authorities according to the nature of the source. Auto emissions standards are primarily a federal responsibility, whereas the majority of standards for stationary sources are set

by local authorities. This fragmented responsibility means that relatively little attention has been given to the problem of determining the most efficient division of abatement efforts between stationary and mobile sources. Several studies have argued that total abatement costs could be reduced by relaxing auto emissions standards and increasing the strictness of standards for emissions from stationary sources.[8]

Fourth, in recent years the Environmental Protection Agency has pushed state and local authorities to adopt mandatory vehicle inspection and maintenance programs. Evidence is accumulating that emissions control systems do not operate effectively over the entire 50,000-mile life that is required by current standards and guaranteed by manufacturer warranties. The primary focus of the debate is over the proper placement of responsibility for this problem: on drivers, on vehicle manufacturers, or on manufacturers of emissions control systems. Mandatory inspection and maintenance programs place primary responsibility on the automobile owner, who will have to pay for the inspection and for any repairs (other than replacement of an emissions control device that is still under warranty). The primary responsibility for designing an inspection program falls on state and local governments, much as they now are responsible for stationary source inspection programs.

For a mandatory inspection program to be within reasonable costs to motorists, inspection of the vehicles must be simple. Comprehensive vehicle inspection requires placing a vehicle in an enclosed chamber and running it over a driving cycle for a substantial period of time. Obviously, such a comprehensive approach makes sense only for a statistical sample of vehicles, not for the entire fleet. It therefore places the primary responsibility for meeting standards on automobile and emissions control manufacturers. This has implications with respect to income distribution and political acceptability as well as efficiency.

A final controversy surrounding the automobile is the promotion of alternative fuels. The two principal contenders are diesel fuel and methanol. Diesel fuels permit automobiles to use less energy per mile of operation. Historically, diesel fuel has also been somewhat cheaper than gasoline; this probably would not continue, however, if demand for diesel caused it to become the principal determinant of the output mix in oil refining. Diesel also produces a very different combination of emissions: it would generally reduce the emissions of components of photochemical smog, but would significantly increase emissions of particulate matter, a possible carcinogen.

Although it is more expensive than gasoline, methanol can be produced domestically from grains and other vegetable products. It is

an especially clean fuel. Automobiles powered by methanol produce one bad pollutant—formaldehyde—which can be controlled relatively cheaply. Unfortunately, at the federal level the virtues of a methanol strategy have been debated primarily in the context of energy policy. Environmental implications have received secondary attention.

Recent debate has focused increasingly on whether the emissions ceiling for autos ought to be relaxed and whether it ought to be applied uniformly everywhere. This battle involves some major economic interests, each of which takes a predictable stand.

The automobile manufacturers support a relaxation in emissions standards that, they claim, will reduce the cost of an automobile by $300. In a normal year in which 10 million new passenger cars are sold, this amounts to a cost saving of $3 billion.[9] Automobile manufacturers obviously hope that a $300 cost reduction will yield more profits, and perhaps, if they reduce prices, more sales. Moreover, because emissions are more costly to control from large cars, they may hope to narrow the price difference between large and small cars and thereby recapture some of the market lost to foreign manufacturers.

The Manufacturers of Emissions Controls Association (MECA), a trade association representing the sellers of emissions control equipment, has lobbied vigorously to retain the current standards. Obviously, they stand to lose revenues if the automobile producers return to a simpler emissions control technology. MECA has argued that the current technology, including electronic ignition, three-way catalysts, and oxygen sensors, has contributed to improved performance and fuel economy. In congressional testimony, MECA has argued that the upper limit of cost savings from relaxation of the statutory standards would be $150 per car, not $300.

Because environmental regulation is a responsibility shared by all three levels of government, a major organized interest in the debate over automobile controls consists of regulators from state and local government. Local regulators have opposed any relaxation of automobile emissions control standards. A relaxed standard for automobile emissions would force local regulators in highly polluted areas to write more rigorous standards for stationary sources, imposing economic and political costs on state and local governments.

Environmental groups also oppose any change in automobile regulation. In part, their opposition is strategic: they prefer not to open the Clean Air Act to amendment on the fear that they have more to lose than to gain in the present political climate. In part, their opposition is based on an underlying philosophy of "technology-forcing" regulation designed to minimize emissions. They maintain that eco-

nomic efficiency arguments should be given relatively little weight. Their objective is simply to make air cleaner everywhere. Relaxing a standard when it is technically possible to satisfy it is therefore certain to be opposed by environmental groups.

Within the automobile industry, the issue of averaging emissions across models, rather than having a uniform standard for all cars, is controversial. General Motors, with many model lines, is enthusiastic about the proposal. Ford argues that it deserves study. Chrysler and American Motors fear that it will disadvantage smaller, single-line companies. Economic efficiency would probably require at least some flexibility across model lines, but with Chrysler already surviving on government-guaranteed loans, any change in this direction will be scrutinized carefully by Congress.

Inspection and maintenance programs also receive considerable attention. Auto manufacturers prefer the mandatory, universal system that shares responsibility for poor performance among automobile owners, emissions control device manufacturers, and state and local governments. Another interest group has entered this debate: the people who inspect and repair vehicles. This group favors a system whereby every automobile will be inspected and owners will need to have repairs made before passing the test. Only environmentalists favor comprehensive testing. They also want to have a mandatory, uniform system to encourage proper vehicle maintenace.

Automobile owners are unrepresented in these debates. To the extent that they care about air quality, they are represented by environmental groups. To the extent that they care about the cost of new autos, they are represented by auto manufacturers. But nowhere is there systematic, comprehensive representation of their interests. Such representation would require taking account of the full costs (including the energy-related costs) of operating an automobile, as well as the proper trade-offs between automobile controls and standards for other sources. Drivers would presumably be interested in more cost-effective methods for reducing total emissions from autos, such as substituting fuel injection for carburetion and making a concomitant relaxation of tailpipe emissions controls.

None of the organized interests is concerned with policy at this level. MECA wants to enhance the tailpipe device business, whereas auto manufacturers want to reduce the costs of complying with regulation. Environmentalists want more controls at every stage, and state and local regulators are jockeying with EPA, each trying to get the other to bear as much of the political cost of developing and enforcing regulations as possible. Farmers and oil companies are squaring off

over the issue of promoting methanol, but this debate has not yet spilled over into environmental policy.

Political debates about auto emissions standards have often reflected bitter divisions among the interest groups that are able to find representation in the process. But the environmental area is not the only one that has this characteristic. Energy regulation also has its private factions, as we shall see in the discussion of natural gas deregulation that follows. And like the Clean Air Act standards, which created such new interest groups as MECA, natural gas regulation has created its own interests with a stake in perpetuating the status quo.

Natural Gas Deregulation

Public regulation of natural gas markets has frequently sparked rancorous political battles. Discord in natural gas policy debates and the resulting inability of the policy process to arrive at solutions that serve the public's general interest have been the consequences of the extreme complexity and detail of natural gas regulation. The prevalence of special categories, clauses, exceptions, and exemptions has consistently split otherwise "natural" consumer, producer, and distributor interest groups into multiple factions, each with vested interests in particular components of policy. The difficulty of mediating these interests has thus far prevented policy makers from wiping the slate clean and undertaking fundamental regulatory reform.

The regulation of natural gas markets takes place at several stages. The federal government regulates selling prices charged by gas producers. Natural gas pipelines are subject to both federal and state regulation, typically of the rate-of-return variety, of their sales to retail gas utilities; and state public utility commissions regulate the operations and prices of local distribution systems. The federal role began with pipeline regulation under the Natural Gas Act of 1938 and was extended to field price setting for gas sold in interstate commerce in 1954. By the mid-1970s, these federal price controls on natural gas had produced the classic results described in economic textbooks: demand at regulated prices exceeded available supply. Consumers in interstate markets faced severe shortages. The political consequence was the Natural Gas Policy Act of 1978 (NGPA).

The central thrust of the NGPA is ostensibly some measure of deregulation of natural gas prices designed to prevent the recurrence of shortages. The act passed by Congress, however, may actually have worsened the allocation of natural gas resources.[10]

The act extended price controls to previously unregulated intra-

state markets. This discouraged production but simultaneously increased the relative profitability of selling to interstate buyers. It also created an incredible array of gas categories, each of which receives a distinct price and schedule of gradual price escalation. The NGPA also removed price ceilings on gas from deep wells and, by 1985, on gas from newly developed properties. This combination of partial deregulation and the disparate effects of the NGPA's complicated provisions across interest groups have resulted in the current move toward further regulatory reform. Proposals for "reform," however, range from immediate and full decontrol to blocking further implementation of the NGPA and reimposing tight ceilings on all gas.

The NGPA has produced a number of significant and well-recognized anomalies and economic distortions.[11] The act was originally designed to allow the price of new natural gas to rise gradually to market-determined levels. Gas and oil are close substitutes, and market prices for gas are at parity with prevailing oil prices. The very sharp rise in oil prices in 1979–1980, however, left NGPA escalation rates too low. Thus, there is a strong likelihood that natural gas prices could make a sharp and politically unpalatable jump in 1985. Industrial, residential, and commercial consumers are accustomed to artificially cheap natural gas. They are thus motivated to oppose further moves toward deregulation. The magnitude of per company stakes has put utilities near the front of industrial sector opposition to decontrol, while a homogeneity of interests has brought residential and commercial consumer interest lobbies into the fray.

One of the most anomalous results of the NGPA is that, in at least some regions of the country, the price of natural gas has already been pushed to and perhaps even above free market levels. This development has taken place rapidly and has been felt as a shock by customers who have seen prices rise by as much as 50 percent over the last year.[12] The NGPA caused this because its multifarious price caps interact in curious ways with pipeline regulation.

Pipelines are regulated by setting both long-distance and local distribution rates on the basis of average costs. A pipeline purchasing gas from producers is assured that a particularly high price paid for a given portion of its overall acquisitions can be recouped by forcing up average costs. Thus, a pipeline with access to low-cost, low-ceiling-price old gas can effectively cross-subsidize its purchases of high-cost gas (for example deregulated deep gas) so long as the average-cost price ultimately charged consumers is competitive with alternative fuels. Because the total profits of a pipeline or a distribution company increase with total deliveries and because there is a shortage of low-

priced gas at ceiling prices, pipelines have the economic incentive to engage in precisely such cross-subsidization.

Pipeline bidding for deregulated gas has had two general effects. First, gas supplies that are deregulated have increased in price. Deregulated gas prices at the wellhead in 1982 reached twice the level they would attain in a completely unregulated market.[13] Second, where gas prices are regulated, competition among buyers has led to contract terms favorable to producers. Most notable among these are take-or-pay provisions, which require pipelines to pay for specified quantities of gas even if the pipeline cannot take such gas because of a lack of customers. These provisions primarily affect high-cost gas supplies and have dramatically distorted the purchasing decisions of affected pipelines. The recent recession depressed demand for gas, particularly in the industrial sector, and reduced the load requirements of many pipelines. Some pipelines as a result have turned away low-cost gas rather than pay for, but give up, high-cost gas. This behavior by pipelines has been strongly criticized by customers, and the resulting high gas prices have served consumer advocates as arguments for recontrol. The behavior, however, has simply been a privately rational response to regulatory incentives.

The competition for natural gas supplies has been, and after 1985 will continue to be, unequal across pipelines. Pipelines with relatively large endowments of low-priced gas supplies have been given a "cushion" of revenues with which to cross-subsidize high-priced purchases. This cushion is based on the difference between the delivered price based on the average cost the market will bear and the cost of low-priced endowments. As a result of past regulation, pipeline cushions tend to be largest in the interstate market (that is, at the time the NGPA was signed, intrastate prices and thus ceilings were generally higher than interstate prices). The result is not only a competitive disadvantage for intrastate pipelines, but a distortion in the allocation of gas among consumers. Customers with lower-valued uses served by large-cushion pipelines are able to obtain gas while customers with higher-valued uses served by small-cushion pipelines go unsatisfied.[14] These consequences will gradually be dissipated under the NGPA as low-cost older gas supplies are used up, but it would take well into the late 1980s to accomplish this. Meanwhile, public policy will continue to lead interstate and intrastate pipelines and their respective customers to have divergent interests in natural gas debates.

Producers of natural gas gain from higher prices, and the NGPA has held the national average of producer prices below the market level. It should not be concluded, however, that producers unani-

mously support immediate and full deregulation. Those producers currently selling deregulated deep gas at (cross-) subsidized prices, for example, are better off under the NGPA, as would be producers of gas deregulated in 1985. The nation's economy, however, suffers from the inefficiency of regulations that induce producers to explore for and develop relatively expensive supplies while cheaper, but price-controlled, resources lie idle.

Rapidly rising natural gas prices in some markets, subsidies for some types of gas production and price ceilings on others, competitive and revenue boosts to some pipelines at the expense of others—all of these anomalies contribute to current political pressure to alter the NGPA. If the administration and Congress open the issue they will be confronted by a tangled web of interest groups. Some customer groups have coalesced because of their homogeneity and the size of the stakes and can be expected to continue to support moves toward recontrol. Pipeline companies do not benefit from the discouragement of production that price controls bring. Interstate pipelines nevertheless can generally be expected to support the type of gradual deregulation embodied in the NGPA. Intrastate pipelines, however, have an interest in more rapid decontrol. Finally, natural gas producers are well organized politically through trade associations and include some of the largest corporations in the country. They too, however, cannot act with one voice, as the NGPA has split their interests.

Conclusion

Chapter 1 has been designed to whet the reader's appetite. We have summarized half a dozen disparate regulatory areas now on the policy agenda: deregulation of network television, the fifty-five mile-per-hour speed limit, real estate settlement costs, FDA drug licensing reform, auto exhaust standards, and natural gas prices. Despite the disparity of the issues, the underlying pattern is similar. The future of regulation in these areas is being decided in an environment of interest group pressures and representation that is often entirely predictable. If it is predictable, is it possible for the policy maker and the public to assess the likely strengths and weaknesses of the arguments and the information with which they are presented? What about the groups that are not heard at all?

Chapter 2 will explore the basis in political/economic theory for making predictions about interest group motives and participation in deregulatory debates.

Notes

1. For general surveys of the literature on the economics of regulation, see Roger G. Noll and Paul L. Joskow, "Regulation in Theory and Practice: An Overview" in Gary Fromm, ed., *Studies in Public Regulation* (Cambridge, Mass.: MIT Press, 1981); Bruce M. Owen and Ronald Braeutigam, *The Regulation Game* (Cambridge, Mass.: Ballinger, 1978); Sam Peltzman, "Toward a More General Theory of Regulation," *Journal of Law and Economics*, vol. 19, no. 2 (1976), p. 211; Leonard W. Weiss and Michael W. Klass, eds., *Case Studies in Regulation* (Boston: Little, Brown, 1981).

2. Here and throughout this book the term "network" is used synonymously with ABC, CBS, and NBC, the three major national advertiser-supported commercial television broadcast networks. There are other sorts of television networks, such as the Public Broadcasting System or commercial and noncommercial satellite-distributed cable network services, but the FCC rules here in question were designed to apply only to these three companies, and as yet no others have fallen within the FCC's technical definitions. Also, network "regulation" and "deregulation" are defined for our purposes solely in terms of program supply, that is, they refer only to regulations that affect the relationship between the networks and program suppliers and syndicators. They do not include other FCC regulations such as those that affect network-affiliate relationships or political broadcasting.

3. Syndication is the licensing and distribution of feature films and television programs to local television stations on a station-by-station basis.

4. Federal Communications Commission Network Inquiry Special Staff, *New Television Networks: Entry, Jurisdiction, Ownership and Regulation* (1980).

5. For example, see Charles Lave, "Potential Energy Savings in Urban Transportation," in *Economic Impact of Energy Conservation*, U. S. House of Representatives, Subcommittee on Advanced Energy Technology and Conservation of the Committee on Science and Technology (Congressional Research Service, 1979).

6. Thomas H. Stanton and John P. Brown, FTC staff, prepared statement for *Hearings on the Real Estate Settlement Procedures Act*, U.S. House of Representatives, Subcommittee on Housing and Community Development of the Committee on Banking, Finance, and Urban Affairs (September 15, 16, 1981), pp. 75-76.

7. Linda E. Demkovich, "Critics Fear the FDA Is Going Too Far in Cutting Industry's Regulatory Load," *National Journal*, vol. 14, no. 29 (July 17, 1982), pp. 1249-52.

8. Lawrence J. White, *The Regulation of Air Pollutant Emissions from Motor Vehicles* Washington, D.C.: American Enterprise Institute, 1982), chapter 6; Robert Crandall, Theodore Keeler, and Lester Lave, "The Cost of Automobile Safety and Emissions Regulation to the Consumer: Some Preliminary Results," *American Economic Review: Papers and Proceedings*, May 1982, pp. 324-27.

9. U.S. Senate, Committee on Environment and Public Works, *Hearings: Clean Air Act Oversight, Automobile Emissions Standards*, part 4, June 23, 24, and 25, 1981; U.S. House of Representatives, Subcommittee on Health and Environment of the Committee on Energy and Commerce, *Hearings on H.R. 4400 and H.R. 2310: Mobile Source Provisions*, September-December, 1981, and January 21, 1982.

10. See U.S. Department of Energy, *A Study of Alternatives to the Natural Gas Policy Act of 1978* (1981) and Glenn C. Loury, *An Analysis of the Efficiency and Inflationary Impact of the Decontrol of Natural Gas Prices* (Washington, D.C.: Natural Gas Supply Association, 1981).

11. Milton Russell, "Natural Gas Deregulation: Overview of Policy Issues" (unpublished, Resources for the Future, 1982) provides an excellent summary of these effects of the NGPA.

12. Bureau of Labor Statistics, "Consumer Prices: Energy and Food" (monthly).

13. See Catherine Good Abbot, "Is This a Natural Gas *Market*?" (U.S. Department of Energy, Office of Policy Planning and Analysis, October 15, 1982).

14. Robert C. Means, Office of Regulatory Analysis, Federal Energy Regulatory Commission, "Analysis of the Bidding Disparity between Interstate and Intrastate Pipelines" (Natural Gas Deregulation Seminar, American Enterprise Institute, March 12, 1982) discusses the relative positions of interstate and intrastate pipelines under the NGPA.

87

THE POLITICAL ECONOMY OF GOVERNMENT REGULATION

edited by
Jason F. Shogren
John A. Walker College of Business
Appalachian State University
Boone, North Carolina 28608 USA

Kluwer Academic Publishers
Boston/Dordrecht/London

3

BOOTLEGGERS AND BAPTISTS IN THE MARKET FOR REGULATION

Bruce Yandle

In the Beginning

Regulation of individual behavior by higher authorities is as ancient as the Garden of Eden and as recent as yesterday's Federal Register. Adam and Eve chaffed against the iron-clad specification standard they confronted, accepted the advice of an independent counselor, engaged in noncompliance activities, and suffered the consequences. They were required to leave a pristine environment where entry was barred and move to a significantly deteriorated competitive location where labor productivity was lower and future regulations would be crafted by their fellow man.

Still today, many people are frustrated by complex environmental rules, seek wise counsel as to how to deal with them, and sometimes pay high penalties when they fail to satisfy the regulator. On the other hand, other people complain about the lack of rules, seek more of them, and lobby fiercely for stricter enforcement.

Yet a third group of people from within the ranks of the frustrated and penalized silently accept regulation and welcome the support of those who seek more of it. Indeed, a careful

examination of most any successful regulatory episode suggests that there are winners and losers at the margin, but also that the hats worn by the participants are rarely all black or all white. We understand that regulation, like taxation, redistributes wealth and carries costs. We also know that regulatory reform, like tax reform, alters the former redistribution effects and may relieve some of the burdens of regulation, provided the key parties that originally sought the regulation somehow support the changes.

What Theory Tells Us

The economic literature on regulation and efforts by special interest groups to gain favors from government--whether they be members of the steel industry or the Sierra Club--illuminates some of the dimensions of the demand for regulation, at least for some of the people in a regulation story. We now understand that some regulated firms view regulation favorably, once they realize government intervention is inevitable. Regulation is not necessarily a government-designed hair shirt that constantly limits the desired actions of these firms. It is more like an old tweed suit that not only fits but also feels good. Indeed, those in the tweed suits get upset when efficiency lovers suggest that the government-imposed suit should be thrown away. They do not want deregulation. Regulation has tilted the economic game in their direction and now protects their position.

Their long-suffering compatriots, scratching in their hair shirts, feel differently about the matter, but may not be very successful in a showdown, because they lack the support of another group: Those who seek regulation for reasons that have nothing to do with anti-competitive tweed suits. In other words, there are coalitions that work and others that do not.

The modern theories of regulation that carry us beyond the noble public interest story gain considerable yardage in explaining important aspects of many regulations. A widely cited theory of regulation developed by Richard Posner causes us to focus on transfers of wealth from politically weak to stronger groups.[1] Fundamental work by Nobel Laurette George Stigler and Sam Peltzman ask us to consider special interest groups, to look for differential effects across those groups, and to view regulation as a market process with demanders and suppliers.[2] Stigler and Peltzman also argue that most successful regulation will generate some benefits to consumers, even though the favored producer group will likely gain the most. Gordon Tullock and Nobel Laurette James Buchanan, who are the founders of the Public Choice school in economics and political science, call our attention to rent-seeking behavior, where government's power to limit competition and output always beckon and where seekers of government favors tend to spend the value of their expected gains while chasing them.[3] Tullock and Buchanan observe regulation as a way to restrict output, raise price, and foreclose markets to new competition.

Fred McChesney notes that politicians can act as agents in the regulatory game, profiting regardless of the result.[4] He suggests that politicians can propose harsh rules that cause the affected individuals to organize politically and lobby for relief, which strengthens the politician's position. Once the rules are in place, the organized group will continue to lobby, supporting those politicians who are sympathetic to their cause. Gary Becker directs our attention to coalitions that favor and oppose actions by government to redistribute wealth, whether that is done by taxing and spending or by rent-generating regulations that impose a tax to be shared by one and all in the form of deadweight

cost to the economy.[5] His analysis focuses on political actions and reactions that are induced when groups seeking governmental favors are countered by others who bear the net cost of those actions.

A Neglected Point

A host of empirical work now lends strong support to elements of each of these compatible theories. However, while there is room in the new theories for a neglected focus, none of the theories emphasizes the potential importance of having public interest support for successful regulatory ventures. It is as though the public interest theory--the pure notion that political agents are dedicated to serving a collective public interest--died from over-exposure to better theories and left no heirs. I argue that politicians in a democracy must find ways to dress their actions in public interest clothing. Highly visible special interest benefits just cannot be transferred in the raw.

To make the point, consider these questions. Would the Act to Regulate Commerce that produced the Interstate Commerce Commission in 1887 have passed without the support of Populists who thought they were getting the best of the railroads? Would the English Factory Acts that arguably gave an advantage to capital intensive firms have been passed without the movement against child labor? Would federal meat inspection and the associated limits on the importation of foreign beef have made it on the law books without stories of poison food and an associated public outcry? I say "No" to these questions. But saying no is not enough. We need to know more about how a "moral majority" becomes valuable in the political economy of regulation. Think about coalitions of groups that support government action where the coalition includes

some that seek directly enhanced wealth and others that wish for an improved vision of society.

This chapter focuses on that particular coalition of interest groups that always seems present in a successful regulatory episode. The chapter addresses economic and social regulation, the rules that dictate methods of production and consumption, where the focus is on functions within firms and those that affect industry behavior with respect to price and entry.[6] The chapter presents a theory of durable regulation--the theory of Bootleggers and Baptists--and provides summaries of regulatory events that appear to be explained by the theory.

The Theory of Bootleggers and Baptists

Bootleggers and Baptists have historically supported a form of social regulation that closes corner liquor stores on Sunday. The two groups are very distinctive, even though we can refer to their joint effort as forming a coalition. Think about their differences. Bootleggers are generally not accepted in polite Baptist company, certainly not when wearing tags that identify their occupations. Of course, some bootleggers may be Baptist, but the brethren don't advertise that in the Sunday bulletin. Now, consider their common interest--Sunday closing laws that shut down the corner liquor stores. The bootleggers want to eliminate direct competition. The Baptists want to reduce indirect competition and diminish the consumption of alcoholic bevereges.

But when we say they both lobby, we must add quickly that the lobbying occurs in markedly different ways. The bootleggers do not organize walks, parades, letter writing campaigns or sit-ins at state capital buildings. They confront the politicians more furtively, yet more positively. The Baptists bring something to the anti-competitive effort that cannot be delivered by

bootleggers. They add public interest content to what otherwise would be a strictly private venture. The Baptist element, which I ask you to think of as a generic term, adds a moral ring to what might otherwise be viewed as an immoral effort, the passing of money (and electability) to politicians to obtain a political favor.

Probing deeper into this notion, think about the design of the regulation delivered to bootleggers and Baptists. The common regulation does not consist of higher taxes on alcohol, that is, the use of economic incentives, even though efficiency-driven economists might term that approach the more efficient one. Nor does it address the Sunday consumption or possession of alcolohic beverages. It is command and control regulation that focuses solely on the sale of the product. If a diminution in the consumption of the liquid were the over-riding goal, a public interest theory would likely predict consumption to be the offense. Going further, a pure public interest argument might conclude that higher taxes on the undesired beverages would address the problem. Of course, monitoring and enforcement costs have to be considered. There is, after all, a supply side to all regulatory problems. But as Becker reminds us, the wealth-redistributing regulation obtained is probably the most efficient in that set; which is to say that both the bootleggers and the Baptists have to be satisfied with the final equilibrium.

Interestingly, regulations of the Sunday sale of booze tie together bootleggers, Baptists, and the legal operators of liquor stores. The bootleggers buy from the legal outlets on Saturday, sell at higher prices on Sunday, and the Baptists praise the effort to enforce the regulatory cartel. Meanwhile, the political suppliers of the regulation reap the support of all the groups, and the Internal Revenue Service

works to prevent market entry by those who would produce alcoholic beverages on homemade stills.

What might cause the coalition to crumble, so that we might observe the repeal of Sunday closing laws? To answer that question, we must consider some elements of regulatory demand. First, the Baptist appeal works so long as most of the Baptists recognize and accept the over-riding moral argument, so long as the group continues to represent a politically valuable interest group, and so long as group leaders are able to marshall resources from the members. There is always a potential free rider problem in such ventures. What we term the public interest is defined by public opinion, but delivering political support is fraught with practical problems.

Next, the bootleggers must earn a high enough return from their endeavor to buy political favors. If entry occurs in their market--by means of more illicit stills, lower cost transportation from locations that have no restrictions, or by the expansion of lower cost private clubs that offer the restricted beverage to their members, or if demand for the product simply diminishes, the bootleggers will be pushed from the picture. Once the restriction either ceases to be binding or loses its moral support, we predict regulatory reform and the possible replacement of bootleggers by another politically powerful interest group, such as the private club owners.

While bootleggers and Baptists are dominant figures in the theory, there are always other groups who bear the costs of the restriction. They too can become more powerful, especially as the costs of the restriction rise, and exert enough force to overcome the political demands of the dominant group. Opportunity cost tends always to raise its head.

An examination of several recent regulatory episodes will illustrate some of these analytical points. To give a flavor of some research and

findings, four regulatory stories can be considered: State regulations of Sunday retail sales, or Blue Laws; federal regulation of flammable sleepwear; state regulation of gambling, or lottery laws; and a state/federal episode involving seatbelts and airbags.

State Blue Laws

State Blue Laws are kissing cousins of Sunday Closing Laws that make a market for bootleggers.[7] Both regulations date back to colonial times, probably reflecting religious preferences of the time. Indeed, the term "Blue Law" takes its name from the color of the paper on which early colonial statutes were written. The modern period finds Blue Laws in a gradual state of decline. For example, in 1970, 25 states had restrictive statutes. By 1984, only 14 states remained in the fold, and others were threatening to modify or repeal outright the remaining vestiges of the centuries-old institution. The systematic disappearance of Blue Laws provides an opportunity for researchers to examine the shifting support for the law and so to identify what might motivate the political economy that delivers the rules.

Research on this topic focused on theoretical notions about the demand for Blue Laws and from that developed statistical models that might test the theory. To capture changes in Blue Laws, data for 1970 and 1984 were examined. The theoretical arguments played on the theme of bootleggers and Baptists. We argued that Blue Laws preserve retailers' revenues while distributing those revenues over fewer operating hours and increasing the average per customer purchase. Since their organizing costs are already covered, we argued that unions might be better positioned to bargain for higher wages, though we recognized that most retail establishments are not unionized. We also argued that unionized labor forces have more predictable and uniform work hours and holidays,

which means restrictions on Sunday shopping carry lower opportunity costs for unionized communities. Union workers were first members of the bootlegger group that might favor Blue Laws.

We then thought about bootleggers on the side of repealing Blue Laws. Among various retailers, we predicted large drug stores would fight Sunday restrictions, since they must operate their core businesses on Sunday, have variable costs covered, but are limited by Blue Laws in selling a portion of their inventories defined as "nonessential." On the other hand, we predicted large general retailers would support the laws in the pre-mall 1970 period, since those establishments were geared to compete for downtown shoppers who wouldn't likely flock to the cities on Sundays. In the later period, we predicted large retailers would be either indifferent or opposed to Blue Laws. The large stores might be termed backsliding bootleggers.

The opportunity cost of shopping entered our analysis in another way. Historically, women have specialized in shopping. As the average workers' real wage fell, more women entered the work force, the opportunity cost of Blue Laws rose. The percent of women in the labor force proxied for neither bootleggers or Baptists, but simply served in our research as an indicator of the cost of the restriction. However, in the earlier period of the analysis, we argued that the widely fractured population of women workers faced a high cost of organizing politically. Later, and due to other causes, women became a more identifiable interest group.

The chief Baptist element in our analysis was the Baptists themselves. We used the percent of the population Southern Baptist, an organized interest group that polices free riding with sanctions delivered by conventions, as a proxy for a moral majority that favored Sunday restrictions.

The statistical counterpart of our theory used a "yes, a state has Blue Laws/no, it does not" indicator as the dependent variable and included the arguments mentioned as independent variables. Our statistical findings, which are reported in Table 3-1, indicate that the share Baptist has a strong positive association with Blue Laws in 1970, but none in 1984. The Baptist effect seemed to dissipate over time. The percent union is not significant in either period. However, the number of retail stores with more than 100 employees is positively associated with Blue Law status in the 1970 pre-shopping mall period, but has no association in 1984, when larger retail stores were generally found in suburban shopping centers. Drug stores are negatively associated with Blue Laws in 1970, but not associated in 1984.

Briefly stated, we found support for both parts of the theory. Large retail stores and women appear to be bootleggers in the early period. Indeed, work opportunities for women in the early period were associate largely with downtown retail stores and offices. We speculate that having a Sunday holiday appears to have been more important than having opportunities to shop on Sunday. Drug stores bore a cost in the early period, and Baptists played the expected Baptist role. As time passed and the nature of female employment changed, women apparently bore more of the regulation's cost. The Baptist influence, which may have been delivered in large part by women, eroded, and all other significant opposition faded. Blue Laws were repealed.

Table 3-1. Regression Results. Dependent Variable: Blue Law Indicator

	1970		1984	
Variable	1	2	1	2
Intercept	2.414 (0.871)[a]	2.535 (0.979)	0.1204 (0.095)	0.3487 (0.436)
Date	-0.001 (1.178		-0.0004 (0.295)	
Bapt.	0.0325 (3.291)	0.0317 (4.271)	0.0090 (0.503)	
Party	-0.0018 (0.285)		-0.0019 (1.995)	0.0156 (3.767)
Union	0.0092 (1.015)	0.0115 (1.388)	0.0063 (0.662)	
Women	0.0168 (2.762)	0.0185 (3.311)	-0.0243 (1.535)	
SML	-0.00004 (1.501)	-0.0004 (1.688)	-0.0001 (0.864)	
Large	0.0038 (2.712)	0.0026 (2.673)	0.0005 (0.671)	
Drug	-0.0053 (2.595)	-0.0051 (2.581)	0.0004 (0.264)	
Tour	-0.0006 (0.831)		-0.00006 (0.662)	
R^2	.37	.38	.20	.27
F	4.20	5.47	2.37	10.24

[a]Absolute value of t-statistics in parentheses.

Note: Variables in model, not mentioned in text include: Party: The share of legislative seats controlled by the majority state party and TOUR: which is annual dollars of tourism spending.

Flame Resistant Sleepwear

An episode involving the Consumer Product Safey Commission's (CPSC's) 1971 imposition of a flammability regulation for children's sleepwear is particularly interesting, since the agency later, in 1977, banned the chief chemical agent used by industry for meeting the flammability regulation.[8] The chemical, Tris, was found to be a carcinogen. We learned from an examination of background data that Asian-produced sleepwear had taken a substantial part of the U.S., U.K, Canadian, and European markets prior to the CPSC regulation. The U.S. and U.K. adopted flammability standards at the same time, and foreign penetration of their markets fell markedly. That did not occur in the other developed markets that had no flammability standards. We were suspicious.

In this research, we argued that domestic sleepwear producers gained increased market share from the flammability regulation, but did not likely gain much in the way of profits. Entry is relatively quick and easy in that end of the apparel industry. Thinking more about industry supply curves, we argued that certain producers of synthetic fibers gained from the rules, since cotton fiber and fabric could not meet the CPSC standard. Cotton's market share fell to zero after the regulation. Finally, we argued that the producers of Tris, the chemical selected to meet the standard by virtually all in the yarn and fabric industry, gained from the standard. The product was patented, and five U.S. firms were licensed to produce and market it. They had the most inelastic supply curve of all. Focusing on the demand side, we argued that the demand for the flammability treatment was very inelastic, since it was mandated by law and there were few if any substitutes. That made the burden of the restriction quite palatable.

Although we contend that cotton producers bore

the brunt of the industry cost in the episode, we argued they were guarded from losses by a long-standing government-sponsored cartel. The U.S. Department of Agriculture protects cotton producers through acreage and price controls. Perhaps that is partly why a rule destroying a small market for cotton goods could make it through the political thicket.

Using an import penetration model and then financial markets analysis of portfolios of apparel, fabric, fiber, and chemical firms, we tested for effects. The results of the import penetration model, reported in Table 3-2, provided strong support for the theory that import penetration fell markedly with the imposition of the flammability rule. We also found strong support for the notion of financial gains for fiber and Tris producers. But we found no evidence of gains for apparel and fabric producers, while recognizing that most of the apparel firms were too small to be accounted for in our financial markets tests.

This research suggests the bootleggers were the owners of specialized capital in the chemical and fiber industries. But who were the Baptists? They were the parents of children, and other consumer groups, who pressured the CPSC to develop an all-encompassing flammability standard, a rule that would spread the cost of a desired feature across all consumers in the market. Along these lines, a 1971 report in Chemical and Engineering News stated: "An unlikely coalition of mothers and some chemical companies is pleased with the newly promulgated standard."[9] By our theory, the coalition was not an unlikely one.

Of course, the ban on Tris unraveled all this. Unfortunately, the CPSC rule had the effect of spreading a cancer risk across an entire population of young children. There was an

Table 3-2. Regression Results. Dependent Variable: Imported Sleepwear/Domestic Sleepwear

Variable	Coefficient	T-Statistic
Intercept	-0.1970	-0.565
Income	1.2248E-07	3.148
Price	-0.3783	-2.129
Dum 72	-0.2117	-4.847
Dum 75	-0.0760	-2.093
Ban	-0.3482	-4.688
R^2 : .68		
F : 8.30		
Durbin-Watson: 1.825		

Note: The independent variables adjust for real total disposable INCOME, the ratio fo the domestic apparel PRICE index to the all item CPI, the July 1972 flammability regulation--DUM 72, a 1975 regulation that expanded coverage of the previous rule--DUM 75, and the later BAN on chemically treated sleepwear that came in 1977.

understandable public outcry to the news about Tris. The ban that ended that part of the episode resulted in large financial losses for the chemical industry and smaller ones for the fiber industry. The Tris episode was a case of regulatory failure for all parties, except possibly the political agents who managed it.

The ending of the story illustrates another point: When the bootleggers lose the Baptists, the regulation goes away.

State Regulation of Gambling

Government sponsored gambling can be traced

back at least as far as Caesar Augustus, who instituted lotteries for the purpose of rebuilding Rome. They were used by Queen Elizabeth to help fund the Virginia Company's founding of Jamestown. But in more recent times, state-operated lotteries have emerged as a durable source of revenues for state governments.[10]

On its face, gambling is a moral issue to many religious groups. And proposals to institute state lotteries are always opposed by denominational groups. The Baptist element that opposes this state regulation is apparent. That being so, why have lotteries become so popular in recent years? Is this a case like Blue Laws, where moral influence seems to have been swamped by other effects? At present, 28 states have lotteries, up from one in 1964, and their net proceeds in 1986 amounted to more than $5 billion.

Our research on this topic sought to explain state intervention to operate lotteries across states in the face of moral opposition. In our theory, we argued that marginal analyses were made during each legislative session, which is to say that laws could be passed or repealed each year. We first noted that when lottery revenues are viewed as tax revenues they are highly regressive. Put differently, state-operated lotteries provide an opportunity to transfer income from lower-to higher-income taxpayers, which gives the first bootlegger clue.

We also observed that the ever-present demand for gambling, which is relatively inelastic, can be satisfied by either private or public means, and that private provision occurs legally and illegally. If states are to enter the market successfully, they must find ways to limit their competition. That led us to say that states operating a monopoly lottery will generally have a larger police operation than other states, all else equal.

In the analysis, the demand for repeal, as

observed in the nonlottery states, was driven by the Baptist element, which we proxied by the percent of the population holding that faith. Demand for lotteries was driven by higher income people, proxied by average per capita income, with stronger support coming where state debt per capita was higher and where states had a constitutional requirement for a balanced budget. The number of police, weighted by population, entered the analysis to determine its relationship with lottery status. We also included state taxes per capita in our model that explained the occurence of lotteries, suggesting that taxes were a substitute for lotteries. The higher those taxes, the less likely a lottery would exist, all else equal.

Focusing on current data in our statistical testing, our estimate, reported in Table 3-3, found the percent Baptist to be negatively associated with lotteries, debt per capita and income per capita positively related, and police per capita positively related. The presence of a balanced budget requirement was not quite significant, though its sign was positive. We also found state taxes per capita negatively associated with lotteries, which supports the notion that lotteries are a substitute for taxes.

This analysis suggests how other forces can overwhelm a moral element. That is, the bootleggers overwhelmed the Baptists. But we cannot say that Baptists are no longer influential. A related question remains to be resolved. Most states that pass lotteries earmark the funds for some popular social purpose--such as for education. Quite possibly, the bootleggers gain the support of the Baptists by providing an apparent link to a public interest cause that offsets the gambling stigma.

Table 3-3. Regression Results. Dependent Variable: Lottery States (Yes = 1; No = 0)

Variable	Coefficient	T-Statistic
Intercept	0.3582	0.758
Income	0.854	2.212
Police	0.572	1.676
Debt	0.1584	2.046
Tax	-0.1028	3.376
Balance	0.2738	1.382
Baptist	-0.0254	4.393
R^2 : 47		
F : 8.36		

Note: Independent variables are: Per capita state INCOME, Number of POLICE per 1000 population, State DEBT per capita, Occurence of constitutional BALANCE budget requirement (yes=1, no=0), Percent of state population that is Southern BAPTIST.

Federally Mandated Air Bags

A last regulatory episode dealt with a very complicated effort by the federal government to mandate the installation of passive restraints in automobiles.[11] The episode began officially in 1969 with an Advance Notice of Proposed Rulemaking issued by the National Highway Safety Administration (NHTSA), followed by a 1971 final rule and a mass of actions, reactions and delays of implementation; and ending with a 1984 Department of Transportation action requiring

states to settle the issue by voting. After two decades, this regulatory issue is still unresolved.

Air bags became a meaningful topic of conversation in the late 1960s when a bulge of people entering the 16-25 age group contributed to a significant increase in auto fatalities. Of course, that was the period when multiple forces contributed to the development of a regulatory binge in Washington, with many new agencies being formed, additional laws passed, and thousands of new regulations placed on the books.

The air bag had been used in to protect test pilots in the development of aircraft, and one of the bag manufacturers approached NHTSA about requiring that the device be used in all new automobiles. They believed that air bags would protect drivers who chose not to use seat belts.

Interest in air bags increased, and Ford Motor Company joined Eaton Manufacturing Company to demonstrate the first working air bag at an engineering society meeting in early 1968. Soon thereafter, Ford became discouraged about the bag's prospects, noting serious problems for out-of-position passenger's and the probability that passengers would be seriously injured by inflating bags. It was also clear that seat belts still would be needed in combination with bags to meet the rules NHTSA was contemplating. General Motors then became the leading proponent for air bags and demonstrated its ability to build bag-equipped autos early on. With that, NHTSA proposed its passive restraint rule, and GM indicated it would strive to meet the standard with air bags. The other major auto producers sued NHTSA. Eventually much more was learned about air bags, auto safety, and the consumers' willingness to buy bag-equipped cars. Along the way, Sam Peltzman shook the cage of safety scholars with his finding that cars equipped with safety equipment could induce a lulling effect

that caused drivers to sustain more instead of fewer injuries.[12] That did not slow the regulatory juggernaut, but other political and market events did.

Who were the supporters of air bags? Obviously, the holders of air bag patents. But the auto insurance industry was the strongest and most persistent advocate of all. Why might this be so? Surely, other groups interested in safety and health would be counted first.

Auto insurers can gain from passive restraints in several ways. First, the insurers could determine risk easier where passive restraints are used. Unlike belts, the mere presence of bags insures protection to the head and upper body in the event of a head-on accident. Second, they could earn one-time gains from contracts written on the basis of higher risks, which would be reduced by the installation of passive restraints. Third, a reduction in head-related injuries--the most expensive of all--would reduce the cost of extensive injury-related litigation, which would reduce price and expand markets. In addition to these reasons for supporting the regulation, as a regulated industry, insurance firms could gain demand for their product by means of the publicity accompanying a long and controversial regulatory proceeding and encounter little competitive response in doing so. Along with the insurance firms, General Motors was a potential winner, at least in the beginning of the episode. Its competition was behind in developing bag-equipped cars. A rule requiring bags would raise competitors' costs.

Who were the Baptists? All those who responded to the promise of safer cars and reduced fatalities. As some might put it, how could anyone be opposed to auto safety? Ralph Nader's Center for Auto Safey, funded largely by the auto insurance industry, was chief among the organizations that prompted this support.

Financial markets analysis was used to estimate abnormal returns for portfolios of auto manufacturers, air bag manufacturers, and insurance firms. The analytical approach forms a portfolio of the relevant stocks and compares its performance to the entire New York Stock Exchange's performance in association with specific events. Sudden changes in the returns to the portfolio relative to the exchange are then identified. In all, 10 key events that might have increased or decreased the wealth of the portfolio shareholders as the rules were imposed, delayed, modified, and finally put to state vote. While there were mixed results for a number of the events, we observed that air bag producers gained substantial wealth when the passive restraint rule was introduced. Ford and Chrysler suffered cumulative losses, but neither GM nor AMC suffered a loss in association with the initial rule.

When NHTSA delayed the rule in 1970, both the insurance industry and air bag portfolios sustained abnormal losses. However, the delay did not generate gains for the auto portfolio, partly because the delay was accompanied by additional safety standards that related to padding and interior design. In yet another event, in 1980 when Congress passed legislation delaying the standard again but requiring passive restraints to be met by producers of smaller cars first--a barrier to the flood of imported cars, both the automakers and air bag producers experienced abnormal gains.

In other work, we examined the later state votes on mandatory seat belts, which could result in the elimination of passive restraint requirements. A vote for mandatory belts was a vote against mandatory passive restraints. Our estimating equation for explaining whether or not a mandatory seat belt referendum was passed, used 1986 data, and the results are reported in Table 3-4. As indicated there, we found that the number

of auto frames produced in a state was positively associated with passage. By the time of the vote, auto producers were generally opposed to passive restraints. Earlier efforts to market bags had not been successful and the required installation of passive restraints in smaller cars would significantly raise price. The presence of an air bag manufacturer in a state had strong negative partial effects, and the number of employees in a state's fire, marine, and casualty insurance sector, weighted by population, was negatively associated with passage of mandatory seat belt laws. They wanted passive restraints.

Table 3-4. Regression Results. Dependent Variable: State Passage of Mandatory Belt Law (Yes = 1; No = 0)

Variable	Coefficient	T-Statistic
Intercept	-1.4184	-2.621
Auto	0.24 E-6	2.098
Insur	-0.1814	-2.428
Air	-0.5453	-2.343
Phys	0.0021	1.001
NOF	-0.1914	-1.646
Educ	0.0280	3.190
High	-22.104	-3.756
Inc	0.00005	1.150
R^2 : .50		
Durbin-Watson: 1.92		

Note: Independent variables are: Total number of AUTO assemblies produced annually in each state; the total number of workers in the insurance industry weighted by population for each state, an AIRbag dummy variable (yes=1, no=0) for states with an airbag producer; number of PHYSICIANS per 100,000 people, a dummy variable for no-fault insurance states, NOF; the percentage population having received a high school diploma; and per capita INCOME by state.

In terms of bootleggers and Baptists, the passive restraint research suggests that auto insurance companies, the producers of air bags, and at one point, the leading auto producer were early bootleggers. Public interest groups took the moral high ground and gained financial support from some bootleggers. As time passed, and import penetration increased, a carefully molded passive restraint rule was seen as a way to restrain to competition. The ranks of the bootleggers increased and moved in lock step with the Baptists. Eventually, belts and bags competed in a political economy with bootleggers and Baptists. Meanwhile, technology advanced, prices of passive restraints fell, and more producers began to offer the item as standard equipment in higher priced cars.

Concluding Thoughts

The theory of Bootleggers and Baptists argues that ideology matters in the political economy of regulation, but it matters in a very specific way. When considering the effective demand for regulation and the final form taken by specific rules, we must look for an important group of demanders who deliver public interest content to the regulatory cause. In the first place, there is considerable competition for political favors, and a politician must be able to explain his actions. That being so, we should expect to find strong public interest statements about the virtues of regulation that can be ratified by important social groups and figures. We should recognize that groups like the Environmental Defense Fund and the Sierra Club, just to name two, are vital to the passage of clean air legislation. We should also recognize that the support of those groups can be quite valuable to polluters who seek a particular form of regulation, a form that may raise their

competitor's costs or in other ways improve the future profits of an industry group.

The struggle for regulation that best serves the bootlegger-Baptist coalition occurs at the federal level. It is difficult to gain very much in a competitive environment across 50 states. In a similar way, the outcomes predicted for the theory seldom apply to actions taken by courts. The theory best explains legislative and regulatory actions where the political process can be affected through lobbying, campaign contributions, and efforts by politicians to satisfy constituent groups.

The evolution of environmental regulation in the U.S. allows us to observe just how much a special interest theory of regulation might explain. In the chapters to follow, direct references will be made to interest groups struggles, but while it may be apparent from the stories, members of the groups will never be referred to directly as "bootleggers" or "Baptists."

Notes

1. See Posner (1974).
2. See Peltzman (1976) and Stigler (1971).
3. See Tullock (1967) and Buchanan (1980).
4. See McChesney (1987).
5. See Becker (1983).
6. See Yandle and Young (1986).
7. This discussion is drawn from Price and Yandle (1987).
8. This section is drawn from Shuford and Yandle (1988).
9. See "Flammability Rule Argued," Chemical and Engineering News, April 9, 1971, 9.
10. This section is based on Martin and Yandle (1988).
11. This section is based on Kneuper (1987).
12. See Peltzman (1975). For more on this see Crandall (1986).

References

Anderson, Douglas D. Regulatory Politics and Electric Utilities. Boston: Auburn House Publishing, 1981.

Becker, Gary S. "A Theory of Competition Among Pressure Groups for Political Influence." The Quarterly Journal of Economics 98(August 1983): 371-399.

Buchanan, James M. "Rent Seeking and Profit Seeking." in Toward a Theory of the Rent-Seeking Society, ed. by J. M. Buchanan, R. D. Tollison, and G. Tulluck, College Station: Texas A&M Press, 1980.

Crandall, Robert. Regulating the Automobile, Washington: Brookings Institution, 1986.

"Flammability Rules Argued." Chemical and Engineering News (April 1971).

Irwin, Manley R. "The Telephone Industry." Walter Adams, ed. The Structure of American Industry. New York: Macmillan Publishing Company, 1982, 298-318.

Kneuper, Robert A. "The Political Economy of Mandatory Passive Restraints: An Investigation of Auto Safety Regulation." Master's thesis, Department of Economics, Clemson University, 1987.

Martin, Robert and Bruce Yandle. "Lotteries as Transfer Mechanisms." Unpublished manuscript, Department of Economics, Clemson University, 1988.

McChesney, Fred S. "Rent Extraction and Rent Creation in the Economic Theory of Regulation." The Journal of Legal Studies 16(1987): 101-117.

Peltzman, Sam. "The Effects of Automobile Safety Regulation." Journal of Political Economy. 83(Aug.-Dec. 1975): 677-725.

_____. "Toward a More General Theory of Regulation." Journal of Law and Economics. 19(August 1976): 211-240.

Posner, Richard A. "Theories of Economic Regulation." Bell Journal 5(Autumn 1974): 335-358.

Price, Jamie and Bruce Yandle. "Labor Markets and Sunday Closing Laws." Journal of Labor Research

8(Fall 1987): 407-413.

Shuford, Gordon and Bruce Yandle. "Consumer Protection, Private Interests Effects, and Government Liability: The Tris Episode." Center for Policy Studies, Clemson University, 1988.

Stiger, George J. "The Theory of Economic Regulation." Bell Journal of Economics 2(Spring 1971): 3-21.

Tullock, Gordon. "The Welfare Costs of Tariffs, Monopolies, and Theft." Western Economic Journal 5(June 1967): 224-232.

Yandle, Bruce. "Bootleggers and Baptists: The Education of a Regulatory Economist." Regulation (May/June 1983): 12-16.

_____. "Intertwined Interests, Rent Seeking and Regulation." Social Science Quarterly 65(December 1984): 1004-1012.

_____ and Elizabeth Young. "Regulating the Function, Not the Industry." Public Choice 51(1986): 56-69.

2

A Review of the Economics of Regulation: The Political Process

Robert E. McCormick

Introduction

The early days of the Reagan administration brought a new breed of civil servant to Washington. Many of them were economists; that is nothing new, but these people were different. They were educated and well-versed in the new economic theory of regulation; they were public-choice scholars. They did not believe that government service was performed for the public good. They did not believe that for government to run properly, all that was necessary was to have the "right" people in office. Instead, these economists and lawyers saw the world in a different light: Politicians did things to garner votes; they were responsive to their electorate; they were not benevolent despots. To the newcomers, regulation was used to redistribute income, not to correct market failures costlessly and perfectly. Theirs was not a philosophical view, it was a hard-edged empirical approach to the world, and it was built on twenty-five years of exacting interdisciplinary academic research. The purpose of this paper is to review and explore that literature, which formed the background of many involved in shaping the Reagan administration's policy on the regulation front.

We start this discussion with predation, a venerable concept in the literature of industrial organization. In fact, few topics have received as much theoretical attention and so little empirical scrutiny. For example, see the exchanges between McGee (1980), Areeda and Turner (1975), and Williamson (1978). Basically, price predation is an economic unicorn de-

Thanks go to James Buchanan, Matt Lindsay, Michael Maloney, Roger Meiners, William Shughart II, and Mark Wasserman for help on an earlier draft. The usual caveat applies.

pending on whether you want it to be or not—there is no consensus of opinion on the matter. Therefore, it is a bit surprising that the theory of predation has actually grown into an area where there is considerable agreement. This is the strategic use of governmental processes to disadvantage consumers and rivals.[1] Predation works by manipulating government regulations and the court system rather than through price cutting. Malevolence can mean higher profits through reduced output and higher price. Regulation fashioned in this manner affords regulated firms several advantages: Services are often provided at less than factor cost, and the monopoly police power of the state offers a unique opportunity to adjust the behavior of rivals. Most prominently, cartel enforcement is made relatively inexpensive, and the scrutiny of antitrust authorities is avoided.[2] Lobbying and other vote-supplying activities are the price that must be paid. Whether the strategic use of regulation is profitable then becomes a capital budgeting problem not unlike most other decisions the firm has to make about purchasing inputs (Salop 1981).

The literature in economics on the strategic use of regulation is relatively new; however, in many respects, all analysis of government fits the description. Moreover, there is a growing literature on strategy in general.[3] To define the topic that broadly here would impose large digestion costs for which most readers have neither the time nor the demand. Hence, for tractability, I adopt a more narrow rent-seeking definition. The strategic use of regulation is any attempt by a firm or a collection of firms or others with similar interest to alter the political or legal structure of the economy to their advantage.[4] This approach purposely ignores the question of good and bad influence. Even more importantly, it does not require deliberate aggression. There are two advantages to ignoring motives: First, they are very hard to determine; and second, from an economic standpoint, only the effects are relevant.

Strategy may take the form of trying to coerce legislation, affecting a bureaucratic ruling, or instigating a law suit. However, from an analytical viewpoint, these actions are the same and serve one or more purposes: To restrict the entry of rivals, to prevent nonprice competition, to differentially impose costs on members of an industry, or to restrict the production of substitute goods and services. Whether these actions are legal is, for the most part, irrelevant from the point of view of economic analysis. That question has received considerable attention elsewhere and is interesting in its own right, but brevity requires that I ignore the issue here.

One of the themes in the literature on the strategic use of regulation is the importance of the self-interested politician. This contrasts with most analysis in industrial organization and predation in particular. Here the politician, his motives, and his constraints are often the center of attention.

Although the role is often subsumed, nonetheless the politician is there. And most importantly, he is not disguised as a public-spirited individual benevolently maximizing some well-behaved social welfare function. Instead, the politician like all other actors is a rational, self-interested, maximizing agent. Of course there are exceptions to this principle. Kelman (1981) is the polar case of ignoring economic incentives in regulatory rule making, and most analysis of antitrust law and enforcement takes a benevolent view of politicians and the law. For example, Easterbrook (1983, p. 24) says, "The antitrust laws, in contrast, are designed to preserve the functioning of competitive markets that, at least presumptively, produce allocative efficiency." The myth of the public spirited politician dies slowly.[5]

Incorporating politicians into the behavioral system adds an apolitical market to the analysis and makes regulations endogenous. This makes it easier to predict many aspects of regulation, such as its inception and the industries that will be affected. This is accomplished by focusing analysis on various groups in the economy; for example, consumers and producers and their competing interests. These groups supply votes and campaign contributions to politicians who in turn supply regulation. The outcome of this process ultimately turns on the relative organizational costs across groups, the structure of political institutions, and the extent of competition in the political market. This is the setting for the strategic use of regulation.

This chapter is divided into six sections. The section 2 briefly reviews the emerging theory of rent seeking and its application to regulation. Section 3 reviews the economic theory of regulation, with particular emphasis on standards and cost-imposing regulation. Section 4 highlights the importance of heterogeneous interest groups in affecting regulation—the regulatory triad. The emphasis is on the ability of some firms to disguise their private pursuits as public interest whether there is a legitimate market failure or not. In some cases mutual interests bring together the strangest bedfellows—the Sierra Club and Eastern coal-mining interests. Section 5 focuses on a relatively uncharted area, the strategic use of antitrust laws to prevent competition. The chapter closes with some suggestions for the direction of future research.

Rent Seeking

The economic analysis of rent seeking was recently surveyed by Tollison (1982). The theory has important implications for the analysis of the strategic use of regulation—especially normative analysis. Tullock's (1967) seminal article demonstrates that transfers are typically not a zero-sum

event. For this reason the economic cost of many activities is often far greater than conventionally assumed.

The normative problem arises because it is impossible to differentiate rent seeking from profit seeking except in the context of a normative model. On the one hand, rent seeking refers to (wasteful) competition for rents created by gifts, grants, or government transfer activities. Profit seeking, on the other hand, refers to those activities that are by definition efficient: Research and development expenditures, piano practice, or committing resources to enter an industry where price exceeds average cost are examples of behavior that create value. By contrast, standing in line for free cheese, taking a politician to lunch in hopes of securing his vote on a bill that provides a subsidy, or arguing before the ICC with an eye toward receiving a certificate are examples of behavior that simply consume rents artificially created by government. That is, rent-seeking activities produce nothing real or consumable, these only result in a transfer. Behaviorally the two are indistinguishable, and it is only morally that these can be made distinct.[6]

This approach has important implications for the strategic use of regulation because it can be viewed in the same light. Consider the case of some vertical restraint on trade, such as the prohibition on resale price maintenance (RPM). Suppose that one accepts the agency cost or public good explanation for RPM; that is, RPM is a device used by manufacturers to force retailers to provide complimentary goods, such as service and information, at the point of sale that they would not otherwise rationally offer. In this case, if a firm brings a law suit or lobbies Congress for a change in the law to allow RPM, then it can be argued that this strategic use of regulation is value increasing *even* if it happens to disadvantage some rivals.[7] In this case the strategic use of regulation increases the real output of society. It is not difficult to construct other examples where the opposite conclusion is reached. Based on this approach, it is imperative to know the firm's motives in order to judge its actions—a difficult chore at best.

The problem is pervasive. Spence (1977) makes a similar point in the context of firm size and capacity decisions. He argues that it is hard to tell which capacity decisions are predatory—designed to limit entry—and which are efficient—driven by competition. Courts have faced the problem in terms of influencing the political process and made their judgment, which has been labeled the Noerr-Pennington doctrine—firms may lobby the government even if it disadvantages their rivals. Fischel (1977) analyzes the antitrust implications of Noerr-Pennington and concludes that lobbying is legal (efficient?), while price conspiracy is illegal (inefficient?).[8] In the Pennington case the doctrine was extended to attempts to influence administrative agencies.[9] The economic theory of rent seeking posits that

competition for rents will drive the expected value of the rents to zero at the margin. Moreover, this competition consumes (costs) the economic value of the rents.[10] Transfers are not free. Of course, there is also competition for profits. Therein lies the conundrum. Government action can create rents through regulation, laws, and court decisions. Firms seek these rents or profits in a variety of ways, most notably here, through attempts to influence government decisions. Whether this behavior is efficient or not is beyond the current state of the literature. It all depends on the nature of the regulation and whether or not the rents are artificially created by government.[11] It is fair to say that the economics of rent seeking implies that normative analysis of the strategic use of regulation is rendered virtually impotent, at least for the moment.

The Economic Theory of Regulation

Part one of this section briefly reviews the economic theory of regulation. Part two focuses on the empirical literature of regulation. Part three looks at heterogeneous interest groups within an industry and across industries.

ECONOMIC THEORIES

It is difficult to trace the evolution of the public-interest model of regulation, although Pigou certainly plays a prominent role.[12] This theory argues that regulation corrects market failures stemming from natural monopolies, externalities, economies of scale, public goods, informational asymmetries, or some other problem in property rights assignments. How this benevolence is accomplished through the political process is almost never addressed in public-interest theory. It is plausible that the public-interest theory of regulation was never meant to be descriptive but instead prescriptive. Nevertheless criticism of the public-interest theory of regulation argues that this do-good approach to the behavior of public officials is analytically embarrassing in light of the propensity of most people to pursue their self-interest. The theory can be partly rescued by realizing that alternatively, *constraints* on politicians' behavior can force them, in quest of votes and wealth, to design and implement laws with general welfare-improving characteristics. This is the spirit of Becker (1983) and to a lesser extent Barro (1973) and Becker and Stigler (1974).

Dissatisfaction with the paternalistic view of government implicit in the public-interest theory of regulation has led to the economic theory of regulation (Stigler 1971).[13] Regulation is demanded by special-interest groups to limit entry, raise price, or otherwise reduce output where the

private costs of cartelization are too high. These laws are supplied by politicians. Subsequent contributions have emphasized cross-subsidization, Posner (1971), and the imperfections of such a cartel, Peltzman (1976). For the most part, industries are assumed to be homogeneous. The battle over rents is a simple struggle between consumers and producers. In the last part of this section, this simple one-on-one perspective is criticized and analysis of heterogeneous interest groups is presented.

The economic theory of regulation is descriptive. It attempts to predict the effects of regulation on price and output, the onset of regulation, the pattern of regulation, and deregulation.[14] For the most part, the theory is void of normative analysis, but there is the presumption, based on the considerable weight of the evidence, that regulation in practice bears little resemblance to the vision of Pigou. That is, regulation is industry-inspired and profitable. The moral connotations of this fact are usually left to the reader.

THE EMPIRICAL LITERATURE

The empirical literature on regulation has one predominant theme: Regulation is often beneficial to the regulated firms. This benefit accrues in one of several fashions. In the simplest form a regulatory agency, such as the ICC, acts as an (imperfect) cartel manager for members of the industry, disallowing entry, apportioning and policing output, regulating price, preventing nonprice competition, and regulating the provision of substitutes. It is widely held that in their original forms, the CAB and ICC were at least operated in this way, if not designed for that purpose. Recent research suggests that the story is more complicated. Boyer (1981) argues that the ICC engages in substantial redistribution of rents across modes of transportation. That is, some rules aid railroads at the expense of truckers and vice versa. Moore (1978) presents evidence that truck drivers benefit from ICC regulation through higher wages. He estimates that union members obtain rents on the order of $1 billion. Certificate owners receive transfers totaling about $2 billion primarily because of restrictions on entry of new firms. These numbers suggest that the strategic use of regulation can be a profitable enterprise.

Taxicab and jitney regulation appears to fit the same mold. Kitch, Issacson, and Kasper (1971) estimate the value of rents created by taxicab regulation in Chicago to be more than $40 million. Eckert and Hilton (1972) contend that jitney regulation was designed to eliminate competition with railroads in the mass transit markets; jitneys were a low-cost, high-quality substitute for railroad transportation, so trains "sought protection from municipal governments, which . . . proved unanimously willing to provide it" (p. 304).

The CAB regulation of airlines had the same characteristics: Entry was barred and price regulated. It is hard to control all margins of competition however. Nonprice competition from within the industry eroded much of the cartel profits. Airlines competed in scheduling and the number of flights by adding capacity to the point where expected profits were zero. Douglas and Miller (1974) argue that this process resulted in average load factors equaling break-even load factors. In turn this impled a " 'ratchet effect' of regulation and reaction, in which price increases, thought by the CAB as necessary to raise profits, only resulted in a new equilibrium with greater levels of excess capacity" (p. 55). Airlines also competed in terms of in-flight service. The CAB responded by regulating meals, flight attendants, and liquor service. What has not been adequately explained is why the CAB restricted these latter forms of nonprice competition but did not regulate the obviously more costly methods of competition through increased capacity or flights per day. One explanation is prominent: Excess capacity benefits airplane producers, pilots, engineers, and attendants, so that the political clout of these groups may have forestalled capacity constraints.[15] The strategic use of regulation implies that if entry is restricted, output is reduced, price is above cost, and nonprice competition sets in. Regulated firms will seek ways of preventing this nonprice competition. The degree to which they are successful depends on the impact of competition on input suppliers and diverse consumer groups.

In another area of long-standing government involvement, Jarrell (1978) presents evidence that state regulation of electricity production was sought to prevent competition where rivalry had brought low prices. In fact, regulation proceeded first in jurisdictions with the *lowest* prices—another nail in the coffin of the public-interest theory of regulation.[16]

It would be a mistake to think that the conventional, hands-on type of regulatory programs, such as electricity and transportation, are the only ones where the economic approach of supply and demand of political action are at work. Marvel and Ray (1983) argue that nontariff barriers to trade implemented after the Kennedy round of tariff agreements were primarily in industries that were vulnerable to foreign competition. Similarly, the literature on broadcasting maintains that regulation of cable television (CATV) has primarily been motivated to protect the interests of local over-the-air broadcast franchises. FCC Chairman Burch has said that CATV regulation could be translated "into the short-hand of protectionism for over-the-air broadcasting, but we feel that is a public interest consideration as well" (Besen 1974, p. 41). Greenberg (1967) and Besen (1974) support the view that the primary beneficiaries of regulation were television stations in the top fifty markets. Comanor and Mitchell (1971) argue that CATV regulations in 1966 and 1968 differentially impacted

small firms and drove them out of business. A similar argument is made about antidumping laws: They are a means of preventing foreign competition. In one notable case Outboard Marine Corporation, the sole U.S. producer of golf carts, wanted the U.S. price to be used to determine whether a foreign producer was selling below cost.[17] That is, the corporation wanted it declared illegal for foreigners to sell below its own price—the ne plus ultra of the strategic use of regulation. In fact, they were unsuccessful.

These few examples are by no means the only types of regulation subject to strategic planning by firms.[18] However, most of the recent research in this area stresses the diversity of interests *within* a particular industry. This is the subject of the next section.[19]

HETEROGENEOUS INTERESTS

The economic theory of regulation falls into one of four analytical categories: Producers versus consumers, cross-subsidization, producers versus producers, and the regulatory triad—producers and an unrelated public-interest group against consumers.[20] The first two categories of analysis have not proven satisfactory in explaining regulation as a general phenomenon, although their usefulness is without doubt in such areas as transport regulation. The simple approach has been weak in its ability to explain why so many industries decry regulation. The answer seems to lie in the fact that industries are not human beings. They are a heterogeneous collection of firms and factors of production whose interests may radically diverge on a particular topic.

Since firms are not homogeneous, input price increases will not have symmetric effects. For example, let there be two different production technologies yielding the same minimum average cost. Let one be capital intensive and the other labor intensive. An increase in wage rates will cause average costs to increase more for the latter than the former (Williamson 1968). Some of the labor-intensive firms will leave the market until price is again equal to average cost. Since average cost for labor-intensive firms increased more than for their capital-intensive rivals, it follows that price increases more than average cost for capital-intensive firms. A profit potential exists if capital-intensive firms can somehow increase wage rates. Presumably labor-intensive firms cannot switch technologies for free.

Consider the simple case of an industry with specialized resources and different firm sizes. The industry supply curve will be positively sloped. Profits are zero at the margin, but inframarginal firms (specialized factors of production) earn rents. Again, suppose regulation imposes costs on *all* firms in the industry, but not symmetrically. The supply curve will shift upward. If costs are heaviest on the marginal firms (factors), then supply

121

will become more inelastic, and price will increase more than cost for some firms. Price increases more than cost for some firms because some rivals are eliminated; therein lies the demand for regulation. There are many ways of achieving success (Salop and Scheffman 1983): Capital-intensive firms can join with a labor union to support an industrywide collective bargaining agreement and adopt a wage sufficiently high to exclude some rivals (Maloney, McCormick, and Tollison 1979). Alternatively, the capital-intensive firms can seek regulation to restrict the use of the input that will raise its price (Marvel 1977, Maloney and McCormick 1982, and Neumann and Nelson 1982). Thus the strategic use of regulation can be an effective means of increasing profits.

Examples of this principle in practice are common in the literature. Marvel (1977) argues that just such a scheme explains the English Factory Laws passed in the early 1800s. Water- and steam-powered mills had different costs of production. According to his argument, water-powered mills depended on abundant rainfall for operation. Laws restricting child labor imposed costs differentially on these water-powered mills because it became more costly for them to operate when the weather was right.[21] That is, steam-powered mills sought regulation as a means of reducing output, raising price, and increasing profits at the expense of their water-powered peers.

Maloney and McCormick (1982) make this argument about environmental quality laws. The current practice of regulating environmental quality through standards rather than emission fees is hard to explain without taking into account the interests of the regulated.[22] Moreover, many details of environmental quality regulation are best explained by noting the potential for intraindustry transfers, as described in Figure 1.1. In two examples of the theory using financial market analysis, cotton dust regulation and the PSD ruling, regulation was associated with large increases in value for some of the regulated firms.[23] Yandle (1980) reports that in the negotiating stages of miles-per-gallon (MPG) regulation, GM lobbied for a standard more stringent than was actually implemented. He also reports that the standard was expected to have differentially large costs on both Chrysler and AMC, especially the latter.

Horwitz and Kolodny (1981) argue that regulation of accounting standards is also the focus of strategic planning.[24] After 1975 the SEC and the FASB required research and development outlays to be expensed. Evidence is presented that large companies in high-technology industries benefited from this ruling because small high-tech companies reduced their R & D expenditures, and some were forced to exit.

Ippolito (1979) argues that insurance regulation appears to benefit small writers at the expense of large direct writers. In most cases, the large direct

writers are out-of-state firms, such as Allstate, whereas the smaller American Agency firms are predominately locally owned and operated. It should come as little surprise that regulation is designed to favor local voters at the expense of foreign disenfranchised firms. In addition, there is an effective cross subsidy to high-risk drivers via assigned risk pools. Maurizi, Moore, and Shephard (1981) report that ophthalmologists and optometrists (especially the latter) have successfully used state regulation to eliminate competition from their optician rivals. The result has been higher prices for eyeglasses. Car prices are also higher because of state regulation of automobile franchises according to R. Smith (1982). He attributes the regulation to lobbying by in-state retailers who gain at the expense of out-of-state manufacturers.[25]

The story goes on. Schneider, Klein, and Murphy (1981) report evidence that the cigarette television advertising ban has actually increased the consumption of cigarettes (because warning ads were simultaneously dropped), and a relative price change has resulted. The cost of introducing new low-tar brands has increased, raising the value of existing brands. Higgins and McChesney (1986) find evidence that the FTC's ad substantiation doctrine benefits some large ad agencies presumably while harming other small ones. The costs imposed by ad substantiation fall more heavily on small ad agencies, who find it more difficult to substitute ads not subject to FTC review. The researchers also report that large firms are vocally opposed to deregulation. Linneman (1980) claims that the 1973 mattress safety standard had little impact on the average quality of mattresses because 80 percent of the mattresses produced already satisfied the standards. However, many small producers were adversely affected because of the increased costs of production. Some exited, and consequently "large, significant, and predictable income redistributions from small to large producers resulted from the 1973 flammability standard" (Linneman 1980, p. 478). He also claims that there was a cross subsidy from low-income to high-income families.

A few more examples should suffice to demonstrate that almost no area of regulation is free from strategic planning by firms to disadvantage rivals for higher profits. There is evidence that large textile producers in the United States not only profited from the OSHA-imposed cotton dust standard but supported its passage (Maloney and McCormick 1982, and McCormick 1983). Oster (1982) specifically argues that many regulatory programs "may be used by groups in the industry as a competitive weapon against other groups" (p. 604). Evidence is presented that this force was important in implementing generic drug laws at the state level. Landes (1980) presents evidence that laws passed in 1920 regulating maximum hours worked reduced the number of hours worked by women and their

total employment. Moreover, the entry of foreign-born women was deterred: Unable to work long enough hours to make the trip profitable, many foreign women chose not to immigrate to the United States. Both of these had the effect of raising the wages of men. Federal regulation of financial institutions differentially disadvantages thrift institutions to the advantage of commercial banks (Tuccillo 1977). Consumer protection regulations at the state level are, in part, motivated by intraindustry transfers (Oster 1980). Johnson and Libecap (1982) discuss the conflict between onshore and offshore shrimp fishermen in the design of fishing regulations in Texas. Hours-of-operation regulation in Canada benefits small stores at the expense of large ones (Morrison and Newman 1983). Building codes restrict the entry of "foreign" labor and prevent use of efficient mass production techniques while increasing the demand for local labor (Oster and Quigley 1977). There is little doubt that whiskey-labeling regulation has been used by certain elements in that industry, bonded producers and Scotch importers, to prevent competition from blended products (Urban and Mancke 1972). Labeling requirements in fact deceived customers into thinking that domestic-blended whiskey had not been aged. On the subject of deregulation, Spiller (1983) presents evidence that there are substantially different effects across firms subject to CAB deregulation based on location and routes.

With a few exceptions, the literature does not claim or present evidence that firms actually sought regulation to hurt their rivals. As stressed earlier, most of this literature is positive or descriptive and looks primarily at the effects of regulation, but there is a growing body that takes a stronger stance (Oster 1982). These redistributional effects are not accidental: General Motors knew what it was doing when lobbying for a stringent MPG standard; Burlington was not stupid when supporting cotton dust standards; the Eastern coal-mining industry was not throwing money away when it lobbied Congress and the EPA for a standards-based approach to sulphur oxide emission reductions (Ackerman and Hassler 1981). However, not all take this view; for example, Noll and Owen (1983) cling to the notion that these redistributional effects are an unintended by-product of regulation.

One conclusion seems inescapable: With the abundant evidence presented, it is hard to argue that managers of firms do not anticipate some, if not most, of the effects of regulation. Rational expectations implies that they will, on average, be correct about the impact of regulation. Given the magnitude of the wealth estimated to be redistributed via regulation, job security implies that managers spend a nontrivial amount of time working a regulatory margin, not just to fight it off, but as an input to their production processes.

The Regulatory Triad

Regulation often brings together groups who have little in common. Yandle (1983) calls this the bootleggers and Baptist phenomenon, reminiscent of restrictions on the sale of alcohol in the South. Industry or a subset desires regulation to capture consumer wealth or disadvantage rivals. An independent group seeks regulation to correct what it perceives as a social ailment requiring government intervention. Private interest joins the public interest, and together they present a stronger political force pitted against the interests of consumers or rivals. In many cases only public interest generates sufficient political support to allow regulation to proceed.

The 1962 Drug Amendments were passed shortly after the Thalidomide incident, even though the bill had languished in committee for years. Peltzman (1974) argues that the amendments created a barrier to entry and raised the price of old drugs. Temin (1979) reports that the Food and Drug Act of 1938 was also passed following a drug accident. Elixir Sulfanilamide contained a poison that killed more than one hundred people in September 1937. Weiss (1964) notes that the 1906 meat inspection laws were passed five months after *The Jungle* was published. He, too, finds the industry in bed with the muckrakers: "I find that members of the industry . . . are as ready to recall the mythology of *The Jungle* as any group has ever been" (p. 120).

Maloney and McCormick (1982) and Ackerman and Hassler (1981) argue that environmental quality regulation is the product of a coalition between public-interest groups and industry. This suggests the potential for a whole new approach to the analysis of regulation. What are the private interests behind mandatory seat belt laws or air bags? What was the role of the U.S. airline industry in limiting U.S. landings of the Concorde? Clarkson, Kadlec, and Laffer (1979) claim that regulation was primarily to blame for Chrysler's recent financial difficulties; what was GM's role in this affair? Nuemann and Nelson (1982) argue that labor unions were a major force in implementing coal mine safety regulation, but not for the obvious reason. They claim that the purpose was to purge nonunion production (small mines) from the industry. Did the remaining firms, some of which gained from the regulation, join hands with the union to support the law? J. Smith (1982) rejects the public-interest theory of regulation of liquor stores. She claims that religious groups and others join with producers to effect regulation. Does this justify a different look at the Parker doctrine or licensing in general?

There is abundant evidence in the economics literature that when the flag of public interest is raised to support regulation, there is always a

private interest lurking in the background. There is hardly a regulatory program anywhere that does not benefit some industry or subset, most often at the expense of rivals or consumers. Antitrust authorities are mistaken to assume that just because a legitimate public-interest group supports regulation there cannot be anticompetitive results.

ANTITRUST LAW

Given the volume of research on the importance of private interests in affecting government in general and regulation in particular, it is surprising that there remains one large research area still haunted by the ghost of Pigou. This is the analysis of antitrust law. For example, Posner (1982) links the passage of antitrust law with other public-interest laws, such as statutes against murder. According to Joskow and Klevorick (1979), "The primary objective of antitrust policy is to promote full and fair market competition and to reap the benefit that competition brings with it" (p. 220). But there is another side; McGee (1980) states: "It may pay competitors to complain that someone is preying on them" (p. 300). Earlier in his writing Posner (1976) does an about face saying that the antitrust law is used to "harass competitors that have lower costs and otherwise frustrate the fundamental goals of antitrust policies" (p. 27). He may mean that the law was intended for one purpose but used for another. This implies a mistake in judgment by the forces behind the law. Rational expectations will not allow this explanation to apply to all laws or regulatory programs. As Stigler (1971) notes, "The fundamental vice of such criticism is that it misdirects attention" (p. 17). Just because the law or its advocates say that the law was intended for some purpose does not mean that is the actual purpose. Survey data is notoriously unreliable. Courts have recognized the problem and adopted the public posture that competition is to be protected, not competitors. It remains to be seen whether that is the case or not.

Bork (1978) bites the bullet. Antitrust law is a fertile breeding ground for firms to sow anticompetitive seeds where free-market forces fail to do so. Stone (1977) claims that 80 to 90 percent of all FTC investigations are begun at the request of the public. It would be nice to know how many of these are brought by firms competing with the alleged violator.

In sum, the power of antitrust law can be used by firms to limit the behavior of their rivals. It has not yet been sufficiently demonstrated whether the bulk of antitrust cases are pro- or anticompetitive. Smith (1982) says, "Mounting evidence . . . suggests that the correspondence between the stated objectives of regulatory legislation and the actual effects is sufficiently weak to lead one to seek out an alternative model of regulation which stresses the gains and losses to plural special interest groups" (p.

319). There remains a great deal of research to be accomplished. Shughart and Tollison (1985) survey the positive analysis of antitrust, which is but a small first step in this direction.[26]

Conclusions

The first of two major themes in this chapter stresses that anticompetitive or strategic use of regulation is pervasive. There is a lot of wealth at stake, and managers would be remiss in their fiduciary responsibilities if they ignored profits available through (legal) manipulation of governmental processes. The decision to invest resources in lobbying to prevent the entry of rivals, to form a regulatory cartel, or to impose costs on existing rivals does not differ materially from other decisions managers make on a daily basis.

Secondly, even though one may want the law to accomplish a specific goal, in practice it frequently achieves a different one. More importantly, if the law repeatedly accomplishes something other than its avowed purpose, then it is time to abandon the pretense that people cannot rationally anticipate ultimate effects. It will no longer suffice to adopt the position that "prior to regulation, no interest group . . . existed. But regulation created such an interest, one that subsequently fought hard against deregulation" (Noll and Owen 1983, p. 35). Although this view may be correct in selected circumstances, it cannot provide a general explanation for regulation. Hardly anyone would claim that managers do not rationally forecast markets for their new products. Sometimes they are wrong, but on average they are correct. What is the nature of regulation that exempts it from this same principle?

This points the way for fruitful research. What is the role of the politician as a regulatory entrepreneur? Does he function as a leader throwing regulatory stones into the water, hoping to cast a ripple of wealth redistribution on some innocent bystanders who then surface and offer votes or other thanks? Or in contrast, do managers and politicians work hand-in-hand developing wealth redistributions? To be specific, what has been the role of coal and oil producers and existing electricity producers in regulating the production of nuclear power? Evidence in the economics literature leads me to believe that these interest groups have not quietly watched as the NRC has slowly but surely put nuclear power on the back burner in the United States. Instead, intuition suggests that they have played a much more active part in eliminating their rival, especially when they have such a strong political ally as the antinuclear movement.

In sum, the literature has an abiding theme: Real political clout is one of the most important, if not the most important, determinants of regula-

tion. The reformers of regulation who came to Washington with newly elected President Reagan knew this, or at least they should have. The literature is compelling. This leaves us with one conclusion: As long as there is a monopoly on police power, there will be a strategic use of regulation. The original research presented in the rest of this volume helps answer some issues not previously addressed in the economics literature. While many of us may wish that more economically sensible policy changes had been made during the Reagan administration, at least we are all a bit wiser now in understanding the mechanisms at work.

Notes

1. I use the phrase strategic use of regulation only because it is used by so many others. I attach no special importance to strategic behavior as distinct from any other kind of behavior. In fact one of the main themes of this paper is to argue that there is no fundamental difference between the so-called strategic use of regulation and the day-to-day operations of the firm.

2. This is not always true. Whether regulation is outside the scope of antitrust enforcement depends on the regulation and who is doing the regulating. State regulation must meet certain procedural requirements as detailed in several Supreme Court decisions—Parker v. Brown, 317 U.S. 341 (1943) and California Retail Liquor Dealers' Association v. Midcal Aluminum, Inc., 445 U.S. 97 (1980), among others. Moreover, certain cartel arrangements are specifically excluded from scrutiny by legislative mandate. The FTC cannot expend resources to investigate agricultural cooperatives and Federal Marketing Orders, and labor unions are exempt by statute in the Clayton Act and the Norris-LaGuardia Act.

3. See Caves (1980) for a review of strategy and industry structure.

4. Compare this with Bork's definition of predation: "Predation may be defined, provisionally, as a firm's deliberate aggression against one or more rivals through the employment of business practices that would not be considered profit maximizing except for the expectation either that (1) rivals will be driven from the market, leaving the predator with a market share sufficient to command monopoly profits, or (2) rivals will be chastened sufficiently to abandon competitive behavior the predator finds inconvenient or threatening" Bork (1978, p. 144).

5. I take it as given that the reader is familiar with the public-choice literature and its cynical approach to politicians' behavior. Those who wish to become more acquainted with this literature can see Mueller (1976) for a survey. More recent work includes Landes and Posner (1975), Crain (1977), McCormick and Tollison (1978), and Becker (1983).

6. See Buchanan (1980) for a more thorough elaboration on this subject. The point can be made obvious with a simple example. Imagine two children who both spend time learning to spell and write. The first uses his skills to become a successful playwright. The second uses his skills as a lobbyist for the sugar industry, obtaining quotas on imported sugar. The first is profit seeking and the

second is rent seeking, but in the classroom the two activities are identical. Rent seeking as a cost turns on the individual's definition of waste.

7. The problem can be analyzed with externality theory but with a twist. Welfare analysis holds that pecuniary externalities do not disturb Pareto optimality. That is, my demand for cars, though it may affect the price you pay, does not cause price to diverge from its social cost. However, my demand for laws can harm you or help you, as in the case of tying arrangements, and the question of Pareto optimality depends on the effect of the law. In this case a pecuniary externality can disturb a Pareto optimality.

8. Eastern Railroad Presidents Conference v. Noerr Motor Freight, Inc., 365 U.S. 127 (1965). The court concluded that there was an "essential dissimilarity" between the two activities. This position appears to have been abandoned, or at least modified, in California Motor Transport Co. v. Trucking Unlimited, 404 U.S. 508 (1972).

9. United Mine Workers v. Pennington, 381 U.S. 657 (1965). The doctrine was extended even further to include such things as boycotts in NAACP v. Clairborne Hardware Co., 102 S. Ct. 3409 (1982). From a legal perspective the problem is compounded by the courts' indecisiveness on the antitrust character of state regulations. The Parker doctrine effectively exempts state regulations from antitrust scrutiny (Parker v. Brown, 317 U.S. 341 [1943]). However, recent decisions have altered this course and imposed necessary guidelines for exemption. For example see California Retail Liquor Dealers' Association v. Midcal Aluminum, Inc., 445 U.S. 97 (1980). Page (1981) analyzes the court's position.

10. Tullock (1980) addresses the question of whether the entire rent is consumed or not. However, his analysis is not based on rent seekers employing a Nash equilibrium strategy.

11. See also Anderson and Hill (1983). Their analysis further muddles the issue. Even competition for scarcity rents that are efficiently created can impose a social cost. This implies that what has previously been considered efficient enforcement of property rights by government may have hidden costs that make the efficiency claim suspect.

12. For example, see Pigou (1932).

13. See Posner (1974) for an old review of the economic theories of regulation and McGraw (1975) for a slightly more recent version.

14. For a sampler see Jarrell (1978), Pincus (1977), Guttman (1978), or any issue of the *Journal of Law and Economics* or the *Rand Journal of Economics*.

15. For additional analysis of the impact of regulation on the airline industry, see Keeler (1972), Jordan (1970), and La Mond (1976). La Mond analyzes airline regulation in the state of California. He concludes that the California Public Utility Commission behaved in much the same way as the CAB, protecting intrastate airlines from competition with one other.

16. Earlier research on electricity prices (Stigler and Friedland 1962) has reported no impact of regulation on prices. There is also a strand of the literature that looks at the political environment of regulation. That is, whether the method of selecting regulators has any impact. See Eckert (1973) on taxicabs and Crain and McCormick (1984) on utilities for example.

17. Outboard Marine Corp. v. Pezetel, 461 F. Supp. 384, 474, F. Supp. 168 (D. Del. 1978, 1979). See Schwartz (1980).

18. It is almost impossible to list all the research employing the economic theory of regulation, but two more examples should suffice to demonstrate that no quarter is given nor any asked. Benham and Benham (1975, p. 423) argue that regulations restricting the flow of information "may be one of the most effective politically acceptable methods available for constraining the behavior of suppliers and consumers in the desired direction [decreased competition and higher price]." Plott (1965) reports that the Oklahoma Dry Cleaning price-setting board works hand-in-hand with the industry trade association.

19. Not all of the empirical literature finds that regulation is profitable to the regulated firms. For example, Schwert (1977) reports that the value of New York Stock Exchange seats fell in the period preceding passage of the SEC Act in 1934. One explanation for this empirical anomaly is that a private cartel was already in place, but this begs the question of why regulation was ever passed in the first place. Tests of the public-interest theory of SEC regulation have not proved very successful.

20. For analysis of cross-subsidization, see Posner (1971) and Tuccillo (1977), among others. There is also the bureaucratic largesse approach to regulation as developed in Niskanen (1971) and Tullock (1965). These models of bureaucracy claim that regulation proceeds to maximize the size of the bureau.

21. Anderson and Tollison (1984) claim this is only part of the story. They argue that adult male laborers were the primary force behind, and beneficiaries of, the laws and that switching technologies for water-powered mills could be accomplished at low cost.

22. See also Buchanan and Tullock (1975).

23. For a description of the PSD ruling, see Alabama Power Co. v. Costle, 636 F. 2d 323 (1979) at 346. Schwert (1981) details the use of financial market analysis to assay the effects of regulation. Basically, according to the efficient markets hypothesis of modern finance theory, a security price incorporates at every moment all information available to the market. Therefore these prices, common stocks most notably, can be used as a benchmark to measure the impact of unexpected changes in regulation.

24. For a general theory of the setting of accounting standards that focuses on the strategic planning of firms, see Watts and Zimmerman (1978, 1979). Watts and Zimmerman (1979) report evidence that large accounting firms systematically support FASB rulings that will increase the wealth of their most important clients.

25. See Maloney, McCormick, and Tollison (1984) for more on this exportation effect of regulation. Easterbrook (1983) also discusses the problem in the context of federalism. He takes the view that state regulation should not be the concern of antitrust authorities because there is competition across states. His structuralist approach to competition does not allow for comparative advantage or immobile specialized resources across states.

26. See Clarkson and Muris (1981), Weingast and Moran (1983), and Stigler (1985) for instance.

References

Ackerman, Bruce A., and Hassler, William T. *Clean Coal/Dirty Air.* New Haven: Yale University Press, 1981.

Anderson, Terry L., and Hill, Peter J. "Privatizing the Commons: An Improvement?" *Southern Economic Journal* 50 (October 1983): 438–50.

Anderson, Gary M., and Tollison, Robert D. "A Rent-Seeking Explanation of the British Factory Acts." In *Rent Seeking and Appropriability,* edited by David Collander. Cambridge, Mass.: Ballinger Press, 1984.

Areeda, Phillip, and Turner, Donald F. "Predatory Pricing and Related Practices under Section 2 of the Sherman Act." *Harvard Law Review* 88 (1975): 679–712.

Barro, Robert T. "The Control of Politicians: An Economic Model." *Public Choice* 14 (Spring 1973): 19–42.

Becker, Gary S. "A Theory of Competition among Pressure Groups for Political Influence." *Quarterly Journal of Economics* 98 (August 1983): 371–400.

Becker, Gary S., and Stigler, George J. "Law Enforcement, Malfeasance, and Compensation of Enforcers." *Journal of Legal Studies* 3 (January 1974): 1–18.

Benham, Lee, and Benham, Alexandra. "Regulating Through the Professions: A Perspective on Information Control." *Journal of Law and Economics* 18 (October 1975): 421–48.

Besen, Stanley M. "The Economics of the Cable Television 'Consensus'." *Journal of Law and Economics* 17 (April 1974): 39–52.

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

Boyer, Kenneth D. "Equalizing Discrimination and Cartel Pricing in Transport Rate Regulation." *Journal of Political Economy* 89 (April 1981): 270–86.

Buchanan, James M. "Rent Seeking and Profit Seeking." In *Toward a Theory of the Rent-Seeking Society,* edited by J. M. Buchanan, R. D. Tollison, and G. Tullock. College Station: Texas A & M Press, 1980.

Buchanan, James M., and Tullock, Gordon. "Polluter's Profits and Political Response: Direct Controls versus Taxes." *American Economic Review* 65 (March 1975): 139–47.

Caves, Richard E. "Industrial Organization, Corporate Strategy, and Structure." *Journal of Economic Literature* 18 (March 1980): 64–92.

Clarkson, Kenneth W., Kadlec, Charles W., and Laffer, Arthur B. "Regulating Chrysler out of Business?" *Regulation* 3 (September/October 1979): 44–49.

Clarkson, Kenneth W., and Muris, Timothy J., eds. *The Federal Trade Commission Since 1970.* Cambridge: Cambridge University Press, 1981.

Comanor, William S., and Mitchell, Bridges M. "Cable Television and the Impact of Regulation." *Bell Journal* 2 (Spring 1971): 154–212.

Crain, W. Mark. "On the Structure and Stability of Political Markets." *Journal of Political Economy* 85 (August 1977): 829–42.

Crain, W. Mark, and McCormick, Robert E. "Regulators as an Interest Group." In *The Theory of Public Choice II,* edited by J. M. Buchanan and R. D. Tollison. Ann Arbor: University of Michigan Press, 1984.

Douglas, George W., and Miller, James C. III. *Economic Regulation of Domestic Air Transport Theory and Practice*. Washington, D.C.: Brookings Institute, 1974.

Easterbrook, Frank H. "Antitrust and the Economics of Federalism." *Journal of Law and Economics* 26 (April 1983): 23–50.

Eckert, R. D. "On the Incentives of Regulators: The Case of Taxicabs." *Public Choice* 14 (Spring 1973): 83–99.

Eckert, R. D., and Hilton, George. "The Jitneys." *Journal of Law and Economics* 15 (October 1972): 293–326.

Fischel, Daniel R. "Antitrust Liability for Attempts to Influence Government Action: The Basis and Limits of the Noerr-Pennington Doctrine." *University of California Law Review* 45 (Fall 1977): 80–123.

Greenberg, Edward. "Wire Television and the FCC's Second Report and Order on CATV Systems." *Journal of Law and Economics* 10 (October 1967): 181–92.

Guttman, Joel L. "Interest Groups and the Demand for Agricultural Research." *Journal of Political Economy* 86 (June 1978): 467–84.

Higgins, Richard, and McChesney, Fred. "Truth and Consequences: The Federal Trade Commission's Ad Substantiation Program." *International Review of Law and Economics* 6 (1986): 151–68.

Horwitz, Bertrand, and Kolodny, Richard. "The FASB, the SEC, and R & D." *Bell Journal* 12 (Spring 1981): 249–62.

Ippolito, Richard. "The Effects of Price Regulation in the Automobile Industry." *Journal of Law and Economics* 22 (April 1979): 55–90.

Jarrell, Gregg A. "The Demand for State Regulation of the Electric Utility Industry." *Journal of Law and Economics* 21 (October 1978): 269–98.

Johnson, Ronald N., and Libecap, Gary D. "Contracting Problems and Regulation: The Case of the Fishery." *American Economic Review* 72 (December 1982): 1005–22.

Jordan, William A. *Airline Regulation in America*. Baltimore: Johns Hopkins Press, 1970.

Joskow, Paul L., and Klevorick, A. K. "A Framework for Analyzing Predatory Pricing Policy." *Yale Law Journal* 89 (1979): 213–70.

Keeler, Theodore E. "Airline Regulation and Market Performance." *Bell Journal* 3 (Autumn 1972): 399–414.

Kelman, Steven. *What Price Incentives? Economists and the Environment*. Boston: Auburn House, 1981.

Kitch, Edmund, Issacson, Marc, and Kasper, Daniel. "The Regulation of Taxicabs in Chicago." *Journal of Law and Economics* 14 (October 1971): 285–350.

La Mond, A. M. "An Evaluation of Intrastate Airline Regulation in California." *Bell Journal* 7 (Autumn 1976): 641–57.

Landes, Elisabeth M. "The Effect of State Maximum-Hours Laws on the Employment of Women in 1920." *Journal of Political Economy* 88 (June 1980): 476–94.

Landes, William, and Posner, Richard. "The Independent Judiciary in an Interest-Group Perspective." *Journal of Law and Economics* 18 (December 1975): 875–901.

Linneman, Paul. "The Effects of Consumer Safety Standards: The 1973 Mattress Flammability Standard." *Journal of Law and Economics* 23 (October 1980): 461–80.

Maloney, Michael T., and McCormick, Robert E. "A Positive Theory of Environmental Quality Regulation." *Journal of Law and Economics* 25 (April 1982): 99–124.

Maloney, Michael T., McCormick, Robert E., and Tollison, Robert D. "Achieving Cartel Profits through Unionization." *Southern Economic Journal* 46 (October 1979): 628–34.

Maloney, Michael T., McCormick, Robert E., and Tollison, Robert D. "Economic Regulation, Competitive Governments, and Specialized Resources." *Journal of Law and Economics* 27 (October 1984): 329–38.

Marvel, Howard. "Factory Regulation: A Reinterpretation of Early English Experience." *Journal of Law and Economics* 20 (October 1977): 379–402.

Marvel, Howard, and Ray, Edward J. "The Kennedy Round: Evidence on the Regulation of International Trade in the United States." *American Economic Review* 73 (March 1983): 190–97.

Maurizi, Alex, Moore, Ruth L., and Shephard, Lawrence. "Competing for Professional Control: Professional Mix in the Eyeglass Industry." *Journal of Law and Economics* 24 (October 1981): 351–64.

McCormick, Robert. "The Cotton Dust Standard: The Private Interests." Unpublished, 1983.

McCormick, Robert E., and Tollison, Robert D. "Legislators as Unions." *Journal of Political Economy* 86 (February 1978): 63–78.

McGee, John S. "Predatory Pricing Revisited." *Journal of Law and Economics* 23 (October 1980): 289–330.

McGraw, Thomas K. "Regulation in America: A Review Article." *Business History Review* 49 (Summer 1975): 159–83.

Moore, Thomas G. "The Beneficiaries of Trucking Regulation." *Journal of Law and Economics* 21 (October 1978): 327–44.

Morrison, Steven A., and Newman, Robert J. "Hours of Operation Restriction and Competition among Retail Firms." *Economic Inquiry* 21 (January 1983): 107–14.

Mueller, Dennis C. "Public Choice: A Survey." *Journal of Economic Literature* 14 (June 1976): 395–433.

Niskanen, William A. *Bureaucracy and Representative Government*. Chicago: Aldine-Atherton, 1971.

Noll, Roger, and Owen, Bruce M. *The Political Economy of Deregulation*. Washington: American Enterprise Institute, 1983.

Nuemann, George R., and Nelson, Jon P. "Safety Regulation and Firm Size: Effects of the Coal Mine Health and Safety Act of 1969." *Journal of Law and Economics* 25 (October 1982): 183–200.

Oster, Sharon M. "An Analysis of Some Causes of Interstate Differences in Consumer Regulations." *Economic Inquiry* 18 (January 1980): 39–54.

Oster, Sharon M. "The Strategic Use of Regulatory Investment by Industry Subgroups." *Economic Inquiry* 20 (October 1982): 604–18.

Oster, Sharon M., and Quigley, John M. "Regulatory Barriers to the Diffusion of Innovation." *Bell Journal* 8 (Autumn 1977): 361–77.

Page, William H. "Antitrust, Federalism, and the Regulatory Process: A Recon-

struction and Critique of the State Action Exemption after Midcal Aluminum." *Boston University Law Review* 61 (November 1981): 1099–1138.

Peltzman, Sam. *Regulation of Pharmaceutical Innovation*. Washington, D.C.: American Enterprise Institute, 1974.

Peltzman, Sam. "Toward a More General Theory of Regulation.," *Journal of Law and Economics* 19 (August 1976): 211–40.

Pigou, Arthur C. *The Economics of Welfare*. London: Macmillan, 1932.

Pincus, Jonathan J. *Pressure Groups and Politics in Antebellum Tariffs*. New York: Columbia University Press, 1977.

Plott, Charles R. "Occupational Self-Regulation: A Case Study of the Oklahoma Dry Cleaners." *Journal of Law and Economics* 8 (October 1965): 195–222.

Posner, Richard A. "Taxation by Regulation." *Bell Journal* 2 (Spring 1971): 22–50.

Posner, Richard A. "Theories of Economic Regulation." *Bell Journal* 5 (Autumn 1974): 335–58.

Posner, Richard A. *Antitrust Law: An Economic Perpsective*. Chicago: University of Chicago Press, 1976.

Posner, Richard A. "Economics, Politics, and the Reading of Statutes and the Constitution." *University of California Law Review* 49 (Spring 1982): 263–91.

Salop, Steven C., ed. *Strategy, Predation, and Antitrust Analysis*. Washington: Federal Trade Commission, 1981.

Salop, Steven C., and Scheffman, David T. "Raising Rivals' Costs." *American Economic Review* 73 (May 1983): 267–71.

Schneider, Lynne; Klein, Benjamin; and Murphy, Kevin. "Government Regulation of Cigarette Health Information." *Journal of Law and Economics* 24 (December 1981): 575–612.

Schwartz, Louis B. "American Antitrust and Trading with State-Controlled Economies." *Antitrust Bulletin* 25 (1980): 513–55.

Schwert, G. William. "Public Regulation of National Securities Exchanges: A Test of the Capture Hypothesis." *Bell Journal* 8 (Spring 1977): 128–50.

Schwert, G. William. "Using Financial Data to Measure Effects of Regulation." *Journal of Law and Economics* 24 (April 1981): 121–58.

Shughart, William F, II, and Tollison, Robert D. "The Positive Economics of Antitrust Policy: A Survey Article." *International Review of Law and Economics* 5 (July 1985): 39–57.

Smith, Janet K. "An Analysis of State Regulations Governing Liquor Store Licenses." *Journal of Law and Economics* 25 (October 1982): 301–20.

Smith, Richard L., II. "Franchise Regulation: An Economic Analysis of State Restriction on Automobile Distribution." *Journal of Law and Economics* 25 (April 1982): 125–58.

Spence, A. Michael "Entry, Capacity, Investment, and Oligopolistic Pricing." *Bell Journal* 8 (Autumn 1977): 534–44.

Spiller, Pablo T. "The Differential Impact of Airline Regulation on Individual Firms and Markets: An Empirical Analysis." *Journal of Law and Economics* 26 (October 1983): 655–90.

Stigler, George J. "The Theory of Economic Regulation." *Bell Journal* 2 (Spring 1971): 3–21.

Stigler, George J. "The Origin of the Sherman Act." *Journal of Legal Studies* 14 (January 1985): 1–12.

Stigler, George J., and Friedland, Claire. "What Can Regulators Regulate? The Case of Electricity." *Journal of Law and Economics* 5 (October 1962): 1–16.

Stone, Alan. *Economic Regulation and Public Interest: The Federal Trade Commission in Theory and Practice*. Ithaca: Cornell University Press, 1977.

Temin, Peter. "The Origin of Compulsory Drug Prescriptions." *Journal of Law and Economics* 22 (April 1979): 91–106.

Tollison, Robert D. "Rent Seeking: A Review." *Kyklos* 35 (1982): 575–602.

Tuccillo, John. "Taxation by Regulation: The Case of Financial Intermediaries." *Bell Journal* 8 (Autumn 1977): 577–90.

Tullock, Gordon. *Politics of Bureaucracy*. Washington, D.C.: Public Affairs Press, 1965.

Tullock, Gordon. "The Welfare Costs of Tariffs, Monopolies, and Theft." *Western Economic Journal* 5 (June 1967): 224–32.

Tullock, Gordon. "Efficient Rent Seeking." In *Toward a Theory of the Rent Seeking Society*, edited by J. M. Buchanan, R. D. Tollison, and G. Tullock, College Station: Texas A & M Press, 1980.

Urban, Raymond, and Mancke, Richard. "Federal Regulation of Whiskey Labelling: From the Repeal of Prohibition to the Present." *Journal of Law and Economics* 15 (October 1972): 411–26.

Watts, Ross L., and Zimmerman, Jerold L. "Towards a Positive Theory of Accounting Standards." *Accounting Review* 53 (January 1978): 112–34.

Watts, Ross L., and Zimmerman, Jerold L. "The Demand for and Supply of Accounting Theories: The Market for Excuses." *Accounting Review* 54 (1979).

Weingast, Barry R., and Moran, Mark. "Bureaucratic Discretion or Congressional Control?" *Journal of Political Economy* 91 (1983).

Weiss, Roger. "The Case for Federal Meat Inspection Examined." *Journal of Law and Economics* 7 (October 1964): 107–20.

Williamson, Oliver E. "Wage Rates as a Barrier to Entry: The Pennington Case." *Quarterly Journal of Economics* 82 (February 1968): 85–116.

Williamson, Oliver E. "On Predatory Pricing." *Yale Law Journal* 87 (1978): 1337–1668.

Yandle, Bruce. "A Cost-Benefit Analysis of the 1981–1984 MPG Standard." *Policy Analysis* 6 (Summer 1980): 291–304.

Yandle, Bruce. "Bootleggers and Baptists." *Regulation* 7 (May/June 1983): 12–16.

Reprinted by permission from the Virginia Law Review Association.
This article appears in Vol. 75, March 1989.

STRUCTURE AND PROCESS, POLITICS AND POLICY: ADMINISTRATIVE ARRANGEMENTS AND THE POLITICAL CONTROL OF AGENCIES

Matthew D. McCubbins, Roger G. Noll,** and Barry R. Weingast****

IN 1977, Congress substantially revised the Clean Air Act,[1] the nation's flagship legislation on environmental policy. Many changes were considered, and among those that Congress adopted was an intricate redefinition of the procedures to be used by the Environmental Protection Agency (EPA) in making rules.[2] The Clean Air Act Amendments of 1970 (1970 Amendments), which had moved responsibility for air pollution regulation from the Public Health Service in the Department of Health, Education, and Welfare (HEW) to the newly minted EPA, set up the EPA's rulemaking procedures as "informal" with few procedural requirements and considerable decisional flexibility.[3] After extensive debate in both the 94th and 95th Congresses,[4] Congress changed this to a new hybrid process (more formal than "informal rulemaking" but less formal than "formal rulemaking") that requires a more elaborate written record and a clearer statement of agency intentions and of the bases for its decisions.[5]

As a reading of the committee reports and floor debates about these and similar proposals makes clear, legislators regard the choice of administrative structure and process as vitally important.[6] The legislative history of admin-

* Associate Professor, Department of Political Science, University of California, San Diego.

The authors thank Sharyn O'Halloran for her assistance on this Article. Mr. McCubbins and Mr. Weingast acknowledge the support of the National Science Foundation under grants SES-8811022 and SES-8617516.

** Department of Economics, Stanford University.

*** Senior Research Fellow, Hoover Institution, Stanford University.

[1] Clean Air Act Amendments of 1977, Pub. L. No. 95-95, 91 Stat. 685 (codified as amended at 42 U.S.C. §§ 7401-7642 (1982 & Supp. IV 1986)).

[2] See id. § 305, 91 Stat. at 774 (codified at 42 U.S.C. § 7607(d) (1982)); see also H.R. Rep. No. 294, 95th Cong., 1st Sess. 27, *reprinted in* 1977 U.S. Code Cong. & Admin. News 1077, 1105 ("This section establishes comprehensive procedures for most informal rulemaking under the Clean Air Act in lieu of the Administrative Procedure Act.").

[3] Clean Air Act Amendments of 1970, Pub. L. No. 91-604, § 3, 84 Stat. 1676, 1677 (codified as amended at 42 U.S.C. § 7607(d) (1982)).

[4] See infra notes 121-56 and accompanying text.

[5] See Clean Air Act Amendments of 1977, Pub. L. No. 95-95, § 302, 91 Stat. 685, 770 (codified at 42 U.S.C. § 7607(d) (1982)).

[6] As used here, "process" refers to the rules and standards that apply to policy decisions by an agency and guide judicial review, whereas "structure" refers to the allocation of resources

istrative reforms contains discussion of two issues that also surface in the scholarly literature. As emphasized by public administration and organization theory scholars, legislators are concerned about agency efficiency in collecting and evaluating relevant information to guide administrative decisions, and in implementing these decisions.[7] Also, as emphasized by scholars of administrative law, legislators debate fairness and the protection of individual rights when considering alternative institutional arrangements.[8]

In addition, the choice of structure and process is guided by political concerns. Specifically, legislators see the choice of administrative structures and processes as important in assuring that agencies produce policy outcomes that legislators deem satisfactory. Structure and process are regarded as important in determining the relative influence of different interests in the decisionmaking process, as well as the balance of influence between the President and Congress. The purpose of this Article is to contribute to the positive political theory of the structure and process of administrative agencies.

In a previous article,[9] we explained our view of how elected political leaders design administrative procedures. We asserted that administrative procedures are one means of guiding agencies to make decisions that are consistent with the preferences of the legislative coalition (including the President) that succeeded in passing the agency's enabling legislation, but without requiring the members of that coalition to monitor, or even be aware of, the nature of the agency's actions.[10] This Article advances the discussion in three ways.

First, we explicitly take into account that elected political officials may differ over their desired policy outcomes. In so doing, this Article more clearly identifies the kinds of internal coalitional problems that structure and process might be expected to solve. Specifically, structure and process can be viewed as embodying an ex ante agreement among legislators and the

---

and decisional authority among agencies and within an agency. Examples of process are rules of standing and evidence and the assignment of burdens of proof, whereas a flow chart depicting the sequence of actions and identifying the associated actors would reveal examples of structure. Most often, structure refers to "veto gates"—those points in the process where policy can be killed—and which actors control them.

[7] See, e.g., J. Charlesworth, Governmental Administration 45-86 (1951); J. Millett, Government and Public Administration: The Quest for Responsible Performance 63-250 (1963); J. Pfiffner & R. Presthus, Public Administration 443-536 (1960); J. Shafritz & A. Hyde, Classics of Public Administration (1987); J. Shafritz & J. Ott, Classics of Organization Theory (1987).

[8] See, e.g., J. Shafritz & A. Hyde, supra note 7; Berger, Administrative Law After Forty Years, 33 Fed. B. News J. 297 (1986).

[9] McCubbins, Noll & Weingast, Administrative Procedures as Instruments of Political Control, 3 J.L. Econ. & Org. 243 (1987).

[10] Id. at 253-55.

President that limits the ability of each to engage in ex post opportunistic behavior. Because each has an incentive to influence the agency to make policies that are not consistent with the coalition's legislative agreement, all can be better off if they employ means to limit their own opportunism.

Second, this Article identifies and analyzes a key problem with the use of traditional methods of oversight and rewards or punishments to assure agency compliance with the policy preferences of the winning coalition. Specifically, this is the problem of "history dependence" or "reactive enforcement" in legislative processes. For reasons elaborated below, the outcome of a legislative attempt to rectify an act of noncompliance by an agency will not, in general, reproduce the policy outcome that was sought by the winning coalition, even if the preferences of the members of the legislative body remain unchanged. Thus, effective political control of an agency requires ex ante constraints on the agency (that is, a means of restricting the agency's decisionmaking before it actually makes policy choices), one source of which is manipulation of its structure and process.

Third, to illustrate the principal lines of argument, this Article traces the organizational and procedural history of air pollution regulation in the United States. The discussion of air pollution regulation is not intended to be comprehensive; indeed, several excellent books have been written on the subject.[11] Instead, we provide examples of structural and procedural reforms considered by Congress, many of which were never enacted, that exemplify our general conceptual model.

I. Structure, Process, and Bureaucratic Compliance

In examining the problem of assuring agency compliance with the desires of the political coalition enacting and overseeing legislation, we make use of three related theoretical advances of the 1970s: principal-agent theory,[12] perfect equilibrium,[13] and structure-induced equilibrium.[14]

Principal-agent theory applies to circumstances when one person (the principal) arranges for another (the agent) to take an action that is beneficial

---

[11] See, e.g., B. Ackerman & W. Hassler, Clean Coal/Dirty Air (1981); R. Melnick, Regulation and the Courts (1983).

[12] See Holmstrom, Moral Hazard and Observability, 10 Bell J. Econ. 74 (1979) (considering the role of imperfect information and moral hazard in principal-agent relationships).

[13] See Selten, Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games, 4 Int'l J. Game Theory 25 (1975) (discussing the concept of a perfect equilibrium where the equilibrium is robust to slight mistakes).

[14] See Shepsle & Weingast, Structure-Induced Equilibrium and Legislative Choice, 37 Pub. Choice 503 (1981) (explaining the properties of legislative institutions necessary for the existence of equilibrium and offering an alternative view of institutions based upon majority rule).

to the principal but costly to the agent, under circumstances when the principal cannot perfectly and costlessly enforce an ex ante promise by the agent to act in the best interests of the principal. For the purposes of this Article, the principal in the problem of bureaucratic compliance is the coalition enacting a new policy and establishing a structure and process for implementing it. The agent is the bureau that is to implement the policy.

The standard solution to a principal-agent problem is: first, for the two actors to agree on a compensation schedule that the principal will implement on the basis of the outcome of the agent's actions; and second, for the principal to engage in costly monitoring to enable the principal to assess the agent's performance. The counterpart in political processes is oversight: investigations into the performance of an agency, sometimes in the context of the annual budgetary process, and occasionally as part of the reauthorization of an agency's programs. In addition, both Congress and the President have "watchdog" agencies to monitor agency performance, such as the Office of Management and Budget and the General Accounting Office.

An important disadvantage of the investigative oversight process is its cost. The time of political officials and their staffs is a valuable resource, and normal oversight consumes it. An alternative is to set up a system in which someone else (that is, a third party outside of the principal-agent diad) monitors the agent and reports acts of noncompliance. In political processes, the object of legislation is to deliver benefits to the policy's target group. If politicians make it easy for this group to detect and report noncompliance, they need not use their own resources in monitoring the agency. Instead, a politician who was a member of the coalition that enacted a program can rely on "fire alarms" sounded by the targeted beneficiaries as a mechanism to trigger formal investigations and/or legislative responses to noncompliance.[15]

A fire alarm converts the oversight job of a politician from active monitor to reactive servant of affected constituencies, and fits nicely with the observation that the role of a modern legislator is more like that of an ombudsman than a policymaker.[16] But the effectiveness of the fire alarm depends on the credibility of political officials when they threaten to punish an agency that is not complying with the wishes of its overseers. That is, if an agency prefers to adopt a policy that differs significantly from the preference of the coalition that enacted its program, it can be dissuaded from doing so only if its political overseers are able to undo its actions and/or punish it.

---

[15] See McCubbins & Schwartz, Congressional Oversight Overlooked: Police Patrols Versus Fire Alarms, 28 Am. J. Pol. Sci. 165 (1984).

[16] See M. Fiorina, Congress: Keystone of the Washington Establishment 42-49 (1977); Fiorina & Noll, Majority Rule Models and Legislative Elections, 41 J. Pol. 1081, 1101 (1979).

For several reasons, agencies may not regard threats of punishment as credible, especially if the threats are from Congress. Most of these reasons are discussed in our earlier article,[17] and we will not reprise them here. Instead, we focus on two points. First, the President's role in appointing the top administrators of agencies offers him an advantage in influencing policy, especially for agencies in the executive branch where the top administrators serve at the pleasure of the President. Hence, we expect that an enduring theme in Congress will be to build in protections against undue influence by the President. Second, Congress and the President cannot rely on their ability to correct a noncomplying decision by an agency through legislative action, whether by changing its mandate, its structure, its procedures, or its budget. The reason is to be found in the theory of structure-induced equilibrium and, in particular, in the dependence of legislative processes on the details of the status quo that legislative action is supposed to alter.

A. Potential Manipulation by the Agency

To see why Congress and the President cannot rely solely on the threat of new legislation to force agency compliance, we consider the following example.[18] Although our example is very simple, it illustrates the general proposition that different status quo policies produce different legislative outcomes.

To begin, we assume that the game between politicians and agencies is played only once—that is, the agency chooses a policy that the politicians can either accept or reject by passing new legislation. Consider a three-person unicameral legislative body picking a policy in a two-dimensional policy space, as depicted below in Figure 1. Each member is assumed to have common knowledge about the structure of the game and the preferences of all other players. Anticipating the analysis of air pollution regulation in Part II of this Article,[19] the dimensions could be the stringency of environmental policy (for example, how much pollution to abate) and the extent to which antipollution policy will be structured to advantage established (as opposed

---

[17] See McCubbins, Noll & Weingast, supra note 9, at 248-53.

[18] The example that follows is in the spirit of the model developed by Hammond, Hill, and Miller and provides the analysis necessary to understand the problem of manipulation of politicians by their agents. As we show below, however, the appropriate interpretation of this model is that it characterizes the problem to be solved, not the actual pattern of interaction between an agency and politicians. See T. Hammond, J. Hill & G. Miller, Presidential Appointment of Bureau Chiefs and the "Congressional Control of Administration" Hypothesis (March 30, 1986) (unpublished paper delivered at a meeting of the American Political Science Association).

[19] See infra notes 36-148 and accompanying text.

to entering) production facilities. The points labeled *1*, *2*, and *C* represent the most-preferred policies (ideal points) of the three legislators.

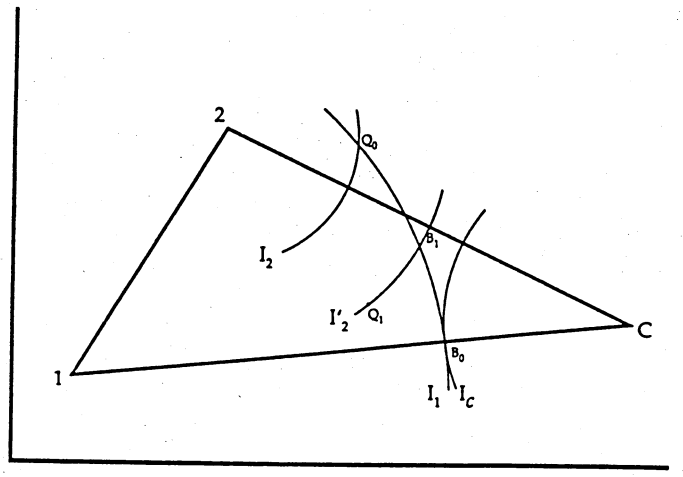


Figure 1

We also assume that the loss of welfare to each member as the policy moves away from that person's ideal point is proportional to the distance from that ideal point (that is, their indifference contours can be represented as circles). If this is the case, the triangle defined by the lines connecting the three ideal points contains all of the Pareto optimal outcomes—that is, the policy choices that cannot be changed without making one of the three members worse off. The point Q_0 represents the status quo, which is the policy that will be in place if the legislature takes no action. Finally, member *C* is a one-person committee that has jurisdiction over the policy under consideration. We assume for simplicity that member *C* possesses a monopoly on making proposals to change the status quo, and that these proposals are considered by the legislature (that is, *C* together with members *1* and *2*) under a "closed rule" wherein the members simply vote for *C*'s proposal or the status quo.[20]

Assuming that each member seeks to move policy closer to his or her ideal point, the best strategy available to member *C* is to propose an alternative that comes as close to point *C* as possible while still obtaining the support of at least one other legislator, thereby defeating the status quo by a vote of at

[20] Closed rules (i.e., where a member's ability to offer amendments to a bill on the floor is restricted) are not common in the House of Representatives and are virtually never adopted in the Senate. Although the details of the rules pertaining to a bill clearly influence the nature of legislative outcomes, the basic processes at work are similar for the closed rule and produce qualitatively similar results. See Weingast, Floor Behavior in Congress: Committee Power Under the Open Rule, 83 Am. Pol. Sci. Rev. (forthcoming Sept. 1989).

least two to one. *C* can find this point by locating the points in the triangle that either *1* or *2* regards as equally desirable as the status quo. We have drawn indifference contours (I_1 and I_2) for both legislators *1* and *2* that pass through status quo point Q_0. As is apparent from inspection, point B_0 is the best alternative for member *C*. If *C* proposes a bill slightly closer to member *1* than point B_0, the proposal will defeat Q_0 by a vote of two to one. Members *1* and *C*, therefore, can be regarded as having formed a coalition to enact B_0, which then becomes the new status quo.

The problem of a noncomplying agency is depicted on the diagram as follows. Suppose that the agency implements not B_0 but Q_1, which thus becomes the new status quo point. Then, in the next session of the legislature, the best proposal that *C* can hope to enact is bill B_1, which is passed by a new coalition between *C* and *2*. (I_2' is the indifference contour for legislator *2* when the status quo point is Q_1.) Although B_1 is preferred to B_0 by member *2*, members *1* and *C* are worse off than at outcome B_0. Moreover, by slight modifications in the diagram, one can depict circumstances in which agency noncompliance makes any given member, or any combination of two members, worse off than at outcome B_0.

The important insight from this example is that, even with perfect monitoring of agency noncompliance, no legislative remedy is available to the original coalition that will restore its original agreement.[21] By establishing a new status quo, a noncomplying agency has broken apart the coalition that gave rise to its initial mandate. Of course, this phenomenon is quite general—agency noncompliance in other directions away from B_0 inflicts a different pattern of gains and losses and causes different coalitional responses. Hence, each party, including the legislator who is not a member of the original coalition, stands some chance of loss from agency noncompliance. If legislators are risk averse,[22] unpredictability in the nature of agency noncompliance will be regarded by all as undesirable. Thus, they will all have an incentive to develop some means of assuring compliance other than correction of errors after they are observed.

So far, our analysis has focused entirely on members of a legislature, but with a few changes it can be altered to represent the policy tug-of-war among the House, the Senate, and the President. For purposes of exposition, we

---

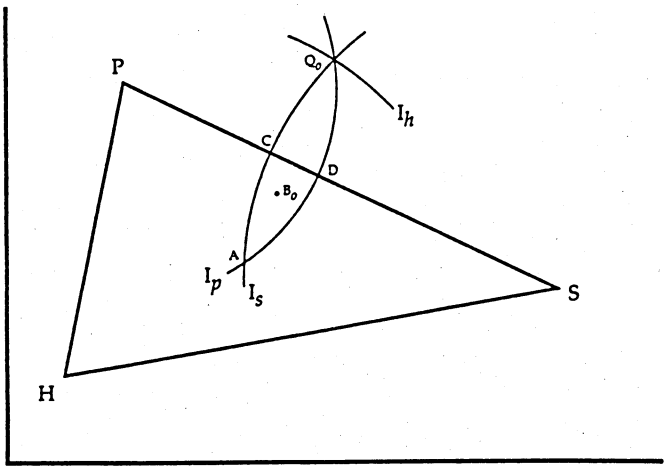
[21] The unqualified nature of this conclusion depends on the one-shot assumption noted above. When the relationship is repeated, more complicated patterns of interaction can arise.

[22] Legislators are likely to behave as if they are risk averse, even if they are personally risk neutral, if their constituents punish unpredictable policy choices or their reelection probability is nearly unity. See L. Cohen & R. Noll, Intertemporal Policy Preferences of a Legislator (Center for Economic Policy Research, Stanford University 1984). See generally, D. Mayhew, Congress: The Electoral Connection 13-17 (1974) (explaining how legislative behavior is affected by the quest for reelection).

will ignore the problem of intrachamber choice depicted in Figure 1, and assume that each legislative body is homogeneous, with a unique collective ideal policy and set of indifference curves. The point of this analysis will be to illustrate the significance of the quite different institutional rules governing policy formation among the two branches of the legislature and the President. The key points for our analysis are that the House and Senate bargain over the contents of a legislative proposal, and that the President has veto power. We will ignore the possibility of a veto override in order to retain simplicity in the example.

Figure 2 depicts the ideal points of the House (*H*), Senate (*S*), and President (*P*), the status quo point (Q_0), and the indifference curve of each that

Figure 2



passes through Q_0. The lens-shaped figure defined by points Q_0, *A*, *C*, and *D* defines the possible policy outcomes that are preferred by all three. Because each has a veto (that is, the bill must pass both chambers and be signed by the President), the new policy outcome must be located within this "lens." Point *D* represents the best possible feasible outcome for the Senate, Point *A* is the best that the House can do, and Point *C* is the most desirable feasible outcome for the President. The half-lens defined by points *A*, *C*, and *D* contains all possible final bills.[23] The bargaining process and the strength of the

[23] The proof of this assertion is as follows: (1) for any proposal outside of the triangle *PHS* there are proposals inside the triangle that make all three members better off (recall that *PHS* contains all Pareto optimal policies); (2) for any proposal within *PHS*, but outside the half-lens, at least one member is worse off than if Q_0 is retained, and that member will veto the proposal.

three participants will determine which point within the figure *ACD* will be selected. For example, if the House and Senate can convince the President that no second bill will emerge from them if their proposal is vetoed, then they can bargain over points along the President's indifference curve passing through Q_0 (here, curve *AD*). However, if the President, knowing that Q_0 is regarded as undesirable by the House and Senate, expects a veto to lead to a bill closer to *P*, a veto threat can successfully force the House and Senate to propose a policy near *C*. If none of the three participants is in a dominant position, yet none of them is sufficiently weak to be forced to gain nothing from the bill, some intermediate point, such as B_0, will be enacted.

Once a policy is enacted, the agency must implement it. In so doing, the agency may shift the policy outcome away from the legislative intent (here B_0). Two important observations can be made about the consequences of policy drift. First, as long as the agency stays within the triangle *PHS*, *no legislative correction or punishment is possible.* Any policy outcome within the triangle (that is, within the Pareto optimal set) must be preferred by at least one of the three parties to the original agreement to enact B_0. Hence, because all three actors have a veto, one of the three will not agree to an action that forces the agency back to B_0.[24] Second, if the agency causes policy to drift outside the triangle, all three can agree that a correction and punishment are deserved; however, the new policy is not likely to be B_0. The reason is that the set of feasible policies (the half-lens *ACD*) is uniquely determined by the exact nature of the status quo. Hence, the agency's implemented policy outcome is not likely to cause the President and the legislature to agree on B_0 as the correction, unless the agency has specifically attempted to implement the former status quo, Q_0.

The issue of protecting against agency noncompliance has three components. First, if political actors are risk averse, all three will prefer greater certainty in policy implementation as compared to random noncompliance (that is, noncompliance that may drift away from the preferred outcome of each of the three). Second, each of the three wants to minimize the chance that one of the other two will influence the agency against its interests. Nonetheless, all have an ex post incentive to spend resources persuading the agency to sway policy their way. This is a negative sum game, so ex ante all three actors regard such expenditures as wasteful. Third, none of the parties wants to let the agency choose which political actor to favor. The power to

---

[24] Note in this situation that the President is in a more powerful position than the Congress. By exercising the power to fire heads of agencies and to issue executive orders, the President can influence policy without obtaining the agreement of the House and Senate. This opportunity for effective ex post response to noncomplying behavior implies that Congress is likely to be more concerned about structure and process than is the President.

choose is the power to manipulate, holdup, and extract. Politicians would not willingly subject themselves to such behavior by the agency.

B. Structural Constraints Placed on an Agency

The logic of the preceding argument is that the most effective means for achieving policy stability are constraints on the flexibility of agencies, rather than reliance on rewards, punishments, and oversight. If the agreement within each legislative body and among the President and the two houses of Congress can be clearly articulated in terms of policy outcomes, the best solution is legislative specificity: writing into the law precisely what the agency is to achieve, and how it is to do so. If the best policy from the perspective of the winning coalition depends on arcane information or is uncertain because of frequent changes in the state of knowledge about the problem that the policy is supposed to ameliorate, however, legislative specificity cannot identify the policy outcome that is embodied in the legislation.[25] This does not mean that the agency necessarily must be free to violate the spirit of the coalitional agreement. An alternative means of achieving the policy outcome that the coalition would have adopted in the absence of uncertainty is to constrain an agency's policies through its structure and process by enfranchising the constituents of each political actor—members of Congress and the President—that is a party to the agreement to enact policy B_0.

There are two fundamental ways in which an agency's structure and process can influence its policy decisions. First, because policy decisions depend upon what information is available to the agency, structure and process determine the quantity, quality, and completeness of available information and the extent to which policy decisions must be supported by this information. Political principals can control the influence of a constituency by using structure and process to affect the dependence of the agency on information the constituency supplies.

One example is the role of agency resources in enfranchising poorly organized constituents. An agency that has sufficient resources to generate its own information about the consequences of its decisions, available funds to subsidize the participation in its processes of various poorly organized interests, and a relatively lenient standard for judicial review of its actions (for example, arbitrary and capricious), will be far less dependent on highly organized, well-represented interests than an agency that lacks resources and faces a

---

[25] See Mashaw, Prodelegation: Why Administrators Should Make Political Decisions, 1 J.L. Econ. & Org. 81 (1985).

high standard for upholding its decisions in court.[26] Similarly, if cause and effect relationships are uncertain, the assignment of the burden of proof (for example, must a product be proven safe before marketing, as is the case with drugs, or proven dangerous to be prevented from being marketed, as with chemicals) will also affect the influence of different constituencies (for example, chemical companies versus environmentalists), and hence the ultimate policy outcome.

The second way structure and process encourage compliance is by preventing the agency from presenting political principals with a fait accompli and instead forcing it to warn them well in advance of any potentially noncomplying decision. A common observation about administrative procedures is that they cause delay by requiring agencies to follow intricate and often cumbersome decisionmaking processes.[27] The courts are undoubtedly a major source of these procedures, reflecting their attempts to protect individual rights of due process. But there is more to administrative delay than that. Even without court-sponsored procedures, Congress and the President typically want administrative procedures because of the political role they play. Indeed, legislation often imposes procedural complexities that go beyond what the courts have required.

As argued above, when an agency presents politicians with a fait accompli, politicians may find it difficult, if not impossible, to respond. Legislation can reverse the agency, but not before a new constituency is mobilized in support of the new policy. Moreover, some members of the coalition giving rise to the original legislation may actually prefer the agency's decision and oppose reversing it. As long as the agency has been careful to choose a policy within the set of Pareto optimal policies, complete reversal is essentially impossible, and only a partial correction is likely even if the new policy is not Pareto optimal.

---

[26] For a more complete development of these ideas, see Noll, The Political Foundations of Regulatory Policy, *in* Congress: Structure and Policy 462 (1987).

[27] The "problem" of delay has been much criticized by various private and government reports. See, e.g., Federal Regulation: Roads to Reform, 1979 A.B.A. Comm'n L. & Econ. 92 (criticizing "cumbersome" administrative procedures that result in delay); Staff of Senate Comm. on Gov't Affairs, 95th Cong., 1st Sess., Delay in the Regulatory Process ix (Comm. Print 1977) (A Committee poll of about one thousand lawyers practicing regularly before eight major commissions showed that "undue delay" was the most frequently cited major problem with federal regulation. The Committee concluded that "[d]elay is a fundamental impediment to the effective functioning of regulatory agencies."); The President's Advisory Council on Executive Org., A New Regulatory Framework: Report on Selected Independent Regulatory Agencies 5 (1971). Interestingly, none of these reports discusses why Congress and the President have persistently ignored recommendations to end delay by streamlining regulatory processes.

Administrative procedures erect a barrier against an agency carrying out such a fait accompli by forcing the agency to move slowly and publicly, giving politicians (informed by their constituents) time to act before the status quo is changed. Consider, for example, the requirements of formal rulemaking.[28] Before it can issue a change in policy, an agency subject to formal rulemaking must first announce that it is considering a policy change and solicit the views of all relevant parties.[29] Often, it is required to announce a "provisional" rule and again solicit still more comments.[30] In addition, formal rulemaking requires the agency to conduct a trial-type hearing, allowing interested persons to testify and to cross-examine witnesses.[31] Only then can it implement a new policy by issuing the long-sought new rule. Finally, the agency must produce a record setting forth substantial evidence in favor of its finding and reasons for rejecting alternative findings.[32]

These procedures allow politicians to prevent deviations before they occur.[33] The members of the coalition enacting the policy can adopt a blanket agreement to inhibit all possible deviations while the nature of the deviation is still in doubt and the coalition has not yet formed that might support the deviation. Delay gives the old coalition time to mobilize its constituents before the agency undermines it by enunciating a noncomplying policy that changes the status quo.

---

[28] Formal rulemaking is required by the Administrative Procedure Act (APA) where the agency-enabling legislation requires rules to be made "on the record after an opportunity for an agency hearing." Administrative Procedure Act, § 553, 5 U.S.C. § 553(c) (1982); see, e.g., Food, Drug, and Cosmetic Act, § 301, 21 U.S.C. § 371(e) (1982) (procedures for formal rulemaking by the Food and Drug Administration (FDA)); see also United States v. Allegheny-Ludlum Steel Corp., 406 U.S. 742, 757 (1972) (ruling that the APA's formal rulemaking provisions, 5 U.S.C. §§ 556-557, govern rulemaking proceedings only if the agency's enabling statute, in addition to providing for a hearing, prescribes explicitly that it be "on the record"); United States v. Florida E. Coast Ry., 410 U.S. 224 (1973) (holding that no hearing was required under the language of the Interstate Commerce Act).

[29] See 5 U.S.C. § 553(b)-(c) (1982); see, e.g., 21 U.S.C. § 371(e)(1)-(2) (1982) (FDA procedures).

[30] See, e.g., 21 U.S.C. § 371(e)(2).

[31] See 5 U.S.C. § 556(d); see, e.g., 21 U.S.C. § 371(e)(3).

[32] See 5 U.S.C. §§ 556(e), 557(e); see, e.g., 21 U.S.C. § 371(c)(3).

[33] Moreover, agencies are prevented from mobilizing a new constituency behind a different policy prior to any decision. Because such actions provide clear evidence that the agency was prejudiced and had made its "decision" prior to completing its procedures and weighing all the evidence, they are easily reversed by the courts. Cf. Crowell v. Benson, 285 U.S. 22, 48 (1931) (finding that ex parte investigation and facts not put into evidence will not support an agency order); NLRB v. A. Sartoris & Co., 140 F.2d 203, 205 (2d Cir. 1944) ("[I]f an administrative agency ignores all the evidence given by one side in a controversy and with studied design gives credence to the testimony of the other side, the findings would be arbitrary and not in accord with the legal requirement.").

One low cost route to achieving this end is by attaching appropriations riders that prevent agencies from spending money on ongoing cases or investigations that could cause deviations from the status quo. Because appropriations bills encompass large collections of specific programs, they facilitate cooperation among politicians to thwart all deviations even though each member might benefit from some of them.[34] If numerous actions preventing deviations are included in the same bill, the overall incentive to stabilize policy and maintain agreements can override individual incentives on any one issue.

To summarize, the potential for agency deviations from intended policies that are difficult for politicians to punish or correct leads them to devise institutions that limit an agency's ability to deviate. We have argued that two aspects of structure and process play this role. The first creates a decisional environment that causes the agency to be responsive to the constituency interests that were represented in the enacting coalition. The second delays agency policymaking. Although costly to all, delay enables politicians to act to prevent deviations while the coalitional agreement is still the status quo.

There is an important distinction between this view of structure and process and that emphasized in the literature about administrative law and organizational design. Typically, the normative content of that literature deals with achieving efficient outcomes within a process that protects the rights of people who are likely to be affected by a policy decision. Obviously, these issues constitute an important part of a rational political leader's problem in constructing an agency for policy implementation. Protecting constitutional rights is necessary to prevent the courts from invalidating policy and, all else equal, greater efficiency means more deliverable policy output for affected constituents.

The theoretical arguments presented here provide a third piece of the puzzle of agency design. If policies are inherently conflictual, they necessarily will produce winners and losers. An agreement to change policy is an agreement to favor some constituencies over others. Hence, part of the challenge of agency design is for the members of the coalition to use structure and process to cause the decisions of the agency to be more responsive to the constituencies that the policy is intended to favor and to maintain the political compromises negotiated at the time of enactment. Specifically, we would expect agency design to exhibit three characteristics.

---

[34] For examples of the use of appropriations riders and informal controls through appropriations hearings, see W. Cary, Politics and the Regulatory Agencies 35 (1967); M. Kirst, Government Without Passing Laws 64 (1969); Weingast & Moran, Bureaucratic Discretion or Congressional Control? Regulatory Policymaking by the Federal Trade Commission, 91 J. Pol. Econ. 765, 775 (1983).

First, the agency's structure and process should create a political environment that *mirrors* the politics at the time of enactment; that is, interests that are active participants in the debate over the original legislation should be given representation through the structure and process of the agency so that each will be protected against undesirable policy drift. Specifically, the enabling legislation should seek to combine sanctions with an institutional structure to create pressures on agencies that replicate the political pressures applied when the relevant legislation was enacted. Here, the point of administrative procedures is not to pre-select specific policy outcomes but to create a decisionmaking environment that mirrors the political circumstances that gave rise to the establishment of the policy. Although political officials may not know what specific policy outcome they will want in the future, they will know which interests ought to influence a decision and what distributive outcomes will be consistent with the original coalitional arrangement.

Second, the structure and process of an agency should *stack the deck* in favor of the groups who, among those significantly affected by the policy, are also favored constituents of the coalition that caused the policy to be adopted. And third, agency policies should exhibit an *autopilot* characteristic in the sense that as the preferences of the constituencies enfranchised in the agency's structure and procedure change, so too will the agency, freeing Congress and the President from having to enact new legislation to achieve that end.

The implication of this is not that policy is necessarily stable, but that it will change only to the extent that either the preferences of the agency's enfranchised constituencies change or a constituency simply withers away and no longer takes advantage of its structural and procedural advantages. In either case, the agency's political overseers are not likely to care about a drift in policy. Because their preferences reflect the preferences of their constituencies, the disappearance of a constituency or a change in its policy desires will make the old policy obsolete in any case. Thus, the problem of noncompliance is not that policy drifts, but that it drifts in ways that are harmful to the constituents of a member of the coalition that enacted the policy.

C. Noncompliance and the Courts

The federal courts, as well as agencies, can be a source of noncomplying policy outcomes. Vague legislative mandates and weak standards for judicial review give courts an opportunity to shape policy as they see fit. Hence, political actors face a similar problem in trying to limit judicial decisions to the intent of the political coalition that gave rise to the policy.

Although we have not developed a comprehensive theory of political control of the policies promulgated through judicial review, some elements of

the theory pertaining to agencies also apply to the judiciary. Most importantly, ex post reestablishment of a coalitional agreement, after a judicial opinion has upset the status quo, is likely to be difficult.[35] Like an agency, a court is safe from legislative reversal as long as its new policy is within the Pareto optimal set established by the preferences of the House, the Senate, and the President. As with agencies, one potential means of protecting against judicial readjustment of policy is to use either explicit legislation or administrative procedures in an attempt to constrain judicial decisions. The problematic aspect, however, is precisely how judicial opinions can be constrained, especially those rendered by the United States Supreme Court, where the only check upon nakedly noncomplying policymaking is legislative correction. Unlike agencies, or even lower courts, the Supreme Court lacks an external standard for its own decisions other than legislative response. This suggests that elected political officials are likely to be less effective with the Supreme Court than with agencies in using ex ante constraints to prevent direct contravention of their policy preferences.

II. Regulating Air Pollution

The history of federal air pollution legislation provides a natural experiment for illustrating and testing the ideas of Part I of this Article. Since 1955, the methods of regulating air pollution have changed many times, including four times during the decade from 1961 through 1970.[36] Importantly for our purposes, much of this revision consisted of changes in the structure and process through which policies were developed and enforced. Moreover, many of the policy changes in the 1970s followed court decisions that were, arguably, unanticipated by members of the policymaking branches. Thus, these changes exemplify the type of policy drift, as in Figure 2, that is likely to be uncorrectable by members of the legislative coalition.

In this Part we conclude that the policy choice made by the enacting legislative coalition in the 1970 Amendments was unrecoverable after the courts intervened by requiring prevention of significant deterioration (PSD), and that because of the structure of preferences, the courts' policy could not be substantially amended by the policymaking branches. What remained for members of the policymaking branches, then, was to reform the structure and process of EPA decisionmaking to protect themselves from further

---

[35] Marks's recent work studies in detail the problem of judicial review of legislation, focusing on the conditions when politicians can reinstate and judicial change in policy. See B. Marks, A Model of Judicial Influence on Congressional Policymaking: Grove City College v. Bell (Nov. 1988) (Hoover Institution Working Papers in Political Science P-88-7).

[36] See infra notes 40-54 and accompanying text.

seemingly random policy shocks. We examine some aspects of these reforms in Part III of this Article.

The argument we make for the above conclusion takes four steps. First, in order to identify the underlying dimensions of the policy choice, we explore the early evolution of federal air pollution regulation. This exploration also allows us to identify the status quo policy as of the passage of the 1970 Amendments. Second, we argue that the courts' interpretation and the subsequent EPA promulgation constituted an unanticipated policy change. Third, we identify the preferences of the members of each policymaking branch and show how their preferences shaped their response to the court's intervention. And fourth, we describe the nature of the new policy equilibrium.

A. First Step: The Evolution of Federal Clean Air Policy

Two abiding policy issues have been at center stage in the history of air pollution regulation: federalism and the tradeoff between stringency and economic development (that is, what level of pollution will be tolerated and who will bear the costs—established industries and areas or new industries and undeveloped areas). The federalism question pertains to the appropriate division of responsibilities among federal, state, and local authorities. The core of the economic tradeoff is that, generally speaking, economic development goes hand-in-hand with increased air pollution. Hence, if polluted areas are to be made cleaner, the cost is some reduction in economic activity, and if pristine areas are to remain clean, the cost is a sacrifice of some economic growth.

Between 1955 and 1970, both issues were slowly being resolved. The responsibility for air pollution control slowly migrated to the federal government, reflecting a gradual evolution of the views of state and local government officials. The latter, though by no means unanimously or unambiguously favoring an expanded federal role, came to understand two primary difficulties in relying on decentralized (that is, state as opposed to federal) regulation. The first was the presence of informational economies of scale. Key aspects of environmental regulation involve assessing the damages caused by pollution, the relationship between emissions and air quality, and the costs of abatement. Not only is the current state of knowledge on these issues complex and voluminous, it is also evolving rapidly. Hence, a regulator faces a formidable task in developing and maintaining a reasonably complete knowledge base for informing regulatory decisions. Federalization centralizes this responsibility and avoids duplicating essentially the same informational activities in numerous jurisdictions. Indeed, the first federal legislation in this area, the 1955 Air Pollution Control Act, dealt exclusively

with this problem.[37]

The second difficulty arising from decentralized regulation was the possibility that localities would compete for industries by offering more relaxed regulation. The problem with decentralized air pollution regulation can be regarded as a "prisoner's dilemma" in the following sense. All localities may prefer clean air, and if all localities simultaneously enact rigorous standards few industries will have any incentive to relocate. But the community that acts first may impose such sufficiently high costs on its local industries that they close their facilities, either moving elsewhere or being displaced by companies located in communities that have not yet acted.[38] The two key features of this problem are the difficulties of coordinated action and the necessity of overcoming the incentive facing each community to be a little more lax than the others in regulating its industries to give itself a competitive edge. Federalization of regulation attacks these problems. A federal regulator can impose regulations simultaneously on all communities and can inflict punishments on communities that do not make reasonable efforts to enforce these standards.[39]

The first step in federalizing regulatory authority was a modest one—the Clean Air Act of 1963 authorized the Secretary of HEW to take legal actions in a very limited range of circumstances.[40] A more significant step was taken in 1965, when HEW was given the authority to establish national emissions standards for automobiles.[41] This was only a partial assertion of federal jurisdiction because the national standards were to be a minimum—

---

[37] The 1955 Air Pollution Control Act authorized HEW and the Public Health Service to conduct and sponsor scientific studies of air pollution and to coordinate information collection to assist state and local authorities in controlling air pollution. See Air Pollution Control Act, ch. 360, § 1, 69 Stat. 322, 322 (1955) (codified as amended at 42 U.S.C. § 7401 (1982)).

[38] The argument about industry relocation is developed in Pashigian, Environmental Regulation: Whose Self-Interests are Being Protected?, 23 Econ. Inquiry 551 (1985). Pashigian shows that the political problem is even more complicated if not all localities or regions want the same level of pollution control (e.g., if growing regions such as the Southwest are willing to sacrifice control for growth while slow-growing or stagnant areas prefer strong controls). Id. at 552-54.

[39] To achieve this, of course, requires giving federal officials coercive powers either over officials at other levels of government who are responsible for regulating air pollution or directly over the sources of pollution.

[40] The Secretary was authorized to take administrative or legal actions to deal with air pollution that posed a significant threat to human health if state and local actions proved ineffective, if the problem involved interstate pollution (after consultation with state officials), or if the Governor of the affected state or a state air pollution agency requested help from the Secretary. Clean Air Act, Pub. L. No. 88-206, § 5(c), 77 Stat. 392, 396 (1963) (codified as amended at 42 U.S.C. § 7401 (1982)).

[41] See Motor Vehicle Air Pollution Control Act, Pub. L. No. 89-272, § 202(a), 79 Stat. 992, 992-93 (1965) (codified as amended at 42 U.S.C. § 7521 (1982)).

they did not displace any more stringent standards that might be adopted by a state.

Shortly thereafter, the issue of nearly complete assertion of federal authority was raised by President Lyndon Johnson, who in 1967 proposed that HEW be given the authority to set national uniform emissions standards for specific pollutants.[42] Congress did not go as far as the President proposed, but it did further federalize regulation in two ways. First, it preempted state regulation of automobile emissions in all states except California.[43] Second, it authorized HEW to regulate air pollution in any locality when a state had failed to do so and when pollution in the area posed a threat to public health.[44]

The 1970 Amendments further reallocated responsibilities to the federal government.[45] By 1970, state and local government officials generally subscribed to the view that centralized responsibility for setting air quality standards was desirable.[46] This was largely a response on the part of state governments to their own failure to tackle the political issues involved.

The expanding role of the federal government brought to the fore the issues of stringency and cost. The 1970 Amendments authorized the newly created Environmental Protection Agency to set uniform national ambient air quality standards and emissions standards for new stationary sources of pollutants.[47] They also authorized the EPA to set emissions standards for both new and old sources of hazardous air pollutants.[48] The 1970 Amendments required states to adopt plans that would enable them to implement the EPA's requirements and authorized the EPA to impose a plan on a state that did not develop an adequate plan on its own.[49] The EPA was further empowered to take legal action against violators of its standards, leading to fines and/or imprisonment.[50] The 1970 Amendments defined new source

---

[42] See 23 Cong. Q. Almanac 875 (1967).

[43] See Air Quality Act of 1967, Pub. L. No. 90-148, § 2, 81 Stat. 485, 501 (codified as amended at 42 U.S.C. § 7543 (1982)); see also 23 Cong. Q. Almanac 875 (1967) ("The [Act] provided that automobile exhaust standards could be issued only by the Federal Government, except for California, which was permitted to enforce its own (and more stringent) control standards. No other state was given this authority.").

[44] Air Quality Act of 1967, Pub. L. No. 90-148, § 2, 81 Stat. 485, 491-97 (codified as amended at 42 U.S.C. § 7415 (1982)).

[45] See Clean Air Amendments of 1970, Pub. L. No. 91-604, §§ 107-116, 84 Stat. 1676, 1678-89 (codified as amended at 42 U.S.C. §§ 7407-7416 (1982)).

[46] See 28 Cong. Q. Weekly Rep. 973 (1970).

[47] Clean Air Amendments of 1970, Pub. L. No. 91-604, §§ 109, 111, 84 Stat. 1676, 1679-80, 1683-84 (codified as amended at 42 U.S.C. §§ 7409, 7411 (1982)).

[48] Id. § 112, 84 Stat. at 1685-86 (codified as amended at 42 U.S.C. § 7412).

[49] Id. §§ 110, 113, 84 Stat. at 1680-83, 1686-87 (codified as amended at 42 U.S.C. §§ 7410, 7413).

[50] Id. § 113, 84 Stat. at 1686-87 (codified as amended at 42 U.S.C. § 7413).

performance standards (NSPS) in terms of the technology required, not in terms of emission levels.[51] They defined NSPS as a "standard for emission of air pollutants which reflects the degree of emissions limitation achievable through the application of the best system of emission reduction which (taking into account the cost of achieving such reduction) the [EPA] Administrator determines has been adequately demonstrated."[52]

Finally, the EPA was given the authority to regulate the composition of fuels for mobile sources of pollutants if the fuels posed a threat to public health or damaged pollution control devices.[53] What remained for state and local authorities was to write emissions standards for existing stationary sources that achieved national air quality requirements and to enforce these standards and the standards for new sources promulgated by the EPA.[54]

The slow evolution of federal regulation of air quality reflected the underlying politics of the period. First, the growing constituency for environmental policy put ever-increasing pressure on all levels of government to make progress in cleaning the air. Second, improving the air required imposing costs on sources of pollution. Moreover, because significant improvements required that firms develop new technologies for abatement, in some cases industry faced a fixed development cost that was unlikely to affect its production costs and, therefore, was unlikely to be completely recovered later in price increases or productivity advancements. Hence, significantly more stringent environmental regulation was regarded as an unrecoverable wealth transfer away from the owners and employees of affected industries. And if the regulations were imposed discontinuously—in the form of a one-time substantial cost shock—some firms could be bankrupted and some facilities closed. Needless to say, political officials would be wary of imposing such shocks on their constituents. Thus, the problem in designing an institution for promulgating air pollution control policies was to make progress on the air quality front while not disrupting local economies.

Because the policy problem was shrouded in uncertainty concerning the costs of pollution and its abatement—including the amount of abatement that was technologically feasible—the regulatory process would benefit from flexibility in responding to new information. But with flexibility comes the possibility of noncomplying behavior by either of two sets of regulators: the federal officials promulgating national standards or the state and local offi-

---

[51] Id. § 111(a)(1), 84 Stat. at 1683 (codified as amended at 42 U.S.C. § 7411(a)(1)).

[52] Id.

[53] Id. § 211(c)(1), 84 Stat. at 1698-99 (codified as amended at 42 U.S.C. § 7545(c)(1)).

[54] Id. § 110, 84 Stat. at 1680-83 (codified as amended at 42 U.S.C. § 7410); see also H.R. Conf. Rep. No. 1783, 91st Cong., 2d Sess. 45, *reprinted in* 1970 U.S. Code Cong. & Admin. News 5374, 5377-78 (report accompanying H.R. 17,255 describing the states' role in the adoption and enforcement of implementation plans).

cials making and enforcing implementation plans. Hence, capturing the benefits of flexibility while minimizing the risk of noncompliance required the development of a process that would prevent both a drift away from progress toward cleaner air and disruption of local economies.

B. Second Step: Policy Innovation by the Courts and the Evolution of PSD

Soon after the passage of the 1970 Amendments, the EPA promulgated regulations regarding degradation of air quality in areas with clean air.[55] The EPA concluded that the intent of Congress was for the agency to establish a uniform nationwide air quality standard that would not require a strict rule against degradation of air quality in pristine regions.[56] The Sierra Club appealed this decision to the United States District Court for the District of Columbia, which eventually ruled that the EPA had to reject state implementation plans (SIPs) that did not prevent the degradation of air quality in clean air regions.[57] The ruling caused the EPA to reinstate the nondegradation policies it had previously discarded. Because of the ruling, the EPA replaced the uniform national ambient air quality standards it had promulgated pursuant to the 1970 Amendments with a multiplicity of standards based not on health or welfare considerations but on each region's actual air quality.[58]

The legislative history of the 1970 Amendments provides scant basis for believing that Congress intended the EPA to promulgate rules to prevent significant deterioration. Judge Pratt of the District of Columbia District Court based his *Sierra Club* ruling on five items in the legislative history of the 1970 Amendments,[59] of which only two were actually legislative in origin, and of these two only one actually preceded the 1970 Amendments.[60] In his decision, Judge Pratt referred to a sentence in the 1970 report of the Senate Committee on Environment and Public Works (Senate Public Works Committee) stating that the EPA should disapprove SIPs for clean air areas that did not "provide, to the maximum extent practicable, for the continued

---

[55] See National Primary and Secondary Ambient Air Quality Standards, 40 C.F.R. § 50 (1972); Requirements for Preparation, Adoption, and Submittal of Implementation Plans, 40 C.F.R. § 51.

[56] These views were encapsulated in a regulation promulgated by the EPA in the wake of the 1970 Amendments requiring only that states not allow air degradation to rise above applicable secondary standards. See 40 C.F.R. § 51.12(b).

[57] Sierra Club v. Ruckelshaus, 344 F. Supp. 253 (D.D.C. 1972), aff'd per curiam by an equally divided Court sub nom. Fri v. Sierra Club, 412 U.S. 541 (1973).

[58] R. Melnick, supra note 11, at 71.

[59] 344 F. Supp. at 255.

[60] For an excellent discussion of the legislative history relied on in the *Sierra Club* litigation, see R. Melnick, supra note 11, at 76-80.

maintenance of such ambient air quality."[61] This language could be interpreted as implying PSD, but the report went on to state:

> Once such national goals are established, deterioration of air quality should not be permitted *except* under circumstances where there is no available alternative. Given the various alternative means of preventing and controlling air pollution—including the use of the best available control technology, industrial processes, and operating practices—and care in the selection of sites for new sources, land use planning and traffic controls—deterioration need not occur.[62]

A more reasonable interpretation of this passage is that the Senate Public Works Committee did not intend to prevent degradation of clean air areas, but rather intended to require polluters to use the "best available control technology" (BACT).[63] This interpretation would also be consistent with other discussions of the 1970 Amendments in the report.

Had Congress intended a nondegradation policy, one would expect that the legislative response to *Sierra Club* would have been a relatively noncontroversial one, featuring at best a few perfecting changes in legislation or some discussions of the details in oversight hearings. Instead, the debate about PSD continued for five years before a very detailed set of PSD rules was finally enacted as a compromise between very different bills passed by the House and Senate.[64] It is thus reasonable to conclude that the court's decision was a policy innovation that was unanticipated by the policymaking branches.

1. Implementation of PSD

In the absence of strict guidelines from either the court[65] or Congress, the EPA responded to the ruling in *Sierra Club* by promulgating regulations to

---

[61] *Sierra Club*, 344 F. Supp. at 255 (quoting S. Rep. No. 1196, 91st Cong., 2d Sess. 2 (1970)).

[62] S. Rep. No. 1196, 91st Cong., 2d Sess. 11 (1970) (emphasis added).

[63] For example, the Senate bill proposed that new stationary sources be "designed, built, equipped, operated, and maintained so as to reduce emissions to a minimum." The performance standards were to be met through the "application of the latest available emission control technology or through other means of preventing or controlling air pollution." Id. at 15-16.

[64] See Clean Air Act Amendments of 1977, Pub. L. No. 95-95, §§ 160-178, 91 Stat. 685, 731-51 (codified at 42 U.S.C. §§ 7470-7508 (1982)). Compare S. 252, 95th Cong., 1st Sess., 123 Cong. Rec. 18,517 (1977) with H.R. 6161, 95th Cong., 1st Sess., 123 Cong. Rec. 16,966 (1977).

[65] Indeed, as Melnick points out, the court failed to define what constitutes "significant" deterioration. See R. Melnick, supra note 11, at 77.

prevent significant deterioration.[66] These regulations incorporated three basic features.

First, the EPA defined significant deterioration and established guidelines for the development and submittal of SIPs with respect to PSD. In no instance did the proposed rules define significant deterioration in terms other than concentrations of sulfur dioxide and particulate matter.[67] Further, the determination of what actually constituted significant deterioration was left to the states under this promulgation, with the EPA Administrator retaining the authority to assess whether they conformed to the federal standard.[68]

Second, the EPA provided for the designation of air quality areas under three classifications: class I applied to "areas in which practically any change in air quality would be considered significant"; class II applied to "areas in which deterioration normally accompanying moderate well-controlled growth would be considered insignificant"; and class III applied to "areas in which deterioration up to the national [ambient air quality] standards would be considered insignificant."[69] The proposed regulations specified the allowable increments in poilutant concentrations over baseline air quality concentrations for class I and class II regions. Areas designated as class III were limited to concentrations of particulate matter and sulfur dioxide no greater than the national ambient air quality standards.[70] The deterioration increments in class I and II areas were established by emissions ceilings that could not be exceeded by any new major source. The determination of allowable air quality increments permitted reductions of emissions from existing sources to be taken into account in determining the unused portion of the allowable air quality increment.[71] States could reclassify areas, subject to EPA review and approval, to allow for the introduction of sources not otherwise compatible with the initial classification if the resulting deterioration would not be significant.[72] EPA approval of proposed redesignations would be based on the record of the state's hearings.[73]

Third, the EPA modified the review procedures for new polluting sources by adding considerable detail to the NSPS requirements. New sources, wherever located, would be reviewed to determine their impact on air qual-

---

[66] See 40 C.F.R. § 52.21 (1975).

[67] Id. § 52.21(c)(2).

[68] Id. § 52.02.

[69] Environmental Protection Agency Approval and Promulgation of Implementation Plans: Prevention of Significant Air Quality Deterioration, 39 Fed. Reg. 30,999, 31,003 (1974) (EPA Administrator's Aug. 15, 1974 introduction to proposed changes to 40 C.F.R. § 52.21).

[70] 40 C.F.R. § 52.21(c)(2)(i)-(ii) (1975).

[71] Id. § 52.21(d)(2)(i)-(ii).

[72] Id. § 52.21(c)(3)(ii).

[73] Id. § 52.21(c)(3)(ii)(d).

ity. Preconstruction review would be applied to proposed facilities in nineteen specific major source categories.[74] The list of sources subject to review was expanded to include two additional source types: fuel conversion plants and primary lead smelters.[75] The regulations further specified that no construction or modification of new sources could commence unless the EPA Administrator determined that the effect on air quality would not lead to a violation of air quality increments.[76] All new or modified sources were required to meet an emissions limit representing the level of emissions reduction achieved by the application of BACT.[77] In determining whether the new source met the BACT requirement, the EPA Administrator would consider fuels and raw materials available, the respective costs of the application of new control techniques, process changes, or alternative fuels, and any applicable state and local emissions limitations.[78]

A source that was modified but did not increase pollution other than sulfur oxides or particulate matter, or was modified to utilize an alternative fuel, was not subject to the new regulations. This provision exempted modifications of such sources as coal cleaning plants, pulp mills, cement plants, primary zinc smelters, iron and steel mills, primary aluminum ore reduction plants, and primary copper smelters from the requirements of the regulation.[79] As a general rule, though, the EPA disapproved SIPs that did not prevent "significant deterioration of air quality in any portion of any State where the existing air quality is better than the national ambient air quality standards."[80]

In effect, the regulations promulgated by the EPA required that air quality not be allowed to deteriorate, even if the air quality of a region greatly exceeded national standards. Except in areas with especially bad air quality, established sources generally were not required to achieve as great a degree of emissions abatement as new sources. State implementation plans for nonattainment areas were required to establish regulatory strategies for existing sources so that an air quality region could achieve national standards. New sources were required to adopt the best possible control technologies. Moreover, in nonattainment areas they were further required to abate other existing sources so that the net effect of the construction of new

---

[74] Id. § 52.21(d)(1)(i)-(xix).

[75] Id. § 52.21(d)(1)(xvii)-(xviii).

[76] Id. § 52.21(d)(2)(i).

[77] Id. § 52.21(d)(2)(ii). In this case, BACT was defined to mean "any emission control device or technique which is capable of limiting emissions to the levels proposed or promulgated." Id. § 52.01(f). Where no standards of performance existed, BACT was to be determined on a case-by-case basis considering six enumerated factors. Id.

[78] Id. § 52.01(f)(1), (4), (5).

[79] Id. § 52.21(d)(1)(i)-(vii).

[80] Id. § 52.21(a).

sources and abatement of existing sources was an improvement in air quality. In both PSD regions and nonattainment areas, NSPS were required; however, in PSD regions a new source did not necessarily have to abate other sources to achieve zero net emissions. But to escape the zero (or close to zero) net emissions requirement, the owner of the new source had to show that the proposed new facility would not cause a significant deterioration of air quality.[81]

C. *Third Step: The Political Response to the Court*

The court's ruling in *Sierra Club*, together with the EPA's interpretation of the court's requirements in its promulgated regulations, established a new status quo policy.[82] Here we explore the logic of the proposition that Congress was blindsided by the court on the PSD issue and examine whether subsequent developments are consistent with this proposition. If PSD was a surprise, its effect was an unanticipated movement from a legislative policy agreement (B_0 in figures 1 and 2) to some new status quo point. Here the source of the movement was judicial policymaking, not agency noncompliance. Implicit in this account is that the EPA was correct when it concluded that it lacked legislative authority to develop PSD rules.

Further, if the court's change in policy is also in the set of Pareto optimal outcomes, no significant changes in this policy can be enacted because each institutional actor holds a veto power over proposed changes in the status quo. Thus, the subsequent legislative enactment of basically the same PSD requirements that emerged from the court and the EPA's subsequent rulemaking provides *no evidence* that PSD was the original intention of Congress in 1970.[83]

1. *The Preferences of the Players*

To shed further light on the PSD issue requires that we examine the policy preferences of the important actors involved in forging the 1970 Amendments. To do so, we examine not only the specific issue of PSD but also the passage of the 1970 Amendments that were determined by the court to require PSD, and Congress's legislative response in 1976 and 1977 to the adoption of PSD.

---

[81] Id. § 52.21(d).

[82] Sierra Club v. Ruckelshaus, 344 F. Supp. 253 (D.D.C. 1972), aff'd per curiam by an equally divided Court sub nom. Fri v. Sierra Club, 412 U.S. 541 (1973); see 40 C.F.R. § 52 (1975).

[83] This conclusion parallels Marks's exploration of what kinds of inferences can be drawn about politicians' preferences and intentions based on their subsequent reactions to judicial changes in policy. See Marks, supra note 35.

The incentives of political actors are forged through the electoral connection.[84] Different politicians confront different reelection problems,[85] so their incentives and actions will be shaped somewhat differently. Indeed, this was the intent of the Constitutional Convention in creating different electoral systems—different constituencies and different terms of office—for Congressmen, Senators, and the President. By giving the members of the policymaking branches somewhat different ties to the people, the founding fathers sought to ensure that the theoretical separation of powers was maintained in practice.[86]

The differences in constituencies and terms of office among elected officials are likely to cause several types of differences in their concerns about policies. First, all else equal, smaller constituencies generally have less diversity of interest and enterprise. This implies that representatives with smaller constituencies are more likely to see an issue in terms of a single overriding interest. Thus, House members are more likely to be spokesmen for certain industries or groups than Senators whose constituencies are, with a few exceptions, larger than those represented by House members from their states. Because Senators face larger constituencies and have more media attention (and because many of them aspire to the Presidency), they are more likely to be worried about national goals and the larger picture rather than specific industries. Smaller states are likely to be less diversified, so their Senators are more likely to be champions of a single interest. The President, representing a national constituency, will be the least tied to specific industries, groups, or sectors.

Second, the shorter term for members of the House implies that they will apply a higher discount rate than will the President or Senators when considering the long-term costs and benefits of government policy. Because House members are essentially continuously standing for reelection, they are more likely to favor programs that have immediate payoffs and delayed costs.

In order to evaluate the positions of members of the House and Senate and of the President, we need to identify the underlying dimensions of the PSD

---

[84] This line of argument is now standard in the political science literature on Congress. See, e.g., R. Fenno, Home Style: House Members in Their Districts 171-213 (1978) (suggesting how members of Congress first build, then maintain, a "reliable reelection constituency"); M. Fiorina, Congress: Keystone of the Washington Establishment 39 (1977) (asserting that "the primary goal of the typical congressman is reelection"); D. Mayhew, supra note 22, at 5 (1974) (describing United States Congressmen as "single-minded seekers of reelection").

[85] They face different districts with different primary and reelection constituencies. Because of this, their relationships with their constituents will differ and the actions they take will be predictably different. See R. Fenno, supra note 84, at 1-29.

[86] See The Federalist No. 39, at 254-55 (J. Madison) (C. Van Doren ed. 1945).

tradeoff. The two dimensions are who is to bear the costs of air pollution regulation and the magnitude of the environmental benefits. This translates into how much of the cost is to be imposed on established firms and how much is to be born by new enterprises and undeveloped areas, together with how stringent the requirements will be.

The position of the President is most difficult to determine ex post because he was less involved in the legislative process. Because of their national constituency, Presidents are inclined to be more concerned about economic efficiency and less about protecting specific industries. In the case of air pollution legislation, this suggests that the President will pay less attention to the relocational aspects of environmental regulation, and more to minimizing the costs of achieving a given policy objective.

President Nixon was relatively favorably inclined toward environmental regulation, having created the EPA through a Reorganization Plan[87] and having proposed much of the ultimate contents of the 1970 Amendments.[88] Under Nixon, however, the EPA rejected PSD in 1971. No doubt this reflected the President's view, for the EPA decision was not followed by an executive proposal to correct the issue in the Clean Air Act so that PSD rules could be promulgated. Indeed, the Nixon Administration's response to the court ruling was to propose an amendment to the Clean Air Act eliminating the PSD requirement.[89] It seems reasonable to conclude that Nixon favored somewhat looser overall environmental regulation than that favored by the median (Democratic) member of Congress, but did not favor the elaborate use of air pollution regulation to preserve the existing pattern of industrial development, especially at high cost.

The preferences of relevant House and Senate actors are somewhat easier to discern. The arguments presented in Part I of this Article lead us to begin with an examination of the policy preferences of the oversight committees in both chambers, for congressional rules of procedure give them considerable influence in determining the ultimate legislative outcome.[90] In the House of

---

[87] Reorg. Plan No. 3 of 1970, 3 C.F.R. 1072 (1970), *reprinted in* 5 U.S.C. app. at 1132 (1982), *and in* 84 Stat. 2086-87 (1972).

[88] See The President's Message on the Environment, Feb. 10, 1970, 116 Cong. Rec. 32,908.

[89] See 39 Fed. Reg. 42,509, 42,510 (1974) ("The Administration has submitted for consideration an amendment to the Act which would eliminate the requirement for preventing significant deterioration of air quality.").

[90] The influence of committees is standard wisdom in congressional literature. See D. Mayhew, supra note 22, at 85-97. The powers underlying this influence include the following three procedures: (1) the monopoly right, held by the oversight committee, to introduce legislation (i.e., a first-mover advantage); (2) the right of committee members to revise amendments proposed by members not on the committee; and (3) the practice of populating conference committees with members of the committees that have jurisdiction over the bill. These rules give committees two vetoes over legislative actions—one prior to legislative

Representatives, jurisdiction over environmental legislation resides in the Committee on Interstate and Foreign Commerce (House Commerce Committee). The responsibilities of this committee include oversight of many regulatory agencies and the Department of Commerce. In both 1970 and 1977 its Chairman was Harley O. Staggers (Democrat, West Virginia). It is hardly surprising, then, that the House has been steadfast in using NSPS for electric utilities to protect the West Virginia coal industry.

The overall membership of the House Commerce Committee exhibits two additional features. Historically, it has been rather pro-business, having had oversight responsibilities for the host of regulatory agencies that gave rise to the capture-cartel theory of regulation,[91] and most of its members have been from industrialized, populous states. Of the thirty-six members in 1970, fourteen were from the industrial Northeast and Midwest and four more were from eastern coal-mining states. In addition, five were from California or Texas, sunbelt states with serious urban air pollution problems.[92] By 1977, the House Commerce Committee was even more unbalanced: twenty-two of its forty-three members were from the industrial Northeast and Midwest, three more were from eastern coal areas, and nine were from California or Texas.[93] Thus, approximately two-thirds in 1970 and three-fourths in 1977 of the oversight committee members could be expected to be deeply concerned about the economic dislocation effects of air pollution regulation.

For the entire House, concern for protecting specialized local economies is likely to be strong. Conversely, long-term policies, with costs now but benefits down the road, are likely to be less popular in the House than in the Senate. Hence, the propensity of the oversight committee to be concerned about established industries, though more targeted on traditional manufacturing than the House as a whole, would likely be broadly consistent with the general inclinations of all House members. From these observations, it is

---

consideration and one afterwards in conference. These vetoes offer committee members special advantages. For a discussion of the effects of these rules, see Shepsle & Weingast, The Institutional Foundations of Committee Power, 81 Am. Pol. Sci. Rev. 85 (1987).

[91] In contrast to the traditional view that bureaus attempt to serve the general welfare are theories that view bureaus as servants of well-defined interests, either because they were set up to serve specific clients (the cartel theory), or because, through the years, they become vulnerable to being taken over by some special interest (the capture theory). For a survey of capture-cartel theories, see Noll, Government Regulatory Behavior: A Multidisciplinary Survey and Synthesis, *in* Regulatory Policy and The Social Sciences 9, 24-28 (R. Noll ed. 1985).

[92] This was somewhat counterbalanced in 1977 because Senator Malcolm Wallop of Wyoming (a major site of western coal) was on the Public Works Committee's Subcommittee on Environmental Pollution. See 1977 U.S. Code Cong. & Admin. News lxxxv.

[93] See 1970 U.S. Code Cong. & Admin. News cxiii.

likely that the idea of PSD was congenial to members of the House in both 1970 and 1977.

In the Senate, jurisdiction over environmental policy resides in the Senate Public Works Committee. The jurisdiction of this committee is wholly different from that of its House counterpart, for public works is a traditional federal pork barrel. Its primary task is to authorize federal construction projects like rivers, harbors, and reclamation. Because these programs are especially important in less populated states, the membership of the Senate Public Works Committee differs from that of the House Commerce Committee.

In 1970, only four of the fifteen members of the Senate Public Works Committee were from the industrial Northeast and Midwest (counting Senator Thomas Eagleton of Missouri—a debatable classification), and none were from California or Texas. Its main point of commonality with the membership of the House Commerce Committee was eastern coal: three Senators represented eastern coal states, including the Chairman, again from West Virginia, Jennings Randolph. Among the remainder of the Senate Public Works Committee were members from Maine, New Mexico, Alaska, Kansas, and Oregon, states with at best minor problems with air pollution from industrial sources; and from Virginia, Florida, and North Carolina, sunbelt states with some track record of attracting new industry away from the Northeast.[94] By 1977, membership on the committee had shifted even further away from northeastern influence. Although Randolph was still Chairman, only two of the fifteen members represented eastern coal, and only two members were from the industrial Northeast. Largely rural states had a clear majority, with nine members from the following states: Maine, Vermont, Alaska, North Dakota, Iowa, Colorado, Idaho, New Mexico, and Wyoming.[95]

It is difficult to imagine that a committee with this composition would have strongly favored protecting eastern manufacturing industries; only Randolph's position as Chairman makes it plausible that the Senate Public Works Committee favored protecting eastern coal.[96] Indeed, its membership and jurisdiction indicate that the committee would have preferred economic development in clean air areas rather than a strict PSD policy.

The Senate as a body differs from the House in two important ways. First, in comparison with the House, the Senate overrepresents the sparsely populated states of the Great Plains and the West[97]—the kinds of states that

---

[94] See 1977 U.S. Code Cong. & Admin. News cxvii.

[95] See 1970 U.S. Code Cong. & Admin. News xc.

[96] See 1977 U.S. Code Cong. & Admin. News lxxxv.

[97] Relative to the House, malapportionment in the Senate is extreme. For example, a Senator from California represents over fifty times as many voters as a Senator from Wyoming.

favor federal development projects and that are the mainstay of the Senate Public Works Committee. Second, in most states, Senators have a more diversified constituency and so are less responsive to the needs of particular industries. In addition, their longer terms politically enable Senators to take a longer view of the benefits and costs of policy actions. On balance, these factors make the Senate somewhat less inclined to stress stringent regulation of air quality in relatively less polluted areas, and less inclined to worry about possible adverse effects of the regulation on specific plants and industries. To the extent the latter issue was important, the Senate would be less inclined to favor using environmental regulation to keep traditional manufacturing industries in the Northeast.

The implication of this analysis is that the Senate would not have been likely to have sought PSD rules in 1970. It is unlikely that members of the upper chamber would have favored a policy that sacrificed opportunities for pollution control in nonattainment areas in order to use environmental regulation to slow the relocation of industry. Thus, although the Senate might have been inclined to go along with nationwide NSPS, it would not have been inclined to favor tough PSD rules because it represented a long-term commitment to cleaner air everywhere.

2. *Some Evidence on Preferences*

The conjectured preferences for the two chambers of Congress and the President described above are sufficient to yield a single rank-ordering over the two issues involved in the legislation. With respect to allocating the costs of regulation, the members of the House, on average, would be expected to prefer to place more of the burden on new industry. Their Senate counterparts would be less inclined to sacrifice growth for the benefit of established firms, and the President would be even less so inclined. Within each chamber, eastern and Great Lakes representatives, on average, should prefer to burden new firms to a greater extent than would members from the West and Southwest.

With respect to stringency, members of the House, on average, represent more urban (and therefore more polluted) areas and so would be expected to favor more stringent regulations than the average member of the upper chamber, who in turn is more inclined toward stringent regulation than the President. This conjecture is somewhat controversial as it has often been assumed that it was the Senate rather than the House of Representatives that sought tougher environmental legislation. The median Senators, however,

---

See Bureau of the Census, Dep't of Commerce, Statistical Abstract of the United States 252 (1988).

represent clean air areas. This implies that stringent regulation does them little good and may in fact impede growth in their states.

Before proceeding with our analysis we need to define stringency more carefully. In practice, stringency is a complex concept. One aspect is the number of pollutants regulated, and the criteria for establishing maximum concentrations of each. Here we regard a policy as more stringent if it regulates more pollutants or adopts more rigorous criteria for setting ambient air quality standards. In addition, stringency has both a shortrun and longrun aspect. Because of the costs to established firms, members of the House generally preferred to delay stringent regulation (and thus delay or evade the costs to old plants), whereas Senators were more likely to prefer results sooner rather than later. But the House sought greater longrun stringency. The standards it adopted, though taking a longer time to become effective, would eventually result in cleaner air than those proposed by the Senate. In this sense, we regard the House as favoring more stringent standards.

Stringency is determined not only by the standards enacted, but also by technical and arcane definitions of the technology required (for example, "best available technology" versus "continuous emissions reduction"). Further, requirements to take into account the costs of proposed regulations, or to consider alternative means of accomplishing an end, affect the ability of the EPA Administrator to promulgate stringent regulations, at least in the short run. Provisions allowing for exemptions and waivers also reduce stringency by limiting the applicability of standards. Thus, comparing the proposals offered in each chamber involves an examination of much more than just the substantive definitions of the standards so often discussed in the literature.

An examination of the proposals and amendments to the Clean Air Act offered in each chamber in the 1970s confirm these conjectures. The amendments to the Clean Air Act offered in the House in 1970 were contained in H.R. 17,255;[98] the Senate's amendments were presented in S. 4358.[99] In comparing these two bills, three differences are apparent.

The first major difference pertains to the comparative treatment of old and new sources. The House proposed to protect against the relocation of existing plants and would have sheltered existing firms from severe cost burdens. For all practical purposes, the House sought to exempt old sources from regulation, while subjecting new sources to stringent control.[100] To this end, the House bill required that any new source be designed and equipped to prevent and control emissions to the fullest extent compatible

---

[98] H.R. 17,255, 91st Cong., 2d Sess., 116 Cong. Rec. 19,224 (1970).

[99] S. 4358, 91st Cong., 2d Sess. (1970).

[100] H.R. 17,255, 91st Cong., 2d Sess. § 5, 116 Cong. Rec. 19,225 (1970).

with "the available technology and economic feasibility."[101] In interpreting this condition, the report of the House Commerce Committee stated that the promulgation of federal emissions standards for new sources would "preclude efforts on the part of States to compete with each other in trying to attract new plants and facilities without assuring adequate control of extrahazardous or large-scale emissions therefrom."[102] The Senate was less insistent about the technology used and instead defined new source performance standards in terms similar to national ambient air quality standards.[103] The technology bias of the House, of course, illustrates its members' desire to prevent industrial relocation.

A second difference between the chambers was their treatment of economic considerations in setting standards. For example, the House explicitly required the EPA to consider the costs of proposed pollution control techniques in setting national ambient air quality standards[104] and NSPS,[105] whereas the Senate left cost considerations implicit in setting national standards and deemphasized such considerations in NSPS.[106] Environmentalists have generally fought the consideration of costs in setting standards, so this difference forms the basis for the belief that the Senate bill was more stringent.

The Senate and House differed in another way that offset the differences over cost considerations. This third difference pertains to the structure and process of standard-setting. The House proposed to delegate far greater discretion to the EPA than the Senate; the House also proposed a mechanism for more effective, and thus more stringent, regulation than would have emerged under the Senate proposal. Whereas the House proposed to delegate the authority to make almost every decision to the Administrator,[107] the Senate proposed to delegate numerous decisions to others. For example, the Senate sought to delegate authority to grant exemptions for implementation plans to the President[108] and to the courts.[109] By creating more veto points, the Senate gave opponents of stringent standards a greater likelihood of both delay and success.

---

[101] Id., 116 Cong. Rec. 19,226.

[102] H.R. Rep. No. 1146, 91st Cong., 2d Sess. 3, *reprinted in* 1970 U.S. Code Cong. & Admin. News 5356, 5358.

[103] S. 4358, 91st Cong., 2d Sess. § 6 (1970).

[104] See H.R. 17,255, 91st Cong., 2d Sess. § 5, 116 Cong. Rec. 19,225 (1970).

[105] Id. § 5, 116 Cong. Rec. 19,225.

[106] S. Rep. No. 1196, 91st Cong., 2d Sess. 16 (1970).

[107] H.R. 17,255, 91st Cong., 2d Sess. § 4, 116 Cong. Rec. 19,224 (1970).

[108] S. 4358, 91st Cong., 2d Sess. § 7 (1970).

[109] Id. § 6.

The Senate also was more detailed, and hence more constraining, in its definition of the Administrator's duties. For example, the Senate proposed that the EPA establish air quality standards for three categories of pollutants[110] whereas the House offered no such restriction. Moreover, the House left open the means by which new sources were to comply with national standards, calling for the lowest possible emissions given cost efficiency, and thereby giving more flexibility to the Administrator.[111] The Senate wanted NSPS to be met through the application of the best available emissions control technology.[112]

Finally, the Senate proposed many more ways in which states or sources could exempt themselves from the requirements of the Clean Air Act or waive the implementation of standards, at least temporarily. The Senate would have allowed state implementation plans to contain less rigorous standards than the national standards for up to three years,[113] whereas no such escape was allowed in the House bill. The House proposed that the Administrator be permitted some limited discretion to exempt new facilities from compliance for reasons of public health, research, or national security or, in the case of federal facilities, if determined to be in the "paramount interest of the United States."[114] The Senate, however, would have allowed the Administrator to grant extensions,[115] the President to exempt federal facilities,[116] and governors to petition the courts for exemptions.[117] The Senate would have also allowed waivers if the Secretary found that the source had achieved maximum air pollution control using the best available technology.[118] Relief could be obtained from the courts when substantial effort could be shown.[119]

Taken together, the greater burden imposed by the Senate on the Administrator, the more limited authority given the Administrator, and the greater ability for others to grant exemptions and waivers of the standards estab-

---

[110] Id. §§ 5-6 (including agents covered by air quality criteria that affect public welfare, agents emitted by stationary sources, and agents that "available material evidence indicates are hazardous to the health of persons").

[111] H.R. 17,255, 91st Cong., 2d Sess. § 5, 116 Cong. Rec. 19,225 (1970); see H.R. Rep. No. 1146, 91st Cong., 2d Sess. 3, 10, *reprinted in* 1970 U.S. Code Cong. & Admin. News 5356, 5358, 5365-66.

[112] S. Rep. No. 1196, 91st Cong., 2d Sess. 16.

[113] Compare S. 4358, 91st Cong., 2d Sess. § 6 (1970) with H.R. 17,255, 91st Cong., 2d Sess. § 4, 116 Cong. Rec. 19,225 (1970).

[114] H.R. 17,255, 91st Cong., 2d Sess. §§ 5, 10, 116 Cong. Rec. 19,225, 19,228.

[115] S. 4358, 91st Cong., 2d Sess. § 6 (1970).

[116] Id. § 7.

[117] Id. § 6.

[118] Id.; see S. Rep. No. 1196, 91st Cong., 2d Sess. 57-58.

[119] S. 4358, 91st Cong., 2d Sess. § 6 (1970).

lished by the Administrator lead us to conclude that the Senate sought a less stringent air pollution policy. The effect of the Senate provisions was to make it more difficult for the EPA to implement stringent policies effectively.[120]

3. *The Congressional Response to Policy Innovation*

During 1976 and 1977, both the House and the Senate considered numerous responses to the PSD policy that the EPA promulgated in 1974. The actions they ultimately chose reflect the preferences we conjectured on the basis of electoral considerations. The aspects of the House proposals in 1976 and 1977 that provided a measure of protection for existing firms, particularly manufacturing firms in the Northeast and Midwest, are well-documented elsewhere,[121] and we will not review them here. Our focus is on the differences between the policy preferences of the House and Senate as revealed in the bills each produced in these years. In general the House bills proposed PSD regulations that were more stringent than those the Senate produced.

In 1976, the House and Senate bills offered different definitions of PSD. Both chambers defined PSD in relation to the maximum allowable increments of pollution over baseline concentrations for each designated control region, but differed over which pollutants would be subject to PSD requirements and how those requirements would be met. In the House, increments were established for all pollutants for which there existed a national ambient air quality standard.[122] The EPA was to implement this by requiring standards of performance that reflected "the degree of emission reduction achievable through the application of the best technological system of continuous emission reduction."[123] The Senate, however, specified deterioration increments only for sulfur oxides and particulates, and recommended that these standards be met through the relatively weaker requirement of the "best available control technology."[124] Continuous emissions reduction is a more stringent requirement than BACT in two respects: first, the continuous emissions reduction requirement was meant to forbid intermittent nonattain-

---

[120] Cf. McCubbins, The Legislative Design of Regulatory Structure, 29 Am. J. Pol. Sci. 721 (1985) (examining the manner in which substantive discretionary authority available to an administrative agency is fashioned through the scope of regulatory activities granted to the agency, the instrumentality by which the agency can implement its policy choices, and the procedures required for agency decisionmaking).

[121] See B. Ackerman & W. Hassler, supra note 11, at 29-57.

[122] See H.R. 10,498, 94th Cong., 2d Sess. § 108(a), 122 Cong. Rec. 30,780 (1976).

[123] Id. § 111, 122 Cong. Rec. 30,784.

[124] S. 3219, 94th Cong., 2d Sess. § 6, 122 Cong. Rec. 30,763 (1976).

ment by sources even for short periods of time; second, it was not restricted to "available" technology.

As in the regulations the EPA promulgated in 1974, the House bill specified three classes of clean air areas subject to PSD. Class I included national parks and wilderness areas, and class II encompassed all other attainment areas.[125] The states could also designate (or redesignate) some areas as class III, which were to be clean air areas in which industrial development would be permitted and in which air quality levels would be allowed to deteriorate to the national standard.[126] But before doing so, the states were required to give notice and to hold public hearings.[127]

The 1976 Senate proposal specified only two classes of attainment areas. Like the House bill, class I included parks and wilderness areas and class II included all remaining areas.[128] But the Senate bill rejected the policy promulgated by the EPA and accepted by the House that some clean air areas should be set aside for industrial development where deterioration to the national standard would be allowed, as in the proposed class III areas in the House bill. The Senate Public Works Committee report stated that the wording of the Senate bill was intended to reject the policy that class I areas should remain pristine. Western Senators with large fractions of their states contained in national parks and wilderness areas continually sought to redefine and weaken the class I definitions.[129]

The House allowed for compliance date extensions in its rules, but the Senate, as in the 1970 Amendments, allowed for many more exemptions and waivers.[130] Further, the House imposed somewhat tighter procedural requirements for granting waivers—the EPA Administrator could grant compliance date extensions only with notice and a public hearing "on the record," and the governor of the affected state was given the power to veto such actions.[131]

In its report accompanying the Senate bill,[132] the Senate Public Works Committee stated that its intent was to restrict the EPA's role in preventing significant deterioration.[133] Its proposal was to limit the EPA's role in

---

125 H.R. 10,498, 94th Cong., 2d Sess. § 108(a), 122 Cong. Rec. 30,781 (1976).

126 Id.

127 Id.

128 S. 3219, 94th Cong., 2d Sess. § 6, 122 Cong. Rec. 30,763 (1976).

129 S. Rep. No. 717, 94th Cong., 2d Sess. 25 (1976).

130 Compare H.R. 10,498 §§ 103(a), 106(a), 112, 202, 122 Cong. Rec. 30,776, 30,778, 30,785, 30,787 (1976) with S. 3219 §§ 7(a), 9(a), 15, 122 Cong. Rec. 30,764-65, 30,768 (1976).

131 H.R. 10,498 § 103(a), 122 Cong. Rec. 30,776 (1976); see also H.R. Conf. Rep. No. 1742, 94th Cong., 2d Sess. 90-91 (1976) (discussing the compliance date extension procedures contained in the House bill).

132 S. 3219, 94th Cong., 2d Sess., 122 Cong. Rec. 30,763 (1976).

133 S. Rep. No. 717, 94th Cong., 2d Sess. 2-3 (1976).

implementing a new PSD policy to: (1) approving the new source review established by the state;[134] (2) seeking injunctive relief or other measures necessary to prevent the issuance of a permit for new sources that did not comply with air quality requirements;[135] (3) resolving interstate disputes;[136] and (4) notifying a state when it believed an adverse impact might occur in a class I area.[137] In essence, the EPA's role was restricted to assuring compliance with the law. Lastly, the Senate proposed to transfer the authority to grant exemptions for coal conversion from the EPA to the Federal Energy Administration, presumably to reduce the influence of environmental policy on the granting of exemptions while increasing the importance of energy policy.[138]

The conference committee appointed in 1976 to reconcile the two bills accepted all of the amendments to the 1970 Clean Air Act that each house offered individually, and compromised on all amendments over which the two houses differed.[139] The conference accepted the Senate's weaker definition for class I areas,[140] while also adopting the House proposal to allow the redesignation of areas to class III.[141] The conference settled on the Senate's plan to implement BACT but accepted many of the House's procedural demands and all of the proposed exemptions of both chambers. Despite these modifications, the bill failed to pass.[142]

In 1977, both chambers proposed bills similar to their 1976 proposals. The House, however, proposed regulations even more stringent than they had the year before. They added nitrogen dioxide to the list of pollutants subject to national standards,[143] added "visibility" as a criterion for setting standards in class I areas,[144] and placed additional burdens and standards of proof on new sources in meeting compliance requirements.[145]

---

134 Id. at 27.

135 Id.

136 Id.

137 Id.

138 S. 3219, 94th Cong., 2d Sess. § 15(b)(7), 122 Cong. Rec. 30,768 (1976).

139 See H.R. Conf. Rep. No. 1742, 94th Cong., 2d Sess. 85-125; see also 122 Cong. Rec. 34,380-88 (1976) (statement of the bill's floor manager, Senator Edmund Muskie of Maine, summarizing the conference report provisions).

140 See H.R. Conf. Rep. No. 1742, 94th Cong., 2d Sess. 103 (1976).

141 Id.

142 The Senate attempted to take up consideration of the conference report on the final day of the 94th Congress, but did not reach a final vote on the agreement. See 122 Cong. Rec. 34,415-17 (1976).

143 H.R. Rep. No. 294, 95th Cong., 1st Sess. 41, *reprinted in* 1977 U.S. Code Cong. & Admin. News 1077, 1119.

144 Id. at 13, *reprinted in* 1977 U.S. Code Cong. & Admin. News at 1090-91.

145 Id. at 190, *reprinted in* 1977 U.S. Code Cong. & Admin. News at 1268-69.

By contrast, the 1977 Senate proposal sought less stringent requirements than those of their relatively weaker 1976 proposal. For example, the Senate sought to exempt small polluters.[146] Ultimately, the provisions of the conference report for these sections were much the same as in 1976, and in the end, although both chambers proposed significant changes, the final legislation differed little in substantive detail from the original 1974 EPA promulgation.

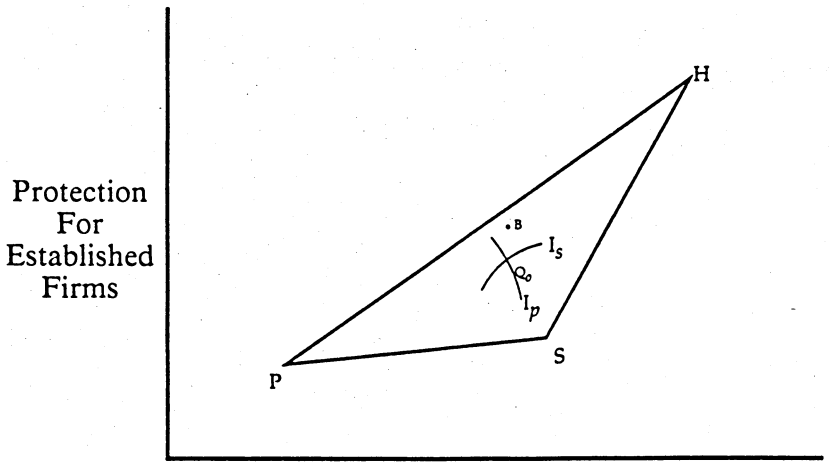


Figure 3

D. Fourth Step: The New Equilibrium

The inference to be drawn from this discussion of the history of the Clean Air Act between 1970 and 1977 is as follows. If the relative positions of the House, Senate, and President are as described, and if the 1970 Amendments were not intended to produce PSD, the 1972 court decision in *Sierra Club v. Ruckelshaus*[147] upset a policy equilibrium. The consequences of this unanticipated change in policy can be illustrated by returning to our simple two dimensional model of policy choice. The preferences just established for each institutional actor are illustrated in Figure 3. The House favors greater

---

[146] Compare S. 252, 95th Cong., 1st Sess. § 6, 123 Cong. Rec. 18,517 (1977) with S. 3219, 94th Cong., 2d Sess. § 6, 122 Cong. Rec. 30,763 (1976) (providing for a small polluter exemption). Consideration of the Senate bill (S. 252) was indefinitely postponed early in the 95th Congress, and the House bill (H.R. 6161) was passed in lieu of that Senate bill. See Clean Air Act Amendments of 1977, Pub. L. No. 95-95, 91 Stat. 685 (codified as amended at 42 U.S.C. §§ 7401-7642 (1982 & Supp. IV 1986)).

[147] 344 F. Supp. 253 (D.D.C. 1972), aff'd per curiam by an equally divided Court sub nom. Fri v. Sierra Club, 412 U.S. 541 (1973).

stringency than the Senate favors, which in turn favors more stringent regulation than the President. The House, more than the Senate or the President, prefers that new firms bear the costs of pollution control. The ideal points on the stringency-cost tradeoff among the House, the Senate, and the President are represented in the figure by *H*, *S*, and *P*.

In Figure 3, the original 1970 policy is represented by point Q_0. The court decision in *Sierra Club* and the EPA's promulgated regulation moved the policy to point *B*. The new policy was in the Pareto optimal set formed by the preferences just identified. There are two implications of this change in policy. First, as in our previous analysis, a return to the original policy, Q_0, is not possible. Indeed, the new policy cannot be significantly amended.

Second, the House was the principal beneficiary of the change, for it favored a combination of relatively strict long-term goals for air quality (which reflected the desires of its urban constituency for cleaner air) *and* an implementation strategy that would minimally disrupt established economic patterns of development (which reflected its relatively greater sensitivity to narrow industrial interests). PSD strengthened standards in areas that already had clean air. It also slowed progress towards air quality goals by giving greater protection to established industries. A clear implication of EPA's promulgated PSD rules, together with the provisions of the 1970 Amendments regarding new sources, was that existing facilities would be protected against the possibility that stringent air pollution regulation would hasten their obsolescence.[148] Consider the requirements for new sources. Generally, it is less expensive to achieve a given degree of pollution abatement per unit of production in a new facility than in an old one. The reason is that the abatement system can be included in a new plant's original design, but must be fit into an older plant. Hence, a "least cost" strategy for achieving a given air quality target will often involve replacing older facilities with newer ones. New source performance standards protected existing firms against this eventuality by imposing less demanding requirements on them.

This strategy would be far less effective if it applied only to areas with severe air pollution problems. In areas where emissions could increase significantly without causing air quality to fall short of national standards, new facilities could be built that might face even lower costs of control than those facing established facilities in areas with poor air quality. Nationwide NSPS combined with PSD rules impose on these new facilities a set of regulations almost as demanding as they would face if they located in an area with dirty air. Hence, NSPS and PSD implement the twin policies of improving air quality and protecting established industries against a nationwide redistribution of economic activity due to air pollution regulation. Consequently, the

---

[148] See B. Ackerman & W. Hassler, supra note 11, at 10-12.

court's decision and EPA's subsequent PSD rules moved air pollution policy toward the ideal point of the House.

III. The Use of Structure and Process to Enforce Congressional Intent

The changes in structure and process proposed by each chamber in response to the policy innovation of PSD reflect the basic differences in the electoral considerations of the members of each house. These proposed reforms had unambiguous consequences for environmental policy. In this Part we show, first, that the structure and process enacted by Congress for the implementation of air pollution policies by the EPA mirrored the conflict just described between the House and Senate; and second, that the procedures each chamber proposed, and those each eventually adopted, sought to stack the deck in favor of the interests that chamber represented. As discussed in Part II above, the House proposed procedures to enfranchise local industry, whereas the Senate favored procedures that reduced the force of EPA regulations. In almost every case neither house agreed to the changes in PSD proposed by the other, thereby thwarting this avenue of upsetting the status quo.

We examine proposed and enacted changes in the procedures for EPA rulemaking, the standards of proof, and the standards of review of EPA decisions. We then explore proposed changes in the system of civil representation for the EPA. We describe the debates surrounding the creation of the National Commission on Air Quality (NCAQ), and show how members of each chamber sought to use the NCAQ to alter and constrain the EPA Administrator's ability to implement the new PSD provisions. Lastly, we examine an attempt to enfranchise small independent auto repair shops into EPA rulemaking on auto warranties.

A. Administrative Procedures

In its proposed 1977 Clean Air Act Amendments, the House sought to make EPA rulemaking dockets more formal, protracted, and detailed. The House proposed raising the evidentiary standard from the "arbitrary and capricious" standard of the 1970 Amendments to a "substantial evidence" test in 1977.[149] In addition, the House proposed a more restrictive standard under which courts could invalidate EPA rules—namely, that the agency's error had been so "serious and related to matters of such central relevance . . . that there is a substantial likelihood that the rule would have been significantly changed if such errors had not been made."[150] The House also speci-

---

[149] H.R. 6161, 95th Cong., 1st Sess. § 305(a), 123 Cong. Rec. 16,967 (1977).

[150] Id.

fied elaborate rules regarding the content of rulemaking dockets, and required that more elaborate public hearings and defenses of agency decisions be undertaken.[151]

More elaborate procedures are generally regarded as favorable to regulated industries. Because industries possess much of the information relevant to regulatory decisions, elaborate processes give them more power by increasing the importance of that information. Another contributing factor is that industries, with greater economic stakes in regulatory issues, are more likely to devote the resources necessary to be effectively represented in expensive proceedings. In this case, established industries (old sources) are more likely to be advantaged by cumbersome proceedings. In the development of NSPS, most of the facilities to which these standards would apply do not exist at the time the regulations are promulgated, and for that reason they are less likely to have their interests effectively represented. Indeed, the House hearings indicate that industrial groups representing established industries strongly supported most of the procedural changes that the House Commerce Committee eventually reported.[152] This is consistent with the view that the House did not want to disrupt existing patterns of industry through air pollution regulation.

In the Senate, no procedural changes of these types were adopted in 1976 or 1977, and in conference much of what the House had proposed was removed.[153] For example, the recommended "substantial evidence" standard for judicial review was deleted, and the provision requiring cross-examination of witnesses at public hearings on proposed rulemaking was replaced

---

[151] See H.R. Conf. Rep. No. 564, 95th Cong., 1st Sess. 177, *reprinted in* 1977 U.S. Code Cong. & Admin. News 1502, 1558. For example, the 1977 Amendments contained provisions that specified what documents were to be included in the docket: (1) a statement of the basis and purpose of the rule including all data, information, and documents pertinent to the rulemaking procedures; (2) all written comments and documentary information that was of central relevance to the rulemaking; (3) the transcript of any public hearings on the proposed rule; and (4) draft proposals and accompanying documents submitted by the Administrator to the Office of Management and Budget, interagency comments on the proposal, and EPA responses to those comments. Clean Air Act Amendments of 1977, Pub. L. No. 95-95, § 307(d)(3)-(4), 91 Stat. 685, 778-79 (codified at 42 U.S.C. § 7617 (1982)); see also H.R. Rep. No. 294, 95th Cong., 1st Sess. 318-25, *reprinted in* 1977 U.S. Code Cong. & Admin. News 1077, 1397-1404 (summarizing House Commerce Committee's reasons for proposed changes in administrative procedures).

[152] See Clean Air Act Amendments of 1977: Hearings on H.R. 4151 and H.R. 4758 Before the Subcomm. on Health and Environment of the House Comm. on Interstate and Foreign Commerce, 95th Cong., 1st Sess. 958, 1003, 1044-45 (1977) (statements of Kenneth Tucker, Vice President, Int'l Council of Shopping Centers; Robert Arquila, President, Nat'l Ass'n of Home Builders; and Edward Weber, Ass't Gen'l Counsel, Republic Steel Corp.).

[153] H.R. Conf. Rep. No. 564, 95th Cong., 1st Sess. 177-78, *reprinted in* 1977 U.S. Code Cong. & Admin. News 1502, 1558-59.

by one that permitted only written rebuttals.[154] Perhaps most significantly, the Senate refused to accept the House proposal of a one-house legislative veto over EPA rules.[155] Moreover, although the Senate adopted several House provisions that made EPA's procedures somewhat more elaborate, it insisted that this not affect EPA review of state implementation plans.[156] Because SIPs are the primary vehicles for regulating existing stationary sources, they contain most of the regulations for established facilities. Preservation of a simple, flexible process for reviewing SIPs meant that specific industries would be less advantaged procedurally in the review of the regulations pertaining to them. To the extent that an industry was advantaged procedurally, it was also advantaged in more general rulemaking, such as the establishment of national ambient air quality standards, NSPS, or emissions standards for mobile sources and hazardous pollutants, all of which affected Senate as well as House constituents.

B. Agency Representation Before the Courts

Another proposed change involved the relationships between the Department of Justice and the EPA in the latter's representation before federal courts. The 1970 Amendments authorized the Attorney General to represent the EPA in all cases appearing in federal courts.[157] Of course, the relationship between the Justice Department and the EPA is not parallel to a normal attorney-client relationship. The Justice Department is itself a policymaking agency, and historically it has been especially closely tied to the President. Thus, when the EPA's staff promulgates a rule that is appealed by an interested party, granting the Attorney General authority to defend the rule creates a de facto veto power that could undermine the defense of the EPA's policy decisions. Of course, one of the most important examples of EPA representation was the *Sierra Club* case,[158] in which the Justice Department defended the EPA's decision not to adopt PSD rules.

In 1976, the House Commerce Committee bill[159] included a provision stripping the Attorney General of his authority to represent the EPA in court, and would instead have authorized the Administrator to do so unless

---

[154] Id.

[155] Id. at 188-89, *reprinted in* 1977 U.S. Code Cong. & Admin. News at 1569.

[156] Id. at 177, *reprinted in* 1977 U.S. Code Cong. & Admin. News at 1558.

[157] Clean Air Amendments of 1970, Pub. L. No. 91-604, § 12(a), 84 Stat. 1676, 1707 (1970) (codified as amended at 42 U.S.C. § 7605 (1982)).

[158] Sierra Club v. Ruckelshaus, 344 F. Supp. 253 (D.D.C. 1972), aff'd per curiam by an equally divided Court sub nom. Fri v. Sierra Club, 412 U.S. 541 (1973).

[159] H.R. 10,498, 94th Cong., 2d Sess., 122 Cong. Rec. 30,798 (1976).

he requested that the Attorney General prosecute the case.[160] On the House floor, Representative M. Caldwell Butler of Virginia proposed an amendment to delete this section of the Committee's bill, stating that one agency should be responsible for all federal litigation and that creating a separate litigation staff in the EPA was a "wasteful duplication of resources and efforts."[161] Reflecting the power of House rules to protect committee proposals, Committee member Richardson Preyer of North Carolina immediately offered a substitute for the Butler amendment that restored most of what the Committee had proposed.[162] Preyer's substitute gave the Justice Department the right to represent the EPA before the Supreme Court and made the Department a party to decisions about whether the EPA should appeal; however, the EPA could represent itself before the courts of appeals (but not the Supreme Court) as long as the Justice Department agreed that an appeal could be made.[163] The House adopted Preyer's substitute to the Butler amendment.[164] In 1977, when the Committee again reported a bill to amend the Clean Air Act, the Preyer amendment was included in the bill.[165]

The Senate, on the other hand, simply never considered the question of EPA representation in either its 1976 or 1977 bills. In neither year did the Senate Public Works Committee propose to change the system of representation, and in neither year was such an amendment proposed on the Senate floor.[166] In the final version of the 1977 Clean Air Act Amendments (1977 Amendments), the Justice Department's responsibility to represent the EPA remained, subject, however, to the provisions of a memorandum of understanding between the EPA and the Department of Justice that retained the Department's overall supervision, but gave the EPA's lawyers a formal role in developing litigation strategy.[167] The 1977 Amendments state that litigation was to be conducted in accordance with this memorandum of

---

[160] Id. § 311, 122 Cong. Rec. 30,798; see also H.R. Rep. No. 1175, 94th Cong., 2d Sess. 272-77 (1976) (discussing the legislative intent of this proposal).

[161] 122 Cong. Rec. 30,499 (1976).

[162] Id. at 30,499-500.

[163] Id.

[164] Id. at 30,503.

[165] H.R. Rep. No. 294, 95th Cong., 1st Sess. 28, 332-337, *reprinted in* 1977 U.S. Code Cong. & Admin. News 1077, 1411-16.

[166] Neither conference report contained mention of a Senate proposal regarding EPA representation. See H.R. Conf. Rep. No. 1742, 94th Cong., 2d Sess. 124 (1976); H.R. Conf. Rep. No. 564, 95th Cong., 1st Sess. 173, *reprinted in* 1977 U.S. Code Cong. & Admin. News 1502, 1554.

[167] 42 U.S.C. § 7605(b) (1982); see also H.R. Conf. Rep. No. 564, 95th Cong., 1st Sess. 173-76, *reprinted in* 1977 U.S. Code Cong. & Admin. News 1502, 1554-57 (discussing congressional reaction to memorandum of understanding between the EPA and the Department of Justice).

understanding.[168]

This episode illustrates a number of important points. The controversy over representation indicates that all participants regarded the issue as important to the development of the details of environmental regulation. Moreover, the House was much more concerned about the Justice Department's policy role in environmental matters than were either the Senate or the Administration. In the end, the latter two gave in on one element: the special expertise of the agency demanded that its role in developing cases involving the agency be guaranteed. But the status quo under which the Justice Department supervised litigation whenever the Attorney General (and presumably the President) wanted was not changed. The implication is that the House had different policy objectives than either the Senate or the President—objectives that were threatened by the lack of autonomy at the EPA in representing itself in court.

It is impossible to ascertain precisely what policy issues so concerned the House. But one was that the EPA, after initiating the process regarding PSD, first rejected the PSD proposal in accordance with the Justice Department's recommendation, and then had this position vigorously defended by the Justice Department in court. One plausible inference from this sequence of events, as Melnick has previously argued,[169] is that the EPA initially wanted to adopt PSD, but decided against doing so after the Administration intervened. If so, the attempt by the House to insulate the EPA from the Justice Department, and the failure of the Senate to comply, is consistent with our hypotheses about the relative policy preferences of each. It also supports the view that the strong role given the Justice Department in the 1970 Amendments constituted a procedural protection for the Senate and the President against precisely the kind of drift in agency policy that the House preferred—and ultimately obtained from the courts.

C. The Use of Studies to Control Agency Rulemaking

Congress often constrains the ability of an agency to control the flow of information relevant to its rulemaking activities as it did in the 1977 Amendments. This is most often accomplished by requiring that some third party conduct an independent study, or in some cases a joint study with the agency, the results of which affect the agency's rulemaking in some fashion.[170] These requirements have two effects. First, the third party

---

[168] Clean Air Amendments of 1977, Pub. L. No. 95-95, § 304(a), 91 Stat. 685, 772 (codified at 42 U.S.C. § 7605 (1982)).

[169] See generally R. Melnick, supra note 11, at 71-80 (recounting the history of the *Sierra Club* litigation).

[170] See Clean Air Amendments of 1977, Pub. L. No. 95-95, § 403(a), 91 Stat. 685, 792 (codified at 42 U.S.C. § 7548 (1982)); Toxic Substances Control Act, § 25, 15 U.S.C. § 2624

enfranchises an important set of constituents interested in the agency's decisions.[171] Second, the structure establishes a fire alarm oversight system.[172] Because the study group has access to all of the information available to the agency, and because their mandate is to inform Congress, the study group is in a perfect position to provide third party oversight for members of Congress.

Both aspects of third party studies serve to restrict agency decisionmaking. Because a study group gives members of congressional committees greater ability to oversee the actions of their agencies, its creation induces agency bureaucrats to comply with congressional wishes. The higher the probability that noncompliant behavior will be discovered, the more effective politicians can be in structuring the agency's incentives. On occasion, Congress makes this incentive more direct by formally linking the agency's decisionmaking to the group's findings.[173] For example, Congress may require consultation between the agency and a study group, occasionally requiring that a study be conducted jointly by the two. Or it may require the agency to obtain the study group's agreement on proposed actions. In other instances, Congress may require that the study group's report be part of the agency's proceedings, thereby forcing the agency to respond to the report in its decisionmaking. This also constrains the agency, for the report can then be used as evidence against the agency in court.

For example, the EPA was authorized to undertake several studies in the 1977 Amendments.[174] Congress also chose to delegate some studies jointly to the EPA and other agencies and in some instances to bypass the EPA and

---

(1982) (authorizing a study on establishing a standard classification system of chemicals and related substances and a standard storage and access system for retrieving the information).

[171] For example, with respect to establishing air quality criteria, Congress provided that the Administrator of the EPA may "establish a standing consulting committee for each air pollutant . . . which shall be comprised of technically qualified individuals representative of State and local governments, industry, and the academic community." 42 U.S.C. § 7408(b)(2) (1982). Congress also required the EPA to utilize third parties in promulgating a national primary air quality standard for nitrous oxides. To assist the EPA, the 1977 Amendments provided for a seven-member independent scientific review committee composed of "at least one member of the National Academy of Sciences, one physician, and one person representing State air pollution control agencies." Congress mandated that the committee "shall complete a review of the criteria published under section 7408 [of Title 42] . . . and shall recommend to the Administrator any new national ambient air quality standards and revisions . . . as may be appropriate." Id. § 7409(d).

[172] See supra note 15 and accompanying text.

[173] See, e.g., Clean Air Amendments of 1977, Pub. L. No. 95-95, § 403(f), 91 Stat. 685, 793 (codified at 42 U.S.C. § 7548 (1982)) (authorizing the EPA and the National Academy of Sciences to study the effects of emissions on the "public health and welfare" and the technological feasibility of meeting emissions standards).

[174] Id. §§ 403-405, 91 Stat. at 792-95 (codified at scattered sections of 42 U.S.C.).

to rely completely on a third agency to undertake studies for the EPA and Congress. For example, Congressmen concerned about the effect a ban on aerosols would have on the cosmetics and household products industries required the Secretary of Labor to study the effects of a ban on the use of halocarbons in aerosol containers.[175] On the same topic, the Administrator was required to establish a coordinating committee for regulating halocarbons. Congress required that the committee include the National Oceanographic and Atmospheric Administration (NOAA), the National Aeronautic and Space Administration (NASA), the Federal Aviation Administration, the Department of Agriculture, the National Cancer Institute, the National Institute of Environmental Health Sciences, the National Science Foundation (NSF), and the Department of State.[176] This coordinating committee was to report to "the appropriate committees of the House and the Senate," not to the EPA.[177] NOAA, NASA, NSF, the Department of Agriculture, and HEW each were also delegated responsibilities to undertake continuing studies on ozone and halocarbons.[178] The National Academy of Sciences (NAS) was asked to conduct three separate studies.[179]

In other instances, Congress has required the EPA to consider a study group's findings. For example, the EPA was ordered to include a statement of a rule's basis and purpose in any notice of proposed rulemaking.[180] Congress added a requirement to the 1977 Amendments that such statements

> shall also set forth or summarize and provide a reference to any pertinent findings, recommendations, and comments by the Scientific Review Committee established under section 7409(d) of this title [42] and the National Academy of Sciences, and, if the proposal differs in any important respects from any of these recommendations, an explanation of the reasons for such differences. All data, information, and documents referred to in this paragraph on which the proposed rule relies shall be included in the docket on the date of publication of the proposed rule.[181]

---

175 42 U.S.C. § 7453(e) (1982).

176 Id. § 7453(f).

177 Id. § 7453(g).

178 Id. § 7454.

179 Clean Air Amendments of 1977, Pub. L. No. 95-95, §§ 106(a), 403(a), 405(a), 91 Stat. 685, 691, 792-93 (codified at scattered sections of 42 U.S.C.).

180 Id. § 305(a), 91 Stat. at 774 (codified at 42 U.S.C. § 7607 (1982)).

181 42 U.S.C. § 7607(d)(3) (1982).

1. The National Commission on Air Quality

The use of studies to enfranchise the interests of each party to an enacting coalition, to create a more effective fire alarm oversight system, and to mirror the policy conflicts at the time of legislative enactment are exemplified in the debate surrounding the creation of the National Commission on Air Quality. This debate illustrates the strategic use of structure and process to constrain agency decisions.

The Senate proposed to make the "achievement and maintenance of national ambient air quality standards" and "prevention of significant deterioration" the NCAQ's principal areas of investigation.[182] Secondarily, the Senate would have authorized the NCAQ to examine unregulated pollutants and to address the adequacy of abatement research and development and the ability of federal, local, and state agencies to implement the purposes of the Clean Air Act.[183]

The 1976 Senate bill proposed a sixteen-member NCAQ comprised of twelve members of the public (including four governors), to be appointed by the President with the advice and consent of the Senate, and four nonvoting Congressmen. The four governors or their appointed representatives would have provided state input into the congressional advisory reports. In effect, the NCAQ would have provided a means for the states to check EPA actions. It would also have been required to seek consultation from federal, state, and local agencies. To facilitate congressional oversight, the chairman and ranking minority member of both the Senate Public Works Committee and the House Commerce Committee would have been nonvoting ex officio members.[184]

The 1976 House amendment to the Senate bill would have established a similarly structured NCAQ composed of eleven members, including the chairman and ranking minority member of the Senate Public Works Committee and the House Commerce Committee.[185] The House conference committee report reflected many of the same concerns as the Senate committee report, except that this version reflected a greater concern with the effect of EPA regulations on the automobile industry. The final version of the bill required the NCAQ to focus its studies on the "extent to which the reduction of hydrocarbon emissions is an adequate or appropriate method to achieve primary standards for photochemical oxidants"[186] as well as the means of achieving and maintaining "national ambient air quality standards

---

[182] S. 3219, 94th Cong., 2d Sess., § 38, 122 Cong. Rec. 30,773 (1976).

[183] Id.

[184] Id.

[185] Id. § 108(f), 122 Cong. Rec. 30,783 (House amendment to Senate bill).

[186] H.R. Conf. Rep. No. 1742, 94th Cong., 2d Sess. 76 (1976).

and . . . the prevention of significant deterioration of air quality,"[187] the Senate's primary concern.

The strategic use of structure and process—in this case the requirement of a study—to constrain agency decisions is exemplified by the attempts in both chambers to amend the provisions creating the NCAQ. In the 1976 Senate debate, the central area of controversy was the relationship between the proposed NCAQ study to the implementation of PSD requirements as defined in section 6 of the Senate bill.

Early in the debate, the Chairman of the Senate Public Works Committee, Jennings Randolph of West Virginia, offered an amendment concerning timetables and procedures for the NCAQ that assumed the immediate implementation and enforcement of the nondeterioration provision. The NCAQ was to give priority to a study of the implementation of the PSD provisions in the Clean Air Act and was required to submit a report of its conclusions to Congress within two years.[188] The report was to focus on whether the increments of change in air quality were appropriate to prevent significant deterioration of air quality in class I and class II areas. The NCAQ's authority in the area of significant deterioration, however, was to be constrained because the Randolph amendment gave the EPA authority over the funds needed to support the NCAQ study.[189] It would be directed to provide information to assist Congress in determining future air pollution programs.[190] According to Senator Randolph, the NCAQ was to assist Congress in closely watching the results of its efforts. Randolph added: "We must not forget to do this. We must monitor the program. We must have the oversight hearings. We must be very careful to see that the intent of Congress is carried out, as well as the actual language of the law."[191]

In debating the timetable Randolph proposed for the NCAQ study, Senator Frank Moss of Utah, an opponent of PSD, offered an amendment that sought to delete the provision related to prevention of significant deterioration so that there would be no legislation on that subject until the NCAQ's study was complete.[192] Moss's action was understandable, given that he represented a state with many class I areas (parks), and therefore sought to limit the applicability of PSD. The Moss amendment would have further altered the context and timetable of the reports to be submitted. Investigations were to consider the effects of "existing or proposed national ambient air quality standards," as well as the effects of any "existing or proposed policy of

---

187 Id.

188 122 Cong. Rec. 23,961 (1976) (provision of the Randolph amendment).

189 Id.

190 Id.

191 Id. at 23,962 (statement of Senator Randolph).

192 Id. at 25,148.

prohibiting deterioration of air quality in areas identified as better than required."[193] The NCAQ would be directed to submit a report of appropriate recommendations and results of studies one year after enactment. A second report was to be submitted "with regard to all other Commission studies and investigations, together with any appropriate recommendations, not later than three years after the date of enactment" of the bill.[194]

The proposed Moss study would have addressed the concerns that nondegradation would have an adverse impact on the economy and land use. The primary intent of the Moss amendment was to strike the nondeterioration proposal so that there would be no legislation until the Commission's report was complete. It was argued that the delay in legislation would provide Congress with the opportunity to obtain adequate information before arbitrarily setting emission standards. During the time the study was to be conducted, existing EPA regulations would remain in force.[195] The Moss amendment was supported by the National Construction Industry Council,[196] the National Rural Elect Cooperation Association,[197] President Gerald Ford,[198] and the Builders and Construction Trades of the AFL-CIO.[199] Nevertheless, the Senate rejected the amendment by a 63-31 vote.[200]

Two other Senators from areas that would have been hard hit by PSD requirements introduced similar amendments. Senator James Allen of Alabama offered an amendment to the Randolph amendment that would have provided for the enactment of the significant deterioration proposal (similar to the Randolph amendment discussed above), but would have suspended implementation and enforcement of PSD rules until one year after the Commission's report.[201]

In another attempt to delay PSD regulations, Senator William Scott of Virginia proposed that "during the period of the [NCAQ] study . . . nothing in [the Clean Air Act] shall be construed to require or provide for the establishment of Federal standards more stringent than primary and secondary air quality standards."[202] The Senate rejected both the Scott and the Allen amendments by substantial majorities,[203] but approved the Randolph

---

193 Id.
194 Id.
195 Id. at 24,536.
196 Id. at 25,150.
197 Id. at 25,156.
198 Id. at 25,150.
199 Id. at 15,155.
200 Id. at 25,192.
201 Id. at 25,541.
202 Id. at 25,159.
203 The Scott amendment was defeated by a vote of 74-17, id. at 25,178, and the Allen amendment was defeated by a vote of 59-23, id. at 25,550.

amendment by the overwhelming margin of 83-1.[204]

The House also debated the role of the NCAQ. The original House bill did not provide for such a commission,[205] so Representative Paul Rogers of Florida, Chairman of the House Subcommittee on Health and the Environment, offered an amendment calling for the formation of the NCAQ.[206] The Rogers amendment was similar to the Randolph amendment in the Senate in that it would have implemented the new PSD regulations and a study by the NCAQ on air quality deterioration concurrently.[207]

Like Senator Randolph, Representative Rogers faced a series of amendments from colleagues who sought to delay implementation of the new PSD regulations until the NCAQ study was completed. Claiming that the House proposal for new nondeterioration policies was based on "scanty information," Representative Bill Chappell of Florida offered an amendment—similar to the Moss amendment in the Senate—that would have left the existing EPA regulations on nondeterioration unaffected until the NCAQ study was concluded.[208] Representative Chappell argued that his amendment would provide Congress and the EPA with an opportunity to learn more about the effects of the nondeterioration proposal on the economy and the environment.[209] But like the Senate, the House rejected all attempts to delay implementation of the study, and rejected the Chappell amendment by a vote of 199-156.[210] It approved the Rogers amendment by a margin of 301-57,[211] and the final version of the House bill contained the NCAQ proposal.[212]

Ultimately, the composition of the NCAQ in the 1977 Amendments reflected a compromise between the House and Senate over which constituent group interests to incorporate into EPA decisionmaking. As enacted, the bill provided:

> Such Commission shall be composed of eleven members, including the chairman and ranking minority member of the Senate Committee on Public Works and the House Committee on Interstate and Foreign Commerce . . . and seven members of the public appointed by the President . . . . Not more than one-third of the members of the Commission may have any interest in any business or activity regulated

---

204 Id. at 25,552.

205 See H.R. 10,498, 94th Cong., 2d Sess., 122 Cong. Rec. 29,216 (1976).

206 122 Cong. Rec. 29,234 (1976).

207 Id.

208 Id. at 29,244.

209 Id. at 29,255.

210 Id.

211 Id. at 29,242.

212 Id. at 29,243.

under this Act.[213]

Interestingly, in the final version of the bill, members of Congress on the Commission did not serve ex officio, but instead had an active role in shaping the Commission's studies and recommendations.[214]

The focus of the 1977 Clean Air Act Amendments also reflected a compromise between the Senate and House proposals. Under the 1977 Amendments, the Commission was to focus on both emissions from mobile sources and the implications of significant deterioration.[215] The Commission's jurisdiction over the issue of nondegradation, however, was severely limited by the requirement that the NAS conduct studies on the same subject.[216]

The debate over the creation of the NCAQ illustrates three important points. First, the composition of its membership was of critical importance, for it determined which constituents were represented and, therefore, which constituents had the ability to affect the EPA's activities through the NCAQ's study. This aspect of the NCAQ was one important difference between the Senate and House provisions, with each favoring a panel that represented their interests. Second, the NCAQ demonstrated congressional intent to use studies to oversee agency activity. Again, the House and Senate differed as to the principal mission of the NCAQ, each wanting it to focus on different aspects of EPA policy. Third, the attempts to use the NCAQ to forestall the implementation of PSD rules shows how procedures, in this case a simple study, could be used to affect agency decisionmaking and the policies that resulted from it.

D. Warranties

Another major point of controversy in the 1976 and 1977 debates was over the establishment of emission standards for mobile sources. The history of warranties for auto emissions control devices provides additional insight into the use of procedures for policy purposes, as well as the differences between the policy objectives of the House and Senate. The 1970 Amendments required automobile manufacturers to provide a warranty on pollution control devices for five years or 50,000 miles.[217] Soon thereafter, it became apparent that manufacturers would respond by requiring that their emission control devices be maintained and repaired by their own authorized

---

[213] Clean Air Amendments of 1977, Pub. L. No. 95-95, § 313, 91 Stat. 685, 787, repealed by Act of July 2, 1980, Pub. L. No. 96-300, § 1(c), 94 Stat. 831 (1980).

[214] Id., 91 Stat. at 787.

[215] Id., 91 Stat. at 785.

[216] Id., 91 Stat. at 787-88.

[217] Clean Air Amendments of 1970, Pub. L. No. 91-604, § 6(a), 84 Stat. 1676, 1692 (codified at 42 U.S.C. § 7521(d)(1) (1982)).

mechanics—for the most part the dealers who sold their cars. Not surprisingly, independent automobile mechanics expressed concern that by linking the warranty to dealer maintenance, they would lose a substantial portion of their regular repair business to the authorized dealers. The independent mechanics received some support from the Federal Trade Commission (FTC), which warned the EPA that the warranty provisions would be anticompetitive if they unfairly tied a car owner to the dealer who sold the car.[218]

The House, ever sensitive to the concerns of any industry, responded by making three proposals in its 1976 bill. First, it limited the warranty to eighteen months or 18,000 miles, thereby shortening the period of tied maintenance.[219] Second, it authorized the FTC to undertake a study of the competitive effects of its warranty requirements, thereby formally requesting that the agency elaborate on its earlier warnings to the EPA. Third, it authorized the EPA to increase the warranty to five years or 50,000 miles, but only if the FTC agreed. This gave the FTC a veto power over the EPA's decision. Moreover, it required that the EPA find that the longer warranty had "no significant anticompetitive effects."[220]

The Senate's actions were similar to those of the House in some ways but not others. The Senate Public Works Committee proposed to keep the five year/50,000 mile requirement; however, it also authorized the FTC to study its anticompetitive effects. Moreover, the Committee asked the FTC to report to Congress, not to the EPA.[221] But it did not allow the FTC to veto EPA decisions, nor did it set up a procedure under which the EPA could alter the terms of warranties.

In the floor debate, Senator Lloyd Bentsen of Texas proposed an amendment almost identical to the House proposal, which expressed his deep concern about "the 400,000 independent repair shops in this country, 1,700 independent parts manufacturers, and 22,000 independent parts distributors" that he believed would be adversely affected by retaining the five year/50,000 mile warranty instead of adopting the eighteen month/18,000 mile warranty proposal.[222] Bentsen's proposed amendment was defeated by a 51-45 vote.[223]

The 1977 House and Senate bills were essentially identical to their proposals of a year earlier, and so a conference committee was again called upon to resolve the dispute. The tradeoff was clear enough. Automobile manufac-

---

[218] S. Rep. No. 717, 94th Cong., 2d Sess. 69 (1976).
[219] H.R. Rep. No. 1175, 94th Cong., 2d Sess. 235 (1976).
[220] Id. at 236.
[221] S. Rep. No. 717, 94th Cong., 2d Sess. 72 (1976).
[222] 122 Cong. Rec. 24,299 (1976).
[223] Id. at 24,313.

turers could not fairly be required to give a warranty on a device that others would maintain and repair; however, an important constituency was deeply offended by the government requiring that all its customers do business with its competitors. The Senate was willing to study the issue, but in this case wanted the more stringent provision. The House wanted to allow the FTC, an antitrust watchdog, to veto any proposal for more than a minimum warranty period. Knowing the proclivities of the FTC, which had already expressed its position on the matter, one could rationally expect that this was tantamount to enacting the weaker warranty.

Congress compromised on an intermediate warranty of two years or 24,000 miles.[224] Interestingly, although both the House and the Senate had passed bills requesting an FTC study, the conference report (and the 1977 Amendments) contained no such provision. In addition, the bill did not require that the EPA revise the warranty requirement and, of course, did not give the FTC a veto power over such a proposal.

CONCLUSION

The theory developed in this Article focuses on the problem of controlling bureaucratic agents and has two implications. First, it shows how an ex post legislative solution is a cumbersome and generally ineffective tool against bureaucrats who deviate from the policies preferred by politicians. Within limits, politicians are unlikely to be able to reverse a bureaucratic decision so long as bureaucrats deviate in a way that makes one of the major political actors better off than they were under the status quo. But this does not imply that the situation for politicians is hopeless. The second implication of the theory is that the main avenue for controlling bureaucrats is to place ex ante procedural constraints on the decisionmaking process. If these constraints allow politicians to respond to agency deviations *prior* to the agency's implementation of a change in policy, they can provide effective control over agency decisions.

We showed how administrative procedures play this critical role, and that they do so in several ways. First, by imposing a particular sequence on agency decisionmaking, they create an "early warning" system that alerts politicians (and their constituents) that an agency may attempt to change course. Second, the process itself imposes delay, affording ample time for politicians to intervene before an agency can present them with a fait accompli. Third, procedures allow politicians to adjust the set of pressures from the environment the agency faces, and in so doing, to stack the deck in favor of certain constituents.

---

[224] H.R. Conf. Rep. No. 564, 95th Cong., 1st Sess. 168, *reprinted in* 1977 U.S. Code Cong. & Admin. News 1502, 1549.

The Article ends with a case study of air pollution control regulation. This study shows that politicians were concerned with procedures as a means of influencing agency decisionmaking throughout the 1960s and 1970s. The discussion of the court's unanticipated change in policy, in its 1972 decision in *Sierra Club v. Ruckelshaus*,[225] shows how difficult ex post legislative changes are, and in particular, that politicians were unable to reestablish the status quo. The analysis in Part III shows that the 1977 Amendments to the Clear Air Act, the main legislative reaction to *Sierra Club*, were primarily procedural. The review of the considerations of this legislation reveals how different politicians attempted to use procedure to benefit their own constituents. Most of these attempts were unsuccessful because other legislators saw that the proposed changes would thwart their own policy objectives.

[225] 344 F. Supp. 253 (D.D.C. 1972), aff'd per curiam by an equally divided Court sub nom. Fri v. Sierra Club, 412 U.S. 541 (1973).

This work is distributed as a Policy Paper by the

CENTER FOR ECONOMIC POLICY RESEARCH

100 Encina Commons
Stanford University
Stanford, California 94305
(415) 329-1211

July 1984

CEPR Publication No. 30

THE POLITICS OF DERUSTICATION: COURT-ORDERED REDISTRICTING AND ITS POLICY EFFECTS

by

Mathew D. McCubbins* and Thomas Schwartz*

*Department of Government, University of Texas, Austin, Texas 78712.

We thank Peter Aranson, Bruce Cain, Charles Cnudde, John Ferejohn, Norman Frohlich, Melvin Hinich, Roger Noll, Joe Oppenheimer, Peter Ordeshook, Benjamin Page, William Riker, Kenneth Shepsle, Terry Sullivan, Barry Weingast, and especially Roderick Kiewiet for helpful discussions.

For all the talk about civil rights, consumerism, environmental protection, transportation deregulation, and other policy trends, one of the biggest policy stories of the past two decades has received scant attention. It is a story that subsumes many others and is yet to end. We are referring to the continuing reallocation of the net benefits of federal-government activity from rural to nonrural Americans. A host of policy initiatives and pullbacks, dissimilar in other respects, fits this pattern. The principal cause, we suggest, was court-ordered congressional redistricting based on the one-man, one-vote rule. This led immediately to the derustication of Congress, bringing about a nonrural House majority for the first time, and ultimately to a series of policy changes that benefited metropolitan citizens at some cost to rural citizens.

Previous empirical research found little or no evidence of policy changes attributable to redistricting (Jacob 1964, Dye 1965, Hofferbert 1965, Brady and Edmonds 1967, Fry and Winters 1970, Erikson 1973). Although Pulsipher and Weatherby (1968), Hanson and Crew (1973), and Frederickson and Cho (1974) reported some policy effects, Newcome and Hardy (1980) pointed out problems of multicolinearity in these studies and showed that respecification of the models wiped out positive results. All this work focused on policy effects at the state level--and, as Erikson (1973) has pointed out, some of it was based on rather strong assumptions about policy preferences. We instead examine the effects of congressional redistricting on national policy.

After describing the relevant court rulings and their immediate effects on congressional representation, we explain the content and timing of these decisions on the basis of general judicial goals and constraints. Next we show how a rural-to-nonrural reallocation of policy benefits would predictably

follow the imposition of the one-man, one-vote rule. We go on to test our reallocation hypothesis by regressing certain budget changes that benefited nonrural Americans against a measure of court-ordered changes in congressional districts. Having described our engine of change and showed that it is at work, we then survey several salient policy areas, such as agriculture, health-and-safety regulation, and the deregulation of transportation, and argue that changes in these areas over the past two decades can be described as part of a rural-to-nonrural reallocation.

1. Court-Ordered Redistricting and its Effects on Representation

In Baker v. Carr, 1962, the U.S. Supreme Court ruled that federal courts could hear challenges to legislative apportionments. This reversed a contrary ruling by a lower court on a challenge to Tennessee's apportionment law, which gave rural areas more state legislators per population than urban areas had. Within a year, federal courts decided apportionment cases in twenty-five states, finding the legislatures of nineteen states to be malapportioned in violation of the equal-protection clause of the Fourteenth Amendment, while fifteen states passed reapportionment acts. Between 1963 and 1965, eighteen more states reapportioned their legislatures.

Early in 1964, in Wesberry v. Sanders, the Supreme Court found for petitioners in a challenge to Georgia's system of congressional districts. The Court called for districts of equal size, requiring that "as nearly as practicable one man's vote in a congressional election is to be worth as much as another's." In Reynolds v. Sims, 1964, the Court found Alabama's "little federal system," in which state senators represented single counties of unequal size, to be unconstitutional: "Legislators represent people, not

trees or acres. Legislators are elected by voters, not farms or cities or economic interests." The Court strengthened the one-man, one-vote rule in Kirkpatrick v. Preisler, 1969, mandating that any variance between districts, however slight, must be justified or shown to have come about in spite of "good faith" efforts. To enforce these rulings, federal courts engaged in a massive redistricting effort between 1964 and 1970. Details are displayed in Table 1.

The effect on equality of representation was dramatic. As judged by the 1960 census, the ratio of largest to smallest population was 242 to 1 for Connecticut state House districts, 223 to 1 for Nevada Senate districts, 141 to 1 for Rhode Island Senate districts, 9 to 1 for Georgia Senate districts, 4 to 1 for Texas congressional districts, and 3 to 1 for Arizona, Maryland, and Ohio congressional districts. In the 88th Congress, elected in 1962, only nine districts were within one percent of the average district size in their states, and 236 districts deviated at least 10 percent each from the average. But in the 93rd Congress, elected in 1972, fully 385 districts deviated less than one percent from the average.

Particularly striking was the effect of redistricting on rural versus metropolitan (urban plus suburban) congressional representation. The 1920 census showed that urban Americans were by then a majority (suburbs being as yet undreamt of). But as late as 1964, rural districts had a virtual majority of U.S. representatives. By 1972, however, metropolitan districts had a majority, as shown in Table 2. (The 1972 figures in Table 2 do not, however, disentangle demographic changes, reflected in the 1970 census, from changes in districts brought about by redrawing boundaries.)

TABLE 1

Congressional Redistricting, 1964-1970

Year	Congress	Number of states redistricted	Number of districts affected
1964	89th	8[a]	65
1966	90th	24	234
1968	91st	17[b]	237
1970	92nd	6	60

[a]Includes three states that were redistricted by state legislatures prior to Wesberry v. Sanders.

[b]Includes some of the biggest states.

Source: Congressional District Data Book, 93rd Congress, Bureau of the Census: 549.

2. Through the Political Thicket with Gun and Camera

Why did the Supreme Court act as it did, when it did?

Like everyone else, Supreme Court justices pursue goals subject to constraints. Their goals, we assume, are to enunciate and enforce their values, to preserve and enhance their power, and to make their marks on history in a way that brings them honor. In the case at hand, the Warren Court certainly were keen to promote liberal values associated with equality and the rights of citizens; review of district maps certainly would enhance their power; and massive Court-initiated redistricting to achieve equality of population certainly would stand as a major democratic electoral reform in the honored tradition of the Reform Act of 1832. Warren has described Baker v. Carr as the "most vital decision" handed down during his tenure.

Yet even liberal courts can be conservative in the constraints they observe. Here we mention four[1]:

RESPECT FOR PUBLIC AND INSTITUTIONAL OPINION Because the Court can enforce its decisions only when other branches and levels of government recognize its authority, it must have some respect for public opinion and beware of hostile reactions by other institutions, especially when it claims a jurisdiction previously reserved to the legislative branch or to the states. Seeking power and honor, Supreme Court justices try to avoid constitutional crises lest their authority be impaired or their mark on history be a blemish.

TABLE 2

Characteristics of House Districts,
1960-1973

	1960	1964	1968	1972
Metropolitan	153	153	214	233
Rural	212	203	155	130
Mixed	72	79	66	72

Source: Congressional Quarterly Weekly Report 32 (April 6, 1974), p. 878, and Congressional Quarterly Census Analysis: Congressional Districts of the United States.

RESPECT FOR PRECEDENT (STARE DECISIS) The Court is bound, to a degree, by precedent. This constraint doubtless is enforced by conscience as well as by the need for broad acceptance of the Court's authority. Here the most important precedent was Colgrove v. Green, 1946, in which the Court rejected a challenge by urban voters to Illinois's legislative apportionment. Writing for the Court, Frankfurter foreswore the "political thicket" of apportionment cases, owing to their "peculiarly political nature." True, in Gomillion v. Lightfoot, 1960, the Court threw out an Alabama law that redrew city boundaries to dilute black voting strength. But the opinion rested on the narrow Fifteenth Amendment prohibition against racial disenfranchisement.

AVAILABILITY OF ALTERNATIVES The Court tends to avoid solving problems by asserting authority over other branches and levels of government when nonjudicial remedies are available. In Colgrove, Frankfurter wrote: "the remedy for unfairness in districting is to secure State legislatures that will apportion properly, or to invoke the ample powers of Congress."

PRUDENCE IN SETTING PRECEDENT The Court tries to avoid deciding cases on dangerously broad grounds--on grounds that might set a precedent for bad decisions or give rise to cases that the Court had rather not decide at all. In the case at hand, the Court would have preferred not to base redistricting decisions solely on the equal-protection clause of the Fourteenth Amendment, proscribing legislative districts of unequal size solely on the ground that such districts put some groups--urban residents, say--at a political disadvantage. This is because any legislative apportionment works to the foreseeable political disadvantage of some faction or interest group. The Court sought to fashion a shining and durable monument, not a tar baby.

Given these goals and constraints, the following six observations help explain the series of Court actions on redistricting in the 1960's:

First, the problem was growing worse: differences in district size were continuing to increase (Baker 1966: 25 and 34-35; David and Eisenberg 1961 and 1962). Given its values, then, the Court's incentive to act had increased.

Second, legislative remedies, state and federal, were effectively unavailable, despite increasing complaints. Rural legislators had no incentive to destroy their majorities. Even urban legislators did not have an unqualified incentive to increase urban legislative voting strength: why should an individual legislator dilute his own power? In times past, Congress had played a more active role. An 1872 federal apportionment statute required that U.S. Representatives "be elected from districts composed of contiguous territory and containing as nearly as practicable an equal number of inhabitants,. . .no one district electing more than one Representative." This requirement was repeated in post-census apportionment acts through 1911. But after the 1920 census, which for the first time showed rural residents to be a minority, no apportionment act was passed until 1929, when Congress enacted a formula for apportioning House seats to the states, left the application of this formula to the executive branch, and dropped the requirement of equal-size, single-member districts.

Third, by 1962 the nation had grown used to broad judicial interpretation of citizenship rights and to federal-court intervention in areas previously reserved to other branches and levels of government. Therefore, although the Supreme Court was yet to assert its authority over

apportionments, its respect for precedent and for public and institutional opinion had become less severe constraints on apportionment decisions than they may once have been.

Fourth, Colgrove v. Green did not really support the doctrine that federal courts lacked jurisdiction in apportionment cases. A seven-member court had found against petitioners by a four-member majority. Only two other members of the majority joined Frankfurter's opinion, which denied judiciability. The fourth member voted against petitioners on merit but agreed with the three dissenters that federal courts could hear apportionment cases. Although a majority ruled against the petitioners, another majority upheld judiciability. This point was made in oral argument by Charles S. Ryne in the first of two Supreme Court hearings (1960) on Baker v. Carr. Frankfurter agreed from the bench, and Brennan repeated the point in the majority opinion.

Fifth, the Court proceeded with caution, gradually strengthening its stand as its actions found general acceptance. In Baker v. Carr the Court merely ruled (citing Colgrove) that apportionment cases were judiciable, sending the case back to Federal District Court for a hearing on merit. The Supreme Court did not even state principles for the lower court to follow. As Stewart emphasized in a concurring opinion, the Court was not requiring "equality of voice of every voter" and was not deciding the question: "may a State weight the vote of one county or one district more heavily than it weights the vote of another?" The Court deflected potential opposition by avoiding any substantive ruling and by shifting immediate responsibility to lower courts, leaving itself free to overturn lower-court decisions in the face of bitter or widespread opposition.

Not until 1964, by which time state-legislative redistricting was well under way and had met no major negative reaction, did the Court rule on substantive issues. Even then the Court hedged its statement of the one-man, one-vote rule. In Reynolds v. Sims, Warren called for "substantial equality among the various districts, so that the vote of any citizen is approximately equal in weight to that of any other citizen in the State. So long as the divergences from a strict population standard are based on legitimate considerations incident to the effectuation of a rational state policy, some deviations from the equal-population standard are constitutionally permissible. . . ." Warren went on to list contiguity, compactness, and integrity of political units as examples of potentially reasonable goals for which some equality of population might be sacrificed. The Court came much closer to a hard, mathematical standard of equality only in Kirkpatrick v. Preisler, 1969.

Sixth, in the first of the 1964 cases, Wesberry v. Sanders, the Court avoided basing its opinion on the dangerously broad equal-protection clause of the Fourteenth Amendment. Instead it appealed to Article I, Section 2, of the U.S. Constitution, whose "command. . .that representatives be chosen 'by the People of the several States' means that as nearly as practicable one man's vote in a Congressional election is to be worth as much as another's." (We did not list logic as a constraint.)

3. From Votes to Policy

In §§4 and 5 we describe a number of policy changes that have reallocated government benefits from rural to nonrural voters. Our explanation of this reallocation has two steps: First, redistricting based on

the one-man, one-vote rule brought about a nonrural majority in the House of Representatives, which had had a rural majority. Second, this majority made policy changes to benefit its own constituents. Let us consider each step in turn before adding flesh to the bare bones of our reallocation hypothesis:

FIRST STEP: FROM REDISTRICTING TO A HOUSE MAJORITY A majority of Americans have long been nonrural, or metropolitan: they live within Standard Metropolitan Statistical Areas (SMSA's), which consist chiefly of inner cities and suburbs. Drawing single-member congressional districts of equal population makes it likely that nonrural Americans--or, indeed, any given majority of Americans--will be a majority within a majority of districts.

Although likely, this is not logically necessary. A majority of citizens can control just a minority of congressional districts if they are highly concentrated in some districts and highly dispersed among many others.

But such a pattern would be extremely unlikely in the case at hand for four reasons: First, metropolitan residents are not only a majority but an overwhelming one. Second, it is almost impossible for a significant number of metropolitan voters to be highly dispersed among numerous districts because, by definition, metropolitan voters live in geographic clusters. Third, the state legislatures that might have contrived to under-represent metropolitan voters had themselves been redistricted. Finally, courts and even state legislatures observe severe technical limits on the creativity with which district lines can be drawn (see Cain 1984).

SECOND STEP: FROM MAJORITY TO POLICY Our contention is that the shift from a rural to a nonrural House majority brought about policy changes that reallocated net benefits of government activity, to some degree, from rural to

nonrural residents. Our underlying assumption is that if a group of voters with some common, salient interests increases its congressional representation from a minority to a majority, there will be policy changes which benefit that group.

Here are five objections one might raise to this assumption, at least as it applies to nonrural voters[2]:

First objection: The House of Representatives does not decide policy by itself. The Senate and President must concur.

Reply: Since a rural-to-nonrural shift in the House position on an allocation issue is almost certain to bring about a shift of the same direction (although not necessarily of the same magnitude) in the final position, there is no need explicitly to bring in the Senate or President. In a larger study, however, it would be interesting to compare the policy positions and the electoral-support bases of the House, Senate, and President.

Second objection: Many important policy decisions, including ones that affect the rural-nonrural allocation, are not passed as acts of Congress but are made by the bureaucracy.

Reply: We have argued elsewhere (McCubbins and Schwartz 1984) that Congress' oversight system gives it greater control over the bureaucracy than many scholars have thought it had--great enough control to enable congressmen to pursue their (changing) legislative goals through the bureaucracy without passing new legislation (see also Weingast 1983).

Third objection: The rural minority might preserve its position by luring <u>part</u> of the metropolitan majority into a new majority coalition.

Reply: The policy package produced by such a coalition would still allocate more to metropolitan interests, and possibly less to strictly rural interests, than would the policy package produced by a rural majority.

Fourth objection: Owing to the power of congressional committees and their chairmen, it is naive to suppose that Congress operates by pure majority rule when deciding rural-nonrural allocations.

Reply: Since the 1960's the House has become more egalitarian and majoritarian: now a congressman can chair only one committee or subcommittee; seniority is no longer rigidly adhered to; committee chairmen have lost their control over subcommittees, which have increased in number and often are chaired by comparatively junior members; the party caucuses have asserted their control over assignments and chairmanships, voting on chairmen one at a time by secret ballot; and the Speaker's power has increased (Dodd and Oppenheimer 1977, Sundquist 1981). Similar reforms have occurred in the Senate (Ornstein et al 1977). More generally, because the rules and organization of the House are themselves ratified by majority rule, they are not likely to work to the disadvantage of a well-defined majority in any flagrant, systematic way.

Fifth objection: On so-called distributive issues, such as public works, Congress acts on a norm of universality rather than majority rule: distributive measures--measures that provide divisible, locally targeted benefits, singly or in a package--tend to enjoy near unanimous support (Lowi 1964, Ferejohn 1974, Stockman 1975, Fiorina 1978, Shepsle and Weingast 1980, Schwartz 1983).

Reply: This may be true, but it is compatible with our analysis. If rural districts are a majority, any pork barrel of distributive benefits,

containing something for each district, will normally provide more for the rural majority than for the nonrural minority. If nonrural districts are a majority, the same will be true in reverse. Either way, the total rural-nonrural allocation will favor the majority.

SPECIFIC REALLOCATION HYPOTHESES Our hypothesis that the new nonrural congressional majority would retarget policy benefits toward its own constituents allows us to predict a number of more specific policy changes:

- We may expect a decline in rural distributive benefits, notably price-income supports for farmers. Note, however, that support payments can fluctuate with economic conditions and that support programs are entitlements based on formulas: having authorized a price-income-support formula, Congress includes a mere estimate of its cost in the federal budget, allowing the Commodity Credit Corporation to borrow from the Treasury to cover any under-estimate.

- Agriculture programs should be increasingly directed at nonrural interests, such as consumers and the urban poor: they should increasingly provide packages of rural plus nonrural distributive benefits. Seeking to maintain a flow of agricultural benefits in the face of declining legislative power, rural congressmen will have to form coalitions with some metropolitan representatives. Even if strictly agricultural benefits did not decline, such packaging would bring about some reallocation. It is important to note that our reallocation hypothesis is compatible

with an increase in spending for programs run by the USDA, so long as benefits are directed away from farmers.

° For similar reasons, rural congressmen will increasingly have to pay for rural distributive benefits by trading votes with nonrural congressmen.

° Agricultural benefits should shift in the direction of large farmers, for the following reason: Apart from pleasing rural voters, congressmen want the nation to have a strong agricultural sector to enhance our balance of trade and to help ensure that Americans eat well at low cost. But this is better achieved by rewarding the most efficient farmers than by rewarding the most farmers. Owing to the high productivity of agribusiness, the most significant social consequence of redistricting may come to be the disappearance of the Jeffersonian yeomanry.

° Urban and suburban distributive benefits, such as federal housing programs, should increase.

° We may expect police-like regulatory activities to expand. This is because consumer, health, safety, and environmental regulation are demanded by urban and suburban voters.

° Cartel-like regulatory activities, notably transportation regulation, should reflect a reallocation. Because they had brought about metropolitan-to-rural cross-subsidies, such activities should undergo policy shifts that redirect benefits to metropolitan interests.

- We may expect Congress itself to undergo internal organizational changes by which rural congressmen lose power.

- In states in which one party is predominently rural or predominantly metropolitan, there should be an appropriate waxing or waning of partisan fortunes.

- The amount of federal and state revenue aid to units of government below the state level should reflect the populations of those units more than it did before redistricting.

4. Derustication and Federal Appropriations: An Econometric Test

Owing to the obvious problem of measuring certain kinds of policy change, a completely general test of our reallocation hypothesis would be difficult to construct. Insofar as the benefits from federal policy take the form of spending on projects, direct cash outlays, or bureaucratic services, then changes in policy should induce changes in the appropriations for the agencies administering these policies. Thus, budget figures may provide a measure of policy change. In this section we will examine changes in the budgetary fortunes of a set of federal programs as a "first-order" test of our hypotheses.

Funding levels for federal programs are determined through strategic interaction between the president and Congress. The interplay between the two branches over appropriations resembles a sequential bargaining

game. The Office of Management and Budget (OMB) working for the president frames a budget request for each agency's activities. By moving first in the bargaining game, OMB attempts through its choice and packaging of budget requests to control the agenda before Congress. In moving first, OMB's requests will likely be strategic, reflecting the administration's expectation of congressional action. On the other hand, because it moves second, Congress's power to amend OMB's budget requests is almost unconstrained. Congress may strategically anticipate or respond to presidential action on other issues or to a presidential veto, but, as intended by the founders, the power of the purse is primarily in the hands of Congress.

The strategy of either side in this bargaining game can be reactive or accomodative. A reactive strategy overstates or understates an appropriation to offset an understatement or overstatement by the other side. An accommodative strategy is one of bending with the wind: one side overstates or understates its own ideal appropriation to reflect the other side's preferences (Kiewiet and McCubbins 1984b).

Kiewiet and McCubbins (1984a,b) have developed and tested an "electoral-connection" model of the appropriations process that contains many of the elements of the bargaining game hypothesised here. In their model, the president and congressmen desirous of reelection, seek to maximize the electoral support (campaign contributions, votes, etc.) they receive from the various beneficiaries of government goods and services. They assume that groups and constituents base their support of incumbents at least in part upon their retrospective evaluations of the levels of benefits conferred on them by federal programs. The extensive literature on retrospective voting has

generated a great deal of evidence in this regard: voters are seen to have fairly large positive discount factors and are also "myopic," i.e. they discount the future costs of present benefits. The level of support received from each clientele group will be assumed to be a nondecreasing function of the appropriations awarded to the agency which benefits them. We also assume, however, that incumbents experience declining marginal returns in electoral support from the appropriations awarded to each agency. In order to maximize their reelection prospects, incumbents will thus choose appropriation levels which equate marginal returns in electoral support across federal agencies.

In contrast, the costs of these goods and services provided by agencies of the federal government are quite diffuse, as they are financed out of general revenue collected from all U.S. taxpayers. To be sure, congressmen may make tradeoffs between the campaign resources garnered through their support of various federal benefit programs and votes obtained by decreasing taxes. However, the fact that benefits tend to be concentrated and costs diffuse, combined with the fact that the budget constraint faced by the federal government is "soft" (it may and almost always does spend more than it takes in) means that appropriations decisions will not reflect the true tax costs of federal benefit programs.

Within such a model the derustication of Congress predicts a number of specific appropriations changes. For one thing, we would expect programs whose benefits largely flow from the federal treasury will directly reflect policy changes in their budgets. As the number of metropolitan groups within a district increases, programs concerned with providing public works, income subsidies, or grants-in-aid to narrowly targeted metropolitan interests should experience an increase in spending. On the other hand, funding of such

programs for rural constituents should decline as fewer congressmen include such groups in their reelection constituency. From our hypotheses in section 3, we also expect that programs which supply divisible benefits to both rural and urban constituents, as do school lunch programs, should receive greater funding as a result of redistricting. What is more, we should expect a budgetary expansion for regulatory agencies. This is largely because consumer, safety, health, and environmental regulations serve mainly metropolitan interests. Also, policy changes within these agencies, which shift the distribution of regulatory benefits to metropolitan constituents, will lead to expansions in the bureaucratic structure needed to handle these new claimants, thus necessitating an expansion in the operating budgets for these regulatory agencies.

The strategic nature of appropriations politics suggests that presidential budget requests may also be influenced by the increase in metropolitan representation in Congress. Witnessing the increased urbanization of Congress, the president may anticipate the increased demand for metropolitan spending accommodatively, by requesting increased funding for metropolitan programs. That way he appears to lead Congress and can claim credit for the benefits delivered to metropolitan constituents. The president can instead anticipate congressional actions with a reactive strategy, attempting to offset increased metropolitan demands.

To be sure, not all policy changes will show up as budgetary changes. In looking at the U.S. Department of Agriculture's (USDA) share of the federal budget we do not immediately witness much of a decline: The USDA received 6% of the total federal budget in 1965 and maintained 6% of the budget in 1980.

As we suggested in section 3 the size of the USDA budget is no measure of net federal benefits to rural constituents. If USDA programs have been retargeted, as we hypothesized, the size of budget may not reflect the policy change. The reallocation is evident upon disaggregating the USDA's budget. Within its budget new programs which supply benefits to metropolitan interests have emerged since 1965: food stamps has grown from a pilot program funded initially at $80 million in 1966 to almost $6 billion in 1979; school lunch programs have grown from $146 million in 1965 to over $1 billion in 1979; other programs, designed to benefit metropolitan consumer interests have had their budgets increased ten-fold in the last two decades; on the other hand, payments to farmers under the Sugar Act have declined from over $87 million in 1965 to zero in 1979.

On the other hand, funding levels for federal agencies may change for other reasons. In their studies (Kiewiet and McCubbins (1984a,b) found that several political and economic variables determine, at the margin, the budgetary fates of federal agencies. One of these factors affecting congressional and presidential attitudes toward spending is the proximity of election day. As the salience of electoral imperatives (raising money and garnering votes) increases as the election nears, congressmen will become increasingly anxious to channel government benefits to their constituents (or at least, alternatively, less anxious to cut such benefits). Such benefits, furthermore, will also become politically more salient to constituents. This tendency will be reinforced, of course, by congressmen knowing that over time beneficiary groups discount the previous benefits they have received. Maximizing the amount of electoral support earned by the flow of government

goods and services thus indicates "heaping" these benefits later on in the electoral cycle (Kiewiet and McCubbins 1984a:6). On the presidential side, this would take the form of larger budget requests for federal agencies during presidential election years. For Congress, however, any impact of election year considerations should register every even-numbered year. Although congressmen never like to cut programs that serve key interest groups in their districts, we would expect calls for economy in these areas to have the least impact as the second session of a Congress comes to a close.

Further, appropriations strategies should be a function of changing economic conditions. Fluctuations in the nation's economy clearly affect the outcomes of congressional elections (Kramer 1971), and campaign decisions of both incumbents and potential challengers take account of this (Jacobson and Kernell 1981). Contributors, in particular, can respond to higher rates of unemployment or inflation by either witholding funds from incumbents or by funding challengers. Further, many groups have come to accept the Keynesian axioms that the proper governmental action in the face of high unemployment is to increase spending, whereas high inflation should be countered by decreasing federal spending. We thus hypothesize that Congress and the president respond to high rates of inflation by appropriating less to government agencies. Conversely, we will hypothesize that high rates of unemployment lead Congress and the president to award higher levels of appropriations to government programs and agencies.

Another political variable which we would expect to exert a clear, continuous influence on appropriations decisions is party politics: Democrats spend more than Republicans. This hypothesis is compatible with our electoral-connection view of legislative decision making because Republicans

and Democrats have to please somewhat different reelection constituencies (Fiorina 1974; Fenno 1978; Poole and Rosenthal 1983).

Assuming that policy preferences reflect party, then, the nature of the appropriations game suggests that the period of transition from a Democratic to a Republican administration (or vice versa) provides us with an additional opportunity for measuring the extent to which the president's preferences are accommodated by Congress. A few days before leaving office in 1977, for example, President Ford submitted his budget requests for fiscal year 1978 to Congress. The president who ultimately had to sign the ensuing appropriations bills was Carter. During the transition to a Democratic administration congressional accommodation to the president should result, ceteris paribus, in higher levels of appropriations than if a Republican had remained in office. The transition to a Republican administration should have the opposite effect.

A final variable which previous research has shown to be a determinant of presidential and congressional bids--the involvement of the United States in armed conflict. Several studies, using a variety of methods and focusing on several different concerns, have generated considerable evidence that domestic agency budgets are pared back during times of war (Pressman and Wildavsky, 1973, p. 31; Okun, 1970, p. 78; Padgett, 1981; Mowery and Kamlet, 1982).

DATA

We collected OMB estimates and final appropriations figures from fiscal years 1948 to 1979 for three urban public-works programs, four farm programs, and five regulatory programs. Both sets of figures were reported in various regular annual appropriations acts. Many agencies often receive additional

funding in supplemental and deficiencies acts. These figures are almost always quite small, however, and including them in the agencies' yearly appropriations totals would have little effect upon the results of this analysis. The 12 agencies which we analyze are listed in Table 3. These twelve programs are the only ones which fit our hypotheses and for which reliable and long enough budgetary time series were available.

Measuring the extent of redistricting presents special problems. How do we discern changes in the composition of congressional districts due to redistricting as opposed to changes brought about through population growth and migration? Fortunately we can observe a controlled experiment, with treatments provided by federal courts. In ruling on the redistricting cases of the 1960's, the courts required states drawing new boundaries to use 1960 census figures. In describing the new congressional districts for the 89th and 90th Congresses (elected in 1964 and 1966), the Census Bureau used 1960 census information, thereby excluding the effects of migration and growth. For the 91st through 95th Congresses we employed information on boundary changes and statewide demographic changes to control demographic changes within each district and to reckon the changes in district composition that are due solely to redistricting. We discounted statewide demographic changes from the change within each congressional district. Changes in the classification of congressional districts due to redistricting are reported in Table 4.

Table 4 classifies congressional districts into five categories along a single underlying dimension reflecting urbanness. A taxonomy as simple as this one, of course, blurs many important distinctions. Congressmen representing predominately urban districts may still have sizeable and

important rural groups as members of their reelection constituencies, and as such will seek to continue an uninterrupted flow of divisible benefits to these groups. As a district becomes more urban, however, the legislative package sought by its representative will likely contain an increasing share of urban benefits. Therefore, support for urban spending should increase monotonically from rural to urban districts in Table 4, while support for rural spending should decrease monotonically from rural to urban districts.

Table 4 shows that the impact of court-ordered redistricting was far greater and occurred much earlier than the conventional view suggested. During the 88th Congress, 105 congressional seats were rural while 119 were urban. By the 93rd Congress only 17 seats remained rural, whereas 186 seats were then classified as urban. Of these, all but 13 of the decline in rural seats and three of the increase in urban seats were accounted for by redistricting. Contrary to the conventional wisdom, which held that the main effects of redistricting were first generated by the 1970 reapportionment and subsequent redistrictings, Table 4 shows that more than 70 percent of the changes due to redistricting were imposed by the courts in the 1960's.

Since the measures in Table 4 are derived from data on Standard Metropolitan Statistical Areas, our measure of urbanness combines urban and suburban constituencies. Although urban and suburban constituents have different preferences on some issues, these differences are not likely to be great for the issues we have chosen to examine.

METHODOLOGY

The first issue that must be addressed concerns the way in which the two strategic variables (OMB requests and congressional appropriations) should be specified. Our hypotheses about the effects of redistricting concern real levels of funding. The nominal requests and appropriations figures are converted into constant (1972) dollars (deflation will be based upon the Implicit Price Deflator for Federal Government Goods and Services). Furthermore, our hypotheses concern the relative distribution of benefits between rural and urban areas. The real-dollar OMB estimate and congressional appropriations figures will therefore be divided by the total real-dollar appropriations passed in the given fiscal year, producing the two strategic terms to be used in the analysis. (The figures were multiplied by 1000 to avoid rounding problems. Thus, to interpret the coefficients one must divide them by 1000.)

Because of the sequential nature of the bargaining game, in which OMB is choosing its strategy in anticipation of congressional action, it is necessary to use some form of instrumental-variables technique to estimate OMB's forecast of congressional action. The particular technique employed is two-stage least squares, which constructs instruments for the endogenous congressional forecast in the OMB equation. Because final congressional action on the president's budget requests occurs several months after these figures are submitted, the OMB term in the congressional equation is treated as exogenous, and the congressional equation is estimated by ordinary least squares.

TABLE 3

Agencies and Programs

Housing Programs

- Community Development Grants
- Housing Assistance Programs
- Manpower Programs (CETA)

Farm Programs

- School Lunch Programs
- Commodity Programs
- Conservation Reserve Programs
- Sugar Act

Regulatory Programs

- Civil Aeronautics Board
- Federal Power Commission
- Federal Trade Commission
- Interstate Commerce Commission
- Security and Exchange Commission

TABLE 4

Change in Congressional Districts
Due to Redistricting

	Rural	Predominately Rural	Mixed	Predominately Urban	Urban
88th	0	0	0	0	0
89th	-5	-9	+1	+2	+11
90th	-36	-13	+8	+6	+34
91st	-6	-13	-2	+3	+18
92nd	-1	+1	-2	+1	+1
93rd	-27	+12	+12	+3	0
94th	0	0	0	0	0
95th	0	0	0	0	0

A congressional district is defined rural if ten percent or less of the district resided within an SMSA, predominately rural if more than ten percent but not more than forty percent resided within an SMSA, mixed if forty to sixty percent resided within an SMSA, predominately urban, if sixty to ninety percent resided within an SMSA, and urban if more than ninety percent resided within an SMSA.

Probably the most serious estimation problems faced in this analysis derive from the fact that the number of observations for each program is quite limited; the full time series is only 32 years long, and for some programs the number of observations is less than that. Pooling the data across categories of the programs listed in Table 2 thus presents some important advantages. Besides offering a gain in statistical leverage by increasing the number of observations, it also simplifies the test of the hypotheses, as each of the major hypotheses can be tested by estimating a single coefficient for each sample of programs. Problems of nonstationarity and of heteroscedastic error variances, often present in the analysis of time-series and pooled data, are avoided here because the dependent variables are expressed as the relative share each agency received of the total budget.

Another problem that can be anticipated when pooling data in this manner is cross-sectional correlation. Factors that produce error in predicting appropriations for agency i in year t may well produce error in predicting appropriations for agency j in year t. The resultant correlation between error terms for the same year will lead to estimated standard errors that are downwardly biased. It will thus be important to calculate the degree of correlation between residuals for the various agencies to determine whether or not this poses a serious problem.

Yet another problem is that to pool cross-sectional time-series data is implicitly to assume that the coefficients for the explanatory variables are equal across each cross-section. It may be that coefficients differ across programs. An F test can be calculated, using estimated error variances from the unrestricted 2sls and the restricted (i.e., common coefficients) 2sls estimation.

The most obvious methodological problem is how to construct and incorporate a measure of redistricting effects. Constructing a composite redistricting "score" from Table 4 presents many problems. Such a score would require too much from the taxonomy in the Table, and any deviations from monotonicity of support across categories would introduce error in our measure. We instead focus on the cumulative increase in the urbanness of Congress. Because we predict that policy will change as a result of an increase in the number of urban seats attributable to redistricting, our focus on the urban category of Table 4 tests our hypothesis directly.

The use of this measure has important advantages. By restricting ourselves to the urban category, we need not make the strong assumption that the likelihood of support for urban (rural) spending bills increases (decreases) across categories in the Table. By restricting our measure to districts that are overwhelmingly urban (more than ninety percent of the residents live within an SMSA), we can reasonably assume that increases in the number of these districts led directly to increases in the urban coalition in Congress.

One last potential problem is <u>serial correlation</u>. Unfortunately, common methods of diagnosing the severity of this problem (e.g., the calculation of a Durban-Watson statistic) are not applicable in the context of pooled cross-sectional time series data. We will need to examine the errors produced by estimating Equations 1 and 2 below individually for each program.

We tested our hypotheses by estimating the following two equations:

$$\Delta OMB_{it} = a + g_1 \Delta CONG^{*}_{it} + g_2 I^{p}_{t-1} + g_3 U^{p}_{t-1} + g_4 PRES_t + g_5 WAR_t + g_6 E^{p}_{t} \quad (1)$$
$$+ g_7 REDIST + u_{it} ,$$

$$\Delta CONG_{it} = d + b_1 \Delta OMB_{it} + b_2 I^{c}_{t-1} + b_3 U^{c}_{t-1} + b_4 HOUSE_t + b_5 E^{c}_{t} + b_6 TRANS_t \quad (2)$$
$$+ b_7 REDIST_{t-1} + b_8 WAR_t + e_{it}$$

where:

$'OMB_{it}$ = the appropriations (in constant dollars) requested by the president in the form of the OMB estimate for agency i in fiscal year t as a proportion of the total appropriations for that fiscal year

$\Delta CONG_{it}$ = the appropriations (in constant dollars) awarded by Congress to agency i in fiscal year t as a proportion of the total federal budget for fiscal year t ($\Delta CONG^{*}_{it}$ is an instrument reflecting OMB's forecast of $CONG_{it}$)

a and d = constant terms

I^{p}_{it} = the (annualized) percentage change in the Consumer Price Index during the six months prior to the president's submission of agency budget requests to Congress (for FY1965, for example, this measure would register the rate of inflation during the last half of calendar year 1963)

I^{c}_{it} = the (annualized) percentage change in the Consumer Price Index during the first six months of the session of Congress in which appropriations for a given fiscal year are considered (for FY1965 this measure would register the rate of inflation during the first half of calendar year 1964)

U^{p}_{t-1} = the average rate of unemployment during the six months prior to the president's submission of agency budget requests to Congress

U^{c}_{t-1} = the average rate of unemployment during the first six months of the session of Congress in which appropriations for a given fiscal year are considered

$PRES_t$ = a dummy variable that takes the value 1 when the president is a Democrat and 0 when he is a Republican

WAR_t = a dummy variable that takes the value 1 if the U.S. was engaged in a war in fiscal year t and 0 otherwise (these were fiscal years 1952-54 for the Korean Conflict and 1967-74 for Vietnam)

$HOUSE_t$ = the percentage of seats in the House held by Democrats

E^{c}_{t} = a dummy variable that takes the value 1 for budgets considered during congressional election years (the second session of each Congress), 0 otherwise (appropriations decisions concern the upcoming fiscal year. So while congressional elections are held in even-numbered years, the budgets considered by Congress just prior to the election are for odd-numbered fiscal years)

E^{p}_{t} = a dummy variable which takes the value of 1 for budgets considered during presidential election years (the second session of each Congress), 0 otherwise

TRANS = a variable which takes the value of 1 when a Democratic president replaces a Republican, -1 when the transition goes to the Republicans from the Democrats, and 0 otherwise

$REDIST_{t-1}$ = the cumulative measure of urban seat gain (since members elected from a newly redistricted Congress will have their first

impact on the appropriations act for the fiscal year following their taking office--as appropriations acts are passed in the year preceding their expenditure--this distribution is lagged one year. In other words, members elected in November 1964 to the 89th Congress will vote on appropriations for fiscal years 1966 and 1967)

u_{it}, e_{it} = randomly distributed error items

The OMB equation was estimated by two stage least squares, while the congressional equation was estimated by ordinary least squares.

ANALYSIS AND RESULTS

If the data support our hypotheses, the sign of the redistricting terms in the congressional equation for the housing (community development, housing assistance, and CETA), school lunch, and regulatory programs (CAB, FPC, FTC, ICC, and SEC) will be positive, whereas it will be negative for the farm programs (commodity programs, conservation reserve, and Sugar Act payments). The sign of the redistricting terms in the presidential equation may mirror the expectations in the congressional equation if the president is pursuing an accommodative strategy or may be opposite if the president is pursuing a negative strategy vis-a-vis the urbanization of Congress. In both the OMB and congressional equations, the sign of the unemployment term should be positive and the sign of the inflation term should be negative. The signs of the election-year dummies should be positive, as should be the sign of the strategic variables, the party terms, and the administrative transition dummy. We expect the sign of the war dummies to be negative.

The results of this analysis, reported in Table 5, provide strong support for the redistricting hypothesis. The coefficient of the redistricting term in the congressional equation for the housing, school lunch, and regulatory programs is in the predicted direction (positive) and significant. By our measure, court-ordered redistricting did lead to an increase in spending for programs which benefited urban constituents. Though negative in direction, the coefficient of the redistricting term for the farm programs is not significant. It is interesting to note, however, that the coefficient of the redistricting term for the presidential equation is negative and significant. The net effect of redistricting, then, on these farm programs was negative, though the locomotion was transmitted through a different engine, that of presidential anticipation. In interpreting these results, it should be noted that the budgetary consequences of redistricting vary widely across programs: the effect of the cumulative increase in urban seats on the 1979 budget for our sample of housing programs was over $1 billion, whereas the initial effect of a gain of eleven urban seats on the budgets for our sample of regulatory programs was a mere $2 million in 1966.

The results in Table 5 are all the more impressive given the relatively short duration of the experiment. The first court-ordered redistricting affected congressional elections in 1964. The last redistricting with any effect on the urban-rural composition of congressional seats affected congressional elections in 1974.

Changes in the composition of congressional districts led to differing strategies for the president. For school-lunch programs the president submitted slightly higher budget requests the greater the urbanization brought about by redistricting suggesting on this issue that he was trying to

accommodate a forecasted increase in school-lunch appropriations by Congress. On the other hand, budget requests for farm and regulatory programs saw a decline as Congress was derusticated. Farm and regulatory programs are peculiar in that they present national policy issues which also provide localized distributive benefits. We would expect both the president and Congress to seek to influence bureaucratic policy making on these issues. That both the congressional and presidential equations for our sample of regulatory programs are highly predictive and have many significant variables (including the redistricting terms) reflects the tension existing between Congress and the president over setting regulatory policy.

It is possible, however, that the reallocation of regulatory benefits resulting from a change in regulatory policy is not directly responsible for the increase in the budget of our sample of regulatory agencies. It is possible that the increase in the budgets of these agencies came about through a different sort of reallocation. In response to increased pressure from metropolitan consumer groups during the sixties and seventies, Congress required these regulatory agencies to develop new procedures and offices to serve consumer interests. This expansion of the regulatory bureaucracy, designed to serve the new metropolitan interests, would naturally have a trace in the funding levels for regulatory agencies.

Though the results of the analysis are not as unambiguously strong on the remaining coefficients we do find some support for the other electoral hypotheses. Kiewiet and McCubbins (1984a, 1984b), for a different set of domestic federal agencies, found broad support for a congressional "electoral-appropriations" cycle: Congress appropriates about 2.7 percent more to their set of agencies in congressional election years than in

nonelection years. Though housing and farm programs showed little effect of an electoral-appropriations cycle, the election year term in the congressional equations for school lunch and regulatory programs was quite strong.

The coefficients of the strategic variables (OMB_t and $CONG^*_t$) are positive, and with the exception of the congressional term in the OMB equation for the school lunch programs, were significant. It is interesting to note that the size of the strategic terms hover about one, suggesting a great deal of interaction and cooperation between the two branches of government in making public policy. Also, again except for school lunch programs, the size of the forecasted congressional term in the OMB equation is greater than the size of the OMB term in the congressional equation. This implies a greater degree of anticipation and accommodation in the executive branch than in Congress.

The estimated coefficients of the remaining variables show little clear pattern and do not replicate the results in Kiewiet and McCubbins (1984a, 1984b). Though there is some evidence that Democratic presidents prefer to spend more on housing and regulatory programs and less on farm programs than do Republicans, there is little evidence that economic conditions influenced congressional or presidential appropriative decisions in any clear manner for this set of programs. It is also interesting to note that the war dummy was insignificant in almost all of the equations, indicating that no guns vs. butter tradeoff was made for these programs.

Analysis of the residuals indicates that the results of this analysis were not compromised by potential econometric difficulties. First, the correlation between the residuals from one agency and those from another, in each pooled category, tended to be positive, but barely (most of the Pearson r

TABLE 5

The Effects of Redistricting on Federal Appropriations Programs

OMB Equation:

Variables	Housing Coefficient (stand error)[a]	School Lunch Coefficient (stand error)	Farm Coefficient (stand error)	Regulatory Coefficient (stand error)
Constant	-.4408 (.5796)	1.294** (.5041)	.3376 (.2115)	.0014 (.0024)
$\Delta CONG_t^*$	.9551** (.0985)	.1691 (.2157)	.9570** (.0343)	1.014** (.0080)
I_{t-1}^p	-.0083 (.0322)	-.0091 (.0216)	-.0012 (.0104)	.0004** (.0001)
U_{t-1}^p	-.0414 (.1051)	-.0483 (.0702)	-.0037 (.0354)	-.0001 (.0004)
$PRES_t$	.6061** (.2334)	.0963 (.1455)	-.1507* (.0798)	.0030** (.0009)
WAR_t	-.2918 (.3041)	-.4060** (.1782)	-.0012 (.1051)	-.0001 (.0012)
E_t^p	.4164* (.2354)	.2065 (.1754)	-.0911 (.0826)	-.0016* (.0009)
$REDIST_{t-1}$	.0046 (.0095)	.0005** (.0002)	-.0059** (.0019)	-.0001** (.0000)
adj R^2	.905	.703	.951	.992
n	90	32	96	160
SSR	74.90	3.190	10.49	.0037

a. ** = $p < .01$ * = $p < .05$

TABLE 5

The Effects of Redistricting on Federal Appropriations Programs

Congressional Equation:

Variables	Housing Coefficient (stand error)[a]	School Lunch Coefficient (stand error)	Farm Coefficient (stand error)	Regulatory Coefficient (stand error)
Constant	-.2758 (1.016)	.4759 (.9730)	.2907 (.3862)	-.0112 (.0044)
ΔOMB_t	.7178** (.0482)	.4897** (.2053)	.9181** (.0280)	.9754** (.0075)
I^c_{t-1}	-.0238 (.0460)	-.0978** (.0396)	-.0066 (.0181)	-.0005** (.0002)
U^c_{t-1}	-.0642 (.1180)	-.0723 (.0947)	-.0291 (.0466)	-.0001 (.0005)
$HOUSE_t$	.1425 (1.711)	1.139 (1.327)	-.1422 (.6486)	.0170* (.0072)
E^c_t	.1065 (.1798)	.2793* (.1399)	.0149 (.0703)	.0028** (.0008)
$TRANS_t$	.2103 (.2887)	.0292 (.2235)	.0135 (.1142)	.0019 (.0013)
$REDIST_{t-1}$	.0254** (.0073)	.0006** (.0002)	.0031 (.0025)	-.0001** (.0000)
WAR_t	.1460 (.3154)	-.1605 (.2075)	.0477 (.1239)	-.0004 (.0014)
adj R^2	.924	.770	.946	.992
n	92	32	96	160
SSR	57.52	3.26	9.67	.0035

a. ** = $p < .01$ * = $p < .05$

statistics lay between .00 and .50, and the number of associations which were statistically significant was only slightly larger than that which would be expected by chance). Second, a standard F test on the restrictions imposed by pooling within categories revealed no significant evidence of an increase in variance due to the restrictions on the coefficients.

Since Durbin-Watson statistics are not applicable to pooled cross-sectional time series, another method was employed to check for serial correlation. Pooling alone should introduce no serial correlation. It, therefore seems reasonable to assume that the pooled estimation should be free of serial correlation if the nonpooled estimation is. This was, indeed, the case: the Durbin-Watson statistics calculated for the twelve agencies independently varied from 1.6 to 2.8.

5. The Wider Pattern of Derustication

A wide range of policies adopted since the middle 1960's has diminished the net benefits enjoyed by rural Americans while increasing those enjoyed by urban and suburban citizens. Here we discuss five categories of policy in which a rural-to-metropolitan reallocation has occurred:

AGRICULTURE Before redistricting, the benefits of agricultural legislation were directed almost wholly at farmers. The chief example is price-income supports for wheat, feed grains, cotton, wool, sugar, rice, peanuts, tobacco, and dairy products, beginning with the Agriculture Adjustment Act of 1933. Prices were kept high and production restricted even in times of urgent need for food and fiber, such as WWII and the postwar period of European famine.

Things have changed since redistricting. In 1970 Congress imposed a ceiling of $55 thousand per farmer per year on support payments. In the same year, storage and acquisition costs were reduced by means of supply controls. In 1973 the subsidy ceiling was further reduced to $20 thousand, and supports for the cost of production replaced parity payments (a percentage of fair market price). In 1983 these supports took the form of Payments in Kind rather than cash subsidies. The net effect, according to a 1978 study by the Congressional Budget Office, is that price-income supports have declined (Congress of the United States 1978: 43).

Since at least 1970, the benefits of support programs have shifted in the direction of the largest, most efficient farmers at the expense of numerous small farmers, thereby benefiting consumers at the expense of rural voters (Penn and Boehm 1978: 6; Lee 1983: 10).

Loan programs targeted at rural constituents, such as those administered by the Farmers Home Administration (FrHA) and the Rural Electrification Administration, have been steadily reduced or eliminated since the mid 1960's. Budgeted loan authority for the FrHA has fallen from $360 million in 1964 to $24 million in 1979.

No longer are rural residents the sole or even the chief beneficiaries of agricultural legislation, which now is concerned with cost to consumers, soil conservation, land use, environmental damage, income relief, dietary goals, and foreign aid. At the same time, agriculture bills are logrolled with legislation that serves metropolitan interests. The Agriculture and Consumer Protection Act of 1973 provided food stamps for strikers and was logrolled with an increase in the minimum wage. The 1977 Food and Agriculture

Act eliminated the purchase requirement for food stamps, and included separate titles devoted to Food for Peace and nutrition education.

How the new metropolitan majority affects agriculture policy can be seen with a closer look at the passage of the 1977 Act, which required a new coalition of farm, consumer, and labor representatives. (V. Peters 1978). Chairman Thomas Foley (D-Wash.) characteristically allowed the proposal of floor amendments to his House Agriculture Committee's carefully crafted omnibus bill. In response to depressed wheat prices, declining credit, and a bumper wheat crop, Representative Glenn English (D-Okla.) offered an amendment to increase wheat supports. A coalition of relatively junior reform Democrats supported the amendment out of solidarity with their fellows on Agriculture; reciprocity was, of course, expected. The Consumer Federation of America also supported the amendment, but demanded, in return, support for the Consumer Protection Agency. The amendment passed--but only through a trade with nonagricultural interests.

As chairman of the Subcommittee on Domestic Marketing, Consumer Relations, and Nutrition, Brooklyn Democrat Frederick W. Richmond, the millionaire representative of one of America's worst slums, was allowed by Foley to steer a separate food-stamp title, which completely eliminated the purchase requirement, through the entire legislative process. Previously, recipients had been required to buy food stamps with the price determined by need. This title received rural support on the full committee, and the bill as a whole received urban support on the floor. "We did build up a very good working urban-rural coalition," said Richmond. "It was my job to convince urban members of Congress to support the family farmer and convince rural

members to support cities." (Cindy Montgomery, "Richmond Backs Rural Interests," Lincoln Journal, August 24, 1977: 28.)

Representative Steven Symms (R-Idaho) offered an amendment, generally favored by farm interests, to require some payment for food stamps. The Richmond coalition held: only twenty-eight rural members voted for the Symms amendment, which failed. Also generally favored by farm interests, an amendment by Representative Richard Kelley (R-Fla.) to eliminate food stamps for strikers failed as well: nearly half the representatives of rural and mixed districts voted against the amendment. (V. Peters 1978: 28, Table 1.)

Provisions of the food stamp title were also used strategically by Dawson Mathis (D-Ga.), chairman of the Subcommittee on Oilseeds and Rice and the representative of a peanut-producing district. Mathis had planned to offer an amendment empowering the Secretary of Agriculture to implement any number of pilot projects in which food-stamp recipients would be put to work. He dropped this amendment when Richmond successfully killed another amendment, by Representative Margaret Heckler (R-Mass.), that would have reduced the peanut subsidy.

In the end the omnibus agriculture bill was supported by 70 percent of congressmen from urban districts and 60 percent of those from suburban districts (Peters 1978: 28, Table 1).

ANTI-TRUST Much of the force of anti-trust legislation had originally been directed at monopsonies. Faced with low prices, high railroad tariffs, and high costs of the things they had to buy, farmers in the late nineteenth century sought protection against railroads and other middlemen through

anti-trust laws. The Sherman and Clayton Acts, the Packers and Stockyards Act, the Commodity Exchange Act, and the U.S. Warehouse Act were to a great extent designed to reduce the market power of middlemen in the food and fiber industry (Buse and Bromley 1975: Chap. 11). By contrast, more recent anti-trust legislation, such as the FTC Improvement Acts of 1975 and 1980, has been directed at monopolies and general business practices and has deemphasized monopsony regulation.

REGULATION The last decade and a half have seen dramatic changes in regulatory policy. New environmental, health, safety, and consumer regulations have been enacted, and moribund agencies have been reinvigorated. The emergence of consumer and environmental groups in the late 1960's obstructed regulatory decision making in Congress as well as the bureaucracy. Joskow (1974) has argued that regulatory decision making before the late 1960's was infrequent and noncontroversial. In an era of economic prosperity and declining marginal costs for utilities, rate hearings were quiet and uneventful. Then two shocks made the system break down: Rapidly increasing factor prices led utilities to seek repeated rate hikes, which consumer groups opposed, and the emergence of environmental concerns forced regulators to deal with new issues. The existing regulatory apparatus was not designed to handle the new demands made upon it. Weingast (1978, 1980), too, explains regulatory change in the 1970's as the result of new interest groups pressing their demands.

Why did consumer and environmental groups both emerge as organized interests in the late 1960's? Joskow contends that increased energy costs turned regulatory decisions into a zero-sum game between utilities and consumers. Although true, this does not explain the emergence of health,

environmental, and labor interests. Nor does it explain how these newly organized interests quickly gained standing before Congress and the bureaucracy over the opposition of entrenched industry interests.

Our explanation begins with the observation that the newly emergent interests were predominantly nonrural. In transferring political power from rural to metropolitan areas, court-ordered redistricting was the enzyme for the organization of these interests. In Congress, the new metropolitan bloc gained control by the late 1960's. Beyond the outpouring of consumer and environmental regulation, existing regulatory procedures were amended to give these new nonrural interests standing before the bureaucracy. For example, the National Environmental Policy Act of 1969 requires federal agencies to solicit the objections of environmental groups before any action can be taken. The congressional oversight network was extended to accommodate these and other nonrural interests (McCubbins and Schwartz 1984), and each new piece of legislation provided elaborate procedures to enfranchise the new interests in regulatory decision making (McCubbins 1982 and 1983).

TRANSPORTATION Before Baker v. Carr, transportation regulation was designed to provide cross-subsidies, ensuring that transportation was available to residents of sparsely populated areas at near the same rates as were paid by residents of densely populated areas. For ground transportation, the Interstate Commerce Act of 1887, strengthened by the Elkins Act of 1903, the Hepburn Act of 1906, the Mann-Elkins Act of 1910, the Shipping Act of 1916, the Transportation Acts of 1920, 1940, 1942, and 1958, the Motor Carrier Act of 1935, and the Reed Bulwinkle Act of 1948, gave the Interstate Commerce Commission control over rates and routes and encouraged the Commission to hold down rates and establish routes for unprofitable rural destinations. More

recently, the Railroad Revitalization and Regulatory Reform Act of 1976, the Staggers Act (1980), and the Motor Carrier Act of 1980 have "deregulated" ground transportation by encouraging competition. Meanwhile, owing to appointments made in the 1970's, the ICC was in some cases ahead of Congress in its loosening of price, entry, and route requirements. (Alexis 1982 gives a thorough survey.)

The Civil Aeronautics Act (1936) also promoted metropolitan-rural cross-subsidies. Deregulation of the airline industry began even before the Airline Deregulation Act of 1978. In 1977 the Civil Aeronautics Board approved Super Saver fares, and in 1978 they allowed carriers to set fares 10 percent above or 50 percent below CAB standard fares. The 1978 Act allowed the CAB also to liberalize route awards. Between 1978 and 1981, weekly departures between small airports declined significantly (Graham et al 1983: 120). Before deregulation, airline fares were based on distance; they were below cost for short hauls and above cost for long hauls. By 1980, fares varied inversely with both distance and density (Graham et al 1983: 122).

CIVIL RIGHTS Beginning with the 1964 Civil Rights Act and the 1965 Voting Rights Act, civil-rights legislation appealed to urban more than rural voters, especially outside the South. The beneficiaries included rural blacks. But most of them were not voters until after the legislation was passed and enforced, whereas urban blacks and liberals strongly favored such legislation.

INTERNAL ORGANIZATION OF CONGRESS In the 1950's Southerners constituted a majority of the House Democratic Party and acquired the seniority on standing committees by which they dominated decisions long after

non-Southerners had become numerically the dominant wing of the Party. Hard on the heels of redistricting came a secular shift from a Southern rural to a Northern liberal-metropolitan majority, and therewith the conditions for institutional change within Congress.

The institutional changes took on two forms. First, the Democratic Caucus in the 1970's became an operational instrument of the House Party's majority for selecting chairmen of full committees and Appropriations subcommittees. Personnel decisions within the Caucus demonstrated the will of the new metropolitan majority, unseating three senior Southern committee chairmen. Representative W. R. Poage of Texas, first elected to the House in 1937, lost his chairmanship of the House Agriculture Committee to Thomas S. Foley (Wash.), first elected in 1964.

In 1977 the House Democrats ousted an Appropriations subcommittee chairman and voted in substantial numbers against Jamie L. Whitten (Miss.), the chairman of the Subcommittee on Appropriations for Agriculture and Related Agencies. Whitten ultimately kept his job by relinquishing jurisdiction over consumer and environmental matters. As a result of these personnel changes Southerners in 1977, though constituting 30% of the House Democratic Party, held but five of twenty-two committee chairmanships. By 1977 almost half of all Democrats, including many junior members, headed subcommitees, and often it was they rather than the chairmen of their full committees who shepherded bills through the entire legislative process.

Second, the new metropolitan majority brought about rule changes. In 1973 the Caucus created a new twenty-four-member Steering and Policy Committee, chaired by the Speaker, to develop policy and to oversee the Party's legislative strategy. In 1975 this committee was assigned the

committee-on-committees functions, formerly exercised by Democrats on the Ways and Means Committee. The Speaker's powers to move legislation were expanded by the Caucus' decision that he should appoint all Democratic members of the Rules Committee. Together these institutional changes enabled the new metropolitan majority to exercise its will on national policy issues.

6. Conclusion

Court-ordered redistricting turned a predominantly rural Congress into an overwhelmingly nonrural one. This institutional change led to a series of policy changes that reallocated the net benefits of government activity from rural to nonrural citizens. We explained the latter change in terms of the former and tested our hypothesis through a two-stage least-squares regression of changes in the relative fortunes of federal agencies against a measure of redistricting.

To be sure, our test applied only to a small set of policy changes, not all the changes discussed in §5. There is plenty of room for further research here. Having shown, however, that our hypothesized engine of change is at work (and having discussed how it works), we would be surprised if it were not at least partly responsible for other instances of policy derustication.

In their enforcement of the one-man, one-vote rule, the federal courts have conducted a massive experimental test of our system of representative democracy: Does policy monotonically reflect congressional representation? Would a large-scale change in the congressional representation of certain interests produce a like change in policy? The result of the experiment was positive: representative democracy seems to work as it should.

Notes

1. Here we ignore three obvious constitutional constraints which in practice have proved not to be very constraining: the threat of impeachment (not used as a policy tool), the words of the Constitution (long subject to the loosest of constructions), and the case-and-controversy requirement of Article III (greatly weakened by broad powers of certiorari, selectively used).

2. Echoing some Tory criticisms of the 1832 Reform Act, Aranson (1982) has offered an ingenious theoretical argument that redistricting per se can have no policy effects, but he makes extremely strong assumptions about capital mobility, fungibility, and the incentives of legislators.

3. Different measures of redistricting can be constructed from Table 4 by using different categories or by combining categories. The results of the analysis we report in Table 5 are robust to simple changes in the redistricting measure. We chose to report the results with redistricting changes measured as changes in urban seats because it was the simplest and most straightforward measure.

Data Sources

Presidential budget requests (in the form of OMB estimates) and final appropriations figures are reported in the Annual Senate Document Appropriations, Budget Estimates, Etc., the section entitled "Itemized Comparisons of Budget Estimates and Appropriations Arranged by Senate Acts."

The deflator used to convert the appropriations and estimate figures into constant dollars was the Implicit Price Deflator for Federal Government Purchases of Goods and Services. The time series for this deflator was taken chiefly from The National Income and Product Accounts of the United States, 1929-74, Statistical Tables. Data after 1974 are taken from monthly issues of the Survey of Current Business. Both are published by the Bureau of Economic Analysis, U.S. Department of Commerce.

The unemployment and consumer-price-index figures were taken from issues of the Monthly Labor Review, Bureau of Labor Statistics, U.S. Department of Labor.

Data on changes in the urban-rural composition of the House were collected from the Congressional District Data Book and its supplements, published by the Bureau of the Census, U.S. Department of Commerce. Other demographic, boundary, and redistricting information was collected from the CQ Census Analysis: Congressional Districts of the United States, CQ Press, 1964; CQ Weekly Report No. 34, August 21, 1964; Congressional District Atlas (Districts of the 92nd Congress), published by the Bureau of the Census; and Congressional Districts in the 1970's and its revisions by CQ Press.

Court Cases

Baker v. Carr, 369 U.S. 186 (1962)

Colgrove v. Green, 328 U.S. 549 (1946)

Gomillion v. Lightfoot, 364 U.S. 347 (1960)

Kirkpatrick v. Preisler (1969)

Reynolds v. Sims, 377 U.S. 533 (1964)

Wesberry v. Sanders, 376 U.S. 1 (1964)

References

Alexis, Marcus. 1982. "The Applied Theory of Regulation: Political Economy at the Interstate Commerce Commission." Public Choice 39: 5-28.

Aranson, Peter. 1982. "Political Inequality: An Economic Approach." In Peter C. Ordeshook and Kenneth A. Shepsle, eds. Political Equilibrium. Boston: Kluwer-Nijhoff: 133-150.

Baker, Gordon E. 1966. The Reapportionment Revolution. New York: Random House.

Bicker, William. 1971. "The Effects of Malapportionment in the States--A Mistrial." In Nelson Polsby, ed. Reapportionment in the 1970's. Berkeley: University of California Press.

Brady, David, and Douglas Edmonds. 1967. "One Man, One Vote--So What?" Transaction 4: 941-46.

Brewster, David E., Wayne D. Rasmussen, and Garth Youngberg, eds. 1983. Farms in Transition. Ames, Iowa: Iowa State University Press.

Buse, Rueben, and Daniel Bromley. 1975. Applied Economics: Resource Allocation in Rural America. Ames, Iowa: Iowa State University Press.

Cain, Bruce. 1984. The Reapportionment Puzzle. Berkeley: University of California Press.

Congress of the United States Congressional Budget Office. 1978. Public Policy and the Changing Structure of American Agriculture. Washington: U.S. Government Printing Office.

David, Paul, and Ralph Eisenberg. 1961 and 62. Devaluation of the Urban and Suburban Vote. Two volumes. Charlottesville, VA: University of Virginia Bureau of Public Administration.

Davis, Otto A., M.A.H. Dempster, and Aaron Wildavsky. 1966. "A Theory of the Budgetary Process." American Political Science Review 60: 529-47.

_____. 1974. "Toward a Predictive Theory of Government Expenditures: U.S. Domestic Appropriations." British Journal of Political Science 4: 419-52.

Dodd, Lawrence, and Bruce Oppenheimer. 1977. "The House in Transition." In Lawrence Dodd and Bruce Oppenheimer, eds. Congress Reconsidered. New York: Praeger: 21-53.

Dye, Thomas. 1965. "Malapportionment and Public Policy in the States." Journal of Politics 27: 586-601.

Erikson, Robert S. 1973. "Reapportionment and Policy: A Further Look at Some Intervening Variables." Annals of the New York Academy of Sciences 219: 280-290.

Fenno, Richard F. 1966. The Power of the Purse: Appropriations Politics in Congress. Boston: Little, Brown.

_____. 1973. Congressmen in Committees. Boston: Little, Brown.

_____. 1975. "If, as Ralph Nader Says, Congress is 'the Broken Branch,' Why Do We Love Our Congressmen So Much?" In Norman Ornstein, ed. Congress in Change. New York: Praeger.

Ferejohn, John. 1974. Pork Barrel Politics. Stanford: Stanford University Press.

Fiorina, Morris. 1974. Representatives, Roll Calls, and Constituencies. Lexington, Mass.: Lexington Books.

_____. 1977. Congress: Keystone of the Washington Establishment. New Haven: Yale University.

_____. 1978. "Legislative Facilitation of Government Growth: Universalism and Reciprocity Practices in Majority Rule Institutions." Social Science Working Paper No. 226. California Institute of Technology.

Frederickson, H. George, and Yong Hyo Cho. 1974. "Legislative Apportionment and Fiscal Policy in the American States." Western Political Quarterly 27: 5-37.

Fry, Brian R., and Richard F. Winters. 1970. "The Politics of Redistribution." American Political Science Review 64: 508-522.

Graham, Daniel R., Daniel P. Kaplan, and David S. Sibley. 1983. "Efficiency and Competition in the Airline Industry." Bell Journal of Economics 14: 118-138.

Hanson, Roger A., and Robert E. Crew, Jr. 1973. "The Policy Impact of Reapportionment." Law and Society Review 8: 69-92.

Hardin, Charles M. 1978. "Agricultural Price Policy: The Political Role of Bureaucracy." In Hardwiger and Browne 1978: 7-14.

Hardwiger, Don F., and William P. Browne, eds. 1978. The New Politics of Food. Lexington, Mass.: D.C. Heath.

Hausman, J. A. 1978. "Specification Tests in Econometrics." Econometrica 46: 1251-1270.

Hofferbert, Richard. 1966. "The Relationship Between Public Policy and Some Structural and Environmental Variables in the American States." American Political Science Review 60: 73-82.

Jacob, Herbert. 1964. "The Consequences of Malapportionment: A Note of Caution." Social Forces 43: 256-71.

Jacobson, Gary, and Samuell Kernell. 1981. Strategy and Choice in Congressional Elections. New Haven: Yale University Press.

Joskow, Paul. 1974. "Inflation and Environmental Concern: Structural Change in the Process of Public Utility Price Regulation." Journal of Law and Economics 17: 291-327.

Kiewiet, D. R., and Mathew D. McCubbins. 1984a. "Congressional Appropriations and the Electoral Connection." Unpublished manuscript, California Institute of Technology.

_____. 1984b. "Congress, the President, and the Appropriations Process." Forthcoming, Legislative Studies Quarterly.

Kramer, Gerald. 1971. "Short-Term Fluctuations in U.S. Voting Behavior, 1896-1964." American Political Science Review 65: 131-43.

Lee, John E., Jr. 1983. "Some Consequences of the New Reality in U.S. Agriculture." In Brester et al 1983: 3-22.

Lowi, Theodore. 1984. "American Business, Public Policy, Case Studies, and Political Theory." World Politics 16: 677-715.

McCubbins, Mathew D. 1982. "Rational Individual Behavior and Collective Irrationality: The Legislative Choice of Regulatory Form." Ph.D. dissertation. California Institute of Technology.

_____. 1983. "A Formal Model of the Form of Regulatory Intervention." Working Paper on Institutional Design and Public Policy, No. 1, Department of Government, University of Texas at Austin.

McCubbins, Mathew, and Thomas Schwartz. 1984. "Congressional Oversight Overlooked: Police Patrols vs. Fire Alarms." American Journal of Political Science.

Meier, Kenneth J., and William P. Browne. 1983. "Interest Groups and Farm Structure." In Brewster et al 1983: 47-56.

Mowery, David, and Mark Kamlet. 1982. "Coming Apart: Fiscal and Budgetary Policy Processes in the Johnson Administration." Public Budgeting and Finance 11: 16-34.

Newcomer, Kathryn E., and Richard J. Hardy. 1980. "Analyzing Political Impact: Selection of a Linear Trend Model." Policy Studies Journal 8: 928-941.

Okun, Arthur. 1970. The Political Economy of Prosperity. New York: W. W. Norton.

Ornstein, Norman, Robert Peabody, and David Rhode. 1977. "The Changing U.S. Senate: From the 1950's to the 1970's." In Lawrence Dodd and Bruce Oppenheimer, eds. Congress Reconsidered. New York: Praeger: 3-20.

O'Rourke, Timothy. 1980. The Impact of Reapportionment. New Brunswick: Transaction Books.

Padgett, J. 1981. "Hierarchy and Ecological Control in Federal Budgetary Research." American Journal of Sociology 87: 75-129.

Peltzman, Samuel. 1976. "Toward a More General Theory of Regulation." Journal of Law and Economics 18: 211-40.

Penn, J. B., and William T. Boehm. 1978. "Research Issues Re-emphasized by 1977 Food Policy Legislation." Agricultural Economics Research 30: 6.

Peters, John G. 1978. "The 1977 Farm Bill: Coalitions in Congress." In Hardwiger and Browne 1978: 23-26.

Poole, Keith, and Howard Rosenthal. 1983. "The Polarization of American Politics." Unpublished manuscript. Carnegie-Mellon University.

Porter, Laurellen. 1978. "Congress and Agricultural Policy, 1977." In Hardwiger and Brown 1978: 15-22.

Pressman, Jeffrey, and Aaron Wildavsky. 1973. Implementation. Berkeley: University of California Press.

Pulsipher, Allan, and James Weatherby. 1968. "Malapportionment, Party Competition, and the Functional Distribution of Governmental Expenditures." American Political Science Review 62: 1207-19.

Rae, Douglas. 1971. "Reapportionment and Political Democracy." In Nelson Polsby, ed. Reapportionment in the 1970's. Berkeley: University of California Press: 91-112.

Schwartz, Thomas. 1983. "The Pork-Barrel Paradox." Working Paper on Institutional Design and Public Policy, No. 14, Department of Government, University of Texas at Austin.

Shepsle, Kenneth, and Barry Weingast. 1980. "Political Preferences for the Pork Barrel: A Generalization." Working Paper No. 57. Washington University, St. Louis. Center for the Study of American Business.

Stigler, George. 1971. "The Theory of Economic Regulation." Bell Journal of Economics and Management Sciences: 3-21.

Stockman, David. 1975. "The Social Pork Barrel." The Public Interest 39: 3-30.

Sundquist, James. 1981. The Decline and Resurgence of Congress. Princeton: Princeton University Press.

Tufte, Edward. 1978. Political Control of the Economy. Princeton: Princeton University Press.

Weingast, Barry. 1978. "A Positive Model of Public Policy Formation: The Case of Regulatory Agency Behavior." Working Paper No. 25. Washington University, St. Louis. Center for the Study of American Business.

_____. 1980. "Congress, Regulation, and the Decline of Nuclear Power." Public Policy 28: 231-55.

_____. 1983. "A Principal-Agent Perspective on Congressional-Bureaucratic Relations." Paper delivered at the Fifth Carnegie Conference on Political Economy, Carnegie-Mellon University, Pittsburgh, PA.

Wildavsky, Aaron. 1974. The Politics of the Budgetary Process. Boston: Little, Brown.

CENTER FOR ECONOMIC POLICY RESEARCH
100 Encina Commons
Stanford University
Stanford, California 94305

CEPR Publication Series

1. Paul A. David, "Microelectronics and the Macroeconomic Outlook," Two Papers for the OECD Working Party on Information, Computer and Communications Policy, April 1982.

2. Timothy F. Bresnahan, "The Impact of Proposed Emissions Rollbacks on the Automobile Industry," April 1982.

3. Ronald I. McKinnon, "A Program for International Monetary Stability," January 1983.

4. Timothy F. Bresnahan, "Automobile Price Indexes by Continent of Origin," November 1982.

5. Ronald I. McKinnon, "Why U.S. Monetary Policy Should Be Internationalized," April 1983.

6. Robert M. Coen and Bert G. Hickman, "Tax Policy, Federal Deficits and U.S. Growth in the 1980's," April 1983.

7. Ronald I. McKinnon, "Dollar Overvaluation Against the Yen and Mark in 1983: How to Coordinate Central Bank Policies," May 1983.

8. Lawrence J. Lau, "The Measurement of Productivity: A Lecture in Memory of Professor Ta-Chung Liu," July 1983.

9. Bert G. Hickman, "Growth, Inflation and Unemployment in the United States," August 1983.

10. Michael J. Boskin, Kenneth Cone and Sule Ozler, "The Federal Budget and Deposit Insurance," November 1983.

11. Michael J. Boskin and Bradford Barham, "Measurement and Conceptual Issues in Federal Budget Treatment of Loans and Loan Guarantees," November 1983.

12. Claudio R. Frischtak, "Choice of Technology and Economic Growth: A Reassessment," December 1983.

13. M. Elisabeth Paté-Cornell, "Probabilistic Assessment of Warning Systems: Signals and Response," January 1984.

14. Bruce D. Spencer, "Avoiding Bias in Estimates of the Effect of Data Error on Allocations of Public Funds," April 1984.

15. Carson E. Agnew and Richard G. Gould, "Frequency Coordination and Spectrum Economics," Revised March 1984.

CEPR Publication Series (con't)

16. Timothy F. Bresnahan and Peter C. Reiss, "Dealer and Manufacturer Margins," CEPR Technical Report, February 1984.

17. B. Douglas Bernheim, "Dissaving After Retirement: Testing the Pure Life Cycle Hypothesis," CEPR Technical Report, revised March 1984.

18. Bruce D. Spencer, "Sensitivity of Benefit-Cost Analysis of Data Quality," CEPR Technical Report, May 1984.

19. I.P.L. P'ng, "Liability, Litigation, and Incentives to Take Care," CEPR Technical Report, revised May 1984.

20. M. Elisabeth Paté-Cornell and Timothy S. O'Grady, "Reduction of Fire Risks in Oil Refineries: Costs and Benefits of Camera Monitoring," CEPR Technical Report, January 1984.

21. Ben S. Bernanke and James L. Powell, "The Cyclical Behavior of Industrial Labor Markets: A Comparison of the Pre-War and Post-War Eras," revised May 1984.

22. Yoram Weiss, "The Effect of Labor Unions on Investment in Training: A Dynamic Model," CEPR Technical Report, December 1983.

23. Paul A. David, "The Reaper and the Robot (The Diffusion of Micro-electronics-Based Process Innovations in Historical Perspective)," revised May 1984. (Same as TIP Working Paper No. 1)

24. Paul A. David and Trond E. Olsen, "Anticipated Automation: A Rational Expectations Model of Technological Diffusion," revised April 1984. (Same as TIP Working Paper No. 2)

25. Josep M Vegara-Carrio, "Software Technological Change: An Economic Approach," May 1983. (Same as TIP Working Paper No. 3)

26. Michael J. Boskin, Marc S. Robinson, Terrance O'Reilly and Praveen Kumar, "New Estimates of the Value of Federal Mineral Rights and Land," revised August 1984.

27. B. Douglas Bernheim, "Life Cycle Annuity Valuation." August, 1984.

28. Arthur T. Denzau, William H. Riker and Kenneth A. Shepsle, "Strategic Behavior in the Theory of Legislatures," CEPR Technical Report, revised June 1984.

29. John E. Chubb, "The Political Economy of Federalism," June 1984.

30. Mathew D. McCubbins and Thomas Schwartz, "The Politics of Derusti-cation: Court-Ordered Redistricting and Its Policy Effects," July 1984.

The Political Economy of Regulated Price Structures: The Impact of Divestiture on State Telephone Rates

Susan R. Smart *

In the 1980s, state regulatory commissions had to respond to an abrupt change in the structure of the telecommunications industry–the breakup of AT&T and the creation of the Regional Bell Holding Companies. Prior to the introduction of competition in the long distance market, federal and state regulatory policies were compatible. In the tradition of the "universal service" goal, prices for residential and rural customers were kept below the cost of providing service. The universal service doctrine reflected the view that telephone service should be available to everyone at an affordable price. This was accomplished in part by transferring revenues from services which fell under federal control (interstate long distance) to the jurisdiction of state regulators.

The move toward competition left state regulators in the unpleasant situation of regulating generally underpriced local exchange services while receiving shrinking revenues from interstate long distance. Joskow (1974) has pointed out that regulators are particularly concerned with increasing rates. He argues that con-

*Assistant Professor of Business Economics and Public Policy at Indiana University. I would like to thank Roger Noll, Tim Bresnahan and Frank Wolak for their helpful comments. This research was supported by a grant of the John and Mary Markle Foundation to the Center for Economic Policy Research.

sumers are unlikely to voice concern when prices are stable or decreasing, but, when prices increase, consumers may exert political pressure to roll them back. The telephone rate increases necessary due to the diminished opportunities for cross subsidization made such consumer outcry likely.

In this paper, I discuss two decisions which state regulatory commissions have made since divestiture. The first is whether to restrict activities which bypass the services of the local exchange, thereby decreasing its revenues by introducing competition. The second is the post-divestiture structure of local service rates. Bypass, by decreasing the local operating company revenues, decreases the regulators' opportunities for cross subsidization. Despite the expressed concern of regulators that increasing prices may adversely affect certain classes of telephone users,[1] political factors will influence the decisions they make in restricting bypass technologies and establishing rate structures.

In the next section, I outline the background of the policy equilibrium which existed prior to divestiture. These initial conditions are important determinants of regulatory response to divestiture. Section two develops the theory of the political factors affecting regulatory policy. Section three describes the data used, and in section four a model is developed to look at the resulting political decisions and the effects of these decisions on the subsequent pattern of telephone rates within the states. In section five, the estimation and results are presented, and in the last section I offer conclusions.

[1]For example, see Noll (1986).

1 Policy Background

In the early 1980s, state policy makers were faced with potentially very large increases in monthly telephone rates. Noll (1986) and Noll and Smart (1991) describe three effects which put pressure on these rates. First, divestiture reversed the historical trend of an increasing subsidy from interstate long distance service to local access. Second, the divestiture settlement barred the Bell Operating Companies (BOCs) from participating in many competitive markets. Third, in order to provide long distance carriers and information service providers equal access to the local exchange, the BOCs were forced to make large investments to upgrade local plant facilities. The combination of these factors led state regulators to believe that large increases in local rates would be necessary in order to support local telephone companies. Thus, regulators in states with relatively high subsidy levels (low prices) prior to divestiture faced the greatest potential for subsequent consumer outcry. State policy and the structure of rates prior to divestiture, therefore, bears directly on post-divestiture policy.

Before the breakup of AT&T, interstate long distance service supported a substantial fraction of the costs of the local exchange.[2] By 1981, this fraction had reached 26 percent. The transfer of funds from interstate to intrastate service was accomplished in a process known as Separations and Settlements. The costs of the local exchange were separated between interstate and intrastate service. The fraction of local plant costs paid for from interstate revenues differed by state. The formula for calculating the fraction depended on the relative use of local and long distance calling for each state. Beginning in 1952, this formula began weight-

[2]Wynns (1984) provides a thorough discussion of this process.

ing long distance usage more heavily, collecting a subsidy from long distance toll revenues to offset part of the costs of local service. As the costs of long distance service fell throughout the 1960s and 70s due to technological innovation, the percentage of local plant costs allocated to the interstate jurisdiction (the subscriber plant factor or SPF) rose rapidly. Through the Separations and Settlements process, states with particularly high local plant costs received larger subsidies from long distance toll revenues.

Because interstate long distance service falls under federal regulatory jurisdiction, federal cooperation was necessary in order for states to use long distance revenues to offset local exchange costs. Since divestiture, federal regulators have moved toward competitive, cost-based pricing which does not allow for subsidization of the local exchange. Instead, the Federal Communications Commission (FCC) has adopted a plan to collect a portion of the fixed costs of the local exchange by imposing a flat monthly charge to subscribers. To prevent sudden large increases in monthly rates, the FCC is gradually increasing these charges. As of 1990, each residential consumer pays $3.50 per month in "long distance access charges" to cover part of the fixed costs of the local exchange. Long distance carriers continue to pay per minute fees to the local exchange for the use of their connection facilities. As the subscriber flat rate charges increase, these per minute fees fall. Consequently, long distance rates are also falling.

State regulators, concerned about increasing local service rates, had to make decisions concerning competition within their states. The alternative to allowing basic rates to increase in order to cover the fixed costs of access was to find revenues elsewhere to support lower basic rates. Prohibiting competition for certain services

could allow them to collect revenues to offset rate increases for some classes of users.

One source of revenue for the local exchange is the intrastate long distance toll market. At the time of divestiture, states were divided into Local Access and Transport Areas (LATAs) which corresponded roughly with a metropolitan area. The federal divestiture settlement only required competition in the interstate long distance market.[3] Most states elected to make intrastate interLATA service competitive but to restrict or prohibit intraLATA competition. Mathios and Rodgers (1990) have examined the effect of entry and competition in this market and found that indeed intraLATA toll rates are higher in those states which do not permit entry in this market, thereby constituting a possible source of continuing subsidy for basic service.

In this paper, I discuss restrictions on other classes of service which may affect the revenues of the local exchange carrier. These restrictions can prohibit or tax services offered in the private sector which have traditionally been provided by the local exchange carrier. In particular, I consider restrictions on the following:

- Metropolitan Area Networks (MANS) allow interoffice communication in a limited service area. These may combine fiber optic and microwave facilities. Local Area Networks (LANS) allow intraoffice communication among computers over nonLEC cables.
- Interexchange carriers can carry the long distance traffic of large users without paying access charges by bypassing the local exchange and connecting directly to the customers.
- Customer-owned bypass systems could also provide a direct connection to an interexchange carrier.

[3]InterLATA long distance within a state could be a state regulated monopoly, but the BOCs are prohibited from providing this service. The BOCs may provide intraLATA service.

- Digital Termination Systems (DTS) allow digital data communication in a limited service region using microwave transmission.[4]

Because the nature of these activities is to construct facilities which duplicate the function of the local exchange plant, these activities fall under the general heading of bypass technologies.[5]

2 Regulatory Decision on Bypass and Rates

The decision facing state regulators since the breakup of AT&T has been how the inevitable rate increases will be passed along to different classes of users. Those expected to be hardest hit are the residential and the smallest business users.[6] These are the users who have no alternative to local exchange service. It is important to keep in mind that rates for these services have not traditionally been set in relation to cost. The expected increases in rates in the wake of divestiture resulted from an anticipated move toward cost-based pricing, which would reduce opportunities for cross subsidization. So, while there is pressure to move toward costs, regulators may look for ways to maintain some level of subsidy for certain services, lessening the impact of divestiture on rates for basic service. In his study of public utility price regulation, Joskow observed that state regulators may pay little attention to rates as long as they tend to be falling. When, however, prices are increasing and therefore likely to become a salient political issue, regulators

[4]Regulation of DTS illustrates the conflict between state and federal regulators. The FCC has now preempted state restrictions of DTS.

[5]This type of bypass is often called "uneconomic" if the service could be supplied at a lower *cost* by the local company. Because the *price* charged is not based on cost, some low cost consumers may find it profitable to bypass. Regardless, both economic and uneconomic bypass reduce operating company revenue.

[6]For the purposes of this paper, a small business is one for which single- line telephone service is adequate.

react to keep rates low. If regulators who restrict bypass opportunities use the associated revenues to keep basic rates lower, they may reduce the probability that rates become a political issue.

Cross-subsidization for certain classes of users implies that elsewhere in the rate structure prices are being held above cost. Consumers of telecommunication services in these markets would pay for any continued subsidization of basic rates. Regulators considering the bypass issue may be reluctant to raise rates for those large users for whom bypass is an alternative to traditional LEC services, despite their interest in keeping prices for basic service down. In deciding whether to do so, state regulators must have some mechanism for balancing these conflicting interests.

In analyzing regulators' decisions, it is important to recognize that they gain authority through a political process. Politicians are tied to the interests of their constituencies through the reelection process. A common assumption in positive political theory is that the objectives of elected officials can be summarized in the goal of reelection.[7] This single objective encompasses all of the official's reasons for holding office, for without the office, none of a politician's other goals could be achieved. This argument can also be extended to regulators who are appointed rather than elected[8] on the grounds that elected political officials act as policy-making principals in an agency relationship with regulators.

The theory of the political control of agencies states that agencies will represent the preferences embodied in the coalition agreement giving rise to legislation. In a legislature, committee oversight and the budget process provide incentives to

[7]See Downs (1957).

[8]Eleven states in the sample have Commissioners elected in general elections.

agencies not to ignore the wishes of legislators, whereas the role of the executive in appointing agency leaders empowers it in maintaining influence over policy.[9] As applied to state regulatory commissions, this theory implies that a governor's appointees to a commission are likely to serve the political interests of the governor in order to retain favor and influence.

Commissioners who are directly elected differ from those appointed by the governor in two respects. First, direct election obviously reduces the influence of the governor, but not the state legislature. The latter still controls budgets and legislation concerning regulation. Second, commissioners are elected in a single-issue election, the single issue being utility regulation. An elected commissioner must acquire the support of a majority of voters on this single issue. Consequently, a candidate for an elected commission may be more sensitive to interest groups with a stake in utility regulation than would be the governor, whose election campaign covers a wider range of issues. Because it is straightforward to verify that promises made by a commissioner during a campaign are carried through, these interest groups may provide future campaign support on the basis of the extent to which such promises were, in fact, implemented.

Through the political process, large, organized interest groups have more influence on policy than individuals. Organized groups are more effective in influencing election outcomes and, therefore, can also influence state regulators concerned with election outcomes. In order for a group to organize for the purpose of affecting policy, the expected gains from organizing must offset the costs. Therefore, the interests which have the greatest per capita stake in telecommunications regulation

[9]See, for example, McCubbins, Noll and Weingast (1989).

are most likely to organize to affect policy.

The discussion above suggests that regulators who are making a decision whether to restrict bypass service are influenced by those groups whose members stand to lose or gain the most from such a restriction. Bypass is an option for only the largest users of telecommunications services. Large businesses, particularly those that are heavy users of telecommunications services, therefore, are likely to influence regulatory policy regarding restriction of bypass. For this reason, in states with a larger proportion of large businesses, regulators should be less likely to restrict bypass technologies. The other influential organized group in this debate is the telephone company itself. Local telephone companies are on the other side, urging regulators to protect them from competition. Providers of competitive services form another organized group; however, they generally are from outside the state and so have less political influence.

Small business consumers and residential consumers are also affected, though indirectly, by restrictions on bypass, since bypass restrictions may be used to maintain some level of subsidy to these users. Unlike large, organized users, these consumers have high organization costs relative to the potential benefit to the individual from organizing. An individual consumer is unlikely to incur the cost of becoming informed on the consequences for residential rates of a restriction on bypass technology.

The prevention of bypass through the use of state restrictions may not serve well as a long term policy for maintaining LEC revenues. Advances in technology require repeated decisions by state regulators to block new methods for bypass. Cellular telephones, for example, formerly "car phones," are now used more widely,

providing an alternative to the local exchange. Conflict between state and federal regulators may also limit states' abilities to use restrictions to promote cross subsidization. The federal preemption of DTS restrictions mentioned above is one example.

Even though a long-term policy of restricting competition may fail, regulators may still find it in their interest to make a gradual transition to higher rates. Kahn (1984) suggests that the length of the transition in a state should depend on the costs of providing access and on the possibilities for bypass. Where access costs are high, rates must increase more in order to meet these costs, and the benefits of a long transition period are higher. Where few bypass alternatives exist, a longer transition period during which rates for some services are priced above cost will be easier to maintain. Thus, Kahn suggests longer transitions in rural states and relatively short transitions in large cities where access costs are lower and bypass alternatives are more prevalent.[10] Regulators may exploit this relationship when weighing the relative merits of bypass restrictions.

The decision to restrict bypass is only one of the factors determining the amount of cross subsidization possible in a state. Additional factors may also be important in determining rates. For example, Noll and Smart found that dividing the states into size classes illuminated inherent pricing differences among different states. Using data on monthly basic service rates for residences and small businesses, they showed that states which have very large cities had smaller rate increases immediately after divestiture than did other states. On the other hand, states which have no large cities had higher than average rate increases. This re-

[10]Kahn (1984), p. 150.

sult suggests that states with a large urban area may have additional resources for cross subsidization. These findings indicate that other state characteristics may influence the state's ability to subsidize certain classes of users. Thus, states which have such resources may exhibit a longer transition period in spite of the factors discussed by Kahn. The state characteristics associated which these resources are discussed in section five.

Evans and Garber (1989) have developed a theoretical model which predicts a state regulator's response to the bypass threat when bypass opportunities are exogenous to the regulator's decision making. In a model in which regulators balance the interests of consumers and the local exchange carriers in setting residential prices, they predict that enhanced opportunities for bypass will result in higher residential prices. Regulators, however, have considerable influence over bypass opportunities; the extent of bypass and basic service rates are simultaneously determined by state regulators. As a result, the interests of potential users of bypass technologies must also be considered by state regulators. The endogeneity of the bypass decision is made explicit in the empirical model which follows.

3 The Model

The estimation employs a simultaneous equations, switching regressions model, introduced by Maddala and Lee (1976) and Heckman (1978). These authors derived a general class of models involving systems of equations in which one or more dependent variable is discrete. In this paper I estimate a special case of this general class of models.

In the model used here, the rate for each class of telephone service is a continu-

ous endogenous variable. The decision to restrict bypass is also endogenous, and is a discrete choice limited dependent variable. The disturbances in these equations may be correlated due to unmeasured factors which affect both bypass restrictions and rate setting. This leads to the following specification.[11]

$$y_1 = \beta_1 y_2 + \gamma_1' x_1 + u_1 \tag{1}$$

$$y_2^* = \beta_2 y_1 + \gamma_2' x_2 + u_2 \tag{2}$$

$$y_2 = 1 \text{ if } y_2^* > 0 \tag{3}$$

$$y_2 = 0 \text{ if } y_2^* \leq 0$$

where

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}\right).$$

The errors u_1 and u_2 have a joint normal distribution with mean 0, covariance σ_{12}, $\text{var}(u_1) = \sigma_1^2$ and $\text{var}(u_2) = \sigma_2^2$.

In this specification, y_1 is a vector of monthly basic rates for telephone service within a particular size class of cities. The state regulators' valuation for blocking bypass, y_2^*, is unobserved. The observed bypass restriction decision, y_2, takes the value one for states in which bypass is restricted and the value zero otherwise. Both the bypass restriction decision and the rate for telephone service, y_1, are assumed to depend on exogenous variables, x_1 and x_2, as well as other endogenous variables.

[11]Maddala (1983), p. 120 examines estimation of these types of models.

In the model under consideration, I assume that $\beta_2 = 0$, which has the following interpretation.[12] First, the decision to restrict bypass is based only upon factors exogenous to the system. Then rates are determined, and these depend on the current regulatory environment, specifically on whether bypass is currently restricted, as well as pre-divestiture rates and other exogenous variables. This is called a recursive model in the literature, but this definition differs from the standard simultaneous equations recursive models in which the errors are not correlated across equations.

The dichotomous nature of the regulatory decision also limits the identification of the parameters of the model. Rewriting the system above in terms of the parameters which can be estimated (and including the restriction that $\beta_2 = 0$) yields

$$y_1 = \beta_1 y_2 + \gamma_1' x_1 + u_1 \tag{4}$$

$$y_2^{**} = \frac{y_2^*}{\sigma_2} = \frac{\gamma_2'}{\sigma_2} x_2 + \frac{u_2}{\sigma_2} \tag{5}$$

Because the identifiable parameters are β_1, γ_1, $\frac{\gamma_2}{\sigma_2}$, $\frac{\sigma_{12}}{\sigma_2}$ and σ_1, the imposition of the restriction that $\sigma_2 = 1$ can be made without loss of generality.

[12]The model outlined above requires restrictions on the parameters in order to ensure the internal consistency of the model. Without this restriction, the relationship between y_2 and y_2^* might be violated. To see this, consider the reduced form for y_2^* when there is only one rate class and hence β_1 and y_1 are scalars.

$$y_2^* = \beta_1\beta_2 y_2 + \beta_2\gamma_1' x_1 + \gamma_2' x_2 + \beta_2 u_1 + u_2$$

Note that

$$y_2 = 0 \text{ if } y_2^* \leq 0 \Rightarrow \beta_2 u_1 + u_2 \leq \beta_2\gamma_1' x_1 + \gamma_2' x_2$$
$$y_2 = 1 \text{ if } y_2^* > 0 \Rightarrow \beta_2 u_1 + u_2 > \beta_2\gamma_1' x_1 + \gamma_2' x_2 - \beta_1\beta_2$$

Consistency therefore requires that $\beta_1\beta_2 = 0$. This is termed the "principal assumption" in the literature. This ensures that one and only one of the preceding inequalities holds and rules out a contradiction such as $y_2^* > 0$ and $y_2 = 0$. (See Heckman (1978) and Schmidt (1982) for further discussion.)

Consistent estimates of these parameters can be obtained using a two-stage procedure similar to that used by Nelson and Olsen (1978) to examine a simultaneous equation Tobit model. In the first stage of this procedure, consistent estimates of the parameters of the probit equation (2) are obtained using maximum likelihood. These estimates are then used in the second stage. Rewrite equation (4) as

$$y_1 = \beta_1 F_2(\gamma_2' x_2) + \gamma_1' x_1 + w \tag{6}$$

where $w = u_1 + \beta_1(y_2 - F_2(\gamma_2' x_2))$.

The new error term, w, has mean 0 and is uncorrelated with the regressors. Therefore, OLS provides consistent estimates of the parameters β_1 and γ_1. Amemiya (1978) derived the asymptotic variance-covariance matrix for models of this type.

More efficient estimates of the parameters may be obtained by maximum likelihood. This estimation is practical when there is only one discrete variable in the system of equations. Because each discrete variable requires another integral in the likelihood function, the increase in efficiency from maximum likelihood is offset by increasing computational difficulty when there are additional discrete endogenous variables.

Because in the current application the decision to restrict bypass is the only discrete endogenous variable, a maximum likelihood model is developed to examine the relationship between bypass restrictions and the telephone rate-setting decision. In this model, both residential rates (y_0) and business rates (y_1) are continuous variables, while the decision to regulate bypass is a discrete choice (y_2).

For the three equation case,

$$y_0 = \beta_0 y_2 + \gamma_0' x_0 + u_0 \tag{7}$$

$$y_1 = \beta_1 y_2 + \gamma_1' x_1 + u_1 \tag{8}$$

$$y_2^* = \gamma_2' x_2 + u_2 \tag{9}$$

$$y_2 = 1 \text{ if } y_2^* > 0$$

$$y_2 = 0 \text{ if } y_2^* \leq 0$$

where

$$\begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} \sim N(0, \Sigma), \text{ and } \Sigma = \begin{bmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{02} \\ \sigma_{01} & \sigma_1^2 & \sigma_{12} \\ \sigma_{02} & \sigma_{12} & 1 \end{bmatrix}$$

To estimate this model by maximum likelihood, the following joint densities are used.

$$g(y_0, y_1, y_2 = 1) = \int_{-\gamma_2' x_2}^{\infty} f(u_0, u_1, u_2) du_2$$

$$g(y_0, y_1, y_2 = 0) = \int_{-\infty}^{-\gamma_2' x_2} f(u_0, u_1, u_2) du_2$$

where $f(u_0, u_1, u_2)$ is the joint density function of (u_0, u_1, u_2). This leads to the likelihood function

$$L = \prod_{y_2=0} \int_{-\infty}^{-\gamma_2' x_2} f(u_0, u_1, u_2) du_2 \prod_{y_2=1} \int_{-\gamma_2' x_2}^{\infty} f(u_0, u_1, u_2) du_2$$

which can be simplified by noting that $f(u_0, u_1, u_2) = f(u_0, u_1) f(u_2 \mid u_0, u_1)$. The distribution of u_2 given u_0 and u_1 is:

$$(u_2 \mid u_0, u_1) \sim N\left(\begin{bmatrix} \sigma_{02} \\ \sigma_{12} \end{bmatrix}' \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix}^{-1} \begin{bmatrix} u_0 \\ u_1 \end{bmatrix}, 1 - \begin{bmatrix} \sigma_{02} \\ \sigma_{12} \end{bmatrix}' \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{02} \\ \sigma_{12} \end{bmatrix} \right). \tag{10}$$

Making this substitution in the likelihood function yields:

$$L = \prod_{y_2=0} \int_{-\infty}^{-\gamma_2' x_2} f(u_0, u_1) f(u_2 \mid u_0, u_1) du_2 \prod_{y_2=1} \int_{-\gamma_2' x_2}^{\infty} f(u_0, u_1) f(u_2 \mid u_0, u_1) du_2 \qquad (11)$$

The log of the likelihood function is then:

$$\ln L = \sum \ln[f(u_0, u_1)] + (1 - y_2) \sum \ln \left(\Phi \left(\frac{-\gamma_2' x_2 - E(u_2|u_0,u_1)}{\sqrt{Var(u_2|u_0,u_1)}} \right) \right) + y_2 \sum \ln \left(1 - \Phi \left(\frac{-\gamma_2' x_2 - E(u_2|u_0,u_1)}{\sqrt{Var(u_2|u_0,u_1)}} \right) \right)$$

in which Φ is the standard normal distribution function and u_2 has been transformed into a standard normal disturbance using equation (10) .

In the estimation, equations (7) - (9) can be estimated separately for each rate class in order to test the hypothesis that both σ_{02} and σ_{12} are zero. The acceptance of this hypothesis would allow the estimation to employ a simplified form of the likelihood function (11).

In this simplified form, $f(u_2 \mid u_0, u_1) = f(u_2)$ and

$$L = \prod_{y_2=0} \int_{-\infty}^{-\gamma_2' x_2} f(u_0, u_1) f(u_2) du_2 \prod_{y_2=1} \int_{-\gamma_2' x_2}^{\infty} f(u_0, u_1) f(u_2) du_2.$$

The simplified log of the likelihood function is

$$\ln L = \sum \ln[f(u_0, u_1)] + (1 - y_2) \sum_{y_1=0} \ln \left(\Phi \left(\frac{-\gamma_2' x_2}{\sigma_2} \right) \right) + y_2 \sum_{y_0=1} \ln \left(1 - \Phi \left(\frac{-\gamma_2' x_2}{\sigma_2} \right) \right) \qquad (12)$$

Note that maximizing equation (12) is equivalent to estimating equations (7) and (8) jointly and estimating equation (9) separately. Thus, the rate equations

would be estimated jointly, taking into account the correlation between the equations' errors, and the bypass decision equation would be estimated separately using maximum likelihood (probit).

4 Data

The data on monthly telephone rates for residential and small business subscribers is taken from the *Bell Operating Companies' Exchange Service Telephone Rates* published by the National Association of Regulatory Utility Commissioners. For residential consumers I collected the rate which allows unlimited local calling for a flat monthly fee. The rate collected for businesses is for single-line service, also allowing unlimited local calls for a flat monthly fee. Single-line service is adequate for only the smallest businesses. Rates were collected from each telephone company in the Bell system on December 30 in 1985 and 1988. I will use these rates to analyze the effects of bypass restrictions immediately following divestiture and the longer-term effects of these prohibitions. Rates for 1980 and for 1981 were collected from each company on June 30 of the corresponding year. Since political concerns suggest that regulators will attempt to avoid increasing rates too rapidly, the predivestiture rate structures represent one of the constraints faced by policymakers.

The unit of observation for telephone rates is a company. Many states have only one local exchange company. The most common exception to this rule is an independent company operating in a single market. These companies are not included in the sample. Within some states, however, two Bell companies operate, and they do not charge the same rates. These companies are included in the

sample.

In general, companies report a range of rates for each type of service, which differ depending on the size of the exchange being served. Corresponding to each rate, therefore, is a range of exchange sizes to which the rate applies. In order to make comparisons among companies, rates were collected for each company according to size of exchange. Ten exchange sizes were chosen—the smallest exchange, an exchange with 1,000 terminals, and one each with 5,000, 25,000, 50,000, 100,000, 250,000, 500,000, 750,000 and one million terminals or more. Rates for each of these were taken from the appropriate range of exchange sizes for each company. The exchanges with 5,000 or fewer terminals are the rural areas and small towns of a few thousand population. The federal government's Metropolitan Statistical Areas generally correspond to exchanges with 25,000 or more terminals. Typically, the number of terminals is approximately half of the population of an exchange. Thus, a metropolitan area with 100,000 terminals will have a population of approximately 200,000. The nation's largest cities have 1 million terminals or more.

Many states have no large exchanges and therefore do not quote rates. For example, Tables One and Two contain the rates charged for basic service in Arkansas and Florida, respectively. Arkansas has no large exchanges, and some rates are therefore missing. Florida has rates for all exchange sizes.

The unit of observation for restrictions on bypass technology is the state. Information on state regulation of bypass is taken from Appendix B of Huber's *The Geodesic Network* (1987) prepared for the Department of Justice. A state is defined as restrictive if it imposed restrictions on one of the forms of bypass noted

TABLE ONE: Arkansas rates for basic service, 1985										
Size of Exchange	smallest	1,000	5,000	25,000	50,000	100,000	250,000	500,000	750,000	1,000,000
Residential single line service	11.18	11.18	12.58	13.98	13.98	15.38				
Business single line service	23.36	23.36	26.26	29.21	29.21	32.16				

TABLE TWO: Florida rates for basic service, 1985										
Size of Exchange	smallest	1,000	5,000	25,000	50,000	100,000	250,000	500,000	750,000	1,000,000
Residential single line service	8.55	8.55	8.95	9.75	9.75	10.15	10.95	11.80	12.20	12.20
Business single line service	20.35	20.35	21.35	23.45	23.45	24.40	26.30	28.55	29.65	29.65

at the end of section one or if no carrier in the state has applied for these forms of bypass. In the models estimated here, alternate definitions of what constitutes effective restriction of bypass technology produced similar results.

Included in the estimation of the bypass equation (9) are variables measuring the percent of state employment in industries which are most affected by restriction of bypass technology. As discussed in Section 2, these are the telecommunications and finance industries. The variables Telecomm and Finance measure employment in these industries. Also included are variables which describe the pre-divestiture regulatory equilibrium in each state. States with higher local plant costs and a higher fraction of those costs paid for by interstate long distance prior to divestiture face potentially large increases in local rates. Regulators in these states are,

therefore, more likely to need to restrict bypass in order to keep rates down. Also, states with lower pre-divestiture rates in rural exchanges should be more likely to need to restrict bypass in order to continue cross subsidization. A dummy variable, Elected Comm., is also included to determine the differences between elected and appointed commissioners with respect to the issue of bypass.

The bypass restriction decision then enters as an explanatory variable in the rate estimation equations. In addition, I include variables which incorporate distributional aspects of the rate decision and variables which relate to the availability of resources for subsidization within the state. I also control for the pre-divestiture level of rates. Regulators may be concerned with the effects of pricing decisions on certain segments of the population. I include variables which measure the age distribution of the population and the percent of the population living in poverty. These variables may influence the setting of residential rates. Rates for small businesses may likewise be influenced by the percent of state employment in small enterprises. Per capita income should influence the states' ability to set lower rates for both classes of service. Consumers with higher incomes demand more services from the BOCs. Following Noll and Smart, size class variables are included which indicate the size of the largest exchange in the state. States with very large cities have greater resources to use in cross subsidization because a lower proportion of their population live in relatively high cost rural areas.

The remaining variables and their definitions are summarized in Table Three. A table of summary statistics is provided in Table Four.

TABLE THREE	
Variable	Definition
Bypass	Dummy variable equal to one if the state has any of the following restrictions on bypass: – does not allow competing metropolitan or local area networks – does not permit interexchange carriers to construct bypass systems – does not allow customer-owned bypass systems or taxes them – does not allow competing Digital Termination Systems – has other restrictions on local facilities-based competition
Urban-Met. pop	% of population living in urbanized metropolitan areas, 1980
Income	Per capita income in thousands of dollars, 1985
Per<17	% of population under the age of 17, 1980
Per>65	% of population over the age of 65, 1980
Poverty	% of the population living in poverty, 1979
Finance	% of employment in the Fin., Ins. and Real Estate sector, 1985
Small business	% of employment in firms with 10 or fewer employees, 1985
Small state	Dummy variable equal to one for states with no exchange larger than 100,000 terminals
Medium state	Dummy variable equal to one for states with largest exchange either 250,000 or 500,000 terminals
Rur. Res. 80-81	Pre-divestiture residential rate in smallest exchange averaged over 1980-1981
Residential 80-81	Pre-divestiture residential rate averaged over 1980-1981 averaged over 1980-1981
Business 80-81	Pre-divestiture business rate averaged over 1980-1981
Elected Comm.	Dummy variable equal to one if regulatory commission elected
Telecomm.	% of state employment in telecommunications sector
SPF	Subscriber Plant Factor–percentage of local plant costs allocated to the interstate jurisdiction, 1981
Bell Atlantic	Dummy variable equal to one if the state is in this Regional Bell Holding Company (RBHC)
Southwest Bell	Dummy variable equal to one if the state is in this RBHC
Bell South	Dummy variable equal to one if the state is in this RBHC
Nynex	Dummy variable equal to one if the state is in this RBHC
Pac Tel	Dummy variable equal to one if the state is in this RBHC
US West	Dummy variable equal to one if the state is in this RBHC

TABLE FOUR						
DESCRIPTIVE STATISTICS						
	Variable Name		Mean	Stan. Dev.	Min	Max
State Characteristics	Bypass		0.592	0.497	0.000	1.000
	Population		0.048	0.051	0.005	0.264
	Percent<17		28.627	2.245	22.500	37.000
	Percent>65		11.198	1.786	7.500	17.300
	Rural Res. 80-81		6.624	1.633	3.550	10.155
	Elected Comm.		0.224	0.422	0.000	1.000
	Telecomm.		1.118	0.374	0.000	1.770
	Income		13.137	2.081	9.187	18.168
	Poverty		12.712	3.540	8.000	24.500
	Finance		5.575	1.108	3.904	9.271
	SPF		30.500	9.406	16.900	62.100
	Small Bus.		29.287	5.498	21.174	46.531
	Urban-met pop.		53.238	23.466	12.613	100.000
	Bell Atlantic		0.143	0.354	0.000	1.000
	Southwest Bell		0.102	0.306	0.000	1.000
	Bell South		0.184	0.391	0.000	1.000
	Nynex		0.143	0.354	0.000	1.000
	Pac Tel		0.041	0.200	0.000	1.000
	US West		0.286	0.456	0.000	1.000
	Small state		0.224	0.422	0.000	1.000
	Medium state		0.367	0.487	0.000	1.000
Rural	Residential rate	(1985)	10.375	2.635	5.610	19.100
		(1988)	10.396	2.550	6.480	19.100
Exchanges	Business rate	(1985)	25.177	7.677	12.160	43.560
		(1988)	25.627	7.264	13.800	43.560
Sm. City	Residential rate	(1985)	12.337	2.890	7.450	21.000
		(1988)	12.181	2.865	7.450	21.000
Exchanges	Business rate	(1985)	32.262	8.766	16.300	50.740
		(1988)	31.781	8.122	16.300	50.960
Med-large city	Residential rate	(1985)	13.770	3.203	7.800	23.390
		(1988)	13.427	2.860	9.350	23.390
Exchanges	Business rate	(1985)	37.870	10.033	22.000	58.230
		(1988)	37.090	9.222	22.000	58.230

5 Estimation and Results

In this section I report results from estimating equation (12).[13] The state regulators' valuation for bypass depends on the regulatory structure in the pre-divestiture equilibrium and on the size of interest groups affected directly by the bypass decision. In addition, the decision depends on whether the regulatory commission is elected or appointed. The results of the estimation of the bypass equation are shown in Table Five.[14] Parameter estimates from the probit are reported in column one. Columns two and three report probability derivatives. In column two, the dummy variable for elected commissioners is equal to one, and all other variables are evaluated at their means. In column three, the dummy variable for elected commissioners is equal to zero.

As expected, the variables describing the regulatory equilibrium within the state prior to divestiture predict well the probability that regulators will block competition with the local exchange carriers in their state. The variable Rur. Res. 80-81 measures rates in the smallest exchange in the state prior to divestiture. States with the lowest rural exchange prices were more likely to impose bypass restrictions, which in turn allowed them to keep rates lower. Pre-divestiture prices $1.00 lower increased the probability of observing bypass restrictions by approximately ten percent. The variable SPF, which measures the fraction of the costs

[13]Before estimating equation (12), I first estimated the unrestricted version of the likelihood function. Using first the Wald test and then the likelihood ratio test, I was unable to reject the hypothesis that both σ_{02} and σ_{12} are equal to zero. Results of this estimation for the smallest exchange size are reported in Table Eight. Similar results were obtained for the larger exchange sizes. Note that the coefficient estimates in the simplified model are roughly the same as those in the unrestricted model.

[14]Teske (1990) has also looked at the effect of interest groups on state regulatory decisions regarding competition. He did not consider the regulation of bypass technology in his paper.

of the local plant which were allocated to the interstate jurisdiction in 1981, provides an indication of the federal subsidy that is declining since divestiture. States with the highest subsidy faced the prospect of the largest increases in local rates following divestiture. A one percent higher SPF prior to divestiture increased the probability of restricting bypass by two percent.

TABLE FIVE

	Probit Est.	Probability Derivatives	
	Bypass	Elected Comm.	Appointed Comm.
Variable name	Estimate	Estimate	Estimate
Intercept	4.909 **	1.447 *	1.516 **
	(2.124)	(0.764)	(0.638)
Rur. Res. 80-81	-0.384 **	-0.113 **	-0.118 **
	(0.157)	(0.056)	(0.048)
Finance	-0.601 **	-0.177 *	-0.185 **
	(0.256)	(0.092)	(0.076)
Elected Comm.	-1.494 **	-0.440 ***	-0.461 ***
	(0.623)	(0.081)	(0.156)
Telecomm.	-0.325	-0.096	-0.100
	(0.707)	(0.208)	(0.217)
SPF	0.068 *	0.020 *	0.021 **
	(0.036)	(0.012)	(0.0099)
N	49		
Percent predicted	69		
Log L	-23.41		

* significant at .10 level
** significant at .05 level
*** significant at .01 level

The variables measuring the relative strengths of organized interest groups are Finance and Telecomm. The financial sector, one of the largest users of telecommunications services[15], is the sector for which bypass may be an attractive alternative to the services of the local exchange. The results indicate that large, well organized interest groups are able to influence policy outcomes. A one percent increase

[15]See, for example, the discussion by Carnevale (1989).

in the percentage of employment in the finance industry decreases the probability of bypass restrictions by almost twenty percent. The variable Telecomm, which covers the entire telecommunications sector, has no effect on the bypass decision. Because this measure includes all employment in the telecommunications sector and does not pinpoint the employment of the regulated local telephone companies, it is a poor measure of the political strength of the BOCs.

The coefficient on elected commissioners indicates that states which elect commissioners are less likely to restrict bypass technology. These regulators depend on direct campaign support and may be more responsive to interest groups which can mobilize support for a candidate or provide campaign contributions.

Table Six reports the results of the estimation of the equation governing the rate setting decision in 1985.[16] For clarity in presentation, I focus on four of the ten exchange sizes which were collected from each state. From each state, I analyze the rate from a rural exchange, the rate from a small metropolitan area with 50-100,000 terminals, the rate from a medium-sized exchange with 250-500,000 terminals, and the rate from the largest metropolitan areas with 750,000 terminals or more. Again, all four exchange sizes are not available in each state. In fact, only thirteen companies have rates in the largest exchange class. The hypothesis that these rates may be pooled with medium-sized exchange rates cannot be rejected, however, and for this size class pooled regression results are shown. The dependent variable is the log of the telephone rate charged in each exchange.

State regulators who have chosen to limit bypass competition within their

[16]Table Nine shows the results of the complete estimation, which includes dummy variables which control for the geographic region of the telephone companies. The main results are presented in Table Six.

TABLE SIX

1985

Variable name	Rural Exch. Residential exchanges		Rural Exch. Business exchanges		Small Cities Residential exchanges		Small Cities Business exchanges		Med. & Large Cities Residential exchanges		Med. & Large Cities Business exchanges	
Bypass	-0.139	**	-0.233	***	-0.081		-0.083		-0.027		-0.0019	
	(0.064)		(0.069)		(0.055)		(0.068)		(0.065)		(0.087)	
Intercept	1.912	***	3.987	***	1.619	***	3.592	***	2.140	***	5.247	***
	(0.687)		(0.397)		(0.599)		(0.466)		(0.656)		(0.407)	
Ln(Res 80-81)	0.358	***			0.375	***			0.077	**		
	(0.088)				(0.095)				(0.037)			
Ln(Bus 80-81)			0.377	***			0.337	***			0.032	
			(0.074)				(0.104)				(0.033)	
Pop<17	0.023	*			0.030	***			0.035	***		
	(0.013)				(0.011)				(0.011)			
Pop>65	0.012				0.019	*			0.029	***		
	(0.012)				(0.011)				(0.0094)			
	(0.025)		(0.024)		(0.021)		(0.024)		(0.026)		(0.026)	
Poverty	-0.0026				-0.0074				0.00064			
	(0.0074)				(0.0063)				(0.0071)			
Small Bus.			-0.874				-0.126				-0.692	
			(0.533)				(0.527)				(0.586)	
Urb-met pop	-0.00028		0.0017		-0.0050	***	-0.0048	**	-0.0097	***	-0.012	***
	(0.0019)		(0.0022)		(0.0016)		(0.0021)		(0.0017)		(0.0022)	
Income	-0.075	***	-0.128	***	-0.043	**	-0.067	***	-0.022		-0.048	**
	(0.025)		(0.024)		(0.021)		(0.024)		(0.026)		(0.026)	
Small state	-0.053		0.036		-0.166	**	-0.155	*				
	(0.095)		(0.102)		(0.093)		(0.074)					
Med. state	0.112	*	0.149	**	-0.021		-0.0057		-0.088	**	-0.175	***
	(0.060)		(0.065)		(0.059)		(0.046)		(0.045)		(0.059)	

* significant at .10 level
** significant at .05 level
*** significant at .01 level

states have been successful in keeping rates lower than they would have been in the absence of bypass restrictions, especially in the rural exchanges. The coefficient on the variable Bypass indicates that in rural exchanges residential rates are 14 percent lower where bypass is restricted. In a rural exchange in a state which has no large cities, rates are 14 percent or about $1.29 lower. For business single-line service, rates are 23 percent, or about $5.08, lower.[17] Large cities have an even higher margin due to the bypass restriction, up to $1.50 for residential rates though the subsidy is smaller in percentage terms since these rates are higher. It is interesting to note that small businesses in rural exchanges and small cities receive

[17]The dollar approximations are for a state with no large cities in the Ameritech RBOC with all other variables set at their mean values. (See Table Four for means.)

a larger subsidy from bypass than do residences, both in absolute and percentage terms. Business customers face the largest price increase when bypass competition is permitted.

Rural exchanges are the primary beneficiaries of bypass restrictions. For small and medium-large exchanges, the magnitude of the subsidy is smaller and the coefficient is no longer significant. This does not imply, however, that customers in these exchanges receive no subsidy. Included in the regression are rates charged prior to divestiture, the average rate from 1980-81. States with lower rates and higher subsidies prior to divestiture continued to have lower rates in 1985. For both residential and business rates in rural areas and small cities, the influence of past prices on current prices is similar, an elasticity estimate of approximately .35. For the largest exchange size class, however, the influence is much smaller. Again, the structure which existed prior to divestiture is an important determinant of the post-divestiture structure of prices, as suggested by Joskow's theory.

The next group of variables incorporates distributional aspects of the politics of state regulators. In spite of the concern voiced by Congress and state regulators for the effect rate increases may have on certain segments of the population, states with the highest percentage of their population over the age of 65 actually charged significantly higher rates in the largest exchange sizes. In addition, the proportion of the state's population living in poverty had no affect on the rates charged in any exchange.[18] Interest-group politics does a better job of explaining the post-divestiture rates. Though imprecisely estimated, the sign of the coefficient on the percent of state employment working in small businesses indicates that rates

[18] "Lifeline" programs available to low income consumers in some states target this group, so that the desire to hold price down for the poor may not affect the basic access rate.

are lower in states with more small business. States with a high fraction of the population living in urbanized metropolitan areas also charge significantly lower rates in all sizes of urban exchanges.

The remaining variables control for factors within the state which may affect the resources available for subsidization. States with high per capita income are able to charge lower rates for residential and small business customers in both rural and urban exchanges.[19] The size class dummy variables indicate the size of the largest exchange in the state. States which have exchanges with 750,000 or more terminals (the excluded dummy variable) charge the highest rates in all but the rural exchanges. In these states rural rates are as low as those in states with only rural exchanges and small cities. In states with mid-size cities, rates are higher in the rural exchanges and lower in the large urban exchanges. These results are consistent with the results of Noll and Smart who found that only in states with mid-size cities did the differential between rural and urban exchanges become smaller after divestiture. States with large urban exchanges continue to charge low rates in rural exchanges despite their large urban population centers.

Table Seven[20] represents the results of estimating the rate equations with 1988 data. The variable Bypass has not changed and is exogenous to the 1988 rate decision. The coefficient on Bypass measures the difference in 1988 rates between states which restricted bypass immediately after divestiture and states which allowed bypass to occur.

[19]Mathios and Rodgers demonstrate that in states which restrict intraLATA competition, rates for intraLATA long distance are higher. If the income-elasticity of demand for long distance is positive, wealthier people make more long distance calls and so generate more long distance subsidy for basic access.

[20]Regional dummies were also included in the 1988 estimation. The estimated coefficients are very similar to those reported in Table Nine.

TABLE SEVEN

1988

	Rural Exch.		Small Cities		Med. & Large Cities	
Variable name	Residential exchanges	Business exchanges	Residential exchanges	Business exchanges	Residential exchanges	Business exchanges
Bypass	-0.0422	-0.0719	-0.0396	-0.0039	0.0249	0.1141
	(.0718)	(.0696)	(.0620)	(.0670)	(.0525)	(.0728)
Intercept	1.0451	3.2691 ***	1.5232 **	3.5804 ***	2.1473 ***	4.4736 ***
	(.8412)	(.4092)	(.6996)	(.4350)	(.6157)	(.3490)
Ln(Res 80-81)	0.3107 ***		0.3109 ***		-0.0006	
	(.0980)		(.1033)		(.0107)	
Ln(Bus 80-81)		0.3305 ***		0.2337 ***		0.0336
		(.0719)		(.0971)		(.0199)
Pop<17	0.0291 *		0.0239 *		0.0284 ***	
	(.0167)		(.0136)		(.0110)	
Pop>65	0.0146		0.0082		0.0225 ***	
	(.0150)		(.0126)		(.0091)	
Poverty	0.0028		0.0024		0.0059	
	(.0088)		(.0076)		(.0071)	
Small Bus		-0.0076		-0.0047		-0.0011
		(.0059)		(.0053)		(.0071)
Urb-met pop	-0.0022	-0.0001	-0.0056 ***	-0.0058 ***	-0.0104 ***	-0.0122 ***
	(.0021)	(.0021)	(.0017)	(.0019)	(.0013)	(.0020)
Income	-0.0232	-0.0679 ***	-0.0169	-0.0397 **	-0.0019	-0.0155
	(.0264)	(.0222)	(.0208)	(.0193)	(.0197)	(.0212)
Small state	0.0162	0.1348	-0.0549	-0.0230		
	(.0980)	(.0965)	(.0769)	(.0834)		
Med. state	0.1707 ***	0.2433 ***	0.0464	0.0980 *	-0.0504	-0.0802
	(.0605)	(.0594)	(.0477)	(.0520)	(.0340)	(.0512)
R^2	.534	.684	.710	.744	.831	.730
N	47		44		44	
Log L	105.410		115.204		130.058	

* significant at .10 level
** significant at .05 level
*** significant at .01 level

Two aspects of these results stand out. First, states which restricted bypass following divestiture no longer have lower rates. In rural areas and small cities the coefficient on Bypass is smaller than in 1985, and in all exchange sizes it is insignificant. Second, as Kahn suggested, the transition to higher rates has been more gradual in the rural exchanges as a result of the bypass restrictions. In the early post-divestiture years rural exchanges benefitted from the bypass restrictions, and pre-divestiture rates are still important determinants of rural rates. This is not true of large cities, additional resources for cross-subsidization existed.

6 Conclusion

The results demonstrate that some state regulators have been successful in limiting rate increases, especially in the smallest exchange sizes. In rural areas, those states which restricted bypass have rates for residential service which are 14 percent lower, and the percentage difference for small business service, 23 percent, is even larger. State regulatory decisions have been affected by political concerns within their states, most notably large organized interests, and by the structure of rates which existed in their states prior to divestiture. State regulators are willing to restrict competition in order to increase opportunities for cross subsidization, but only if the political costs of the restriction are low. These political costs apparently depend on the pre-divestiture rate structure, for lower rates imply that imposing cost-based rates exacts a higher political cost. Political factors are less influential in determining the pattern of rate increases, indicating that subsidies are spread across residential and small business consumers. The exception is pricing in urban areas, which is affected by the size of the urban population. The availability of resources for providing a subsidy to local service also influences the final rate structure.

While state regulators have been willing to restrict competition, the effects of these restrictions on basic service rates have been temporary. By 1988, price differences due to the restriction of bypass activities had practically disappeared. The decline in the price differences between these types of states may be due to several factors. Perhaps states prefer to restrict bypass and continue benefits to basic service but are prevented from doing so by technological innovation. Perhaps conflict with more procompetitive federal regulators has prevented effective

restrictions. In any case, regulators in these states may have planned a more gradual transition to the new higher rates without expecting to maintain low rates in the long term. Finally, it is possible that in states which did not restrict bypass rates rose too rapidly following divestiture. In this case part of the decrease in the difference between the rates in the two types of states could be due to falling rates in the states which did not restrict bypass. æ

TABLE EIGHT
SIMULTANEOUS ESTIMATION

Variable name	Bypass			Small Residential exchanges			Small Business exchanges		
	Estimate	Standard error		Estimate	Standard error		Estimate	Standard error	
Intercept	3.874	2.271	*	2.077	0.688	***	4.304	0.421	***
Rural Res. 80-81	-0.394	0.190	**						
Residential 80-81				0.327	0.096	***			
Business 80-81							0.324	0.085	***
Finance	-0.709	0.306	**						
Elected Comm.	-1.873	0.756	**						
Telecomm.	-0.044	0.932							
SPF	0.118	0.049	**						
Bypass				-0.176	0.086	**	-0.242	0.103	**
Pop<17				0.025	0.012	**			
Pop>65				0.011	0.012				
Income				-0.085	0.023	***	-0.138	0.022	***
Poverty				-0.0042	0.0076				
Urban-met pop.				-0.00028	0.0017		0.0015	0.0021	
Small Business							-0.0088	0.0054	
Small state				-0.038	0.089		0.026	0.108	
Medium state				0.108	0.0528	**	0.137	0.062	**
Bell Atlantic				0.236	0.102	**	0.388	0.117	***
Southwest Bell				-0.0019	0.107		-0.101	0.119	
Bell South				-0.163	0.099		-0.242	0.110	**
Nynex				0.180	0.107		0.348	0.122	***
Pac Tel				0.347	0.139	**	0.427	0.157	***
US West				0.030	0.085		0.140	0.107	
σ_0^2	0.016	0.0039	***						
σ_1^2	0.023	0.0053	***						
σ_{01}	0.015	0.0040	***						
σ_{02}	-0.039	0.051							
σ_{12}	-0.051	0.065							
N	45								
R^2				0.724			0.773		
Percent predicted	75.6								
Log L	95.09								

TABLE NINE: Complete estimation of Table Six Results

	Rural Exch.		Small Cities		Med. & Large Cities	
Variable name	Residential exchanges	Business exchanges	Residential exchanges	Business exchanges	Residential exchanges	Business exchanges
Bypass	-0.139 **	-0.233 ***	-0.081	-0.083	-0.027	-0.0019
	(0.064)	(0.069)	(0.055)	(0.068)	(0.065)	(0.087)
Intercept	1.912 ***	3.987 ***	1.619 ***	3.592 ***	2.140 ***	5.247 ***
	(0.687)	(0.397)	(0.599)	(0.466)	(0.656)	(0.407)
Res 80-81	0.358 ***		0.375 ***		0.077 **	
	(0.088)		(0.095)		(0.037)	
Bus 80-81		0.377 ***		0.337 ***		0.032
		(0.074)		(0.104)		(0.033)
Pop<17	0.023 *		0.030 ***		0.035 ***	
	(0.013)		(0.011)		(0.011)	
Pop>65	0.012		0.019 *		0.029 ***	
	(0.012)		(0.011)		(0.0094)	
Poverty	-0.0026		-0.0074		0.00064	
	(0.0074)		(0.0063)		(0.0071)	
Small Bus		-0.874		-0.126		-0.692
		(0.533)		(0.527)		(0.586)
Urb- met pop	-0.00028	0.0017	-0.0050 ***	-0.0048 **	-0.0097 ***	-0.012 ***
	(0.0019)	(0.0022)	(0.0016)	(0.0021)	(0.0017)	(0.0022)
Income	-0.075 ***	-0.128 ***	-0.043 **	-0.067 ***	-0.022	-0.048 **
	(0.025)	(0.024)	(0.021)	(0.024)	(0.026)	(0.026)
Small state	-0.053	0.036	-0.166 **	-0.155 *		
	(0.095)	(0.102)	(0.093)	(0.074)		
Med. state	0.112 *	0.149 **	-0.021	-0.0057		
	(0.060)	(0.065)	(0.059)	(0.046)	-0.088 **	-0.175 ***
Bell At	0.175	0.372 ***	0.170	0.107	(0.045)	(0.059)
	(0.108)	(0.117)	(0.106)	(0.085)	-0.032	-0.063
SW Bell	-0.044	-0.098	-0.154	-0.073	(0.074)	(0.099)
	(0.114)	(0.118)	(0.108)	(0.090)	-0.365 ***	-0.473 ***
Bell South	-0.134	-0.201 *	-0.066	-0.080	(0.072)	(0.094)
	(0.108)	(0.112)	(0.103)	(0.086)	-0.222 ***	-0.255 **
Nynex	0.127	0.313 ***	0.334	0.149 ***	(0.090)	(0.115)
	(0.114)	(0.121)	(0.113)	(0.090)	0.165	0.322
Pac Tel	0.284 *	0.388 ***	0.161	0.183	(0.157)	(0.201)
	(0.153)	(0.163)	(0.148)	(0.121)		
US West	-0.0057	0.141	0.074	-0.0100	-0.160 ***	-0.162 *
	(0.090)	(0.105)	(0.095)	(0.070)	(0.061)	(0.089)
σ_0^2	0.022 ***		0.013 ***		0.010 ***	
	(0.0045)		(0.0029)		(0.0023)	
σ_1^2	0.026 ***		0.021 ***		0.019 ***	
	(0.0053)		(0.0044)		(0.0041)	
σ_{01}	0.020 ***		0.013 ***		0.012 ***	
	(0.0045)		(0.0032)		(0.0028)	
R^2	.634	.737	.750	.732	.806	.743
N	48		45		42	
Log L	114.035		120.92		124.077	

References

[1] Amemiya, Takeshi. "The Estimation of a Simultaneous Equation Generalized Probit Model". *Econometrica*, 46(5):1193–1205, September 1978.

[2] Carnevale, Mary Lu. Finance World is Newest Turf in Phone Battle. *The Wall Street Journal*, November 15, 1989.

[3] Downs, Anthony. *An Economic Theory of Democracy.* Harper & Row, New York, 1957.

[4] Evans, Lewis and Steven Garber. "The Rational Regulator's Pricing Response to Enhanced Opportunities to 'Bypass' the Public Telephone Network". *Journal of Regulatory Economics*, 1:271–291, 1989.

[5] Heckman, James J. "Dummy-Endogenous Variables in a Simultaneous Equation System". *Econometrica*, 46(6):931–59, July 1978.

[6] Huber, Peter W. *The Geodesic Network: 1987 Report on Competition in the Telephone Industry* . U. S. Government Printing Office, Washington, D. C., January 1987.

[7] Joskow, Paul L. "Inflation and Environmental Concern: Structural Change in the Process of Public Utility Price Regulation". *Journal of Law and Economics*, 17:291–327, October 1974.

[8] Kahn, Alfred E. "The Road to More Intelligent Telephone Pricing". *Yale Journal on Regulation*, 1:139–57, 1984.

[9] Maddala, G. S. *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, 1983.

[10] Maddala, G. S. and Lung-Fei Lee. "Recursive Models with Qualitative Endogenous Variables". *Annals of Economic and Social Measurement*, 5(4):525–45, 1976.

[11] Mathios, Alan and Robert P. Rogers. *"The Impact and Politics of Entry Regulation on Intrastate Telephone Rates". Journal of Regulatory Economics*, 2:53–68, March 1990.

[12] McCubbins, Matthew D., Roger G. Noll and Barry R. Weingast. "Structure and Process, Politics and Policy: Administrative Arrangements and the Political Control of Agencies ". *Virginia Law Review*, 75(2), March 1989.

[13] National Association of Regulatory Utility Commissioners. *Bell Operating Companies Exchange Service Telephone Rates.* Washington, D. C., December 1985–1988.

[14] National Association of Regulatory Utility Commissioners. *Exchange Service Telephone Rates.* Washington, D. C., June 1980–1981.

[15] *National Association of Regulatory Utility Commissioners, Annual Report, 1986.* Washington, D. C., 1987.

[16] Nelson, F. S. and L. Olsen. "Specification and Estimation of a Simultaneous Equation Model with Limited Dependent Variables". *International Economic Review*, 19:695–710, 1978.

[17] Noll, Roger G. "State Regulatory Response to Competition and Divestiture in the Telecommunications Industry". In Ronald E. Grieson, editor, *Antitrust and Regulation*, Lexington Books, Lexington, MA, 1986.

[18] Noll, Roger G. and Susan R. Smart. "Pricing of Telephone Services". In Barry Cole, editor, *After the Breakup: Assessing the New Post-AT&T Divestiture Era*, Columbia University Press, 1991.

[19] U. S. Bureau of Economic Analysis. *Survey of Current Business.* August 1986.

[20] U. S. Bureau of Labor Statistics. *Employment and Earnings.* May 1986.

[21] Schmidt, Peter. "Constraints on the Parameters in Simultaneous Tobit and Probit Models". In Charles F. Manski and Daniel McFadden, editors, *Structural Analysis of Discrete Data with Econometric Applications*, The MIT Press, Cambridge, MA, 1982.

[22] Teske, Paul E. In *After Divestiture: The Political Economy of State Telecommunications Regulation.* State University of New York Press, Albany, NY, 1990.

[23] United States Department of Commerce, Bureau of the Census. *County Business Patterns.* U. S. Government Printing Office, Washington, D. C., 1985.

[24] United States Department of Commerce, Bureau of the Census. *State and Metropolitan Area Data Book, 1982.* U. S. Government Printing Office, Washington, D. C., 1982.

[25] United States Department of Commerce, Bureau of the Census. *State and Metropolitan Area Data Book, 1986.* U. S. Government Printing Office, Washington, D. C., 1986.

[26] Wynns, Peyton L. *The Changing Telephone Industry: Access Changes, Universal Service, and Local Rates.* The Congress of the United States, Congressional Budget Office, Washington, D. C., June 1984.

DATA APPENDIX

In order to simplify the analysis, the paper focuses on only four exchange sizes for each state. In this appendix, I offer the regression results for each of the ten exchange sizes shown in Tables One and Two.

The appendix tables for the smallest exchange size and for exchanges of 1,000 terminals are very similar to the results for rural exchanges shown in Tables Six and Seven. In rural exchanges, states which restrict bypass have lower rates in 1985 than do states which allow bypass. The 1988 rates, however, are not significantly different in those states which initially restricted bypass.

The appendix tables for exchanges with 25,000, 50,000 and 100,000 terminals are very similar to the results for small city exchanges shown in Tables Six and Seven. The difference in rates between states due to bypass restrictions is smaller and more imprecisely estimated in these larger exchange sizes. The appendix table for exchanges with 250,000 terminals has an anomalous result. The effect of bypass in residential exchanges in 1988 is both larger than expected and significant. The remainder of the table conforms to the expected pattern.

The final table of the appendix contains results for the three largest exchange sizes, corresponding to the medium and large cities in Tables Six and Seven. Due to the small number of observations in these largest exchanges, these three exchange sizes have been

pooled. The appendix table allows for both the coefficient on bypass and the intercept to vary for each exchange size. In Tables Six and Seven I was able to restrict these interaction terms to equal zero, but a likelihood ratio test shows that with the data used in the appendix the hypothesis that these terms are equal to zero can be rejected.

Two results in this final table differ from those in the paper. First, the bypass coefficient is larger in the exchange with 500,000 terminals (and significant in one case). But, as expected, the bypass increment terms show that as the exchange size gets larger, the effect of bypass is smaller. Second, the coefficient on rates from 1980-81 is much larger, perhaps due to the data disaggregation.

The full range of data support the results and conclusions based on the data used in the text of the paper.

Size of Exchange: Smallest				
	1985		1988	
Variable name	Residential exchanges	Business exchanges	Residential exchanges	Business exchanges
Bypass	-0.1389 **	-0.2333 ***	-0.0702	-0.1108
	(.064)	(.069)	(.070)	(.070)
Intercept	1.9115 ***	3.9871 ***	1.2082	3.4977 ***
	(.687)	(.397)	(.825)	(.412)
Res 80-81	0.3583 ***		0.3268 ***	
	(.088)		(.097)	
Bus 80-81		0.3769 ***		0.3404 ***
		(.074)		(.074)
Pop<17	0.0232 *		0.0285 *	
	(.013)		(.017)	
Pop>65	0.0117		0.0118	
	(.012)		(.014)	
Poverty	-0.0026		0.0014	
	(.007)		(.008)	
Small Bus		-0.0087		-0.0099 *
		(.005)		(.006)
Urb-met pop	-0.0003	0.0017	-0.0016	0.0008
	(.002)	(.002)	(.002)	(.002)
Income	-0.0751 ***	-0.1277 ***	-0.0342	-0.0839 ***
	(.025)	(.024)	(.025)	(.022)
Small state	-0.0527	0.0362	-0.0039	0.1113
	(.095)	(.102)	(.098)	(.100)
Med. state	0.1116 *	0.1491 **	0.1636 ***	0.2335 ***
	(.060)	(.065)	(.061)	(.062)
Bell At	0.1746	0.3722 ***	0.1859	0.3883 ***
	(.109)	(.117)	(.115)	(.116)
SW Bell	-0.0444	-0.0981	0.0154	-0.0598
	(.114)	(.118)	(.120)	(.116)
Bell South	-0.1340	-0.2008 *	-0.0850	-0.0683
		(.112)	(.118)	(.114)
Nynex	0.1075	0.3132 **	0.1095	0.2546 **
	(.114)	(.121)	(.120)	(.120)
Pac Tel	0.2844 *	0.3883 **	0.3452 **	0.3851 **
	(.153)	(.163)	(.161)	(.160)
US West	-0.0057	0.1408	0.0810	0.2283 **
	(.090)	(.105)	(.095)	(.104)
σ_0^2	0.0219 ***		0.0247 ***	
	(.005)		(.005)	
σ_1^2	0.0260 ***		0.0257 ***	
	(.005)		(.005)	
σ_{01}	0.0196 ***		0.0194 ***	
	(.005)		(.005)	
N	48		48	
Log L	114.0		106.1	

Size of Exchange: 1000				
	1985		1988	
Variable Name	Residential exchanges	Business exchanges	Residential exchanges	Business exchanges
Bypass	-0.1272 **	-0.2178 ***	-0.0611	-0.0991
	(.063)	(.065)	(.070)	(.066)
Intercept	1.7817 **	3.9025 ***	0.9822	3.4285 ***
	(.682)	(.368)	(.820)	(.377)
Res 80-81	0.3908 ***		0.3686 ***	
	(.090)		(.099)	
Bus 80-81		0.3988 ***		0.3607 ***
		(.074)		(.073)
Pop<17	0.0241 *		0.0312 *	
	(.013)		(.016)	
Pop>65	0.0121		0.0130	
	(.012)		(.014)	
Poverty	-0.0025		0.0015	
	(.007)		(.008)	
Small Bus		-0.0091 *		-0.0104 *
		(.005)		(.006)
Urb-met pop	-0.0009	0.0009	-0.0022	0.0000
	(.002)	(.002)	(.002)	(.002)
Income	-0.0695 ***	-0.1215 ***	-0.0274	-0.0785 ***
	(.025)	(.023)	(.025)	(.021)
Small state	-0.0763	0.0053	-0.0240	0.0866
	(.093)	(.097)	(.097)	0.0937
Med. state	0.0986 *	0.1335 **	0.1500 **	0.2182 ***
	(.059)	(.062)	(.060)	(.058)
Bell At	0.1700	0.3662 ***	0.1808	0.3800 ***
	(.107)	(.110)	(.114)	(.109)
SW Bell	-0.0349	-0.0850	0.0124	-0.0668
	(.112)	(.111)	(.118)	(.109)
Bell South	-0.1207	-0.1817 *	-0.0733	-0.0532
	(.106)	(.106)	(.117)	(.107)
Nynex	0.1284	0.3147 ***	0.1036	0.2509 **
	(.112)	(.114)	(.119)	(.112)
Pac Tel	0.2723 *	0.3639 **	0.3325 **	0.3582 **
	(.151)	(.154)	(.160)	(.151)
US West	-0.0015	0.1454	0.0832	0.2308 **
	(.089)	(.099)	(.094)	(.098)
σ_0^2	0.0211 ***		0.0243 ***	
	(.004)		(.005)	
σ_1^2	0.0231 ***		0.0226 ***	
	(.005)		(.005)	
σ_{01}	0.0180 ***		0.0180 ***	
	(.004)		(.004)	
N	48		48	
Log L	117.2		109.4	

Size of Exchange: 5000				
	1985		1988	
Variable Name	Residential exchanges	Business exchanges	Residental exchanges	Business exchanges
Bypass	-0.1080 *	-0.1913 ***	-0.0459	-0.0821
	(.065)	(.069)	(.068)	(.065)
Intercept	2.1686 ***	3.6242 ***	1.2301	3.1826 ***
	(.658)	(.401)	(.791)	(.385)
Res 80-81	0.4733 ***		0.4180 ***	
	(.092)		(.100)	
Bus 80-81		0.4721 ***		0.3983 ***
		(.081)		(.077)
Pop<17	0.0129		0.0238	
	(.012)		(.016)	
Pop>65	0.0029		0.0082	
	(.011)		(.014)	
Poverty	-0.0026		0.0014	
	(.007)		(.008)	
Small Bus		-0.0064		-0.0073
		(.005)		(.005)
Urb-met pop	-0.0019	-0.0011	-0.0033	-0.0016
	(.002)	(.002)	(.002)	(.002)
Income	-0.0736 ***	-0.1109 ***	-0.0269	-0.0658 ***
	(.025)	(.024)	(.025)	(.020)
Small state	-0.1366	-0.1308	-0.0822	-0.0176
	(.101)	(.110)	(.102)	(.099)
Med. state	0.0465	0.0579	0.1155 *	0.1852 ***
	(.061)	(.066)	(.061)	(.058)
Bell At	0.1831 *	0.3734 ***	0.1863	0.3717 ***
	(.109)	(.116)	(.112)	(.107)
SW Bell	0.0162	-0.0200	0.0464	-0.0061
	(.111)	(.115)	(.115)	(.106)
Bell South	-0.1312	-0.1826	-0.0862	-0.0571
	(.108)	(.111)	(.115)	(.105)
Nynex	0.1754	0.3791 ***	0.1381	0.2893 **
	(.115)	(.122)	(.118)	(.112)
Pac Tel	0.2263	0.1929	0.2165	0.1428
	(.179)	(.192)	(.185)	(.176)
US West	0.0094	0.1253	0.0831	0.1993 **
	(.089)	(.102)	(.091)	(.095)
σ_0^2	0.0220 ***		0.0235 ***	
	(.005)		(.005)	
σ_1^2	0.0258 ***		0.0217 ***	
	(.005)		(.005)	
σ_{01}	0.0202 ***		0.0176 ***	
	(.005)		(.004)	
N	47		47	
Log L	115.1		109.7	

Size of Exchange: 25,000				
	1985		1988	
Variable name	Residential exchanges	Business exchanges	Residential exchanges	Business exchanges
Bypass	-0.0982	-0.1382 *	-0.0427	-0.0493
	(.063)	(.072)	(.064)	(.061)
Intercept	1.9129 ***	2.9975 ***	1.1948	2.6272 ***
	(.606)	(.476)	(.726)	(.425)
Res 80-81	0.5836 ***		0.5220 ***	
	(.108)		(.115)	
Bus 80-81		0.5356 ***		0.4728 ***
		(.107)		(.096)
Pop<17	0.0118		0.0185	
	(.011)		(.014)	
Pop>65	0.0023		0.0048	
	(.011)		(.013)	
Poverty	-0.0045		-0.0001	
	(.006)		(.008)	
Small Bus		0.0007		-0.0028
		(.005)		(.005)
Urb-met pop	-0.0027	-0.0023	-0.0039 **	-0.0029
	(.002)	(.002)	(.002)	(.002)
Income	-0.0608 **	-0.0849 ***	-0.0208	-0.0435 **
	(.024)	(.025)	(.023)	(.019)
Small state	-0.1189	-0.1299	-0.0517	-0.0103
	(.099)	(.115)	(.095)	(.095)
Med. state	0.0498	0.0542	0.1208 **	0.1918 ***
	(.059)	(.069)	(.056)	(.056)
Bell At	0.1290	0.2664 **	0.1503	0.3028 ***
	(.106)	(.122)	(.104)	(.102)
SW Bell	0.0229	-0.0459	0.0706	0.0197
	(.110)	(.122)	(.109)	(.101)
Bell South	-0.1441	-0.1747	-0.0879	-0.0335
	(.105)	(.118)	(.107)	(.101)
Nynex	0.1383	0.3226 **	0.1180	0.2521 **
	(.113)	(.131)	(.110)	(.109)
Pac Tel	0.2021	0.0948	0.2059	0.0851
	(.175)	(.201)	(.172)	(.168)
US West	-0.0093	0.0361	0.0663	0.1325
	(.087)	(.106)	(.085)	(.090)
σ_0^2	0.0210 ***		0.0201 ***	
	(.004)		(.004)	
σ_1^2	0.0284 ***		0.0197 ***	
	(.006)		(.004)	
σ_{01}	0.0212 ***		0.0155 ***	
	(.005)		(.004)	
N	47		47	
Log L	117.4		116.1	

Size of Exchange: 50,000								
	1985				1988			
Variable name	Residential exchanges		Business exchanges		Residential exchanges		Business exchanges	
Bypass	-0.0862		-0.0795		-0.0427		-0.0078	
	(.055)		(.065)		(.062)		(.062)	
Intercept	1.7862	***	3.5958	***	1.6611	**	3.6610	***
	(.581)		(.445)		(.693)		(.409)	
Res 80-81	0.4049	***			0.3110	***		
	(.099)				(.108)			
Bus 80-81			0.3182	***			0.1909	**
			(.101)				(.093)	
Pop<17	0.0261	**			0.0220			
	(.011)				(.014)			
Pop>65	0.0154				0.0071			
	(.010)				(.013)			
Poverty	-0.0094				0.0008			
	(.006)				(.008)			
Small Bus			0.0015				-0.0014	
			(.005)				(.005)	
Urb-met pop	-0.0051	***	-0.0056	***	-0.0060	***	-0.0067	***
	(.002)		(.002)		(.002)		(.002)	
Income	-0.0468	**	-0.0638	***	-0.0186		-0.0377	**
	(.022)		(.023)		(.021)		(.018)	
Small state	-0.1725	**	-0.2495	**	-0.0855		-0.1125	
	(.080)		(.096)		(.082)		(.084)	
Med. state	-0.0136		-0.0440		0.0363		0.0747	
	(.047)		(.057)		(.048)		(.049)	
Bell At	0.1114		0.1628		0.1298		0.2023	**
	(.086)		(.102)		(.090)		(.090)	
SW Bell	-0.0542		-0.1610		-0.0275		-0.1357	
	(.091)		(.103)		(.096)		(.091)	
Bell South	-0.0779		-0.0646		-0.0730		0.0315	
	(.087)		(.100)		(.095)		(.090)	
Nynex	0.1568	*	0.3635	***	0.1170		0.2465	**
	(.091)		(.109)		(.099)		(.102)	
Pac Tel	0.1537		-0.0445		0.1350		-0.0345	
	(.142)		(.168)		(.149)		(.149)	
US West	-0.0015		0.0659		0.0504		0.0864	
	(.071)		(.090)		(.074)		(.081)	
σ_0^2	0.0132	***			0.0145	***		
	(.003)				(.003)			
σ_1^2	0.0190	***			0.0149	***		
	(.004)				(.003)			
σ_{01}	0.0129	***			0.0109	***		
	(.003)				(.003)			
N	44				43			
Log L	122.1				116.0			

Size of Exchange: 100,000								
	1985				1988			
Variable name	Residential exchanges		Business exchanges		Residential exchanges		Business exchanges	
Bypass	-0.0754		0.0076		-0.0264		0.0332	
	(.077)		(.084)		(.076)		(.092)	
Intercept	1.8889	***	3.8229	***	2.0072	***	3.5929	***
	(.549)		(.460)		(.594)		(.514)	
Res 80-81	0.3918	***			0.3154	***		
	(.105)				(.104)			
Bus 80-81			0.3220	***			0.2600	**
			(.093)				(.101)	
Pop<17	0.0257	**						
	(.010)							
Pop>65	0.0149				0.0118			
	(.011)				(.012)			
Poverty	-0.0106	*			0.0045			
	(.006)				(.012)			
Small Bus			-0.0068				-0.0020	
			(.006)				(.007)	
Urb-met pop	-0.0045	**	-0.0059	***	-0.0048	***	-0.0060	***
	(.002)		(.002)		(.002)		(.002)	
Income	-0.0468	**	-0.0555	**	-0.0239		-0.0370	*
	(.023)		(.023)		(.020)		(.021)	
Small state	-0.1532		-0.2257	**	-0.0589		-0.1622	
	(.095)		(.105)		(.086)		(.105)	
Med. state	0.0036		-0.0200		0.0661		0.0453	
	(.053)		(.060)		(.049)		(.061)	
Bell At	0.0296		-0.0251		0.0557		0.0197	
	(.092)		(.104)		(.086)		(.108)	
SW Bell	-0.1208		-0.3083	***	-0.0780		-0.2720	**
	(.104)		(.111)		(.097)		(.115)	
Bell South	-0.1270		-0.1247		-0.0911		-0.0760	
	(.103)		(.113)		(.104)		(.126)	
Nynex	0.0459		0.1594		0.0429		0.1080	
	(.109)		(.122)		(.096)		(.121)	
Pac Tel	0.0985		0.2085		0.1806		0.3074	
	(.172)		(.191)		(.159)		(.192)	
US West	-0.0810		-0.0334		0.0063		-0.0422	
	(.081)		(.101)		(.076)		(.108)	
σ_0^2	0.0131	***			0.0116	***		
	(.003)				(.003)			
σ_1^2	0.0168	***			0.0183	***		
	(.004)				(.004)			
σ_{01}	0.0130	***			0.0117	***		
	(.003)				(.003)			
N	37				37			
Log L	111.5				104.7			

Size of Exchange: 250,000								
	1985				1988			
Variable name	Residential exchanges		Business exchanges		Residential exchanges		Business exchanges	
Bypass	-0.0893		0.0169		-0.1280	*	-0.0060	
	(.059)		(.080)		(.062)		(.084)	
Intercept	2.3860	***	3.5416	***	2.0859	***	3.5672	***
	(.639)		(.594)		(.655)		(.622)	
Res 80-81	0.4444	***			0.3603	***		
	(.095)				(.101)			
Bus 80-81			0.4415	***			0.3468	**
			(.119)				(.126)	
Pop<17	0.0133				0.0178			
	(.011)				(.012)			
Pop>65	-0.0013				0.0082			
	(.010)				(.011)			
Poverty	-0.0104				0.0013			
	(.007)				(.007)			
Small Bus			-0.0127	*			-0.0035	
			(.007)				(.008)	
Urb-met pop	-0.0069	***	-0.0089	***	-0.0082	***	-0.0089	***
	(.002)		(.0021)		(.001)		(.002)	
Income	-0.0334		-0.0282		-0.0183		-0.0213	
	(.023)		(.025)		(.020)		(.021)	
Small state								
Med. state	-0.0774	*	-0.1217	**	-0.0532		-0.0851	
	(.038)		(.053)		(.037)		(.053)	
Bell At	-0.0013		-0.0363		0.0072		-0.0480	
	(.062)		(.090)		(.063)		(.092)	
SW Bell	-0.1661	**	-0.3322	***	-0.1516	*	-0.3514	***
	(.076)		(.104)		(.078)		(.108)	
Bell South	-0.2113	***	-0.1993	*	-0.3218	***	-0.2740	**
	(.075)		(.104)		(.079)		(.113)	
Nynex	0.0921		0.1416		0.0637		0.0770	
	(.125)		(.176)		(.118)		(.167)	
Pac Tel								
US West	-0.0879		-0.0523		-0.0904		-0.1581	
	(.053)		(.095)		(.054)		(.099)	
σ_0^2	0.0050	***			0.0050	***		
	(.002)				(.002)			
σ_1^2	0.0109	***			0.0111	***		
	(.003)				(.003)			
σ_{01}	0.0056	***			0.0050	***		
	(.002)				(.002)			
N	27				27			
Log L	92.4				88.7			

Size of Exchange: 500,000 & 750,000 & 1 million pooled				
	1985		1986	
Variable name	Residential exchanges	Business exchanges	Residential exchanges	Business exchanges
Bypass	-0.2437 **	-0.0108	-0.2007	-0.1225
	(.111)	(.087)	(.124)	(.145)
Byp. increment for 750,000	0.0019	0.0018	0.0656	0.0392
	(.076)	(.083)	(.109)	(.132)
Byp. increment for 1 mil.	0.1252	0.0326	0.1574	0.1084
	(.108)	(.117)	(.152)	(.182)
Intercept	1.5232	3.1442 ***	-4.2911 ***	-0.4414
	(2.255)	(.764)	(1.038)	(.811)
Int. increment for 750,000	-0.0267	-0.0113	-0.0540	-0.0452
	(.057)	(.063)	(.082)	(.099)
Int. increment for 1 mil.	-0.1221	-0.0254	-0.2243 **	-0.1851
	(.076)	(.082)	(.104)	(.125)
Res 80-81	0.8066 ***		1.8435 ***	
	(.195)		(.170)	
Bus 80-81		0.5313 ***		1.3340 ***
		(.164)		(.144)
Pop<17	0.0147		0.0023	
	(.027)		(.014)	
Pop>65	-0.0082		0.0056	
	(.032)		(.013)	
Poverty	-0.0163		0.0946 ***	
	(.030)		(.012)	
Small Bus		-0.0214 ***		-0.0102 **
		(.006)		(.004)
Urb-met pop	-0.0079 ***	-0.0097 ***	-0.0115 ***	-0.0079 **
	(.003)	(.002)	(.003)	(.004)
Income	-0.0080	0.0023	0.1776 ***	0.0442
	(.068)	(.025)	(.039)	(.039)
Med. state	-0.0507	-0.0259	0.0232	-0.0270
	(.071)	(.060)	(.078)	(.093)
Bell At	-0.0229	0.0322	-0.4377 ***	-0.0845
	(.099)	(.093)	(.126)	(.146)
SW Bell	-0.0588	-0.3559 ***	0.0410	-0.1420
	(.150)	(.078)	(.106)	(.113)
Bell South	-0.3369 **	-0.1815 *	-0.8966 ***	-0.3200 *
	(.129)	(.108)	(.155)	(.173)
Nynex	-0.1762	-0.1437	-0.8412 ***	-0.4189
	(.235)	(.176)	(.230)	(.266)
US West	-0.0360	-0.0174	0.1494	-0.0027
	(.078)	(.079)	(.103)	(.121)
σ^2_0	0.0089 ***		0.0183 ***	
	(.003)		(.005)	
σ^2_1	0.0109 ***		0.0265 ***	
	(.003)		(.007)	
σ_{01}	0.0086 ***		0.0217 ***	
	(.002)		(.006)	
R^2	.818	.856	.519	.575
N	36		35	
Log L	123.6		129.9	

BIBLIOGRAPHY

Abler, David G., 1989. "Vote Trading on Farm Legislation in the U.S. House," American Journal of Agricultural Economics, (Vol. 71): 583-591.

Abler, David G., 1991. "Campaign Contributions and House Voting on Sugar and Dairy Legislation," American Journal of Agricultural Economics, (Vol. 73): 11-17.

Bates, Robert H., 1985. "Governments and Agricultural Markets in Africa," in The Role of Markets in the World Food Economy, edited by D. Gale Johnson and G. Edward Schuh. (Boulder, Colorado: Westview Press): 153-183.

Bonnen, James T. and William P. Browne, 1989. "Why is Agricultural Policy So Difficult to Reform?" in The Political Economy of U.S. Agriculture, edited by Carol S. Kramer, (Washington, D.C.: Resources for the Future): 7-33.

Gardner, Bruce L., 1989. "Economic Theory and Farm Politics," American Journal of Agricultural Economics, (Vol. 71): 1165-1171.

Johnson, D. Gale, 1990. World Agriculture in Disarray, (New York: St. Martin's Press): Chapter 1.

Pasour, E.C., Jr., 1988. Agriculture and the State: Market Processes and Bureaucracy, (San Francisco, CA: The Independent Institute): Chapters 4 and 5.

Rapp, David, 1988. How the U.S. Got Into Agriculture and Why it Can't Get Out, (Washington, D.C.: Congressional Quarterly, Inc.): Chapters 1, 4, and 5.

Variyam, J.N., J.L. Jordam & J.E. Epperson, 1990. "Agricultural Policy Preferences," American Journal of Agricultural Economics, (Vol. 72): 257-267.

World Bank, 1986. "Congress in Policy" and "Executive Branch," U.S. Agricultural Policy Guide, (1-36 and 51-81 in 1988 edition.)

Young, Linda, Mary Marchant, and Alex McCalla, 1991. "The Political Economy of Agricultural Trade: A Review of the Literature on Domestic Policy Behavior and International Price Formation," U.S. Department of Agriculture, Economic Research Service, Agriculture and Trade Analysis Division.

WHY EXPERTS ON THE ECONOMICS OF AGRICULTURE HAVE CHANGED THEIR POLICY TUNE*

BY
BRUCE L. GARDNER
ASSISTANT SECRETARY FOR ECONOMICS
U. S. DEPARTMENT OF AGRICULTURE

Agricultural economists have viewed the farm economy in basically the same way throughout the post-World War II period--one of the few truly competitive sectors, subject to inexorable decline in employment, sluggish adjustment in factor markets, and instabilities of prices and incomes--and their normative stance has been fundamentally unchanged also. Yet a significant change in the thrust of commodity policy recommendations has occurred. Until about 1950, the most carefully considered recommendations focused on judicious intervention by government to improve the functioning of markets. By 1990 recommendations for such intervention had virtually ceased. Why?

The most promising hypotheses are:

1. Changes in (our knowledge of) the factual situation in agriculture.
2. Changes in economists' beliefs about the economic behavior of farmers and the commodity markets.
3. Changes in the economic theory used to organize our understanding of the agricultural economy.
4. Changes in our knowledge of governmental action.
5. Changes in the self-interest of the advisors.

---

*Presented at the D. Gale Johnson Festschrift Symposium, Chicago, Illinois, May 3, 1991.

The first four hypotheses involve developments in positive economics. The body of this paper marshals evidence, albeit not rigorously, on each of these hypotheses and on self-interest, as well as on the possibility of normative change that I have already disparaged.

THE EVOLUTION OF ADVICE ON COMMODITY POLICY

The menu of commodity policy recommendations in the United States took shape in the 1920s and has consisted of the same basic items since that time:

1. Laissez faire (no intervention)
2. Supporting prices received by farmers
   a. via governmental willingness to acquire stocks
   b. via production control
   c. via direct payments
3. Attempts to find creative ways to accomplish the desirable consequences of price supports without (as much of) the undesirable effects.

Economists have been split among these alternatives since first they were heard from in policy debate. The lack of convergence to a single position is apparent in the recent survey of members of the American Agricultural Economics Association by Pope and Hallam (1986). Pope and Hallam asked for preferences between laissez faire and existing policies for grains and cotton, milk, and tobacco, and more generally whether laissez faire was preferred to governmental intervention. A slight majority preferred laissez faire for milk, a large majority for

tobacco, with a 50-50 split for grain and cotton. But 58 percent favored government intervention and 38 percent laissez faire on the more general question (Pope and Hallam, Table 1). It is nonetheless striking that economists' writings on agricultural policy have recently been largely unenthusiastic about commodity policies centered on supporting farm prices. In Congressional testimony for the 1985 and 1990 Farm Bills, agricultural economists from outside the government appeared on many occasions, as for previous farm bills. But very few argued for intensifying price support activity and most cautioned against it. This was especially striking in 1985 when the economic situation in agriculture was commonly described as a crisis.

The most favorable views of price support activity were expressed during the agricultural price depression of the years just after World War I. In 1922, the Secretary of Agriculture convened a National Agricultural Conference at which university and government agricultural economists participated. They made many activist recommendations. On the subject of "price adjustment", they stated that "whereas the prices of agricultural products are far below the cost of production, so far below that relatively they are the lowest in the history of our country; therefore, it is the sense of this committee that the Congress and the President of the United States should take such steps as will immediately reestablish a fair exchange value for all farm products with that of other commodities" (Quoted in Taylor 1952, p. 578).

In the period immediately following World War II, there again came forth a spate of expert committee undertakings to provide advice on price policy for agriculture. The views of agricultural economists at this time are well represented in the American Farm Economic Association's Readings on Agricultural Policy (1949). The authors are uniformly skeptical of price supports established above the level at which supply equals demand. However, some intervention is recommended in commodity markets. A typical recommendation is: "in order to give the farmer the orientation and incentive to make shifts in the proper direction and to assure him against drastic declines in the return from specific commodities, the government should announce in advance a support schedule of prices for each agricultural commodity. The support schedule for any year should be fixed within a range of 70-90 percent of the average price of the commodity of the previous 3-5 years, adjusted for changes in the index of prices paid by farmers for articles used in living and production" (Report of the Committee on Parity Concepts, in AFEA, 1949, p. 139).

This recommendation belongs to the same family as the proposal put forth by T. W. Schultz and D. Gale Johnson, spelled out most completely in Johnson's Forward Prices for Agriculture (1947). In that book Johnson recommended a policy of establishing the price for each commodity for the upcoming crop year at the level at which anticipated supply equals demand. If

the market prices fell sufficiently far below the forward price level, then payments would be made to farmers.

Three conditions were listed as necessary for the forward price policy to succeed: First, farm people must be willing to accept direct and indirect subsidies; second, there must be created a public spirited and technically competent administrative agency to determine price expectations and to administer the forward prices; and third, Congress must be willing to formulate workable general principles of price policy and then permit the administrative agency freedom of action (p. ix). Johnson believed that if these conditions were realized "the economic arguments for the system of forward prices, supplemented by compensatory price payments during depression periods, is a convincing one" (p. ix).

Later work by the same author provides an example of a more complete disenchantment with commodity policy, however sophisticated it may attempt to become. In World Agriculture in Disarray (1973) Johnson set forth the following general framework for agricultural policy in an industrial country. First, provide rural farm youth with the same access to education as existing in urban areas; second, a variety of measures to facilitate off-farm mobility and migration; third, income payments to low income farm families who could not expect an improvement in income through migration or job mobility; fourth, "gradual reduction of current high price supports and subsidies until an internationally agreed level of protection for agriculture has been achieved" (p. 222).

Sixteen years later, in the revised version of this book, Johnson lists the same four elements of an appropriate agricultural policy except that the system of payment is to be unrelated to current and prospective agricultural output, and is to be made a transition mechanism for all farms as well as the immobile farms singled out in 1973. Also, the recommendation for reduction in price supports is now stated as follows: "Consistent and gradual reduction of price supports and subsidies that affect output to levels that approach the prices that would prevail under a liberal world trading regime" (p. 12.23).

Similar instances of evolution of professional opinion away from commodity programs are the writings of Willard Cochrane and Luther Tweeten. Cochrane (1985) discusses his own changes of mind from recommending intervention to the view that "we should eliminate the price and income support features of the commodity programs as quickly as possible" (1985, p. 1007).

The case of Luther Tweeten is especially interesting in that he absorbed the tradition of D. Gale Johnson and T. W. Schultz and combined it with the land-grant university orientation of Earl Heady. Tweeten's Foundations of Farm Policy (1971) was the most complete published summing up of the views developed in the 1940's and 1950's, coupled with the rural development and human capital issues of the 1960's. Tweeten is quite circumspect about making recommendations. His summary judgement is "in formulating realistic policies, it is well to recognize that commodity programs do not raise the net income of farm people over the long

run. The principal purposes of commodity programs are to create an orderly economic environment for agriculture and to hold a strategic reserve of farm production capacity. The stability function is so important that a free market is now mostly an academic exercise and is unlikely to become an actual policy for agriculture." (Tweeten, 1971, p. 357-58). This view is very close to that of Johnson (1947), although Johnson would never have used the terms "free market" or "academic exercise" in the disparaging sense that Tweeten did.

Tweeten (1989) was intended originally as a revision of Foundations of Farm Policy but ended up being changed so significantly that the book required a new title (Tweeten, 1989, p. vii). The new book contains a summary paragraph much like the one quoted from Tweeten (1971) with changes as follows. "In formulating realistic policies, it is well to recognize that commodity programs do not raise the net income of farm people over the long run. Farmers have demonstrated that they are capable of adjusting to changing conditions. Adequate size, reasonably well managed family farms have and likely will continue to earn a favorable return on resources. A greater market orientation in farming threatens neither the family farm nor food supplies" (Tweeten, 1989, p. 419). The statement is again circumspect, but the omission of the point that the stability function of the programs dominates free-market considerations is notable.

Notwithstanding the current split in agricultural economists' views that Pope and Hallam found, the positions of Johnson, Cochrane, and Tweeten indicate a general trend from 1947 through 1989--a movement from a position that appropriate commodity price regulation including price supports is called for to a position that such supports should be eschewed. This latter position attained the status of conventional wisdom in the context of the Uruguay Round of GATT negotiations in 1986-90. The positions staked out by the World Bank (1986) and in Sanderson (1990) represent this viewpoint well.

CAUSES OF OPINION CHANGE

What is behind the change of view? To begin with the fifth hypothesis listed earlier, could it be a change in the predominant economic interests of the experts who are making recommendations? As self-proclaimed seekers of objective truths, the very idea may make us uneasy. But it is a question we do not hesitate to ask about other peoples' behavior. Pope and Hallam report evidence that it is indeed the case that agricultural economists' opinions on policy issues are not independent of who they work for. Those employed in the Atlantic region report significantly greater support for existing tobacco policy, and in the Midwest, grain policy, as compared to economists from other regions. In general, "employment matters" (Pope and Hallam, p. 585).

Nonetheless, it is apparent that other forces besides economic interests are at work in observed changes of view. The

profession as a whole has roughly the same set of interests now as 40 years ago; at least, there is no apparent shift of economic interest significant enough to cause a major change of view. And, our paradigm case study of Gale Johnson, and the Willard Cochrane and Luther Tweeten examples as well, involves the evolution of opinion with no evident change of economic interest.[1] Moving beyond self-interest, the large alternatives are normative evolution (of goals or values) versus the accumulation positive economic knowledge.

The authors of the writings on policy reviewed do not evince normative evolution during the post-World War II period. In the case of D. Gale Johnson, the stated goals of policy are very much the same in 1947 and 1989. The goals in 1947 are listed as, first, efficiency including maximizing the total returns from given resources and making provisions for economic growth and, second, income goals including the provision of a minimum level of living for all, mitigating gross inequality in income distribution, and taking steps to enable any particular income group receive a per capita income on a par with any other comparable groups in society (Forward Prices for Agriculture, pp. 18-21). I find no evidence in later writing of any appreciable change in these goals. Luther Tweeten (1971, 1989) exhibits a similar consistency. The general position is that of standard welfare economics. Efficiency and equity are sharply distinguished. While economists cannot prescribe on income distributional matters as objective scientists, writers on farm

commodity policy throughout the post-World War II period take it as a criticism of a commodity program if its benefits go predominantly to higher income as opposed to lower income people.

With respect to positive economics, the situation is different. In the instance of Gale Johnson, it is noteworthy first that there has been a change in the broad conception of how the economy operates in the 1947 as compared to the later writings. Throughout, his general view has been "neo-classical", the view that prices matter, that competition is a desirable and productive state of affairs, and that intervention in markets causes problems. But, as he states in the illuminating introduction to the 1975 Arno reprint of Forward Prices, Johnson, as most economists, was still heavily influenced by the cataclysm of the Great Depression in the immediate post-World War II period. This muted the general confidence in laissez faire that really only became a Chicago hallmark, in the macroeconomic sphere, with Milton Friedman's influence in the 1960s and more broadly in the 1970s and 1980s, as doubt began to grow about the government's competence in and the necessity for macroeconomic steering of the economy.

Placing the prospect of serious recession if not depression in the forefront also had implications for agricultural policies more narrowly. Although agricultural economists more recently have urged the incorporation of macroeconomics into our analysis of the farm economy (e.g., Schuh (1976), Chambers (1984), Rausser (1985)), it was the 1940s that marked the high-water mark of

macroeconomic influence upon agricultural policy analysis. In the 1949 AFEA Readings, of the 15 selections on "Price and Production Adjustment", five give heavy emphasis, even predominance, to macroeconomic considerations (Schultz (1945), (1948); Committee on Parity Concepts (1947); Committee on Postwar Policy of the Association of Land-Grant College and Universities (1944); Nicholls (1945)).[2]

The relevance of macroeconomics to commodity policy turns on instability. Nicholls (1945) begins his discussion of agriculture's problem not with production or export demand instability but with "The Problem. Industrial capitalism is apparently peculiarly subject to progress by fits and starts, reflected in wide fluctuations in the general price level, employment, and national income. Agriculture is particularly vulnerable in such an economy..." (p. 166). The problems due to macroeconomic instability are more severe than sector-generated instability due to foreign or domestic output fluctuations because the macroeconomic downturns generate low prices along with low output and reduced off-farm opportunities, and because the downside of the cycle lasts longer.

Commodity policy contributes to a remedy as follows: "agricultural purchasing power would be maintained by payments (for continued production) equal to the difference between the going market price of their product and some specified percentage of the predepression price" (Nicholls, p. 167). The Committee on Parity Concepts, endorsing a similar concept, suggested 70 to 90

percent of the average price for the previous 3-5 years, indexed to USDA's index of prices paid by farmers (p. 139). In Nicholls' approach the base price would be a three-year pre-depression price. Essentially the same approach was also recommended by Schultz (1945) and Johnson (1947).

In today's terms, the proposed program of the 1940's is most nearly equivalent to the marketing loan program that was enacted for soybeans in 1990. (The proposed program would even more precisely resemble a target price program with a current output payment base, and no ARPs). Applying the proposal to current conditions, the three-year soybean price would be $5.79 for the 1987-89 crops (excluding the 1988 drought year), which adjusted for the prices paid index would give a producer support price of about $6.40 in 1991. This is a higher support price, even without the prices paid adjustment and the high 1988 price, than either the House or Senate Agriculture Committee contemplated in their extravagant period before settling upon the $4.92 level (including a 2 percent "loan origination fee") that was finally enacted for the 1991 crop.

In short, fear of the agricultural consequences of depression led the dominant market-oriented wing of economists in 1945-50 to recommend producer price guarantees more generous than even the prairie populists were recommending in 1990. And while it became apparent in late 1990 that the U.S. economy had entered a recession, arguments linking this event to a need for price support was conspicuous by its absence.

An item of detail that Johnson and other writers in 1945-50 shared was a marked preference for governmental storage or direct payments as opposed to production controls (items 2a or 2c above in preference to 2b). In Forward Prices, Johnson rules out production control "because it is ineffective as a method and because the effects on resource allocation are undesirable" (p. 142). He also notes as an important drawback that "the realm of choice available to the farmer is obviously limited by production control...Production control through allotments, quotas, and rationing could almost completely supplant the individual farmer as an entrepreneur and place almost all entrepreneurial functions in a governmental agency" (p. 251). Following Wallace (1962) the more prevalent view has been that the choice of production control or direct payments is an empirical matter turning on relative elasticities of supply and demand. Although Schuh (1976, for example) led a school of thought that also decried production controls, his position was based on the high elasticity of demand for exported commodities. Johnson's resource allocation argument is different, being based upon underproduction already being a problem in depression periods so that direct payments help correct a pre-existing problem while production controls worsen it--a kind of second-best argument. Again, macroeconomic instability was much more a driving force in the 1940's than it is today.

What caused the change in outlook? The experience summarized in Table 1. Prices received by farmers rose at the

same rate, on average, in recessionary and nonrecessionary years. Farm relative to nonfarm income is actually slightly higher in the years of economic downturn. Nonetheless, real farm income falls more rapidly in these years along with aggregate GNP. These two comparisons are not in conflict. They indicate that both farm and nonfarm income fell in recessionary years, but that the difference between economic performance in recessionary and nonrecessionary years was larger for the nonfarm than the farm sector. These data make it clear that the business cycle has not been a driving force behind fluctuations in farm prosperity in the post-World War II period.

In 1966, Vernon Ruttan stated that "it is no longer reasonable to suggest that the farm problem is 'primarily a product of business fluctuations and unbalanced expansion of the economy' [quoting Schultz, 1945]" (Ruttan, p. 4). In the 1980s those economists who were still emphasizing macroeconomics, e.g., in the Chambers and Rausser papers cited earlier, were concerned not with instability arising from business cycles but rather from monetary policies intended to stabilize the economy but perhaps destabilizing agriculture.

Another area of potential change in positive economics is in economists' perceptions of the basic characteristics of agriculture that Brandow (1972) identifies as defining a classical model of the "farm problem". He traces this model principally to Schultz (1945). The key features of the model are inelastic product supply and demand, rapid technical change,

factor immobility. These features are still presented in textbook discussions as the essence of the agricultural economy. The development of the literature based on this model is too involved a subject to describe here, but I have attempted this task elsewhere (Gardner 1991). My point in that paper is that while evidence confirming the basic sectoral characteristics has accumulated, the model has fallen out of use for its chief original purpose--to explain low incomes in agriculture relative to the nonfarm economy. The model has fallen out of use because farmers' incomes have ceased to be low.

In any case the model of agricultural inelasticity and fixity--apart from its role in explaining low farm incomes--generated no apparent commodity policy implications. The main such implication would have been that price supports would cause only small deadweight losses, compared to intervention in markets with more elastic supply or demand, because quantity responses to price distortions would be small. On the other hand, inelasticity does provide reasons for special attention to price instability in agriculture, since even modest quantity shocks would cause large price and income changes. This creates an economic environment of uncertainty and income instability for farmers with associated economic problems. It is on uncertainty and instability that the analytical basis for commodity policy is focused in Schultz (1945) and Johnson (1947), as well as Tweeten (1971).

On the aspects of the farm economy pertaining to instability and the dynamics of agricultural production and resource use, the prevailing opinion--and, I believe, the opinions of Gale Johnson, Luther Tweeten and others who have moved away from moderately interventionist policy recommendations--has undergone fateful changes. To investigate the changes with some specificity and yet remain within a manageable scope, the discussion is limited to views expressed in Forward Prices for Agriculture.

Here are statements on five topics from that book:

1. "By and large, the market system and the firm have been unable to adapt themselves to output fluctuations arising from natural causes. A major purpose of price policy is to aid in such an adaptation" (p. 35).
2. "Capital rationing leads to too many small-scale farms--farms too small to utilize the labor capacities of the families living on them. Put the other way round, capital rationing is at least in part responsible for too many people being employed in agriculture by affording too large a number of presumed business opportunities" (p. 71).
3. "Under conditions of certainty private storage, if not monopolistically controlled, will lead to exactly the same results as a public storage policy with the objective of maximization of total utility...If expectations are uncertain, there arise significant

differences in storage activities of private individuals and the public agency" (p. 165).

4. "The extent of the effect of mistaken price expectation on resource allocation is influenced by two considerations. First, there is generally a certain amount of contagion in the formation of price expectations...second, where there is no contagion in the formation of price expectations different producers will not have the same price expectation, and some expectations must be in error. It might be assumed that, with so many producers involved, in some average sense the expected price of the group as a whole would be approximately correct. Even if this were true, the allocation of resources would be inferior to the best allocation. Some producers' expected price would be above the true price, and these producers would market too much; other producers' expected price would be less than the true price and their production too small" (p. 46).

5. "When the pressure of fixed obligations upon the farm...becomes rather sharp, there usually occurs a type of disinvestment in soil resources which is manifested by erosion and depletion" (p. 102). "It is apparently true that soil conservation for example has presented a more acute national problem during periods of depression than when farm incomes were high" (p.

215). "If price policy is effective in evening out the cyclical swings of prices...it is likely that a more even income flow which would result from forward prices would mean somewhat less necessity for the exhaustion of human and material resources..." (p. 102).

On each of these topics, research and experience in the economics of agriculture has changed the plausibility of either the views expressed or the policy implications that can be drawn.

With respect to the first topic, output fluctuations and reactions to them, the quoted statement is used to support the general principle, "Pricing policy should be utilized to attain a considerable degree of stability in the output of individual crops and livestock product" (Johnson, 1947, p. 35). On this topic the most notable change in economists' perception is the gradual disappearance of belief in the inherent cyclicality of farm commodity markets. No so long ago discussions of cattle cycles, hog-corn cycles and other "cobweb" cycles pervaded the lore of agricultural economics. Regularity in such cycles assumed an inability of farmers to learn from experience that left open the door for productive governmental intervention. If farmers cannot learn that, say, after two years of high hog prices you get two years of low prices, then there is a role for the government to (1) disseminate this information, and (2) perhaps set a producer price at the mean level. Granted, the policy implications are not strong ones in that (1) price forecasts could be a beneficial policy even without cycles, and

(2) forward prices might be an undesirable policy even with cycles. Nonetheless, it is likely that the decline of belief in cycles in commodity markets has blunted arguments for a governmental role in pricing.[3]

Johnson's discussion of capital rationing is one of the few--and to this day one of the most thoroughly thought out--arguments for the structure of farming, i.e., the mix of small and large farms, being distorted by market failure. But the idea that capital rationing has needlessly prevented farms from reaching their optimal size has never received convincing empirical support. The idea persists in theoretical literature because scholars have found models of asymmetric information to be readily assimilable to the relationship between borrowers and lenders (Innes, 1989). But there is still no body of evidence that this phenomenon helps explain the size of U.S. farms.

The implication of capital rationing for commodity policy is largely a negative one. Capital rationing is a problem that can't be cured by commodity price policy.[4] Capital rationing might more readily argue for intervention in farm credit markets, but this line was not pursued by Johnson, nor by Schultz--who also at this time argued for capital rationing as an important phenomenon. Today, the perceived failure of farm credit markets is that under the protective umbrella of guarantees against bad loans, government-backed lending institutions granted too much, not too little credit in the 1970s and 1980s.

Government's role in commodity storage is one of the few areas where advances in the conceptual framework or theory we have to work with has been influential in changing policy recommendations. This can be seen in the progress from Johnson's discussion of the economics of storage to the work of Gustafson (1958) which was supervised by Johnson and placed his views about the optimality of competitive private storage on a firmer basis, and then to Newbery and Stiglitz (1981) and Runge and Myers (1985) who further formalized both the arguments for and pitfalls in private storage as compared to public.

Gustafson's work showed the generality of the idea that private speculative storage would lead to the same stabilization results as optimal public storage. At the same time, the market conditions which Johnson saw as preventing this private sector optimality from being realized in practice have come to seem less and less pertinent. Measures by the Federal Government to institute storage regimes have been converted too regularly into attempts to redistribute income in favor of commodity suppliers. The programs were also to difficult to manage properly, at least in retrospect (Schnittker, 1973). Thus, these policies became discredited. Consequently, in the 1990 Farm Bill discussion there was virtually no disagreement with the proposal that the Commodity Credit Corporation (CCC) should be given the tools to enable it to support prices without accumulating stocks and that subsidized storage regimes, notably the Farmer Owned Reserve program, should be scaled back and structured so that farmers and

not public agencies would make the decisions about when commodities should be marketed and at what price.

The fourth topic, the consequences of mistaken price expectations, has not played a role in political discussions of commodity policies. Neither has the line of thought that inefficiency results from differing expectations been followed up in the empirical literature of agricultural economists. Several interesting theoretical papers have appeared, notably Holthausen (1979) and Feder, Just and Schmitz, (1980). They showed that rational utility maximizing producers, even if risk averse, will in the presence of futures markets behave as if the futures price is the expected price. They will not react to equate marginal costs to different prices even if they have differing expectations.

The essential economics of this outcome is that the differences in expectations that exist are more efficiently acted upon in the futures markets than by commitment of physical resources to production. If I expect the price of soybeans to be $5.00 while the comparable futures price is $6.00, I should sell forward at $6.00 and deploy my soybean growing resources accordingly; and, I should act on my $5.00 expectation by selling even more than I produce, how much more depending on my risk aversion and my portfolio of other assets. While this idea has never been established as empirically important its reasonableness causes one to question whether differing price

expectations is a good reason for publicly established forward prices.

Even if Johnson's hypothesis that differing price expectations lead to inefficiency were true, the main policy remedy might be the provision of information rather than establishing government determined prices. But, it is this and other points about the farmer's inability to adapt to uncertainty that enable Johnson to place efficiency, in the sense that "the nation should maximize the total return from a given body of resources" (1947, p. 18) among the first of the goals that agricultural price policy can help to achieve.

While instances exist, as just cited, where the conceptual literature subsequent to 1950 has weakened the case for governmental intervention in commodity markets, the general thrust of the theory of policy has not changed. The central idea has always been the theory of second-best, even if not explicitly under that label. Johnson (1947) emphasized that "in agriculture only one of the half dozen assumptions of perfect competition is fulfilled, namely a proliferation of the number of firms. Other assumptions of major significance, particularly mobility of resources, freedom of action, and perfect knowledge and foresight are only crudely approximated." This is taken to imply that "some type of positive action is required to improve the allocation of resources" (1947, p. 7).

In recent years such market failures have been expressed in different theoretical language, notably in models of asymmetric

information and incomplete markets. But while the list of reasons why competitive markets don't have the strong welfare implications that the simplest theory attributes to them has continued to expand, the more recent authors, with ever more rigorous demonstrations that markets fail, nonetheless and perhaps paradoxically couch their conclusions so as to carefully distance themselves from recommendations that government should intervene. Thus, works such as Newbery and Stiglitz (1981) or Innes (1989), while theoretically as supportive of forward prices or similar policies as Johnson (1947), in fact do not make proposals for intervention anywhere near as strong or specific as those Johnson made. The reason, as with the second thoughts about stabilization policy through governmental storage, seems to be a general disenchantment with the ability of government to act as an efficiency-seeking corrector of market failures as opposed to a purely politically driven distributor of incomes.

The fifth quotation, on soil erosion and conservation, is included to remind ourselves that conservation issues are not just a recent concern, and again as an instance of how a changed set of experiences can alter a seemingly common sense observation. While Johnson associated depression and low farm incomes with soil erosion problems, in the 1970's the opposite assertion was commonly made--that high prices led to more intense production methods and thus increased soil conservation problems. I do not know of evidence, or even systematic attempts to find

evidence, that either of these hypotheses is empirically important.

A topic perhaps more important than any of these five is implicit in the political events that have led to economists pulling back from recommendations for intervention in commodity markets; namely, the apparent pursuit of politically motivated income redistribution to the exclusion of any competing purposes of commodity policy. The key issue in positive economics with respect to this topic is the extent to which commodity price supports have been successful in achieving income gains for farmers.

Gale Johnson was an early skeptic on the ability of price supports to improve farm income. Johnson (1976) summarizes his 1947 position as incorporating the view that price policy was not an effective means for reducing the difference between the levels of income of farm as compared to non-farm families. "Low incomes in agriculture, I argued, were due primarily to a lack of resources and thus increasing prices would contribute relatively little to the incomes of the lower income segments of the farm population. Increasing the incomes of the poorest of the farm families required improving the education and health of farm people, and increasing the ease with which farm people could transfer to non-farm jobs." (1976, Introduction, pages unnumbered). As I read the 1947 text, in places it is more optimistic about the income-increasing possibilities. Johnson states: "The possibility exists that a forward price system could

be used to augment the incomes of some or most farmers. This could be done by a continuous and conscious establishment of forward prices above the estimated equilibrium level and the accumulation of storage stocks. The accumulation of storage stocks would lead largely to a change in the time pattern of income, though, if subsidies were used to dispose of the stocks (as was the case with wheat and is currently being done with cotton), a net increase in income would result." (1947, pp. 251-52). In World Agriculture in Disarray, Johnson is firmer in rejecting the possibilities of income increases through price supports. It is true, however, that the main discussion of price policy and farm income, in Chapter 8 of Forward Prices is skeptical about low income farm people being aided by commodity policy.

It now appears to be the generally accepted view that commodity price supports can increase the rental value and asset value of land or policy determined rights to produce commodities. But beyond increasing the rents received by owners of these assets, the policies will not affect farm income significantly in the long run.

This absence of income effect does not occur in the models used by government agencies such as the Economic Research Service (ERS) of the USDA, or the Congressional Budget Office, or outside analysts such as the Food and Agricultural Policy Institute (FAPRI). The models used in these analyses find quite important net income effects from reductions in target prices, for example.

But, these simulations are not backed by direct econometric evidence that farm income has been affected by price support levels, even in the short run. Of course, identification of the effects of price supports on farm income is hindered by the fact that the redistribution to farmers is largest precisely when prices and hence farm incomes are lowest, so that a simple correlation between farm income and program income support activity tends to show a negative effect.

Nonetheless, I take the conjunction of economic reasoning and simulation in the writings of Gale Johnson and his followers on the one hand, and the supply-demand simulations on the other, as indicating that price support policies do in fact increase farmers' incomes as long as these programs are in effect, in even the longest run. But the income gains take the form of increased rental returns on land used to produce the supported commodity.

The Role of Theory. Since 1950, economics generally, as well as agricultural economics, has become more theoretically precise and quantitatively oriented. Have there been policy consequences? One area mentioned earlier where the conceptual apparatus available has improved substantially in the theory of optimal commodity storage. On the general policy issue of the optimality of private storage, subsequent developments have strengthened the case argued by Johnson that private storage is socially optimal and extended it to the circumstances of uncertainty. Johnson saw market failure here where more recent writers do not. But more recent work (e.g., Newbery and

Stiglitz) cites its own theoretically more precisely delineated market failure.

More important in a practical policy context were the advances by Gustafson in using dynamic programming methods to specify optimal storage policies for a given set of circumstances. But no storage regime based on these methods has ever been put in place. The experience of government failure to establish anything resembling an optimal public storage regime has carried more weight in determining economists' views.

The theory of behavior under uncertainty generally, and of optimal insurance, futures and cash contracting in particular, has advanced substantially also. The general policy consequence, I think, is to cause economists to be less inclined than Johnson was in 1947 to see uncertainty as a source of market failure calling for governmental intervention. Market failures still appear in plenty in the writings of agricultural economists, but they are today more subtle and less amenable to translation into a commodity policy remedy.

Overall, the immediate postwar positive writings in commodity policy would be found wanting in rigor and econometric evidence for the conclusion reached by a reviewer today. But the added rigor of today's writing has not added to their policy relevance--more likely the opposite--and empirical evidence bearing on the choice of optimal commodity policy has improved quite little in view of the size of the investment made in agricultural economics as a branch of scientific inquiry.

28

CONCLUSIONS

Returning to the four hypotheses listed at the beginning of this paper, the conclusions I would tentatively draw are as follows.

With respect to the first hypothesis that changes in the factual situation in agriculture have been important in changes of policy recommendations, the most important changed perception is that farmers are no longer seen as economically hard-pressed. This does not bear on the market failures that economists have both in the 1940's and more recently cited as reasons for intervention. But, what is increasingly clear is that the economic precariousness of farming has been an important reason, indeed a more important reason than the existence of market failures, for agricultural economists' sympathy with government intervention. With the disappearance of belief in this precariousness, support for intervention has waned.

With respect to changes in our understanding of the behavior of farmers and markets, the most important changes have been with respect to the inherent tendency of the overall economy to generate recurrent periods of deep recession and with respect to the inherent tendency of commodity markets to generate cycles fueled by farmers' own imperfect responses to uncertainty.

Changes in economic theory are less important even though some definite improvements have been made. We have sharpened up our conception of what sorts of market failure make most sense as justifying governmental intervention in commodity markets. But

there has been no change in the profession's theory of these failures to account for recommendations of less intervention now than 40 years ago.

Finally, and I believe most importantly, there is the hypothesis of changed perception of governmental action. This has been the most sobering set of observations. The policies that agricultural economists earlier recommended have not been adopted! Instead of doing a tune-up of price supports so that their economic advantages to society can be realized without the distortions that unwise policies can cause, the tendency has been to continue the distortions and to ignore possible improvements from an economy-wide viewpoint. Only when new political interests came into the picture, notably environmental interests and the pressures associated with a general necessity to cut government spending, has there been a significant policy change. I have found some evidence that government is attempting to waste as little as possible of the economic pie in its attempts to slice and reslice it. Unfortunately, given that government is going to carry out significant redistribution of income toward the farm sector, there does not seem to be any feasible way of accomplishing this without wasting a substantial amount of the economy's real income. It is this realization more than any other, I believe, that leads economists today to recommend, when asked, that government desist from pie slicing.

ENDNOTES

1. Willard Cochrane was a government employee in a politically sensitive post when some of his interventionist opinions were unleashed. But those opinions predated this employment.
2. The participation of committees of economists in policy debate is notable following World War II as it was in the 1920s. The Association of Land-Grant Colleges and Universities now eschews commodity policy recommendations and limits political activities to keeping the Universities' feet firmly in the Federal trough. (Similarly ad hoc groups of agricultural economists formerly developed and defended positions on broad policy matters but now limit their joint exertions principally to the pursuit of funding.)
3. It is not clear that Johnson, in the statements quoted, has commodity cycles in mind as prime instances of farmers' inability to adapt themselves to output fluctuations--one of the few places where his point is not clear.
4. Johnson (1947, p. 102) does argue that capital rationing, along with farmers' weak financial condition, is among the reasons why farmers fail to follow the speculative motive for storage in periods of low prices sufficiently to provide significant price stabilization. Capital rationing thus supports Johnson's recommendation of governmental commodity storage activity.

REFERENCES

American Farm Economics Association, Readings on Agricultural Policy, Philadelphia: The Blakiston Company, 1949.

Chambers, Robert G. "Agricultural and Financial Market Interdependence in the Short Run", American Journal of Agricultural Economics, 66 (Feb. 1984): pp. 12-24.

Cochrane, Willard W. "The Need to Rethink Agricultural Policy in General and to Perform Some Radical Surgery on Commodity Programs in Particular", American Journal of Agricultural Economics, 67 (Dec. 1985): 1002-09.

Committee on Parity Concepts, "Outline of a Price Policy for American Agriculture for the Post-War Period", AFEA (1949), pp. 124-140.

Committee on Post-War Agricultural Policy of the Association of Land Grant Colleges and Universities, "Agricultural Prices", in AFEA (1949), pp. 141-148.

Feder, G., Just, R.E., and Schmitz, A. "Futures Markets and the Theory of the Firm Under Price Uncertainty", Quarterly Journal of Economics, 94 (1980): 317-28.

Gardner, Bruce L. "Changing Economic Perspectives on the Farm Problem", Journal of Economic Literature, forthcoming, 1991.

Gustafson, Robert L. "Carryover Levels for Grains", U.S. Department of Agriculture, Technical Bulletin No. 1158, 1958.

Holthausen, Duncan M. "Hedging and the Competition Firm Under Price Uncertainty", American Economics Review, 69 (1979): 989-95.

Innes, Robert.

Johnson, D. Gale. Forward Prices for Agriculture, Chicago, IL: University of Chicago Press, 1947 (reprinted with a new introduction, New York: Arnold Press, 1976).

_____. World Agricultural in Disarray, London:Fontana/Collins, 1973. (draft revision, 1989).

Newbery, D.M.G., and Stiglitz, J.E. The Theory of Commodity Price Stabilization, Oxford: Oxford University Press, 1981.

Nichols, William H. "A Price Policy for Agriculture, Consistent with Economic Progress, That Will Promote a More Stable Income from Farming", in AFEA (1949), pp. 165-173.

Pope, Rulon D. and Hallam, Arne. "A Confusion of Agricultural Economists? - A Professional Interest Survey and Essay", American Journal of Agricultural Economics 68:3, August 1986, (pp. 572-94).

Rausser, Gordon C. "Macroeconomics and U.S. Agricultural Policy" in B. Gardner, ed., U.S. Agricultural Policy, Washington, DC, American Enterprise Institute, 1985, pp. 207-52.

Runge, C.F., and R. J. Myers. "Shifting Foundations of Agricultural Policy Analysis: Welfare Economics When Risk Markets are Incomplete", American Journal of Agricultural Economics, 67 (Dec. 1985), pp. 1010-16.

Ruttan, Vernon. "Agricultural Policy in Affluent Society", Journal of Farm Economics, 48 (Dec. 1966):_____

Sanderson, Fred H., ed., Agricultural Protectionism in the Industralized World, Washington, DC: Resources for the Future, 1990.

Schnittker, John.

Schuh, G. Edward. "The New Macroeconomics of Agriculture", American Journal of Agricultural Economics, 58 (1976):802-11.

Schultz, Theodore W. "Needed! A Production and Marketing Price Policy", in AFEA (1949), pp. 120-123.

Schultz, Theodore W. "Elements of a Price Policy for Agriculture", in AFEA (1949), pp. 174-188.

Schultz, Theodore W. Agriculture in an Unstable Economy, New York: McGraw-Hill Company (1945).

Taylor, Henry C. and Taylor, Ann D. The Story of Agricultural Economics in the United States, 1840-1932, Ames, Iowa: The Iowa State College Press, 1952.

Tweeten, Luther. Foundation of Farm Policy, Lincoln: University of Nebraska Press, 1971.

_____. Farm Policy Analysis, Boulder, Colorado:Westview Press, 1989.

The World Bank. World Development Report, 1986, New York:Oxford University Press, 1986.

Economic Influences on the History of Agricultural Policy

Peter H. Lindert

Working Paper Series
No. 58

Note: The Working Papers of the Agricultural History Center, University of California, Davis, are preliminary materials circulated to invite discussion and critical comment. These papers may be freely circulated, but to protect their tentative character they are not to be quoted without permission of the author.

A revised version of this paper will be published by Cornell University Press in a volume on *Agriculture and Economic Development in the 1980s: Lessons for the Role of Government*, edited by C. Peter Timmer. This research was supported by California Agricultural Experiment Station Project #CA-D-AHC-4573.

November 1989

Agricultural History Center
University of California
Davis, California

"ECONOMIC INFLUENCES ON THE HISTORY OF AGRICULTURAL POLICY"

Peter H. Lindert
University of California - Davis

Table of Contents

ECONOMIC INFLUENCES ON THE HISTORY OF AGRICULTURAL POLICY*

Peter H. Lindert
University of California - Davis

I. INTRODUCTION.

The governments of today's industrial countries distort their economies by protecting their farmers. Third World governments distort their economies even more by taxing their farmers heavily. Why? We lack an answer because the question is relatively new, and because the answer will require the cooperation of several disciplines, not just economics.

Most well-known empirical tendencies in economics are based on long historical experience or on common folklore. Most are sanctified with quotes showing that they were foreseen by Adam Smith back in 1776. Not so with the developmental pattern of agricultural policy. It has emerged only in the 1970s on the basis of postwar data. (Little, Scitovsky, and Scott 1970; Balassa 1971; Johnson 1973; Schultz 1978; Lutz and Scandizzo 1980; Bale and Lutz 1981; Binswanger and Scandizzo 1983). The World Bank, the FAO and other international agencies have played a leading role, as providers of data, and as sponsors and publicists for much of the relevant research (e.g. World Bank 1982, 1986, and the forthcoming team effort edited by Krueger, Schiff and Valdés). A pathbreaking pair of books by Anderson and Hayami (1986) and Tyers and Anderson (forthcoming) has extended the pattern, to other commodities and back into the 1950s for fifteen leading countries.

Two patterns have been noticed: the developmental pattern (the more advanced the nation, the more its government favors agriculture) and an anti-trade pattern, namely that governments tend to tax exportable-good agriculture and protect import-competing agriculture. The developmental pattern implies that successfully developing nations will drift toward subsidizing agriculture, though it

cannot yet tell us whether the most advanced nations will go on protecting their farmers when they have become more prosperous and less agricultural than ever before. The anti-trade bias implies that policy will dampen comparative advantages the world over, cutting dependence on, and the gains from, international trade in agricultural products.

The history and theory of today's international patterns in agricultural policy still elude us.[1] Did these patterns prevail anytime before World War II, or are they postwar ephemera? Why should there be consistent patterns at all, among nations so varied in their politics and geography?

This paper takes steps toward both a history and a theory of the developmental and anti-trade patterns in agricultural policy. Both patterns existed before World War II, but have been hidden by two clouds: English exceptionalism and the Great Depression of the 1930s. Basic economic influences can help explain these patterns, but have been hidden by two other clouds: the belief that each nation's history defies global influences, and the willingness to use common arguments for agricultural protection as if they explained the observed policy patterns.

Part II re-sketches the recent global picture. The available measures are known to contain many flaws. Yet the flaws tend to *understate* policy contrasts between more- and less-developed nations, and between agricultural-exporting and agricultural-importing nations.

Part III follows the history of government treatment of the agricultural sector backward in time. Today's patterns can be traced back to the 1860s, the Great Depression notwithstanding. Different patterns stand out in the history of the less industrial, more impoverished world before the 1860s. Central government (usually, the throne) often squeezed agriculture, especially in food crises, when it sought to force delivery of affordable food to the cities. That is, there was an urban bias like that observed in the Third World today. Agriculture fared better at the hands of governments dominated by landed elites. Ironically, this group included the government of the First Industrial Nation.

Part IV clears the way for a deeper theory of the agricultural policy pattern by

showing that traditional rationales for agricultural protection fail as positive theories of the policy patterns we observe. The goal of food security, concern over the instability of the agricultural sector, sympathy for farmers as poor people, and political nostalgia for the farm sector all fail to explain who taxes agriculture and who supports it, even though these are elements in the story.

Parts V and VI apply a frugal model of economic influences on political struggles between pressure groups. The model underlines two basic reasons why agricultural policy should shift from taxing to subsidizing agriculture:

(1) The shrinkage of the agricultural population raises the per capita subsidy and lowers the per capita cost of a given real transfer.

(2) The relative size and strength of the political alliances defending agriculture do not shrink nearly as fast as the sector itself, because the barriers to political participation by farmers and their closest allies decline faster than the barriers to participation by others. (A similar point was made by Mancur Olson (1985).) Exceptions such as Corn-Law England (1660-1846) help to prove the rule, by being as peculiar in their political structure as in their policies.

In addition, the model brings out two ways in which the switch in agricultural policy should have been more dramatic, over the course of economic development, than changes in the policy toward any other major sector:

(3) The decline in population share, cited in argument (1), is historically more dramatic for agriculture than for other sectors.

(4) Own-price elasticities of supply and (especially) demand are lower for agricultural products than for other sectors of comparable size and trade orientation. Lower price elasticities mean lower net deadweight costs of marginal redistribution through government. In industrialized countries, governments tend to cut deadweight costs further by choosing among types of farm support so as to minimize the relevant elasticities and costs (Gardner 1987, 1988). Government intervention is not resisted as strongly when the main redistribution relates to agricultural markets, both in the early settings in which the debate was usually over how much to tax

agriculture and in the industrialized nations debating how much to subsidize it.

Part VI applies the pressure-group reasoning to some issues and supplements it for others. The two most important supplementary arguments are (1) that there is a secular rise in the farm sector's income sensitivity to agricultural price changes that the government could control, and (2) that the government's own revenue demands help explain the anti-trade bias of policy, especially in newly independent nations.

Part VII performs regression tests of competing explanations for the observed policy patterns, using international data from 1980. Part VIII summarizes.

II. TODAY'S INTERNATIONAL PATTERNS.

A. The NPC studies: Fine Results from Crude Measures.

To compare policies over dozens of countries, recent comparative studies had to do without relevant details and concentrate on a single statistic available from every country. That statistic is the ratio of the domestic producer price to the border price, or "world price," of the same agricultural product, sometimes with adjustments for official exchange-rate overvaluation. If NPC is above unity, farmers are said to be protected at the rate NPC-1. If it falls below unity, they are said to be taxed at the rate 1 - NPC.

Table 1 and Figure 1 show how NPC's from around 1980 suggested both the trade bias and the developmental pattern of agricultural policy. Three important products have been separated from export crops, to illustrate the trade bias. NPC's for exportable crops are less than one, meaning that the domestic producer price is below the world price, usually because of official export restrictions. Producers of importables, by contrast, enjoy higher NPC's for any given level of GNP per capita, as illustrated in the cases of wheat, rice and beef. Even without developing a model of the policy process, one can think of a proximate explanation of the trade bias: governments find it easier to protect producers of importables and tax producers of

exportables because both policies can generate tax revenue instead of requiring subsidies.

The development pattern of agricultural policy also emerges in Table 1 and Figure 1. For each product, there was a rough upward slope relating net protection to GNP per capita around 1980. In this international cross-section, the slope is due almost entirely to the discrete jump in NPC's across the income gap separating developing from developed countries. The raw correlation between NPC and GNP per capita is insignificant among developing countries. It is also lacking among developed countries, partly because two high-income nations, Canada and the United States, were relatively unprotective agricultural-exporting nations in 1980.[2] But there does seem to be a difference between the two groups of countries. No industrial country has resisted the pressure to subsidize its producers of importable goods, though the North American and Australasian governments have kept fairly neutral toward their exportable agricultural products.

There is also a third pattern, a food-commodity pattern hardly mentioned by the literature. Among the 15 industrial nations studied by Anderson and Hayami, the general ranking by NPC for 1980 is:

milk > beef > sugar ≥ grains > (chicken, eggs and pork)

The ranking for 1980 had some glaring exceptions: eggs received heavy price protection in Australia and New Zealand, and beef was much more heavily protected in the European Community than elsewhere. Yet there is a pattern of sorts, running from dairy products down to pork, chicken and eggs. Fruits and vegetables are hard to categorize, but were generally less protected than grains. (Anderson and Hayami 1986, Tyers and Anderson forthcoming). What might explain this general protection ranking , and the exceptions to it?

B. What Better Measures Would Show.

Nominal protection coefficients fall far short of quantifying the impact of all

government policy on the relative income of the agricultural sector. For a start, they omit any taxes or subsidies that do not affect the farmgate price of the product. The seriousness of this omission is clear enough in the case of a small (price-taking) free-trading country, which cannot affect the world price. The small country's government could lay heavy subsidies or heavy taxes on farmers, without budging the NPC from unity. Two micro-measures that avoid part of this defect are the "effective rate of protection" (ERP) widely used in the 1960s and 1970s, and the "producer subsidy equivalent" (PSE) developed by economists at the OECD and the US Department of Agriculture (OECD 1987, USDA 1988).[3] The ERP captures the protective (or taxing) effect of unit subsidies (or taxes) on internationally-traded inputs into the relevant output sector. The PSE may miss some of the subsidies and taxes from trade policies relating to farm inputs, but it has the advantage of quantifying some of those other payments and taxes that do not affect traded-good prices. In principle, the PSE can be refined into a fairly comprehensive measure of the effects of all policies on the rate of producer surplus.

Two other key shortcomings of the NPC also afflict the currently available ERP and PSE measures. One is that these are all measures of effects on nominal incomes within the (farm) enterprise, ignoring any effects of government policies on the cost of living, and thus the real value of those nominal incomes. Obviously, taxes and tariffs (or subsidies) on consumer products imply a direct markdown (markup) in farmers' real incomes. This effect should have been incorporated into all three proximate-protection measures. The other shortcoming is the failure to include systemic (e.g. general-equilibrium) influences of all policies on the market prices of inputs and outputs. Trade policy, for example, affects wage rates, interest rates, rents and exchange rates, and other real prices that are taken as given in the NPC, ERP and PSE measures.

The shortcomings of NPC's as measures of governmental impact on the real purchasing power of farm income recipients cannot be corrected here. We can, however, be fairly confident that the NPC's point in the right direction and probably *understate* the contrast between developing countries' taxation and industrial countries' protection of agriculture. First, the elaborate calculation of PSE's still

leaves the same commodities in the taxed or protected camps in the countries studied (USDA 1988). Second, the ERP studies have shown that the developing countries, while nominally taxing agriculture, are heavily protecting their manufacturing sectors, imposing much greater taxes on agriculture through their purchases of farm inputs and consumer goods. Estimates from six developing countries in the 1960s show that their NPC's understate the absolute and, especially, the relative taxation of agriculture.[4] For 1975-84, Krueger, Schiff and Valdés (1988, pp. 262-3) find that industrial protection and exchange-rate overvaluation reduce the effective protection of agriculture by 11 percent for export crops and 20 percent for importable crops in eighteen developing countries. In addition, the available estimates captured only indirect commodity taxation, missing the frequent tendency of direct taxes to fall more heavily on agriculture in developing countries.

The NPC measures are thus likely to show the correct direction of government agricultural distortions in the postwar era, with some understatement of the net taxation of agricultural products in developing countries. When we turn to history below, the NPC's are likely to be freer of bias, because they capture the effects of an import tariff, the policy tool that dominated before 1930.

III. DO THE PATTERNS HAVE A HISTORY?

Does history confirm the development pattern and the anti-trade bias in policies affecting agriculture? Did today's developed countries tax agriculture more heavily in the past, when their economies were poorer and more agricultural? Did they favor import-competing producers of agricultural products relative to producers of exportables? What were the earlier policies of governments of countries still less developed today?

A. Industrial-Country Trends since 1860.

To sharpen our focus on longer-run national trends, we must first adjust our view for the distorting effect of short-run inflations and deflations. Sudden

inflations slash nominal protection rates and deflations raise them. The effect is largely automatic. Many trade barriers are specific duties, denominated in units of currency per physical unit, so that a jump (drop) in market price lowers (raises) the ad valorem rate before officials have time to react to the new price levels. Yet the effect of inflations and deflations is also partly behavioral. Governments reacted to the great slump of 1929-33 with panicked protectionism, the classic illustration being America's Hawley-Smoot tariff of 1930. These cyclical influences must be factored into any test for long-run trends. It will be possible to argue that there has been a sustained trend toward agricultural protectionism, even though the highest rates of all are not today's rates, but those in the depressed 1930s, followed by those around 1972, just before the inflationary oil shocks.

Japan's history clearly conformed to the developmental pattern. Her tax and trade policies shifted from taxing to subsidizing agriculture over the century since the Meiji Restoration. Direct taxation discriminated against agriculture under the Tokugawa shogunate up to 1868. The Meiji restoration continued to tax agriculture more heavily than industry,[5] but the net tax rate dwindled until the tax rates on agriculture and non-agriculture were roughly equal, and were both low, in the 1930s.

Japan began protecting her domestic rice farmers against import competition as early as 1904. That was a year in which she first became a steady net importer of rice, needed revenues for the war against Russia, and was no longer bound by the free-trade clauses of her earlier treaties with the great powers. Rice protection and imperial self-sufficiency grew until rice farmers received a net protection of 84 percent by 1938, as shown in Figure 2.[6] Postwar Japan has continued the trend toward higher protection, of course, and has reached heights of agricultural protection unmatched by other industrial market economies except Korea and Switzerland (Anderson and Hayami 1986, USDA 1988, Tyers and Anderson forthcoming).

Within the postwar era, independent Korea and Taiwan have compressed the same policy revolution, switching from policies that depressed prices for domestic farmers in the 1950s to heavy protection in the 1980s (Anderson and Hayami 1986, pp. 17-38., USDA 1988).

The United States conformed to the developmental pattern, as well as the usual trade bias, though only if we start its history from the Civil War. Temperate zone agriculture has kept a comparative advantage in the U.S. trade pattern, with cotton and grains serving as traditional exports. American trade policy treated the exportable crops badly before the 1930s. Farmers got little or no aid, and had to face tariffs on imports, mostly industrial, ranging from 20 to 45 percent for the half-century from the Civil War to World War I and still around 15 percent in the 1920s (U.S. Tariff Commission 1934; Benedict 1953; U.S. Census Bureau 1976, Series U211). Significant income support did not arrive until the New Deal. Today's domestic supports remain generous by older American standards, though they are lower here than in Europe and Japan (Anderson and Hayami 1986, p. 26; Petit 1985, pp. 38-56; USDA 1988). Thus for the United States, the drift from taxation toward subsidization since the Civil War took a particular form: little trade protection throughout, declining tariffs on importable industrial goods, and rising direct subsidy payments after 1933.

Like the United States, Britain fits the developmental pattern of agricultural policy as long as we start from the 1860s. The starting point in this case is one of virtual free trade and low tax rates that were nearly neutral toward agriculture. There Britain remained, essentially, until she turned protectionist in 1932.[7] For the rest of the 1930s, there was little shift toward net protection of agriculture, however, because imports from the Commonwealth were not discouraged and because of the significant duties on steel, autos and other industrial goods. Heavy agricultural protection came to Britain mainly after World War II. It peaked around 1972, when she had begun to conform to the Common Agricultural Policy of the EC, imposing higher duties on her traditional agricultural imports from the Commonwealth. The inflationary oil shock of 1973-74 lowered ad valorem rates in Britain as elsewhere, but by 1980 the 1972 peak protection had been regained.

Prussia and unified Germany also followed the development pattern from the 1860s on, as illustrated by her rates of protection for wheat and other foodstuffs in Figures 3 and 4 and Table 2. Germany became an early international leader in the protection of the grain sector with Bismarck's famous tariff package, the "compact of

iron and rye" in 1879. Across the 1880s pressure to protect agriculture, and with it essential German values and national security, continued to build. After a freer-trade interlude under Caprivi, in 1902 the intense pressure from the Junkers and the *Bund der Landwirte* had produced grain duties matching those on protected industries, plus virtual prohibition of foreign livestock and meat.

Protection for industrialists reduced the effective rate of protection for agriculture in Imperial Germany. In 1913, for example, tariffs of about 21 percent for foodstuffs were partially offset by tariffs averaging 10-15 percent for industrial goods (Table 2). Granted, German policy did not bend the overall terms of trade against agriculture, nor did direct taxation discriminate against farmers. In this respect, it differed from the prevalent urban bias of policy in today's Third World. Still, imperial Germany did not favor agriculture as much as its successors.

The protectionist trend resumed under the Weimar republic in the 1920s, and was accelerated by the Nazi march to autarky (Gerschenkron 1943; Kindleberger 1951; Gourevitch 1977; Tracy 1982, pp. 87-110, 193-216). Postwar support for West German farmers was less extreme than under the Third Reich, but remained stronger than the support from the prewar empire or the Weimar republic, peaking around 1965 and declining moderately to 1980 (Figure 4), before rising again.

France also tended to follow the developmental pattern since 1860, though the movement from taxing to subsidizing was unsteady. Starting from virtual free trade in the 1860s, agricultural protection rose after 1881, but these movements were partly offset by contemporaneous increases in industrial tariffs. Even the Méline tariff of 1892, sometimes cited as a triumph for agriculture, was part of a larger package of duties that actually gave almost as much nominal protection to industry (Golub 1944; Augé-Laribé 1950; Tracy 1982, pp. 59-86, 173-92). This tension between moderate agricultural and industrial tariffs remained as late as 1927 (Figures 3 and 4, Table 2). In the 1930s France shared in the global retreat to autarky and self-sufficiency. Postwar French support for agriculture resembles that of West Germany: less autarkic than in the 1930s, but more protective of agriculture than in the 1920s or earlier (Delorme and André 1983, pp. 303-29), with a drop in protection in the 1970s and a rise in the 1980s.

Similar trends were followed by the policies of other European governments for which we have data back to 1913. As shown in Figure 4, postwar farmers received more protection in Italy, Sweden and especially Switzerland than did their prewar and interwar predecessors. So did postwar Austrian farmers (Gulbrandson and Lindbeck 1973, p. 38, Liepmann 1938).

B. What Policy History for the Third World?

If agricultural protection grew in each of the now-industrialized nations since the 1860s, was there a similar trend in the long-independent countries of the Third World? Did they tax agriculture even more heavily in the past? We need to know the origins of their taxes on agriculture in order to judge the prospects for lowering those taxes. Historical data are particularly scarce on this issue, and only the most tentative hunches should be ventured until a careful comparative history can be compiled.

For the postwar era, we have only slight hints about trends in the relative taxation of agriculture in Third World Countries. Over periods of ten years or longer starting from the 1960s, Brazil, Colombia, Korea, Mexico, the Philippines and Taiwan have all shifted to lower taxation or positive subsidization of their farm sectors. From the 1970s to 1981-3, a group of thirteen African nations switched to net subsidization of cereals, with no net trend in the taxation of their export crops (World Bank 1986, pp. 62, 68). On the other hand, preliminary results from eighteen developing countries (Krueger, Schiff, and Valdés 1988, pp. 262-3) show no change in protection rates from 1975-79 to 1980-84. It is still premature to say there was a postwar trend toward lighter taxation of agriculture in the Third World.

Could Third World governments have been drifting toward lighter taxation over a whole century or longer, marching in parallel with the industrial countries? There are limits to this possibility, because there are limits to how severe the net taxation of agriculture could have been in the past. Even before the limits to taxpayer endurance come the limits to tax-collecting competence in less developed settings. Underdevelopment means, among other things, that the government

cannot raise more than a tiny share of national product in tax revenues. Surely the history waiting to be written in this area is a history of how the state gained taxation powers and how they happened to be aimed primarily at agriculture.

The implausibility of a centuries-long drift toward ever lighter taxation of agriculture is illustrated by two sets of data from before World War II. First we have Liepmann's international cross-sections of tariff rates from the first third of this century, shown in Table 2. If the policy trend were monotonically tied to development (say, real GNP per capita) with similar time slopes in every country, we should be able to replicate the recent (1980) cross-sectional development pattern in this earlier data set. Yet any examination of Table 2's tariff rates and income levels in Europe shows the opposite of the postwar correlation between development and protection of agriculture. On the contrary, the higher the level of GNP per capita in 1913 or 1929, the *lower* the nominal protection of agriculture. It seems that the lower-income countries of Eastern and Southern Europe gave heavy protection to both agriculture and manufacturing. An open research question is why the taxed groups in those countries were the export and non-traded-good sectors, rather than farmers as such.

The second set of data from before World War II reaches beyond Europe to the developing countries of Latin America, Asia and Australasia. To compare these developing countries' trade policies with those of Europe and North America, we can combine Liepmann-like international cross-sectional studies of tariff rates (Fetter and Chalmers 1924; Great Britain, Committee on Trade and Industry, 1927; Wright 1935) with good national studies (e.g. Brigden et al. 1929, Diaz Alejandro 1970, Ingram 1971) and long time-series on overall-average tariff rates in many countries (e.g. Mitchell 1983, Wilson 19__).

The miscellaneous historical data sets from the Third World seem to say that the anti-trade pattern was always strong, while the drift toward protecting agriculture was weak or non-existent until the nations reached the threshold of industrial-country status. Before World War II, less developed countries were generally exporters of agricultural products. They followed the anti-trade pattern faithfully, protecting nascent industries with tariff rates of 15-35 percent, while

denying such support to farmers or mineral exporters. Such was apparently the policy of Argentina, Australia, Brazil, Colombia, Mexico, Peru, Thailand, Uruguay, and Venezuela as well as the higher-income agricultural exporters of North America, both around 1913 and in the 1920s. Only a few import-threatened food sectors got help in such less-developed settings (as in the grain sectors of Japan, Mexico and New Zealand in the 1920s). The anti-trade pattern among less developed countries remained fixed and trendless from the earliest nineteenth-century data to the present (except in countries like Egypt, Meiji Japan or Thailand, where great-power pressure delayed the rise of protectionism). There was a counter-cyclical tendency to stiffen the anti-trade policies during the depression of the 1930s and to weaken them during the inflation of the 1970s, but no trend toward protecting agriculture.

C. Europe before 1860.

Indirect clues on the origins of the Third World's willingness to tax agriculture can come from the history of an earlier, less developed Europe. Recall that as of 1860, the only widely-used policy implements were trade policy and direct (e.g. land) taxes, and that these were used least of all in that relative heyday of laissez faire. Any excursion into the more distant European past thus starts from relatively sector-neutral policies around 1860.

Given the limits on government's ability to raise revenues and administer complex programs in earlier settings, we expect to find little systematic policy, except where the state could have a major impact at low fiscal expense. Indeed, we find few outright subsidy expenditures in the earlier history, or anytime before the 1930s, when the arsenal of peacetime sectoral policies suddenly expanded.

For most European nations, as for late Imperial China, the main policy stance in early modern times was one of "provisioning," in which the state was prepared to compel affordable deliveries of necessities, especially of food and especially to the cities (Heckscher 1934, vol. II, pp. 80-111; Kaplan 1976; Outhwaite 1981; Fogel 1989, pp. 28-35). The instruments chosen involved little budgetary outlay, often resulting in

weak enforcement. During grain shortages the throne would issue edicts demanding that grain be delivered at prices at or under a decreed maximum. Sometimes local officials were compelled (at little or no royal expense) to inventory private grain reserves for possible state seizure. Exports were often banned. We may suspect that the provisioning policy worsened the supply in years of relative abundance, in its attempt to improve the supply during times of shortage. The clear intent, in any case, was to make the food supply system pay the price of food insurance for its customers. In this sense, it was a policy that taxed agriculture (and merchants), albeit less harshly than many of the agricultural policies in today's Third World.

While the provisioning policy plagued the agricultural interest in most countries and city-states of early Modern Europe, it was not followed in all cases. The Scandinavian countries and Prussia maintained more protectionist policies, mainly in the form of duties on imported grain. Denmark led the way to grain protectionism in the seventeenth century, followed by the others in the eighteenth (Heckscher 1934, vol. II, pp. 92-93).

The trade disruption of the French Wars (1793-1815) brought a widespread by temporary protectionism to the Continent, until the international movement for free trade gathered momentum and triumphed in the 1860s (Kindleberger 1975). Sliding-scale duties on imported foods were tried in France, Bavaria, Belgium and the Netherlands, while fixed duties prevailed in Austria, Portugal, Scandinavia and the Kingdom of the Two Sicilies, and Spain prohibited most food imports altogether (MacGregor 1850). In general, agriculture got less protection than industry, and its protection was confined to import protection without direct subsidies for production or exports. In this era those most exploited by policy were producers of non-tradables, rather than the agricultural interest as such. There may not have been much change in the net effective protection of agricultural incomes.

By far the most important exception to provisioning, however, was England under the Corn Laws (1660-1846). As Bates (1988) has rightly argued, England in this era was so exceptional that her dominance in the historical view of the English-speaking world has distorted our understanding of long-run trends in

policies affecting agriculture. To put the English exception into perspective, we need to note its timing and determinants.

Tudor and early Stuart monarchs strove to follow the provisioning policy, especially in times of crisis. For the period 1600-1640 their efforts seem to have succeeded in the sense of noticeably reducing the deviations in grain prices from trend (Fogel 1989, p. 34). The regulatory machinery was dismantled under Cromwell. After 1660, the Restoration not only failed to re-assemble the provisioning machinery, but took a very different tack. The provisioning policy had been aimed at eliminating high grain prices during dearth, with little intended effect on prices during normal years. The new Corn Laws, by contrast, were aimed at *raising* grain prices in normal years, turning passive and permissive only in the years of greatest dearth. To this end, the government paid export bounties in non-crisis years for the century when England remained a net grain exporter as often as not (1660-1765). Import duties were also decreed. The import duties were frequently adjusted, especially after their relevance was enhanced by England's permanent shift to being a food importer from 1765 on. What evolved was a system of sliding scales, in which the import duty was set higher, the lower the latest domestic grain prices.

The price-raising effect of the Corn Laws increased over the net-export century 1660-1765, as suggested by Table 3's comparisons of English and Continental wheat prices. After 1765, being a net grain importer allowed England to protect her agricultural interest and raise revenue at the same time, by taxing imports. As Table 3 and Figure 5 show in complementary ways, England's price-propping was effective in the 1765-1846 era, and the import duties seem to be the reason why. The policy remained flexible, however. In an extreme food supply crisis, starvation and rebellion had to be avoided. Thus in the French War era (1793-1815) and in other years of extreme hardship, the Corn Laws were either suspended or implemented with cautious partiality (Barnes 1930, Fay 1932). We are left with a set of questions about the great Corn Law exception: Why was England different? Why was the dear-food policy ascendent from the Restoration to the early nineteenth century? Why relent during crises? And are the answer to these questions easily aligned with

the seeming causes of the 1846 Repeal?

D. Summary of Historical Patterns.

History thus sharpens our perspective on today's comparative pattern of agricultural policies in different ways. The developmental pattern and the anti-trade bias are extended back to the 1860s.[8] Yet they show a different profile in history than in today's global view. Figure 6 underlines the difference by drawing the two patterns schematically. Figure 6a sketches today's general international tendencies in terms of an effective rate of protection, one imagined to capture all general-equilibrium effects. The switch from taxing to subsidizing is common to agricultural importables and exportables, though the former are always less taxed or more protected.

Figure 6b shows the somewhat different story from the history of policies since the 1860s. The right-hand pair of curves traces the rise of agricultural protection in industrial countries, with a lurching detour during the Great Depression. The industrial countries' early taxation of agriculture, however, was less severe in the 1860s or earlier than in today's Third World. The two lower-left-hand curves in Figure 6b show that long-independent developing countries, especially in Latin America, were consistent in their relative taxation of agriculture. They did not drift toward protecting farmers until they became either net agricultural importers or truly industrialized nations. The average experience of still-developing countries shown in Figure 6b combines this static Latin-American pattern with a drift toward taxing agriculture a bit more heavily in a few agricultural-exporting countries (e.g. Egypt, Thailand) whose ability to favor industry was checked until this century by pressure from the great powers.

How do we explain the similarities and differences between history and today's global pattern? How do we deal with exceptions like Corn Law England? For perspectives on both the general patterns and their exceptions, we turn first to some familiar explanations and then to some new ones.

IV. POPULAR EXPLANATIONS OF AGRICULTURAL PROTECTION.

A typical reaction to the developmental pattern of agricultural policy is that the basic explanation is "easy" and "well known." There are in fact several explanations that are well known, commonly invoked, -- and inappropriate. To clear the way for deeper insight into agricultural policy, one should start by noting their common shortcoming: they are all theories of just the urge to protect, and cannot account for the equal or greater prevalence of discriminatory taxation of agricultural producers, or for the switch from taxes to protection. Each well-known answer addresses only a narrower question in an ad hoc way.

A. Food Security.

The most common rationale for farm income supports and food import barriers is the need to assure a safe food supply for future crises. Food security rhetoric was prominent inthe building of the strongest farm supports, especially in Japan and Switzerland, two nations scarred by memories of shortages during World War II.

For all its seeming power, the food security argument misses the mark as an explanation of the developmental, anti-trade, and commodity patterns of agricultural policy. For it to help explain the developmental pattern, one would have to show that high-income countries are more worried about food security than low-income countries, and that food security is more of a concern today, when the food sector is protected, than in the famine-ridden past, when it was taxed. Both propositions are false.

Among the high-income countries, food security concerns simply cannot explain the degree and pattern of agricultural protection we observe. Take the case of Japanese rice, supposedly the focus of food-security fears exacerbated in World War II. Since 1968 the government of Japan has been disposing of surplus rice beyond what the nation wants to stockpile for emergency reserves. Japan has even become a net *exporter* of rice since 1977. In addition, the government has been

paying farmers to take land *out of* rice production since 1969, especially in 1971-73, around 1980 and since 1987. Sometimes the intent has been to encourage shifts to other food crops, but sometimes the subsidy allows shifts to non-food crops and to fallow (Otsuka and Hayami 1985; Australian Bureau of Agricultural and Resource Economics 1989, pp. 101-20). Nor is Japanese rice the only case of a glaring departure from food security goals. Canada, the United States and the EC also subsidize acreage reductions and exports on food crops. EC dairy products, like Japanese rice, were converted from importables to exportables by generous producer subsidies. The extreme example of the same conversion is that of wheat in oil-rich Saudi Arabia: since 1986 the Saudis have subsidized the *export* of surplus home-grown wheat they do not want to store. While food-security rhetoric continues, it is cheap. The types and levels of farm subsidies among the high-income countries reveal that farm income support, not food security, is the main motive.

In the Third World, the set of governments paying regular lip service to the food security goal far exceeds the number that protect the food sector. Out of 25 nation/food-product combinations for which nominal protection coefficients were measured around 1980 (Binswanger and Scandizzo, 1983), only four were cases of genuine protection (NPC above 1.20). Seven others were nearly neutral (NPC between 0.80 and 1.20), and fourteen taxed food producers severely (NPC < 0.80).

One likely reason why food security concerns correlate so poorly with actual policies is the economic point that food security does not call for protecting growers. What it calls for is maximizing crisis supply at the lowest cost. The right form of food insurance depends on the type of crisis that is most likely. The four main possibilities are:

(a) temporary bad harvests,
(b) prolonged bad harvests (e.g. Sahelian drought),
(c) temporary siege or embargo by a hostile power, and
(d) prolonged siege or embargo by a hostile power.

The first two cases are ones in which growing food yields low returns. Whatever

causes the bad harvests is also likely to make subsidies to growers look unpromising relative to the obvious option of stockpiling food at non-crisis prices. While the biblical advice of Joseph to the pharoah is as valuable as ever, protection against imports does little to offset bad harvests. If the concern is hostile interruption of food supplies from other countries, the prescription depends on the likely length of the interruption. If the interruption is not likely to extend beyond the next full crop season or animal-breeding cycle (as in (c)), what the nation must have ready is a stockpile of food, not a stockpile of farmers. Protecting farmers during non-crisis is a food security policy only for the contingency of a prolonged interruption of import supply (case (d)). This case fits few of the high-income countries that are doing the protecting. It also requires a strained argument about supply dynamics: only years of non-crisis subsidy can build up a food-growing capacity to be mobilized in the crisis, and neither years of stockpile management nor a strong financial position can do the job.[9]

B. Agricultural Price Instability.

The primary sector is afflicted with particularly unstable prices, and perhaps also with particularly unstable producer surplus.[10] Policies to raise farm incomes are often defended as insurance against the problem of price instability.

Yet the policy and the problem are a mismatch in three ways. First, they are a conceptual mismatch, because instability calls for stabilization policies, not a perennial income redistribution -- an attack on the statistical second moment, not on the first. Another conceptual problem is that stabilizing prices may destabilize producers' surplus or may lower it, depending on the nature of the market disturbances.

Second, they are also an historical mismatch. To match the problem and the demand for support policies over time, one should show rising instability of producer surplus (or at least domestic prices) in less regulated settings, followed by lower instability after support policies were enacted. This is not the trend pattern (Scandizzo and Diakosawas 1987, pp. 58-103, 164-5). Only the Great Depression of the

1930s links instability and public demand for support policies. It was, incidentally, an era in which government commodity-market interventions themselves may have have worsened the price instability in international markets.

Finally, the problem of instability lacks a strong raw correlation with farm support policies when we look across the spectrum of commodities. The most protected commodities are not the ones with the greatest price deviations from trend since 1900. Among those with the worst instability in the terms of trade, only sugar, highly protected and highly unstable in price, favors the instability explanation of support policies. Even in the sugar case, it is not clear that we can rule out reverse causation, from the policy to the instability, especially since generous sugar protection started well before 1930, and got more generous with rising price instability over time. Furthermore, among the primary products with the worst price instability were three tropicals that are now heavily *taxed:* rubber, cocoa and coffee (Scandizzo and Diakosawas 1987, p. 71). While state intervention does lower the price variability of these crops somewhat (Krueger, Schiff, and Valdés 1988, p. 365), this stabilization seems less central to farmer's wellbeing than the fact that they are often denied half their returns. One must ask whether protecting farmers against insecurity is the prime policy motive in such cases.

The price instability argument needs to be more carefully specified and tested. It will be re-packaged as part of the "rising income sensitivity to price" theme in Part VI, and tested in Part VII.

C. Agricultural Protection as Poor Relief.

A common intuition, though one seldom made explicit by scholars, is that agriculture gets special policy protection because farmers would have income far below the average if they were not supported by government. It is plausible to view the rise of generous farm supports as part of the broader twentieth-century mandate to fight poverty. The idea is not refuted by the fact that median income is as high for U.S. farm households as for other U.S. households, and almost as high for Japanese farm earners as for all Japanese earners (Hayami 1988, pp. 20-1, 92-3). It could still be

true that the average farm earner would be poorer than the average earner if the generous postwar supports were removed. Nor can the political *perception* of this point be dismissed, even though the actual benefits of farm supports eventually accrue to the wealthier owners of farm land, not to the average farmer.

The main stumbling block in the path of the poor-relief interpretation of the developmental pattern in agricultural policy is the fact that higher-income settings are not the ones in which farmers have their lowest earning potential aside from government support. The best crude indicator of the part of farm earning power that is less directly tied to government subsidy is gross product per employed person in agriculture, compared with gross product per employed person outside of agriculture.[11] This indicator did not fall over time or development, during the shift from agricultural taxation to agricultural protection. In today's international cross-section, this relative labor productivity of agriculture is slightly higher in the most industrialized, and farm-protecting, countries, failing to show their greater need to offset agricultural poverty. The long sweep of history is similar: the relative labor productivity of agriculture did not decline in the United Kingdom since 1856 (Feinstein 1972), or in the United States since 1900 (*Historical Statistics*), or in Japan since the Meiji reign (Hayami 1988, pp. 20-1, 92-3).

Despite the clear link between farm poverty and the demand for support policies in the depths of the Great Depression, there is reason to doubt that trends in the threat of farm poverty explain trends in policy. Indeed there is an equally plausible counter-argument suggesting that farm poverty should weaken political support for farmers. The greater is their "social distance," and specifically their income distance, from the rest of the population, the less sympathy there may be for subsidizing farmers on self-insurance grounds ("I could end up in their shoes"), as suggested in Part IV. Part VII will test the relative-poverty argument statistically.

D. Nostalgia.

Modern nonfarm voters respect the farm life enough to pay for its support through government (though not enough to live on a farm). While the nature of

the respect varies, there is no denying that distant glow.[12] Yet it too fails to help explain when and how government support for agriculture rose. Reverence for the agricultural life has such a long rhetorical history that the sentiment must have been widespread centuries before the onset of farm protection and subsidies. One could argue that it lacked political power until the nostalgic urban classes reached critical size. The size argument, however, clashes with the consensus that smaller group size is a political advantage. Also, if the spread of nostalgic pro-farm sentiment is correlated with economic development and with protective policies, the best bet is that it is an endogenous by-product of the development process, and not an independent explanatory force. If so, scholars should develop testable theories of the determinants of that sentiment, so that we can predict when it is strong and when it is weak.

V. A SIMPLE PRESSURE-GROUP MODEL.

To bring order to the economic influences on agricultural policy, let us consider a modest and frugal theoretical framework. It should be modest because the political process is complex. It should be frugal to concentrate on those few basic forces that transcend that complexity. Yet a framework there should be, to keep our imaginations logical.

The best starting point is basic intuition. We know that in the long run groups' economic interests are likely to be reflected dimly in the political arena. Most of us imagine the following:

> (1) The political process involves fights over government actions that would redistribute economic wellbeing. The fights arise whenever one side sees enough chance of gain to justify fighting over redistribution.
>
> (2) To participate in the political process requires resource expenditures. The expenditures consist of time costs (campaigning, demonstrating, lobbying, holding office), commodity costs (spending on political campaigns), physical dangers (as in illegal demonstrations), emotional costs and the opportunity

cost of using up one's voting rights.

(3) People care about their own material wellbeing, but also about the wellbeing of others, especially others they identify with. They do not necessarily "vote their own pocketbooks."

(4) Over the long run, what people expect to get out of a redistributive outcome is positively related to what they actually get if it happens.

(5) Redistribution through government has economic side-effects, because economic incentives are altered (e.g. imports of foreign goods are affected by a tariff).

To build an economic theory of long-run patterns over a wide range of political settings, democracies and dictatorships alike, one could stick to these intuitions and not try to adorn them with more specific assumptions. That cautious strategy is followed here.[13]

To analyse economic forces that transcend specific political contexts, let us choose a model that assumes little about the political process. Let us imagine only that government reacts to competing pressures by deciding how much to redistribute from a taxed group (a T group consisting of N_T individuals) to a subsidized group (an S group consisting of N_S individuals). The amount of revenue taken from the taxed group and given to the subsidized is G. On the average, each subsidized individual gets G/N_S from the redistribution and the average taxed individual pays G/N_T. The amount redistributed depends on the resources spent by the two camps fighting over the redistribution. Extra expenditures by the "for" camp raise G, though with diminishing returns. Extra expenditures against the redistribution lower G, but again with diminishing impact.

Government redistributions usually have side-effects on both affected groups, because they change incentives. Consider, for example, raising the income tax to cover government losses on price-support wheat purchases from farmers. The taxed group loses more than just the income tax it pays. If its productive effort is discouraged, there is a further loss because effort is diverted to less valuable

activities. The subsidized wheat farmers may gain more or less than the payments they receive from the government. They would gain more to the extent that the government program raises the market price of wheat sold to private buyers, but they would gain less to the extent that growing more wheat has a rising marginal cost, taking away part of the price markup. Meanwhile, both groups would be hurt to the extent that they are consumers of food or taxpayers sharing in the government's losses on surpluses destroyed or sold abroad at a lower world price.

Since the side-effects are crucial to the arguments that follow, it is useful to label them separately from the redistribution itself. Define D_S and D_T as the side costs (or deadweight costs) felt by the taxed and subsidized groups. These depend on the amount redistributed, with derivatives D'_S and D'_T. The subsidized group's total gain from the redistribution is thus $G - D_S$, where D_S could be positive or negative or zero. The burden on the taxed group is $G + D_T$, where D_T is probably positive. Together the side-effects $D_S + D_T$ add up to the national deadweight loss from the redistribution.

The side-cost magnitudes depend on the costs of administering the transfer itself. They also depend on the incentive-related costs of the redistribution. The more the redistribution affects real quantities exchanged, the greater its deadweight cost. The degree of incentive-related cost depends critically on the price elasticities of the activities being taxed or subsidized.[14] Just which elasticities matter most depends on the choice of policy instrument. To illustrate, here are the elasticities that are crucial in determining the rates of deadweight cost for some common agricultural policy tools:[15]

Policy instrument	In a small open economy, deadweight loss from extra redistribution, $D'_S + D'_T$, is (especially) raised by a higher absolute price elasticity of:
Production tax, subsidy or deficiency payment	domestic supply
Consumption tax or subsidy	domestic demand

Import tariff, subsidy or quota	import demand (and therefore domestic demand and supply)
Export tax, subsidy or quota	export supply (and therefore domestic demand and supply)
Input (e.g. acreage) subsidy, tax, or quota	domestic factor supply and demand (and therefore product demand)

The basic formula is that higher elasticities mean higher deadweight costs. Part VI will apply this point to explaining policy differences between sectors and between agricultural products.

Each economic agent cares about what the redistribution does to the two groups. In a simple linear case, we imagine that the gain for the subsidized group (e.g. farmers) raises her utility by $b \cdot (G - D_S)$, and the burden on the taxed group lowers her utility by $c \cdot (G + D_T)$, where b and c are positive caring coefficients. Common intuition suggests a systematic pattern in how individual agents care about effects on others. A "social distance" hypothesis presented elsewhere (Kristov, Lindert, and McClelland 1989) argues that people care more about the fortunes of persons who are like them in some social attribute (income, ethnicity, gender, region, tastes, etc.). The more a group resembles the individual in relevant attributes, the more the individual cares about the material wellbeing of the group. Thus greater proximity to the groups that would be subsidized raises b, and greater proximity to the groups to be taxed raises c. A narrower variant of the social distance hypothesis also accords with a common conjecture: in the self-insurance variant, individuals care about the effects on an affected group only because they could end up in that group. On this interpretation, which has been applied to redistribution between income classes (Lindert 1989, pp. 38-43), the coefficients b and c are just the individual's probabilities of ending up in the subsidized and taxed groups.

Individuals interact in a political struggle by joining the camps that seem to promise them the greater utility. Some might join neither camp, because the initial fixed cost of joining exceeds the expected benefits from either outcome. Agents can be inactive either because they are truly indifferent (e.g. their b and c are zero) or

because they are denied voice by a repressive system that inflicts harsh penalties on them if they speak out. Allowing for such inactive groups, and for groups that are not affected, defines nine groups in any fight over a redistribution:

	Affected-group status: To be subsidized (S)	Unaffected (U)	To be taxed (T)
Political camp			
Actively in favor (F)	group SF	group UF	group TF
Inactive (I)	group SI	group UI	group TI
Actively against (A)	group SA	group UA	group TA

If large parts of the population are kept politically inactive, e.g. by force, the F and A camps may spend much of their political resources fighting over how to exploit an inactive TI group. Examples would be the fiscal treatment of foreigners in most countries, or of blacks in South Africa, or of the majority of the population in an elitist dictatorship. In such a setting the dominant political solution is likely to resemble the behavior of a price discriminating monopolist, exploiting those low-elasticity groups who cannot exit more than the high-elasticity groups who can (Kristov, Lindert and McClelland 1989; Lindert 1989, pp. 24-5, 27-31). If, on the other hand, all groups have voice, the tax burden will fall on a group that was represented but lost the fight.

Each political fight pits the aggregate expenditures of those favoring the redistribution (the F camp) against the expenditures of those against it (the A camp). Each side's expenditures are subject to eventually diminishing returns. The outcome, an amount of redistribution G, depends on three kinds of parameters, in ways summarized in Table 4. It depends on how those side costs (D_S and D_T) vary with G: the greater the cost to either side, the lower the amount of redistribution, since raising D_S weakens the enthusiasm of those in favor, and raising D_T stiffens the opposition. It depends on the strength of caring (b and c). And it depends on group sizes. In general, the smaller the group, for given caring intensities (b and c), the more effective its lobby, since a small group can get a large enough per-member gain to mobilize every member while a large group suffers from poor organization

and free-riding (as in Olson 1965, 1985).

These predictions are tentative and dependent on a number of conditions not fully identified here. Nonetheless, they have at least the breadth we would seek when trying to explain patterns that roughly fit so many nations' experiences. They deserve to be applied and tested.

VI. INTERPRETATIONS.

A. Sector Size and the Developmental Pattern.

The front-running hypothesis for explaining how farmers evolved from beasts of burden to favored pets is the "small is powerful" argument. Its logic is compelling. Part of the logic is that appeasing a smaller group is cheaper. When farmers were half the voting population, a $100 pure transfer to each farmer would cost the average non-farmer $100, and there would be strong resistance. When farmers had become only 5 percent of the population, giving $100 to each farmer cost the average non-farmer only $5.26 (= 5/95). The rest of the logic is that mobilizing a smaller group is also cheaper. As long as there are fixed costs to becoming politically active (and the distributions of individual stakes are similarly shaped in the two opposing camps), the larger group will suffer more free-riding and inactivity (Olson 1965, 1985). Such simple arithmetic cannot be the whole story, but it must have played at least some role in every country's drift toward farm subsidies.

To give the sector-size argument its due, the first step is to keep it from explaining too much. By itself it implies that the optimal size of a political lobby is just one person. Everbody implicitly rejects such an implication, preferring instead to believe that there is a small optimal lobby size -- from six to eight percent of the population, according to regression results for fifteen industrial countries (Anderson and Hayami 1986, p. 44). But what is it that makes very small lobbies weak? Surely, beyond some point extra smallness of the group costs it more in sympathy from others (lowers their b coefficient) than it gains from being cheaper to mobilize and to appease. Beyond some point of shrinkage of N_S, that is, the caring coefficients (b's)

must drop by faster percentages than G/N_S rises. A key empirical question opened up is whether the peaking out of protection rates in the most industrialized countries in the 1970s was due to agriculture's having dropped below optimal lobby size as Anderson, Hayami and Honmo seem to imply. It could be that the limits of powerful smallness have been reached. Or it could just be that macroeconomic instability cut the protection temporarily in the 1970s, in view of the rise of protection in the 1980s. So far, none of the available regression tests is robust enough for a clear answer, and a new test will be tried in Part VII.

The other initial correction to be made in the argument that small is powerful is to note that it gives us a slope without an intercept in the relationship of agricultural protection to development. It is persuasive as a reason why a country shifts *toward* protection as it develops, but it does not predict when or whether the country crosses the switch-point of zero net taxation of agriculture.

The sector size argument is able to fend off another challenge, however. One might ask how we know that sympathy for the group (the b coefficients again) has not dropped off as fast as the sector shrinks, cancelling any advantage of sectoral shrinkage from the start. There are counter-tendencies raising sympathies for farmers until very late in the history of industrializing countries. As Mancur Olson (1985) has argued, the farm population joins the political arena more effectively as its membership converges toward the rest of the population in literacy and conquers the communications costs of being spatially dispersed. Thus they do not shrink nearly so fast as a share of potential political participants as they shrink as a group to be taxed or subsidized.[16]

Thus there are two reasons why a smaller farm sector gains power, and one reason why the sympathy for them does not fall as fast among political participants as the shrinkage of their numbers might suggest. The smaller farm sector is cheaper to appease, and cheaper to mobilize. Voter sympathy for farmers does not decline as fast as their numbers decline because they converge toward the rest of the population in communication ability (and, one might add, the ability to invest in political action instead of just concentrating on staying alive).

B. Rising Income Sensitivity to Prices.

Farmers' voice may also have been raised by a long-term upward trend in their sensitivity to shifts in prices, shifts that government could control. Here, I suspect, lies a central explanation to the developmental drift from taxation to protection. At first glance, the phrase "rising income sensitivity to prices" might seem to be no more than a return to the traditional price instability concern, on which Part IV.B cast some initial doubts. Yet farmers' income sensitivity relates to the political process through more links than just the instability of prices in the outside world. Their incentive to have the government intervene weighs the cost of political action against the benefits from government help. The benefits side depends on each term in this identity:

farmers' % income sensitivity to government price intervention	=	shares of farm full income from selling each i^{th} factor service	x	magnification effect of commodity prices on farm factor income	x	% commodity price change change caused by government intervention
$\hat{y}$	=	$\Sigma_i \quad \theta_i$	x	μ_i	x	$\hat{p}$

The traditional argument about price instability simply asserts that variation in the free-market price dictates the $\hat{p}$ that farmers think government could bring about by shielding them from the market. This is a plausible argument (subject to the caveats in Part IV.B), and it will receive some support from regressions in Part VII. It is only the third link in the logical chain, however. Let us turn to the first two links, the historical drift in the θ and μ multipliers.

Price sensitivity has been raised by a simple basic fact of economic development: farm products decline as a share of farmers' own consumption. In early settings, the farm sector consumed a large share, say half, of its own product, including housing (e.g. Gregory King (1695) in Laslett 1973, pp. 65, 210; Brady 1972, p. 80). A ten percent drop in the price of all farm products would lower the farm sector's real income by only five percent. Today, by contrast, the farm sector sells 90 percent or more of its gross income in exchange for non-farm products (i.e. θ rose to

0.90), so that a ten percent drop in the price of all farm products would lower the sector's real income by nine percent. The primary source of the rise in the "openness" of the farm economy is, of course, Engel's Law that the share spent on non-food rises with the level of income. Another source is the decline in transport and communications costs, which integrates individual farms into the market economy.

The rise in market exposure (θ_i's) can be quantified at the level of individual farms for the twentieth century, with only rough suggestions of an earlier rise. In the United States, the marketed share of gross farm product had risen to 78 percent by 1910, as shown in Figure 7. An earlier increase in exposure to the vagaries of the market is believed to have been a key to the puzzle of U.S. farm protest in the midst of favorable farm-income trends in the late nineteenth century (Mayhew 1972, McGuire 1981). Within the twentieth century, Figure 7 shows a further rise in the share marketed by individual farm households, to around 90 percent of household gross income.[17]

The second rising-sensitivity argument relates to the μ_i, the multipliers that relate commodity-price changes to farm factor incomes. Could it be that the farm sector's μ's rise with economic development, while the effects of the same commodity prices on the non-farm sector become smaller? Kym Anderson (1987) found this effect to be very strong in a quantified general-equilibrium with three sectors and four factors, only one of which was mobile between sectors. The more advanced the economy, he argued, the greater the percentage impact of the terms of trade on farm land returns and the smaller its percentage impact on non-farm factor incomes. Yet Anderson's result is not tested, and not confirmed by other plausible models. It remains a tantalizing possibility awaiting a test.[18]

Given that the world prices against which the government might shield farmers have not clearly risen in volatility, as noted in Part IV.B, we are left with only one clear reason for predicting that economic development raises farmers' stake in a given percentage of price protection: the argument that θ rises with development. Even by itself, however, this argument seems to have power.

C. The Role of Government Revenue Demands.

The anti-trade bias, and possibly the developmental pattern, of agricultural policy must owe a large debt to the fiscal demands of the state. The state's relentless search for sources of revenue makes at least three contributions to understanding the historical and global patterns of agricultural policy.

In any country, the easiest way to explain the anti-trade bias is to acknowledge that the state is a peculiarly strong-voiced special interest lobby. There is always at least some pressure to raise revenue by taxing imports or exports or both.[19] While the degree of pressure to tax trade can vary with the strength of the protected and taxed groups, and the strength of the national security argument, the bedrock explanation for the tendency to favor producers of importables over producers of exportables is the state's revenue interest. Indeed this point is quite general, not specific to agriculture.

A second role of the state's demands is the "infant-government" effect. In the least developed settings, a nascent state apparatus must raise revenue where it can. It will concentrate on those tax bases that are large, easily monitored, and politically voiceless. In such settings, agriculture qualifies in all respects. It generates a large share of taxable product. Its products are visible, especially in the case of exports, which (like extractive mineral exports) need only be monitored and taxed in a few key ports. And in most cases the agricultural population is, again, poorly mobilized for lobbying. Thus the wllingness to impose heavy burdens on agriculture may have been partly a by-product of early state-building, first in early modern Europe and as late as the 1970s in the Third World.[20]

Finally, the role of state revenue demand *may* be on the rise again. The twentieth-century rise of government debt may reach the limits of creditworthiness even in the most advanced industrial nations. The rapid pace of deficit finance set in the 1980s may have hastened the day of a general budgetary crisis that cuts deeply into industrial-nation subsidies to agriculture. In that eventuality, the anti-trade bias of agricultural policy will become stronger. Farm supports will be concentrated

on revenue-enhancing import protection, and aid to food consumers will be concentrated on export taxation. Among industrial-country farmers, the Japanese will be the most securely protected, even though the need to shift to non-expenditure modes of support will limit their effective protection.

D. Agriculture versus Other Sectors.

Why is the switch to protection more pronounced for agriculture than for other declining sectors? Why were other primary sectors (mining, forests, fisheries) not taxed so heavily when national income per capita was low, or subsidized so much today? Or why not textiles and footwear, two classic declining industrial sectors?

Part of the explanation is easy. Whatever validity the small-is-powerful argument may have, it should have had more effect in agriculture than in other sectors. The reason is that agriculture shrank faster as a share of the electorate, the labor force and GNP than other sectors, thanks largely to Engel's Law. Even if policies were equally responsive to changes in the size of any sector, they would have changed more in agriculture, which shrank more.

An elasticities idea from Part V also seems to help here. As befits the empirical pattern to be explained, the idea is broad. It also has the ability to help explain both the heavier taxation of agriculture in early history and its outstanding subsidies later. And it is testable.

Once other forces, such as sector shrinkage, tip the political balance from taxing a sector to subsidizing it, the resulting swing in policy will be more pronounced the lower are the relevant long-run price elasticities. Lower price elasticities mean lower deadweight costs from each extra dollar of revenue redistributed from one group to another. Lower deadweight costs per dollar redistributed tip the political scales in favor of extra redistribution. Abstract as the idea may sound, there is a mechanism for its affecting political debate. Opponents of a redistribution will be more ardent, and more persuasive to others, the greater the perceived waste and distortion caused by the redistribution. Correspondingly, those

in favor of the redistribution will fight less ardently the greater their own participation in the net national deadweight loss, which in turn would be raised by higher elasticities.

Both the early taxation and the later protection of agriculture may have been reinforced by having lower price elasticities, and therefore lower deadweight costs to check the forces in favor of (first) taxation and (later) subsidies. The premise that price elasticities are lower for agricultural products seems correct, even though econometric attempts to estimate elasticities are notoriously shaky, often leading to underestimates (Schultz 1953, pp. 186-94; Askari and Cummings 1976; Timmer, Falcon and Pearson 1983, pp. 104-9; Carter and Gardiner 1988). On the demand side, the own-price elasticities of demand for most agricultural products (especially foods) are lower than for other sectors of the same size, basically because income elasticities are lower for these products than for non-agricultural products taking similar shares of consumer budgets. On the supply side, low elasticities again seem more dominant in agriculture, perhaps because of the greater reliance on land, a relatively inelastic input. Here may lie a key argument in explaining why the agricultural policy pendulum swings further than policies toward such other sectors as mining, forestry, fisheries, or textiles. Whether or not it corresponds to one's initial intuition, the argument has the virtue of testability. One can calculate the predicted rates of deadweight loss implied by the relevant price elasticities and test to see whether this loss-rate variable has the predicted negative sign in regressions explaining rates of either protection or taxation. Part VII tests the ability of the deadweight-cost effect to explain policy differences among commodities within the agricultural sector.[21]

E. The Case of the Corn Laws.

The greatest historical exception to the rule that the now-industrialized countries drifted from taxing to subsidizing agriculture is one that does fit the present interpretive framework, once we note the key peculiarity of that exceptional case. England, as we have seen, stood out among early modern nations in its desire

to make grain expensive at home from the 1660s to the 1840s, when she became the nation most willing to abandon its agricultural interest in favor of free trade. The case makes sense when the exogenous political oddity of Corn Law England is noted. After 1660, and especially after 1688, England stood out among Western European nations as a nation dominated by the landed aristocracy. Indeed, even in the nineteenth century, her agricultural landownership was outstandingly concentrated and outstandingly correlated with overall wealth, by international standards (Spring 1977, p. 6; Lindert 1986, 1987). Corn Law Repeal was possible in the 1840s largely because the Reform Act of 1832 and other changes in political representation tipped the scales just enough in favor of industrial interests.

While this core political explanation of English exceptionalism is exogenous to the framework of Part V above, that framework helps explain when and why the Corn Laws were differentially applied, and reinterprets the peculiar rise and fall of English poor relief in the eighteenth and nineteenth centuries as a complementary part of a fiscal system dominated by a landed agricultural interest until the reins were seized by others. Over time and space, the Corn Laws and the Old Poor Law wre flexibly applied in ways that fit Part V's price-discriminating-monopoly portrayal of how a political elite exploits the rest of society. Parliamentary struggles over the design of the Corn Laws pitted the obvious landed interest in dear corn against the dangers of exit by a starving or rebellious pauper host. Thus the export bounties and import duties were suspended in isolated years of dearth, when their enforcement would have raised mortality and unrest the most. They were also suspended or only slightly applied throughout the French War era, when food scarcity and the revolutionary threat were perennials. In such times of maximum danger of exit, poor relief also hit its peak. The spatial pattern of poor relief also fits the price-discriminating-monopoly pattern: it was most generous in the off-season and in the rural southeast of England, where labor-hiring landlords and farmers had the greatest stake in retaining a supply of peak-season hands with poor relief partly paid for by other taxpayers (Boyer 1985, 1986). When and where the need to feed the poor was less, the Corn Laws were stiffened and poor relief was dropped. After the Reform Act of 1832 and related changes in local government shifted power toward

urban and industrial interests, both the Corn Laws and the Old Poor Law were in retreat.[22]

VII. REGRESSION TESTS.

The interpretations in Part VI were based on data that are qualitative, or only spotty in their quantification. Using a consistent quantitative data set, we can further test some of the same interpretations plus some others.

A. Sample and Testing Strategy.

Until better historical data have been gathered, the best data base for formal testing consists of a global cross-section around 1980. For 1980 it is possible to compare more countries and commodities, and to try out more hypotheses, than was done in the regression analyses of Honmo and Hayami (in Anderson and Hayami 1986). Combining three sources (Binswanger and Scandizzo 1983, Anderson and Hayami 1986, Tyers and Anderson forthcoming) yields net protection coefficients for a set of 247 nation/commodity cases. Of these, 113 cases are drawn from 13 developed countries, and 134 from 39 developing countries (LDCs). Fourteen different agricultural commodities are represented.

For the present, we must stick to those nominal protection coefficients (NPCs) as our dependent variable. They cast agricultural policy as a single shadow, the international profile revealed by comparing domestic and border prices. There is no better alternative until we have more estimates of "producer subsidy equivalents." Thus, in what follows, we will talk as though the only relevant policies were tax rates on imports and exports, simply because we are forced to use (logs of) NPCs as the dependent variable.

The testing strategy adopted here is to nest some of the competing hypotheses in a single linear equation, and to subject each of the hypothesis to two kinds of competition: competition against the other hypotheses, and competition against a set of what I will call "merely descriptive" variables. A merely descriptive

Selected Themes and Examples from the Literature on
the Political Economics of Environmental Quality

It is useful to understand the differences between environmental quality issues and other public interests. Literature which differentiates environmentally-motivated collective action on the basis of characteristics like imperfect information and high, perceived personal disutility of inaction includes:

Mitchell, Robert Cameron. "National Environmental Lobbies and the Apparent Illogic of Collective Action," in (C. S. Russell, editor) Collective Decision Making: Applications from Public Choice Theory. Baltimore: The Johns Hopkins Univ. Press, 1979, pp. 87-121.

Smith, V. Kerry. "A Theoretical Analysis of the 'Green Lobby'." Amer. Pol. Sci. Review 74(1985): 132-47.

The influence of regulated parties on the choice of environmental regulatory instrument was initially explored by Buchanan and Tullock (included in this collection of readings). Yandle's more recent treatise on that subject (also included in this collection) takes a less theoretical, more pragmatic approach in arriving at similar conclusions regarding the preferences of regulated industries for "command-and-control" policies. For a broad overview on this theme, see also:

Hahn, Robert W. "The Political Economy of Environmental Regulation: Towards a Unifying Framework." Public Choice 65(1990): 21-47.

For empirical support of hypotheses regarding regulated industry preferences and influence, see:

Pashigian, B. Peter. "Environmental Regulation: Whose Self-Interests Are Being Protected?" Economic Inquiry 23(1985): 551-584.

Application of political economic concepts to the relationships among agriculture, environmental quality, and regulation is in an early stage of development. For a conceptual treatment of the role of political preferences in the evolution of agricultural and agriculture-affecting resource and environmental policies, see:

Rausser, G.C. and W.E. Foster. "The Evolution and Coordination of U.S. Commodity and Resource Policies," in (R.E. Just and N. Bockstael, editors) Commodity and Resource Policies in Agricultural Systems. Berlin: Springer-Verlag, 1991, pp. 17-45.

The articles included in this collection of readings by Reichelderfer and by Cropper, Evans and Portney are popularized reports of forthcoming empirical studies. Cropper et al. find evidence of interest group influence on Federal pesticide regulatory decision making. Reichelderfer identifies a series of economic factors that correlate with the revealed strength of interest group influence on Federal legislation affecting both agriculture and environmental quality.

Cropper et al.'s findings may be contrasted with several other investigations of the political economics of pesticide policy:

van Ravenswaay, Eileen O. and Pat T. Skelding. "The Political Economics of Risk/Benefit Assessment: The Case of Pesticides." Amer. J. Agr. Econ. 67(1985): 971-977.

Wise, S. and S.R. Johnson. "A Comparative Analysis of State Regulations for Use of Agricultural Chemicals," in (R.E. Just and N. Bockstael, editors) Commodity and Resource Policies in Agricultural Systems. Berlin: Springer-Verlag, 1991, pp. 48-71.

Polluters' Profits and Political Response: Direct Controls Versus Taxes

By JAMES M. BUCHANAN AND GORDON TULLOCK*

Economists of divergent political persuasions agree on the superior efficacy of penalty taxes as instruments for controlling significant external diseconomies which involve the interaction of many parties. However, political leaders and bureaucratic administrators, charged with doing something about these problems, appear to favor direct controls. Our purpose in this paper is to present a positive theory of externality control that explains the observed frequency of direct regulation as opposed to penalty taxes or charges. In the public-choice theory of policy,[1] the interests of those who are subjected to the control instruments must be taken into account as well as the interests of those affected by the external diseconomies. As we develop this theory of policy, we shall also emphasize an elementary efficiency basis for preferring taxes and charges which heretofore has been neglected by economists.

I

Consider a competitive industry in long-run equilibrium, one that is composed of a large number of n identical producing firms. There are no productive inputs specific to this industry, which itself is sufficiently small relative to the economy to insure that the long-run supply curve is horizontal. Expansions and contractions in demand for the product invoke changes in the number of firms, each one of which returns to the same least-cost position after adjustment. Assume that, from this initial position, knowledge is discovered which indicates that the industry's product creates an undesirable environmental side effect. This external diseconomy is directly related to output, and we assume there is no technology available that will allow alternative means of producing the private good without the accompanying public bad. We further assume that the external damage function is linear with respect to industry output; the same quantity of public bad per unit of private good is generated regardless of quantity.[2] We assume that this damage can be measured and monitored with accuracy.

This setting has been deliberately idealized for the application of a penalty tax or surcharge. By assessing a tax (which can be computed with accuracy) per unit of output on all firms in the industry, the government can insure that profit-maximizing decisions lead to a new and lower industry output that is Pareto optimal. In the short run, firms will undergo losses. In the long run, firms will leave the industry and a new equilibrium will be reached when remaining firms are again making normal returns on investment. The price of the product to consumers will have gone up by the full amount of the penalty tax.

* Center for Study of Public Choice, Virginia Polytechnic Institute and State University. We wish to thank the National Science Foundation for research support. Needless to say, the opinions expressed are our own.

[1] Charles Goetz imposes a public-choice framework on externality control, but his analysis is limited to the determination of quantity under the penalty-tax alternative.

[2] This assumption simplifies the means of imposing a corrective tax. For some of the complexities, see Otto Davis and Andrew Whinston and Stanislaw Wellisz.

No one could dispute the efficacy of the tax in attaining the efficient solution, but we should note that in this setting, the same result would seem to be equally well insured by direct regulation. Policy makers with knowledge of individual demand functions, the production functions for firms and for the industry, and external damage functions, could readily compute and specify the Pareto-efficient quantity of industry output.[3] Since all firms are identical in the extreme model considered here, the policy makers could simply assign to each firm a determinate share in the targeted industry output. This would then require that each firm reduce its own rate of output by X percent, that indicated by the difference between its initial equilibrium output and that output which is allocated under the socially efficient industry regulation.[4]

Few of the standard arguments for the penalty tax apply in this setting. These arguments have been concentrated on the difficulties in defining an efficient industry output in addition to measuring external damages and on the difficulty in securing data about firm and industry production and cost functions. With accurately measured damage, an appropriate tax will insure an efficient solution without requiring that this solution itself be independently computed. Or, under a target or standards approach, a total quantity may be computed, and a tax may be chosen as the device to achieve this in the absence of knowledge about the production functions of firms.[5]

[3] See Allen Kneese and Blair Bower, p. 135.

[4] No problems are created by dropping the assumption that firms are identical so long as we retain the assumption that production functions are known to the regulator.

[5] This is the approach taken by William Baumol, who proposes that a target level of output be selected and a tax used to insure the attainment of this target in an efficient manner.

In the full information model, none of these arguments is applicable. There is, however, an important economic basis for favoring the penalty tax over the direct control instrument, one that has been neglected by economists. The penalty tax remains the preferred instrument on strict efficiency grounds, but, perhaps more significantly, it will also facilitate the enforcement of results once they are computed.[6] Under the appropriately chosen penalty tax, firms attain equilibrium only at the efficient quantity of industry output. Each firm that remains in the industry after the imposition of the tax attains long-run adjustment at the lowest point on its average cost curve only after a sufficient number of firms have left the industry. At this equilibrium, there is no incentive for any firm to modify its rate of output in the short run by varying the rate of use of plant or to vary output in the long run by changing firm size. There is no incentive for resources to enter or to exit from the industry. So long as the tax is collected, there is relatively little policing required.

This orthodox price theory paradigm enables the differences between the penalty-tax instrument and direct regulation to be seen clearly. Suppose that, instead of levying the ideal penalty tax, the fully informed policy makers choose to direct all firms in the initial competitive equilibrium to reduce output to the assigned levels required to attain the targeted efficiency goal for the industry. No tax is levied. Consider Figure 1, which depicts the situation for the individual firm. The initial competitive equilibrium is attained when each firm produces an output, q_i. Under regulation it is directed to produce only q_0, but no tax is levied. At output q_0, with an unchanged number of firms, price is above

[6] See George Hay. His discussion of the comparison of import quotas and tariffs on oil raises several issues that are closely related to those treated in this paper.

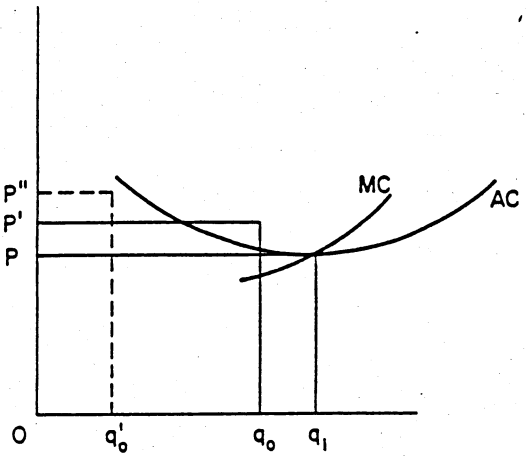


FIGURE 1

marginal cost (for example price is at P'). Therefore, the firm is not in short-run equilibrium, and would if it could expand output within the confines of its existing plant. More importantly, although each firm will be producing the output quota assigned to it at a somewhat higher cost than required for efficiency reasons, there may still be an incentive for resources to enter the industry. The administrator faces a policing task that is dimensionally different from that under the tax. He must insure that individual firms do not violate the quotas assigned, and he must somehow prevent new entrants. To the extent that the administrator fails in either of these tasks, the results aimed for will not be obtained. Output quotas will be exceeded, and the targeted level of industry production overreached.

If the administrator assigns enforceable quotas to existing firms and successfully prevents entrants, the targeted industry results may be attained, but there may remain efficiency loss since the industry output will be produced at higher average cost than necessary if firms face U-shaped long-run average cost curves. Ideally, regulation may have to be accompanied by the assignment of full production quotas to a selected number of the initial firms in the industry. This policy will keep these favored firms in marginal adjustment with no incentives for in-firm adjustments that might defeat the purpose of the regulation. But even more than under general quota assignment there will be strong incentives for firms to enter the industry and to secure at least some share of the rents that the restriction of industry output generates. If the response to this pressure should be that of reassigning quota shares within the unchanging and targeted industry output so as to allow all potential entrants some share, while keeping all firms, actual and potential, on an equal quota basis, the final result may be equivalent to the familiar cartel equilibrium. No firm will be earning more than normal returns, but the industry will be characterized by too many firms, each of which produces its assigned output inefficiently.

II

When we examine the behavioral adjustments to the policy instruments in the manner sketched out above, a theory of policy emerges. Regulation is less desirable on efficiency grounds even in the presence of full information, but this instrument will be preferred by those whose behavior is to be subjected to either one or the other of the two policy instruments. Consider the position of the single firm in the fully competitive industry, depicted in Figure 1. Under the imposition of the tax, short-run losses are necessarily incurred, and the firm reattains normal returns only after a sufficient number of its competitors have shifted resources to other industries. The tax reduces the present value of the firm's potential earnings stream, whether the particular firm remains in the industry after adjustment or withdraws its investment and shifts to alternative employ-

ment. In terms of their own private interests, owners of firms in the industry along with employees will oppose the tax. By contrast, under regulation firms may well secure pecuniary gains from the imposition of direct controls that reduce total industry output. To the extent that the restriction is achieved by the assignment of production quotas to existing firms, net profits may be present even for the short term and are more likely to arise after adjustments in plant. In effect, regulation in this sense is the directional equivalent of cartel formation provided that the individual firm's assigned quota falls within the limited range over which average cost falls below price. Such a range must, of course, exist, but regulatory constraints may possibly be severe enough to shift firms into positions where short-term, and even possibly long-term, losses are present, despite increased output price. Such a result is depicted by a restriction to q_0' in Figure 1, with price at P''.

Despite the motivation which each firm has to violate assigned quotas under regulation, it remains in the interest of firms to seek regulatory policy that will enforce the quotas. If existing firms foresee the difficulty of restricting entry, and if they predict that governmental policy makers will be required to accommodate all entrants, the incentive to support restriction by regulation remains even if its force is somewhat lower. In final cartel equilibrium, all the firms will be making no more than normal returns. But during the adjustment to this equilibrium, above-normal returns may well be available to all firms that hold production quotas. Even if severe restriction forces short-term losses on firms, these losses will be less than those under the tax. Rents over this period may well be positive, and even if negative, they will be less negative than those suffered under the tax alternative. Therefore, producing firms will always oppose any imposition of a penalty tax. However, they may well favor direct regulation restricting industry output, even if no consideration at all is given to the imposition of a tax. And, when faced with an either/or choice, they will always prefer regulation to the tax.

III

There is a difference between the two idealized solutions that has not yet been discussed, and when this is recognized, the basis of a positive hypothesis about policy choice may appear to vanish. Allocationally, direct regulation can produce results equivalent to the penalty tax, providing that we neglect enforcement cost differentials. *Distributionally*, however, the results differ. The imposition of tax means that government collects revenues (save in the case where tax rates are prohibitive) and these must be spent. Those who anticipate benefits from the utilization of tax revenues, whether from the provision of publicly supplied goods or from the reduction in other tax levies, should prefer the tax alternative and they should make this preference known in the political process. To the extent that the beneficiaries include all or substantially all members of the community, the penalty tax should carry the day. Politicians, in responding to citizenry pressures, should heed the larger number of beneficiaries and not the disgruntled members of one particular industry. This political choice setting is, however, the familiar one in which a small, concentrated, identifiable, and intensely interested pressure group may exert more influence on political choice making than the much larger majority of persons, each of whom might expect to secure benefits in the second order of smalls.

There is an additional reason for predicting this result with respect to an innovatory policy of externality control. The penalty tax amounts to a legislated change in property rights, and as such it will be

viewed as confiscatory by owners and employees in the affected industry. Legislative bodies, even if they operate formally on majoritarian principles, may be reluctant to impose what seems to be punitive taxation. When, therefore, the regulation alternative to the penalty tax is known to exist, and when representatives of the affected industry are observed strongly to prefer this alternative, the temptation placed on the legislator to choose the direct control policy may be overwhelming, even if he is an economic theorist and a good one. Widely accepted ethical norms may support this stance; imposed destruction of property values may suggest the justice of compensation.[7]

If policy alternatives should be conceived in a genuine Wicksellian framework, the political economist might still expect that the superior penalty tax should command support. If the economist ties his recommendation for the penalty tax to an accompanying return of tax revenues to those in the industry who suffer potential capital losses, he might be more successful than he has been in proposing unilateral or one-sided application of policy norms. If revenues are used to subsidize those in the industry subjected to capital losses from the tax, and if these subsidies are unrelated to rates of output, a two-sided tax subsidy arrangement can remove the industry source of opposition while still insuring efficient results. In this respect, however, economists themselves have failed to pass muster. Relatively few modern economists who have engaged in policy advocacy have been willing to accept the Wicksellian methodological framework which does, of course, require that some putative legitimacy be assigned to rights existent in the status quo.[8]

[7] For a comprehensive discussion of just compensation, see Frank Michelman.

[8] For a specific discussion of the Wicksellian approach, see Buchanan (1959).

IV

To this point we have developed a theory of policy for product-generated external diseconomies, the setting which potentially counterposes the interest of members of a single producing industry against substantially all persons in the community. External diseconomies may, however, arise in consumption rather than in production, and these may be general. For purposes of analysis, we may assume that all persons find themselves in a situation of reciprocal external diseconomies. Traffic congestion may be a familiar case in point.

The question is one of determining whether or not persons in this sort of interaction, acting through the political processes of the community, will impose on *themselves* either a penalty tax or direct regulation. We retain the full information assumption introduced in the production externality model. For simplicity here, consider a two-person model in which each person consumes the same quantity of good or carries out the same quantity of activity in the precontrol equilibrium, but in which demand elasticities differ. Figure 2 depicts the initial equilibrium at E with each person consuming quantity Q. The

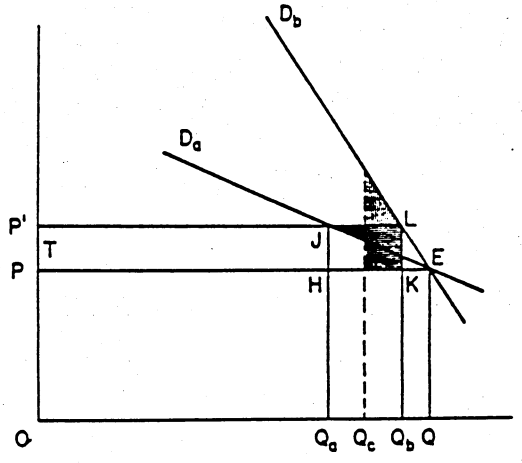


FIGURE 2

existence of the reciprocal external diseconomy is discovered. The community may impose an accurately measured penalty tax in the amount T, in which case A will reduce consumption to Q_a and B will reduce consumption to Q_b. Total consumption is reduced from $2Q$ to (Q_a+Q_b), but both A and B remain in equilibrium. At the new price P', which includes tax, neither person desires to consume more or less than the indicated quantities. The government collects tax revenues in the amount $[2(PP'JH)+HJLK]$. Alternatively, the community may simply assign a restricted quantity quota to each person. If the government possesses full information about demand functions it can reduce A's quota to Q_a, and B's quota to Q_b, securing results that are allocatively identical to those secured by the tax. However, under the quota, both A and B will find themselves out of equilibrium; both will, if allowed quantity adjustment, prefer to expand their rate of consumption.

It will be useful to examine the ideal tax against the quota scheme outlined above, which we may call the idealized quota scheme. If individuals expect no returns at all from tax revenues in the form of cash subsidies, public goods benefits, or reductions in other taxes, both A and B will clearly prefer the direct regulation. The loss in consumers' surplus under this alternative is small relative to that which would be lost under the penalty tax. Each person willingly trades off marginal quantity adjustment for the more favorable inframarginal terms offered under direct regulation, given our assumptions that both instruments achieve the same overall externality control objective.

Under extreme fiscal illusion, individuals may ignore benefits from tax revenues, but consistent methodological precept requires that we allow persons to recognize the benefit side of the fiscal account, at least to some degree. Let us allow all revenues under the penalty tax to be returned in equal shares to all taxpayers. Under this arrangement, each person expects to get back one-half of the amount measured as indicated above for Figure 2. Simplifying, each expects to get back the amount $PP'JH$, which he personally pays in, plus one-half of the amount measured by the rectangle $JHKL$, all of which is paid in by B. From an examination of Figure 2, it is clear that individual A will favor the penalty tax under these assumptions. The situation for individual B is different; he will prefer direct regulation. He will secure a differential gain measured by the horizontally shaded area in Figure 2, which is equal to the differential loss that individual A will suffer under this alternative. The policy result, insofar as it is influenced by the two parties, is a standoff under this idealized tax and idealized quota system comparison.

For constitutional and other reasons, control institutions operating within a democratic order could carcely embody disproportionate quota assignments. A more plausible regulation alternative would assign quotas proportionate to initial rates of consumption, designed to reduce overall consumption to the level indicated by target criteria. The comparison of this alternative with the ideal tax arrangement is facilitated by the construction of Figure 2 where the initial rates of consumption are equal. In this new scheme, each person is assigned a quota Q_e, which he is allowed to purchase at the initial price P. We want to compare this arrangement with the ideal tax, again under the assumption that revenues are fully returned in equal per head subsidies. As in the first scheme, both persons are in disequilibrium at quantity Q_e and price P. The difference between this model and the idealized quota scheme lies in the fact that at Q_e, the marginal evaluations differ as between the two persons. There are unexploited gains from

trade, even under the determined overall quantity restriction.

It will be mutually advantageous for the two persons to exchange quotas and money, but, at this point, we assume that such exchanges do not take place, either because they are prohibited or because transactions costs are too high. Individual *A* will continue to favor the tax alternative but his differential gains will be smaller than under the idealized quota scheme. In the model now considered, *A*'s differential gains under the ideal tax are measured by the blacked-in triangle in Figure 2. Individual *B* may or may not favor the quota, as in the earlier model. His choice as between the two alternatives, the ideal tax on the one hand and the restriction to Q_c at price *P* on the other, will depend on the comparative sizes of the two areas shown as horizontally and vertically shaded in Figure 2. As drawn, he will tend to favor the quota scheme, but it is clearly possible that the triangular area could exceed the rectangular one if *B*'s demand curve is sufficiently steep in slope. In any case, the choice alternatives for both persons are less different in the net than those represented by the ideal tax and the idealized quota.

While holding all of the remaining assumptions of the model, we now drop the assumption that no exchange of quotas takes place between *A* and *B*. To facilitate the geometrical illustration, Figure 3 essentially blows up the relevant part of Figure 2. With each party initially assigned a consumption quota of Q_c, individual *A* will be willing to sell units to individual *B* for any price above his marginal evaluation. Hence, the lowest possible supply price schedule that individual *B* confronts is that shown by the line *RL* in Figure 3. The maximum price that individual *B* is willing to pay for additional units of quota is his marginal evaluation, shown by *SL*. The gains-from-trade are measured by the triangular area *RLS*. The distribution of

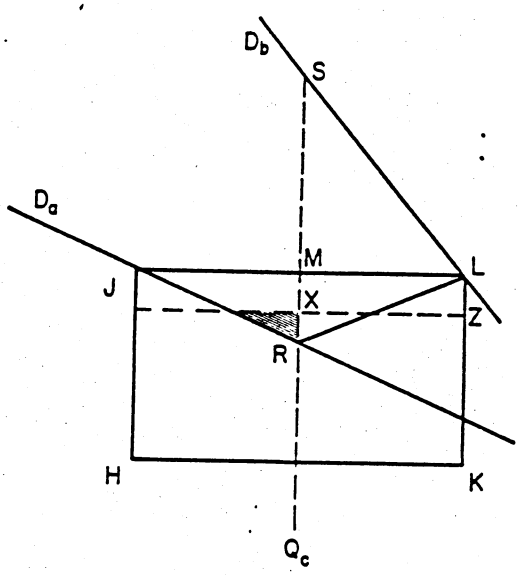


FIGURE 3

these gains will, of course, be settled in the strict two-man setting by relative bargaining skills, but let us assume that individual *B*, the buyer, wants to purchase consumption quota units from *A*, but also to do so in such a way that individual *A* will come to prefer this system over the tax. To accomplish this, he must insure that *A* gets a share of the net gains at least equal to the area *RML* on Figure 3. Individual *B*, the buyer, retains gains of *MSL* under this division of the spoils. But in this arrangement, both persons are indifferent as between the policy alternatives. The system is on the Pareto frontier, and the quota scheme plus the exchange process produces allocative and distributive results identical to those generated under the ideal tax. This becomes the analogue of the Coase theorem in the context that we are examining.[9]

V

These somewhat inconclusive results may seem to provide anything but a posi-

[9] See Ronald Coase. For a related extension of the Coase theorem, see Buchanan (1973).

tive theory of policy akin to that presented with respect to production externalities. The comparisons are, however, a necessary stage in developing such a theory. Recall that we have made these comparisons under the most favorable possible assumption concerning anticipated return of revenues under the penalty tax. In the real world, individuals will not anticipate that these will be returned dollar-for-dollar, and they will tend to place at least some discount on the value of benefits that they expect.

Let us say that each person expects an aggregate benefit value of only 80 cents on the dollar from tax revenues collected under the penalty tax. Consider what this single change does to the results of the last comparison made, that which involves proportionate quota assignments along with a free market in quotas. In this case, individual *B*, the buyer, can offer individual *A*, the seller, more than the amount required to make him prefer the quota alternative, while himself continuing to secure differential benefit under this alternative. Individual *A*'s differential gains from the ideal penalty tax are reduced to the shaded area in Figure 3. By paying individual *A* the amount measured by *RML*, he has improved *A*'s position relative to the penalty tax. And, in the process, he has retained for himself a differential gain measured by the area *MXZL*. Both persons in full knowledge of the alternatives will prefer the quota system, and political leaders will presumably respond by opting for regulation.

The same reasoning can readily be extended to apply to any quota system. In the idealized quota assignment first considered, we demonstrated that one person would favor the penalty tax and the other the quota. Individual *A*, who favors the penalty tax, loses no consumer's surplus, and he does expect to secure an income transfer through the return of tax revenues. When we modify the assumptions concerning expectations of the value of returned revenues or benefits, however, this conclusion need not hold. Individual *A* will, of course, expect to get back in benefits some part of the tax revenues paid in by *B* that is in excess of that contributed by *A* himself. If, however, individual *A* applies the same discount factor to all revenues collected, the deadweight loss may more than offset the income transfer effect. Examination of Figure 2 indicates that under the 80 percent assumption, one-fifth of the area measured by *PP'JH* will represent deadweight loss to *A* from the revenues that he pays in. This deadweight loss may well be larger than the measure of the income transfer that he expects, which amounts to 80 percent of the horizontally shaded area in Figure 2. Once we introduce any plausible discount factor into the expectation of individuals concerning the return of tax revenues, it is relatively easy to demonstrate situations under which both persons may be led by private self-interest to favor the direct regulation alternative.

VI

We have developed a positive theory of externality control policy for both the production and consumption interactions under highly abstract and simplified models which allow us to isolate influences on policy formation which have been neglected. Decisions on the alternative policy instruments in democratic governments are surely influenced by the preferences of those who are subjected to them. The public-choice approach, which concentrates attention on the individual's choice as between policy instruments, allows us to construct hypotheses that explain the prevalence of direct regula-

tion.[10] For economists who continue to support the penalty tax alternative, the analysis suggests that they had best become good Wicksellians and begin to search out and invent institutional arrangements that will make the penalty tax acceptable to those who are primarily affected.

[10] Much of the analysis developed in this paper can be applied more or less directly to policy alternatives proposed in the energy crisis of late 1973 and early 1974. For such application, see Buchanan and Nicolaus Tideman.

REFERENCES

W. J. Baumol, "On Taxation and the Control of Externalities," *Amer. Econ. Rev.*, June 1972, *62*, 307–22.

J. M. Buchanan, "Positive Economics, Welfare Economics, and Political Economy," *J. Law Econ.*, Oct. 1959, *2*, 124–38.

———, "The Coase Theorem and the Theory of the State," *Natur. Resources J.*, Oct. 1973, *13*, 579–94.

——— and N. Tideman, "Gasoline Rationing and Market Pricing: Public Choice in Political Democracy," research pap. no. 808231-1-12, Center for Study of Public Choice, Virginia Polytechnic Inst. and State Univ., Jan. 1974.

R. H. Coase, "The Problem of Social Cost," *J. Law Econ.*, Oct. 1960, *3*, 1–44.

O. A. Davis and A. Whinston, "Externalities, Welfare, and the Theory of Games," *J. Polit. Econ.*, June 1962, *70*, 241–62.

C. J. Goetz, "Political Equilibrium vs. Economic Efficiency in Effluent Pricing," in J. R. Conner and E. Loehman, eds., *Economic Decisionmaking for Environmental Control*, Gainesville 1973.

G. A. Hay, "Import Controls on Foreign Oil: Tariff or Quota?," *Amer. Econ. Rev.*, Sept. 1971, *61*, 688–91.

A. V. Kneese and B. T. Bower, *Managing Water Quality: Economics, Technology, Institutions*, Baltimore 1968.

F. J. Michelman, "Property Utility, and Fairness: Comments on the Ethical Foundations of 'Just Compensation' Law," *Harvard Law Rev.*, Apr. 1967, *80*, 1165–1257.

S. Wellisz, "On External Diseconomies and the Government-Assisted Invisible Hand," *Economica*, Nov. 1964, *31*, 345–62.

Anal # 20215

Why Environmentalists Should Be Efficiency Lovers

by Bruce Yandle

CENTER FOR THE STUDY OF AMERICAN BUSINESS

Formal Publication Number 105

April 1991

Introduction

The 1990s: Will This Be the Decade of the Environment?

By any measure we might apply, the 1990s will be a decade of struggle with environmental issues. Why so? Why now?

We can tick off major accidents and activities that heightened ecological awareness in the late 1980s: the Valdez oil spill, the spewing of radiation from Chernobyl, massive timber cuttings in Brazil, the frequent appearance of medical wastes on the nation's seashores, the high-pitched debates on global warming, and the sorry state of environmental quality now seen in Eastern Europe. But while we can agree that these events and conditions capture our attention, we can find similar threatening disasters in times past, events that were not followed by a national ecological movement.

More is required for defining a period of environmental struggle than a series of environmental calamities. A political focal point is needed. A struggle for change must be focused on political decision makers, and the level of activity becomes more intense when political decision makers have the power to write and enforce all-encompassing federal rules. The U.S. Congress, working through the U.S. Environmental Protection Agency (EPA), offers that attraction. They have monopoly power in regulating the environment.

The timing of recent environmental disasters coincides with and reinforces political action. For example, the Clean Air Act has recently been revised with tighter emissions standards. The Comprehensive Environmental Response, Compensation and Liability (Superfund) legislation will be up for reauthorization in 1991. International conferees are discussing global warming, and new efforts are being made to bring a coordinated approach to the control of pollution that crosses national boundaries.

Always alert to the mood of the public, Madison Avenue has also responded. Previously unadorned soft drink bottles are now labeled environmentally safe. Deodorants and hair spray are hailed for their ecologically neutral propellants or for having no

---

Bruce Yandle is Alumni Professor and Director of the Strom Thurmond Institute of Government and Public Affairs at Clemson University. An abbreviated version of this paper was given at the *Key Issues Forum on Environmental Regulation* at the Center for the Study of American Business.

372

propellant at all. Indeed, the enthusiasm of ad copy writers has waxed to the point that the Federal Trade Commission (FTC) is now scrutinizing environmental claims made for consumer goods that might, in the FTC's opinion, impose losses on unwitting buyers who think they are buying ecologically safe products when in fact they are not.[1]

The banter about environmentally friendly products is joined by efforts to organize recycling stations across communities nationwide and by the promulgation of rules to guide the safe construction and operation of landfills.

Playing its historic role as bellwether state, California is taking the lead in developing tighter auto-emissions standards and other environmental rules that exceed and will influence the standards set at the federal level.[2] In an effort to reduce driving and thereby control auto emissions, California's South Coast Air Quality District is discussing some rather drastic steps. Consumers of hamburgers at fast food shops may no longer enjoy drive-up service. Gasoline-powered lawnmowers could have emissions controls imposed on them and driving restrictions may be imposed on commuters to reduce the large flow of cars into and out of cities.

Massive expenditures on pollution control do not necessarily result in equally significant improvements in environmental quality.

Some might argue that the magnitude of environmental degradation and the high and lofty goals of the environmental movement are sufficient to command general support for new packets of legislative solutions. Environmental degradation is real. With sufficient exposure, toxic chemicals can be deadly to humans. Oil spills definitely do destroy aquatic life and property values. Contaminated aquifers can raise the cost of providing community drinking water. If cleanliness is next to godliness, how can we have too much?

The outcomes of more than 20 years experience with federal programs to improve the quality of air, water, and earth warn us that things are not so simple. We have learned that lofty goals embodied in highly detailed federal legislation do not translate smoothly into effective solutions. Massive expenditures on pollution control do not necessarily result in equally significant improvements in environmental quality.

Can Economic Logic Help?

Agreeing that the 1990s will be a decade of environmental struggle is far different from concluding that this will be a decade of progress in cleaning and protecting the environment. Significant progress will come when the resources spent generate the largest gain possible. And that has not happened.

Economic logic does not question environmental goals. The logic tells us we should seek the biggest bang for each environmental buck. The application of economic thinking forces us to focus on incentives, how things are done. We have to consider the process as much as the end goal itself. If outcomes matter, economic logic can help.

A review of the record since 1970 tells us we have gotten far too little from the massive effort to control pollution. Indeed, the record since 1970 leaves a strong and somewhat paradoxical impression that environmental quality could have been protected more effectively had it not been for the federal monopolizing effort to do so.

It is difficult to see how any group that prizes environmental quality can be happy with the outcome. Surely those who have borne the costs of environmental cleanup are not happy. Somehow, the application of economic logic has taken a back seat to other motivations.

If the 1990s are to be a decade of progress, those who truly love the environment will become lovers of efficient management of environmental quality. They will find allies among those who worry about the wasteful use of scarce resources. Economic logic will have to move to the front seat.

Some Key Questions

Why is the record of environmental accomplishments so disappointing? And if there is agreement that the results are worse than anticipated, then why are we today still repeating the failed experiments of the past? Is there a better path than the one we have followed?

These questions lie at the heart of this report. Environmental purity is an elusive goal. And like most things perfect, a pure environment will never be achieved. Even so, we should expect to observe measurable progress from extensive cleanup efforts.

We should also expect environmentalists who seek to gain as much environmental quality as possible to be among the first to support cost-effective pollution control. Lacking that, they should be the first to sound the alarm when ineffective actions are taken. Oddly enough, that has not been the case. Indeed,

environmental organizations have more often than not supported the most costly forms of regulation rather than the most effective.

The next two sections of the study describe some recent history and explain why so little progress has been made. The track record for air and water pollution is presented in section two, while a discussion of incentives that cause major players in the struggle to prefer high-cost outcomes is developed in section three. The cleanup of hazardous wastes presents a different kind of environmental challenge and another graphic picture of high cost with little accomplishment. The background to and record of Superfund regulations are discussed in section four and an alternate policy proposal is presented.

The cleanup of hazardous wastes presents a graphic picture of high cost with little accomplishment.

An even more intriguing question is, "Why are we repeating the failed experiments of the past?" If we understand that past efforts came at great cost and lacked effectiveness, continuing on the same path implies rational choice. In other words, the political mechanism has purposefully disregarded lower-cost options in favor of higher-cost, less-effective approaches. The last major part of the paper addresses this tendency to continue programs that are acknowledged to be inefficient.

It should be made clear at the outset that this paper is not an attempt to convince readers that government is misguided or that nothing should be done to protect the environment. To the contrary, the stories recounted here reflect evidence that people place a high value on a cleaner environment and that government is carefully guided in its undertakings. Saying that government is carefully guided does not explain how the guidance system works, however.

The analysis presented here assumes that politicians consciously, but never perfectly, carry out the wishes of those who value political services most highly. Another way of expressing this same thought is to say that environmental rules are traded in political markets. The politician-broker who fails to satisfy market participants will be replaced by others who perform more effectively. The struggle over environmental regulation takes place in a political market arena.

Large Expenditures for Relatively Little Progress

A Look at the Record

The assertion that relatively little progress has been made toward important environmental goals requires some justification. It is not equivalent to saying that no progress has been made. America's air and water are cleaner today than two decades ago, even though the population has grown by 50 million.

For example, many rivers are cleaner today than they were in 1972, the year Congress passed the first major piece of federal water pollution control legislation. The air we breathe is also purer in important ways. In spite of growth of human and automobile populations, urban air has not become more polluted. Indeed, EPA's latest *National Air Quality and Emissions Trends Report*, which reflects 1987 data, shows that since 1978, emissions of lead oxide have fallen by 88 percent, carbon monoxide by 32 percent, sulfur dioxide by 35 percent, and nitrogen oxides have decreased 12 percent.[3] We can point to hundreds of tons of contaminated soil that have been cleaned and note outright bans on harmful insecticides and environmentally offensive chemicals.

The issue, however, is whether the amount of environmental progress is commensurate with the cost. The efforts are massive, while the results are modest.

To determine effectiveness of the massive cleanup efforts, we must attempt to isolate their effects from other things that could independently affect environmental quality. With success in that undertaking, we can then consider the trend line that describes environmental progress before the outset of tremendously costly federal programs and determine if the additional activities increased, decreased, or simply maintained progress.

A host of factors beyond federal regulations can affect environmental quality. To illustrate, if the nation sustains a serious recession, as in 1982, output of goods, services, *and pollution* declines. Similar considerations need to be given to energy costs. Since the mid-1970s, the higher relative price of fossil fuels has led to sharp reductions in the use of energy for producing goods and services. Those substitutions have contributed to reductions in sulfur dioxide and hydrocarbon emissions. All else equal, the environment has improved. The point: to isolate the independent effects of federal regulation, we have to control for industrial production and energy consumption, since those activities are major sources of pollution.[4]

What Is the Baseline?

Where do we begin if we wish to determine the effectiveness of federal environmental programs? Obviously, environmental protection did not begin in 1970 with the establishment of the U.S. Environmental Protection Agency. For decades before, people acting through state and municipal governments as well as through multi-state compacts took purposeful steps to control pollution and protect the natural environment.[5] As might be expected, given the great variety of economic activities and demographic characteristics of people across the 50 states, every form of environmental regulation was applied, including none at all. Rules were strict in some cases and loose in others. The regulatory process was diverse and competitive.

Competition for industrial plant locations and the need for politicians to satisfy voter-citizens added discipline to independent state actions. No single state legislature could set dramatically higher or lower environmental standards without suffering the effects that came when citizens and industrialists reacted. Consumers and investors can vote with their feet.

The year 1970 marked a move that modified the 50-state laboratory, where sharply different regulatory approaches were applied to heterogeneous conditions and differing community preferences for environmental protection. These were replaced with an environmental control monopolist—the EPA—which was soon equipped with a uniform congressional mandate to be applied across all states.

As might be expected, evidence on the effects of the transition with respect to air-pollution control suggests strongly that community leaders lost power to control the environmental quality in their own areas.[6] Less attention was paid to things that matter to ordinary people, more was devoted to matters that were "of national importance." Instead of focusing on problems across town, as might be the case when community leaders are responsible for environmental control and the community bears the cost of cleanup, the national government focused on average ambient conditions for larger regions.

What About the EPA Trend Line?

What can be said about achievement of environmental protection in the pre-EPA period? What about the trend line that reckons environmental quality against the passage of time? Researchers have attempted to estimate improvements in air quality while adjusting for the effects of industrial production and

other activities that can induce independent effects on outcomes. The task is understandably daunting.

It is impossible to adjust completely for all activities that have a bearing on environmental quality. At some point, the environmental rules themselves become entangled in the mix of industry and other factors supposedly being held constant. As a result, the findings offer evidence, not conclusions.

Keeping that caveat in mind, a review of the research findings makes a weaker case for federal than for state control in terms of environmental improvement.[7] After completing a 1983 survey of EPA's record compared with earlier state efforts and making his own estimate of the agency's effect, Brookings Institution economist Robert W. Crandall summarized his findings:

> Every tabulation since 1972 shows less relative improvement than was achieved in the 1960s. Sulfur dioxide concentration appears to have fallen 11.3 percent per year from 1965 through 1971 but no more than 4.6 percent per year in the 1970s. Similarly, the average concentration of TSP (total suspended particulates) fell 2.3 percent per year in the 1960-71 period, but only 0.6 percent per year from 1972 to 1980.[8]

Crandall's research took into account GNP growth and other factors that could have increased or decreased pollution.

While the summary is discouraging for advocates of federal programs, Crandall's final thoughts hit even harder: "Therefore, these data suggest that pollution reduction was more effective in the 1960s, before there was a serious federal policy dealing with stationary sources, than since the 1970 Clean Air Amendments."[9]

Similar but less conclusive findings surface when improvements in water quality are examined. The difficulty in drawing firmer conclusions relates to adjusting for factors other than federal regulation that could affect water quality.

Former Council of Environmental Quality economist H. Myrick Freeman III notes both the lack of progress and the assessment problem in his 1990 review of federal water programs.[10] Freeman indicates that "average water quality was not bad in the United States in 1972," but that some localities had serious problems.[11] He goes on to point out that the lack of meaningful progress is due to the slow implementation of major elements of the federal program.

Implementation has been slow because costs have been high, says Freeman. He uses EPA and Council on Environmental Quality estimates to compare the benefits of water quality improvement to costs imposed by the federal programs.[12] Freeman

finds the annual cost of programs in the neighborhood of $25-$30 billion (1984 dollars). The estimates of benefits range from $5.7-$27.7 billion. Accepting the highest possible estimate of benefits and lowest estimate of costs, yields at best, a break-even situation. The point: cost-effective control based on market incentives would have generated more water quality improvement over the past decade and a half.

The annual cost of clean water programs is in the neighborhood of $25-$30 billion (1984 dollars); estimates of benefits range from $5.7-$27.7 billion—at best a break-even situation.

Although research findings are somewhat tentative on the relative merits of air- and water-pollution control by states versus control through federal programs, the picture for hazardous wastes is even less clear. There were few state regulatory efforts to deal with hazardous waste before the 1970s. Rules of liability, common law remedies, and local ordinances were the controlling mechanisms. EPA's Superfund program emerged in the early 1980s to address the cleanup of old hazardous waste sites. All assessments of Superfund show it to be a high-cost, low-output program. Indeed, support for the program on a cost-effectiveness basis is conspicuously absent.

These results do not tell us that federal regulation was a mistaken venture that should have been avoided. This fragmented evidence does suggest, however, that the cost of obtaining environmental quality rose as federal control expanded. Instead of making it easier for citizens to improve the environment, the federal rules made it more difficult.

Incentives, Regulation and Politics

Why would federal efforts prove to be less cost-effective than state controls? And why would environmental groups support such ineffective programs?

Conventional wisdom suggests that federal regulation will achieve more than the messy diversity of state regulations and municipal ordinances will achieve. This rather orthodox logic is

based on a vision of government that serves a broad public interest. It argues that, in spite of political pressures that deflect the regulatory process off its public-interest path, the centralized federal machinery will be effective when viewed over a long period of time.

Unfortunately, the public-interest logic fails to consider the economic incentives of environmental groups, politicians, industrialists, and others who seek federal as opposed to state regulation. The logic overlooks completely the motivation of those who operate the regulatory process once it is centrally administered. The public-interest perspective also undervalues the role played by local citizens in policing public-sector performance.

By overlooking incentives that affect human action, the public-interest story fails to recognize the enormous amount of wealth that can be garnered for specific groups through the use of detailed technical standards enforced nationally by the federal government. It underemphasizes the fact that environmental quality is perceived by people locally and that political voice needs to be located near the problem.

When considering motivation, the public-interest story does not take into account the incentives of environmental organizations, that they too seek to expand "markets" and revenues. Politicians are viewed as agents of the public interest rather than brokers who must find ways to balance their actions across competing interest groups.

The role played by incentives surface when we consider the choice between the uniform federal regulations of the post-1970s and the earlier heterogeneous state and local rules. In either case, competitive forces cause compliance costs to be borne by consumers. The earlier state-controlled regulation presented a major problem to industrial firms—uncertainty. Industrial firms could not readily predict production costs at any given location. Firms located in highly polluted locations with strict state laws frequently had to compete with firms in less-polluted locations with less-restrictive laws. The cost differences could be overcome by expanding production in the lower-cost location, but expansion elsewhere was costly on other grounds.

Workers with higher-wage contracts, politicians with specialized knowledge about particular constituent groups, mayors of older cities with fragile tax bases, and individuals who just did not want to move or lose the benefits of a somewhat stable social structure, opposed the loss of industry to the newer regions. Those who simply favored a clean environment that could be paid for by a vast body of consumers and taxpayers argued for

federal environmental protection, a uniform baseline to be met eventually at all locations.

While these incentives played across the various private interest groups, another set of incentives propelled organized environmental groups that understandably sought to become a national influence.[13] Organizing to affect 50 state legislatures is far more costly than organizing to affect the U.S. Congress. Economies of scale to be found in educating and influencing legislators and in soliciting support placed the environmentalists on the side of federal rules. In addition, a large federal bureaucracy dedicated to protecting the environment provides a training ground and a funding source to feed the expanding national organizations.[14]

The combined force of the organized interest groups was politically overwhelming. National industries, established regions, local government officials, elected representatives, and members of environmental organizations could gain what they desired individually from a centralized approach to environmental protection. In short, federal uniformity generated far more special interest benefits than state diversity. All parties seemed to agree on the need for federal controls, even the polluters.

Idealistic (naive) observers of the shift to federal environmental programs thought the goal of environmental purity was in sight. But victory was declared when rules were written, not when environmental quality was delivered.

There was no efficiency lobby, no systematic effort to force government to minimize the cost of the environmental journey. Instead, the federal programs emerged in a vacuum purged of economic considerations.

What Happened to Uniformity?

This explanation of why environmental regulation was shifted to the national government still does not tell us why the federal effort was so costly and so unproductive. To appreciate why federal regulation took such an ineffective turn, we must first review some of the incentives that govern the behavior of polluters. The regulatory options available to environmentalists, politicians, and federal regulators then should be compared to these basic motivations.

The notions are simple, thus discussion is appropriately brief. But while brief, the review helps to identify the extent to which desires for obtaining as much environmental quality as possible was overwhelmed by the other influences in the final selection of control strategies.

Polluter Incentives

Let us focus on major industrial firms that are the largest potential sources of pollution. Taken as a group, the firms that produce pollution have a number of incentives that cause them to avoid actions that damage the environment. Fundamentally, all privately owned firms are concerned about the bottom line—profit. It is this concern that leads them to protect the environment. How is this so?

Firms that produce pollution have a number of incentives that cause them to avoid actions that damage the environment.

Private firms have to be concerned about long-run profitability, which is determined ultimately by consumers in the marketplace, the owners and employees of the firm, and other major stakeholders. Brand name capital (reputation), which results from past successes, also conditions future outcomes.[15] All else equal, a firm with a good reputation will outsell firms with no reputation.

A good name is an investment that can depreciate when a firm acts irresponsibly in the eyes of consumers and shareholders. The larger the firm and the more extensive its product line, the larger is the bond represented by its brand name. Larger multi-product firms that become environmentally careless have much at risk. Actions by one division impose reputational costs on others. Firms will take actions to protect their brand name, which is to say they will attempt to control pollution.

Firms also have to attract workers in competitive markets where their plants are located. If the firm is a known polluter that shows little regard for the environment, employees who live with the environmental outcome will, at a minimum, demand higher wages.

Of course, a firm is actually an abstraction; fundamentally, it is a group of people. Directors, managers, officers, workers, and shareholders all live in the environment affected by firms. The larger the firm and the more extensive its operations, the more likely key individuals will be sensitive to the firm's environmental record.

Even if these forces had no affect on polluter decision makers, there are other constraints that do. Firms can be sued for damages caused by pollution. If an aquifer supplying drinking water is polluted by careless discharge of chemical wastes, the polluter

can be sued and forced to pay damages to those who are adversely affected.

Granted, there are problems in identifying causation, but common-law cases involving water, air, and land pollution go back for centuries. Because of common-law liability, firms carry general liability insurance to spread the potentially large costs over a longer period. Firms that are careless pay higher insurance premiums than firms with good records.

These market incentives that affect private polluters are not so effective with public-sector organizations, however. For example, the Department of Defense cannot be sued as easily for damages caused by sudden spills of hazardous waste as can a private chemical company. The Department of Defense, which happens to be a major polluter, does not bear the burden of paying premiums for environmental liability insurance nor run the risk of being run out of business by a more careful firm. Managers of public-sector organizations also are less likely to be fired for actions that might tarnish the reputation of their organization.[16]

None of this suggests that organizations—public or private—will achieve levels of environmental quality desired by any given individual or community. However, knowledge and use of incentives—the levers that direct behavior—can help policymakers to achieve their goals.

Affecting Polluter Behavior with Regulation

Recognition that incentives matter differently to the various parties to environmental control efforts is at the center of an evaluation of environmental regulation. Those seeking to improve environmental quality through regulation will focus on built-in incentives and call for programs that generate productive responses.

To illustrate, we begin with a story from fantasy land; that is, the assumptions included are not to be believed. Suppose a concerned interest group, say, a federation of national environmental organizations with power to bring about change, desires to produce as much environmental protection as possible. Given the resources they can muster, the group will insist that one simple principle be applied when policies are considered: the option that promises the largest amount of pollution control per dollar invested will be selected in preference to any more costly option. In other words, love of the environment translates to love of efficiency. Of course, this assumes that the interest group has information about costs, present and future strategies to be

taken by decision makers, and the results obtained. To have a goal other than efficient control would raise questions regarding the motivation of the group.

With full information, the task is conceptually simple. The interest group can make least-cost choices and instruct regulators to implement them. Command-and-control regulation will prevail.

However, the controls selected will not likely be applied uniformly across all pollution sources. Sources that produce small but costly-to-control amounts of pollution might not be controlled at all. Controls will be tighter where the payoff in pollution reduction is high. They will be lax where the payoff is small. Uniform reductions across all sources will occur only if all sources are identical with respect to control cost and the benefits obtained.

Since the interest group cannot possibly have complete information of the control costs for all sources of pollution and the various ways millions of decision makers may alter their behavior in response to environmental rules, other options must be considered for the purpose of minimizing cost. Somehow incentives must be offered to bring the best information to bear on each control problem.

Two more approaches generally surface. The interest group can insist that taxes or fees be placed on each unit of pollution discharged to the environment, which implies that discharge will be measured and taxes calculated effectively. To avoid the charges or minimize them, polluters will search for and implement lower-cost control techniques, alter production processes, and redesign products.

Alternately, the interest group can simply state their objective, the environmental goal to be achieved, set a performance standard and describe the penalties that befall those who fail to achieve the goal. Operators of plants are then left to their own devices for achieving the goal and will minimize costs while doing so. If one plant has lower control costs than a neighbor and environmental authorities allow for trades to be made among polluters, the neighbor can meet the performance standard by paying the low-cost operator to reduce more pollution. Since low-cost control capability is an asset, firms will invest in efforts to become leading producers of environmental quality.

The latter alternative implies a number of things. First, environmental quality will be monitored and data will be gathered and reported systematically to determine whether or not environmental goals are being achieved. Unfortunately, such routine

activities seem to be unexciting to environmentalists and place an unwelcome burden on the regulators. Next, the approach assumes that pollution fines would be calculated, collected, and adjusted for the sole purpose of regulating environmental pollution.

There is a fourth approach, but it can hardly be called regulation. The environmental group can work to define property rights to environmental assets. Future users of the environment will then have to bargain with the owner or pay damages when accidents occur. But we wish to focus on regulatory alternatives.

The three regulatory approaches—nonuniform command-and-control, emissions fees, and performance standards—impose differing administrative and compliance costs. Command-and-control allows no flexibility on the part of polluters who ordinarily have incentives to search out and implement lower cost controls. For that reason, command-and-control is the most costly option when measured in terms of pollution reduction per dollar spent.

Command-and-control is the most costly option in terms of pollution reduction per dollar spent. However, it is the least-cost option for the regulators who will administer the control program.

However, command-and-control is the least-cost option for the regulators who will administer the control program. They can simply state what must be done for each and every polluter and follow up to see that it is done. Pollution does not have to be measured and monitored. If the controls are in place and operating, pollution will be reduced. Unfortunately for those who care about results, this strategy emphasizes inputs, not outputs.

Emissions fees are less attractive and more complex for all parties—politicians, the regulators, and the regulated. They are taxes, and wary politicians have difficulty placing environmental control in the hands of the Internal Revenue Service. Taxes are technically difficult to design and operate. Somehow appropriate fees have to be determined, pollution has to be measured regularly so that charges are levied, and environmental quality has to be monitored to determine the results obtained from the fee schedule. When important conditions change, the fees have to be altered. However, fees induce cost-effective behavior on the part of polluters, while also providing additional revenue that can be used for public purposes.

The cost of obtaining environmental quality with fees will be less than through command-and-control. At the same time, the

job of the regulator is more complicated. They have to focus on output—environmental quality, not inputs. Managers at firms which are pollution sources also resist taxes for another reason: the taxes are imposed on all existing polluters for all pollution currently produced, so current and future costs will be higher.

Performance standards seem to be the simplest of all. Interest groups concerned about environmental degradation must surely have data on environmental quality. Otherwise, their concerns are baseless. Knowing the current situation and the desired outcome, the regulator can simply state the goal for major sources of pollution and sit back and wait. Those mainly concerned about environmental quality, the true environmentalists, will obviously monitor progress. Polluters will have complete flexibility when choosing how to reduce pollution. Costs will be minimized. However, it is costly for regulators to know current situations and to monitor progress.

This quick review of control strategies suggests strongly that less-than-perfectly-informed efficiency lovers will choose performance standards or emissions fees, just as will lovers of environmental quality.

Command-and-control will seldom be used, if environment lovers have their way. Using the command-and-control strategy efficiently means designing regulations for each and every source on the basis of control costs. Doing that requires regulators to have unattainable knowledge of outcomes before they happen. It is the highest-cost option.

However, if regulators, not environmentalists, have their way, command-and-control will rule the day—management costs are lowest for this alternative. Sadly, for those who are concerned about environmental quality, command-and-control emphasizes the methods used to control pollution, not the amount of pollution discharged to the environment. As a result, regulators tend to be technicians, not protectors of the environment.

Applying command-and-control, technology-based standards that require uniform reductions across all sources has been the strategy used by federal regulators for the last 20 years. Has regulator convenience simply swamped the influence of efficiency lovers? Is this the whole story? Not quite.

The Polluter-Politician Influence

Lovers of the environment, efficiency proponents, and regulators were the three players considered in the preceding review of options for controlling pollution. What about the interests of polluters and their political agents? How do they stand on the op-

tions? And what about the regulators? Does being the only game in town, a true monopolist in the field of regulation, make a difference?

Owners and managers of industrial firms generally will pay a premium to reduce uncertainty. Indeed, reduction of regulatory uncertainty was one of the forces that pushed regulation from the states to the federal government. Owners and managers also seek to avoid taxes, especially when the taxes apply to an input they have previously used at no charge. In addition, it is cheaper to monitor legislation at one national capitol than 50 state governments. A dislike of uncertainty and taxes and a search for lobbying efficiency causes businessmen who operate firms nationwide to place emissions fees low on the list of preferred control strategies. That leaves performance standards and command-and-control as the preferred options.

The embodied flexibility of performance standards makes them relatively attractive to decision makers at major sources of pollution, but uncertainty enters the picture if the performance goal can be set arbitrarily by regulators. All things considered, performance standards could easily be ranked second, if not first, in the minds of polluters. Command-and-control regulation emerges as the contender for first or second place.

How could the command-and-control strategy be modified to make it solidly number one in the minds of operators of polluting plants, as it is in the mind of regulators? As it turns out, the technology-based, command-and-control standards applied by EPA for water and air pollution are not uniform for all plants. New plants are required to meet higher standards than older ones. The higher new-source standards make it more costly for new competitors to enter markets, as well as for older firms to expand. EPA becomes an unwitting cartel manager. Only a monopoly regulator could accomplish such a feat.

Does being a monopoly regulator matter? Obviously, EPA could never enforce national rules if there were competing regulators. In addition, the absence of the spur of competition causes regulators to be less innovative and less concerned with costs and achievement of goals than would otherwise be the case. One additional problem arises when citizens have no way to compare the services they receive from one agency with another. There is no benchmark, no way to know if EPA is really doing a great job. Thus, when pollsters ask if EPA should be supported with a larger budget and tighter legislative standards, the public answers with a resounding, "Yes."

Comparing Costs

To say that performance standards will deliver more environmental control per dollar than command-and-control, source-by-source regulation calls for some evidence. What do we know about relative costs?

Research on air-pollution control provides the best documentation of potential cost savings to be achieved by freeing polluters to choose their own methods for reaching performance standards. Table 1 contains summary results for ten studies that focused on different industries, locations, and pollutants. The last column in the table provides key information on cost-effectiveness. That column shows the ratio of the cost of control when EPA command-and-control regulations are implemented relative to the cost obtained when polluters are allowed to hit the same control target by reducing more pollution where it is cheapest, and less where it is costly.

As indicated, the ratios range from as high as 22 to 1 to as low as 1 to 1. Even the worst case for flexible control shows that costs could be almost halved if producers were simply allowed to minimize cost. Operating expenditures for air-pollution control by all U.S. industries run approximately $7 billion annually.

Manufacturers annually spend about $6 billion on new air-pollution control equipment. If the cost reduction was in the range of just 20 percent, the savings would be large enough to fund the equivalent of one year's purchases of new air-pollution-control equipment for the U.S. petroleum industry.

Recognize that the savings occur annually. Each year of efficient pollution control could generate a one-year advance in the control of pollution for a major polluting industry. Efficiency should be an environmentalist's best friend.

Research on the relative costs of water-pollution control is much less extensive than that on air pollution, but still shows the potential for considerable savings when moving to performance standards that allow market-like responses. Ratios of the cost of command-and-control to performance standards run as high as 2.6 to 1 and as low as 1 to 1. No firm would deliberately choose to harm itself by selecting the high-cost path.

The Wastefulness of Superfund

The discussion of the regulatory dynamics associated with efforts to control air and water pollution covers a major part of the pollution-control story, but certainly not all of it. The nation's

Table 1

Empirical Studies of Air Pollution Control

Study	Pollutant	Industry	Ratio of Command-and-Control to Least Cost
Atkinson-Lewis	Particulates	Power and other	6.0
Diemer-Eheart	Sulfur dioxide	Power	2.0
Hahn-Noll	Sulfates	Steel, Petroleum, Power	1.1
Krupnick	Nitrogen oxide	Steel, Chemicals, Oil (200 sources)	6.0
Maloney-Yandle	Hydrocarbon	DuPont Chemical	4.2
McGartland	Particulates	All sources (Baltimore)	4.2
Palmer, *et al.*	CFC	Refrigeration (Plastics)	2.0
Roach, *et al.*	Sulfur dioxide	Power	4.3
Seskin, *et al.*	Nitrogen oxide	Power, Steel, Oil (Chicago)	14.4
Spofford	Sulfur dioxide	Power, Steel, Oil (Delaware Valley)	1.8
Spofford	Particulates	Power, Steel, Oil	22.0

Source: Tom Teitenberg, *Environmental and Natural Resource Economics*, second edition (Glenview, Ill.: Scott Foresman and Company, 1988), pp. 346-347 and correspondence with Teitenberg.

current efforts to control pollution follow a general blueprint that outlines future goals for cleanup and then prescribes in detailed fashion how the goals must be achieved.

The approach for handling hazardous waste follows this basic blueprint. Precise standards are prescribed for current producers, importers and exporters, and operators of transportation, storage, and disposal firms that might handle hazardous wastes.

Legislation dating back to 1972 addresses the issues. Chief among the laws are the 1974 Resource Conservation and Recovery Act (RCRA), the 1976 Toxic Substances Control Act, and the 1984 Amendments to RCRA. The legislation spawned rules that

cover handling hazardous wastes from birth to grave. Command-and-control rules the day.

Love Canal in Niagara Falls, New York, provided the impetus for passage of the 1980 Superfund law. Love Canal refers to an old toxic waste storage site used by Hooker Chemical Company for years. A long chain of events, dating back to the 1940s and involving threats of condemnation and the use of eminent domain, led Hooker to sell the hazardous waste site to the Niagara Falls School Board in 1953 for $1.[17]

At the time of transfer, so the seller claims, the deed to the land contained restrictions that described the hazardous waste site and proscribed its limited future use. As the story goes, all cautions were pushed to one side. The location of the hazardous waste site became part of a residential community. A combination of heavy rain and careless construction ruptured the sealed canal, leading to the migration of wastes from the storage canal to the basements of homeowners. The rest of the story is history.

Lifted by a tide of public opinion calling for control of old waste sites, Congress established a $1.6 billion cleanup fund, later raised to $8.5 billion. The fund is fueled primarily by uniform taxes imposed on chemical feedstocks and crude oil that are unrelated to the environmental record of the purchasers of the inputs. The 1980 law had the immediate and politically valuable effect of generating a large demand for huge pork barrel projects. Soon, old abandoned landfills joined more serious toxic waste sites on a growing list of Superfund targets.

EPA was instructed to use the fund for cleaning up abandoned sites, while simultaneously identifying those responsible for the site and suing them for the cost of cleanup. It makes little difference whether the owners of the waste site, those whose waste was deposited there, or those who transported the waste violated laws at the time a waste site was operating. The law treats all participants as potential violators. Any one or all of the parties that contributed to a waste site in any way, by any amount, at any time in its history can be held liable for the complete cost of cleanup—technically known as joint-and-several strict liability. Insurers that may have provided liability coverage at any time in the history of the site are also subject to claims from insured firms that become responsible for cleanup.

Superfund gives the impression that concerned citizens can clear away hazardous and other waste sites without bearing any cost once EPA designates a site for targeted action. So long as the perceived benefits are greater than zero, citizens will pre-

dictably lobby for cleanup. However, Superfund's record of accomplishments leaves little satisfaction for concerned citizens. As the Office of Technology Assessment (OTA) put it in 1988: "OTA's research, analysis, and case studies support the view shared by most observers—including people in affected communities and people in industry paying for cleanups—that Superfund remains largely ineffective and inefficient."[18]

The ineffectiveness and inefficiency of Superfund springs largely from its liability rules. Given the rule of joint-and-several liability, each potentially liable party will seek to shift the cost burden to some other party. In addition, each responsible party will sue its previous insurers in an effort to avoid the cost of cleanup. As it turns out, Superfund has become a Mecca for litigation, but is a desert for waste mitigation. Indeed, after making an extensive 1989 review of the program, OTA called the Superfund adversarial condition the "Superfund Syndrome."[19]

Since 1980, EPA has cleaned fewer than 35 sites. Some 1,200 sites have been targeted for cleanup, and more than 9,000 additional sites have been identified that could be added in the next 10 years.[20] In 1988, the average cost of cleanup was about $20 million. Just the toxicological assessment of a site, a step required to determine how to mitigate the waste dump, often runs $5 million. Based on these costs, some $24 billion will be needed to clean the currently targeted sites. OTA estimates that future costs could reach $500 billion unless major program changes are made.[21]

In many cases, more is being spent on litigation costs than would be required for hazardous waste cleanup.

In many cases, more is being spent on litigation costs than would be required for cleanup. In one extreme case, as many as 270 insurance firms are being sued by the companies involved in cleaning just one site.[22] Oddly enough, little attention has been focused on results, and no organized group has lobbied Congress to modify the joint-and-several liability rule that spawns so much costly litigation. Instead, the focus is on EPA programs, gaining increased budgets for staff, and finding ways to add sites to those that are not being cleaned under the current program.

Why has Superfund continued to operate in its current form, even after amendments were added? The old notion of somehow getting a free lunch from government supports an unmodified Superfund program. Citizens, who have little incentive to be informed of the resources wasted in administering the program, like the idea of getting some amount of benefit at little or no direct cost. Local hazardous sites, which in most cases pose very low risks, can be cleared away at almost zero local cost. Environmentalists favor the program because it is national in scope and focuses on emotionally charged issues. Large firms face a public relations nightmare if they suggest reform; they would be accused of self-dealing and environmental insensitivity, or worse. Regulators support the program because of its growth potential. When combined, all these motivations may be enough to keep Superfund in its present garb. With modification this high-cost, low-benefit program could accomplish far more.

Is there another approach, one that might lead to a more efficient use of resources? Two thoughts come to mind. First, old hazardous waste sites are a community problem, not a national problem. The benefits accrue to neighborhoods of people, not to some amorphous mass of unidentifiable Americans. That suggests cleanup decisions should be made locally and that funds for cleanup should come largely from the community. If a community realized that $20 million is needed to mitigate an old landfill, the community might decide to rebuild its public schools or build a community water system instead. The benefits of alternate methods for improving life would be weighed against the costs.

If the notion of federal action is persuasive on other grounds, high-cost litigation could be avoided by funding Superfund as a public works project with local matching of costs, which comes close to one of President Carter's original Superfund proposals. The public works idea was altered to a program funded by taxes in 1979.[23]

Since parties to Superfund projects are not, for the most part, past law violators, there is only one good reason to make them parties to suits: the federal government is unable to obtain the revenues for the program from other sources. Calling for a public works program opens debate about benefits and costs among taxpayers. If taxpayers are unwilling to fund the program as a public works project it seems reasonable to conclude that the case for Superfund is not compelling.

The End of the Story

The first part of the story is complete. We return to the question: why has so little been accomplished? The answer: the most costly form of environmental control has been applied, in spite of the availability of lower cost options.

Why has the most costly form of control been selected? The answer is that three major players in the story prefer it: (1) Regulators prefer command-and-control because life is made simpler for them; (2) Polluters can reduce uncertainty in decision making and the approach raises barriers to entry; and (3) Politicians gain from the opportunity to receive higher political payoffs.

Federal, as opposed to state, regulation gives environmental organizations a national market, increased political power, and funding.

What about the fourth major group—the environmentalists? Why aren't they efficiency lovers? In the final analysis, they apparently want environmental control at the federal level more than they want efficiency. Federal, as opposed to state, regulation gives environmental organizations a national market, increased political power, and funding. Federally enforced command-and-control requirements satisfy these goals. Efficient pollution control implies a decentralization of decision making that weakens environmentalist influence.

Are We Repeating Failed Approaches?

The logic of environmental protection implies that performance standards should be the dominant form of private-sector environmental regulation in a market economy, if protecting the environment is the principal goal. Command-and-control should be the fall-back strategy for privately owned firms and the dominant form of regulation for public enterprises.

We have just the reverse regulatory environment. Command-and-control dominates the private sector. Market flexibility is the fall-back position. Public enterprises enjoy more flexibility, more or less free of command-and-control solutions.

Opportunities for altering the arrangement are presented each time major environmental laws are written, amended, and revised. Why do the ineffective approaches of the past continue to be repeated? While the final Clean Air legislation changed significantly from the initial administration package, a review of President Bush's proposals for amending the Clean Air Act and the congressional action that followed may help to answer the question.

Amending the Clean Air Act

Seeking to be recognized as a friend of the environment, President Bush chose to announce his proposed Clean Air Act revisions from among the charred Douglas firs of Yellowstone National Park in June 1989.[24] The President's proposals called for reductions in pollutants that contribute to the formation of acid rain, additional steps to be taken toward the elimination of toxic discharges, and actions to control urban smog.[25] The latter proposals focused on automobile emissions and contained severe command-and-control rules.

Taking a page from the efficiency story, President Bush called for new flexibility in controlling conventional (nontoxic) pollutants. After working through a regulatory maze, plant operators would be able to minimize cost when controlling multiple sources of emissions. Allowing operators to demonstrate alternate methods for controlling emissions means that someone has to keep score on environmental quality. The goal of a cleaner environment, not the method of control, has finally entered the regulatory strategy.

The tradeable permit features of the acid rain provisions are an important acknowledgement of the potential effectiveness of a market-based approach to reducing pollution. The actual effectiveness of this system of tradeable permits may not be great, however, due to the labyrinth of state controls over public utilities and their treatment of sulfur dioxide permits in allowable rates.

The Bush proposals for the use of market-like controls, which reflected massive efforts by policy analysts, became a major influence in the formation of final legislation. At first blush, it appears that twenty years of failed command-and-control regulation took its toll. Unbelievably high control costs had stymied efforts to clean the air. What initially was speculation by economists about costs became facts reflected in operating statements and balance sheets. But this interpretation of the move to sanity has to do with efficiency. What about the special interest groups who were previously willing to bear high regulatory costs?

A more important element in explaining the call for flexibility was a reversal in special-interest support. By 1989, much of the nation's industrial base had been transformed from smokestack industries to high tech. Declining industries have little to offer in the struggle for legislation. Major corporations, previously protected from new competition by command-and-control regulation, found their positions reversed. Now they were seeking to

enter new product markets and industries. Costly regulation that had formed a comfortable protective tweed jacket became a hair shirt.

The positions taken by politicians at the local, state, and national level were also altered. Always sensitive to key industries and employment in their areas, politicians saw dramatic changes occurring as their economies restructured. Smokestack industries declined in the face of increased global competition. The services sector expanded. Incomes increased. The demand for environmental quality rose, and the cost—in terms of dislocated workers—fell.

Instead of being concerned primarily about plants running away to the Sunbelt, politicians in the frostier regions realized that international competition would ultimately determine the shape of the economy. Efficiency and cost-effectiveness became much more attractive positions to take when representing old-line firms seeking to survive in a global economy, which is to say the political cost of arguing for regulatory relief for old-line industries declined.

By the time of the 1990 Clean Air amendment debates, environmental groups were well aware that command-and-control had not worked effectively.

The clarion call for efficiency is seen in John Quarles' 1989 testimony before a congressional subcommittee in behalf of a coalition of labor unions, industry, and agriculture that was concerned about the forthcoming Clean Air Act amendments.[26] Quarles, a former Assistant Administrator of EPA, emphasized a number of concerns about the pending legislation but repeatedly called for regulations that reflected risk to human beings, not just risk to the environment.

Speaking for the coalition, Quarles urged Congress to set a maximum incremental cost for removing a ton of any given pollutant and to require the use of cost-effective controls. Arguing for market-like approaches that give firms incentives to minimize cost, Quarles noted that pending legislation applied those concepts as an exception, not the rule. Although unsuccessful in molding the final legislation, the testimony underlined the importance of maintaining economic growth through the use of common-sense rules that generate more clean air at lower costs.

By the time of the 1990 Clean Air amendment debates, environmental groups were well aware that command-and-control

had not worked effectively. They too became cautious supporters of alternate regulatory strategies, at least for nontoxic pollutants. At the same time, the groups held a general distrust for incentive-based regulation that failed to hold polluters' feet to the fire.

With various special-interest groups becoming more sensitive to costs and lack of progress toward environmental goals, President Bush and some key legislators were in a position to call for simple sanity, at least conceptually if not in fact: Let polluters find the simplest way to meet the requirements of law.

The path for the Bush proposals had been cleared partly by an important congressional report produced in 1988. "Project 88: Harnessing Market Forces to Protect Our Environment" was sponsored by Senators Timothy E. Wirth (D-CO) and John Heinz (R-PA).[27] The special report, which was authored by Harvard Business School economist Robert Stavins, reviewed past experience, emphasized the logic of market forces and flexible controls, and called for basic revisions in the Clean Air Act. To the extent that the report was a trial balloon, it soared high and traveled a great distance.

Command-and-Control Enters Again

The Bush administration proposals and the final congressional compromise for the Clean Air amendments included a heavy dose of market-like flexibility for broad categories of pollutants affecting old-line industries. True environmentalists should be thankful. However, the final package also included some of the most far-reaching command-and-control legislation ever proposed.

Command-and-control is found in two areas: toxic pollutants and auto emissions. The proposal for regulating toxic chemicals comes from EPA estimates that 1,500 to 3,000 fatal cancers can be associated with the 2.7 billion tons of toxic chemicals emitted into the air each year. But once again, technology is the basis for control, not improvements in health or reductions in risk. In congressional debate on the proposal, Representative Henry Waxman (D-CA) called for risks from toxic chemicals to be reduced to less than one in a million and for the closing of plants that could not meet his proposed standard in 16 to 22 years.[28]

The cost of meeting the proposed toxic emissions standard was estimated to be $2 billion annually, or $650,000 to $1.3 million per cancer avoided. Is this a reasonable allocation of resources to address cancer risk?

Two pieces of information provide some much needed perspective. First, a proposal by former Council on Environmental Quality economist Paul Portney calls for an annual expenditure of $25 million to reduce radon exposure. This proposal would lead to 10,000 fewer lifetime cancers from an easily identified population of 100,000 people, or a cost of $2,500 per cancer avoided.[29]

The second relevant piece of information relates to the findings of biochemist Bruce Ames.[30] Ames' research indicates that most cancers do not result from exposure to manmade carcinogens, but from human habits of consumption and behavior. The level of carcinogens and the risk exposures found in common foods and from treating drinking water exceed the levels associated with many toxic emissions targeted for control. The point is that instead of focusing on pollution as though all pollutants impose equal costs on human health, resources should be focused where they can do the most good.

The final Clean Air legislation predictably relies on technology-based standards to reduce toxic emissions, but this time includes a cost-effectiveness crumb.

The final Clean Air legislation predictably relies on technology-based standards to reduce toxic emissions, but this time includes a cost-effectiveness crumb. Command-and-control is the basic approach, but polluters can use alternate methods once they have reduced 90 percent of the controlled emissions. Even the crumbs can be large: the last 10 percent of the pollutants to be controlled is by far the most costly 10 percent to remove.

On the plus side of the regulatory ledger, the 1990 amendments continue the offset program, which allows expanding firms in nonattainment regions to acquire emissions reductions from other firms that may be able to abate these emissions at lower cost. The new law also allows electric utilities to trade emissions reduction credits to achieve the tighter controls on sulfur dioxide.

The final law contains a large dose of mandatory automobile "medicine" that will raise costs for all American motorists. Although everyone will pay, the auto-emissions problem is not a general one. Some 20 cities in the United States have truly severe ozone-related air quality problems. When the American Petroleum Institute examined EPA's data from ozone monitors and averaged the readings from all monitors in particular areas,

they found that with the exception of Los Angeles, all other noncompliance cities were meeting the ozone standard 99.47 percent of the time.[31]

The proposed rules require stricter auto-emissions standards for the national fleet and deny manufacturers the right to average emissions reductions across fleets. The most costly control opportunity for reducing emissions is treated exactly like the lowest-cost opportunity. Fuel refineries are required to produce and sell clean-burning gasoline. Auto producers must install onboard gasoline vapor controls while filling stations must install redundant vapor recovery systems at service stations. Finally, the auto industry is required to produce and market special California cars.

While progress toward cost-effective regulation is made at some margins, command-and-control always seems to dominate new regulating activities. Why is this the case?

The stories retold in this paper point to one answer: command-and-control will be used when large amounts of wealth can be protected or transferred by means of regulation. Regulatory relief in the form of performance standards and other incentive-based systems will emerge where international competition is severe, where industries are in a state of decline, and where pollution problems become so routine and lacking in controversy that they no longer attract public attention.

Some Final Thoughts

This paper has focused on one central question: why are national environmental goals so difficult to achieve? Answering the question required an examination of the federal environmental saga, which began in 1970 when a system of regulatory federalism was converted to a federal monopoly.

Several features of the story deserve to be emphasized:

- People who truly value the environment and wish to see it managed with owner-like concern will be strong supporters of efficient regulation. Efficiency translates to lower-cost methods for producing cleaner air, water, and land. With lower costs, greater progress can be achieved.
- Polluting firms in competitive markets have natural incentives to protect the environment. These incentives can be used when regulating private firms; they are not available for public-sector organizations. Differing methods of regulation are called for when seeking to affect diverse organizations.

- Competitive regulation among states has a natural tendency to generate more environmental quality per dollar than provided by a monopoly regulator. However, competitive regulation will not predictably provide uniform outcomes.
- The incentives faced by the major players in the struggle to affect federal regulation cause them to prefer higher-cost methods to control pollution that emphasize inputs—technology—instead of improvements in environmental quality. The political economy of the process predicts frustration.
- Command-and-control regulation will be the dominant mode of federal regulation so long as significant amounts of wealth are at stake. Severe international competition and economic dislocations can shift the regulatory mode away from command-and-control to a search for cost-effectiveness.

As stated at the outset, the 1990s will be the decade of the environment. The forecast is based on two initial conditions: there are environmental problems, and the political apparatus is yawning to respond to them.

There are prospects for improvements in environmental efficiency and quality. Interaction between true environmentalists, those who care deeply about real environmental progress, and pure efficiency lovers eventually focuses attention on the practical problem of getting the job done. In some cases, it may take 20 years or more before the practical aspects of the situation outweigh the emotional and political. Ultimately, efficiency and environmental progress coincide. But that coincidence of forces will always be threatened by new environmental issues and new opportunities for special interests to seek their own objectives while dressed in the garb of environmental protection.

Notes

1. On this, see "Firms Backing Off Environmental Claims," *The State* (Columbia, S.C.), December 30, 1990, p. 3-G.
2. See Charles McCoy, "California Air Quality Regulators Mull Plan to Radically Lower Emissions," *The Wall Street Journal*, December 16, 1989, p. A4; and Robert Reinhold, "California Stepping Up Pressure on Auto Industry to Cut Pollution," *The New York Times*, December 16, 1989, p. 1. The expansion of California's standards to other states and regions is described in David Stipp, "Massachusetts Follows California's Lead and Adopts Tough Auto-Emission Rules," *The Wall Street Journal*, December 28, 1990, p. 34.
3. See *Press Release*, Clean Air Working Group, Washington, D.C., March 23, 1989.
4. For additional discussion of these points, see Murray Weidenbaum, *Protecting the Environment: Harnessing the Power of the Marketplace* (St. Louis: Washington University, Center for the Study of American Business, 1989).
5. For a review of the pre-federal activity, see Bruce Yandle, *The Political Limits of Environmental Regulation* (Westport, Conn.: Quorum Books, 1989).
6. See Robert Quinn and Bruce Yandle, "Expenditures on Air Pollution under Federal Regulation," *Review of Regional Studies*, Vol. 16, No. 3, Fall 1986, pp. 11-16.
7. For a sample of the research, see Paul W. McAvoy, *The Regulated Industries and the Economy* (New York: W.W. Norton & Company, 1979), pp. 94-104; and Robert W. Crandall, *Controlling Industrial Pollution: The Economics and Politics of Clean Air* (Washington D.C.: The Brookings Institution, 1983), p. 19. Also see Lester B. Lave and Gilbert S. Owenn, *Clearing the Air: Reforming the Clean Air Act* (Washington D.C.: The Brookings Institution, 1981), p. 44.
8. See Robert W. Crandall, *Controlling Industrial Pollution.*
9. Ibid.
10. See H. Myrick Freeman III, "Water Pollution Policy," Paul R. Portney, ed., *Public Policies for Environmental Protection*, (Washington, D.C.: Resources for the Future, 1990), pp. 97-151.
11. Ibid., p. 120.
12. Ibid., pp. 125-126.
13. For details on the growth of environmental organizations that followed federal legislation, see Paul B. Downing, *Environmental Economics and Policy* (Boston: Little, Brown and Company, 1984), pp. 263-268.
14. On this point, see James T. Bennett and Thomas J. DiLorenzo, *Destroying Democracy, How Government Funds Partisan Politics*, (Washington, D.C.: Cato Institute, 1985), pp. 137-172.

15. See Benjamin Klein and Keith Leffler, "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, Vol. 89, 1981, p. 615.

16. See Murray Weidenbaum, *Protecting the Environment*, pp. 4-5. For a recent discussion of government immunity to environmental suits, see David Rapp, "Taking the King to Court," *Governing*, December 1990, p. 69.

17. For background, see Amy Dockser Marcus, "Love Canal Suit Threatens to Make Old Errors Costly," *The Wall Street Journal*, October 24, 1990, p. B-1; "People Are Just Beginning to Return to Love Canal," *The Greenville (S.C.) News*, December 23, 1990, p. 8E; Eric Zuesse, "Love Canal: The Truth Seeps Out," *Reason*, February 1981, pp. 17-33; "Superfund: A Game of Chance," *Natural Resources and Environment*, Vol. 1, No. 3, Fall 1985; Raymond A. Rae, "Hazardous Waste Pollution: The Need for a Different Statutory Approach," *Environmental Law*, Vol. 12, 1982, pp. 662-71; and Elizabeth Whelan, *Toxic Terrors* (Ottawa, Ill.: Jameson Press, 1985).

18. See "Are We Cleaning Up?" (Washington, D.C.: Office of Technology Assessment, 1988), p. 1.

19. *Coming Clean* (Washington, D.C.: Office of Technology Assessment, 1989), p. 3.

20. Ibid., p. 6.

21. Ibid.

22. See *Financing the Cleanup of Hazardous Wastes: The National Environmental Trust Fund* (New York, N.Y: American International Group, 1989).

23. See "Carter Proposes $1.6 Billion Fund to Fight Chemical Waste Hazards," *The New York Times*, June 14, 1979, p. 1.

24. See George Bush, "Remarks by the President in Announcement of the Clean Air Act Amendments" (Washington, D.C.: Office of the Press Secretary, The White House, June 12, 1989).

25. See U.S. Environmental Protection Agency, "Clean Air Amendments of 1989: Section-by-Section Analysis" (Washington, D.C.: U.S. Environment Protection Agency, External Affairs [A-102], July 1989). Also, see Kenneth Chilton and Anne Sholtz, *Battling Smog: A Plan for Action* (St. Louis: Washington University, Center for the Study of American Business, 1989).

26. See National Environmental Development Association, Washington, D.C., testimony of John Quarles, Chairman, National Development Association's Clean Air Act Project, before the Subcommittee on Health and the Environment, House Committee on Energy and Commerce, May 24, 1989.

27. See Robert Stavins, "Project 88: Harnessing Market Forces to Protect Our Environment: Initiatives for the New President. A Public Policy Study Sponsored by Senator Timothy E. Wirth and Senator John Heinz" (Washington, D.C., December 1988).

28. See George Hager, "Bush's Plan for Cleaner Fuels Scaled Back by House Panel," *Congressional Quarterly Weekly Report*, Vol. 47, October 14, 1989, pp. 2700-2701.

29. See Paul Portney, "Reforming Environmental Regulation: Three Modest Proposals," *Issues in Science and Technology*, Winter 1988, pp. 74-81.

30. See Bruce N. Ames, Renae Max, and Lois Swirkey Gold, "Ranking Possible Carcinogenic Hazards," *Science*, Vol. 236, April 17, 1987, pp. 271-280. For a survey of other related research and a thorough discussion of the development of false concerns regarding environmental risks, see Margaret Maxey, *Managing Environmental Risks: What Difference Does Ethics Make?* (St. Louis: Washington University, Center for the Study of American Business, 1990).

31. For detailed discussion, see Melinda Warren and Kenneth Chilton, *Clearing the Air: Regulating Ozone in the Public Interest* (St. Louis: Washington University, Center for the Study of American Business, 1988), p. 9.

Reprinted with permission of Resources for the Future from Resources, No. 102, Winter 1991.

The expanding role of environmental interests in agricultural policy

Katherine H. Reichelderfer

Despite the potential of modern agriculture for harming the natural environment, agricultural activities in developed economies have been regulated quite differently with respect to environmental consequences than have the activities of the mining, manufacturing, energy, and construction sectors from which similar environmental damages arise. Understanding the factors that have influenced how agriculture is regulated to achieve environmental goals provides important clues to the success of future agroenvironmental policy efforts and to the changing role of agriculture in a growing economy.

The contributions of modern agricultural practices to the depletion of natural resources and the degradation of the environment have been well documented. Modern agriculture is associated with depletion of underground water sources, degradation of soil resources, contamination of surface and ground water with substances that run off or percolate from agricultural land, destruction of wildlife habitat, and endangerment to biodiversity. Agriculture is really no different from other industries in that it generates waste materials. But unlike other sectors of the economy—in which pollution has increasingly been controlled through federal standards, fees and fines, restrictions, or (more recently) market-based incentives—agriculture is unique in having engendered relatively less federal government intervention with respect to its environmental consequences. When intervention has occurred, it has been achieved—more often than in other industries—through mechanisms that increase rather than decrease producers' incomes.

Federal agricultural resource and environmental programs have existed in the United States since the 1930s. As originally established and traditionally maintained, these programs have been largely voluntary and have relied on the use of positive incentives to achieve their goals. For instance, agricultural landowners have long had access to the Agricultural Conservation Program, the Soil Conservation Service, and the Great Plains Conservation Program, which, along with similar programs, offer technical and financial assistance for voluntary initiation of soil and water conservation planning and implementation at the farm level. The current Conservation Reserve Program—like its predecessor, the Soil Bank Program of 1956—allows farmers to receive annual rental payments from the federal government for retiring land on which cultivation may pose environmental hazards. Such programs mutually benefit the environment and the farmers who choose to participate in them.

Federal agricultural resource programs have traditionally relied on positive incentives to encourage resource conservation.

Only since 1985 have some penalties been added to the incentives offered to farmers for resource conservation. The 1985 Food Security Act prohibits farmers from receiving benefits through commodity, farm credit, and related farm programs if the farmer drains wetlands or cultivates erodible land without having a conservation plan in place. While involving penalties of a sort, these compliance programs are also voluntary. Any participant in a farm program is free to drop out of the program rather than comply with its environmental requirements. As conditions in agricultural markets improve (making farm program safety nets less necessary) or the level of farm program benefits declines, the penalty for noncompliance with environmental guidelines can rapidly diminish.

Despite the historical tendency for U.S. policy to treat farmers as willing stewards of the land, some environmental legislation since the early 1970s has directly affected farming. In particular, the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), under which pesticides are regulated, has reduced the number and variety of alternative substances available to farmers for pest control. FIFRA is unique in two ways: it addresses the safety and environmental effects of the use of products rather than the making of products; and it is among the few federal environmental acts that mandate a balancing of benefits and risks in decisions to ban or restrict a product. As a result of these characteristics, FIFRA has had limited economic effects on farmers. The need to balance the benefits of a pesticide's use against the risks that use poses has meant that, in most cases, uses have been banned only when there were close chemical or nonchemical substitutes available. (For an examination of how the U.S. Environmental Protection Agency [EPA] balances the risks and benefits of pesticides, see the article "An analysis of EPA pesticide regulation" in this issue.) Only those farmers who were especially dependent on a pesticide for a use that was banned have suffered economic losses. The farming sector at large may even have gained from the increases in revenue that come about when a regulatory action decreases the production of only some farmers and subsequently stimulates rises in commodity prices that benefit all producers. Paradoxically, costs have been restricted mainly to consumers, whose public health FIFRA was designed to protect.

Agriculture has thus far been overlooked by or excused from meeting the requirements of most other environmental policies. Federal policy regarding water quality and toxic substances has focused on point

sources of pollution, postponing the more difficult problem of nonpoint sources, mainly agricultural. For instance, the Clean Air and Water Quality acts impose technology-based standards that affect the location, configuration, operating conditions, and costs of virtually all industrial and public utility facilities, yet they place no limits on effluents or emissions from agricultural and other nonpoint sources of air and water pollution. Similarly, industries and municipalities spend an estimated $23 billion to $30 billion annually to comply with the 1972 Federal Water Protection Control Act, yet that act authorizes federal subsidies to help states plan and farmers adopt water quality management strategies for which there are no associated standards.

The unique treatment of agriculture is apparent in a range of resource conservation and environmental policies. During the energy crisis of the 1970s, agriculture was routinely exempted from controls on the price and availability of fuels. At present, agricultural landowners whose practices have rendered land unusable (through accumulation of salts, heavy metals, or other toxic substances in the soil) are not subject to any law equivalent to that which requires land users to return areas scarred by surface mining to their original condition at private cost. Thus, while the centralized or command-and-control approach to environmental policy has been given precedence in nonagricultural sectors, incentive-based and subsidy approaches have predominated in the agricultural sector. Why is this so?

Is agriculture special?

In some respects, unique approaches to minimizing potential environmental damages from farming might seem warranted. First, there is more uncertainty about the nature of nonpoint sources of pollution than there is about readily observable point sources. Contaminants from nonpoint sources cannot easily be traced either to agricultural activities (some could originate naturally or in a golf course or home garden) or to a specific parcel of land or land operator. Thus regulations based on limitations on or requirements for certain agricultural practices—with or without associated fees, fines, or taxes—are more difficult to design than are regulations for point sources of pollution, which can be monitored.

Second, in farming, individuals are making use of privately owned resources. In other industries, where environmental concerns focus on the private use of public goods such as air and water for discharge, there are few counterparts to the property rights issues involved in decisions about how farmers use their own land. Questions about whether farmers' property rights might be violated by environmental regulation that acts upon the public's right to an undegraded environment are also complicated by a pervasive paternalism toward American farmers. The special reverence with which small farms and family farms are regarded is not common to most other groups of producers, especially in the manufacturing sector, and creates a public desire to resolve environmental problems without hurting farmers.

Finally, agriculture in the United States and other developed countries is highly protected through a network of farm income and price support policies. The distortions created in agricultural markets by such intervention can offset regulatory incentives for changes in agricultural technology that are environmentally beneficial. In other words, the maintenance of farm income through production and price controls makes regulations that raise the cost of environmentally damaging farm practices weaker in agriculture than in other markets that remain unregulated.

Despite these constraints, a range of policy options for more efficient control of agricultural sources of pollution is readily identifiable. For instance, the sales price of agricultural chemicals known to pose risks could be taxed at rates consistent with their social costs. Farm income support could be linked to environmental stewardship instead of to commodity production levels. Markets for permits to use certain agricultural chemicals in closed biosystems could be established. The fact that such options have not been implemented suggests that there are other factors influencing the direction that agroenvironmental policy has taken to date. Research at Resources for the Future demonstrates that it is largely broader political and economic trends that have most influenced past patterns and that are likely to change future policy approaches to environmental regulation in American agriculture.

Critical factors

Trends in the value of gains and losses as perceived by public interests on the one hand and agricultural interests on the other, and the subsequent influence of competing interests on the policy process, best explain policy choices for environmental regulation of agriculture. How the public and its representatives view and value the goods arising from agricultural activities depends on many factors, one of which is economic growth.

Rising per capita income in the developed economies increases the level of demand for goods such as environmental quality, recreation, and aesthetics at a greater rate than it does the level of demand for basic goods like food and fiber. Demographics reinforce this demand as an aging population with greater leisure time exerts pressure for clean recreational and retirement sites. Consistent with these trends is a generally increasing valuation by the public of the environmental costs arising from agricultural production. As perceived costs rise, the proclivity to protect agriculture may decline in relation to the demand for environmental regulation of agriculture.

Whether federal legislation favors agricultural or environmental protection depends partly on relative farm income.

At the same time, the size of the agricultural sectors of developed economies tends to decrease as the economy continues to grow. The decline in the number of farmers implied by this phenomenon actually increases rather than decreases the political influence of agricultural interests; as the size of the agricultural community decreases, each member of that community has a larger personal stake in decisions

about agroenvironmental policy. Thus economic growth can create tension and increased competitiveness between groups that have invested in agriculture and those that demand higher levels of environmental protection.

The response of legislators to these ofttimes competing interests is in part a function of how well farmers are faring in relation to the rest of the economy. There seems to be a strong correlation between relative farm income and the passage of agroenvironmental legislation, as well as the form that legislation takes. When farmers are perceived as being richer than the rest of us, it is more likely that restrictive legislation in the manner of FIFRA will be passed. Conversely, when farmers are suffering financially in relation to the rest of the economy, as in 1985, legislation addressing agroenvironmental problems is more likely to take the form of a subsidy that enhances farm income (see figure 1).

Because of farm program payments, fluctuations in the extent to which the capacity to produce agricultural goods corresponds to demand for those goods vary in a different way than does relative farm income over time. Yet the willingness of federal legislators to enact laws protective of agriculture or the environment also appears to be related to the size of surpluses. The larger those surpluses, the more likely it is that legislation favors environmental interests over agricultural ones.

The political strength of environmental interest groups lobbying to represent public interests in agroenvironmental quality is also an important factor in whether legislation favors agricultural protection or environmental protection. The number of environmental groups involved in agricultural policy, their membership, and the resources available to them have grown dramatically over the last two decades. As environmental groups become increasingly efficient at exerting pressure, the degree to which environmental interests influence policymaking may rise. Independent of the activities of these groups, the rapid accumulation of information on the levels and possible consequences of environmental contaminants from agricultural sources is likely to raise the public's demand for environmental regulation of agriculture.

Implications for the future

Many of the factors that have affected the level and direction of U.S. agroenvironmental policies in the past are still in evidence or are gaining in influence today. The long-term outlook for the economy is continued growth, implying a continued general shift of public preference toward environmental regulation of agriculture. Relative farm income is on the rise, reinforcing trends that place greater weight on environmental interests in policymaking. Furthermore, the size and influence of environmental and other public interest groups concerned with agroenvironmental policy is growing.

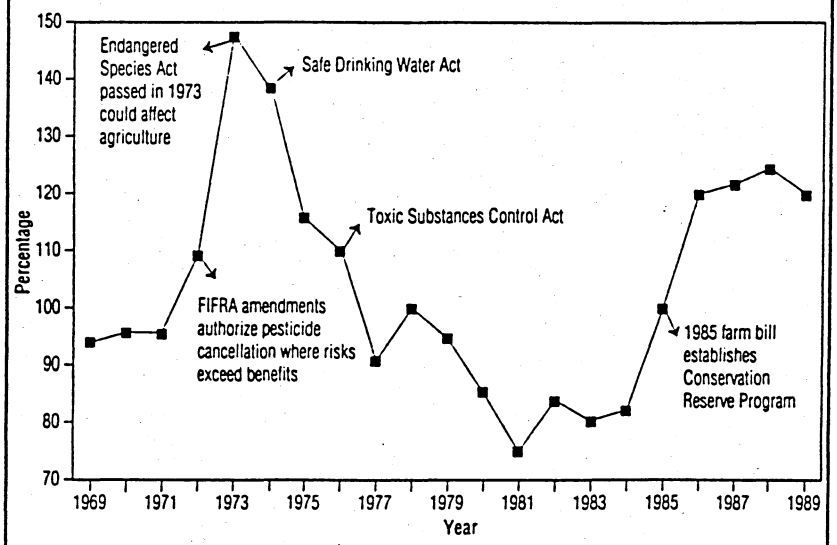
Other factors may accelerate the shift toward regulation of agriculture for the purpose of environmental protection. One is the changing composition of the House of Representatives, which with each redistricting in recent years has lost some proportion of representation from rural and farming-dependent regions. Others include increasing agricultural productivity, shifts in agricultural trade patterns, and the proliferation of environmental regulation at the state level.

U.S. agricultural productivity increased an average of 2 percent per year during the 1980s. As the efficiency of production continues to increase, the costs to the public of agricultural programs will rise (unless demand increases at the same rate—a phenomenon not expected in the short run). These rising costs will likely decrease the political strength of agricultural interests relative to that of taxpayers, implying a future decrease in agricultural protection relative to environmental protection.

As for trade, current negotiations under the General Agreement on Tariffs and Trade (GATT) are attempting to decrease the level of subsidies to domestic agricultural producers while exempting agricultural programs that are oriented toward environmental protection or conservation from similar cuts. If successful, GATT reforms could promote agriculture as an industry that must be more responsive to environmental concerns.

Regardless of the outcome of GATT negotiations, continued or increased reliance by U.S. agricultural producers on the export market will reinforce pressure for reforms in the agricultural sector. This is because the costs to the public of agricultural support tend to be greater in the relatively price-sensitive export market, and because the direct and indirect costs of environmental degradation associated

Figure 1. Average farm income as a percentage of average U.S. income



Note: Laws potentially costly to agriculture are more common when farm income is high relative to average income.

with production are not passed on to foreign consumers.

Finally, the number of environmental standards established, laws enacted, and programs implemented at the state level increased dramatically during the 1980s. This increase was partly in response to federal mandates for states' development of customized environmental protection efforts, and partly a result of public clamor and responsive legislatures in progressive states. At present, a fair proportion of state environmental legislation specifically targets or has direct implications for agriculture. In California, Proposition 65 may restrict some uses of agricultural pesticides otherwise allowed under FIFRA. In Connecticut, liability has been imposed on individuals (including farmers) shown to have contaminated drinking water sources. In Iowa, fertilizers and pesticides are taxed to raise revenues for improvements in water quality.

Great variation in the environmental laws of individual states can create problems for agricultural industries that operate nationally. If and when such variation becomes a serious constraint, the agribusiness industry itself may exert pressure for federal provision of some uniformity—a move that suggests the possibility of increased centralization of agroenvironmental policy in the future.

As the U.S. economy grows, new information on the environmental effects of agriculture is made available, and existing environmental legislation is applied to nonpoint pollution sources, the level of environmentally motivated government intervention in agriculture will begin to approach that in other industries. This is not likely to happen overnight or in a continuous fashion. Just as a generally growing economy experiences periodic recessions and expansions, the influence of economic factors on environmental regulation of agriculture is likely to wax and wane. An example of this is the recent defeat of the Big Green initiative in California, public support for which was seen to diminish in direct response to the developing recession in the state's economy.

There is little chance that agricultural protection will be overrun by environmental protection in the near future; only that the level of agricultural protection relative to environmental protection will decline. The form that new legislation takes will depend on the unique characteristics of agriculture, the public's view of agriculture, and the influence of private interests. However, in the future it is increasingly likely that the agricultural sectors of the United States and other developed countries will be affected by a centralized form of environmental regulation. Moreover, federal budget deficit problems in the United States will make it increasingly difficult to address agroenvironmental problems chiefly through subsidy programs, as has been typical in the past. The choice that farmers, agribusiness, and policymakers face is whether to increase environmental regulation of agriculture through a command-and-control approach or a market-based one. Experience in other industries suggests that the more efficient market-based approach has greater potential for creating a climate under which production that is sensitive to environmental protection is also good for agricultural business. ■

Katherine H. Reichelderfer is a senior fellow in the National Center for Food and Agricultural Policy at RFF.

An analysis of EPA pesticide regulation

Maureen L. Cropper, William N. Evans, and Paul R. Portney

Does the U.S. Environmental Protection Agency balance risks and benefits in regulating pesticide use? A recently completed study of the agency's decisions regarding cancellation of some registered food uses of pesticides suggests that it does. The study also finds that the agency's regulation of pesticides is influenced by special interest groups—a fact that some economists and risk analysts may find discouraging and others encouraging.

Pesticides are at least partly responsible for the large increases in agricultural productivity that the United States has enjoyed since World War II. However, their use may pose risks to the environment—to ground or surface water or to wildlife habitat—as well as to workers who apply them and to consumers who eat pesticide residues on food. It is the job of the U.S. Environmental Protection Agency (EPA) to regulate pesticide use to manage these risks. Specifically, the EPA decides whether a pesticide can be used and, if so, what residues may safely remain on foods. According to the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), the agency makes the first decision—whether to allow a pesticide to be used—by assessing whether the pesticide imposes "unreasonable adverse effects on the environment." Once it approves a pesticide for use, the EPA must act to "prevent any unreasonable risk to man or the environment, taking into account the economic, social, and environmental costs and benefits of the use of [the] pesticide." This implies that a pesticide should be banned only if the risks of its use outweigh the benefits.

In the past the EPA has been criticized for its decisions to ban or not ban pesticide uses. Environmental groups cite the agency's failure to ban pesticides, such as dicofol, that pose risks to wildlife. At the same time, some economists allege that the agency pays too much attention to pesticide risks to farmworkers and consumers. They claim that the EPA has reduced the risk of cancer to these groups only at very high cost. In the same vein, farmers have been quick to point out that banning a pesticide can be very costly to them, especially when few substitute pesticides are available.

These criticisms raise several questions about the EPA's recent decisions regarding pesticides. First, in deciding whether or not to ban a pesticide, has the agency balanced the health risks of pesticide use against the benefits, as it is required to do under the law? Have the costs of banning a pesticide been considered as well as the risks, or does the EPA always ban pesticides when risks of cancer to farmworkers or consumers exceed some threshold level, regardless of how much such a ban costs? In the area of risk regulation, the notion that substances posing high risks to any one person should always be banned—even if the cost is high—is a common one. The other side of this argument is that substances posing low risks should never be banned, even if it is inexpensive to do so. Has the EPA acted in accordance with this argument in regulating pesticides, or has it balanced costs against benefits at all risk levels?

Second, has the EPA been responsive to the interests of environmental groups in regulating pesticides? When organizations such as the National Audubon Society or the Environmental Defense Fund comment publicly in support of cancelling a registered use of a pesticide, do their comments increase the chances that the EPA will ban the pesticide? In light of the history of U.S. pesticide regulation, the EPA's responses to such comments are particularly interesting. Before the EPA was created, pesticides were regulated by the U.S. Department of Agriculture. Transfer of responsibility for pesticide regulation to the EPA was prompted in part by the view that the Department of Agriculture was not sufficiently responsive to environmental and consumer groups.

Third, what other political considerations have influenced pesticide regulation? Has participation in the regulatory process by growers' organizations or their representatives decreased the likelihood that a pesticide will be banned? Are pesticide decisions sensitive to political concerns, given that the administrator of the EPA is a political appointee?

In an attempt to answer these questions, researchers at Resources for the Future and the University of Maryland undertook a study of the EPA's decision to ban or not ban each of the 245 registered food uses of the nineteen cancer-causing, food-use pesticides that went through the agency's special review process between 1975 and 1989 (see table 1). The EPA cancelled 39 percent of these food uses. The study explains the pattern of cancellations as a function of the risks and benefits of pesticide use, as well as of political variables.

Registering pesticides

If the EPA suspects that use of a pesticide imposes unreasonable adverse effects to human health or the environment, it must subject the pesticide to a special review before banning it. During this review, the agency examines the risks and benefits of the pesticide's application for each crop on which the pesticide is used. The EPA next makes a preliminary judgment, crop by crop, regarding cancellation of the pesticide. Then follows a period during which members of the public, including environmental groups and growers' organizations, may comment on the proposed decision. At the end of the comment period a final decision (Notice of Final Determination) is issued.

During the special review process, the EPA considers not only the ecological effects of pesticides, such as whether a particular pesticide is toxic to wildlife or is likely to contaminate ecologically fragile environments, but also the risk of cancer to persons who mix and apply pesticides and to consumers who ingest pesticide residues on food. Evidence that a chemical is carcinogenic usually comes from tests on animals; these tests produce an estimate of the relationship between dosage of a pesticide and lifetime risk of cancer. This estimate is extrapolated to humans and multiplied by an estimate of human dosage (exposure) to calculate the incremental lifetime risk of cancer to a farmworker or consumer from that exposure.

Incremental lifetime cancer risks are typically much higher for pesticide applicators than for consumers of food products. For example, for the pesticides studied by researchers from RFF and the University of Maryland, the highest lifetime cancer risk for pesticide applicators is 0.10 for ethylene dibromide (EDB) when used in spot fumigation—that is, as a result of applying this pesticide, the applicator's lifetime risk of cancer in-

Table 1. EPA Pesticide-Use Decisions Studied by RFF/University of Maryland

Active ingredient (generic names)	Year of decision	No. of food-use registrations	No. of final cancellations
Dibromochloropropane (DBCP)	1978	12	12
Amitraz	1979	2	1
Chlorobenzilate	1979	3	2
Endrin	1979	8	4
Pronamide	1979	4	0
Dimethoate	1980	25	0
Benomyl	1982	26	0
Diallate	1982	10	0
Oxyfluorfen	1982	3	0
Toxaphene	1982	11	7
Trifluralin	1982	25	0
Ethylene dibromide (EDB)	1983	18	18
Ethalfluralin	1983	3	0
Lindane	1983	8	0
Silvex	1985	6	6
2,4,5-T	1985	2	2
Dicofol	1986	4	0
Alachlor	1987	10	0
Captan	1989	65	44
Totals		245	96

Photo courtesy of the Southern Agricultural Chemicals Association

Health risks to pesticide applicators appear to weigh heavily in decisions by the U.S. Environmental Protection Agency to ban a pesticide use.

ers' organizations and environmental groups—the two special interests having the most clearly defined objectives—it is clear that growers were more likely to comment when losses to producers from pesticide cancellation were high, while environmental groups were more likely to comment when a pesticide posed a danger to wildlife. In addition, whether or not special interests commented on proposed decisions was influenced by who the current administrator of the EPA was. When Anne Burford was the agency's administrator, no environmental groups commented on the 75 food-use decisions proposed during her tenure, possibly because they thought they would not receive a sympathetic hearing. By contrast, growers' organizations, anticipating more sympathetic treatment, were more likely to comment during Burford's tenure. In fact, half of all comments by growers' organizations occurred during the two years that Burford was administrator of the EPA. Thus Burford's tenure at the EPA seems to have had a negative effect on the likelihood of pesticide cancellation, due to the fact that no environmental groups intervened during her administration, whereas grower organizations were more likely to have intervened.

An RFF/University of Maryland study indicates that the EPA has balanced the benefits of pesticide use against health risks at all levels of risk.

Cause for comfort and concern

The results of the RFF/University of Maryland study offer both comfort and concern to persons interested in environmental regulation. With respect to comfort, it appears that the EPA is indeed capable of making decisions that balance risks and benefits, as the law requires. The study demonstrates that risks to human health, the environment, or both increased the likelihood that a particular food use of a pesticide would be cancelled by the EPA, while the larger the benefits associated with a particular use, the lower the likelihood of cancellation.

On the other hand, the study's results provide some cause for concern. For instance, researchers found that the implicit value of a statistical life in the 245 regulatory decisions studied is $35 million. This value is based on a risk estimate that, by design, is much more likely to be too high than too low. In other words, the value that the EPA implicitly placed on reducing risks to pesticide applicators may be considerably in excess of $35 million per life saved. Since there are a variety of policy measures, environmental and otherwise, that are capable of reducing cancer cases at much lower costs, it might be possible to reduce the cancer rate through a reallocation of resources.

It is less clear how to view the study's findings concerning the influence of interest groups on the cancellation of uses of pesticides on food. Clearly, intervention in the regulatory process—by both business and environmental groups—affects the likelihood of restrictions on pesticide use. Other factors being equal, intervention by environmental groups has about twice the impact on the likelihood that a pesticide will be banned as intervention by growers does; however, the combined impact of growers and academics acting on behalf of growers is approximately equal to that of environmental groups.

To those who view pesticide regulation as the proper province of economists and risk analysts alone, these findings may be discouraging. On the other hand, those taking the view that regulation—like government taxation or spending—is inherently a political act may find it encouraging that affected parties not only participate actively in the regulatory process but do so quite effectively. ■

Maureen L. Cropper is associate professor of economics at the University of Maryland, College Park, and a senior fellow in the Center for Risk Management at RFF. William N. Evans is assistant professor of economics at the University of Maryland, College Park. Paul R. Portney is vice president of and senior fellow at RFF. This article is based on research coauthored with Stephen J. Berardi and Maria M. Ducla-Soares, students at the University of Maryland, College Park.

A list of biographical references follows:

1. Cuzan, Alfred. "Appropriators versus Expropriators: The Political Economy of Water in the West," Chapter 1 in **Water Rights: Scarce Resource Allocation, Bureaucracy, and the Environment,** edited by Terry L. Anderson, Pacific Institute for Public Policy Research, 1983: 13-44.

2. Rucker, Randal R. and Price V. Fishback. "The Federal Reclamation Program: An Analysis of Rent-Seeking Behavior," Chapter 2 in **Water Rights:** (same as item 1):45-80.

3. Gardner, B. Delworth. "Water Pricing and Rent-Seeking in California Agriculture," Chapter 3 in **Water Rights:** (same as item 1):83-112.

4. Berkman, Richard L. Berkman and W. Kip Viscusi. **Damming the West,** Grossman Publishers, New York, 1973.

5. Gardner, B. Delworth. "The Untried Market Approach to Water Allocation," Chapter 7 in **New Courses for the Colorado River,** edited by Gary D. Weatherford and F. Lee Brown, University of New Mexico Press, Albuquerque, 1986:155-176.

6. Ingram, Helen M. Lawrence A Scaff, and Leslie Silko. "Replacing Confusion with Equity: Alternatives for Water Policy in the Colorado River Basin," **New Courses for the Colorado River,** (same as item 5):177-200.

7. Shanks, Bernard. "Dams and Disasters: The Social Problems of Water Development Policies," in **Bureaucracy vs. Environment,** edited by John Baden and Richard L. Stroup, The University of Michigan Press, Ann Arbor, 1981:108-123.

8. Howitt, Richard E., Dean E. Mann, and H. J. Vaux, Jr. "The Economics of Water Allocation," in **Competition for California Water,** edited by Ernest A. Engelbert and Ann Foley Scheuring, University of California Press, Berkeley, 1982:136-162.

Life in the Desert with Obsolete Water Institutions[1]

B. Delworth Gardner
Professor of Economics
Brigham Young University

I approach my topic this morning with some trepidation. Few topics are more controversial in an arid state than water policy. Mark Twain, that great American humorist, was at least half right when he said that "whiskey is for drinkin' and water is for fightin'. He was right about the water half.

But in pointing out some of the problems with water policy I am really striking as well at the hundreds of similar government programs that are designed to redistribute income and which in the process destroy national wealth and reduce our collective standard of living. I believe that slow rates of economic growth over the past couple of decades are largely the result of these types of policies. But I must not get ahead of my story.

During the period of settlement of the West, water institutions evolved that served the region well. As everyone knows, human beings and most animals cannot live for more than a few days without water. Perhaps it is this "indispensability" that gives water a privileged, almost mystical position among natural resources. Almost every State constitution declares the corpus of water to be owned by the people of the state and appropriators can obtain only a use-right to it. In addition, in arid regions most of the important food crops must be grown under

---

[1] Address presented at University Forum, Brigham Young University, May 28, 1991.

irrigation. Since the majority of people were involved in agriculture, control of irrigation water amounted to the control of most of society's wealth. So when our pioneer forefathers and others came to these arid mountain valleys, a top priority was to harness water resources and apply them to parched land so that crops could be grown and life sustained. My own great, great grandfather, Archibald Gardner, came in 1847 and built mills in several communities in Utah and Wyoming. These mills utilized water power to turn the great wheels that produced flour and lumber.

The United States Constitution was silent about water as a Federal concern, so water law and administration were left to the states. In the Eastern and Southern regions of our country where European man first settled, water was not needed for irrigation and the common-law riparian doctrine was utilized to allocate it. This doctrine permitted those located along a river to have a right to use water so long as other riparian users were not negatively affected. But this doctrine obviously would not work for mining and irrigation in the West where water had to be moved away from land contiguous to the water source, sometimes for long distances through canals and ditches to the crops and mines that utilized it.

In this arid region the doctrine of prior appropriation was developed. Certain procedures had to be complied with like filing for a right with a regulatory agency of state government. Two important ideas are contained in the appropriation doctrine:

one, the notion of "beneficial use" which required that water be put to some use deemed to be beneficial, such as culinary use, or watering livestock, or irrigating crops. It was thought that this doctrine would prevent appropriation in advance of actual use and thus would minimize speculation in water. The second idea was "first in time, first in right." Those who applied for water rights earliest in time received a superior right. Senior rights were to be fully satisfied before junior appropriators could receive any water. This law greatly reduced risk and provided security of tenure for water users. They knew exactly what resources they had and orderly development could and did occur. Of course, establishing a legal right was one thing; protecting it against theft was quite another. Water had to flow through the property of many landowners, sometimes many miles to its ultimate destination and use. Ditch riders were hired and conflict was often intense, and thus Twain could rightly say that water was for fightin'.

Irrigators organized themselves into mutual irrigation companies. Significantly, these were private organizations which issued shares on the basis of acres irrigated. These industrious farmers used their own money and labor resources to develop irrigation facilities, divert water, and produce wealth for themselves. Some of these original mutual companies have delivered water for more than a century. The overall result was efficient development and allocation of water and rapid economic growth. Unfortunately this happy outcome was destined not to

last.

Near the end of the 19th Century, national attention turned to the frontier in the West for a variety of reasons. Vast herds of cattle and sheep swept onto the public lands and the ranges soon were overgrazed and became seriously depleted. Watersheds for developing communities were destroyed and flooding was a serious problem. Galloping on the scene came Theodore Roosevelt's progressives and scientific management of resources was supposed to cure all the region's problems. The National Forests were reserved and grazing permits were issued to local stockmen. The National Parks were created beginning with Yellowstone in Wyoming in 1872. Two immensely important institutional developments occurred in water policy during this period that were to have profound impacts: the National Reclamation Act of 1902 and the Colorado River Compact of 1922.

The Reclamation Act authorized federal development of water projects and established the Bureau of Reclamation which became the principal builder and contractor of new projects. The Colorado River Compact produced a rigid allocation of water between the upper basin and the lower basin states and among the states in each region. One serious problem with the compact was that the volume of the average annual flow of the Colorado was overestimated at the time it was divided, and therefore in later years there was not enough water to satisfy the various state allocations. Conflict was inevitable, and the states became very jealous of their allocations and reluctant to see their water

moved elsewhere.

But by all odds the most important factor affecting the West was the discovery that water development could be used to redistribute income and wealth from the nation's taxpayers to the Western region through federal projects. The argument was used that this was equitable, since the water subsidy was the way the rest of the nation could compensate the West for the removal of the forests and the parks from the private economy.

Almost from the beginning federal water pricing was mishandled. Under the Reclamation Act rules were established for repayment of the development costs by the beneficiaries. But farmers either could not or would not meet these repayment obligations. Moreover, it was argued by irrigation interests that many of the project costs were not their responsibility but should be paid by the Treasury because of the commitment of the government to provide broad public goods like flood control and recreation. These were considered to be nonreimbursable costs. Even deducting these costs, however, irrigators could not pay the costs directly attributable to irrigation. Interest costs, always the largest class of costs on these long-lived projects, were therefore waived for irrigators. Repayment periods were lengthened to fifty years in order to reduce annual repayment obligations. Electric power users were required to subsidize irrigators to make it appear that entire projects were economically feasible. And finally, when irrigators could not repay even these remaining reimbursable costs, the government

came up with the "ability-to-pay" rule. This procedure utilized cost and returns budgets to show that farmers could afford to pay only a small fraction of the total separable costs.

Very significantly, when it was discovered that water subsidies could be used to transfer income and wealth from the taxpayers of the nation to Western irrigators, a water lobby developed that became immensely powerful in Washington in producing new projects for the region.

There can be little question that the effect of these changes in repayment rules along with the power of the water lobby was to produce premature and excessive water development and associated high environmental costs. What evidence supports such a charge?

The most convincing is a simple comparison of costs and values of federal water. On the newer projects, the separable costs attributable to irrigation vary from $250 to $500 per acre-foot.[2] (An acre-foot is the volume of water spread over one acre of land one foot deep.) I must reemphasize that these separable costs do not include any of the joint project costs such as dams, but do include interest costs on the original irrigation component of the project.

---

[2] Ray Coppock, Robert Hagan, and William Wood, "Introduction," **Competition for California Water,** edited by Ernest A. Engelbert, (Berkeley: University of California Press, 1982.) Zack Willey points out that the costs of proposed new California state water projects range from approximately $450 per acre-foot to $850 per acre-foot in 1985 dollars. See Zack Willey, **Economic Development and Environmental Quality** (Berkeley, California: Institute of Governmental Studies, University of California 1985): p. 26.

The average value of irrigation water ranges from around $5 per acre-foot in some areas to as much as $50 per acre-foot in others which have favorable climates and growing conditions. How do we know? We have a few water markets where transfer prices are observable. One of the oldest of these markets with reliable data is located just across the mountains from here in the Delta area of Utah. Economists can also simulate these values from economic optimization models.

But what do farmers repay to the government? Under the ability-to-pay rule farmers pay from $20 per acre-foot on down to practically nothing, depending on how the budgets are constructed. Let's give farmers the benefit of the doubt and assume that they repay $20 per acre-foot to the government.

The difference between water's value, say $50, and the repayment price of $20 is of course $30 per acre-foot. Further, a good average is the delivery of three acre-feet of water to each acre in the project. Hence, the farmer's profit from subsidized water might be $90 per acre per year. Assuming a conservative forty year contract, the capitalized value of this $90 of annual water profit at 5% is just over $1,500, a pretty good estimate of the appreciation of land values in good agricultural areas when land is irrigated with subsidized water.

It is this increment in wealth for land owners lucky enough to receive subsidized water that drives the politics of the system. Good, you might say! So the government subsidizes water development that creates wealth in the region. What can be wrong

with that? Plenty it turns out.

I come now to the most important point of this talk, the economic efficiency implications of irrigation subsidy. Society as a whole cannot be wealthier if it requires $250 to $500 worth of scarce labor and capital resources to produce an acre-foot of water that is worth only $50 to the users. Let's assume a separable water cost of $400 per acre foot. The difference between the $400 of cost and water's value of $50 is of course $350. What happens to this $350? Who captures it? The answer is nobody does. It simply vanishes into thin air. It is deadweight loss suffered by all of us taxpayers resulting from excessive and premature water development.

And sadly, this investment in existing dams and irrigation facilities cannot be economically recovered. Society must write off the loss as a costly mistake. The best we can do is learn from experience and not repeat these errors. I will return later to what should now be done.

The environmental costs of the irrigation subsidy are also very large. These are currently a serious national concern. They result from both excessive water applications per acre and too many acres irrigated. Salinity is probably the costliest of these problems. You cannot irrigate without soluble salts entering the waste water. If the water is applied again and again the water becomes more and more salty. After a certain level these salts are toxic to plants. If the drainage is poor, the salts rise to the top of the soil profile and nothing grows.

Warren Hall, an eminent hydrologist, has made an eloquent statement of the salinity problem:

"Salt problems are particularly insidious. They do not come charging at you with trumpets blowing and battle flags flying, a sight to set stirring the heart of activists in any century. Rather, they slip in almost unnoticed. They invariably seem to promise to step aside and behave themselves in return for small additional concessions. Then one day, as witnessed by many dead civilizations, they assert their supreme command of the situation. Time is of no concern, for they are supremely confident of their ultimate victory. . . They have quietly destroyed, without fuss or fanfare, more civilizations than all of the mighty armies of the world."[3] Waterlogging of soils also occurs from excessive irrigation if drainage is poor as water builds up in the soil profile and stops all plant growth.

There is also the loss of free-flowing rivers and streams and wildlife habitat that result from excessive irrigation development. Chemical contamination of surface and ground water is also becoming an increasingly serious environmental problem. The pesticides and fertilizers used in agricultural production percolate into water supplies and pollute them. Let me summarize what I think I have said thus far.

When all is taken into account and evaluated as best we know how, the costs of federal water developed over the past half

---

[3] Hall, Warren A. "Statement: Salty Solutions to Salty Problems." **Salinity in Water Resources** (Boulder: Merriman 1974).

century, but especially in recent years, are far greater than the benefits. What a sorry ending from such promising beginnings.

What should now be done?

It is most critical to get water prices right, but this is more complex than it sounds. The direct beneficiaries of new water development should pay the full costs just as our pioneer forefathers did. These costs should not be shifted to local property owners or other taxpayers. They should not be shifted to power users. If farmers will not pay the full costs, then this is very strong evidence that irrigation development is not producing net wealth. This reform alone would stop any uneconomic new development with all its associated environmental costs.

For old projects already built, the remedy is not so easy. The price of water should be at least as high as the avoidable O & M costs. If this price will not be paid, society would be ahead by simply abandoning the project. But if irrigators can afford to pay these avoidable costs, then the existing small net benefits alluded to above will not be lost.

In addition to the pricing problem, we seem to be operating with a set of obsolete institutions that were suitable for a time when water needed to be harnessed and developed. Now we need institutions that will provide incentives for water to be conserved and transferred to higher valued uses. For the most part, our existing legal and administrative institutions simply will not do the job. A functioning water market where free

trading is permitted is the only reliable institution available for this purpose.

There is an urgent need to eliminate impediments to water market transfers to higher-valued uses and users. The federal government should permit recontracting by those who now have long-term contracts with repayment obligations transferred with the water. State transfer-approval agencies, such as the State Engineer in Utah, should not deny transfer petitions unless the denial serves some urgent public purpose. The beneficial use doctrine should be abandoned. If a prospective user is willing to pay the market price for a water transfer, then the purpose must be beneficial by any reasonable economic interpretation.

These changes would remove the chief obstacles to the formation of water markets and allow them to work their magic in facilitating water movement. Instream uses such as recreation, now excluded in many states from the appropriation process, could legitimately compete for water supplies. Markets could greatly augment the effective water supply by providing incentives for conservation and the use of efficient water application technologies. There is still much technical inefficiency in irrigation. If farmers could sell part or all of their existing water they would take steps to use it efficiently and conserve it.

As a society we must find a way to diminish the power of the water lobby, and water markets would contribute to that too. Spending huge amounts for political favors is irrational and

unnecessary if resources are allocated by impersonal markets.

And huge amounts are being spent in political lobbying. Soon after he came to power, President Jimmy Carter's administration clearly saw the inefficiency of several proposed Western water projects. The President proposed to eliminate 18 of the least cost-effective scheduled for future construction.[4] Little did he realize what it might be like to take on the water lobby. It is significant that although Carter's proposed water cuts accounted for a mere 1.7 percent of the total proposed expenditures in the omnibus Public Works Appropriation Bill, hearings on the bill focused overwhelmingly on the water projects. Of the 413 witnesses who appeared before the subcommittee to testify on the bill, over 80 percent were present to defend the threatened water projects. Of these, 107 were members of the House of Representatives, a full quarter of that body. Nearly three quarters of the pages devoted in the Congressional Record to the omnibus bill were focused on the water projects.

President Carter backed off in the face of this ferocious Congressional opposition. He agreed to accept only $60 million in cuts, affecting only 9 of the 18 projects originally proposed for elimination. But Congress wasn't finished with the matter. The following spring the Public Works Subcommittee restored 8 of the 9 deleted projects. Carter vetoed the bill but was

---

[4] Peter Samuel, "Hill Bent on Spending," **Reason**, Vol. 16, No. 6, (November 1984): 42-46.

overridden by Congress. So 17 of the 18 projects designated for elimination survived and probably will be built.

Let me talk briefly about another case that is much closer to home, a fascinating example of how good economics is sacrificed on the altar of special-interest politics. Hearings were held last fall by the Senate Subcommittee on Water and Power.[5] The topic of the hearings was completion of the Irrigation and Drainage System (I&D) of the Bonneville Unit of the Central Utah Project. The Central Utah Project Completion Act being debated would authorize $150 million in Federal funds for construction of this system. It must be emphasized that no work on this system has yet begun so no capital has been sunk in canals, ditches, and pumps. Senator Bill Bradley, the subcommittee chairman, had requested the General Accounting Office (GAO) to prepare a benefit-cost analysis on this unstarted project work.

The primary purpose of the Bonneville Unit's I&D system is to supply irrigation water to farmlands in central and southern Utah, but it will also provide a small amount of municipal and industrial water to cities in Juab and Utah Counties. According the Bureau's plans, about 40 percent of the system's water will provide supplemental water to presently irrigated land and thereby allegedly stabilize agricultural production. Most of the

---

[5] United States Senate, Hearing before the Subcommittee on Water and Power, Committee on Energy and Natural Resources, **Central Utah Project Completion Act,** (Washington: U.S. Government Printing Office, September 18, 1990):148-168.

remaining system water will be used to irrigate presently unirrigated land to offset land being taken out of agricultural production by urbanization and industrialization in the state.

The Bureau had not calculated a benefit-cost ratio on the I&D system, but had on the project as a whole. To calculate the ratio desired, the GAO extracted from the Bureau's 1988 benefit-cost analysis of the Bonneville Unit only those benefits and costs associated with the I&D system. The ratio was 0.8. (This number says that the national economy will benefit by 80 cents for every dollar of assigned cost to the system.) However, recall that these separable costs do not include any interest charges on construction, no share of the joint costs of the larger Bonneville Unit, and therefore are only a fraction of the real economic costs of the system, probably less than 10 percent.

Using standard economic principles, the GAO analysts found what they considered to be errors in the way the Bureau had calculated the benefits and the separable costs. The result was a large overstatement of the benefit-cost ratio. The Bureau had included as benefits the indirect profits earned by food processors, transporters, and retailers for delivering increased farm production to final consumers. GAO correctly surmised that these secondary benefits should not be considered as a benefit to the economy as a whole since these benefits would have been earned if the project funds had been spent in other alternatives. Further, the Bureau had not counted the labor of the farmers as a cost of the project. This assumes that farmers would have no

alternative employment over the 100-year life of the project had it not been built. This assumption is clearly erroneous since farmers have been moving out of agriculture to other and perhaps even better employment on a continuing basis for hundreds of years. And finally, GAO increased the annual costs to recognize salinity impacts downstream resulting from the I&D system that the Bureau had neglected to count. By diverting water from the Colorado River the salt concentrations in the remaining water of the Colorado would be higher, imposing costs on downstream users in both the United States and Mexico. These various adjustments by GAO resulted in a reduction of the benefit-cost ratio from 0.8 to 0.3, or 30 cents of benefits for every dollar of costs.

Strangely, it is not even clear that this I&D system will benefit the citizens of Utah on average. The cost of constructing the system has been estimated to be $328.5 million, of which only $150 million will be borne by the Federal Government. This leaves more than half of the costs to be funded by the Bureau's local cost-sharing sponsors. If these sponsors were the farmers and other direct beneficiaries and they would contract to bear these costs then they must expect to be better off from having the project. However, as illustrated above, under the ability-to-pay repayment rule used by the Bureau, the farmers will not bear the majority of these costs. Rather, it is intended that revenues be raised through a local ad valorem tax on other economic interests. In other words, the same sort of hiding and shifting of costs is occurring at the state level

between the direct beneficiaries and the taxpayers as is going on at the federal level. I haven't made the calculations, but I am not at all convinced that even Utah will be better off if this system is constructed.

The reactions of Western Senators at the hearings to the GAO report were predictably hostile. Interestingly, they did not dispute any of the GAO data, analytical procedures or conclusions; they simply found them irrelevant. In fact, one of this State's Senators argued that he and the other politicians knew all along that some components of the CUP were economically infeasible, but that the project as a whole was economically justified. Besides, the federal government had promised this system to the people in poor rural counties and now the credibility of the government was at stake. The government must deliver on its promises.

Of course, there is moral value in keeping promises, especially for the government. But what as citizens are we to think of a government where the politicians knowingly, in advance, promote water projects for special interests that cost the taxpayers of the nation $300 per acre-foot when that water is worth $20 or less to the interest group. I don't know about you, but talk of credibility and keeping promises under such circumstances seems pretty hollow.

The problem is that this plundering of the national treasury is pandemic in our political system. This point came up at these same hearings. A Senator from the South argued that the poor

people of Utah should have their subsidized water since farmers in his state were subsidized to grow tobacco. We call this mutual back-scratching log-rolling.

Another visible and similar phenomenon, is the proposed closing of unneeded and obsolete military bases. And it is just as predictable that Congress will fight tooth and nail to protect those in their districts and that sufficient logrolling will occur to keep most of them. Is it any wonder with such squandering of the nation's capital that we scarcely grow at all and increases in standards of living have been so modest over the past two decades.

It would be too bad, however, if I closed this discussion on this very pessimistic note. Some encouraging things are happening with water, but candor requires me to tell you that they are rare and face almost insuperable obstacles. Where water was developed by private parties, it is indeed moving to more valuable uses, especially in Utah, New Mexico, and Colorado which have the most flexible water transfer institutions. But in many states, objections of a relatively minor nature are blocking water transfers worth millions of dollars.

People are beginning to think about markets. They are beginning to talk about trades that make economic sense. Officers of the Imperial Irrigation District in Southern California have been talking with officials from the Metropolitan Water District headquartered in Los Angeles. The MWD is proposing to repair leaks in Imperial's irrigation system that

result in a loss of about 100,000 acre-feet of water annually. Then MWD will pay Imperial for the water saved and move it to urban Southern California at far lower costs than would be spent for alternative supplies. Another proposal would transfer water from Little Snake River farmers near Baggs Wyoming to San Diego in Southern California. The plumbing is the Colorado River and the Colorado Aqueduct. Urban water in Southern California is probably worth at least $1,000 per acre-foot while in agriculture in Baggs Wyoming, the most it could be worth is $10 per acre-foot. Baggs' farmers could be made wealthy by selling out, no other Upper Basin water users would be worse off, and consumers of water in Southern California would be much better off. It is difficult to identify any losers. However, there is almost a consensus among water planners that without approval of all the Colorado River Basin States such proposals will not get off the ground because of the likelihood of court actions brought by somebody. But who would sue? Probably state officials and politicians who would be worried that their water promised under the Compact would be flowing to despised Southern California. Hence, it is highly problematic whether political and institutional constraints can be overcome to permit the transfer.

One last word. As citizens we should pressure our elected representatives and regulatory officials to approve such obvious accretions to national wealth. It is my belief that political obstacles that prevent resources from moving to higher valued uses is the principal impediment to economic growth and

improvements in our standards of living and our quality of life. We have to find a way to get rid of them. The economic welfare of generations to come as well as our own hangs in the balance.

Chapter 1

APPROPRIATORS VERSUS EXPROPRIATORS
The Political Economy of Water in the West

Alfred G. Cuzán

A POLITICAL ARGUMENT FOR THE PRIVATIZATION OF WATER IN THE WEST

Economists and philosophers since Locke and Smith have explained the economic laws of property and exchange. A solid body of scholarship, both classical and modern, suggests that aside from enforcing property rights, reducing transaction costs and, in some instances, providing for so-called public goods that are difficult to charge for, government can do little to improve the efficiency of free markets. This relatively unrestrained system, what Oppenheimer called the economic means of appropriation and exchange, constitutes one of the most effective mechanisms for harnessing the energies of enterprising human beings to increase production and raise living standards the world over.[1]

The same cannot be said for the other type of social system identified by Oppenheimer—the political means of expropriation and taxation. The indiscriminate use of laws and regulations, even in democracies, generally results in a net loss in efficiency as rent-seek-

1. Franz Oppenheimer, *The State* (Indianapolis: The Bobbs-Merrill Company Publishers, 1914); Ludwig von Mises, *Human Action* (New Haven, Conn.: Yale University Press, 1949).

ing groups team up with the bureaucracy and politicians to reap most of the benefits from public policy. The social costs of acting with institutions that raise revenues through taxes, register popular preferences with infrequent acts of voting, allocate resources in political struggles among small groups and manage them through a centralized bureaucracy generally exceed whatever benefits are bestowed on the public or the small, active minorities who exercise the most influence or control over the government's actions.[2]

The contrasts between the economic and the political means in the field of water resources has been of interest to economists and lawyers for some time. In their path-breaking book, *Water Supply*, Hirshleifer, DeHaven, and Milliman noted the many failures of water policies at the local, state, and federal levels, arguing that a better method would be to treat water like any commodity, subject to appropriation and exchange in a market economy.[3] Water policy in the nineteenth century embodied this approach, so the authors' recommendation actually was a reversal of the policies that had accumulated since 1880. Other writers have arrived at similar conclusions.

My purpose in this paper is not to repeat the economic arguments against government controls, but rather to make a political case for privatization. By examining the machinery of government and the dynamics of politics, I will show that the internal laws that regulate the political means necessarily promote centralization and bureaucratization of the water industry, not in the interest of equity or efficiency, but as a consequence of organized efforts by a ruling class to expropriate income and wealth from the public. This constant relation, as Mosca calls it, is probably the most solidly established law in political science.[4] In order to escape its perverse power, society must choose the economic laws of the market instead.

I begin with a discussion of Locke's positive theory of property, which explains the way water in fact was appropriated under nearly anarchical conditions in the West during the mid-nineteenth century. I then examine the evolution of public water policies since 1860, demonstrating that they have consistently expanded and centralized

2. See William C. Mitchell, *The Anatomy of Public Failure: A Public Choice Perspective* (Los Angeles: International Institute for Economic Research, Original Paper 13, June 1978).

3. Jack Hirshleifer, James C. DeHaven, and Jerome W. Milliman, *Water Supply* (Chicago: University of Chicago Press, 1969).

4. Gaetano Mosca, *The Ruling Class* (New York: McGraw-Hill, 1939).

the power of those who influence, control, and benefit from the political means. This is followed by an analysis of the organizational instruments that have planned, promoted, and implemented these policies in a comparative study of the federal Bureau of Reclamation and the Water and Power Department of the city of Los Angeles. These two agencies are remarkably similar in their political origins, territorial growth strategies, and evolution toward centralized bureaucratic management. I conclude with a suggested rule for reappropriating water in the United States.

THE AMERICAN DOCTRINE OF PRIOR APPROPRIATION: A LOCKEAN EXPERIMENT IN NATURAL LAW

In *The Second Treatise of Government*, John Locke explains how increasingly scarce common-pool resources come to be appropriated in a State of Nature according to the principles of priority of right and beneficial use, a model that can readily be applied to the history of the American doctrine of prior appropriation.[5] Locke assumes that, in the beginning, the earth and its products constitute a great common to mankind while individuals have a property in their own persons. Nature compels individuals to apply their labor to take what they need from the commons in order to survive.

A person's right to anything in the commons is established by the simple act of taking it or enclosing it with his or her own labor. With respect to water, Locke observes, "Though the water running in the fountain be everyone's, who can doubt that in the pitcher is his only who drew it out? His labor has taken it out of the hands of nature where it was common and belonged equally to all her children and has thereby appropriated it to himself."[6] Note that the act of *removing* a portion of the commons establishes an individual's property over it, eliminating the ambiguity associated with the concept of "mixing" one's labor with the earth: "the *taking* of what is common

5. Compare Locke's *The Second Treatise of Government* (Indianapolis: Liberal Arts Press, 1952) with the following two articles: Armen A. Alchian and Harold Demsetz, "The Property Rights Paradigm," *Journal of Economic History* 33 (1973): 16–27; and Terry L. Anderson and P. J. Hill, "The Evolution of Property Rights: A Study of the American West," *Journal of Law and Economics* 18 (1975): 163–179.

6. Locke, *Second Treatise*, p. 18.

and removing it out of the state nature leaves it in which begins the property, without which the commons is of no use" (emphasis added).[7] Locke calls this principle the "original law of nature."

In some instances, the act of discovery, itself being an act of labor, is sufficient to establish a prior right even before actual enclosure. The hunter, for example, who pursues a hare "has thereby removed her from the state of nature wherein she was common, and has begun a property."[8] Thus, a man who first discovers a gold mine or a new source of water establishes a right to it by virtue of having found it.

Under conditions of unlimited supply or relative abundance the appropriation of any part of the commons harms no one, as Locke observes with respect to water: "Nobody could think himself injured by the drinking of another man, though he took a good draught, who had a whole river off the same water left him to quench his thirst; and the case of land and water, *where there is enough for both*, is perfectly the same."[9] (emphasis added). However, under the more usual condition of relative scarcity of either land or water, where each succeeding appropriation leaves fewer or less valuable resources in the commons, an individual's right is limited to only so much as he or she can use to any advantage before it spoils; additional resources exceed his or her share and belong to others. This rule applies not only to the products of the earth, such as wildlife and water, but also to land: "As much land as a man tills, plants, improves, cultivates, and can use the product of, so much is his property. He by his labor does, as it were, enclose it from the common. . . ."[10]

Individuals may give away, barter, or exchange for money anything that they appropriate. It is only if one allows resources or products of the earth to rot or remain unproductive that "this part of the earth, notwithstanding his enclosure, was still to be looked on as waste and might be the possession of any other. . . ."[11]

Therefore, in a modern economy, one need not work resources directly in order to retain title. All one need do is to ensure that what one owns does not go to waste, for example, the stockholder in a water company, the buyer of canal bonds, and the speculator in water rights. Locke would not have found the commercialization of

7. Ibid.
8. Locke, *Second Treatise*, p. 19.
9. Ibid., p. 20.
10. Ibid., p. 20.
11. Locke, *Second Treatise*, p. 23.

water irreverent or objectionable in any way; on the contrary, he would have heartily agreed with the authors of *Water Supply* on this point.

Like a modern economist, Locke argued that as resources that were once part of the great commons of mankind rise in value, individuals strive to appropriate them. Far from harming those who fail to acquire a portion of the shrinking commons, the process of appropriation benefits them as well by increasing the productivity of the resource:

> He that encloses land, and has a greater plenty of the conveniences of life from ten acres than he could have from a hundred left to nature, may truly be said to give ninety acres to mankind; for his labor now supplies him with provisions out of ten acres which were by the product of a hundred lying in common.[12]

It is not my purpose here to evaluate Locke's theory as a normative principle of justice. Rather, I want to use the positive aspect of the theory to explain the process by which the waters of the West were originally appropriated outside the established legal framework, or as Locke would say, "out of the bounds of society."

Locke's theory of property rests on a natural law of appropriation regulated by economic forces. As resources held in common become more scarce, the most enterprising members of the community—the industrious and rational—apply their labor to enclose and put them to use. In order to minimize the cost or inconvenience associated with disputes over title or the size of possession, the appropriators, by voluntary consent, reach agreement on two fundamental rules for dividing the commons: (1) first come, first served, or priority of right acquired by virtue of discovery or possession, and (2) a person's right is limited to what he or she puts to beneficial use.

It is precisely these two principles upon which the American doctrine of prior appropriation rests. This body of rules was developed by communities of California miners in mid-nineteenth century. Around 1850, thousands of men came from around the world to search for gold in what was then largely a wilderness. They became squatters on the federal public domain, outside the established legal order, with no government to impose order or settle their disputes. In effect, they found themselves in a "state of nature." As Locke

12. Locke, *Second Treatise*, p. 23.

would have predicted, a first come, first served principle was adopted in the establishment of rights over what had been held in common—minerals and water—along with limitations on what any one individual could own.

> Following a tradition of collective action on the mining frontiers of other continents, the miners formed districts, embracing from one to several of the existing 'camps' or 'diggings' and promulgated regulations for marking and recording claims. The miners universally adopted the priority principle, which simply recognized the superior claims of the first arrival. But the . . . miner's codes defined the maximum size of claims, set limits on the number of claims a single individual might work, and established regulations designing certain actions—long absence, lack of diligence, and the like—as equivalent to the forfeiture of rights. A similar body of district rules regulates the use of water flowing in the public domain.[13]

In order to ensure that no valuable mineral rights were wasted, local custom sanctioned claim jumping whenever "the prior claimant had abandoned his claim, had failed to diligently work it, had staked it without following local regulations, or held more claims than district rules permitted."[14]

These customs spread throughout the West as miners, irrigators, and cattlegrowers took possession of the most valuable portions of the public domain without legal authorization from territorial, state, or federal governments. The first person to work a mine, graze a herd on a meadow, or divert water from a stream acquired a prior right to what he or she took, and these appropriations were for the most part respected by subsequent settlers.

Elwood Mead describes how cattlegrowers on the public domain divided the grasslands among themselves:

> There was no law by which men could legally secure control of the land they occupied. All the land laws dealt with farming land. There was no provision for leasing or settling the grazing land in tracts large enough to be of any service. Hence the range stockmen simply took possession of the country. Each man chose a location which suited him, fixed in a rough way the boundaries of his domain, and helped to create a public sentiment which made it unpleasant, if not dangerous, for a late comer to attempt to share with him the territory he had so marked out. In this way range rights came to have the force of law.[15]

13. Charles W. McCurdy, "Stephen J. Field and Public Land Law Development in California, 1850–1866: A Case Study of Judicial Resource Allocation in Nineteenth-Century America," *Law and Society* (Winter 1976): p. 236.
14. Ibid., p. 243.

The custom originated by miners and cattlegrowers had the greatest impact with respect to water used for irrigation. The American doctrine of prior appropriation, or arid region doctrine, was adopted by state constitutions, legislation, and judicial rulings. It sanctioned the diversion of water for use on lands far from the natural watercourse on the basis of priority of right and wrought what Webb called "a revolution in the law of water" displacing the riparian doctrine partially or totally over most of the West.[16] The greater the relative scarcity of water, the more it was appropriated. Thus, the doctrine became most firmly established in the most arid portions of the region—Arizona, New Mexico, Utah, Nevada, Colorado, and Idaho.

The new body of law effectively separated rights to water from rights to land. Companies mobilized private capital to build irrigation works and transport water to where it was most productively used. Writing in 1903, Mead called corporate irrigation "the leading factor in promoting agricultural growth of the Western two-fifths of the United States."[17] Mead credited canal companies with promoting efficient irrigation practices through contractual arrangements, advising the state of Colorado to study canal companies' management of the water they appropriated.

By 1910, over 13 million acres of land were irrigated privately in the 17 western states.[18] In a very short time, thousands of farmers had multiplied by many times the productive value of this region. Mead, who found much to object to in private water development, nevertheless acknowledged that "in the last third of the nineteenth century the arid West became one of the greatest irrigated districts of the globe. In mileage of ditches and in acres of land irrigated it surpasses any country of Europe or Africa and is second only to India and China in Asia. . . ."[19]

15. Elwood Mead, *Irrigation Institutions* (New York: Macmillan Co., 1903), pp. 28–29.
16. Walter Prescott Webb, *The Great Frontier* (Austin, Tex.: University of Texas Press, 1951), pp. 254–259.
17. Mead, *Irrigation Institutions*, p. 57.
18. Alfred R. Golzé, *Reclamation in the United States* (Caldwell, Id.: Caxton Printers, 1961), p. 12.
19. Mead, *Irrigation Institutions*, p. 349.

By the turn of the century, this remarkable phase of private development of the West, a reflection of the laissez faire era of the late nineteenth century, was in its twilight. A wave of so-called reform was being mobilized by men such as Major Powell and his nephew Arthur P. Davis, Elwood Mead, F. H. Newell, William Stewart, Theodore Roosevelt, and others. They sought to regulate the water industry with the police power, expropriate rights with the navigation acts, develop the American desert with public funds while simultaneously promoting democracy west of the hundredth meridian with the land grants.

EXPROPRIATION OF WATER BY STATE AND FEDERAL GOVERNMENTS: A HISTORY OF WATER POLICIES IN THE ARID LANDS

The history of water policies since 1860 is one of expropriation of property rights by federal, state, and local governments. This trend is manifested in the transformation of the appropriation doctrine into administrative law at the state level; control over canal company water rates by state and county governments; state controls over underground water pumping; the takeover of private water companies by municipal governments; and the imposition of federal controls over rivers, dams, reservoirs, and irrigation works for reclamation and power development. In this section, the evolution of state and federal controls is reviewed in general terms. Part IV compares in detail Bureau of Reclamation policies with those of the city of Los Angeles.

From Rights to Permits: The Evolution of the Water Law at the State Level

In much of the arid West, the appropriation doctrine was being applied by irrigators and miners before state or even territorial governments were organized. A potent political force, appropriators were able to resist or overcome efforts by the newly organized governments to establish bureaucratic controls over them. However, over a number of decades, an accretion of state and judicial controls, promoted by reformers and the federal government, transformed the doctrine of appropriation into a system of administrative law.

In one of its first acts, the California legislature in 1851 sanctioned the local customs by which water and mineral rights had been established.[20] Over the next decade, a series of state supreme court decisions lent additional sanction to the appropriation doctrine, which took its place along if not above the riparian doctrine.

The initial victory of the appropriators was relatively short-lived, however. No sooner had their rights been recognized than the political means were organized to expropriate them. State constitutions and statutes universally adopted the principle of public ownership over water. Appropriators were granted the right to use the water (usufructury right) but ownership over the resource itself (the *corpus* of water) was declared state property. The inherent tension between these two principles eventually was resolved in favor of state controls over water.

The first impetus to state controls over water were water rights disputes. As the population and water demands in the West grew, the decision or transaction costs associated with the establishment and the adjudication of rights increased. In periods of drought, disagreements between senior and junior and upstream and downstream appropriators presented opportunities for judicial and administrative interventions. These were taken partly at the behest of irrigators themselves, who wanted the state to subsidize, at least partly, the settling of conflicting or competing claims. For example, following a series of conflicts over the Cache La Poudre River in 1874, Colorado irrigators "met in convention to demand legislation for public determination and establishment of rights of appropriation, and then state superintended distribution of water in accordance with the thus settled titles. . . ."[21] These recommendations were incorporated into law in 1879.

The transformation of the appropriation doctrine into administrative law, begun with judicial or administrative interventions to settle disputes, established the following requirements and restrictions:

1. Requirements for the filing of new claims, first at the county, then the state level

20. McCurdy, "Public Land Law Development," p. 239; Samuel C. Wiel, *Water Rights in the Western States* (San Francisco: Bancroft-Whitney Company, 1980), p. 12.

21. Moses Lasky, "From Prior Appropriation to Economic Distribution of Water by the State Via Irrigation Administration," *Rocky Mountain Law Review* 1:3: 173. See also Golzé, *Reclamation in the United States*, p. 10.

2. Limitations on the size of excessive claims and legal specifications on the duty of water (the amount applied to an acre of crops)
3. Attachment of water rights to specific land tracts
4. The disallowal of ownership to water by canal companies that did not irrigate lands of their own
5. Regulation of canal company rates by states and counties
6. State encouragement of the formation of irrigation districts with the power to tax, condemn property, and sell bonds to finance construction of irrigation works and buy out water companies
7. Legislative determination of what constitutes beneficial use, along with the ranking of uses by classes
8. Prohibition on sale of water rights beyond state or irrigation district boundaries
9. Administrative allocation of water during periods of drought
10. The establishment of a centralized bureaucracy headed by a state engineer or water commissioner to administer policies and judicial decrees and, in some states, undertake irrigation projects.

Thus, rights established extralegally in a quasi-anarchistic setting gradually were transformed to the status of permits or licenses held at the sufferance of state officials. As early as the first decade of the twentieth century, Professor Moses Lasky, perhaps prematurely, declared in a three-part article in the *Rocky Mountain Law Review*, that the principle of appropriation had all but vanished.[22] In his view, Wyoming and Colorado were leading a new "revolution" in western water law. The thrust of these changes was *away* "from various forms of extreme individualism and vested property rights of substance in water to the same goal, the economic distribution of state-owned water by a state administrative machinery through state-oriented conditional privileges of user. . . . The transition has been via administration, and in administration is most clearly noted."[23] Fifty years later this trend was confirmed by Milliman.[24]

22. Lasky, "From Prior Appropriation," p. 269.
23. Ibid., p. 162.
24. J. W. Milliman, "Water Law and Private Decision-Making: A Critique," *Journal of Law and Economics* II (1959): 41–63. See also Hirshleifer, De Haven, and Milliman, *Water Supply*, chap. IX.

The transformation of a system of water rights acquired independently of the government into one of permits secured from a state bureaucracy undermined the security of titles, making it difficult to transfer water to its most productive uses. This system increased the unreliability of water supply, prevented the integration of irrigation works and river systems, and led to a cumbersome structure of regulations and decrees. In a description of the distribution system of a California irrigation district, Mason Gaffney noted:

> The Kaweah water distribution system has had to grow in a manner analogous to the law itself, with one principle hanging on another back to the ancient and ultimate fountainheads of authority. It is questionable whether circuitous transfers of this sort are desirable at all, even if each individual operation shows a net gain. For as one ditch is latched on to another, more and more interests become vested in an increasingly absurd tangle and the hope of rationalization recedes even further into the realm of inattentive visions.[25]

Moreover, vesting property rights over water in irrigation districts and mutual water companies rather than in individuals had resulted in practical prohibition of its sale. Thus, much water "is effectively withdrawn from commerce in a mortmain grip as deadly as that fastened on the lands of medieval Europe."[26]

The cumbersome and antiquated system of controls described by Gaffney has made it difficult for much of the water to be transferred privately through marginal adjustments made by continuous sales. Instead, the system generates pressures for a step-wise program of monumental water schemes by state and federal agencies, which subsidize low-value crops such as alfalfa and pasture with a policy that produces "too much water, too late."[27]

Also contributing to political pressures for large projects has been the tendency of state governments to grant permits to a volume of water that is greater than what is, in fact, available. Because the water rights of the most junior appropriators are worthless unless additional volume is made available, support is generated for dams and water transfer projects, which raise the value of junior rights.[28]

25. Mason Gaffney, "Diseconomies Inherent in Western Water Laws: A California Case Study," *Western Resources and Economics Development*, Report #9, Western Agricultural Economics Research Council, 1961, p. 71.
26. Ibid., p. 74.
27. Ibid., p. 78.
28. See Elco Greenshields, *Journal of Farm Economics* (December 1955): 900.

Ironically, one of Mead's criticisms of the appropriation doctrine was that it sanctioned excessive claims, often amounting to several times the known river flow. Yet the imposition of state controls did not end the practice; it simply made the results far more costly. Individual overappropriations were checked by their resources to invest in diversion works. State overappropriations are constrained only by the amount of money the federal government is willing to spend augmenting local supplies.

Reclamation and Navigation: The Imposition of Federal Controls

As indicated in the previous section, state legislatures and courts had begun to legalize private claims on the basis of the appropriation doctrine. At the initiative of Nevada Senator William Stewart, Congress in 1866 followed suit and gave its consent to the state laws and local customs on which private claims to minerals and water rested.[29] From then on, however, public entrepreneurs such as Major John Wesley Powell and future commissioners of the Bureau of Reclamation such as F. H. Newell, Arthur P. Davis, and Elwood Mead seemed to be engaged in a race with appropriators for the control of water and irrigable lands in the region. They claimed that federal control would promote scientific conservation and development of land and water resources; prevent the monopolization of water by corporations and speculators; streamline the system for establishing and enforcing water rights; and encourage the development of rural democracy by war veterans and other deserving pioneers. These policies received the strong backing of at least three presidents, including the two Roosevelts and Herbert Hoover.

In 1878, Major Powell, director of the U.S. Geological Survey, presented to Congress his *Report on the Lands of the Arid Region* in which he urged the federal government to assert vigorous control over its western holdings in cooperation with locally organized districts. While welcoming the impetus given to western development by private efforts, he warned that the separation of water rights from land titles would lead to concentration of ownership. "Monopolies of water will be secured, and the whole agriculture of the country

29. Wiel, *Water Rights in the Western States*, p. 26.

will be tributary thereto—a condition of affairs which an American citizen having in view the interests of the largest number of people cannot contemplate with favor."[30]

During the next decade Powell single-mindedly pursued what can only be characterized as a massive land grab of the West, withdrawing from entry 850 million acres of the public domain. He designed large-scale plans for their orderly surveying, irrigation, and development under federal guidance. His budget grew from $50,000 in 1878 to $156,000 in 1881, reaching over $750,000 in 1890.[31]

Powell's policies elicited a mixed response from irrigation interests. As chairman of the Senate Committee on Irrigation, Senator Stewart managed to increase Powell's budget and support his plans for federal surveys of irrigable lands and reservoir sites. The senator was particularly interested in the development of the waters of the Carson, Walker, and Truckee Rivers in Nevada, which in his view were being allowed to go to waste. In 1889, the two men, along with other members of the Senate Irrigation Committee, toured the arid region seeking support for a federal role in irrigation. Two years later irrigation congresses began to meet to promote federal reclamation policies.

But soon after the trip the friendship between Stewart and Powell cooled as their personalities and ideas clashed. Stewart simply wanted federal subsidies with few strings attached to water rights or land uses. Powell, on the other hand, wanted greater federal control preceded by years of study and planning. As the latter continued to withdraw additional millions of acres from entry, opposition in the Congress grew, led by Stewart and another Nevadan, Congressman Francis G. Newlands, an owner of land and reservoir sites along the Truckee and Carson Rivers. Newlands, who later became a U.S. senator, opposed Powell's long-term planning schemes because they obstructed "practical" federal irrigation projects such as the one that now bears his name. Powell was forced to resign in 1894 and Congress loosened the controls he had imposed.

30. J.W. Powell, *Report on the Lands of the Arid Region of the United States*, House of Representatives Ex. Doc. No. 73, 45th Cong., 2nd sess., April 1878, p. 43.

31. Wallace Stegner, *Beyond the Hundredth Meridian: John Wesley Powell and the Second Opening of the West* (Boston: Houghton Mifflin Company, 1954), pp. 233, 273, 337, 341. For a less romantic view of Powell and others associated with reclamation see Stanley Roland Davison, "The Leadership of the Reclamation Movement, 1875-1902" (Doctoral diss., University of California, 1951).

Eight years later, the Reclamation or National Irrigation Act was enacted, initiating a new era of federal subsidies and controls. It began as a relatively modest effort designed to win the support from the West without generating too much opposition from the East. The government would sell public lands and put the proceeds into a separate reclamation fund out of which projects would be financed to irrigate new lands. Capital costs would be repaid within ten years and no interest would be charged. Expenditures among the western states would be proportional to the amount of revenues generated by sales of public lands within their borders. Farms would not exceed 160 acres, which would promote rural democracy. Irrigation works and the acquisition of water rights would conform with state laws, provided that the water rights be made appurtenant to the land. Finally, no "Mongolian" labor (a statutory reference to Oriental immigrants seeking work) was to be employed in the construction of irrigation works.

Hibbard summarizes what in retrospect turned out to be remarkably accurate predictions by the opponents of the act:

> A New York Congressman estimated that the plan would ultimately cost the country billions of dollars. Dalzell of Pennsylvania believed it a plan to "unlock the doors of the treasury." Mr. Cannon of Illinois dubbed the bill a "direct grant in an indirect way." Payne of New York was of a like mind, while Hepburn of Iowa insisted . . . "that this is a thinly veneered and thinly disguised attempt to make the government, from its general fund, pay for this great work—great in extent, great in expenditure, but not great in results."[32]

As it turned out, the critics were right. Within a few years, the federal treasury had to funnel tens of millions of dollars into the fund; repayment periods were extended first to twenty, then forty, then over fifty years as most projects failed financially.

In his seminal study of the relative costs and benefits of land reclamation in the Southeast and West, Rudolph Ulrich estimated that the costs of bringing desert land into agricultural production were from five to fourteen times as great as the costs of clearing, fertilizing, and controlling water inputs to lands in the Southeast.[33]

32. Benjamin H. Hibbard, *A History of the Public Land Policies* (Madison, Wisc.: University of Wisconsin Press, 1965), p. 442.

33. Rudolph Ulrich, "Relative Costs and Benefits of Land Reclamation in the Humid Southeast and the Semi-arid West," *Journal of Farm Economics* 35 (1953): 62–73.

Thus, federal efforts to make marginal desert lands bloom made no economic sense. Yet the more resources were spent for this purpose, the fewer remained to invest in the South, as Hibbard observed:

> In passing the Reclamation Act in 1902 as a nation we clearly forgot those things which were behind, the millions of unoccupied acres of the Mississippi Valley, consisting mostly of fertile, well-watered land needing only to be drained or cleared. Had we really been concerned over the future food supply as we pretended to be, or, being so concerned, had we calmly asked how to increase it in the cheapest and easiest manner, certain of the Reclamation projects would still be undeveloped.[34]

While support for reclamation projects was being harnessed, the federal government used the navigation acts to prevent private parties from developing reservoir sites or rivers it had already selected for itself. These acts asserted federal control over navigable rivers and all their tributaries on the basis of the commerce clause of the Constitution. They gave the government the power to expropriate property so that private concerns could not impede navigation without making any compensation. In several instances, such as in the Rio Grande and Colorado rivers, this power was invoked in order to block private irrigation projects that the Bureau of Reclamation later built itself, bringing ruin to the private developers.[35]

In his critique of the much-abused navigation doctrine, Charles E. Corker noted that it has proved to be a useful judicial device even when it was evident that the river was not navigable.

> Both the Congress and the courts have been content to treat the word "navigation" as an open sesame to constitutionality. So long as Congress uses the word in statute and the case relates to something moist, the Court takes at face value the declaration that the legislation is in furtherance of navigation. Moreover, the test of what constitutes a navigable stream has been stretched to embrace most of the waters of the United States.[36]

34. Hibbard, *History of Public Land Policies*, p. 449. Hibbard quotes a former director of the Reclamation Service as saying, "The fundamental object was to 'make men, not money'," a project in human or social engineering that socialist dictatorships favor in their attempts to build a "new socialist man."

35. Stegner, *Beyond the Hundredth Meridian*, pp. 310–312; Charles E. Corker, "Water Rights and Federalism—The Western Water Rights Settlement Bill of 1957," *California Law Review* 45 (1957): 604–637; Morris Hundley, Jr., *Water and the West* (Berkeley: University of California Press, 1975), pp. 19–26.

36. Corker, "Water Rights and Federalism," pp. 616–617.

It is evident that the long-term trend of federal policy has been to mobilize financial, administrative, political, constitutional, and judicial resources at its disposal to gain—or, perhaps more accurately, regain—control of western waters. It is as if, having been presented with a fait accompli in 1867, federal officials entered a race against time to gain control of the land and waters that were left unappropriated and recoup their previous losses. This was accomplished with a combination of subsidies designed to persuade irrigators and state governments to surrender or compromise their rights and open-ended constitutional claims to federal powers.

Note that federal and state policies have pursued similar strategies. The appropriation doctrine has been undermined, water rights have been virtually expropriated and converted into licenses or permits, and control over western waters has been centralized in state and federal governments. The tools may have been different, but the results have been the same.

THE IRON LAWS OF POLITICAL SCIENCE: A COMPARATIVE STUDY OF THE BUREAU OF RECLAMATION AND THE CITY OF LOS ANGELES

The long-run trend of public policies to expropriate water rights and centralize control over the resource in federal and state bureaucracies can be explained by two natural laws of politics: *the iron law of political redistribution* and *the law of hierarchical centralization.* Here I present a model of government in which a ruling class of bureaucrats, politicians, and interest groups—Lowi's "iron triangles"[37]—use political means to transfer wealth from the mass of the public to themselves. Such transfers are more efficiently carried out the greater the centralization of the government. I illustrate the model with a comparative study of the federal Bureau of Reclamation and the city of Los Angeles.

37. Theodore Lowi, *American Government: Incomplete Conquest* (New York: Holt, Rinehart and Winston, 1976).

A Political Model[38]

In government, individuals act in order to maximize their own utility with the political means of taxation, expropriation, and control or influence over so-called public resources. Those who specialize in exerting control or influence over specific policies constitute what Mosca calls a ruling class. Since the costs of public policy are borne, directly or indirectly, by the entire society, the ruling class in every policy area succeeds in transferring wealth or income from the mass of the public to itself. I call this phenomenon the *iron law of political redistribution.*

Political actors include the following:

1. *Bureaucrats.* The managers of public enterprises, they are control-maximizing actors who strive to secure as many resources—land, water, budgets, employees, or regulatory power—as possible. The more they control, the greater their utility.

2. *Politicians.* These influence-maximizing actors secure votes and consent from the public and its leaders. Politicians run for offices that are constitutionally or legally authorized to make the fundamental policy decisions on behalf of the public. They are the ultimate legitimators of what government does by enacting its laws and approving its budgets. With these decisions they influence the behavior of the bureaucrats. Politicians tend to specialize at influencing those agencies most relevant to their interests, sometimes to the extent that they actually control the agencies. For example, western congresspersons and senators tend to dominate congressional committees that authorize reclamation projects.[39]

38. This model is part of a theory of politics I have presented in "Political Profit: Taxing and Spending in the Hierarchical State," *American Journal of Economics and Sociology* 40 (1981): 265–275. Among the many contemporary scholarly works that have influenced my thinking, three are of special significance: Anthony Downs, *An Economic Theory of Democracy* (New York: Harper & Row, 1957); Gordon Tullock, *The Politics of Bureaucracy* (Washington, D.C.: Public Affairs Press, 1965); and Randall Bartlett, *Economic Foundations of Political Power* (New York: Free Press, 1973).

39. Helen Ingram, "Patterns of Politics in Water Resources Development," *Natural Resources Journal* 2 (1971): 110. See also Arthur A. Maass, "Congress and Water Resources," *American Political Science Review* 64 (1950): 576–593, and Aaron Wildavsky, *The Politics of the Budgetary Process* (Boston: Little, Brown, 1964).

3. *The Bureaucracy.* This consists of the public employees who carry out public policy under the direction of bureaucrats, who make personnel decisions such as job assignments, hiring and firing, salary, promotions, and so on. In the Bureau of Reclamation, engineers constitute the most important professional group.

4. *Clientele.* This is a collection of individuals each of whom has a substantial stake in the material or symbolic outcomes of public policy, making it economical for them to organize into interest groups. Irrigators, construction companies, real estate developers, banks, suppliers of agricultural inputs, chambers of commerce, naturalists, environmentalists, and others who stand to gain or lose from the bureau's projects make up the clientele of reclamation policies at the federal level. They form national, regional, and local associations on a temporary or permanent basis to press for their preferences and interests. The intensity of their support or opposition is proportional to the expected gains or losses. Group success is in large part a function of their numbers and their density, that is, the degree to which they present a united front. The greater the density and size of the group, the greater its influence. Within limits, these two characteristics serve as substitutes for each other. Examples of interest groups formed to influence reclamation policy are the League of the Southwest, which lobbied for federal projects on the Colorado River during the 1920s, and the National Reclamation Association, formed to protect the bureau from eastern interests opposed to federal subsidies toward irrigation about the same time.[40]

The clientele is organized around the supply of inputs to or outputs from particular agencies. Rents are created whenever an agency increases its purchases from factor owners or expands its supply of goods or services to customers at a price below opportunity costs. The owners of scarce factors, such as land, construction materials, or labor, and the recipients of subsidies, such as irrigators, make windfall gains that are quickly capitalized as property, licenses, or privileges and marketed legally or illegally. In the Westlands Irrigation District of California, for example, windfall gains accruing from the sale of "excess land" (acreage beyond Bureau of Reclamation limitations) have been estimated to average approximately $1.45 million

40. The origins of the National Reclamation Association are discussed in Golzé, *Reclamation in the United States*; for an account of the League of the Southwest, see Hundley, *Water and the West*, ch. 3.

per owner, while sales of nonexcess land have resulted in average gains of $49,000 per owner.[41] Naturally, those who paid the market price for this land or made "windfall" gains oppose any change in federal policy that will result in a capital loss for them.

Since support is generated in the process of rent creation, and since rents are quickly capitalized, it pays politicians and bureaucrats to spread the effect of the policy over time. Thus, many projects are started simultaneously and funded intermittently over a period of time longer than is necessary to complete the project economically. While bureaucratic inertia and rigidities were probably partly responsible for Reclamation Bureau projects taking so long to be completed, the incentive to spread benefits of the project over the careers of key bureaucrats and politicians also pointed in the same direction. As of 1973, the bureau had a backlog of construction projects valued at about $7 billion.[42]

5. *The Public.* This is the large mass of the population who pay the taxes and bear the cost of public water policies. They are the relatively passive "consumers" of reclamation projects, dams, and state water laws. Their political acts rarely go beyond voting for the politicians who influence these decisions, attending an occasional citizens' meeting, writing a letter to a newspaper, or making a small contribution to a party or candidate. In a study of public participation in water policies in the state of Washington, it was found that only 18 percent of respondents had acted politically to influence water policy. This figure may be exaggerated, since only 61 percent of the questionnaires were returned to the investigators, suggesting that activists were overrepresented in the sample.[43]

The public's attention is divided over innumerable policies, each of which receives relatively little attention. Due to high information costs, the votes of the public are largely ideological, cast in response to symbols such as "more water," "develop the desert," and "save the farmers," and rhetoric about water "shortages" and "droughts."

41. E. Phillip Le Veen and George E. Goldman, "Reclamation Policy and the Water Subsidy: An Analysis of Emerging Policy Choices," *American Journal of Agricultural Economics* 60 (1978): 929–934.

42. Statement by Gilbert Stamm, Commissioner of the Bureau of Reclamation, 1973–1977, quoted in *Reclamation Era*, Spring 1973.

43. See John C. Pierce, Kathleen M. Beatty, and Harvey R. Doerksen, "Rational Participation and Public Involvement in Water Resource Politics," in *Water Politics and Public Involvement*, ed. John C. Pierce and Harvey R. Doerksen (Ann Arbor, Mich.: Ann Arbor Science Publishers, 1976), p. 172.

Hence their influence over policy is marginal and diffuse, setting limits on the general level of taxation or signalling gross changes in opinions, attitudes, and the popularity of a particular politician, bureaucrat, or policy. It does not specify in any detail programs, organizational structure, budgets, or personnel.

Politicians and bureaucrats are the public entrepreneurs who make and implement government decisions, initiating new programs and agencies, carrying out administrative reorganizations, making policy proposals, planning projects, and deciding on budgets. Major Powell, Arthur Powell Davis, F. H. Newell, Elwood Mead, Senators Stewart and Newlands, President Theodore Roosevelt, and Secretary of Commerce and later President Herbert Hoover were the principal entrepreneurs of federal irrigation policies. They mobilized support for federal controls over western waters, designed and implemented large-scale reclamation and power-generation projects, and organized and defended the Bureau of Reclamation during its formative years. As we shall see, Ezra F. Scattergood, founder and builder of the Bureau of Power and Light, William Mulholland, chief engineer of the municipal aqueduct, and Mayor George E. Cyer performed similar functions in the city of Los Angeles.

Bureaucrats are usually responsible for the creation and growth of public agencies, while politicians provide the necessary support or unwelcome opposition.[44] Major Powell was the "father" and director for fourteen years of the U.S. Geological Survey, where federal irrigation policies and projects were planned for two decades before the passage of the National Reclamation Act of 1902. F. H. Newell and A. P. Davis, both of whom began their careers under Powell, were the first and second commissioners of the Bureau of Reclamation. The first served twelve years as director and the second nine years.

The entrepreneurs exercise control or influence by accumulating power. Power is net support, or the difference between the support and opposition generated with public policy from other entrepreneurs, the bureaucracy, the clientele, and the public. In government, entrepreneurs allocate and reallocate resources so as to generate maximum support over opposition, which strengthens and expands their

44. A similar argument is presented in J. T. Bennett and M. H. Johnson, *The Political Economy of Federal Growth: 1959–1979* (College Station, Tex.: Texas A&M University, 1980).

control or influence over public resources. As support for or opposition to specific agencies, policies, and individual bureaucrats or politicians shifts, so do the fortunes of organizations, programs, and individuals. Policy and administrative initiatives and changes are taken in the direction that yields an excess of support over opposition and in direct proportion to this difference. If the difference is small, the new agency or program is limited in authority, size or resources. For example, in his discussion of the political struggles that took place in Arizona during the 1940s and 1950s over groundwater pumping, Mann shows that near-equality in the forces supporting and opposing such controls resulted in a relatively weak policy, which was just as weakly enforced.[45]

Water policies at the local, state, and federal levels tend to redistribute income and wealth from the mass of the public to the ruling class of politicians, bureaucrats, and interest groups. Political struggles take place largely within this class as the various actors and entrepreneurs maneuver for position and power. Conflicts are often precipitated by the formation or increase in the power of interest groups that seek to change, initiate, or stop programs, elect, appoint, or remove politicians and bureaucrats, and otherwise change public policy. Though struggles take place within this class and though new groups and individuals may join it, the general tendency is for the class to benefit at the expense of the public, who subsidize the programs, projects, and policies with their taxes. The greater the inefficiency of the projects, the greater the public burden.

The redistributive nature of politics leads to the centralization of the government. Income and wealth are more efficiently redistributed the larger the jurisdiction of the government and the higher the level of decision. The larger the jurisdiction, the bigger the resource base, the smaller the per capita burden of taxes and hence the more passive the public. The higher the level of decision, the fewer the number of decisionmakers and hence the lower the costs of reaching agreement. The implication that federal subsidies to water projects should be greater than state projects, such as those undertaken by the California state government, is supported by fact.[46]

45. Dean E. Mann, *The Politics of Water in Arizona* (Tucson: University of Arizona Press, 1963), pp. 51–61.

46. Gardner Brown, "The Economics of Agricultural Water Use," in Thomas H. Campbell and Robert O. Sylvester, *Water Resources and Economic Development of the West*, no. 3 (Tucson, Ariz.: Western Agricultural Economics Research Council, 1955), p. 17.

4-24

The histories of the Bureau of Reclamation and the Water and Power Department of the city of Los Angeles can be interpreted and explained with the use of the theory developed here.

The Bureau of Reclamation

The Bureau of Reclamation is a direct descendent of Major Powell's plans and projects, even though it was established as a service within the U.S. Geological Survey in 1902, eight years after his resignation. Between 1898 and 1900 the Survey examined 147 reservoir sites, many of which no doubt had been selected by Powell during his tenure as director. Within five years, the bureau had secured congressional authorization for twenty-four projects, with at least one in every western state. Most of the projects were begun in great haste with little attention paid to "economics, climate, soil, production, transportation and markets."[47] It appears that the bureau was more concerned with establishing a political base in the West than in designing economical projects. In this respect, its strategy resembled that of an imperial power securing territorial control with the establishment of "missions" or "forts" over its domain.

It soon became apparent that the reclamation fund as originally established—revenues from the sale of public lands and repayments by irrigators—would not suffice to finance these projects. The costs of the projects turned out to be greater than originally estimated and the ability or willingness of irrigators to repay costs below what had been anticipated. In 1910, President Taft recommended and the Congress approved the issuance of $20 million in certificates payable out of the reclamation fund. In 1914, the repayment period was extended from ten to twenty years. These were the first in a long series of alterations of the original act designed to broaden the tax base with which to finance the projects and reduce the financial obligations of the irrigators.[48]

After two decades of reclamation, the bureau found itself facing increasing opposition in Congress, particularly from eastern interests who did not want to subsidize increased agricultural production in the West. Also, many of the bureau's project recipients were dissatisfied with the agency's long delays in construction and with the results of the projects. In 1932, western governors formed the National Reclamation Association in support of bureau policies and projects.

47. Raymond Moley, *What Price Reclamation?* (Washington, D.C.: American Enterprise Association, 1955).

48. Ibid., p. 7.

President Hoover came to the rescue of the bureau, which was faced with extinction, with the Boulder Canyon project. This massive undertaking, which included Boulder (now Hoover) Dam and the "all-American canal" to transport water from the Colorado to Southern California, expanded the bureau's jurisdiction to include hydroelectric power as a major source of revenue and political support. As Secretary of Commerce, Hoover served as federal representative on a commission made up of representatives of Colorado River states that drafted an interstate compact to divide the waters of the rivers between upper- and lower-basin states.[49] The agreement made it possible for the federal government to undertake the Boulder Canyon project. The bill was passed under the administration of President Coolidge but it followed Hoover's financial recommendations. Construction and contracts were implemented during Hoover's tenure as president. The city of Los Angeles became one of the operators of the dam's power facilities.

The Great Depression was a "boom" period for the bureau. President Franklin Roosevelt gave the bureau vigorous support. Between 1935 and 1937, $800 million of projects were authorized, mostly from the general fund. In 1939, a new act was passed relaxing repayment provisions and extending the repayment period up to forty years or more at the discretion of the Secretary of the Interior.

During World War II, new reclamation projects were postponed as resources were shifted to the war effort. But plans for postwar construction efforts continued apace. It was during this period, in fact, that the bureau became embroiled in long and costly struggles with the U.S. Corps of Engineers for control of water resources in the Central Valley of California and in the Missouri basin.[50] These represent the boundaries of the bureau's territory, which is limited to

49. Hundley, *Water in the West*, pp. 138–214.

50. For a discussion of struggles between the two agencies in California see Arthur Maass, *The Kings River Project* (Indianapolis: Bobbs-Merrill Company, 1952) and Maass, "Congress and Water Resources," *American Political Science Review* 64 (1950): 576–593. On the division of functional responsibilities in the Missouri Basin see Marian E. Ridgeway, *The Missouri Basin's Pick-Sloan Plan: A Case Study in the Congressional Policy Determination* (Urbana, Ill.: University of Illinois Press, 1955), and Carlos Davis Stern, "A Critique of Federal Water Resources Policies: Hydroelectric Power Versus Wilderness Waterway on the Upper Missouri River" (Ph.D. dissertation, Cornell University, 1971), ch. II.

Table 1-1. U.S. Presidents and Bureau of Reclamation Commissioners, 1902–1982.

President	*Commissioner*	*Year Appointed*	*Tenure in Office*	*Mean Tenure in Office*
T. Roosevelt	F.H. Newell	1902	12	
H. Taft				
W. Wilson	A.P. Davis	1914	10	8.75
C. Coolidge				
W. Harding	D.W. Davis	1923	1	
	E. Mead	1924	12	
H. Hoover				
F. Roosevelt	J.C. Page	1936	7	
	H.W. Bashore	1943	2	
H. Truman	M.S. Strauss	1945	8	6.8
D. Eisenhower	W. Dexheimer	1953	6	
J. Kennedy	F. Dominy	1959	11	
L. Johnson				
R. Nixon	E. Armstrong	1970	3	
G. Ford	G. Stamm	1973	4	3.7
J. Carter	R.K. Higginson	1977	4	
R. Reagan	R. Broadbent	1981		
				$\bar{x}$ = 6.7

the western states. The struggles resulted in a division of functions in which the bureau was given control over irrigation projects and the corps over flood control projects. This agreement paved the way for a major expansion in the acreage supplied by the bureau, which doubled between 1945 and 1965 from 4 to 8 million acres.

Today the acreage irrigated partly or totally with bureau-supplied water is roughly 11 million acres, or about 25 percent of the total. But the number of farms directly benefiting from federal water projects is only about 150,000. Thus, "the per farmer stakes can be high indeed. . . . Even a modest farm operation of 160 acres in California may receive a subsidy on water costs, the capitalized value of which

is in excess of $100,000."[51] Yet, roughly two-thirds of the lands supplied with bureau water are devoted to relatively low-value crops such as grains and forage.[52] Thus a small minority organized around irrigation has managed to redistribute income and wealth from the taxpayers to itself while misallocating water resources to relatively inefficient uses.

Table 1-1 shows that the bureau has evolved from a relatively autonomous agency controlled by bureaucrats who founded it to one under increased presidential control. For over three decades, the bureau was run by three men who had been active in promoting federal intervention in western irrigation before 1900: F.H. Newell, A.P. Davis, and Elwood Mead. This era of relative autonomy ended with Franklin Roosevelt, who appointed two commissioners. Since then, every new occupant of the White House except Lyndon Johnson has changed commissioners. Note that the mean tenure in office has declined steadily. Thus, the level of decisionmaking in the bureau has been raised to the maximum, a trend in keeping with the law of hierarchical centralization.

Water and Power in the City of Los Angeles: 1890–1950

The political history of water and power in the city of Los Angeles bears a striking resemblance to that of the bureau, at least on those aspects that are relevant to the theory presented in this paper. What follows is a necessarily brief description based on Vincent Ostrom's *Water and Politics.*[53]

Before 1900, the city of Los Angeles was served by privately owned water and power companies. The water company had a contract to supply the city with water from the Los Angeles River, which the municipal government owned in its entirety on the basis of judicial interpretations of Spanish and Mexican pueblo rights.

51. David Sechler and Robert A. Young, "Economic and Policy Implications of the 160-Acre Limitation in Federal Reclamation Law," *American Journal of Agricultural Economics* 60 (1978): 575.

52. William E. Martin, "Economies of Size and the 160-Acre Limitation: Fact and Fancy," *American Journal of Agricultural Economics* 60 (1978): 923–928.

53. Vincent Ostrom, *Water and Politics: A Study of Water Policies and Administration in the Development of Los Angeles* (Los Angeles: Haynes Foundation, 1953).

Toward the end of the nineteenth century, various reform groups argued that the city should develop its own water. After a long period of agitation, the Los Angeles City Water Company was forced to sell the properties it had developed under a thirty-year lease to supply the city with water. William Mulholland, superintendent of water works for the private company, became the chief water engineer for the city, a position he held for twenty-six years. He came to play the dominant role in the city's water policies.

Shortly after it acquired the local waterworks the city embarked on a vast new project, supported by the Bureau of Reclamation, to bring water to the city from Owens Valley, over 200 miles away beyond the mountains to the north. Valley residents opposed the acquisition of water and land by the city, waging a war against what became the Los Angeles Aqueduct.

This project was begun even though the Los Angeles River could have supplied additional water to serve the urban population at less cost. It was subsequently learned that a syndicate composed of several of the leading civic leaders behind the project had bought large tracts of land in the San Fernando Valley that were later irrigated with water from the aqueduct. In fact, for many years most of the new water was used for irrigation. The owners of the San Fernando properties were able to capitalize a very substantial increase in the value of their property as a result, another instance of the iron law of political redistribution.

Control over water in the Los Angeles River and Owens Valley provided the city with a weapon with which to expand territorially. An aggressive annexation campaign multiplied by many times the original tax base. Also, the city's bureaucrats were one of the leading entrepreneurial forces behind the Boulder Canyon project, the Colorado River project, and the Metropolitan Water District (MWD). An independent agency with taxing powers over an area of more than 3,000 square miles in Southern California, the MWD acts as a water wholesaler to cities and districts of the region. Its biggest project is the Colorado River Aqueduct, which brings water from Lake Davis over 240 miles away. The aqueduct has been relatively inefficient, supplying high-cost water while operating at less than half of capacity between 1940 and 1960.[54] However, it heavily subsidizes agricultural uses by taxing the urban populations, particularly Los Angeles

54. Hirshleifer, DeHaven, and Milliman, *Water Supply* p. 294.

residents. As late as 1951, 15 percent of the water used by the city, most of it imported from Owens Valley and the Colorado River, was sold to irrigators at a price less than half of what it cost the city to buy water from the district.[55] Thus, like the Bureau of Reclamation, Los Angeles City pursued a policy of territorial expansion implemented with inefficient projects paid for by the general taxpayer with the support of organized minorities, including irrigation interests.

The Los Angeles Aqueduct also became a source of electric power for the city. Its engineers took advantage of the drops in elevation from Owens Valley to the coast in order to generate hydroelectric power. In 1922, the city forced the private utility companies to sell their properties to the department, subsequently contracting with the Bureau of Reclamation to generate power from Hoover Dam.

As in the Bureau of Reclamation, policy decisions on water and power in Los Angeles were for decades dominated by the men who built the two systems. For twenty-six years Superintendent Mulholland was the most powerful voice on water policy. In 1929, he was succeeded by van Norman, who had been with the department since the construction of the Los Angeles Aqueduct. Van Norman served as director for thirteen years. Ezra F. Scattergood, founder and builder of the Bureau of Power and Light, served as its director for over thirty years and came to rule a veritable political machine. In 1940, a local newspaper observed that Scattergood's bureau "through its many ramifications, its advertising in many small community newspapers and throwaways, and its influence over the thousands of employees, virtually has constituted the balance of power in municipal elections."[56]

Politicians were unable to gain control over the water and power bureaucrats despite several spirited attempts. It proved more advantageous for politicians to support the bureau than to oppose it, as Mayor George E. Cyer discovered. During the 1920s, Cyer "unquestionably made the greatest contribution of any Los Angeles mayor to the development of the program of the Department of Water and Power; but his contribution was in providing political leadership for the policies formulated within the department."[57]

With the passing of Mulholland, van Norman, and Scattergood, who was forced to resign and given a lucrative consulting contract

55. Ibid., p. 308.
56. Ostrom, *Water and Politics*, p. 75.
57. Ibid., p. 108.

that took him away from Los Angeles, the water and power sections of the department were consolidated and centralized. Ostrom, writing in the 1950s, noted that those reorganizations "have tended to raise the level of decision about many of the operational and policy problems that were formerly resolved at the bureau or system level. The office of the General Manager and Chief Engineer has become a vital center of decision-making and leadership for the entire Department of Water and Power."[58]

Summary

This brief comparative study of the political histories of the Bureau of Reclamation and the Department of Water and Power of the City of Los Angeles has generated a number of parallels between the two agencies:

1. Bureaucratic entrepreneurship and dominance of policy occur during the first thirty to forty years of the agency's life.
2. Entrepreneurship by politicians is limited largely to providing support for the bureau's policies and plans; it is unprofitable for politicians to oppose the bureau consistently.
3. Evolution toward centralization or "raising the level of decision" occurs once the bureau founders pass away.
4. Territorial expansion is brought about with the construction of inefficient projects.
5. Costs of projects shifted to the general taxpayer over as large an area as possible, while benefits were concentrated on small minorities.
6. Finally, it is worth noting the role which economic or meteorological "crises" have played in expanding the power of these two agencies. The drought and depression of the 1890s over much of the West generated significant support for federal reclamation legislation, as did the severe winter of 1886 and the slump in agricultural prices between 1880 and 1900.[59] The Great Depression brought an infusion of public works money into the bureau.

58. Ibid., p. 103.

59. See Stegner, *Beyond the Hundredth Meridian*, pp. 294–304, and Glass, *Water for Nevada*, ch. II.

Ostrom notes the similar effect of drought on the support for public water projects in Los Angeles:

> Every major development in water resources programs and water administration has been closely correlated with drought cycles in Southern California. The drought of 1893–1904 produced the Los Angeles Aqueduct; the drought of the 1920s initiated the Boulder Canyon project, the Colorado River Aqueduct, and the organization of the Metropolitan Water District; and the drought of the 1940s produced the "dry cycle harvest" of annexations to the Metropolitan Water District.[60]

TOWARD THE REAPPROPRIATION OF WATER

The history of water policies in the West over the last hundred years shows very clearly the objectives, methods, and results of the political means in action. Irrigators and other beneficiaries of reclamation have capitalized rents created with public policy at all levels of government. Federal and big-city bureaucrats have built impressive monuments to their engineering skills, breaking world records for size and capacity of various dams. And the political entrepreneurs who made it all possible achieved a type of immortality for their efforts: Lake Powell, Lake Mead, Hoover Dam, Lake Roosevelt, and Lake Davis are now part of the political archeology of the Colorado River, the largest in the Southwest. The economic burden of these policies and projects has been borne by the mass of the taxpaying public, who have had to forego the income that western waters would have yielded in uses other than those dictated by the reclamation ruling class.

Currently, federal and state governments are under pressure to engage in ever larger projects or extend bureaucratic controls further. While economists demonstrate the inefficiencies of large-scale water transfers within California, engineers are now making plans for the transfer of tens of millions of acre-feet from the Pacific Northwest, Canada, and even Alaska. It is currently estimated that the most ambitious of these plans would cost $200 billion and take thirty years to build.[61] At the state level, new or tighter controls are being imposed on underground pumping, sometimes at the insistence of

60. Ostrom, *Water and Politics*, p. 234.

61. Arthur F. Pillsbury, "The Salinity of Rivers," *Scientific American* 245 (July 1981): 64.

428

the Bureau of Reclamation. For example, state control over underground pumping has been made a condition for the Central Arizona Project.[62]

Needless to say, continental water transfers and federally influenced controls over underground pumping would vastly increase the power that the bureau (now the U.S. Water and Power Resources Service) already exercises over the West. In light of the results of the bureau's projects and policies, this policy cannot be to the advantage of the region or the nation, even if a minority in the West and in Washington will undoubtedly continue to benefit.

Today, however, the political means appear to command less enthusiasm and support in the United States than scarcely a decade ago. The ideas of classical liberalism seem to be undergoing a revival in universities and other centers of learning. Perhaps it is only fitting that the American doctrine of prior appropriation be reconsidered as an institution for dealing with the problems of water shortages and conflicting interests over water allocation and use.

In principle, private property over surface and underground water could be reestablished with relatively simple rules of appropriation. The first step would be to establish the physical boundaries of rivers, streams, lakes, and aquifers. In the latter case they could reflect variations in pumping lifts. Next, each basin or watershed would be declared the corporate property of those who currently divert or pump water out of it and of those private or public organizations that manage bodies of water for instream uses. Each individual or organizational share of the basin would be proportional to the capacity of its water-using facilities, including diversion works, pumps, and volume reserved for instream uses. These shares would be bought and sold in an open market. State, federal, and private parks, municipalities, recreational associations, and others interested in nonconsumptive or instream uses of water would be free to purchase as many shares as they wanted. Water would thus be divided among uses and users according to its marginal value to each.

Shareholders would elect a set of officers who would appoint corporate managers. Each corporation would be free to decide whether to conserve, mine, export, or import water. All federal and state irrigation works would be sold or given to these corporations. The only role for government would be to enforce contracts among corporations and ensure that water is not transferred among owners or corporations without the consent of the participants.

62. Mann, *Politics of Water in Arizona*, pp. 50–60.

Such an arrangement is not without precedent. Interstate compacts now divide the water of rivers among states. Mead himself described how Utah streams were incorporated by existing appropriators: "All parties having used water from the stream come to an agreement as to their rights, usually on an acreage basis; then form a corporation and issue to each farmer or to each ditch company stock in proportion to their rights. The stream is then controlled by the water master, who is elected by the members of the corporation."[63] Mead thought that this solution would be practical only on smaller streams but did not explain how he reached this conclusion. There is no reason why this ingenious device cannot be applied not only to streams, rivers, and lakes, but to underground water as well.

63. Mead, *Irrigation Institutions*, p. 233.

Chapter 2

THE FEDERAL RECLAMATION PROGRAM
An Analysis of Rent-Seeking Behavior

Randal R. Rucker
Price V. Fishback

INTRODUCTION

In many regions of the western United States, nature has not been kind enough to provide abundant riparian land. Thus farmers in these regions have been forced to develop methods for bringing water to the land. The earliest efforts of this sort were made by enterprising individuals and groups acting without direct aid from the federal government. At the turn of the century western leaders and farmers sought direct assistance from federal authorities for irrigation works. Congress responded by providing interest-free loans and directing the Department of Interior to administer the projects and subsidies and to enforce certain restrictions. To carry out these directives, the Reclamation Service (later renamed the Bureau of Reclamation) was formed and has since played an important part in development of irrigation in the West.

The policies of the Bureau of Reclamation recently have been reviewed by the courts. A statutory interpretation that had exempted

In addition to the Pacific Institute, which has provided funding, both authors would like to thank the Economics Department of the University of Washington for general support. The authors would also like to thank Lee Alston, Terry Anderson, Robert Higgs, Ron Johnson, Mark Plummer, and George Sheldon for their comments and suggestions on earlier drafts of this paper. We owe a special debt to Sharie Olson for her diligent efforts in typing this paper. The responsibility for any errors remains with the authors.

irrigators in the Imperial Valley from acreage limitations since 1933 was overturned by a 1977 circuit court ruling (*United States* v. *Imperial Irrigation District*), which was reversed by the Supreme Court in 1980 (*Bryant* v. *Yellen*, 100 Supreme Court, 2232, 1980). In *National Lands for the People Inc.* v. *Bureau of Reclamation*, the bureau was ordered to propose new policy rules for the enforcement of acreage limitations and other legislative restrictions regarding residency on and disposal of lands receiving federal irrigation water.[1] Changes in the optimal size of farms have led to questions concerning the need for and proper size of acreage limits. Analysis of the effects of policy changes in earlier periods is useful for understanding the possible effects of recent changes in the bureau's policies.

In this paper we will use a model of rent-seeking behavior to analyze earlier actions of the irrigators and bureaucrats who attempted to increase the size of the subsidy and to alter its distribution. Two dimensions of federal policy will be discussed in detail: (1) the provisions established for repayment of construction costs by the settlers (which determined the size of the subsidy), and (2) the limitations on the amount of land eligible to receive irrigation water from federal projects (which determined the distribution of the subsidy). Since these policies have undergone important changes during this century, we will examine the form, effects, and possible explanations of these changes.

In the next section a rent-seeking model that outlines the incentives of irrigators and bureaucrats is described. The following section discusses the nature of the subsidy provided for federal irrigation projects and how irrigators and bureaucrats gained from changes in the rules that determined its size. Constraints placed by Congress on the subsidy's distribution among irrigators and the efforts of irrigators and bureaucrats to capture the rents are then discussed, followed by a brief summary and conclusion.

THE ANALYTICAL FRAMEWORK: A RENT-SEEKING MODEL

Economists define rents as the returns to an owner of a resource in excess of the opportunity costs of that resource. When rents exist,

1. Nancy Jones, "Proposed Rules for Administering the Acreage Limitation of Reclamation Law," in *Natural Resources Journal* 18 (October 1978): 936–937.

profit-maximizing individuals will compete to capture them by diverting the resources they command into those activities. This rent-seeking behavior will be observed both when resource movements are guided by the invisible hand of the market system and when those movements are restricted by political barriers.[2] Interest groups will also seek to create rents through the establishment of political barriers. Under any institutional setting, competition by rent seekers dissipates the rents to the marginal firm or individual if property rights to them cannot be established and enforced.[3] In this paper the efforts of irrigators and bureaucrats to capture the rents created by the federal irrigation program, which was initiated by Congress in response to pressure from special interest groups, are analyzed within this simple rent-seeking framework. The actions of the members of Congress, the judiciary, and the executive branch are treated as exogenous constraints on the behavior of bureau members.

The forms of the rent-seeking behavior of irrigators and agents of the bureaucracy in charge of federal reclamation policy were largely determined by two important cornerstones of that policy. The first of these was the repayment system established for federal irrigation projects in which irrigators were provided with interest-free loans for construction costs. The fact that this repayment scheme later proved to be extremely flexible encouraged rent-seeking activities designed to increase the value of the rents from federal reclamation. The second cornerstone was the limitation placed on the amount of federally supplied water that an individual landowner was eligible to receive. This excess-land law or acreage limitation was crucial in determining the distribution of the benefits from the federal irrigation program. These limitations and the regulations for their implementation, the most controversial aspects of reclamation policy, have been important in determining the nature of activities designed to increase the share of the rents going to the owners of excess lands.

2. James M. Buchanan's "Rent-Seeking and Profit Seeking" in *Toward a Theory of the Rent-Seeking Society*, ed. James M. Buchanan, Robert D. Tollison, and Gordon Tullock (College Station, Tex.: Texas A&M University Press, 1980) distinguishes between profit-seeking under a market structure and rent-seeking under government actions that interfere with the market adjustment process. Use of the term "rent-seeking" in this paper is consistent with this definition.

3. For a discussion of the various methods of dissipating rents through nonprice adjustments, see Steven N. S. Cheung, "The Theory of Price Controls," *Journal of Law and Economics* 17 (April 1974): 53–71. Cheung emphasizes that complete dissipation of rents occurs only if property rights to those rents cannot be established and enforced.

Irrigators on a project included individuals who owned land before the project was initiated as well as settlers who wanted to homestead public lands or buy the excess private lands. Both of these groups received the subsidy from federal involvement in irrigation, dissipating the rents from the subsidy by expending resources in various ways to increase the size of the subsidy or their share of it. For example, settlers homesteaded public lands before federal project water made them irrigable, while large landowners often spent considerable resources devising methods to avoid the acreage limitation.

Members of the Bureau of Reclamation and the Department of Interior were charged with enforcing and administering the reclamation policy. They are not treated as passive respondents to either the directives of Congress or the pressures from irrigators but rather as rational maximizers who developed objectives and policies of their own within the constraints imposed by Congress. Since bureaucrats cannot directly receive the subsidy on federal irrigation projects, the bureaucratic rent-seeking model suggests that their primary goals were to increase their salaries, job security, and power within the political system. These goals might be attained through increased legislative demand for the bureau's output, expanded administrative control and discretionary power over allocation of irrigation water, enlarged staffs, and increased budgets. For example, more political power and higher salaries are often correlated with larger staffs and increased budgets, while increased discretionary power over the budget gives the bureaucrat more freedom to fund pet projects.[4]

That tradeoffs between complying with congressional constraints and achieving these goals occur is indicated by the wide variety of administrative and enforcement policies adopted for different reclamation projects. Our model suggests that bureaucratic actors attempted to capture a portion of the rents from federal reclamation

4. See Roger L. Faith, "Rent-Seeking Aspects of Bureaucratic Competition," in Buchanan, Tollison, Tullock, *Rent-Seeking Society*, pp. 332–345; and Terry Anderson and P.J. Hill, "Establishing Property Rights in Energy Efficient vs. Inefficient Processes," *Cato Journal* 1, no. 1 (Spring 1981): 87–105. An alternative to the rent-seeking model, if attempts to increase the bureau's discretionary control are seen as a means of increasing the demand for the bureau's budget, is a budget-maximizing model such as that of William A. Niskanen in *Bureaucracy and Representative Government* (Chicago: Aldine-Atherton, 1971), and "Bureaucrats and Politicians," *Journal of Law and Economics* 18 (December 1975): 617–643. This model has been applied to the Department of Interior's policies on land disposal for grazing lands by Gary Libecap in "Bureaucratic Opposition to the Assignment of Property Rights: Overgrazing on the Western Range," *Journal of Economic History* 41 (March 1981): 151–158.

with these policies. The process through which the irrigation rents were created, the constraints within which rents were sought, and the actual forms of the rent-seeking activities are discussed in detail in the following sections.

REPAYMENT OF CONSTRUCTION COSTS: ESTABLISHING THE SUBSIDY AND INCREASING ITS VALUE

The congressional policy of providing direct subsidies to irrigators on federal reclamation projects was established in response to pressures from western interest groups with the passage of the Reclamation Act of 1902. The initial subsidy took the form of a ten-year interest-free loan for construction costs of federal projects. Over time, the value of this subsidy to irrigators was increased significantly by such modifications as extensions in the term of the repayment period, allowances for development periods during which no payments are required, and the adoption of the policy of using power revenues from multiple-purpose projects to repay irrigation costs in excess of irrigators' ability to pay. This section describes the events leading to the establishment of this subsidy and the efforts by irrigators and members of the Bureau of Reclamation to increase its value.

Irrigation of arid lands in the West began before American settlement of the frontier. Indians were irrigating their lands when the Spanish first explored California. As early as 1776, the Spanish padres at Mission San Diego de Alcalca irrigated their grapes and gardens. The first efforts to use irrigation methods by American settlers were made by the Mormons upon their arrival in Utah in 1847. By 1890, settlers in California, Wyoming, and Colorado had irrigated over 3 million acres.[5]

Most of the early private irrigation projects involved little more than the construction of ditches and canals for diverting waters from the rivers onto adjacent farmlands. Opportunities for building projects of this type at low cost were soon exhausted. New projects involved the construction of canals for carrying water to lands further

5. Frederick Merk, *History of the Westward Movement* (New York: Knopf, 1978), p. 507.

from the streams and rivers, and of dams and reservoirs for storing water. Western leaders and landowners sought federal aid for these expensive undertakings arguing that further successful settlement of the public lands of the West required that they be irrigated, and that this irrigation would require direct government assistance.

Congress had previously tried to encourage irrigation with the Desert Lands Act of 1877 and the Carey Act of 1894, which offered tracts of land at low prices to those settlers who irrigated the land. Neither of these acts, which are discussed in detail in the next section, had significant effects on irrigation in the West. The movement calling for direct assistance from the federal government gathered momentum. A series of "Irrigation Congresses" were called to press for federal aid and to develop an irrigation policy for promoting the successful settlement of the West that would be acceptable to representatives from all the western states. Additional support came from an 1897 report prepared by Captain Hiram H. Chittenden of the Army Corps of Engineers, which stated that "a comprehensive reservoir system in the arid regions of the United States is absolutely essential" and that "it is not possible to secure the development of such a system except through the agency of the General Government."[6]

Western leaders were encouraged by the responses to their efforts. By 1900 the Geological Survey and the Department of Agriculture were receiving regular appropriations from Congress for investigating different aspects of the irrigation problem.[7] In the 1900 presidential election the platforms of the major parties included planks favoring the reclamation of arid lands. Despite this early success, a federal reclamation bill introduced early in 1901 failed to obtain congressional approval.

Theodore Roosevelt provided the support that finally resulted in the passage of a federal reclamation bill. In his first presidential message to Congress, he declared himself to be strongly in favor of federal construction of western irrigation projects.[8] A compromise reclamation bill was quickly drawn up, passed by a comfortable margin, and signed into law as the Reclamation Act of 1902. Defenders

6. Quoted in Norris Hundley, *Water and the West* (Berkeley, Calif.: University of California Press, 1975), p. 10.

7. For a description of these appropriations, see Alfred R. Golzé, *Reclamation in the United States* (New York: McGraw-Hill, 1952), pp. 21–23.

8. For this portion of Roosevelt's message, see Frederick H. Newell, *Irrigation* (New York: Thomas Y. Crowell and Co., 1906), pp. 393–396.

of the act declared that federal involvement in irrigation was constitutional since it promoted the general welfare by providing a release for overpopulated areas of the East and by conserving the nation's natural resources. To the argument that output produced using the water from federal projects would provide competition detrimental to farmers in the Midwest and East, defenders responded that most of the produce grown in the West would also be consumed there, and that any surplus could be exported to the Orient. It was also noted that even though 112 million acres of federal land had been disposed of in the 1890s, the prices of agricultural goods were about the same at the end of the decade as at the beginning.

Opponents of the act also argued that the benefits from these projects would not justify the expenses, which would be borne by the nation's taxpayers. To allay these objections, the authors of the act established a revolving Reclamation Fund through which the federal projects would pay for themselves and provide funds for additional projects. The funds that would provide the base for the Reclamation Fund were to be revenues from the sale of public lands in the western states. Later acts supplemented the Reclamation Fund with other sources of funds, including proceeds from sales of oil leases and from potassium royalties, as well as revenues from federal power licences, public power revenues, and the sale of town lots on the projects. According to the 1902 act, the settlers in a given project area were to agree to repay the construction costs within ten years.

The subsidy given to irrigators took the form of an exemption from interest charges on the loan for construction costs.[9] Apparently, this feature received little attention during the discussion of the bill. It was obvious that western interests wanted a subsidy, but why Congress opted for this particular form and not a direct payment is unclear. Since an interest subsidy is more subtle than a direct subsidy, it is possible that this form was chosen to make the subsidy more acceptable to nonwestern congressmen, whose constituents were subsidizing irrigation projects.

The value of the subsidy initially obtained by irrigators was significant. To get an idea of this value, imagine a project where the irrigation works have been completed and the cost of the project was $10,000. If the settlers paid the costs immediately, the present value

9. Another subsidy—the difference between the price charged to cover costs and the market price of the water—would have been received by settlers even if they had been required to repay construction costs with interest (assuming, of course, that the market value of the water was greater than the price required to recover costs).

of the payment would be the full value of construction costs, $10,000. The Reclamation Act of 1902 allowed the settlers to repay construction costs over a period of ten years in equal annual installments. The present value of this stream of payments depends on the discount rate used. A minimum measure of the value of the subsidy would be obtained by using the interest rate for risk-free investments. At a rate of 3 percent (which was the approximate rate of interest on risk-free government bonds at that time), the present value of the payments would have been about $8,530, implying a subsidy of about 14.7 percent of construction costs. However, a more accurate appraisal of the size of the subsidy would take into account the risky nature of investment on irrigation projects. Table 2-1 demonstrates that at a discount rate of 10 percent (which probably is a better approximation of the rate of interest faced by settlers), the value of the initial interest subsidy was almost 39 percent.

After the passage of the Reclamation Act, no time was wasted in allocating responsibilities and initiating projects. Frederick H. Newell, former chief hydrographer for the Geological Survey, was appointed chief engineer in charge of the service. Eleven days after the act was passed, land for six projects and surveys was withdrawn from disposal under other federal acts for use in reclamation projects. The construction of four projects was authorized by the end of 1903, and in 1904 and 1905 sixteen more projects received authorization.

It was not long before the Reclamation Service encountered financial problems and settlers began seeking increases in the size of the subsidy. The revenues from public land sales proved to be inadequate for financing the construction of the service's proposed projects, necessitating a congressional loan to the Reclamation Service in 1910. Settlers on reclamation projects complained that costs had been underestimated—Newell's estimate of $5 per acre was well short of the actual cost of $50 to $100 per acre. Settlers also objected that construction was not being completed on schedule. Settlers seeking to capture rents from federal irrigation water were homesteading land before the projects were completed. They were expected to be residents of the land to maintain title, but much of the land was virtually useless without irrigation water. When projects were not completed on schedule as expected by the settlers, many were faced with the choice of starving or relinquishing their rights.

Defaults on repayments were often attributed to the settler's inexperience and lack of the substantial capital needed to prepare arid

Table 2-1. The Interest Subsidy: Subsidized Proportion of Costs.

Payment Plan	*Rate of Discount*[a]		
	3	*6*	*10*
		(percent)	
10-year repayment period; equal installments	14.7	26.4	38.6
20-year repayment period; equal installments	25.5	42.5	57.5
20-year repayment period; graduated installments[b]	28.9	47.8	64.0
20-year repayment period; graduated installments with grace period and down payment[c]	30.7	50.3	66.7
40-year repayment period; equal installments	42.3	62.5	75.5
40-year repayment period; equal installments with 10-year grace period	57.0	79.0	91.0

a. These subsidies were calculated by subtracting the present value of the payments (for any given schedule) from the construction costs, dividing that difference by the construction costs, and multiplying by 100.

b. Repayment schedule (outlined in the act of August 13, 1914) was 2 percent of construction costs for the first four years, 4 percent for the next two years, and 6 percent for the final fourteen years.

c. Repayment schedule (outlined in the act of August 13, 1914) was 5 percent of construction cost down, followed by a five-year development period, then annual payments of 5 percent for five years and 7 percent for the final ten years.

lands for irrigation. The Reclamation Service described this capital problem on early projects:

> Many of the settlers are attempting what is for them practically an impossibility; they are trying to start a farm-business which requires when fully developed as a "going concern" a capital or investment frequently of from $8,000 to $10,000 or more. They are attempting to do this usually with a capital of perhaps only a fourth as much. A 40-acre irrigated farm in best condition represents practically the investment in time and labor as above stated, of from $100 to $200 per acre or more in improvements, in subduing the soil, and in stocking the farm.[10]

10. U.S. Department of Interior, *Thirteenth Annual Report of the Reclamation Service 1913-1914* (Washington, D.C.: Government Printing Office), p. 14.

Agricultural depressions and bad harvests sometimes left settlers starving on their lands. In other areas organized resistance to the repayment of construction charges emerged, even though agricultural conditions were favorable. As a result, defaults on repayment contracts were common during the first three and a half decades of the Reclamation Service's existence—as of 1923, less than $16 million of the $143 million expended on federal irrigation projects had been repayed.[11] In a letter to Compton I. White, chairman of the House Committee on Irrigation and Reclamation in 1937, Charles West, the Acting Secretary of Interior, described the repayment problem of the Bureau of Reclamation:

> The revolving feature of the fund has been seriously retarded and there are projects where water has been available for 29 years and only six annual construction installments have been paid. There has often been organized resistance to the repayment of these charges, which is still being continued, and this notwithstanding the fact that nearly all of the projects have just passed through a successful year and in some cases the most successful year in their entire history.[12]

During this period water users sought and received increases in the size of the subsidy from Congress in five forms: (1) extension of the repayment period, (2) graduation of the scheduled payments, (3) postponement of the date when the first payment was due, (4) increased flexibility in the repayment schedule, and (5) moratoria on repayments during periods of crop failure. The first increase in the interest subsidy was granted in the Reclamation Extension Act of 13 August 1914, which authorized repayment contracts with twenty-year terms, graduated payment schedules, and five-year grace periods on new projects. The grace period and graduated payments were justified on the grounds that the burden on settlers during the years when they were establishing themselves would be reduced if smaller payments were required at the beginning of the repayment period.

The effects of these changes on the value of the interest subsidy are shown in Table 2–1. The extension of the repayment period to twenty years accounted for most of the increase in the subsidy's

11. U.S. Congress, Senate, "Federal Reclamation by Irrigation," Senate Document 92, 68th Cong., 1st sess., p. xi.

12. Quoted in U.S. Congress, House of Representatives, "Relief to Water Users on Federal Reclamation and Irrigation Projects," House of Representatives Report No. 1440, 75th Cong., 1st sess., pp. 3–4.

value. At a 3 percent discount rate, which provides a minimum estimate of the value of the subsidy, an increase in the term of a contract (calling for equal annual payments) from ten to twenty years increased the value of the subsidy from 14.7 percent to 25.5 percent of construction costs. For existing contracts that were renegotiated as a result of the 1914 act, the new repayment schedule required payments of 2 percent of the remaining construction costs for four years, followed by payments of 4 percent for two years, and 6 percent for the final fourteen years. This graduated scheme increased the value of the subsidy to 28.9 percent of construction costs. The formula for new contracts required a down payment of 5 percent, followed by a five-year grace period, then five annual payments of 5 percent and ten payments of 7 percent. This repayment scheme increased the subsidy's value on new projects to more than 30 percent of construction costs.

Defaults on repayment continued to be a problem. On some projects, these defaults could be attributed to the distress in several agricultural areas in the early 1920s. However, there were other projects "where powerful influences [sought] on various pretexts to evade paying. On one project the water users organization in an appeal for blanket deferment said: 'Not one irrigator on this project can pay anything.' "[13] The Bureau of Reclamation denied the blanket deferment but said that they would listen to individual requests for deferral on that project. Their belief that most of the settlers were able to pay appears to have been confirmed when "thousands of dollars came at once into the reclamation treasury" from individuals who could not give firm reasons for not paying their debt.[14] Urged by representatives from projects with large delinquencies to grant moratoria on past due debts, Congress granted this relief in 1921, 1922, and 1924.[15]

In 1923 a fact-finders' committee was appointed to investigate the reclamation program. In response to their recommendations, a new repayment scheme that allowed payments to vary with the productivity of the land was authorized in the Fact Finders Act of 1924. The annual charge for farms in a given district was to be 5 percent

13. U.S. Department of Interior, *Twenty-Fourth Annual Report of the Bureau of Reclamation for the Fiscal Year Ended June 30, 1925* (Washington, D.C.: Government Printing Office), p. 6.

14. Ibid., p. 6.

15. Ibid., p. 8.

of the average gross crop value for the preceding ten years. In practice payments based on this scheme were so small that repayment periods occasionally extended beyond seventy or eighty years. Authority to negotiate repayment contracts under this plan was quickly repealed in the Omnibus Adjustment Act of 1926.

The crop value repayment schemes were replaced by repayment contracts with forty-year terms in the act of 1926. Many of the contracts written under this act called for repayment on a graduated scale, which increased the value of the subsidy beyond the 42.3 percent subsidy that would have resulted under a schedule with equal installments (see Table 2-1). This rearrangement of payments was designed to relieve settlers during a period of low agricultural prices from 1926 to 1930. Unfortunately, crop incomes continued to decline after 1930 and payment stopped completely on some projects. At the request of the settlers, Congress again granted moratoria on payments from 1931 to 1936. Once again it should be noted that water users were engaging in rent-seeking behavior. Although some farmers were starving, others were less affected by the depression and simply refused to pay the construction charges.

The default problem appears to have been solved with the Reclamation Project Act of 1939, which empowered the bureau to enter into more flexible repayment contracts. There were no defaults on contracts negotiated under this act.[16] It is not clear whether the default record has improved as a result of the longer repayment periods and graduated payment schemes per se, or whether these improvements resulted from the increased subsidy that accompanied these modifications.[16a] Contracts negotiated under section 9(d) of the act were permitted repayment periods of forty years with development periods of up to ten years. The contracts could be written to allow for charges that varied with the productivity of different classes of land within the project area and for annual charges that depended on gross crop values. Table 2-1 shows that the grace period increased the subsidy's discounted value to 57 percent of construction charges. The flexibility of the payments under these contracts increased the subsidy's value even more by reducing the burden of the risk borne by irrigators. Under Section 9(e) of the act, authority was granted to negotiate contracts that did not require complete repayment of construction costs within forty years. In these contracts, water recipients were charged rates sufficient to cover an appropriate share of annual operation and maintenance costs and fixed construction costs. From the viewpoint of irrigators, the problem with these contracts was that the contracted water rights did not become attached to their land when the contract ended. Later legislation assured irrigators of their rights to renew these contracts and provided that payments above and beyond operation and maintenance costs would be credited towards repayment of construction costs if they decided to switch to 9(d) contracts.

Later acts made minor changes in the general rules for repayment and authorized more flexible repayment schemes for specific projects. The Small Reclamation Projects Act of 6 August 1956 authorized interest-free loans for small projects (overall cost of less than $10 million) with a repayment period of up to fifty years. Increased flexibility in repayment contracts was provided by the Variable Plan Amendment of 1958, which permitted adjustments in the installment payments (with the constraint that charges must still be repaid within forty years). Several of the special congressional acts authorizing specific projects have specified repayment periods considerably longer than forty years, including the Kennewick division of the Yakima project (sixty-six years), the Mancos project in Colorado (sixty years), and the Paonia project, also in Colorado (sixty-eight years).[17] However, the Reclamation Project Act of 1939 was the last act in which major modifications in the general rules for repayment of construction costs were made.

A rent-seeking model suggests that since members of the Bureau of Reclamation and the Department of Interior were unable to directly capture the gains from the provision of irrigation water, they would have attempted to appropriate the rents obtainable from administering the program. By winning congressional approval that would lead to increased appropriations and administrative power, the bureaucrats could have increased the potential size of the administrative rents. Where possible, the members would have sought to increase the part of their budget that was not part of the common pool

16. Frederick Warne, *The Bureau of Reclamation* (New York: Praeger, 1971), p. 63.

16a. That is, the repayment period could have been extended and payments shifted beyond the difficult transitional years of a project without increasing the value of the subsidy, simply by levying appropriate interest charges. Since relief measures have generally been accompanied by implicit increases in the value of the subsidies, it would be extremely difficult to empirically identify these separate effects.

17. Golzé, *Reclamation*, p. 248.

of federal revenues. This portion of the budget provides a budgetary base that is not subject to direct congressional appropriations and therefore is not shared with other agencies.[18]

One important part of the bureau's budget was the repayment of construction charges into the Reclamation Fund. Increasing the stream of repayments into the fund was in the bureau's interest for three reasons. First, successful repayment could be pointed to as an indication that the bureau was successfully carrying out its congressional mandate. Repayment demonstrated that a particular project had been successfully irrigated and that the bureau was "paying its own way." Second, these funds did not go into the common pool of federal revenues to be reassigned by Congress. Instead, the funds were automatically allocated to the Department of Interior for new reclamation projects chosen by the bureau and approved by Congress. Third, the fact that settlers on a project were able to make their scheduled payments indicated that they were maintaining a reasonable standard of living. Keeping water users happy benefited the members of the bureau, since complaints from unhappy irrigators caused Congress to view the bureau's activities with disfavor. Moreover, discontented settlers could make a local agent's job unpleasant.

Given that the bureau wanted successful settlement and repayment of costs, it was clearly in their interest to have qualified applicants settling on their projects. The writer in the bureau's annual reports in its first twenty years consistently pointed to the settlers' lack of capital, experience, and perseverance as primary reasons for failures to repay charges. At the suggestion of a fact-finders' committee on reclamation appointed by the Secretary of Interior in 1923, the bureau received the authority to require settlers to meet specified qualifications in the Fact Finders Act of 1924. The value of this authority was enhanced by the fact that the guidelines established in the act were general, allowing the bureau extensive leeway in establishing the qualifications. The secretary was authorized "under regulations to be promulgated by him, to require of each entry to public lands on a project, such qualifications as to industry, experience, character, and capital, as in his opinion are necessary to give reasonable assurance of success by the prospective settler."[19]

18. For a similar argument in a different context, see Anderson and Hill, "Establishing Property Rights," p. 91.

19. Subsection C of the Second Deficiency Act (Fact Finder's Act), 1924, (43 Stat. 702, 43 U.S.C. 433).

Another method of achieving the objectives of successful settlement and repayment was to extend the term of the repayment period. These extensions generally reduced the annual payments assessed against the settlers, thereby increasing the likelihood that payments would be made and improving repayment records. Until the bureau began funding larger multipurpose projects, the administrative control was taken over by water users after a certain percentage of the construction costs were repaid.[20] This gave the bureau additional incentive to lengthen the repayment period and maintain control over these projects. However, extensions of the repayment period were not entirely in the bureau's interest, since they reduced both the present value of the payment stream to be received from settlers and the flow of revenues into the Reclamation Fund.

Because most of the repayment extensions were legislated, it is difficult to directly observe the bureau's desired tradeoffs. The fact that the Department of Interior did not take a general stand against extensions is some indication that at the margin, the bureau preferred contented settlers and impressive repayment records to rapid payment. As the scope of the bureau's undertakings broadened to include huge multipurpose projects, increasing the relative importance of special congressional appropriations, it might be expected that the bureau would be relatively less concerned with maintaining the flow of repayment revenues into the Reclamation Fund and more concerned with presenting a rosy picture of their operations to Congress.

Two specific instances show the nature of the bureau's interest in an impressive repayment record: (1) the bureau's opposition to blanket repayment moratoria in the 1920s and 1930s, and (2) their support of the fact-finders' committee's suggestion to write off construction costs on some projects. In 1924 the bureau opposed repayment moratoria, arguing that many irrigators who were able to make payments were using the agricultural depression as an excuse to postpone repayment. Many settlers who were suffering on project lands had "sacrificed" and made their payments, providing a limited flow of revenue into the Reclamation Fund. Blanket moratoria were expected to encourage the settlers to join "the repudiation ranks," which would temporarily cut off the entire flow of revenues into the fund without increasing the probability of repayment by those settlers. Members of the bureau were in favor of granting moratoria for repayments only on projects of their choice during the 1920s and

20. Warne, *Bureau of Reclamation*, p. 68.

1930s, an additional discretionary power that would have allowed the bureau to avoid problems of nonpayment by successful water users.[21]

Several of the early projects were failures. Irrigation water had not improved the lands' productivity enough to support farming; consequently the bureau did not expect the construction costs ever to be repaid. Faced with increasing objections to the financial failure of their projects, the bureau sought a way to exclude these projects from the Reclamation Fund and improve their collection record. Settlers on the projects and the bureau supported the recommendation of the fact-finders' committee of 1923 that $27 million in construction costs on early projects be written off as nonrecoverable losses. Congress responded in 1926 by passing the Omnibus Adjustment Act, which allowed costs on specific projects to be written off.

In the early 1920s the Bureau of Reclamation was confronted with calls for the end of the reclamation program from other federal agencies.[22] The Department of Agriculture led this opposition to continued construction of irrigation projects, arguing that subsidies given to irrigation farmers worsened conditions for all farmers by creating an "oversupply" of farm goods in an already depressed market, and furthermore that many of the projects had been financial failures. The Department of Interior replied that the lands on the reclamation projects were used primarily for growing specialty crops that were not in oversupply. They also maintained that publicly reclaimed land was providing the basis for a society of independent farmers in the arid regions of the West, a nonfinancial benefit that must be considered when evaluating the federal reclamation program.

This crisis was averted with the passage of the Boulder Canyon Act in 1928, which marked the beginning of the bureau's involvement in the development of multipurpose projects. The benefits of such projects as Hoover Dam in the Boulder Canyon project and Grand Coulee Dam in the Columbia River basin included the provision of public power, delivery of municipal water, flood control, and improved river navigation in addition to irrigation. The movement into

21. See *Twenty-Fourth Annual Report*, p. 9, and Charles West's letter to Compton White cited in note 12.

22. Often members of agencies dissipate rents while competing with other agencies to maintain or increase the legislative demand for their services. For a discussion of this form of competition, see Faith, "Rent-Seeking Aspects," pp. 332–345.

multipurpose projects eventually provided irrigators with additional subsidies.

The bureau competed with the Army Corps of Engineers for the rights to build several flood-control projects, including the Central Valley and Missouri River basin projects. The Army Corps had a slight advantage until 1944, when irrigation on flood-control projects was placed under reclamation law. The two agencies actually combined forces on the Missouri River basin project to prevent the creation of a new competitor similar to the TVA.[23]

The Reclamation Project Act of 1939 authorized the allocation of costs among different classes of project beneficiaries. Revenues from power and municipal water users were to be applied to their respective shares of the total costs. Portions of the costs were also to be charged to flood control and navigation on a nonreimbursable basis (see Section 9 [b]). Costs allocable to the preservation of fish and wildlife and to construction of recreation facilities have also been exempted either partly or wholly from reimbursement.[24] This policy eased the repayment burden on irrigators but did not furnish them with a direct subsidy. The major irrigation subsidy generated by the new allocation of costs was derived from the policy of using revenues from municipal and industrial power users to pay the portion of irrigation costs judged to be beyond their irrigators' ability to pay. The bureau gained from this practice, since the lower charges to irrigators enhanced the probability of repayment of their remaining share of the construction costs. This arrangement, while not authorized in the general reclamation law, was expressly authorized for a number of individual projects and has been practiced on most other federal projects.[25]

23. Merk, *History of Westward Movement*, p. 543, and Mary Montgomery and Marion Clawson, *History of Legislation and Policy Formation of the Central Valley Project* (Berkeley, Calif.: Bureau of Agricultural Economics, 1946), pp. 228–238.

24. For discussions of acts relating to the reimbursement policies for these uses, see Charles Meyers and A. Dan Tarlock, *A Coursebook in Law and Public Policy* (Mineola, N.Y.: Foundation Press, 1971), p. 539, and Robert E. Clark, *Water and Water Rights*, vol. 2 (Indianapolis: Allen and Smith Company, 1967), pp. 147–153.

25. See Clark, *Water Rights*, p. 272, for examples of projects where this policy has been authorized. A second subsidy to irrigators may exist if costs are not allocated "correctly" among the project beneficiaries. Meyers and Tarlock (*Coursebook in Law*, p. 545) believe that costs apportioned to nonreimbursable uses are overestimated, which has the effect of increasing the subsidy to irrigation users by reducing the portion of total costs assigned to them. Whether this type of overestimation actually occurs is difficult to determine, since values of the nonreimbursable benefits are not generally determined through market transactions.

Table 2-2. The Power Subsidy.[a]

Project	Costs Allocated to Irrigation	Costs to be Repaid by Irrigators	Percentage of Irrigation Costs Subsidized
Central Valley California	687,152,000	606,646,000	11.1
Chief Joseph Dam[b] Washington	11,083,200	6,050,000	45.4
Collbran Colorado	6,105,000	1,089,101	82.2
Columbia Basin Washington	745,111,398	135,916,400	81.8
Fryingpan-Arkansas Colorado	69,946,000	50,512,300	27.8
Rouge River Oregon	18,064,000	9,066,500	49.8
San Angelo Texas	8,853,904	4,000,000	54.8
The Dalles Oregon	5,994,000	2,550,000	57.5
Venturia River California	18,273,128	10,746,300	41.2
Washita Basin Oklahoma[c]	10,403,011	8,221,000	21.0

a. On some of these projects, a portion of the subsidy to irrigators came from industrial and municipal users.

b. Includes costs and repayments from Foster Creek and Greater Wenatchee Divisions.

c. Includes costs and payments from Fort Cobb and Fass Divisions.

Source: Reclamation Payments and Payout Schedule, Department of Interior, Bureau of Reclamation (Government Printing Office, 1965).

An indication of the value of this subsidy is given in Table 2-2 for several projects. The third column shows the percentage of the total costs allocated to irrigation that has been subsidized by other project beneficiaries. These figures, which range from 11 percent to over 82 percent, demonstrate that this subsidy has been extremely valuable to irrigators on a number of federal reclamation projects.

By seeking legislation making repayment schedules longer and more flexible, irrigators were able to increase the value of the rents accruing to them from the interest-free loan on construction costs. The rents were increased further as the Bureau of Reclamation expanded its administrative role to provision of multipurpose projects.

The total value of the cost and interest subsidies shown in Tables 2-1 and 2-2 can be calculated for specific projects. For example, the repayment contract for the San Angelo project called for "forty successive equal annual installments commencing with the first year following the last year of a development period which is not to exceed ten years following completion of construction."[26] Assuming that the full ten-year development period was allowed and using a 6 percent discount rate, the irrigators on this project received a subsidy of nearly 90 percent of the construction costs allocated to irrigation. With subsidies of this magnitude it is not surprising that extensive efforts were made to increase the size of the rents.

ACREAGE LIMITATIONS—THE DISTRIBUTION OF THE SUBSIDY

Congress set acreage limits and guidelines for land disposal to promote irrigation of arid lands by small-scale family farmers. The members of the Bureau of Reclamation administered these policies, which determined the distribution of rents from federal water to two groups of water users—settlers on public lands and private landowners. In this section the efforts of the irrigators and bureaucrats to appropriate these rents within the constraints imposed by Congress are described in the context of three aspects of reclamation policy: the rules for settling public lands, the restrictions on the sale of excess lands, and exemptions from the acreage limitation.

Before specific types of rent-seeking behavior can be properly analyzed, the precise nature and intent of the congressional constraints must be examined. A recent court ruling (*United States* v. *Tulare Lake Canal Co.*, 9th Circuit Court, 1976) concluded that "the goals of the reclamation laws were to create family sized farms in areas irrigated by federal projects, to break up and redistribute large private land holdings, to have wide distribution of the subsidy involved and to limit speculative gains."[27] To accomplish these ends, section 8 of the Reclamation Act of 1902 limited the land to which irrigation water would be distributed to a maximum of 160 acres in a single ownership. The nature of this limitation differed from previous

26. U.S. Department of Interior, Bureau of Reclamation, *Reclamation Payments* (Washington, D.C.: Government Printing Office, 1965), p. 343.

27. Jones, *Rules for Acreage Limitation*, p. 936.

federal policies insofar as it limited water rights rather than the amount of land that an individual could own. Once obtained, the water right was tied to the land, not to the individual, and limited to the amount that could be put to "beneficial use." Since the projects were expected to provide water primarily for previously unsettled public lands, the law also established basic requirements for land use that had to be satisfied before ownership of the land would be transferred.

These constraints on the size of holdings and their use were similar to the provisions in nineteenth-century public land laws, which had been designed to combat land monopoly and speculation. The Homestead Act of 1862 limited land ownership to 160 acres and required continuous residency on the land for a five-year period. The Desert Lands Act of 1877, an attempt at stimulating settlement of arid lands, allowed settlers to obtain title for up to 640 acres of land each if they irrigated the land within three years after filing. This act was later amended to reduce the maximum acreage to 320 acres and to establish a more stringent set of requirements to ensure the sincerity of settlers on public lands.[28] The Carey Act of 1894 ceded up to 1,000,000 acres of federal lands to any state where those lands were settled, irrigated, and at least partly cultivated. Ownership of these lands was restricted to 160 acres per person.

These rules proved ineffective in promoting irrigation and preventing large landholdings. Even though over 500,000 acres of land were entered annually between 1877 and 1884, only a minute portion were actually irrigated and patented. Ranchers devised a variety of methods, some of them fraudulent, for skirting the acreage limitations of this act and secured large tracts for grazing.[29] Despite the failure of these earlier restrictions to control land monopoly and speculation, the authors of the Reclamation Act expressed confidence that the stipulations they devised would successfully limit land monopolies. Frank Mondell of Arizona stated:

> It is a step in advance of any legislation we have ever had in guarding against the possibility of speculative land holdings and in providing for small farms and homes on the public lands, while it will also compel the division into small holdings of any large areas . . . in private ownership which may be irrigated under its provisions.[30]

28. For a description of the changes contained in the acts of 1890 and 1891, see Benjamin Hibbard, *History of the Public Land Policies* (New York: Macmillan Co., 1974), p. 431.

29. Several of these methods are described in Hibbard, *History of Land Policies*, pp. 428–434, and Clark, *Water Rights*, p. 15. These sources support the view that the acts of 1877 and 1894 failed in promoting the goals of Congress. Fraudulent activities used to obtain federal timber lands in the Northwest are described in Gary Libecap and Ronald Johnson, "Property Rights, Nineteenth-Century Federal Timber Policy, and the Conservation Movement," *Journal of Economic History* 39 (March 1979): 129–142.

30. Quoted in Paul S. Taylor, "The Excess Land Law: Execution of Public Policy," *Yale Law Journal* 64 (February 1955): 484.

In retrospect this confidence seems to have been unfounded, largely because basic provisions for settlement and establishment of water rights were no more specific that those of earlier acts, and because implementation of the provisions again relied heavily on administrative interpretations by the Department of Interior.

The acreage limitations and rules for securing the property rights to irrigation water led to different forms of rent-seeking activities on lands in public and private ownership. These differences are demonstrated by analyzing public and private ownership separately.

The guidelines established in section 3 of the 1902 act for settlement of public lands authorized the Secretary of Interior to withdraw from entry (except under the provisions of the Homestead Law) any public land he believed to be "susceptible" to irrigation by government works. If and when construction of the project began, the secretary was to announce the construction charges and the maximum irrigable acreage per entry, which was to be that "reasonably required for the support of a family upon the lands in question" and was not to exceed 160 acres or be less than 40 acres. All public lands were to be settled under the Homestead Law on a first come, first served basis. Settlers filing entry between the time the secretary withdrew the land and the time he decided whether a project would be undertaken were required to establish residency within six months of filing, regardless of whether water was available. Before settlers could receive title to the land, they were required to live on (or in the neighborhood of) the land for five years, to reclaim at least half of the irrigable acreage on their claims for agricultural purposes, and to pay off the construction costs levied on their tracts. Claims to federally irrigated lands under the commutation clause of the Homestead Laws, under which settlers obtained patent to federal lands by making a cash payment before the residency period expired, were disallowed so that large land empires could not be amassed.

These settlement guidelines were important determinants of the distribution of and competition for the subsidy benefits among settlers. The acreage limits restricted the size of the entries and therefore the rents available to each settler, while the residency requirements and the first come, first served rule determined the form of the competition for those benefits.

This set of rules led to the dissipation of the rents from public lands at the margin through early settlement.[31] The present value of expected net returns on unhomesteaded public lands prior to irrigation was probably negative. In most cases the expected value of the lands increased dramatically when irrigation water was provided. If settlers had been able to obtain land when water was first provided they would have earned rents on their entries. But under the act of 1902, initial rights to the land went to the settlers who first filed their claims. To keep those rights, they had to establish residence on the land within six months of filing, remain there, and cultivate it for the next five years. They competed for the irrigated lands by settling prior to the completion of the projects. Between settlement and completion of the project, settlers earned negative returns on their lands, which were not profitable to homestead without subsidized irrigation works. When faced with competition, people settled earlier and earlier as long as the present value of the negative costs associated with settling before the project was completed was less than the positive present value of the irrigated homestead.

Early settlement on federal projects was widespread. In 1911 the Reclamation Service reported that considerable numbers of settlers were homesteading "adjacent to every area on which surveys had been made by the Government," whether a project was actually being built or not.[32] Evidence of early settlement on specific projects is found in the following statements from annual reports of the Reclamation Service:

> The Public land under this project (the Payette-Boise Project in Idaho) has been filed on rapidly since its withdrawal on March 5, 1903, for reclamation purposes. At the present time practically every tract that can be irrigated under the project has been entered, even though it is well known that in some parts of the project it will be several years before water can be delivered.[33]

> The lands (on the Minidoka Project in Idaho) were rapidly settled when it became known that the Reclamation Service had undertaken the construction of the project and most of the irrigable areas had been entered under the homestead act before the farm units had been determined and considerably in advance of the delivery of water.[34]

> Many homestead entries were made (on the Lower Yellowstone Project in Montana and North Dakota) at about the time of the withdrawal of lands for the irrigation project on August 23, 1903.[35]

Construction on the Yellowstone project was not authorized until 10 May 1904, and project water was not actually delivered until 1909. On this project settlers homesteaded the lands not only before construction was completed, but before it was even decided whether the project would actually be constructed.

In the discussions prior to the 1902 act, early settlement had been anticipated and "an attempt was made at that time to exclude settlement until the works were built. This was opposed on the grounds that no intelligent man would think of attempting to make settlement in a desert until the water was actually in sight."[36] After discovering that intelligence was less widespread than previously believed, that many settlers were starving and that projects were failing as settlers used up their savings prior to delivery of water, the law was amended in 1910. Settlement of all lands withdrawn by the Secretary of Interior was prohibited under section 5 of the 1910 act until the secretary announced the unit of acreage established, and fixed the water charges and the date of delivery of the water. Section 10 of the Reclamation Extention Act of 1914 prohibited entry prior to the time water was actually ready to be delivered. Dissipation though early settlement appears to have been effectively stopped but the bureau complained of the poor quality of some settlers who obtained land under the homestead rules. The Fact Finders Act of 1924 (subsection 4) gave the bureau more direct control of the distribu-

31. Intramarginal settlers received rents from federally irrigated public lands even in the presence of competition.

32. U.S. Department of Interior, *Eleventh Annual Report of the Reclamation Service* (Washington, D.C.: Government Printing Office, 1912), pp. 9–10.

33. U.S. Department of Interior, *Seventh Annual Report of the Reclamation Service* (Washington, D.C.: Government Printing Office, 1908), pp. 89–90.

34. U.S. Department of Interior, *Ninth Annual Report of the Reclamation Service* (Washington, D.C.: Government Printing Office, 1910), p. 105.

35. Ibid., p. 174.

36. U.S. Department of Interior, *Eleventh Annual Report of the Reclamation Service* (Washington, D.C.: Government Printing Office, 1912), p. 10.

tion of lands when minimum requirements for capital and experience were added to the criteria for entrymen.

It was initially thought that most of the land reclaimed by projects constructed under the act of 1902 would be federal lands and that most of the water from these projects would be delivered to settlers on these lands. However, the Reclamation Service quickly discovered that much of the land on the most promising project sites was already privately owned. As a result it soon became apparent that a large portion of the federally provided water would serve private lands. The distribution of the subsidy benefits on these lands was determined by the acreage limitation and the policies governing the sale of excess lands. The reluctance of landowners to voluntarily sell their excess lands indicated that incentives to dispose of their excess lands were not adequate. To understand why these and other policies failed and why the policies were altered as they were, one must look more closely at the incentives faced by large landowners and members of the Bureau of Reclamation.

Insight into the incentives of landowners can be gained from considering the following expression. Let $V_{\hat{d}}$ be the average dry land (or preproject) value to the landowner of an acre of land, $V_{\hat{w}}$ the average value to the landowner of an acre of land receiving federally subsidized irrigation water, $P_{\hat{d}}$ the dry-land market price, and $P_{\hat{w}}$ the market price of federally irrigated land. Assume that we are looking at the options of a landowner who owns 1000 acres of land and who is allowed to sell excess land at its full market value after the federal irrigation project is completed. The landowner will be able to receive water for 160 acres of land. If the irrigation subsidy has positive value to this individual, the value placed on an acre of land will increase from $V_{\hat{d}}$ to $V_{\hat{w}}$. By selling the other 840 acres of land the individual will receive the "with-water" market price for them, and the total gains from the project will be:

$$G_{sell} = (V_{\hat{w}} - V_{\hat{d}}) \times 160 + (P_{\hat{w}} - V_{\hat{d}}) \times 840$$

Several useful observations can be made from this simple expression. First, if the subsidy has positive value to this landowner (and if the irrigated value of the land is greater than the market price of the irrigated land, that is, if $V_{\hat{w}} > P_{\hat{w}}$), the owner will have the incentive to devise methods for avoiding the acreage limitation provisions. Such methods might include pressing for special legislation to exempt the project from the acreage limitation and encouraging the Bureau of Reclamation to adopt policies whose effect is to relax these restrictions (such as allowing a married couple to own 320 acres). Second, if there are economies of scale to irrigation farming for lots larger than 160 acres, this landowner will have additional incentives to avoid the acreage limitation, since $V_{\hat{w}}$ will be larger the more land is irrigated. Third, if $P_{\hat{w}} < V_{\hat{d}}$, which might occur if this individual is extremely efficient at farming without project water, it will not be in the owner's interest to sell the excess lands, even at the irrigated market value.[37] Finally, this landowner has incentives to delay selling excess land if there is a possibility that $P_{\hat{w}}$ will rise in the future, provided the expected rate of return from holding the land is greater than the return from selling now and reinvesting the proceeds. There are indications that landowners did engage in this type of speculation.[38]

These incentives change when we assume that the landowner is required to sell excess lands at a price that does *not* include the value of the subsidy ($P_{\hat{d}}$). Consider again the landowner who owns 1000 acres of land prior to the reclamation project. According to the law, this individual is eligible to receive federal project water for a maximum of 160 acres, leaving 840 acres for which water is not received. The fact that the owner held excess land before the project was begun indicated the value of that land ($V_{\hat{d}}$) was at least as great as its market value ($P_{\hat{d}}$). Since the owner is restricted to receiving $P_{\hat{d}}$ in a sale of excess acreage, there would be losses from the sale equal to $(V_{\hat{d}} - P_{\hat{d}})$ 840. The owner has no incentive to sell the lands at the preproject market price. The landowner will not be able to get subsidized water on excess lands and will not be willing to sell them to potential buyers who would be able to receive the water. If the owner refuses to sell, nobody receives the subsidy on the excess lands. In this situation both large landowners and potential buyers have incentives to break the law so they can capture the subsidy.

37. One way in which the landowner can farm efficiently without project water is by pumping groundwater for his lands. Opponents of acreage limitations in California's Central Valley argued that large landowners would gain from the project by pumping groundwater replenished by water from irrigated lands. It appears that their arguments were unfounded, as most large landowners chose to join the project. See Warne, *Bureau of Reclamation*, pp. 76–83.

38. For references to these types of activities, see Montgomery and Clawson, *History of Central Valley Project*, p. 138; "Reclamation by Irrigation," pp. 113–114; and Clark, *Water Rights*, p. 211.

Numerous bureau policies allow the subsidy to be obtained and shared by excess-land holders and buyers. One is to permit large landholders to sign a contract that allows them to receive water on their excess lands for a limited time period but forces them to sell those lands at the preproject value (P_d^*) at the end of the period. If the period is long enough, the landowners have incentives to sign the contract and sell their lands at the specified time. The buyers of these lands will pay the preproject value of the lands and will receive subsidized federal water on the project when they obtain the land. Under this policy it is in the interest of the large landowners to delay the actual sale of their lands as long as they can, while the buyers will push for shortened contract periods.

This policy leads to a division of the irrigation subsidy benefits (to lands that are initially privately owned) between the large landowners and the buyers. The landowners get the full benefits of the subsidy on 160 acres of their land, plus the benefits from receiving water for their excess lands until those lands are sold. Settlers who bought the excess land at its preproject price receive the remainder of the subsidy benefits.[39]

The actual division of the benefits is determined by the political power of the two groups. Large landowners would be expected to have the advantage in political competition since they are a smaller, more concentrated group with established ownership (implying that they face lower organization costs) than the group of potential buyers. Under all of the policies for sales of excess lands, the costs of the subsidy are borne by the taxpayers on pure irrigation projects. On later multipurpose projects, power users also paid for a share of the irrigation subsidy.

The incentives of bureaucratic actors associated with federal reclamation were also extremely important in determining the actual distribution of the subsidy when federal projects served privately owned lands. Because their interests often conflicted with those of Congress, these actors made tradeoffs between legislative and constituent pressures in an effort to capture a portion of the rents from the federal reclamation program.

39. One possible method of reducing high rates of returns on those lands—settlers paying "above-market" prices for improvements and structures owned by the landlord that were sold with the excess land—was excluded in contracts from the Central Valley project in California, which included provisions that nonland assets also be sold at assessed prices. U.S. Congress, House of Representatives, *Central Valley Project Documents*, House Document No. 246 (Washington, D.C.: Government Printing Office, 1957), pp. 84–139.

The members of the Bureau of Reclamation and the Department of Interior were expected to administer the acreage limits and land resale policies and ensure that the subsidy was widely distributed by promoting the development of family farms. It therefore appears to have been in the interest of the bureau to enforce the acreage limitation laws established by Congress, since without congressional approval the bureau would be unable to undertake new projects. The bureau also was pressured *not* to enforce these excess-land laws. The most obvious source of pressure for nonenforcement came from large landholders on reclamation projects who were willing to incur considerable expense in their efforts to find methods of avoiding these acreage limitations.

In addition to these pressures, the bureau had other incentives to adopt policies of nonenforcement. There are at least three reasons why nonenforcement of acreage restrictions might have increased the flow of repayment on a particular project, which was strongly correlated with project success desired by both Congress and landowners. First, if there were economies of scale in irrigation for farms larger than the acreage limitation on a particular project, the prospects for repayment were enhanced by allowing large landholders to receive (and pay for) water for their excess lands. In the presence of scale economies, the net income available for payment of construction costs from one 320-acre farm was greater than that available from two 160-acre farms. Second, the Bureau of Reclamation generally had more complete, reliable information on the abilities of current landholders to repay construction costs than on the abilities of prospective settlers, particularly on projects that supplied supplemental water to irrigation farmers. Third, if excess lands were not allowed to receive water and large landowners chose not to sell these lands, the construction charges for the project would have been divided among fewer acres, increasing the construction charges on participating acreage and reducing the probability of successful repayment.

It is apparent that the bureau had to make tradeoffs between its goals of ensuring the repayment of construction costs and promoting the development of family farms through the enforcement of acreage restrictions. It is also apparent that enforcement was not an all-or-nothing activity. There was a wide range of policy options available between the extremes of absolute enforcement, where no excess land received water under any circumstances, and total nonenforcement, where no efforts whatsoever were made to encourage landowners to

comply with the acreage limitation laws. The actual enforcement policies of the Bureau of Reclamation were determined by a variety of factors, including: (1) the visibility of the decision makers and (2) the intensities of opposing pressures from large landowners on one side, and Congress, the president, and the courts on the other. The intensity of pressure from owners of excess lands in their efforts to increase the rents they received from the federal water subsidy was a function of the net value of the subsidy, the proportion of lands held in excess, and the probability (as perceived by the large landholders) of successfully avoiding the acreage limitations. The pressures applied by the different branches of the government varied over time with the convictions of the individuals in office.

The visibility of the decisionmaker was important in determining the different responses of members of the Bureau of Reclamation and the Department of Interior to political and constituent pressures. Since it was costly for the lawmakers to monitor the actions of the administrators of the program, the agents were more likely to enforce the laws the more visible their position.[40] For example, the Secretary of Interior and commissioner of the bureau, whose actions were easily observed by the Congress and the president, were relatively more responsive to pressures from political and judicial sources. On the other hand, it was difficult for the president to observe the actions of local bureau representatives, who therefore were relatively less responsive to pressure from Washington to enforce acreage limitations, especially when large landowners demanded nonenforcement. enforcement.

The preceding discussion lays the groundwork for the analysis of actions taken by large landowners and the Bureau of Reclamation in response to the legislation concerning the disposal of excess lands and enforcement of the acreage limitation. The authors of the Reclamation of 1902 clearly wanted to break up land monopolies, but their attitudes toward the distribution of the subsidy from federal water were unclear. In the initial act there was no limitation placed on the sale price of excess lands. Senator Francis G. Newlands explained the relationship of the acreage limitation of 160 acres to the large landowner's incentives:

40. The lawmaker's problems with monitoring bureaucratic agents is a central theme in Douglass North's "A Framework for Analyzing the State in Economic History," in *Explorations in Economic History* 16: 249-259, and *Structure and Change in Economic History* (New York: W. W. Norton and Co., 1981), pp. 20-32.

> The fact that waters from the irrigation project can be brought within the reach of [a] large holding raises its value. Purchasers of that holding in tracts of 160 acres can secure water rights under this Act. The large landed proprietor is benefited by having the water brought within reach. He has the opportunity of making sales of lands hitherto unsalable and a purchaser can unite the water with the land by buying a water right from the government and thus dedicate the land to future productiveness and so that act not only guarantees against monopoly in the state but will gradually destroy existing monopoly and disintegrate these holdings in the country without injury to any and benefit to all.[41]

Since no limitation was placed on the price at which the excess lands were to be sold, the expected future net gains from the receipt of federally subsidized irrigation water would be capitalized into the sale price of that land. In other words, all of the windfall benefits from the project would be received by the large landowners, assuming estimates of the value of future subsidies were on average correct.[42] Apparently, the original reason for breaking up large landholdings was *not* to ensure that the project benefits were widely distributed among the settlers on the resulting family-sized farms, but simply to break up the large holdings.

By 1914 it was apparent that large landholders would not willingly sell their excess holdings under the rules established by the act of 1902. Large landowners were holding their excess lands out of production while they waited for the prices to rise, which delayed the flow of repayment revenues into the Reclamation Fund. To remedy this, the Reclamation Extension Act of 1914 required the owners of excess lands to agree to dispose of those lands at prices specified by the Secretary of Interior before any construction work was done, or they would not receive project water. As written, the act gave no indication as to whether the price established by the secretary was to include the capitalized value of the federal subsidy. This omission increased the discretionary control of the Department of Interior

41. Quoted in Montgomery and Clawson, *History of Central Valley Project*, p. 134.

42. To the extent that federal estimates of construction costs were used to calculate these subsidy values, the benefits received by a large landowner when the land was sold would be greater than the true windfall gains from the project, since federal estimates of construction costs were consistently lower than actual costs.

over the distribution of the subsidy to buyers and sellers of excess lands.

Ten years later the Fact Finders Report of 1924 concluded that these antimonopoly measures were ineffective. Its authors objected to the distribution of the subsidy established by the bureaucrats administering the law. Some of the prices established for the sale of excess lands included the value of project water, thereby giving all of the subsidy from federal water to the large landholder.[43] In other instances, part of the subsidy was going to so-called speculators, since there were no limits on the activities of middlemen who purchased the land at the secretary's price and resold it at the market price. The authors also objected that the Reclamation Service was allowing large landholders to receive water on their excess lands, which was contrary to the clearly stated intent of Congress. The landowners would agree to sell their lands but the actual sale was often delayed indefinitely. In the interim the service interpreted the law to allow delivery of water to the excess lands.

To correct these shortcomings, Congress passed the Omnibus Adjustment Act of 1926. Section 46 of the act required the following to be included in all repayment contracts between the secretary and organizations: (1) all irrigated land held by a single owner in excess of 160 acres was to be appraised in a manner prescribed by the secretary; (2) the appraised price would reflect the value of the land without the value added by the reclamation project; (3) no excess land was to receive water from a federal project if the owner refused to execute a recordable contract agreeing to sell those lands at a price not to exceed the appraised price; (4) until half the construction charges were paid, the purchase price of any sale of excess land had to be approved by the secretary before that land would be eligible to receive project water; and (5) if it was found that the purchase price of excess lands had been falsely reported, the secretary could cancel the water rights attached to that land.[44]

The apparent purpose of these provisions was to ensure that excess lands would be sold and that the selling price would be limited, thereby achieving a wider distribution of the subsidy. Although the members of the bureau and the Department of Interior were constrained to setting dry land prices, the act contained loopholes that

43. "Reclamation by Irrigation," pp. 113–114.
44. Clark, *Water Rights*, p. 212.

still allowed them to control the subsidy's distribution. As in the 1914 act, landowners were only required to *agree* to sell their excess lands; no requirement that these lands actually be sold within a given period of time was imposed. The secretary established the distribution of the rents by requiring landowners to agree to sell their excess lands within ten years and by interpreting the law to allow the delivery of water to excess lands during that period.[45] Another dimension of water contracts that was not specified in Section 46 was the penalties for selling lands at "speculative" prices. These penalties were determined by negotiations between the bureau and the irrigation districts. The contract with the Kittitas Division of the Yakima project in Washington provided that half of the difference between the sale price and the appraised price go to the irrigation district and be credited against the construction costs of the project; the other half to be retained by the seller of the excess lands.[46] Under this type of an arrangement, the incentives of the landowner-speculator would still be to sell excess land at the market price. If the objectives of the irrigation district and the Bureau of Reclamation include early repayment of construction costs, it may also be to their advantage if excess land is sold at a price above the appraised price.

Several acts passed after 1926 contained provisions relating to specific projects or classifications of projects. For example, the Small Reclamation Projects Act of 1956 provided that the repayment plan on qualifying projects would require the payment of interest "on that portion of the loan which is attributable to furnishing irrigation benefits in each particular year to land held in private ownership by Because the Department of Interior has interpreted this passage as a repeal of the excess-land law for small projects, large landowners involved in projects falling under the jurisdiction of this act can receive water for their excess lands by foregoing the interest subsidy and are not required to sigh recordable contracts for the disposal of

45. Ibid., pp. 213–214. It is interesting to note that the loopholes allowing the delivery of water to excess lands under the acts of 1914 and 1926 were the same. If Congress really had wanted to prevent the monopoly of federal irrigation water, it seems reasonable to assume that it would have required the actual sale of the land.
46. Warne, *Bureau of Reclamation*, p. 73. Under the Columbia Basin Anti-Speculation Act of 1937, the maximum that could be kept by the participants in the sale—50 percent of the price differential if the penalty was paid immediately—was to fall by 1 percent for every month that the penalty payment was delinquent. If the penalty payment was not made within fifty months, the government was to receive all of the price differential.

445

those lands.[47] Efforts to obtain this exemption for all reclamation projects have failed. The Omnibus Adjustment Act was the last act containing general provisions for acreage limitations and sales of excess lands.

Most of the controversy over reclamation policy has stemmed from the bureau's enforcement or lack of enforcement of the acreage limit. The enforcement policies adopted by the bureau on different projects have different implications for (1) the nature of the distribution of rents among irrigators and (2) the net benefits accruing to different members of the bureau. The key factors determining these differences in policies have been the intensity of pressure for nonenforcement from constituent landowners, vacillations in congressional pressure to enforce the law, and the visibility of the decisionmakers' actions to the legislative and executive branches.

From the bureau's perspective, the most desirable method of avoiding enforcement of acreage limitations was legislative exemptions, which gave congressional sanction to nonenforcement policies. Large landowners on such projects were appeased by this guaranteed exemption, and the probability of successful repayment from these established farmers was increased. Administrative exemptions by the Secretary of Interior or the head of the Reclamation Bureau for specific projects were less desirable to both the bureau and large landowners. These exemptions were easily monitored by Congress and the president and could have been politically costly to the bureau if these groups desired enforcement. On the other hand, administrative exemptions enhanced the likelihood of successful repayment and relieved bureau representatives at the local level of pressures from large landowners. From the landowners' viewpoint this type of exemption was less attractive than legislative exemptions, since it was less permanent and its legality less certain. The probability of a new Secretary of Interior changing the policy was greater than the probability of Congress retracting a legislated exemption. Exemption through nonenforcement at the local level, the least visible form, was the least costly to the bureau in the face of congressional pressure to enforce the law. Landowners were even less enamored with this type of an exemption, since it was illegal and less permanent than the others. Each of these methods of avoiding acreage limitations

47. Clark, *Water Rights*, p. 289.

has been employed during the history of the federal reclamation program.

One of the most important general relaxations of the acreage limits was the policy of allowing a husband and wife to receive federal water for 320 acres of land. Administrative rulings supporting this policy were made as early as 1904 and have since been applied to virtually all federal reclamation projects.[48] Nearly 16 percent of all acres irrigated on projects built prior to 1944 were freed from the acreage limitation by this exemption. In the 1946 landownership survey 704,410 acres in 3,187 land holdings were freed from the acreage limit by the husband-wife exemption.[49] This policy increased the likelihood of successful repayment in areas where economies of scale existed and where the land was originally in private ownership. This policy, though visible and therefore potentially costly in political terms, seems to have avoided widespread controversy.

Political pressure to enforce the acreage limits appears to have been relatively low in the 1930s. A 1933 statutory interpretation by Interior Secretary Ray Wilbur exempting landowners in the Imperial Valley of California from acreage limits, was overturned in *United States* v. *Imperial Irrigation District*, but was finally upheld by the Supreme Court in 1980 in *Bryant* v. *Yellen*.[50] Legislative exemption granted to the Colorado-Big Thompson project in 1938 had the support of the bureau and was passed by Congress with no opposition. The justification for exempting both of these projects was that federal water was only supplemental to privately supplied irrigation water. The Truckee River and Humboldt projects in Nevada received this rare congressional exemption because the climate and elevation in these regions made the land so unproductive that 160 acres would not support a family.[51] Under these conditions, it seemed unreasonable to disrupt the regional economies by breaking up the large landholdings.

This attitude toward enforcement contrasts sharply with that adopted during the negotiations with California's Central Valley

48. One exception to this was enacted in the Columbia Basin Project Act of 1943, which expressly stated that a family could only receive water sufficient for one unit of land, the size of which was to be established by the Secretary of the Interior.

49. Bureau of Reclamation, *Landownership Survey on Federal Reclamation Projects* (Washington, D.C.: Government Printing Office, 1946), pp. 16–17.

50. For a discussion of the 1977 case see Jones, *Rules for Acreage Limitation*, p. 936.

51. See Montgomery and Clawson, *History of Central Valley Project*, p. 144.

water users, which began during the late 1930s. Local bureau representatives on the Kings River project (and apparently on other projects in the Central Valley) initially assured irrigators that their lands would be exempted from acreage limitations. This policy was supported by Secretary of Interior Ickes and seemed consistent with policies of the previous decade, since most of the land in the Central Valley was already receiving water from private projects.[52] In 1943, after a policy review with the Department of Interior, the bureau was reorganized. Soon afterward Ickes issued a statement, apparently at the request of President Roosevelt, ruling that Central Valley projects would not be exempted from the excess-land laws.

Surprised by this reversal of policy, the landholders in the Central Valley regrouped and began to pressure Congress for a legislative exemption of their lands. They managed to push a rider to the Rivers and Harbors Bill through the House of Representatives in 1944, but it was defeated in the Senate because of strong opposition from Roosevelt. The landowners also attempted to have their projects placed under the authority of the Army Corps of Engineers to be built under flood-control law, but they were stymied in 1944 when Congress made corps projects subject to the excess-land provisions of reclamation law.

After a bill to exempt projects in California, Colorado, and Texas died in committee in 1947, the landowners increased the pressure on members of the bureau to exempt their lands. In 1948 a rider was successfully attached to a bill passed by Congress that effectively removed two influential supporters of enforcement from the payroll of the Bureau of Reclamation.[53] Despite this pressure, water rights in the Central Valley were eventually administered under the acreage limits and large landowners were forced to sign recordable contracts to sell their excess lands before they could receive water.[54]

During the debates over the Central Valley project, the landowners on the Kings River project negotiated with local bureau representatives over the inclusion of a provision allowing for exemption from acreage limits by prepayment of construction costs. Prior to 1946, less than 1 percent of project lands were exempted under this

52. See Arthur Maass and Raymond Anderson, . . . *and the Desert Shall Rejoice* (Cambridge, Mass.: MIT Press, 1978), pp. 264–265.

53. For descriptions of this incident, see Taylor, *Excess Land Law*, pp. 504–505. President Truman later restored these men to the payroll.

54. See Warne, *Bureau of Reclamation*, pp. 80–81.

clause.[55] Throughout the 1950s contracts containing various forms of this provision were negotiated and approved at the local level, but several Interior secretaries hesitated to commit themselves on the question of the legality of the exemption.[56] Eventually, the issue was left up to the courts, where the provision was ruled illegal in 1977.[57]

One less visible type of nonenforcement occurred on the Kings River project during the negotiations over prepayment of construction charges. Throughout the decade of debate over the legality of the prepayment extension, water from Pine Flat Dam was delivered to excess lands and none of the water users on the project were required to comply with the acreage limitation.[58]

By nature, evidence of the extent of nonenforcement is difficult to obtain. However, a landownership survey published by the Bureau of Reclamation in 1946 indicated that such practices had important effects on several projects at that time. On each of the following projects more than 10 percent of the total irrigable acreage was known to be in violation of acreage limitations: Klamath (18.6 percent), Salt River (12.5 percent), Yuma (14.3 percent), Carlsbad (12.9 percent), Rio Grande (15.1 percent) and North Platte (10.3 percent).[59] On the other hand, the survey also indicated that on 28 out of 52 projects, less than 1 percent of the lands was known to be held in violation of the excess-land laws. This result may imply that some local bureau officials were more effective enforcers of the acreage limit or that the optimal farm size was less than the acreage limit on the projects where the law appears to have been enforced. In the latter case there would be less pressure from landowners to not enforce the acreage limit.

55. Estimated from *Landownership Survey*, pp. 16–17.

56. See Maass and Anderson, *Desert Shall Rejoice*, pp. 267–269.

57. For a detailed history of the legislation and administrative decisions concerning prepayment see Taylor, *Excess Land Law*, pp. 490–512. A description of the 1977 court case can be found in Jones, *Rules for Acreage Limitation*, p. 935.

58. See Maass and Anderson, *Desert Shall Rejoice*, p. 271.

59. See *Landownership Survey*, pp. 16–17. These excess lands did not include lands for which recordable contracts had been signed nor for which construction costs had been repaid.

SUMMARY AND CONCLUSION

In this paper we have analyzed the actions of irrigators and bureaucrats who competed to obtain rents from federal water projects within constraints imposed by Congress. In the Reclamation Act of 1902, potential water users successfully obtained federal subsidies in the form of interest-free loans for the construction costs of irrigation works. Congress established rules to control the distribution of these subsidies, the value of which increased over the next forty years as water users successfully obtained more flexibility and longer extensions in their repayment contracts. Often their efforts were consistent with the interest of members of the Bureau of Reclamation, who were seeking to increase the congressional demand for their output by providing an image of successful settlement and repayment on their projects.

The distribution of rents among landowners was determined by acreage limitations and rules for the disposal of public and private lands. According to the 1902 act settlers could obtain the rent on public lands by settling those lands under the Homestead Act on a first come, first served basis. These rents were dissipated by the settlement of project lands prior to the delivery of irrigation water. To reduce dissipation and increase the financial success of projects, Congress passed legislation to restrict early settlement and eventually gave the Bureau of Reclamation increased discretionary power over who settled the public lands.

Under the 1902 act Congress tried to ensure the dissolution of large landholdings by allowing private landowners to sell their excess lands at market prices. It was soon discovered that rather than selling, landowners were holding land until the prices rose further. To remedy this, Congress gave the Secretary of Interior discretionary power to set the selling price of excess lands. In 1926 Congress tried to restrict the secretary's discretionary power and limit the subsidy to large landowners, but the bureau found new ways to control the subsidy's distribution.

Large landowners continually applied pressure for exemptions to the acreage limitation law, while congressional pressure for enforcement of the law varied over time. Enforcement of the acreage limit was expected by Congress, but success on projects (which was desired by Congress, the water users, and the bureau itself) would have been enhanced by nonenforcement. The more visible decisions made by the heads of the Department of Interior and the bureau to exempt large landowners from restrictions were made when congressional pressures to enforce the law were low. When enforcement pressure increased, the bureau toughened its visible policy stance but pursued less visible exemptions of large landholders through nonenforcement of the acreage limit at the local level. We do not know the timing of this nonenforcement, but we do know that it existed.

This description of the first five decades of the federal reclamation program has demonstrated how efforts were made to influence the magnitude and distribution of the subsidies from this particular government program. The nature of these efforts, as well as their effects, are not unique to the reclamation program. Whenever rents are created by government programs and distributed on the basis of nonmarket criteria, competition will lead to dissipation of those rents. As in the reclamation program, attempts to reduce this dissipation by controlling competition along certain margins will generally lead individuals to direct their energies along other, uncontrolled margins in an effort to appropriate available rents.

BENEFIT-COST ANALYSIS: ECONOMICS VS POLITICS?

by A. Allan Schmid

August 1989

Staff Paper No. 89-81
Michigan State Univ.
Dept. Agr. Econ.

Benefit-Cost Analysis: Economics vs Politics?

by A. Allan Schmid

Benefit-cost analysis (BCA)and economics has long been touted as a rational guide to political choice (size and content of public spending). The economist promised to make an independent assessment without any instruction from the political authority who was assumed to be interested in maximizing economic product. There have been two troublesome attacks against this position. One was an objection on distributive grounds. The hard line defense was to say that distribution was a political matter, but for optimal results, government should redistribute income on a lump-sum basis and not mess around with provision of goods and services. If this line could not be held, some economists were willing to have the benefits to some designated parties weighted.

The second attack is relatively new and comes from those who prize goods and services which do not usually have market prices. These products (both project inputs and outputs) did not get accounted for. This seemed to provide an opening for political pricing, but the gap was putatively closed in the last decade by travel cost methods and contingent valuation surveys. The authoritative faith seemed to be renewed by the promise that economists could cleverly find new ways in which people inadvertently revealed their preferences or these could be measured and aggregated directly and without the bias that politicians introduced when listening to voters.

One telling attack upon this citadel of independent authoritarian analysis came with the publication of Sugden and Williams (1978). They argued for a "decision-maker's approach" which required some explicit input from the politicians. The argument was strongest with respect to the discount rate which Sugden and Williams said could only be a matter of political choice and not a matter of data to be discovered by the analyst (see also DeAlessi, 1969). The decision-making approach was sharply critiqued by E. J. Mishan (1982) who in a review said that much would be lost if the economist could no longer be regarded as independent and authoritative. This critique was made in spite of the fact that earlier Mishan (1981, 163) had put some of his own nails in the coffin by suggesting that for a project to go forward it must meet both tests of willingness to pay and willingness to sell. This makes ranking ambiguous without political input.

An earlier attack on the citadel came from Little and Mirrlees (1974). They argued that the analyst could not independently substitute shadow prices for nominal prices affected by taxes, tariffs and exchange controls without asking for the intent of politicians. If the intent of these policies was corrective of income distribution, then it would be value presumptive for the analyst to replace nominal prices. Earlier Little (1957) made a devastating critique of the use of consumer surplus in making welfare comparisons. The theoretical tool which Mishan (1976)

had called the distinguishing feature of policy oriented BCA, Little called a context for political decision.

There has been a revolution in the theoretical welfare literature in the last 20 years which has not yet been fully felt in applied BCA. One its major ideas is the theory of second best. It says that if the prior income distribution is not politically acceptable, no welfare implications can be drawn from present prices. If costless lump-sum redistributions are not available, then redistribution via projects can not be ruled out as inefficient. The same holds for the prices produced by imperfect capital and labor markets in disequilibrium. The gap between theory and application is nowhere better illustrated than in the work of Boadway and Bruce (1984). They demonstrate the limited applicability of first-best theory in a second-best world with many households with diverse preferences. Yet, in their last chapter devoted to BCA they limit their analysis to "projects which have no perceptible effects on the market prices for goods and factors of production in the economy" and assume that "the economy can be treated as if all persons are identical so that no distributive weights are needed" (p.292). In other words, they apply BCA to a make-believe first-best world.

The basis for persistent disequilibria in capital and labor markets is another closely related theoretical development. Stiglitz (1987) argues that information costs imply that markets will be in disequilibrium even if pure competition and no institutional causes are present. This suggests that full utilization of resources is not simply a matter of enforcing competitive markets. Government projects and regulations are part of the everyday management of the economy and not just to fill the breach of an occasional externality or business cycle slump. For example, disequilibrium mean that everyone will not have the same marginal rate of time preference, thus necessitating some political resolution of the conflicting preferences.

Majone (1989, 15) has labeled authoritarian policy analysis as "decisionism". He says it assumes a unitary decisionmaker and is not applicable when there are two or more actors with different objectives. Decisionism assumes all conflicts have been settled and choice of projects or regulations is a *technical puzzle* to be solved rather than a *political judgment* to be made.

A nascent theory of behavioral economics also has implications for the role of technician and politician. The behavioral sciences remind us that people and their perceptions differ. The analyst is just another observer with her own cognitions and it can't be assumed that conflicts in perception can be solved outside of the political process.

It is now possible for an applied BCA to be formulated which is consistent with second-best welfare theory and a real second-best world. To that end it will be necessary to extend Sugden and Williams decision making approach to a fuller political economy approach where political input has a place appropriate for a democratic society and where no self appointed analyst elite usurps representative government unwittingly or otherwise.

The outline for such an approach can be sketched by considering in turn the major steps in BCA including establishment of a nominal accounting framework, estimation of the production function, direct benefit estimation, evaluations of non-

marginal projects, opportunity cost adjustments for imperfect labor markets, time preference in the context of imperfect capital markets, and preferences for uncertainty adjustments. At each step the iterative interaction between political authorities and analyst can be indicated.

Program Information Structure

The first place where political input is needed for analysis is in choice of the nominal input and output categories. Nothing can be priced unless the physical quantity of a set of characteristics is understood. This choice involves the level of detail and aggregation of product characteristics. When is one product different enough to be given a separate name? One can't do even elementary cost-effectiveness analysis without agreement on whether the cost data is comparing products of equivalent quality. Since people differ on how similar is close enough, some political input is needed. Usually the content of product qualities is chosen by reference to statements of objectives in authorizing legislation. Some dialogue is needed between analyst and politician to establish construct validity. These are matters of judgment for a craftsman and not simply matters for a logician (Majone 1989, 47).

The issue here is analogous to that in industrial organization where one sign of non-competitive behavior is undue product differentiation. But, what is undue for one person is not for another and thus legislative and judicial guidance is needed antecedent to cost comparisons. This is the same issue as that involved in choice of program budget categories which facilitate and hinder comparisons between government agencies.

Estimating the Production Function

There is much technical expertise in experimental design to establish whether the project caused a change in output. But, better designs cost more and some judgment is needed on whether the reduction in threats to internal validity are worth the cost. Ultimately, definitive randomly assigned treatments (projects) are rare, which necessitates some qualitative judgment on the weight of the evidence behind alternative projects. These judgments are one that reasonable people may differ and settling differences is what politics is about. This judgment will be discussed further below in the context of uncertainty.

Pricing Benefits and Costs

Analysts are called on to supply prices when no market references are available. This requires inferences from indirect evidence of willingness to pay which is the essence of the economists technical expertise. But, even here there is a need for political input. A selection of cases will make the point. One of simplest methods is to reason from the price of an analogous good to that of the non-marketed project

good. This returns to the first topic above, namely to establish that the goods are perceived as comparable.

The human capital approach is commonly used in valuing life in the context of safety projects and regulation. But, to use the opportunity cost of life-time earnings is to make a political choice of property rights. It includes a decision to put the potentially harmed person in the position of a buyer of safety rather than a seller entitled to be free of harm. This returns us to the first-best vs. second best question. If income distribution had all been settled or one could obtain the desired distribution outside of projects and regulations via costless transfers, then the human capital approach would be unambiguously Pareto-better. The same point can be made with respect to environmental products.

Differences in human perception create the need for political resolution. For example, the cost saving method commonly used in transportation project evaluation requires a choice between the analysts perception of time saved and that of the actual users of the transportation. The same problem arises in the context of exposure to hazardous events. Sugden and Williams (1978, 179) give the label of "merit goods" to products that people would want if they understood their best interests. Whether this is desirable caring or paternalism requires political judgment.

The use of contingent valuation (bidding games) requires the resolution of political questions. The process can't begin unless it is decided whether willingness to pay or sell is decided. As noted in another context above, this is a basic property rights question which is antecedent to market exchange or any simulation thereof. The framing of the questions, the anchor point, and the degree to which opportunity cost tradeoff is made explicit all are known to affect the resulting prices. Politicians seek membership on the rules committee because control of the agenda affects outcomes. Economists at least since Kenneth Arrow also understand that grouping of issues (order of vote) affects the formulation of winning coalitions. Yet, this is ignored when analysts go off by themselves and make an independent contingent valuation study without political input.

The parallels between surveys (contingent valuation) and a politician sampling and acting upon constituents' preferences is striking. Both processes are subject to the same issues of sampling, framing, and aggregation. While both involve issues of property rights, some economists are willing to rewrite the constitution with a cloak which promises to measure the true value of voter sovereignty. This is not simply a problem of principal-agent because the issue is the aggregation of the preferences of multiple principals and the dynamic learning environment of the principals whose preferences are evolving. An extended argument is not possible here, but there is nothing inherently superior about market prices or prices inferred from indirect evidence of willingness to pay or surveys vs. administrative prices (Schmid, 1989). If the legislature can change property rights and generate alternative prices in the market, it can surely choose those prices directly under one constitution or set of political rules (or surveys) or another.

Non-marginal Projects

Some projects and regulation are large enough to cause a change in prices of the output and inputs (Hoehn and Randall 1989). It has become very popular among applied economists particularly in recreational and environmental projects to estimate consumer surplus as a measure of the willingness to pay for a price decrease. Some theorists on the other hand have been hesitant to endorse it. The problem is again that of the first-best assumption of optimal income distribution in the many-consumer economy (Tresch 1981, 198).[1] If costless lump-sum transfers are not available, Boadway and Bruce (1984, 271) conclude that, "The use of the unweighted sum of household compensating or equivalent variations as a necessary and sufficient indicator of potential Pareto improvement is rife with difficulties." The theory has been worked out mostly in the context of taxes but the implications are the same. Tresch (p. 351) says "it may not be very useful to think of the effects of distorting taxes in terms of deadweight loss. Unambiguous notions of efficiency loss involve the use of the expenditure function, which is best suited to one-consumer economies."

The use of consumer surplus is equivalent to a firm being a perfectly discriminating monopolist. The amount of consumer surplus that such a firm can extract depends on whether other firms are also trying to do it. Not all firms whether private or public can simultaneously extract the consumer surplus that is estimated for each one acting alone assuming no one else is trying. As Samuelson (1963, 197) once noted, some otherwise bankrupt firms would have survived with price differentiation.

The applied economist can scarcely do better than conclude as did Little (1957) when he said "The best criterion for investment decisions must, within wide limits, be determined at dynamic and administrative levels--and not at the level of static welfare theory" (p. 184). Which firms are to be allowed to act in terms of being a discriminating monopolist is fundamentally a distributive question. Consumers don't know what they are sovereign over until they know whether they will be subject to price differentiation.

Opportunity Cost

Disequilibrium in labor markets is the context in which analysts offer up shadow prices to replace nominal prices for the unemployed. All are agreed that decisions should be made in terms of opportunity costs, but whose? The price of labor on public projects cannot be separated from the objectives of macro policy which are a matter of political conflict. A government which intends to put downward pressure on wages may purposely create unemployment (or take advantage of it when it occurs). The last thing they want is for the project agencies to have larger budgets as a result of higher net returns when wages are computed at some shadow price instead of the higher market price.

The distributive issues can't be settled in a separate transaction. In a second-best world, the government is likely to want to pay the nominal wage. For example, in the U.S. the Davis-Bacon Act requires it. This means that if projects are built

using shadow prices that would not otherwise be built, they are the occasion for some taxpayers to make transfers to labor.

Valuation over Time

Is the discount rate for public projects a matter of data to be observed or a public choice to be decided and delivered to the analyst? With a perfect capital market, everyone would have the same time preference at the margin. People with initially different time preferences would borrow and lend, market rates would adjust until all players are in equilibrium. But in disequilibrium, people have different opportunities and differ over the desirability of financing public projects by borrowing or taxation. Some further political resolution of conflicting interests is necessary even if the distribution of factor ownership (wealth) were acceptable and there are many borrowers and lenders (no market power). The literature, of which Sugden and Williams and DeAlessi are representative, seems to have shifted to viewing the choice of discount rate as a political decision rather than a datum to be discovered.[2]

Uncertainty

People have different preferences for the tradeoff of mean values and their variation. There are few markets for recording these preferences. Each person can't independently adjust a portfolio of public projects to obtain their preferred risk exposure. This means a political judgment is needed.

Uncertainty is an area where behavioral economics has a lot to offer. Much of the decision analytics separates the perception of mean values from perception of their variation, but much empirical evidence exists that the perceptions are interrelated. The frame for viewing attitudes toward uncertain events has a lot to do with what is seen. A political compromise among differing perceptions is needed.

Conclusion

The traditional separation of technical analysis and political choice is no longer tenable. Theory and experience point to a more interactive, iterative relationship between analysts and politicians. The distribution of rights affect prices and thus any existing set of prices can't guide the choice of rights including those embedded in rules for appraising public spending and regulations. It can't be assumed that political choice has once and for all chosen the distribution of rights and that the only problem is the technical one of implementation of the preferences of rights holders. The process of public investment and regulation is never wholly exchange facilitating (solving market failure) nor wholly grant making, and the distinction needs continuous political input. The analyst need not apologize for asking more questions of the politicians. And the technical input is no less useful for the fact that as new politicians are elected, public investment priorities change. The value of

analysis is in clarifying the substance of evolving preferences and the necessity to be explicit about whose preferences count.

In closing it should be noted that citizens and politicians do not have a universal, consistent, strong preference for explicitness and clarification of the sources of winners and losers. The failure of applied BCA to reflect the evolution of second-best theory is only partly due to reluctance of economists to relinquish the role of supplying authoritative advice.[3] The other part is a substantial public demand for self (and others) deception and vain glory where we advertise a concern for the poor, human life, environment, or whatever, while acting selectively to the contrary. When politicians stand aside from resolving the conflicts of interest behind the BCA rules, they are able to embrace its results piecemeal--accepting its added legitimacy when it suits them but labeling it academic irrelevance when they reject the results. If politicians were part of the process, they would have to change its rules rather than selectively reject its conclusions.

Notes

[1]Sugden and Williams (1978, 127-31) try to finesse the issues by assuming zero income effects. Reasonable applied analysts seem to differ on the reality of this assumption.

[2]Pearce and Nash (1981, 164) observe that "no single school of thought on discount rates commands consensus among economists...the issue is one of choosing a discount rate in a second-best world, so that behaving *as if* first-best conditions prevailed...does not seem relevant."

[3]Other reasons for not pursuing systematic, explicit choice include Wildavsky's (1969) argument that ambiguity is necessary to prevent political breakdown and ultimately civil war. Leibenstein (1987) makes a related point arguing that slack keeps firms with internal conflicts from coming apart.

Bibliography

Boadway, Robin W., and Neil Bruce. 1984. *Welfare Economics*. Oxford: Basil Blackwell.

DeAlessi, Louis. 1969. "Some Implications of Property Rights Structures for Investment Choices within Government." *American Economic Review* 59:13-24.

Hoehn, John P., and Alan Randall. "Too Many Proposals Pass the Benefit Cost Test." *American Economic Review* 79: 544-51.

Little, I. M. D. 1957. *A Critique of Welfare Economics*. London: Oxford University Press.

Little, I. M. D., and J. A. Mirrlees. 1974. *Project Appraisal and Planning for Developing Countries*. New York: Basic Books.

Leibenstein, Harvey. 1987. *Inside the Firm: The Inefficiencies of Hierarchy*. Cambridge: Harvard University Press.

Majone, Giandomenico. 1989. *Evidence, Argument and Persuasion in the Policy Process*. New Haven: Yale University Press.

Mishan, E. J. 1976. *Cost-Benefit Analysis*, 2d ed. New York: Praeger.

Mishan, E. J. 1981. *Economic Efficiency and Social Welfare*. London: Allen and Unwin.

Mishan, E. J. 1982. "The New Controversy about the Rationale of Economic Evaluation." *Journal of Economic Issues* 16:29-47.

Pearce, D. W., and C. A. Nash. 1981. *The Social Appraisal of Projects*. London: Macmillan.

Schmid, A. Allan. 1989. *Benefit-Cost Analysis: A Political Economy Approach*. Boulder: Westview.

Stiglitz, J. E. 1987. "The Causes and Consequences of the Dependence of Quality on Price." *Journal of Economic LIterature* 25:1-48.

Sugden, Robert, and Alan Williams. 1978. *The Principles of Practical Cost-Benefit Analysis*. Oxford: Oxford University Press.

Tresch, Richard W. 1981. *Public Finance*. Plano, Texas: Business Publications.

Wildavsky, Aaron. 1969. "Rescuing Policy Analysis from PPBS." *Public Administration Review* 29:198-220.