

X

Take our Annual EconTalk Survey

Voting closes February 6th.

[Click here.](#)

## Public Choice

By William F. Shughart II



By William F. Shughart II,

SHARE  
POST:



Public choice applies the theories and methods of economics to the analysis of [POLITICAL BEHAVIOR](#), an area that was once the exclusive province of political scientists and sociologists. Public choice originated as a distinctive field of specialization a half century ago in the works of its founding fathers, [KENNETH ARROW](#), Duncan Black, [JAMES BUCHANAN](#), Gordon

Tullock, Anthony Downs, William Niskanen, Mancur Olson, and William Riker. Public choice has revolutionized the study of democratic decision-making processes.

## Foundational Principles

As James Buchanan artfully defined it, public choice is “politics without romance.” The wishful thinking it displaced presumes that participants in the political sphere aspire to promote the common good. In the conventional “public interest” view, public officials are portrayed as benevolent “public servants” who faithfully carry out the “will of the people.” In tending to the public’s business, voters, politicians, and policymakers are supposed somehow to rise above their own parochial concerns.

In modeling the behavior of individuals as driven by the goal of utility maximization—economics jargon for a personal sense of well-being—economists do not deny that people care about their families, friends, and community. But public choice, like the economic model of rational behavior on which it rests, assumes that people are guided chiefly by their own self-interests and, more important, that the motivations of people in the political process are no different from those of people in the steak, [HOUSING](#), or car market. They are the same human beings, after all. As such, voters “vote their pocketbooks,” supporting candidates and ballot propositions they think will make them personally better off; bureaucrats strive to advance their own careers; and politicians seek election or reelection to office. Public choice, in other words, simply transfers the rational actor model of economic theory to the realm of politics.

Two insights follow immediately from economists’ study of collective choice processes. First, the individual becomes the fundamental unit of analysis. Public choice rejects the construction of organic decision-making units, such as “the people,” “the community,” or “society.” Groups do not make choices; only individuals do. The problem then becomes how to model the ways in which the diverse and often conflicting preferences of self-interested individuals get expressed and collated when decisions are made collectively.

Second, public and private choice processes differ, not because the motivations of actors are different, but because of stark differences in the incentives and constraints that channel the pursuit of self-interest in the two settings. A prospective home buyer, for example, chooses among the available alternatives in light of his personal circumstances and fully captures the benefits and bears the costs of his own choice. The purchase decision is voluntary, and a bargain will be struck only if both buyer and seller are made better off. If, on the other hand, a politician proposes a project that promises to protect the new homeowner's community from flooding, action depends on at least some of his neighbors voting for a tax on themselves and others. Because the project's benefits and costs will be shared, there is no guarantee that everyone's welfare will be improved. Support for the project will likely be forthcoming from the owners of houses located on the floodplain, who expect to benefit the most. Their support will be strengthened if taxes are assessed uniformly on the community as a whole. Homeowners far from the floodplain, for whom the costs of the project exceed expected benefits, rationally will vote against the proposal; if they find themselves in the minority, they will be coerced into paying for it. Unless the voting rule requires unanimous consent, which allows any individual to veto a proposal that would harm him, or unless those harmed can relocate easily to another political jurisdiction, collective decision-making processes allow the majority to impose its preferences on the minority. Public choice scholars have identified even deeper problems with democratic decision-making processes, however.

## The Institutions and Mechanisms of Public Choice

It has been recognized at least since the time of the [Marquis de Condorcet](#) (1785) that voting among three or more candidates or alternatives may fail to select the majority's most preferred outcome or may be prone to vote "cycles" producing no clear winner.<sup>1</sup> Indeed, Kenneth Arrow's "impossibility theorem" shows that there is no mechanism for making collective choices, other than dictatorship, that translates the preferences of diverse individuals into a well-behaved social utility function. Nor has any electoral rule been found whose results cannot be manipulated either by individuals voting

insincerely—that is, casting their ballots strategically for less-preferred candidates or issues in order to block even worse outcomes—or by an agenda setter who controls the order in which votes are taken.

## Elections

Studying collective decision-making by committees, Duncan Black deduced what has since been called the median-voter theorem. If voters are fully informed, if their preferred outcomes can be arrayed along one dimension (e.g., left to right), if each voter has a single most-preferred outcome, and if decisions are made by simple majority rule, then the median voter will be decisive. Any proposal to the left or right of that point will be defeated by one that is closer to the median voter's preferred outcome. Because extreme proposals lose to centrist proposals, candidates and parties in a two-party system will move to the center, and, as a result, their platforms and campaign promises will differ only slightly. Reversing 1964 presidential hopeful Barry Goldwater's catchphrase, majority-rule elections will present voters with an echo, not a choice. If the foregoing assumptions hold, the median voter's preferences also will determine the results of popular referenda. As a matter of fact, anticipating that immoderate proposals will be defeated, the designers of ballot initiatives will strive to adopt centrist language, in theory moving policy outcomes closer to the median voter's ideal point than might be expected if decisions are instead made by politically self-interested representatives.

Modeling the decision to vote in a rational choice context, Anthony Downs pointed out that the act of voting itself is irrational. That conclusion follows because the probability of an individual's vote determining an election's outcome is vanishingly small. One person's vote will tip the scales in favor of the preferred candidate or issue only if the votes of all other voters are evenly split. As the number of voters becomes large, the chances of that happening quickly approach zero, and hence the benefits of voting are likely to be less than the costs. Public choice reasoning thus predicts low rates of voter participation if voters are rational. Indeed, if there is an unsolved puzzle, it is not why turnout in U.S. elections is so low, but why it is so high.

Downs and other public choice scholars also conclude that voters in democratic elections will tend to be poorly informed about the candidates and issues on the ballot. Voter ignorance is rational because the cost of gathering **INFORMATION** about an upcoming election is high relative to the benefits of voting. Why should a voter bother to become informed if his vote has a very small chance of being decisive? Geoffrey Brennan and Loren Lomasky, among others, have suggested that people vote because it is a low-cost way to express their preferences. In this view, voting is no more irrational than cheering for one's favorite **SPORTS** team.

## Legislatures

Ballot initiatives, referenda, and other institutions of direct democracy aside, most political decisions are made not by the citizenry itself, but by the politicians elected to represent them in legislative assemblies. Because the constituencies of these representatives typically are geographically based, legislative officeholders have strong incentives to support programs and policies that provide benefits to the voters in their home districts or states, no matter how irresponsible those programs and policies may be from a national perspective. Such “pork barrel” projects are especially likely to gain a representative’s endorsement when they are financed by the taxpayers in general, most of whom reside, and vote, in other districts or states.

Legislative catering to the interests of the minority at the expense of the majority is reinforced by the logic of collective action. Small, homogeneous groups with strong communities of interest tend to be more effective suppliers of political pressure and political support (votes, campaign contributions, and the like) than larger groups whose interests are more diffuse. The members of smaller groups have greater individual stakes in favorable policy decisions, can organize at lower cost, and can more successfully control the free riding that otherwise would undermine the achievement of their collective goals. Because the vote motive provides reelection-seeking politicians with strong incentives to respond to the demands of small, well-organized groups, representative democracy frequently leads to a tyranny of the minority. **GEORGE STIGLER**, Sam Peltzman,

**GARY BECKER**, and others used that same reasoning to model the decisions of regulatory agencies as being influenced by special-interest groups' relative effectiveness in applying political pressure.

The logic of collective action explains why farmers have secured government subsidies at the expense of millions of unorganized consumers, who pay higher prices for food, and why textile manufacturers have benefited significantly from trade barriers at the expense of clothing buyers. Voted on separately, neither of those legislatively enacted special-interest measures would pass. But by means of logrolling bargains, in which the representatives of farm states agree to trade their votes on behalf of trade **PROTECTIONISM** in exchange for pledges of support for agricultural subsidies from the representatives of textile-manufacturing states, both bills can secure a majority. Alternatively, numerous programs of this sort can be packaged in omnibus bills that most legislators will support in order to get their individual pet projects enacted. The legislative pork barrel is facilitated by rational-voter ignorance about the adverse effects of legislative decisions on their personal well-being. It also is facilitated by electoral advantages that make it difficult for challengers to unseat incumbents, who, accordingly, can take positions that work against their constituents' interests with little fear of reprisal.

## Bureaucracies

Owing to the benefits of specialization and division of labor, legislatures delegate responsibility for implementing their policy initiatives to various departments and agencies staffed by career bureaucrats, who secure their positions through civil service appointment rather than by democratic election. The early public choice literature on bureaucracy, launched by William Niskanen, assumed that these agencies would use the information and expertise they gained in administering specific legislative programs to extract the largest budget possible from relatively uninformed, inexpert legislators. Budget maximization was assumed to be the bureaucracy's goal because more agency funding translates into broader administrative discretion, more opportunities for promotion, and greater prestige for the agency's bureaucrats.

More recently, public choice scholars have adopted a “congressional dominance” model of bureaucracy. In that model, government bureaus are not free to pursue their own agendas. On the contrary, agency policy preferences mirror those of the members of key legislative committees that oversee particular areas of public policy, such as agriculture, **INTERNATIONAL TRADE**, and the judiciary. These oversight committees constrain bureaucratic discretion by exercising their powers to confirm political appointees to senior agency positions, to mark up bureau budget requests, and to hold public hearings. The available evidence does suggest that bureaucratic policymaking is sensitive to changes in oversight committee membership.

## **Other Institutions**

Public choice scholars, such as Gary Anderson, Mark Crain, William Shughart, and Robert Tollison, have not neglected the study of the other major institutions of democratic governance: the president or chief executive officer and the “independent” judiciary. They model the occupants of these positions as self-interested people who, by exercising the power to veto bills, on the one hand, and by ruling on the constitutionality of laws, on the other, add stability to democratic decision-making processes and increase the durability of the favors granted to special-interest groups and, hence, the amounts the groups are willing to pay for them.

## **The Lessons of Public Choice**

One key conclusion of public choice is that changing the identities of the people who hold public office will not produce major changes in policy outcomes. Electing better people will not, by itself, lead to much better government. Adopting the assumption that all individuals, be they voters, politicians, or bureaucrats, are motivated more by self-interest than by public interest evokes a Madisonian perspective on the problems of democratic governance. Like that founding father of the American constitutional republic, public choice recognizes that men are not angels and focuses on the importance of the institutional rules under which people pursue their own objectives. “In framing a government which is to be administered by men over

men, the great difficulty lies in this: you must first enable the government to control the governed; and in the next place oblige it to control itself" (*Federalist*, no. 51).

Institutional problems demand institutional solutions. If, for example, democratic governments institutionally are incapable of balancing the public budget, a constitutional rule that limits increases in spending and taxes to no more than the private sector's rate of growth will be more effective in curbing profligacy than "throwing the rascals out." Given the problems endemic to majority-rule voting, public choice also suggests that care must be exercised in establishing the domains of private and collective choice; that it is not necessarily desirable to use the same voting rule for all collective decisions; and that the public's interest can be best protected if exit options are preserved by making collective choices at the lowest feasible level of political authority.

---

## About the Author

William F. Shughart II is F. A. P. Barnard Distinguished Professor of Economics at the University of Mississippi. He is editor in chief of *Public Choice*, the field's leading scholarly journal.

---

## Further Reading Original Sources

Arrow, Kenneth J. *Social Choice and Individual Values*. 2d ed. 1951. New York: Wiley, 1963.

Black, Duncan. *The Theory of Committees and Elections*. 1958. Boston: Kulwer, 1987.

Buchanan, James M., and Gordon Tullock. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor: University of Michigan Press, 1962. Available online at: <http://www.econlib.org/library/Buchanan/buchCv3.html>

Downs, Anthony. *An Economic Theory of Democracy*. New York: Harper, 1957.

Niskanen, William A. *Bureaucracy and Representative Government*. Chicago: Aldine, Atherton, 1971.

Olson, Mancur. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge: Harvard University Press, 1965.

Riker, William H. *The Theory of Political Coalitions*. New Haven: Yale University Press, 1962.

# Comprehensive Literature Reviews

- Mueller, Dennis C. *Public Choice III*. Cambridge: Cambridge University Press, 2003.
- Rowley, Charles K., and Friedrich Schneider, eds. *The Encyclopedia of Public Choice*. 2 vols. Boston: Kluwer, 2004.
- Shughart, William F. II, and Laura Razzolini, eds. *The Elgar Companion to Public Choice*. Northampton, Mass.: Edward Elgar, 2001.
- 

## Footnotes

**1.** Consider the problem of dividing \$100 among three people. Suppose two of them agree to split that sum, with \$60 going to one and \$40 to the other. The third person, who receives nothing, has an incentive to strike a bargain with the second, offering a split of, say, \$50 each, which makes them both better off than under the initial proposal. Faced with desertion, the first person can destabilize the new coalition by offering to accept \$45, leaving \$55 for one of the others. And so on. The game has three possible (and equally likely) outcomes in which two of the three players accept payments of \$50 each, but the third player can always upset the equilibrium by cutting another deal. The same endless series of changing winning coalitions or vote “cycles” can emerge in elections involving three or more candidates or ballot issues when no one of them is strongly preferred by a simple majority of the voters.

RELATED CONTENT By Amy Willis

### A Conversation with James M. Buchanan, Parts I and II

Nobel laureate James M. Buchanan (1919-2013) was recorded in 2001 in an extended video now available to the public. Universally respected as one of the founders of the economics of public choice, he is the author of numerous books and hundreds of articles in the areas of public finance, public choice, constitutional economics, and economic philosophy. Buchanan devoted himself to the study of the contractual and constitutional basis for the theory of economic and political decision-making. In Pa...

[READ THIS ARTICLE](#)

# Public Choice: Politics Without Romance

James M. Buchanan

*Policy, the quarterly review of The Centre for Independent Studies. Spring 2003*

<http://www.cis.org.au/policy/spr03/polspr03-2.htm>

***Public choice theory demonstrates why looking to government to fix things can often lead to more harm than good, as one of its leading architects and Nobel laureate James M. Buchanan explains***

Public choice should be understood as a research programme rather than a discipline or even a sub-discipline of economics. Its origins date to the mid-20th century, and viewed retrospectively, the theoretical 'gap' in political economy that it emerged to fill seems so large that its development seems to have been inevitable.

Nations emerging from World War II, including the Western democracies, were allocating between one-third and one-half of their total product through political institutions rather than through markets. Economists, however, were devoting their efforts almost exclusively to understanding and explaining the market sector. My own modest first entry into the subject matter, in 1949, was little more than a call for those economists who examined taxes and spending to pay some attention to empirical reality, and thus to politics.

Initially, the work of economists in this area raised serious doubts about the political process. Working simultaneously, but independently, Kenneth Arrow and Duncan Black proved that democracy, interpreted as majority rule, could not work to promote any general or public interest. The now-famous 'impossibility theorem', as published in Arrow's book *Social Choice and Individual Values* (1951), stimulated an extended discussion. What Arrow and Black had in fact done was to discover or rediscover the phenomenon of 'majority cycles', whereby election results rotate in continuous cycles with no equilibrium or stopping point. The suggestion of this analysis was that majoritarian democracy is inherently unstable.

I entered this discussion with a generalized critique of the analysis generated by the Arrow-Black approach. Aren't 'majority cycles' the most desirable outcome of a democratic process? After all, any attainment of political equilibrium via majority rule would amount to the permanent imposition of the majority's will on the outvoted minority. Would not a guaranteed rotation of outcomes be preferable, enabling the members of the minority in one round of voting to come back in subsequent rounds and ascend to majority membership? My concern, then and later, was the prevention of discrimination against minorities rather than stability of political outcomes. The question, from an economist's perspective, was how to obtain a combination of efficiency and justice under majority rule.

## Wicksell's insight

The great Swedish economist Knut Wicksell was the most important of all precursory figures in public choice. In his dissertation, published in 1896, he was concerned about both the injustice and the inefficiency resulting from unfettered majority rule in parliamentary assemblies. Majority rule seemed quite likely to impose net costs or damages on large segments of the citizen or taxpayer

group. Why should members of such minorities, facing discrimination, lend their support to democratic political structures? Unless all groups can benefit from the ultimate exchange with government, how can overall stability be maintained?

These considerations led Wicksell to question the efficacy of majority rule itself. His solution to the problem was to propose that majority rule be modified in the direction of unanimity. If the agreement of all persons in the voting group is required to implement collective action, it would guarantee that all persons secure net gains and, further, that the approved actions would yield benefits in excess of costs. Of course, Wicksell recognized that, if applied in a literal voting setting, a requirement of unanimity would produce stalemate. To recognize this, however, does not diminish the value of the unanimity rule as a benchmark for comparative evaluation. In suggestions for practical constitutional reforms, Wicksell supported changes in voting rules from simple to qualified or super majorities, for example, a requirement of five-sixths approval for collective proposals.

In their analyses, Black and Arrow had assumed, more or less implicitly, that the choices to be voted on exist prior to, and outside of, the decision-making process itself. Wicksell understood the error in this assumption, although he did not recognize the importance of this insight. Neither did Gordon Tullock, who wrote a seminal paper in 1959 using the example of farmer voters, each of whom wants to have his local road repaired with costs borne by the whole community. Tullock showed that majority rule allows for coalitions of such farmers to generate election results that impose unjust costs on the whole community while producing inefficiently large outlays on local roads.

If majority rule produces unjust and inefficient outcomes, and if political stability is secured only by discrimination against minorities, how can democracy, as the organizing principle for political structure, possibly claim normative legitimacy? Wicksell's criterion for achieving justice and efficiency in collective action—the shift from majority rule toward unanimity—seems institutionally impractical. But without some such reform, how could taxpayers be assured that their participation in the democracy would yield net benefits?

### **Constitutional economics**

In implicit response to these questions, Tullock and I commenced to work on what was to become *The Calculus of Consent*, published in 1962. The central contribution of this book was to identify a two-level structure of collective decision-making. We distinguished between 'ordinary politics', consisting of decisions made in legislative assemblies, and 'constitutional politics', consisting of decisions made about the rules for ordinary politics.

We were not, of course, inventing this distinction. Both in legal theory and in practice, constitutional law had long been distinguished from statute law. What we did was to bring this distinction into economic analysis. Doing so allowed us to answer the questions posed previously: From the perspective of both justice and efficiency, majority rule may safely be allowed to operate in the realm of ordinary politics provided that there is generalized consensus on the constitution, or on the rules that define and limit what can be done through ordinary politics. It is in arriving at this constitutional framework where Wicksell's idea of requiring unanimity—or at least super majorities—may be practically incorporated.

In a sense, the analysis in our book could have been interpreted as a formalization of the structure that James Madison and his colleagues had in mind when they constructed the American

Constitution. At the least, it offered a substantive criticism of the then-dominant elevation of unfettered majority rule to sacrosanct status in political science.

Our book was widely well received, which prompted Tullock and me, who were then at the University of Virginia, to initiate and organize a small research conference in April 1963. We brought together economists, political scientists, sociologists and scholars from other disciplines, all of whom were engaged in research outside the boundaries of their disciplines. The discussion was sufficiently stimulating to motivate the formation of an organization which we first called the Committee on Non-Market Decision-Making, and to initiate plans for a journal to be called Papers on Non-Market Decision-Making.

We were unhappy with these awkward labels, and after several meetings there emerged the new name 'public choice', both for the organization and the journal. In this way the Public Choice Society and the journal Public Choice came into being. Both have proved to be quite successful as institutional embodiments of the research programme, and sister organizations and journals have since been set up in Europe and Asia.

Many sub-programmes have emerged from the umbrella of public choice. One in particular deserves mention-'rent seeking', a sub-programme initiated in a paper by Tullock in 1967, and christened with this title by Anne Krueger in 1974. Its central idea emerges from the natural mindset of the economist, whose understanding and explanation of human interaction depends critically on predictable responses to measurable incentives. In essence, it extends the idea of the profit motive from the economic sphere to the sphere of collective action. It presupposes that if there is value to be gained through politics, persons will invest resources in efforts to capture this value. It also demonstrates how this investment is wasteful in an aggregate-value sense.

Tullock's early treatment of rent seeking was concentrated on monopoly, tariffs and theft, but the list could be almost indefinitely expanded. If the government is empowered to grant monopoly rights or tariff protection to one group, at the expense of the general public or of designated losers, it follows that potential beneficiaries will compete for the prize. And since only one group can be rewarded, the resources invested by other groups-which could have been used to produce valued goods and services-are wasted. Given this basic insight, much of modern politics can be understood as rent-seeking activity. Pork-barrel politics is only the most obvious example. Much of the growth of the bureaucratic or regulatory sector of government can best be explained in terms of the competition between political agents for constituency support through the use of promises of discriminatory transfers of wealth.

As noted, the primary contribution of *The Calculus of Consent* was to distinguish two levels of collective action, ordinary or day-to-day politics and constitutional politics. Indeed, the subtitle of that book was 'Logical Foundations of Constitutional Democracy'. Clearly, political action takes place at two distinct levels, one within the existing set of rules or constitution, the other establishing the rules or constitution that impose limits on subsequent actions.

Only recently have economists broken away from the presumption that constraints on choices are always imposed from the outside. Recent research has involved the choice of constraints, even on the behavior of persons in non-collective settings, for instance, with regard to drug or gambling addiction. But even beyond that, what I have called the 'constitutional way of thinking' shifts attention to the framework rules of political order-the rules that secure consensus among members of the body politic. It is at this level that individuals calculate their terms of exchange with the state or with political authority. They may well calculate that they are better off for their membership in the constitutional order, even while assessing the impact of ordinary political actions to be contrary

to their interests.

A somewhat loose way of putting this is to say that in a constitutional democracy, persons owe loyalty to the constitution rather than to the government. I have long argued that on precisely this point, American public attitudes are quite different from those in Europe.

### **Objections to public choice**

There is a familiar criticism of public choice theory to the effect that it is ideologically biased. In comparing and analyzing alternative sets of constitutional rules, both those in existence and those that might be introduced prospectively, how does public choice theory, as such, remain neutral in the scientific sense?

Here it is necessary to appreciate the prevailing mindset of social scientists and philosophers at the midpoint of the 20th century when public choice arose. The socialist ideology was pervasive, and was supported by the allegedly neutral research programme called 'theoretical welfare economics', which concentrated on identifying the failures of observed markets to meet idealized standards. In sum, this branch of inquiry offered theories of market failure. But failure in comparison with what? The implicit presumption was always that politicized corrections for market failures would work perfectly. In other words, market failures were set against an idealized politics.

Public choice then came along and provided analyses of the behavior of persons acting politically, whether voters, politicians or bureaucrats. These analyses exposed the essentially false comparisons that were then informing so much of both scientific and public opinion. In a very real sense, public choice became a set of theories of governmental failures, as an offset to the theories of market failures that had previously emerged from theoretical welfare economics. Or, as I put it in the title of a lecture in Vienna in 1978, public choice may be summarized by the three-word description, 'politics without romance'.

The public choice research programme is better seen as a correction of the scientific record than as the introduction of an anti-governmental ideology. Regardless of any ideological bias, exposure to public choice analysis necessarily brings a more critical attitude toward politicized nostrums to alleged socioeconomic problems. Public choice almost literally forces the critic to be pragmatic in comparing alternative constitutional arrangements, disallowing any presumption that bureaucratic corrections for market failures will accomplish the desired objectives.

A more provocative criticism of public choice centres on the claim that it is immoral. The source of this charge lies in the application to politics of the assumption that individuals in the marketplace behave in a self-interested way. More specifically, economic models of behaviour include net wealth, an externally measurable variable, as an important 'good' that individuals seek to maximize. The moral condemnation of public choice is centered on the presumed transference of this element of economic theory to political analysis. Critics argue that people acting politically -for example, as voters or as legislators-do not behave as they do in markets. Individuals are differently motivated when they are choosing 'for the public' rather than for themselves in private choice capacities. Or so the criticism runs.

At base, this criticism stems from a misunderstanding that may have been fostered by the failure of economists to acknowledge the limits of their efforts. The economic model of behaviour, even if restricted to market activity, should never be taken to provide the be-all and end-all of scientific explanation. Persons act from many motives, and the economic model concentrates attention only on one of the many possible forces behind actions. Economists do, of course, presume that the 'goods' they employ in their models for predicting behaviour are relatively important. And in fact,

the hypothesis that promised shifts in net wealth modify political behaviour in predictable ways has not been readily falsifiable empirically.

Public choice, as an inclusive research programme, incorporates the presumption that persons do not readily become economic eunuchs as they shift from market to political participation. Those who respond predictably to ordinary incentives in the marketplace do not fail to respond at all when they act as citizens. The public choice theorist should, of course, acknowledge that the strength and predictive power of the strict economic model of behaviour is somewhat mitigated as the shift is made from private market to collective choice. Persons in political roles may, indeed, act to a degree in terms of what they consider to be the general interest. Such acknowledgment does not, however, in any way imply that the basic explanatory model loses all of its predictive potential, or that ordinary incentives no longer matter.

### **Impact of public choice**

Public choice theory has developed and matured over the course of a full half-century. It is useful to assess the impact and effects of this programme, both on thinking in the scientific community and in the formation of public attitudes. By simple comparison with the climate of opinion in 1950, both the punditry and the public are more critical of politics and politicians, more cynical about the motivations of political action, and less naive in thinking that political nostrums offer easy solutions to social problems. And this shift in attitudes extends well beyond the loss of belief in the efficacy of socialism, a loss of belief grounded both in historical regime failures and in the collapse of intellectually idealized structures.

As I noted earlier, when we look back at the scientific and public climates of discussion 50 years ago, the prevailing mindset was socialist in its underlying presupposition that government offered the solution to social problems. But there was a confusing amalgam of Marxism and ideal political theory involved: Governments, as observed, were modeled and condemned by Marxists as furthering class interests, but governments which might be installed 'after the revolution', so to speak, would become both omniscient and benevolent.

In some of their implicit modeling of political behavior aimed at furthering special group or class interests, the Marxists seemed to be closet associates of public choice, even as they rejected methodological individualism. But how was the basic Marxist critique of politics, as observed, to be transformed into the idealized politics of the benevolent and omniscient superstate? This question was simply left glaringly unanswered. And the debates of the 1930s were considered by confused economists of the time to have been won by the socialists rather than by their opponents, Ludwig von Mises and Friedrich Hayek. Both sides, to an extent, neglected the relevance of incentives in motivating human action, including political action.

The structure of ideas that was adduced in support of the emerging Leviathan welfare state was logically flawed and could have been maintained only through long-continued illusion. But, interestingly, the failure, in whole or in part, of the socialist structure of ideas did not come from within the academy. Mises and Hayek were not successful in their early efforts, and classical liberalism seemed to be at its nadir at mid-century. Failure came, not from a collapse of an intellectually defunct structure of ideas, but from the cumulative record of non-performance in the implementation of extended collectivist schemes-non-performance measured against promised claims, something that could be observed directly. In other words, governments everywhere overreached. They tried to do more than the institutional framework would support. This record of failure, both in the socialist and welfare states, came to be recognized widely, commencing in the 1960s and accelerating in the 1970s.

Where is the influence of public choice in this history? I do not claim that it dislodged the prevailing socialist mindset in the academies, and that this intellectual shift then exerted feedback on political reality. What I do claim is that public choice exerted major influence in providing a coherent understanding and interpretation of what could be everywhere observed. The public directly sensed that collectivistic schemes were failing, that politicization did not offer the promised correctives for any and all social ills, that governmental intrusions often made things worse rather than better. How could these direct observations be fitted into a satisfactory understanding?

Public choice came along and offered a foundation for such an understanding. Armed with nothing more than the rudimentary insights from public choice, persons could understand why, once established, bureaucracies tend to grow apparently without limit and without connection to initially promised functions. They could understand why pork-barrel politics dominated the attention of legislators; why there seems to be a direct relationship between the overall size of government and the investment in efforts to secure special concessions from government (rent seeking); why the tax system is described by the increasing number of special credits, exemptions, and loopholes; why balanced budgets are so hard to secure; and why strategically placed industries secure tariff protection.

A version of the old fable about the king's nakedness may be helpful here. Public choice is like the small boy who said that the king really has no clothes. Once he said this, everyone recognized that the king's nakedness had been recognized, but that no-one had really called attention to this fact.

Let us be careful not to claim too much, however. Public choice did not emerge from some profoundly new insight, some new discovery, some social science miracle. Public choice, in its basic insights into the workings of politics, incorporates an understanding of human nature that differs little, if at all, from that of James Madison and his colleagues at the time of the American Founding. The essential wisdom of the 18th century, of Adam Smith and classical political economy and of the American Founders, was lost through two centuries of intellectual folly. Public choice does little more than incorporate a rediscovery of this wisdom and its implications into economic analyses of modern politics.

### The Author

**James M. Buchanan**, winner of the 1986 Alfred Nobel Memorial Prize in Economic Sciences, is Distinguished Professor Emeritus of Economics at George Mason University. He is best known for developing the 'public choice theory' of economics. Reprinted from *Imprimis* (March 2003), the national speech digest of Hillsdale College ([www.hillsdale.edu](http://www.hillsdale.edu))

“In summary this theory models the *realities* rather than the romance of political institutions.”

“My primary title for this lecture, “Politics without Romance,” was chosen for its descriptive accuracy. Public choice theory has been the avenue through which a romantic and illusory set of notions about the workings of governments and the behavior of persons who govern has been replaced by a set of notions that embody more skepticism about what governments can do and what governors will do, notions that are surely more consistent with the political reality that we may all observe about us. I have often said that public choice offers a “theory of governmental failure” that is fully comparable to the “theory of market failure” that emerged from the theoretical welfare economics of the 1930’s and 1940’s.”

Excerpts from *The Logical Foundations of Constitutional Liberty*, Volume 1, pages 45-46.

# WHAT AFFECTS VOTER TURNOUT?

André Blais

Department of Political Science, Université de Montréal, Montreal PQ H3C 3J7,  
Canada; email: andre.blais@umontreal.ca

**Key Words** institutions, electoral systems, party systems, closeness

**■ Abstract** Why is turnout higher in some countries and/or in some elections than in others? Why does it increase or decrease over time? To address these questions, I start with the pioneer studies of Powell and Jackman and then review more recent research. This essay seeks to establish which propositions about the causes of variations in turnout are consistently supported by empirical evidence and which ones remain ambiguous. I point out some enigmas and gaps in the field and suggest directions for future research. Most of the research pertains to established democracies, but analyses of nonestablished democracies are also included here.

## INTRODUCTION

The dominant view in the literature is that the existing research on voter turnout has established some robust patterns, that we know relatively well why turnout is higher in some countries than in others, and that the main factors that affect variations in turnout are institutional variables. My verdict is different. Many of the findings in the comparative cross-national research are not robust, and when they are, we do not have a compelling microfoundation account of the relationship. And the impact of institutional variables may be overstated.

## THE PIONEER STUDIES

The study of voter turnout started with Powell's (1982) award-winning book, *Contemporary Democracies*, which posited electoral participation as one of the three main indicators of democratic performance, and two *American Political Science Review* articles by Powell (1986) and Jackman (1987).

Powell's APSR article examined mean turnout in 17 countries in the 1970s. He found turnout to be higher in countries with “nationally competitive districts” and “strong party-group linkages.” Nationally competitive districts enhance turnout because “parties and voters have equal incentive to get voters to the polls in all parts of the country” (Powell 1986, p. 21), and vote choice is simpler when and where groups (e.g., unions, churches, professional associations) are clearly associated with specific parties (Powell 1986, p. 22). Powell's main conclusion is that

American turnout is inhibited by its institutional context, and the main emphasis is on party-group linkages, which is the most powerful variable in his model.

Jackman's (1987) article followed in the same spirit, with an even stronger emphasis on institutions. Jackman looks at mean turnout in 19 countries in the 1970s, and he comes out with much cleaner results, showing that five institutional variables affect turnout: nationally competitive districts, electoral disproportionality, multipartyism, unicameralism, and compulsory voting.

Jackman was inspired by Powell's work, but the specific set of variables he retained was different. Most importantly, Powell's main factor, party-group linkage, was left out, because it was found to have no systematic effect. It was Jackman's set of variables that defined the research agenda.

A number of comments can be made about Jackman's study. First, although the emphasis is on institutional factors, one of them—the number of parties—should be considered as the consequence of the institutional context. Second, two variables—national competitive districts and electoral disproportionality—are aspects of the electoral system. They are correlated with each other (larger districts produce more proportional outcomes), and it is not clear why the two should be incorporated into the same model. Third, Jackman's analysis does not include any socioeconomic variable. A fuller model should integrate the role of the socioeconomic environment.

Powell's 1982 book considers a greater array of countries, 29 in total, although the analysis of turnout is restricted to 23 cases. His model distinguishes three blocs of variables: the social and economic environment, the constitutional setting (institutions in the strict sense of the term), and party systems and election outcomes. In his final path analysis model (figure 6.1, p. 121), Powell (1982) identifies four significant variables: one socioeconomic (gross national product per capita), two constitutional (proportional representation and mobilizing voting laws), and one party system (party-group linkage).

Powell's sequential model, which identifies a distant set of variables (socioeconomic), an intermediate set (institutions), and more proximate factors (party systems and election outcomes), seems quite useful. I review the evidence on the effects of these three types of factors, beginning with the impact of institutions, which has been the focus of most research.

## THE IMPACT OF INSTITUTIONS

Jackman identifies three institutions that appear to foster turnout: compulsory voting, the electoral system, and unicameralism. Other institutional variables have also been proposed.

### **Compulsory Voting**

Jackman (1987) estimates that compulsory voting increases turnout by about 13 percentage points. This pattern has been confirmed by every study of turnout in western democracies, and the magnitude of the estimated impact is almost always

around 10 to 15 points (Blais & Carty 1990; Blais & Dobrzynska 1998; Franklin 1996, 2004; Blais & Aarts 2005). “Compulsory voting increases turnout” can be construed as a well-established proposition.

This raises more questions. Must compulsory voting legislation be accompanied by sanctions in order to be efficient? What kinds of sanctions are more prone to induce recalcitrant citizens to go to the polls? How “tough” must these sanctions be? How strictly must they be enforced? The literature provides precious little to answer these questions.

Perhaps there is more to be learned from the experience of nonestablished democracies. Norris (2002) finds that compulsory voting increases turnout only in “older” democracies, and she speculates that the law may be enforced less strictly elsewhere or that its impact is conditional on the presence of broader norms about the desirability of obeying the law. Fornos et al. (2004) develop a four-point compulsory voting scale, and they report a strong impact of compulsory voting on turnout in Latin America, the region with the highest frequency of compulsory voting laws. They do not sort out, however, the specific contribution of sanctions and their degree of enforcement. Finally, Blais et al. (2003) examine the effect of compulsory voting with and without sanctions in their sample of 61 countries, covering both established and new democracies. They find that compulsory voting makes a difference only when there are sanctions (they do not examine the effect of enforcement).

In summary, we know that compulsory voting increases turnout and that its impact depends on its enforcement. But we do not know how strict that enforcement must be in order to work. We know nothing about the public’s awareness and perceptions of the law and its implementation. And there are no comparative analyses of the determinants of turnout in countries with and without compulsory voting. This is an unfortunate state of affairs. If a sense of duty is a crucial motivation for voting (Blais 2000), most people should be predisposed to vote, and loosely enforced, light fines should be sufficient to produce a high turnout. And according to rational choice, the factors that shape the decision to vote or not to vote should be very different when there is a concrete financial cost associated with abstention. In short we know nothing about the microfoundations of compulsory voting (Achen 2002, but see Bilodeau & Blais 2005).

## **Electoral System**

Jackman (1987) finds that turnout is higher in systems with nationally competitive districts, the reason being that in large districts parties have an incentive to mobilize everywhere while some single-member districts may be written off as hopeless.

Jackman’s four-category ordinal variable takes into account the electoral formula and the size of the districts. Further research has utilized either the same variable, or dummy variables that distinguish electoral formulas, or a summary disproportionality index (also included by Jackman). Studies that have been confined to advanced democracies (Blais & Carty 1990, Jackman & Miller 1995, Franklin 1996, Radcliff & Davis 2000) as well as one study of turnout in post-communist countries (Kostadinova 2003) have confirmed that turnout is higher in

proportional representation (PR) and/or larger districts, whereas research dealing with Latin America reports no association (Pérez-Liñán 2001, Fornos et al. 2004), and an analysis that incorporates both established and non-established democracies concludes that the electoral system has a weak effect (Blais & Dobrzynska 1998). See Blais & Aarts (2005) for a more detailed review of these studies.

There are two possible interpretations of the available evidence. The more optimistic view is that PR increases turnout except perhaps in Latin America, a region where there is some dose of proportionality in every country. The more pessimistic view is that once one moves outside Europe there is no generalized correlation between the electoral system and turnout. I lean toward the second, more skeptical position. On the one hand, the study by Fornos et al. (2004) that comes up with negative results in Latin America is at least as methodologically sophisticated as research on established democracies. On the other hand, as I indicate below when discussing the impact of the party system, those studies that come up with positive results have failed to specify how and why PR fosters turnout.

## Unicameralism

Jackman's (1987) last key institutional variable is unicameralism. He shows that turnout is significantly higher in countries where power is concentrated in one legislature. The reason is that, when there are two chambers, power is usually shared between the two and "elections for the lower house play a less decisive role in the production of legislation where bicameralism is strong" (Jackman 1987, p. 408). The more powerful the body that is being elected, the stronger the incentive to vote. We would expect turnout to be particularly low when and where the legislature has little power. Jackman uses a scale [proposed by Lijphart (1984)] with the highest score for unicameral countries and the lowest score for countries in which the upper house is as powerful as the lower house.

Jackman focused on the division of power between the lower and upper houses, but the same rationale should apply to the division of power between the president and the legislature, between the central government and subnational (or supranational) governments, or between the government and the courts. The general proposition is that the more powerful the body that is being elected, the higher the turnout.

Surprisingly perhaps, the findings about the impact of unicameralism on turnout are mixed. Positive results are reported by Jackman (1987), Jackman & Miller (1995), and Fornos et al. (2004). However, Blais & Carty (1990), Black (1991), Radcliff & Davis (2000), and Pérez-Liñán (2001) indicate no effect. Siaroff & Merer (2002) find support for the hypothesis that turnout is lower where there is a "relevant" directly elected president and where there are strong regional governments. Blais & Carty (1990) and Black (1991) indicate that turnout is not higher in federated countries. All in all, the studies that have looked at specific indicators of the relative power of lower chambers relative to other institutions have not systematically confirmed the conventional wisdom that turnout is higher where the lower chamber has greater leverage.

Perhaps what is needed is a summary measure of the “power” of national lower houses that takes these many dimensions into account. Blais & Dobrzynska (1998) created an “electoral decisiveness” scale that considers the presence or absence of subnational elections in federations, upper house direct elections in bicameral countries, and presidential direct elections. They find a strong positive correlation with turnout, but their results have not been replicated.

Franklin (2004) pays close attention to parliamentary responsibility in his account of turnout change in established democracies. Franklin’s main concern is to explain why turnout increases or decreases over time in different countries. His key variable is competitiveness, which I consider below. He also argues that turnout in legislative elections increases when parliamentary responsibility increases, and that it decreases when parliamentary responsibility is weakened. The former is exemplified by Malta gaining independence in the 1960s and decisions of the legislature no longer being subject to ratification by a British appointed governor. The creation of a government cartel in Switzerland after the 1960s, which made elections meaningless, is an illustration of the latter.

It is hard to believe that turnout is unaffected by the salience of an institution. Yet the empirical evidence on that question is ambiguous. The challenge is to come up with reliable scales that encompass the different dimensions of salience. The measures used in the extant research are not very satisfactory.

## Other Institutional Variables

At least two other institutional factors have been shown to affect turnout: voting age and rules designed to facilitate voting. It is a well-established fact that the propensity to vote increases with age (Wolfinger & Rosenstone 1980, Blais 2000), and so we would expect turnout to be lower when the voting age is 18 instead of 21. Research that examines turnout in contemporary advanced democracies does not incorporate that variable for the simple reason that the voting age is now 18 almost everywhere (Massicotte et al. 2004), and there is thus no variation.

Blais & Dobrzynska (1998), whose sample of elections starts in the 1970s, do include a voting age variable and they find a relatively strong effect; their results suggest that lowering the voting age from 21 to 18 reduces turnout by five points. Voting age is also a key factor in Franklin’s (2004) study of turnout dynamics. He estimates that the lowering of the voting age in most democracies has produced a turnout decline of about three percentage points.

The evidence on the effect of vote-facilitating rules is more limited and ambiguous. Franklin’s (1996) initial analysis suggests that turnout is higher when voting takes place on Sunday, so that people presumably have more time to go to the polling station, and when postal (absentee) voting is available. But these same variables proved incapable of predicting changes in turnout over time (Franklin 2004). Norris (2002) examines the effect of specific rules (number of polling days, polling on rest day, postal voting, proxy voting, special polling booths, transfer voting, and advance voting), and she finds no significant effect. Blais et al. (2003) created a summary scale that reflects the presence or absence of postal, advance,

and proxy voting, and they find a rather strong positive association between the presence of such voting facilities and turnout.

It makes sense to assume that people are more prone to vote if it is easy. Gimpel & Schucknecht (2003), in particular, have shown that turnout is affected by the accessibility of the ballot box. Likewise, there is strong evidence that allowing voters to vote by mail increases turnout (Southwell 2004, Rallings & Thrasher 2006). The question is not whether voting facilities influence turnout but rather which ones matter most, and how great a difference they make. In order to correctly address these questions we need more accurate measures of these voting facilities over time and across countries, which means that we need to know not only whether such facilities exist but also how easy it is to use them. We also need to take into account the endogeneity of election laws; measures to facilitate the vote may be more likely to be adopted in countries where turnout is low or declining (Franklin 2004, p. 148). This is no easy task. For the time being, the verdict must be that we know little about how much difference these rules make.

## Conclusions

The primary focus of cross-national studies of turnout has been on the impact of institutional variables. That impetus was shaped in good part by Jackman's influential article. The general perception in the field (and, I must confess, my own perception before I re-examined the evidence more closely in preparation for this article) is that cross-national differences in turnout can be relatively well explained by institutional variables. The perception is that we have come up with a number of well-established propositions about how institutions influence turnout.

That perception may not be well founded. We can safely assert that compulsory voting increases turnout, but we do not know whether a very light sanction suffices and whether that sanction needs to be enforced. Most of the literature supports the view that PR fosters turnout, but there is no compelling explanation of how and why, and the pattern is ambiguous when the analysis moves beyond well-established democracies. Many studies support the common-sense proposition that turnout increases with the saliency of the election, but many studies report no effect. I find it hard not to believe that turnout is higher when and where it is relatively easy to vote, yet the empirical evidence on the effect of voting facilities is inconsistent. All in all, our understanding of the impact of institutions on turnout is shaky.

## THE SOCIOECONOMIC ENVIRONMENT

We know that at the individual level the propensity to vote is associated with a number of sociodemographic characteristics, particularly age and education (Wolfinger & Rosenstone 1980, Blais 2000). It would be natural to assume, in the same vein, that cross-national variations in turnout are associated with socioeconomic differences between countries. Powell (1982) considers the impact of the socioeconomic

environment and finds that turnout does tend to be higher in more economically developed countries. He also reports that turnout tends to be higher in smaller nations, but the relationship is not statistically significant.

The most influential analyses thereafter have neglected this line of inquiry (see especially Jackman 1987 and Franklin 1996, 2004), perhaps because they deal with a small number of established democracies among which there is little variance in the level of economic development. There is, however, relatively strong support for the hypothesis that turnout is higher in economically advanced countries (Blais & Dobrzynska 1998, Norris 2002, Fornos et al. 2004). The relationship is not linear, the main difference being between the poorest countries and all others (Blais & Dobrzynska 1998).

This raises the question of whether turnout increases or decreases with downturns in the economy. As Radcliff (1992) points out, both effects are possible; economic hardship may induce people to mobilize to redress grievances, but it may also lead them to withdraw entirely from the political process. Given these two contradictory possibilities, the most likely outcome is a nil overall effect, and this is precisely what most studies report (Arcelus & Meltzer 1975, Blais & Dobrzynska 1998, Blais 2000, Kostadinova 2003, Fornos et al. 2004; for an exception see Rosenstone 1982).

Radcliff (1992) argues that economic downturns increase turnout at high and low levels of welfare spending but depress it at intermediate levels. However, some of the findings are perplexing (Blais 2000, p. 34), and they have failed to be replicated (Jackman & Miller 1995, Appendix B, note 3). The conclusion must be that there is no clear relationship between the economic conjuncture and turnout.

In my own research, I have been struck by the fact that the highest levels of turnout are reported in small countries such as Malta (Blais & Carty 1990, Blais & Dobrzynska 1998). The real difference is between very small countries and all others, and the pattern is less clear at the subnational level (see Blais 2000, p. 59). The same pattern has been observed at the local level (Oliver 2000). I have speculated that this might result from stronger social networks in smaller communities, but that hypothesis is inconsistent with the absence of a correlation between turnout and urbanization [see Siaroff & Merer 2002, Fornos et al. 2004; Kostadinova (2003) reports a negative correlation but it is quite weak]. Another interpretation is that voters are more likely to feel that their vote could be decisive in a small country. Still another interpretation, and the one I find the most plausible (although it is contradicted by Rose 2004), is that smaller countries have fewer electors per elected member, which makes it easier for candidates and parties to mobilize the vote.

Not surprisingly, political scientists have paid closer attention to the impact of institutions than to the effect of the socioeconomic environment. Still, extant research shows that turnout is substantially lower in poor countries and exceptionally high in exceptionally small countries. Few other consistent patterns have been reported. Given the prominence of the resource model in the field of political participation (Brady et al. 1995), we would expect more systematic analyses of how poverty and/or illiteracy affect turnout.

## PARTY SYSTEMS AND ELECTORAL OUTCOMES

Powell's (1982) initial analysis indicated that turnout was higher in countries with strong linkages between social groups and parties. That finding was not replicated by Jackman (1987), and subsequent studies have left out this variable. Jackman introduced a new variable, the number of parties, which is now incorporated in most research.

The intuition is that turnout should be higher the more parties there are, for at least two reasons. First, voters have more options to choose from. When there are six or seven parties instead of two or three, voters are more likely to find a party whose platform is reasonably close to their own views on the major issues of the election, and they should be less inclined to feel that none of the options is satisfactory. Second, the more parties there are, the greater the electoral mobilization.

As Jackman points out, party fractionalization may also have negative consequences on turnout. The more parties there are, the greater the likelihood that the government will be made of a coalition of parties. In systems that have coalition governments, electoral outcomes are less decisive, because the final composition of the government depends on the deals that parties are willing (or unwilling) to make. The presence of many parties may mean that voters have little say in the actual selection of the government (Downs 1957).

Because of these possible contradictory consequences, it is not clear whether we should expect the correlation of turnout with the number of parties to be positive, negative, or nonexistent. Moreover, it is not clear that it is the number of parties per se that counts. If it is the decisiveness of the outcome that matters, then we should look at the (anticipated) presence or absence of deals after the election and the most important distinction could be between elections producing single-party majority governments (which are decisive) and those producing minority or coalition governments.

Almost all empirical research has found a negative correlation between the number of parties and turnout (Jackman 1987, Blais & Carty 1990, Jackman & Miller 1995, Blais & Dobrzynska 1998, Radcliff & Davis 2000, Kostadinova 2003). The only exceptions are studies of turnout in Latin America, where there seems to be no relationship (Pérez-Liñán 2001, Fornos et al. 2004).

This is a perplexing finding. It seems to imply that people are not more inclined to vote when and where there are more options to choose from, and/or that party mobilization does not matter much (or that the arrival of new parties does not enhance overall mobilization). Furthermore, the usual interpretation that a higher number of parties reduces turnout because they produce coalition governments (and elections are therefore less decisive) is not empirically supported. Blais & Carty (1990) and Blais & Dobrzynska (1998) report that turnout is not higher in elections that produce single-party majority governments.

The bottom line is that we have a poor understanding of the relationship between the number of parties and turnout. We must reject the simple intuition that having more parties fosters turnout. This is an important nil finding. PR and/or larger

district magnitude increases the number of parties (Taagepera & Shugart 1989, Blais & Carty 1991, Lijphart 1994, Cox 1997). We can say that if PR fosters turnout, it is not because it produces more parties. Because we do not know exactly how and why PR may affect turnout—that is, we do not understand the microfoundations (Achen 2002)—the pessimistic reading that there is no generalized correlation between electoral system and turnout seems justified.

It is also time to question the standard interpretation that the often observed negative correlation between the number of parties and turnout reflects the impact of the decisiveness of elections. That interpretation must be directly tested, which means developing measures of election decisiveness. Such measures have been used in other streams of research (see, in particular, Powell & Whitten 1993); they might have to be amended or refined, but they should be incorporated in future studies. It could also be argued that what really matters is clarity of choice, that is, voters need to know with relative certainty the coalitions that might be formed. If that is the presumed process, then clarity-of-choice indicators must be constructed. As things stand, the fact that turnout appears to be lower when there are more parties is intuitively odd, and the supposition that this is so because more parties mean less decisive elections is only a supposition. (I do not find the interpretation that the number of parties increases information costs very plausible either. Voters do not have to inform themselves about each party. Information costs may increase, however, if and when the party system is in great flux.)

There is one final variable that was not included in the pioneer work of Powell and Jackman but has been incorporated in many subsequent studies: the closeness of the electoral outcome. This variable has produced the most consistent findings. My earlier summary of the evidence still holds: “the verdict is crystal clear with respect to closeness: closeness has been found to increase turnout in 27 of the 32 studies that have tested the relationship, in many different settings and diverse methodologies. There are strong reasons to believe that, as predicted by rational choice theory, more people vote when the election is close” (Blais 2000, p. 60). This is the most firmly established result in the literature. I cannot see how this finding could be wrong.

This does not mean that the issue is settled. It does not suffice to say that closeness fosters turnout; we need to specify the magnitude of the impact. I have been struck, in my own research, by its smallness. My cross-national analysis suggests that turnout is reduced by one or two points when the gap between the leading and the second parties increases by 10 points (Blais & Dobrzynska 1998). Very similar patterns emerge in cross-sectional analyses of constituency level turnout (Loewen & Blais, unpublished)<sup>1</sup> and time-series studies of national turnout in Canada (Nevitte et al. 2000).

It is possible that the impact of closeness is underestimated because the variable is not adequately measured. The standard indicator is the vote gap between the

---

<sup>1</sup> Loewen PJ, Blais A. 2005. *Did C-24 Affect Voter Turnout? Evidence from the 2000 and 2004 Elections*. Typescript.

leading and second parties. The indicator makes sense, although it is not clear whether it is the seat or the vote gap that matters. (I personally think it is the vote gap because voters receive much information about vote intentions from polls and because many have a poor understanding of how votes are translated into seats.) In systems with coalition governments, it may be the closeness of the race between the two major coalitions that matters. More complex measures need to be constructed. Furthermore, it is always assumed that the relationship between closeness and turnout is linear. It could be that what matters is that the outcome not be a foregone conclusion, and that the real difference is between elections where the winner wins by a very wide margin and all others. Or perhaps it is only very close elections that excite voters and boost turnout.

Cross-national research typically looks at the overall closeness of the national election. It could be that what matters is the closeness of the race at the district level. Franklin (2004) uses mean margin of victory at the district level as an indicator of closeness, and this is clearly an avenue worth exploring. We should not assume, however, that closeness must absolutely be measured at the district level. In an analysis of individuals' decision to vote or abstain in the 1996 British Columbia election, we found that the perceived closeness of the race at the provincial level had a greater impact than perceived closeness at the district level (Blais et al. 2000).

Finally, there is the question of whether closeness matters in PR systems. Franklin (2004) takes the radical view that the margin of victory matters only in plurality systems. He may be right, but that is an empirical proposition that should be directly tested. The most difficult question is whether closeness (or competitiveness) should be measured the same way in different electoral systems. Margin of victory is the logical indicator in plurality systems because the probability of casting a decisive vote is directly related to margin of victory. In a PR system, however, the outcome can sometimes be a foregone conclusion even if it is "close." In a five-seat district, for instance, it may be obvious to voters that party A and party B will each win two seats and party C one seat (and that could have been the outcome of the previous three elections), yet party A's lead over party B may be minuscule. Such an outcome would be coded as very close, yet the probability of casting a decisive vote in such a district would be as tiny as in single-member "noncompetitive" districts. The probability of casting a decisive vote is equally minuscule in PR and non-PR systems. We need to think hard about what closeness or competitiveness actually means in PR systems.

## DIRECTIONS FOR FUTURE RESEARCH

The dominant view in the field is that cross-national variations in turnout can be explained mostly by institutional factors that make some elections more salient and competitive than others. That view is well expressed by Franklin (1996, p. 232): "A country with low salience elections and an electoral system that was not very proportional could easily show turnout levels 40% [sic; it should be "40 percentage

points”] below a country with high salience elections and a highly proportional system. Such differences arise purely from differences in the institutional context within which elections are conducted.”

I am not convinced. As I have indicated, the evidence that turnout is higher under PR (which is supposed to produce more competitive outcomes) and in more “important” elections is far from robust. The evidence on the impact of closeness is consistent, but that impact appears to be strikingly weak. Institutions matter less than we are prone to believe. Their impact is conditional on the presence of other factors.

In order to disentangle these more complex relationships, we need to reconsider our research designs and methodologies. The standard approach in the field has been a cross-sectional analysis of variations in turnout across countries. This approach is appropriate for sorting out the effect of variables that tend to be stable over time, such as the socioeconomic environment or the electoral system and compulsory voting. The challenge is to include more cases, as the number of democracies expands, so as to test the robustness of the findings observed among established democracies.

But many variables differ from one election to another, and for these variables the analysis should be explicitly dynamic. In his ground-shaking work *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies Since 1945*, Franklin (2004) confronts the issue of moving variables and clearly points in the direction that future research should take (see also Franklin et al. 2004). Franklin makes two crucial points. First, the logical way to ascertain the impact of a variable on turnout is to examine whether turnout increases or decreases when that variable changes. In other words, the analysis should be dynamic. Second, the impact of any change should be felt mostly on the new cohorts, who have not yet developed a habit of voting (or abstaining).

This approach leads Franklin to perform empirical analyses in which previous turnout is included as a control variable, thus making the analysis explicitly dynamic. Franklin also creates interactive variables between institutional factors and the proportion of the electorate that is new (facing one of its first three elections). In some cases, he also uses cohorts as the unit of analysis, which allows him to directly test the hypothesis that institutional variables have a stronger effect on new cohorts.

This is an impressive accomplishment. Franklin (2004) has challenged us to revisit how to test hypotheses about the influence of institutions or party systems on turnout. However, the study has three serious flaws. First, Franklin omits the main effects associated with new cohorts in his estimations because of the presence of multicollinearity. This is not a compelling justification. Brambor et al. (forthcoming) show that when the theoretical model entails interaction effects, all the constitutive terms must be included and that the problems associated with multicollinearity have been greatly overstated. Second, Franklin frequently refers to how generational replacement affects turnout, yet he confines his analysis to the consequences of new cohorts entering the electorate. He does not tackle the

crucial and difficult question of whether new cohorts vote less because they enter politics in a less competitive context (his argument) or because they belong to a new generation with a different set of values (Blais et al. 2004). Third, the kind of model that Franklin proposes calls for the use of multilevel analysis, in which characteristics of voters interact with characteristics of the electoral context.

Despite these shortcomings, Franklin has indicated the new direction that research in the field has to follow. We must pay closer attention to the dynamics of turnout, we must examine how changes in the party system and/or closeness of the election outcome affect electoral participation, and we should explicitly test whether these factors have a greater impact on new cohorts. In that sense, Franklin's book is as much a pioneer study as Powell's and Jackman's work 20 years ago.

Franklin's central argument, which corresponds to the dominant view in the field, is that the degree of electoral competition is the most crucial determinant of turnout. I remain skeptical. As indicated above, turnout is only weakly affected by the closeness of an election. Extremely close elections typically "boost" turnout by a few percentage points. Furthermore, I have seen no evidence that elections are becoming systematically less competitive over time, and so the recent decline in turnout can hardly be attributed to the lack of competition.

Franklin alerts us to the possibility that the impact of institutional characteristics may vary across types of voters. We must also examine the possibility that their effects vary across systems. For instance, turnout may be differentially related to the number of parties and/or the closeness of the election in PR and non-PR countries. Likewise, what increases or decreases turnout may be quite different in rich and poor countries. Because the number of democracies and democratic elections is greatly expanding, it is now possible to test interaction effects between the socioeconomic environment, institutional variables, and party systems and to separate general patterns that hold everywhere from conditional ones that apply only in some specific contexts.

## CONCLUSION

Cross-national studies of turnout have produced a number of robust findings. We can confidently say that turnout is lower in poor countries and higher in small ones, that compulsory voting fosters turnout, and that turnout increases in closely contested elections. But I am more impressed by the gaps in our knowledge. We have a poor understanding of how compulsory voting enhances turnout, and we have a poor appreciation of how much or little competition matters and of how it plays out in PR systems. It makes sense to believe that turnout is lower in less salient elections but what makes an election more or less salient is still obscure.

We can do better. As the number of democracies and the number of democratic elections are greatly expanding, we can test our hypotheses with more cases and with greater variance in both the dependent and independent variables. This means

we must move beyond established democracies and check whether the patterns that we observe among them hold in new democracies. This is why the work of Fornos et al. (2004), which tests some of the standard hypotheses about the determinants of turnout in a new environment (Latin America), is so useful and important. If some factors, such as the electoral system or district magnitude, appear to have an impact only in some subset of countries, we should develop a more complex theory about when and where they matter more and less—or we should perform additional analyses to check whether the apparent relationship could be spurious.

With the advent of data sets such as the Comparative Study of Electoral Systems, it also becomes possible to examine the conditional impact of institutions on different types of voters (see, e.g., Long & Shively 2005). This opens up a fascinating avenue of research. Franklin (2004), Gerber et al. (2003), and Plutzer (2002) have all argued that there is an important habit component in voting. If they are right, we should expect contextual factors to have a much greater effect on new cohorts. That logically calls for a multilevel analysis linking institutional variables with individual voter characteristics.

**The Annual Review of Political Science is online at  
<http://polisci.annualreviews.org>**

## LITERATURE CITED

- Achen CH. 2002. Toward a new political methodology: microfoundations and ART. *Annu. Rev. Polit. Sci.* 5:423–50
- Arcelus FJ, Meltzer AH. 1975. The effect of aggregate economic variables on congressional elections. *Am. Polit. Sci. Rev.* 69:1232–65
- Bilodeau A, Blais André. 2005. *Le vote obligatoire a-t-il un effet de socialisation politique?* Presented to Colloque int. vote obligatoire, Inst. d’Études Polit. Lille, Oct. 20–21
- Black JH. 1991. Reforming the context of the voting process in Canada: lessons from other democracies. In *Voter Turnout in Canada*, ed. HI Bakvis, pp. 61–176. Toronto: Dundurn
- Blais A. 1991. The debate over electoral systems. *Int. Polit. Sci. Rev.* 12:239–60
- Blais A. 2000. *To Vote Or Not To Vote? The Merits and Limits of Rational Choice*. Pittsburgh: Univ. Pittsburgh Press
- Blais A, Aarts K. 2005. *Electoral system and turnout*. Presented at Int. Expert Meet. “Changing the Electoral System: The Case of the Netherlands,” Amsterdam, Sept. 14–15
- Blais A, Carty K. 1990. Does proportional representation foster voter turnout? *Eur. J. Polit. Res.* 18:167–81
- Blais A, Carty K. 1991. The psychological impact of electoral laws: measuring Duverger’s elusive factor. *Br. J. Polit. Sci.* 21:79–93
- Blais A, Dobrzynska A. 1998. Turnout in electoral democracies. *Eur. J. Polit. Res.* 33:239–61
- Blais A, Gidengil E, Nevitte N, Nadeau R. 2004. Where does turnout decline come from? *Eur. J. Polit. Res.* 43:221–36
- Blais A, Massicotte L, Dobrzynska A. 2003. *Why is Turnout Higher in Some Countries than in Others?* Ottawa: Elections Canada
- Blais A, Young R, Lapp M. 2000. The calculus of voting: an empirical test. *Eur. J. Polit. Res.* 37:181–201
- Brady HE, Verba K, Schlozman L. 1995. Beyond SES: a resource model of political participation. *Am. Polit. Sci. Rev.* 89:271–95
- Brambor T, Roberts Clark W, Golder M. 2006. Understanding interaction models:

- improving empirical analyses. *Polit. Anal.* 14:63–82
- Cox GW. 1997. *Making Votes Count*. New York/London: Cambridge Univ. Press
- Downs A. 1957. *An Economic Theory of Democracy*. New York: Harper
- Fornos CA, Power TJ, Garand JC. 2004. Explaining voter turnout in Latin America, 1980 to 2000. *Comp. Polit. Stud.* 37(8):909–40
- Franklin M. 1996. Electoral participation. In *Comparing Democracies: Elections and Voting in Global Perspective*, ed. L LeDuc, RG Niemi, P Norris, pp. 216–35. Beverly Hills, CA: Sage
- Franklin M. 2004. *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies Since 1945*. Cambridge, UK: Cambridge Univ. Press
- Franklin M, Lyons P, Marsh M. 2004. Generational basis of turnout decline in established democracies. *Acta Polit.* 39:115–51
- Gerber AS, Green DP, Shachar R. 2003. Voting may be habit-forming. *Am. J. Polit. Sci.* 47:540–50
- Gimpel JG, Schuknecht JE. 2003. Political participation and the accessibility of the ballot box. *Polit. Geogr.* 22:471–88
- Jackman RW. 1987. Political institutions and voter turnout in industrial democracies. *Am. Polit. Sci. Rev.* 81:405–24
- Jackman RW, Miller RA. 1995. Voter turnout in the industrial democracies during the 1980s. *Comp. Polit. Stud.* 27(4):467–92
- Kostadinova T. 2003. Voter turnout dynamics in post-Communist Europe. *Eur. J. Polit. Res.* 42(6):741–59
- Lijphart A. 1994. *Electoral Systems and Party Systems*. Oxford, UK: Oxford Univ. Press
- Lijphart A. 2000. Turnout. In *International Encyclopedia of Elections*, ed. R Rose, pp. 314–22. Washington, DC: CQ Press
- Long JK, Shively WP. 2005. Applying a two-step strategy to the analysis of cross-national public opinion data. *Polit. Anal.* 13:327–44
- Massicotte L, Blais A. 1999. Mixed electoral systems: a conceptual and empirical survey. *Elect. Stud.* 18:341–66
- Massicotte L, Blais A, Yoshinaka A. 2003. *Establishing the Rules of the Game: Election Laws in Democracies*. Toronto: Univ. Toronto Press
- McDonald MP, Samuel P. 2001. The myth of the vanishing voter. *Am. Polit. Sci. Rev.* 95(4):963–74
- Nevitte N, Blais A, Gidengil E, Nadeau R. 2000. *An Unsteady State: The 1997 Canadian Federal Election*. Toronto: Oxford Univ. Press
- Norris P. 2002. *Electoral Engineering: Voting Rules and Political Behavior*. New York: Cambridge Univ. Press
- Oliver JE. 2000. City size and civic involvement in metropolitan America. *Am. Polit. Sci. Rev.* 94:361–73
- Pérez-Liñán A. 2001. Neoinstitutional accounts of voter turnout: moving beyond industrial democracies. *Elect. Stud.* 20(2):281–97
- Plutzer E. 2002. Becoming a habitual voter: inertia, resources and growth in young adulthood. *Am. Polit. Sci. Rev.* 96:41–56
- Powell GB. 1982. *Comparative Democracies: Participation, Stability and Violence*. Cambridge, MA: Harvard Univ. Press
- Powell GB Jr. 1986. American voter turnout in comparative perspective. *Am. Polit. Sci. Rev.* 80(1):17–43
- Powell GB, Whitten GD. 1993. A cross-national analysis of economic voting: taking account of the political context. *Am. J. Polit. Sci.* 37:391–414
- Radcliff B. 1992. The welfare state, turnout, and the economy. *Am. Polit. Sci. Rev.* 86:444–56
- Radcliff B, Davis P. 2000. Labor organization and electoral participation in industrial democracies. *Am. J. Polit. Sci.* 44(1):132–41
- Rallings C, Thrasher M. 2006. The turnout ‘gap’ and the costs of voting: a comparison of participation at the 2001 general and 2002 local election in England. *Br. J. Polit. Sci.* In press
- Rose R. 2004. Voter turnout in the European Union member countries. In *Voter Turnout in Western Europe since 1945*, ed. R Lopez Pintor, M Gratschew, pp. 17–24. Stockholm: Int. IDEA
- Rosenstone SJ. Economic adversity and voter turnout. *Am. Polit. Sci. Rev.* 26:25–46

- Siaroff A, Merer JWA. 2002. Parliamentary election turnout in Europe since 1990. *Polit. Stud.* 50:916–27
- Southwell PL. 2004. Five years later: a reassessment of Oregon's vote by mail electoral process. *PS: Polit. Sci. Polit.* 37:89–93
- Taagepera R, Soberg Shugart M. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven, CT: Yale Univ. Press
- Wolfinger RE, Rosenstone SJ. 1980. *Who Votes?* New Haven, CT: Yale Univ. Press

## CONTENTS

---

BENTLEY, TRUMAN, AND THE STUDY OF GROUPS, <i>Mika LaVaque-Manty</i>	1
HISTORICAL EVOLUTION OF LEGISLATURES IN THE UNITED STATES, <i>Peverill Squire</i>	19
RESPONDING TO SURPRISE, <i>James J. Wirtz</i>	45
POLITICAL ISSUES AND PARTY ALIGNMENTS: ASSESSING THE ISSUE EVOLUTION PERSPECTIVE, <i>Edward G. Carmines and Michael W. Wagner</i>	67
PARTY POLARIZATION IN AMERICAN POLITICS: CHARACTERISTICS, CAUSES, AND CONSEQUENCES, <i>Geoffrey C. Layman, Thomas M. Carsey, and Juliana Menasce Horowitz</i>	83
WHAT AFFECTS VOTER TURNOUT? <i>André Blais</i>	111
PLATONIC QUANDARIES: RECENT SCHOLARSHIP ON PLATO, <i>Danielle Allen</i>	127
ECONOMIC TRANSFORMATION AND ITS POLITICAL DISCONTENTS IN CHINA: AUTHORITARIANISM, UNEQUAL GROWTH, AND THE DILEMMAS OF POLITICAL DEVELOPMENT, <i>Dali L. Yang</i>	143
MADISON IN BAGHDAD? DECENTRALIZATION AND FEDERALISM IN COMPARATIVE POLITICS, <i>Erik Wibbels</i>	165
SEARCHING WHERE THE LIGHT SHINES: STUDYING DEMOCRATIZATION IN THE MIDDLE EAST, <i>Lisa Anderson</i>	189
POLITICAL ISLAM: ASKING THE WRONG QUESTIONS? <i>Yahya Sadowski</i>	215
RETHINKING THE RESOURCE CURSE: OWNERSHIP STRUCTURE, INSTITUTIONAL CAPACITY, AND DOMESTIC CONSTRAINTS, <i>Pauline Jones Luong and Erika Weinthal</i>	241
A CLOSER LOOK AT OIL, DIAMONDS, AND CIVIL WAR, <i>Michael Ross</i>	265
THE HEART OF THE AFRICAN CONFLICT ZONE: DEMOCRATIZATION, ETHNICITY, CIVIL CONFLICT, AND THE GREAT LAKES CRISIS, <i>Crawford Young</i>	301
PARTY IDENTIFICATION: UNMOVED MOVER OR SUM OF PREFERENCES? <i>Richard Johnston</i>	329
REGULATING INFORMATION FLOWS: STATES, PRIVATE ACTORS, AND E-COMMERCE, <i>Henry Farrell</i>	353

COMPARATIVE ETHNIC POLITICS IN THE UNITED STATES: BEYOND BLACK AND WHITE, <i>Gary M. Segura and Helena Alves Rodrigues</i>	375
WHAT IS ETHNIC IDENTITY AND DOES IT MATTER? <i>Kanchan Chandra</i>	397
NEW MACROECONOMICS AND POLITICAL SCIENCE, <i>Torben Iversen and David Soskice</i>	425
QUALITATIVE RESEARCH: RECENT DEVELOPMENTS IN CASE STUDY METHODS, <i>Andrew Bennett and Colin Elman</i>	455
FOREIGN POLICY AND THE ELECTORAL CONNECTION, <i>John H. Aldrich, Christopher Gelpi, Peter Feaver, Jason Reifler, and Kristin Thompson Sharp</i>	477
ECONOMIC DEVELOPMENT AND DEMOCRACY, <i>James A. Robinson</i>	503
INDEXES	
Subject Index	529
Cumulative Index of Contributing Authors, Volumes 1-9	549
Cumulative Index of Chapter Titles, Volumes 1-9	552
ERRATA	
An online log of corrections <i>Annual Review of Political Science chapters</i> (if any, 1997 to the present) may be found at <a href="http://polisci.annualreviews.org/">http://polisci.annualreviews.org/</a>	

# Rational Choice Theory and the Paradox of Not Voting

Timothy J. Feddersen

**A**t least since Downs's (1957) seminal work *An Economic Theory of Democracy*, rational choice theorists have appreciated the “paradox of not voting.” In a large election, the probability that an individual vote might change the election outcome is vanishingly small. If each person only votes for the purpose of influencing the election outcome, then even a small cost to vote—like a minor schedule conflict or mildly bad weather—should dissuade anyone from voting. Yet it seems that many people will put up with long lines, daunting registration requirements and even the threat of physical violence or arrest in order to vote. Given the central place of voting within political economy, the lack of an adequate rational choice model of large elections with costly voting presents an obvious problem.

For the most part, theorists have bypassed the turnout problem either by eliminating voters as strategic actors or by assuming that the decision to vote is independent of other strategic choices. The problem with the first approach is that the empirical literature on voting behavior provides considerable evidence of apparently strategic behavior. In primary elections, there is evidence that voters condition their vote choice on the viability of candidates (Abramson, Aldrich, Paolino and Rohde, 1992). In a seminal and comprehensive study, Cox (1997) shows that voting patterns and election outcomes are broadly consistent with patterns of behavior predicted by strategic voting models. For example, under plurality rule (in which the candidate with the most votes wins the election),

■ Timothy J. Feddersen is Wendell Hobbs Professor of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, Illinois. His e-mail address is [⟨tfed@kellogg.northwestern.edu⟩](mailto:tfed@kellogg.northwestern.edu).

multicandidate contests generally boil down to a competition between two candidates.

The assumption that voters behave strategically in the voting booth but not when deciding to vote also seems to contradict empirical data. For example, voter turnout is inversely related to the costs of voting (Wolfinger and Rosenstone, 1980; Powell, 1986; Riker and Ordeshook, 1968). Factors that may impact the costs to vote and turnout include the weather, registration requirements, time required to think about the voting decision, distance to the polling place, and so on. Voter turnout is also correlated with education and income levels (Wolfinger and Rosenstone, 1980). If those with greater education or income have access to a better quality of information about candidates and issues, then game-theoretic models suggest that they should vote in greater numbers (Feddersen and Pesendorfer, 1996, 1999). Blais (2000), in an excellent review of the literature on turnout as it relates to rational choice models, finds that closeness of elections influences turnout. This finding suggests that voters participate because they hope to influence the ultimate outcome of the election. Excellent reviews of the empirical literature are available in Aldrich (1993), Matsusaka and Palda (1993), Blais (2000) and Mueller (2003).

Given the extensive evidence of apparently strategic voter behavior, it is unsettling that there is not a canonical rational choice model of voting in elections with costs to vote. But while a canonical model does not yet exist, the literature appears to be converging toward a “group-based” model of turnout, in which group members participate in elections either because they are directly coordinated and rewarded by leaders as in “mobilization” models or because they believe themselves to be ethically obliged to act in a manner that is consistent with the group’s interest as in “ethical agent” models. To appreciate the development of group-based models, it will be useful to begin with a discussion of the decision-theoretic literature on voting, with a focus on the paradox of not voting. Then we will move to the game-theoretic and group-based models of voting. A conclusion highlights some of the problems that group-based models of voting must address.

## **The Decision-Theoretic Approach**

The traditional starting point for the modern theory of voter turnout is the model of Riker and Ordeshook (1968), which crystallizes insights from the earlier literature such as Downs (1957) and Tullock (1967). Riker and Ordeshook analyze a model of an election with two candidates in which a single voter with preferences between the two candidates must decide whether to vote or abstain. The parameter  $B > 0$  represents the utility difference to the voter between the event that the favored candidate is elected and the event that the other candidate wins. Let  $C > 0$  be the cost of voting, and let  $p_j \in [0, 1]$  be the probability that a single vote for candidate  $j \in \{1, 2\}$  will change the outcome of the election in favor of  $j$ . We say

that  $p_j$  is the probability a vote is pivotal for candidate  $j$ . A voter who prefers candidate  $j$  should vote for  $j$  rather than abstain if and only if

$$p_j B - C > 0 \text{ or } p_j > C/B.$$

These inequalities can never be satisfied if the cost of voting  $C$  exceeds the utility difference of voting for a candidate  $B$ , so it is standard to assume that  $B > C$ , at least for many voters.

Riker and Ordeshook (1968) do not provide an explicit formula for computing the probability that an individual's vote is pivotal. However, they do observe that in large elections, the probability a vote is pivotal is very small. Consider an election in which 5 million voters are expected to cast ballots and candidate 1's expected vote share is 50.1 percent, while candidate 2 is expected to receive 49.9 percent of the votes cast. Myerson (2000) develops a formula in which the number of people who vote is a random number drawn from a Poisson distribution with mean  $n$ .<sup>1</sup> According to Myerson's formula, the probability a vote is pivotal for candidate 2 is  $8.1079 \times 10^{-9}$ . Thus, the benefit to a voter who prefers candidate 2 must be more than 8 billion times greater than the cost to vote. For example, if voting costs \$.01, then the expected benefit of electing one's favored candidate must be greater than \$80 million dollars. Expected benefits at such levels seem unreasonable.

Riker and Ordeshook (1968) modify the calculus of voting by assuming that, in addition to a cost to vote, voters get a consumption benefit  $D > 0$  from the act of voting. They propose that the  $D$  term may represent, for example, a payout from fulfilling one's civic obligation or duty to vote. The earlier inequality becomes

$$p_j B - C + D > 0.$$

<sup>1</sup> In Myerson (2000), the number of people who vote is a random number drawn from a Poisson distribution with mean  $n$ . Let  $\sigma_1$  and  $\sigma_2$  with  $\sigma_1 + \sigma_2 = 1$  be the fraction of those voting who vote for candidate 1 and 2, respectively. Then  $n\sigma_i$  is the expected number of people who vote for candidate  $i \in \{1, 2\}$ . Myerson provides the following formula for estimating the probability a vote is pivotal for candidate  $j \in \{1, 2\}$ :

$$p_j \approx \frac{e^{n(2\sqrt{\sigma_1\sigma_2}\sigma_1 - \sigma_2)}}{4\sqrt{\pi n}\sqrt{\sigma_1\sigma_2}} \left( \frac{\sqrt{\sigma_1} + \sqrt{\sigma_2}}{\sqrt{\sigma_j}} \right).$$

A useful simplification of the formula occurs in the case when the total number of potential voters in a population is a Poisson random variable with expectation  $N$  and the total number of votes cast is a Poisson random variable with expectation  $n$ . Suppose that the number of votes expected to be cast for each candidate is the same, so that  $n\sigma_1 = n\sigma_2 = n/2$ . This scenario maximizes the probability a vote is pivotal. Expected turnout is  $n/N$ . According to Myerson's formula, the probability a vote is pivotal in this case reduces to

$$\frac{1}{\sqrt{2\pi n}}.$$

Assuming  $p_j$  is practically zero, a voter casts a ballot only if the consumption benefit of voting  $D$  exceeds the cost  $C$ . The addition of a consumption benefit of voting assumes away the paradox of not voting, as was soon pointed out by Ferejohn and Fiorina (1975).<sup>2</sup> Thus, the research agenda from this model became to explain why voters might receive a consumption benefit from voting.

## Endogenizing Whether a Vote is Pivotal

Riker and Ordeshook (1968) took the probability a vote is pivotal in an election to be exogenous. Ledyard (1981, 1984) argued that the decision to vote should be embedded within a game. He looks at a model of a voting game in which voters must choose to vote for one of two candidates or else abstain. He assumes that voters only care about influencing the election outcome (that is, no consumption benefit to voting) and all voters have strictly positive costs to vote. He also adds candidates as strategic actors. He shows that when the two candidates take distinct positions, there must be positive turnout in equilibrium. The reason is clear. If nobody is voting, then the probability that a vote is pivotal is large and everyone has an incentive to vote. Ledyard does not characterize the magnitude of turnout when candidates have distinct positions. Ledyard also shows that, in large elections, candidates will converge to the median voter position and turnout will go to zero. While Ledyard's model makes a case for the efficiency of elections, it does not explain turnout.<sup>3</sup>

Palfrey and Rosenthal (1983, 1985) follow up on the Ledyard paper by characterizing the magnitude of turnout in a voting game where candidate positions are fixed and different. Palfrey and Rosenthal (1983) analyze a voting game in which each voter must choose whether to cast a costly vote for his or her preferred candidate or to abstain. The costs to vote are identical for each voter. The authors look for the Nash equilibria of this game. They find two kinds of equilibria: low turnout and high turnout. To generate high turnout in equilibrium, it is necessary to generate a high probability of being pivotal. High pivot probabilities are achieved in equilibrium by having nearly identical numbers of voters supporting each candidate. For example, a high turnout equilibrium could be sustained if 2 million voters vote for candidate 1 and 2 million vote for candidate 2. Given that the variance on the actual number of voters for each candidate is low enough and that the expected number of votes for each candidate is almost the same, very high pivot probabilities can be sustained even with high turnout. A low turnout equilibrium is sustained by having supporters of each candidate randomize between

<sup>2</sup> Ferejohn and Fiorina (1975) present a theory of turnout in which voters participate in order to minimize the maximal regret they might experience by not voting—if their vote were pivotal.

<sup>3</sup> Myerson (2000) presents an elegant reformulation of Ledyard's model as a large Poisson game.

turning out to support their candidate with low probability and abstaining. This results in low pivot probabilities.

Palfrey and Rosenthal (1985) demonstrate that the introduction of uncertainty into their earlier model eliminates the high turnout equilibria. They assume that everyone in the population has a type defined by a cost to vote  $c \in [0, 1]$  and a candidate preference  $j \in \{1, 2\}$ . For any given cost to vote, the actual number of voters with costs below that level is a random variable. They find symmetric Bayesian equilibria characterized by two cost points  $c_1$  and  $c_2$ , such that all voters who prefer candidate  $j$  and whose cost is below  $c_j$  vote for candidate  $j$ , while all others abstain. As the size of the electorate gets large, the equilibrium cost cutpoints converge to zero and turnout converges to zero. The introduction of uncertainty ensures that even if the expected number of votes for each candidate is the same, as the expected number of votes gets large, the probability the election results in an exact tie goes to zero. It follows that in any equilibrium, the cost cutpoints must converge to zero as the population size grows. Hence, only low turnout equilibria exist for large populations.

Ultimately, the game-theoretic approach to costly voting tried to escape the paradox of not voting by showing that in equilibrium, election outcomes would be close and pivot probabilities higher than in the decision theoretic literature. But the introduction of uncertainty about the actual number of voters guarantees that even if elections are expected to be close, the probability a vote is pivotal will be very low and turnout should be near zero.

## Abstention and Asymmetric Information

Feddersen and Pesendorfer (1996, 1999) take an alternative approach to explaining turnout. They assume that voting is costless and then explain why voters with strict preferences between two candidates might abstain.

Feddersen and Pesendorfer (1996) consider an electorate with three types of voters—partisans for candidate 1, partisans for candidate 2 and independents. Partisans always vote for their preferred candidate. There are two states of the world. In state 1, all independents prefer candidate 1 to candidate 2; in state 2, independents prefer candidate 2 to candidate 1. A fraction of the independents are perfectly informed about which state has occurred, and the others have no private information at all. The authors find Bayesian equilibria in which the uninformed independents will abstain with positive probability. Abstention by uninformed independents occurs at a level to balance out the expected difference between the partisans.

To gain an intuition for how this model works, consider the following example. Suppose there are four voters  $\{v_1, v_2, v_a, v_b\}$ . In state 1, all voters prefer that candidate 1 win the election. In state 2, voter  $v_1$  continues to prefer that candidate 1 win the election, but voters  $v_2, v_a$  and  $v_b$  prefer that candidate 2 win the election.

Assume that state 1 is known to occur with high probability, so that all voters would choose to have candidate 1 elected if they don't have any additional information about the state and if they alone had to choose the winning candidate. Voters  $v_1$ ,  $v_2$  and  $v_a$  do not have any additional information; however, assume  $v_i$  knows which state has occurred. Finally, suppose that the group will use majority rule to elect one of the candidates. In case of a tie, a coin flip determines the winner.

The following strategy profile is a perfect Bayesian Nash equilibrium in this example. Decisions for two of the voters are straightforward. Voter  $v_1$  votes for candidate 1, because  $v_1$  prefers candidate 1 in both states of the world. Voter  $v_i$  votes for candidate 1 when state 1 has occurred and votes for candidate 2 otherwise. The more complicated cases involve the decisions of voters  $v_2$  and  $v_a$  who do not know which state has occurred. Voter  $v_2$  will reason as follows: if state 1 has occurred, then voter  $v_2$  knows that candidate 1 will be elected (or at worst in a tie) because of receiving the votes of  $v_1$  and  $v_i$ , so voting for candidate 2 does not matter. However, if state 2 occurs, voter  $v_2$  would prefer candidate 2 and also knows that candidate 2 will receive a vote from voter  $v_i$ . The only time the  $v_2$  vote matters is when  $v_i$  has voted for candidate 2. In this situation,  $v_2$  strictly prefers to vote for candidate 2.

Now consider the strategy of voter  $v_a$ . Voter  $v_a$  understands that voter  $v_1$  will always support candidate 1 and voter  $v_2$  will always support candidate 2, while voter  $v_i$  supports candidate 1 in state 1 and candidate 2 in state 2. In this situation, voter  $v_a$  ensures that the proper candidate wins with probability one by abstaining. If voter  $v_a$  did not abstain, and instead followed the logic of voter  $v_2$ , the result could be that when state 1 has occurred, there are two votes for candidate 1 (voter  $v_1$  and  $v_i$ ) and two votes for candidate 2 (voters  $v_2$  and  $v_a$ ) and, thus, a chance that the candidate preferred by voter  $v_a$  would lose.

This illustration of the model is far from exhaustive, since the full model allows each of these four types to be groups of voters, rather than just four individual votes, but it should help to clarify some of the themes of this model. The model suggests why in equilibrium, uninformed independent voters will abstain. In versions of the model with many voters, the uninformed vote at a level that balances out the votes of partisans so that the informed independent voter is more likely to be decisive. Because uninformed independents abstain and informed independents vote, the model provides an informational explanation for why better-educated individuals are more likely to vote. As the fraction of informed independents decreases, the effect is to decrease turnout and the margin of victory. Thus, closeness and turnout are correlated in the model.

In Feddersen and Pesendorfer (1999), the authors generalize their earlier model. They show that when there is a continuum of preference types and a “fine” state space, there will be almost no abstention. The state space is fine if there are states in which a fully informed electorate would be almost perfectly split. However, Feddersen and Pesendorfer argue that it is hard to interpret a fine state space as corresponding to some empirical phenomena. They suggest that “coarse” state spaces may be more characteristic of low information elections. When the state

space is coarse, the more general model can produce the same kinds of comparative statics as in the 1996 paper.

In both papers, there is no cost to vote. If costs to vote are introduced, then only low turnout equilibria would exist for exactly the same reasons as in the earlier papers: in an election with many voters, no individual voter has a reason to believe his or her ballot will be pivotal. Thus, these papers do not provide a solution to the paradox of not voting, but they do suggest a mechanism linking closeness, turnout and information that depends on the relative magnitude of different events rather than on the absolute probability that a vote is pivotal.

## A Group-Based Voting Model of Mobilization

To generate voter turnout in costly elections, we must have some kind of consumption benefit for voting. Within the rational choice research agenda, there are two formally similar but conceptually distinct approaches that can both explain turnout and provide comparative statics consistent with strategic behavior. In both approaches, potential voters are understood to belong to groups of like-minded people who have the same preferences over the candidates. In both approaches, voters cast a ballot if and only if they receive consumption benefits from doing so—the probability an individual vote is pivotal is not relevant. Rather, the focus is on explaining changes in consumption benefits.

Mobilization models assume that groups of ideologically similar voters are coordinated by leaders who share their policy preferences. Each leader determines the level of turnout within his group by allocating costly resources to voters. It is as if leaders buy the votes of followers. As examples of such groups, one might think of unions, environmental groups or churches in which members share a common policy perspective and which have an existing organizational structure including leaders who communicate directly with members and members who meet regularly with one another.

Uhlenauer (1989) constructs a mobilization model in which voters are assumed to get two kinds of benefits from voting: a consumption benefit from voting and an “investment benefit” corresponding roughly to the earlier concept of a probability that a vote is decisive multiplied by the utility benefit of one candidate over the others. As in Riker and Ordeshook (1968), voters participate because of consumption benefits. However, unlike Riker and Ordeshook, the consumption benefits are determined in part by group leaders. Uhlaner’s model illustrates the incentives for group leaders to generate turnout for candidates that take positions the leaders prefer and that, as a consequence, candidates have incentives to take positions to generate such turnout.

Uhlenauer (1989) implicitly sets up a game between voters, leaders and candidates, but does not derive equilibrium results. Morton (1987, 1991) analyzes a game-theoretic mobilization model with a continuum of voters and many leaders.

Each leader controls the votes of a fraction of the electorate and can increase the proportion of his fraction that votes at increasing cost. The winner of the election is the candidate that receives the most votes. The game between leaders is similar in some respects to a game among a relatively small number of voters with costs to vote—the central difference being that in the mobilization game, leaders have a continuum of pure strategies. As in voting games with costly voting, Morton (1991) finds that if at least one leader strictly prefers one of the candidates, then there is an equilibrium with positive turnout. Morton does not characterize the magnitude of turnout.<sup>4</sup>

Shachar and Nalebuff (1999) examine a model of turnout in U.S. presidential elections. To capture the mechanics of the electoral college, their model is built on two leaders—one for each party—in each of 50 states. Again, leaders determine how much effort to expend to mobilize the voters in their group. Since leaders can affect the behavior of measurable fractions of the population, the probability that a change in effort by a leader can impact an election is positive even in elections with many individual voters.

Turnout occurs in group-based models with costly voting for the same basic reason that it occurs in costly voting games with a small number of voters. In equilibrium, leaders who mobilize their supporters at positive cost must increase the probability that their preferred candidate wins the election as a result. As in the costly voting game, if no leader is mobilizing any supporters, then a single group leader can mobilize a small fraction of supporters and elect that group's preferred candidate at minimal cost.

In the voting game models with individual voters, turnout in equilibrium goes to zero as the number of voters gets large. Mobilization models assume a relatively small number of groups so that each group leader controls a measurable fraction of the electorate and, in equilibrium, remains able to change the election outcome with high probability.

The biggest difficulty for mobilization models is explaining how leaders affect the micro-level decision-making of voters. The proponents of mobilization models do not argue that group leaders actually pay voters directly for votes. Such an explanation would be problematic, at least in the United States, because vote buying is illegal and the secret ballot makes it impossible to determine how someone votes—which is not to say that vote buying does not happen, but that it is relatively rare and not at a level sufficient to explain large-scale turnout. Instead, the mobilization theorists suggest that group leaders modulate social pressure by group members upon each other. For example, Uhlener (1989, p. 392) writes: “Intermediary elites can increase turnout by increasing the consumption benefits of action to their members, whether by providing money, fixing sidewalks, or increasing the normative stakes.” Similarly, Shachar and Nalebuff (1999, p. 535) write: “We

<sup>4</sup> Also see Schram (1991) for a group-based model of turnout and discussion of empirical support for the mobilization thesis.

believe the social pressure is very important. There is a contagion effect. The more people in a social network that encourage a person to vote, the more likely that person is to vote and to encourage others to do the same.”

Mobilization models that rely on social pressure explanations or on “increasing the normative stakes” don’t explicitly model how leaders generate social pressure or increase the normative stakes. Social pressure presumably relies on followers to reward and punish each other at the direction of a leader. However, if exerting social pressure is costly to followers, it is not clear how this solves the problem, since followers will have the same incentive to shirk on exerting social pressure that they do to shirk on voting in the first place. Furthermore, social pressure explanations would seem to require that voters be embedded in networks in which people monitor the voting behavior of others—one cannot punish another for not voting if the first is not aware of the voting behavior of the second. While such monitoring may be present in certain tightly knit communities, it seems unlikely to be widespread enough to explain a substantial fraction of the observed turnout. Finally, if the principle driver of participation is not direct social sanction but rather moral suasion, then the key determinant of participation centers on whether the leader can provide a compelling moral argument. So while mobilization models can provide the correct comparative statics, they are not sufficient at the micro level to explain the individual’s decision to vote.

## **Group-Based Ethical Voter Models**

There is considerable evidence that voters are motivated to vote by a sense of civic duty (Blais, 2000). In addition, there is evidence that voters base their vote choice not on how they are doing personally but on “sociotropic” assessments about the overall macroeconomic health of the economy (Kinder and Kiewiet, 1979; Markus, 1988). Sociotropic voters are thought to be motivated by altruistic or ethical concerns for the welfare of others rather than narrowly defined self-interest.

Ethical voter models provide a calculus of civic duty. In group-based ethical voter models, each voter has preferences over election outcomes where election outcomes are defined broadly to include not only which candidate wins the election but also how many people vote in the election. In some models, voter preferences are explicitly assumed to be sociotropic, but in others the extent to which preferences are “ethical” is not determined. However, the reason for applying the appellation “ethical” to these models is not that voter preferences over outcomes satisfy some normative criteria. Rather, agents are described as ethical for two reasons. First, ethical agents evaluate alternative behavioral rules in a Kantian manner by comparing the outcomes that would occur if everyone who shares their preferences were to act according to the same rule. Second, they receive a positive payoff for acting according to a behavioral rule they determine is best given their preferences and their evaluation of alternative rules.

Harsanyi (1977, 1992) carefully formalizes an ethical voter model.<sup>5</sup> Harsanyi (1977) considers a model in which some agents are rule utilitarians—that is, agents who receive an additional payoff for acting according to a strategy profile with the property that if everyone acts according to this strategy, social welfare (the sum of utilities) will be maximized. Harsanyi provides an example of an election with two candidates and costly voting. One of the candidates is assumed to maximize social welfare if elected, and a fixed fraction of the population is assumed to be voting for the socially inferior candidate. In Harsanyi's framework, a rule specifies the voters who cast a ballot and implies a cost to vote and a probability of winning for the welfare-maximizing candidate. Rule utilitarians get a payoff larger than their cost of voting if they act according to the welfare-maximizing rule. In this model, turnout will occur if the fraction of rule utilitarians in the population is large enough. Harsanyi's model provides a micro-level explanation for turnout that depends not only on the relative magnitude of the cost to vote and the payoff for acting ethically, but also upon the level of support for the inferior candidate.

However, Harsanyi's rule utilitarian model relies on the assumption that a candidate that is inferior by assumption is receiving a substantial fraction of the votes. If all rule utilitarians agree on which candidate is best, then it is not clear why an inferior candidate should receive any votes.

Feddersen and Sandroni (2002) endogenize support for both candidates by introducing preference diversity into the Harsanyi framework while preserving Harsanyi's Kantian calculus of duty. They assume a continuum of voters that can be partitioned into two types: those who believe that candidate 1 will produce a better outcome and those who believe candidate 2 is better. As in Harsanyi's model, voting is costly and each voter has a cost drawn from an interval bounded above zero and below some maximum costs. *Ceterus paribus*, all voters prefer election outcomes with lower social cost of voting to outcomes with higher social costs. Each of the groups is further partitioned into ethical voters and abstainers. Ethical voters receive a payoff greater than their cost of voting for acting ethically. All other voters abstain because their cost to vote is positive and a single vote is never pivotal.

The analysis in Feddersen and Sandroni (2002) focuses on determining the ethical rule for ethical voters of each type. A rule is defined as a cost cutpoint for a type such that all voters of the given type with costs to vote below the cutpoint have an ethical duty to vote for their preferred candidate, while those with costs to vote above the cutpoint have a duty to abstain. Taking as given the behavior of agents with different preference types, agents evaluate rules for their type according to their preferences over the social outcomes produced by the rule. A behavior profile is *consistent* if the behavior of each agent follows from the agent's preferred rule (that is, the rule that produces the best outcome from that agent's perspective if all voters sharing that type act according to that rule). The pair of cost cutpoints

<sup>5</sup> See Margolis (1982) for a less formal ethics-based model of participation.

defined by a consistent profile are identical to the cutpoints that would be chosen in equilibrium in a game in which the turnout for each group was determined by a leader maximizing social welfare for his or her group.

Feddersen and Sandroni (2002) analyze the comparative statics on consistent behavior profiles and show that their model will deliver comparative statics on turnout and margin of victory that are consistent with empirical findings. Turnout and margin of victory are positively correlated but not because of changes in pivot probabilities. Instead, as the relative size of the two groups of voters become more equal, turnout increases for both the larger and smaller group and margin of victory decreases. Turnout is also decreasing in costs to vote. However, unlike mobilization models, leaders, social pressure or selective incentives play no role.

A variety of empirical work provides support for group-based models. The earlier discussion has mentioned the empirical work demonstrating correlations between turnout and costs to vote, margin of victory and civic duty, all of which can be interpreted in group-based terms mentioned above, along with the evidence that effort by elites can increase turnout mentioned in the mobilization section. There is a large empirical literature on mobilization and turnout that readers may be interested in as well.<sup>6</sup> Hill and Leighley (1996) examines the roll of parties and find that mobilization efforts by parties can impact turnout. Leighley (1996) reviews the literature on group membership and mobilization. She finds that mobilization due to intentional efforts by group leaders are restricted to explicitly political groups, while unintentional mobilization occurs as a consequence of membership in both political and nonpolitical groups. In addition, recent work by Coate and Conlin (2002) presents empirical evidence supporting group-based turnout models from a Texas referendum.<sup>7</sup>

## Conclusion

The mobilization and ethical agent models have similar formal structures but rest on different conceptual grounds. Basic technical and conceptual questions remain for both approaches. A central problem centers on the existence of equilibria. If, as is done in both the mobilization and ethical voter models above, the strategy space for leaders is continuous, then the existence of equilibria or consistent profiles is not assured. In group-based models, if the strategy space is finite, then there still may not be equilibria in pure strategies. Unfortunately, mixed strategy equilibria present conceptual problems for group-based models.

In a mixed strategy equilibrium, each leader knows the distribution over pure

<sup>6</sup> See Pollock (1982) and Verba and Nie (1972) for earlier work on mobilization and participation.

<sup>7</sup> See also Schram (1991, 1992) and Schram and Sonnemans (1996a, b) for empirical and experimental evidence in support of the importance of groups in explaining turnout. Thanks to Becky Morton for bringing Schram's work to my attention.

strategies employed by the other, but does not know which particular strategy has been chosen. This setup requires that leaders can convey instructions to followers without the instructions being overheard by the other leader. The possibility of such private communications seems remote in the setting of a large election. The problem is compounded in the ethical agent model because no leader exists to coordinate the voters. All voters of the same type act according to a rule that instructs each whether they are to vote or not. Even if a rule required a measurable fraction of voters to randomize, it is still the case that, without a coordinating device, each voter's randomization would be independent. Since there is a continuum of voters, a rule that requires independent randomizations is outcome equivalent to a rule that does not. Thus, if there is no pair of rules that is each a best response to the other, there is no consistent behavioral profile.<sup>8</sup>

Finally, and perhaps most importantly, neither the mobilization nor ethical agent models provide an explanation for how people join or identify with their groups in the first place. Imagine first that the decision to join a group is independent of political concerns, which may hold true of certain groups (perhaps labor unions), or that the decision to have certain ethical preferences transcends daily political choices and is in that sense independent of political concerns. In these cases, we could be comfortable treating groups and ethics as exogenous to the models. But if the choice to join a group is independent of political concerns, why should we expect group members to have similar policy preferences? On the other hand, if the decision to join a group is partly a function of a desire to affect political outcomes, then the decision to join or to identify with a group should be endogenized. A group-based model of voting, whether based on mobilization or on ethics, must ultimately come to grips with the questions of why people join groups.

<sup>8</sup> Feddersen and Sandroni (2002) provide sufficient conditions for pure strategy equilibria in the case of two groups of voters, however, no results are offered for the case of multiple groups.

## References

- Aldrich, John.** 1993. "Rational Choice and Turnout." *American Journal of Political Science*. 37:1, pp. 246–78.
- Abramson, Paul R., John H. Aldrich, Phil Pao-  
lino and David W. Rohde.** 1992. "Sophisticated" Voting in the 1988 Presidential Primaries." *American Political Science Review*. 86:1, pp. 55–69.
- Blais, Andre.** 2000. *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory*. Pittsburgh: University of Pittsburgh Press.
- Coate, Steven and Michael Conlin.** 2002. "Voter Turnout: Theory and Evidence from Texas Liquor Referenda." NBER Working Paper No. w8720.
- Cox, Gary W.** 1997. *Making Votes Count*. Cambridge: Cambridge University Press.
- Downs, Anthony.** 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Feddersen, Timothy J.** 1992. "A Voting Model Implying Duverger's Law and Positive Turnout." *American Journal of Political Science*. 36:4, pp. 938–62.
- Feddersen, Timothy J. and Wolfgang Pesendorfer.** 1996. "The Swing Voter's Curse." *American Economic Review*. 86:3, pp. 408–24.
- Feddersen, Timothy J. and Wolfgang Pesendorfer.** 1999. "Abstention in Elections with Asymmetric Information and Diverse Preferences." *American Political Science Review*. 93:2, pp. 381–98.
- Feddersen, Timothy J. and Alvaro Sandroni.** 2002. "A Theory of Participation in Elections with Ethical Voters." Working paper.
- Ferejohn, John A. and Morris P. Fiorina.** 1975. "Closeness Counts Only in Horseshoes and Dancing." *American Political Science Review*. 69:3, pp. 920–25.
- Harsanyi, John.** 1977. "Morality and the Theory of Rational Behavior." *Social Research*. 44:4, pp. 623–56.
- Harsanyi, John.** 1992. "Game and Decision Theoretic Models in Ethics," in *The Handbook of Game Theory, Volume 1*. R. Aumann and S. Hart, eds. Amsterdam, The Netherlands: Elsevier, North-Holland, chapter 19.
- Hill, Kim Quaile and Jan E. Leighley.** 1996. "Political Parties and Class Mobilization in Contemporary United States Elections." *American Journal of Political Science*. 40:3, pp. 787–804.
- Kinder, Donald R. and D. Roderick Kiewiet.** 1979. "Economic Discontent and Political Behavior: The Role of Personal Grievances and Collective Economic Judgments in Congres-
- sional Voting." *American Journal of Political Science*. 23:3, pp. 495–527.
- Kramer, Gerald H.** 1983. "The Ecological Fallacy Revisited: Aggregate versus Individual-Level Findings on Economics and Elections and Socio-tropic Voting." *American Political Science Review*. June, 77, pp. 92–111.
- Ledyard, John.** 1981. "The Paradox of Voting and Candidate Competition: A General Equilibrium Analysis," in *Essays in Contemporary Fields of Economics*. George Hoorwich and James P. Quick, eds. Lafayette: Purdue University Press, pp. 54–80.
- Ledyard, John.** 1984. "The Pure Theory of Two Candidate Elections." *Public Choice*. 44:1, pp. 7–41.
- Leighly, Jan.** 1996. "Group Membership and the Mobilization of Political Participation." *Journal of Politics*. 58:2, pp. 447–63.
- Margolis, Howard.** 1982. *Selfishness, Altruism and Rationality*. Cambridge: Cambridge University Press.
- Markus, Gregory.** 1988. "The Impact of Personal and National Economic Conditions on the Presidential Vote: A Pooled Cross-Sectional Analysis." *American Journal of Political Science*. 32:1, pp. 137–54.
- Matsusaka, John and Filip Palda.** 1993. "The Downsian Voter Meets the Ecological Fallacy." *Public Choice*. December, 77:4, pp. 855–78.
- Morton, Rebecca.** 1987. "A Group Majority Model of Voting." *Social Choice and Welfare*. 4:2, pp. 117–31.
- Morton, Rebecca.** 1991. "Groups in Rational Turnout Models." *American Journal of Political Science*. August, 35, pp. 758–76.
- Mueller, Dennis C.** 2003. *Public Choice III*. Cambridge: Cambridge University Press.
- Myerson, Roger B.** 2000. "Large Poisson Games." *Journal of Economic Theory*. 94:1, pp. 7–45.
- Palfrey, Tom and Howard Rosenthal.** 1983. "A Strategic Calculus of Voting." *Public Choice*. 41:1, pp. 7–53.
- Palfrey, Tom and Howard Rosenthal.** 1985. "Voter Participation and Strategic Uncertainty." *American Political Science Review*. 79:1, pp. 62–78.
- Pollock, Phillip H. III.** 1982. "Organizations as Agents of Mobilization: How Does Group Activity Affect Political Participation?" *American Journal of Political Science*. 26:3, pp. 486–503.
- Powell, G. Bingham.** 1986. "American Voter Turnout in Comparative Perspective." *American Political Science Review*. 80:1, pp. 17–43.

- Riker, William and Peter Ordeshook.** 1968. "A Theory of the Calculus of Voting." *American Political Science Review*. March, 62, pp. 25–42.
- Shachar, Ron and Barry Nalebuff.** 1999. "Follow the Leader: Theory and Evidence on Political Participation." *American Economic Review*. 89:3, pp. 525–47.
- Schram, A.J.H.C.** 1991. *Voter Behavior in Economic Perspective, Studies in Contemporary Economics*. Heidelberg: Springer Verlag.
- Schram, A.J.H.C.** 1992. "Testing Economic Theories of Voter Behavior Using Micro-Data." *Applied Economics*. 24:4, pp. 419–28.
- Schram, A.J.H.C. and J. Sonnemans.** 1996a. "Voter Turnout and the Role of Groups: Participation Game Experiments." *International Journal of Game Theory*. 25:3, pp. 385–406.
- Schram, A.J.H.C. and J. Sonnemans.** 1996b. "Why People Vote: Experimental Evidence." *Journal of Economic Psychology*. 17:4, pp. 417–42.
- Tullock, Gordon.** 1967. *Towards a Mathematics of Politics*. Ann Arbor: University of Michigan Press.
- Uhlamer, Carol.** 1989. "Rational Turnout: The Neglected Role of Groups." *American Journal of Political Science*. 33:2, pp. 390–422.
- Verba, Sidney and Norman H. Nie.** 1972. *Participation in America: Political Democracy and Social Equality*. New York: Harper and Row.
- Wolfinger, Raymond E. and Stephan J. Rosenstone.** 1980. *Who Votes?* New Haven: Yale University Press.

## The median voter model in public choice theory

RANDALL G. HOLCOMBE

*Department of Economics, The Florida State University, Tallahassee, FL 32306*

### 1. Introduction

When public choice first began emerging as a distinct subdiscipline, the median voter model, built on the foundations of Black (1958), Downs (1957) and Bowen (1943), was considered one of the most solidly established models in public choice.<sup>1</sup> But while at the beginning of the 1970s the median voter model was accepted almost without question in public choice, by the end of that decade it had been assaulted on so many fronts that it was virtually abandoned. As Mueller (1987: 245), remarked, after noting that the median voter model is descriptive under certain conditions, "subsequent research has shown that these conditions are so unlikely to hold in practice that Black's important theorem stands, from the perspective of today, as a kind of special case illustrating the fragility inherent in majority rule committee outcomes."<sup>2</sup>

The purpose of this paper is to review the role of the median voter model in public choice theory, paying special attention to the subsequent research that Mueller has noted. If by the median voter model one means that the public sector produces what the median voter wants, then doubt about the model's veracity is warranted. But if the model instead is interpreted to mean that under general circumstances the demand for public sector output is a democracy is the median voter's demand, then writing off the median voter model as a special case seems unwarranted.

The thesis of this paper is that the median voter model in the public sector has served in much the same role as the model of pure competition in the private sector. In general, microeconomic theory models the economy by starting with the competitive model and then imposing elements of monopoly on it in various ways. Similarly, public sector demand can be understood using the median voter model as a foundation and then imposing various complications – such as multi-peaked preferences and agenda control – on the fundamental model. If this notion is accepted, it provides a foundation for a theory of political structure that parallels the theory of market structure, with the median voter model fulfilling a role analogous to perfect competition in the market. Just because the median voter model is not descriptive of every political market does not mean that it cannot provide a solid foundation for the analysis of public sector demand. And when the evidence is examined, there appears to be

good reason to accept the median voter model as descriptive in a wide range of majority rule decisions.

## **2. The rise of the median voter model**

Following Downs's and Black's work in the late 1950s the median voter model enjoyed a decade of increasing acceptance along with the growing area of public choice. In the early 1970s articles were published in journals such as the *American Economic Review* and the *Journal of Political Economy* which accepted public sector output as representing the median voter's demands without any more evidence than simply citing the existence of the model.<sup>3</sup> Even at this point some potential problems – most notably the cyclical majority problem popularized by Arrow (1951) – were well-recognized. Acceptance of the model nonetheless could partly be attributed to Black's demonstration that cycles could arise only when preferences were not single peaked, and not always even then. While potential problems were recognized, the problems were judged to be insignificant enough that the median voter model result could be assumed without demonstration in public choice models.<sup>4</sup>

Another reason for the model's acceptance is that the median voter model refers not to one actual model, but to a group of models that all arrive at a common result. That the median voter's preference will emerge as the group preference generated by majority rule was demonstrated by Black (1958) for committee type majority rule decisions, by Bowen (1943) in a referendum setting, and by Downs (1957) under representative democracy. With majority rule producing the same outcome under many different specific institutional implementations of majority rule suggests the robustness of the median voter result.

## **3. The decline of the median voter model**

Perhaps the first assault on the median voter model was Niskanen's (1971) portrayal of budget maximizing bureaucracy. Niskanen (1971: 139) explicitly uses the median voter model as the foundation of the demand side of his model. But the conclusions from Niskanen's model were so at odds with the generally perceived conclusions of the median voter model that this model was considered a critique of the median voter model.

The median voter model is clearly only a model of demand aggregation in the public sector, but since at the time there were no generally accepted models of public sector supply, the median voter model was taken to imply that public sector output was that output most preferred by the median voter. Niskanen's innovation was building a supply side of the market and mating it

to the median voter model of demand in order to develop a complete supply and demand model of the public sector. But in so doing, Niskanen's conclusions were at odds with the conclusion generally accepted as being implied in the median voter model that the public sector produced what the median voter wanted.

The next step in the decline of the median voter model was to build Niskanen's budget maximization hypothesis directly into the voting models themselves. An excellent example is the agenda control model examined extensively by Romer and Rosenthal (1978, 1979a, 1979b, 1982). The agenda control model uses a referendum process much like Bowen (1943) described, but illustrates that a budget maximizing agenda setter can manipulate the alternatives in such a manner as to produce an outcome larger than would be most preferred by the median voter.

Control of the agenda provides a powerful tool for manipulating political outcomes,<sup>5</sup> but within a referendum setting such as that employed by Romer and Rosenthal, the possibility of agenda control begs more questions than it answers. In the Romer and Rosenthal model the referendum acts only as a constraint on the agenda setter, and since most public sector spending decisions are made without approval via referendum, if their model is descriptive of actual referenda, it leaves open the more important question of what constraints the size of the budget in the more typical case when a referendum is not held.<sup>6</sup>

The most devastating blow to the median voter model was dealt by McKelvey (1976), who used a model with assumptions similar to Black's, where individuals have single peaked preferences and motions can be considered sequentially by majority rule.<sup>7</sup> The main difference between McKelvey's model and Black's is that McKelvey expands the model so that political issues are multidimensional rather than single dimensional. But with this change McKelvey illustrated that an agenda setter could start at any point in the issue space and by strategically selecting issues end up at any other point in the issue space, so that there is no unique and stable majority rule outcome. Reactions to McKelvey's model varied,<sup>8</sup> but the point here is that the median voter model, accepted almost without question in the early 1970s, was considered to be just a special case not very relevant to the real world only a decade later.

#### **4. Theoretical issues**

One explanation of the rapid decline of prominence of the median voter model in public choice theory is that it has suffered from excess in the application of theoretical results to the real world. At the height of its prominence its modest claim that under some circumstances the median voter's preference will emerge as the collective preference in a majority rule election system often was taken

to imply that the public sector produces what the median voter wants. There are two reasons why this exaggerated claim might not be true.

First, the median voter model says nothing about the supply side of the public sector. Niskanen's contribution in this regard was noted earlier. But adding a supply side to the model does not refute the model of demand aggregation, and Niskanen explicitly employed the median voter construction. Along these lines, note the significant difference between McKelvey's (1976) model and Romer and Rosenthal's (1978). While both are about agenda control, Romer and Rosenthal use the median voter construction and add a budget maximizing agenda setter to show how supply side considerations can affect the referendum variant of the median voter model. But their model is a median voter model nonetheless, using the same demand side but adding a supply side which was not included in the original formulation of the model. McKelvey, on the other hand, also writes about agenda control, but shows conditions where the median voter framework is inappropriate. Thus, McKelvey rejects the basic median voter framework, whereas Romer and Rosenthal accept it and build their model onto it.

The lack of a supply side in the model is not sufficient reason to reject the median voter model itself. The model could be faulted for being incomplete, or more precisely, the model's proponents could be faulted for taking a purely demand side model and using it as a description of the entire market for public sector output, but previous exaggerated claims for the model should not discredit the model entirely.

The second reason why the exaggerated claims of the median voter model might not be true is that the model itself may not be descriptive of the majority rule decision making process, which is the critique of the model within McKelvey's framework. The claim is not that the model is not correct, but rather that there are plausible conditions under which majority rule will not produce a determinate equilibrium. Considering Arrow's (1951) book, this finding has been a part of the public choice literature since the beginning of public choice.

Putting these two critiques in perspective, they amount to determining when the conditions exist that will produce a median voter equilibrium and when those conditions do not exist. This has not changed. What has changed is that public choice theorists in general now believe that the median voter model is less generally applicable than they formerly believed.

A similar sequence of events occurred after the rigorous proof of the uniqueness and stability of a competitive equilibrium in the market. With a fragile mathematical proof as a target, economists set about proving that the conditions of a competitive equilibrium could never be met, the spirit of which is fully captured in Bator's classic article, 'The Anatomy of Market Failure' (1958). A similar taxonomy on the anatomy of government failure has yet to be written, but the literature to form its foundation is already in place. The median

voter model provides a benchmark in that it describes a unique and stable majority rule equilibrium so that theorists – much as they did with the competitive equilibrium – can demonstrate how easy it is to develop plausible conditions under which the model’s conclusions will not hold. But, that having been done, it is worthwhile to step back and examine the model and its exceptions to see when the model’s conclusions might ever be expected to hold true.

## **5. The applicability of the model**

The median voter model, as a model of public sector demand, should not be expected to show a full public sector equilibrium. The median voter model is simply a model of demand aggregation, so that in the same way that individual demand curves are summed in a private market to find the market demand, the median voter model depicts the market demand when aggregated by majority rule to be the demand of the median voter. The question, then, is when conditions would exist such that the median voter model accurately describes the demand aggregation process.

The model’s developers and its critics agree that it is descriptive when the alternatives to be considered can be ranked on a single dimensioned continuum such that individual preferences for the alternatives are single peaked. This statement requires some elaboration. First, it is not necessarily true that multiple peaked preferences will lead to indeterminacy. In one dimension, multiple peaked preferences are a necessary but not sufficient condition for there to be no median voter outcome. Second, saying the model is descriptive does not necessarily imply that the median voter will get what he or she wants. For example, agenda control in the Romer and Rosenthal model operates by identifying the median voter and then manipulating the alternatives that the median voter faces. Third, the alternatives do not have to be related in any manner other than being able to be placed on the continuum such that there are not multiple peaked preferences. For example, one candidate might run solely on a platform related to an arms control treaty, while another runs on a platform devoted to balancing the government budget, and a third runs on an antiabortion platform. These are three separate issues that produce a three-dimensional issue space, but if the candidates can be ranked in one dimension such that preferences for the candidates are single peaked, then the median voter model applies.

This third point is significant when considering the applicability of the median voter model to national electoral politics. In the 1984 presidential election, for example, the three plausible candidates were Mondale, Hart, and Reagan. The issues were complicated, ranging from abortion to taxation to foreign policy, making it appear as a multidimensional setting. But the voters do not have

a choice of selecting any point in an n-dimensional issue space; rather, they have the choice of only three points in the issue space, and it is plausible that most voters would have single peaked preferences ranking those three points on a single dimensioned continuum from left to right such that Mondale is the furthest left and Reagan the furthest right. This requires only that those preferring Mondale would choose Hart over Reagan and those preferring Reagan would choose Hart over Mondale. If this condition is met, then because only a subset of points in the multidimensional issue space are available, the issue space can be collapsed into one dimension and the median voter model applies.<sup>9</sup> It should be noted that this will always be the case when there are only two choices. Seen in this light, it appears plausible that in general electoral politics the limited number of points in the issue space over which voters have the option of choosing creates a structure whereby candidates can be ranked on a single dimensioned continuum in such a way as to have a determinate median.

This is not proof of the applicability of the median voter model. Indeed, proof is not possible that the assumptions or conclusions of a theoretical model accurately reflect the conditions of the real world. But arguments taken to discredit the median voter model as a general description of majority rule electoral outcomes do so only by showing that in theory there are many conditions that could cause the model's assumptions to be unrealistic. This does not by itself imply that the median voter model is any less relevant to majority rule voting than the model of pure competition is to markets.

## **6. Empirical evidence**

Empirical evidence can be consistent or inconsistent with a hypothesis, but it can never provide irrefutable proof. With this in mind, it is worth noting that most of the evidence published regarding the median voter model is consistent with the model. This by itself should suggest that the model is more than just a special case that is rarely applicable to the real world.

A test of the hypothesis was presented by McEachern (1978), who divided states into three categories: those which require no referenda to approve school spending, those which require simply majority rule referenda, and those which require more than a simple majority to approve. He found no difference between the majority rule states and the non-referendum states, but did find significantly lower spending in the states which required approval of more than a simple majority. He concluded that the non-referendum states were already at a median voter equilibrium as a result of Downsian competition, so majority rule referenda would produce the same result. More than a simple majority rule makes some other voter (e.g., the 60th percentile voter if 60 percent approval is required) the decisive voter, thus differentiating those states.

Empirical tests by Holcombe (1980) and Munley (1984) attempted to identify the median voter from referendum data and compare the median voter's most preferred outcome with the actual referendum outcome. In both cases, the evidence was consistent with the median voter hypothesis. These tests are of further interest because they are done in the same theoretical framework as Romer and Rosenthal used for their agenda control models of referenda, but the evidence was consistent with the Bowen model rather than the Romer and Rosenthal model. Romer and Rosenthal (1982) present a test in which they try to differentiate their agenda setter model from the median voter model. They conclude (p. 577), ‘‘Rather than resolving the question of the simple setter model against the simple median voter model, our results indicate that both may be inappropriate . . .’’ and suggest further investigation.

Studies by Pommerehne and Frey (1976), Pommerehne (1978), and Inman (1978) examine test the median voter model by comparing predictions using the median versus mean incomes and find the median to be the better predictor. Holcombe (1977) describes a referendum system once held in Florida that always picked the median voter outcome. In short, there have been a number of empirical tests of the model over the years, all taking very different approaches, and most have found the evidence to be consistent with the model. Mueller (1979: 106–111) cites several other empirical studies which reinforce the conclusion that, in general, empirical work on the subject is consistent with the median voter model. It is worthwhile noting, at the very least, that the preponderance of evidence in the literature is consistent with the model that some are willing to write off as irrelevant.

## **7. Toward a theory of political structure**

Over the past decade, public choice theorists have become increasingly aware that the institutional structure in the public sector is an important determinant of the characteristics of the resulting outcome. Around 1970, the median voter model was accepted as descriptive of the characteristics of majority rule equilibrium regardless of the specific institutions, and one result of the work on public sector supply models and disequilibrium voting models has been to focus the theorist's attention more closely on the specific political institutions and their impact on the resulting outcome. The disequilibrium and public sector supply models have shown that there are a number of reasons why the median voter model might not be descriptive of reality.

But showing that there are conditions where the median voter model might not be descriptive of reality does not imply that the model is never descriptive or that it should be regarded as a special case rarely relevant to real world collective choices. The earlier sections of the paper argued that there is both theo-

retical and empirical evidence weighing in the model's favor. The evidence does not imply that the public sector produces what the median voter wants; this claim is stronger than the model warrants, but it is the claim that is often attacked under the name of the median voter model. Rather, the evidence suggests that the median voter model is a good description about how demands are aggregated under majority rule in many circumstances.

Recall that some of the critiques of the median voter model use the median voter framework to develop alternatives. This method accepts the general validity of the model as a demand aggregation mechanism while attacking the overly strong conclusion that the public sector produces what the median voter wants. Combined with the evidence on the model, this suggests that the more appropriate role of the model is to provide a polar case of demand aggregation under majority rule in much the same way that the model of pure competition provides the polar case with regard to the market.<sup>10</sup>

In microeconomics, the competitive model provides the polar case upon which imperfections are added to form more realistic models. In public choice theory, the median voter model can provide a similar polar case of majority rule decision making under ideal 'competitive' conditions. But again, complications can arise. These complications do not disprove the median voter model any more than the monopoly model disproves the model of competition. Rather, they provide a method for building on to the basic model a description of how the real world can differ from the competitive ideal under certain conditions.

The median voter model has already been used like this in Niskanen's bureaucracy model and in Romer and Rosenthal's model of agenda control, to cite two examples. Theorists wanting to explore supply conditions or develop models of bureaucracy accept some notion of stable public sector demand implicitly rather than some indeterminate or forever cycling demand. Empirical work in public choice also accepts the notion, although sometimes implicitly, because in most circumstances it would make no sense to run regressions on observed characteristics that are assumed to be unstable. If this implicit recognition were made explicitly, then the median voter model would be in more of a position to provide a foundation for the theory of public sector demand. As it stands, the model is often accepted implicitly while its relevance is explicitly questioned.

## **8. Conclusion**

The changing perception by public choice theorists about the relevance of the median voter model is a result of excessive extrapolation of the conclusions of theoretical models to the real world. Early in the 1970s the median voter model

was often accepted as implying that the output produced in the public sector was what was most preferred by the median voter. This claim is excessive because the median voter model is only a model of demand aggregation under majority rule and has little to say about the supply side of the public sector. In the late 1970s many scholars identified several circumstances under which the model would not apply in theory, but these critiques of the model were often viewed as reasons to abandon the median voter model altogether. The model went from having excessive claims that made the model appear to be more powerful than it really is to excessive claims that made the model appear to be less powerful than it really is. These latter claims were often in response to the earlier claims rather than to the model, appropriately applied.

Pointing out that the model might not be valid under some circumstances in no way implies that the model is never valid. In fact, this paper has reviewed strong arguments, both empirical and theoretical, suggesting that the median voter model is a good approximation of demand aggregation in the public sector for many issues. One paper will not change the opinions of public choice theorists on the median voter model. But the argument given here is that there is a large amount of theoretical and empirical evidence supporting the median voter model as a good foundation for the development of the theory of public sector demand. Once the overly ambitious claims that have been made for the model are set aside, the median voter model is in a good position to provide a base for the development of a theory of political structure that is analogous to the theory of market structure in economics.

## Notes

1. See Buchanan (1975) for a discussion of the emergence of public choice as a subdiscipline in the years following World War II.
2. Rowley (1984) gives a detailed analysis of the generality, realism, and predictive power of the median voter model and concludes that the model falls short on each count. His paper gives a good overview of the research that Mueller refers to.
3. See, for example, Barlow (1970), Bergstrom and Goodman (1973), and Borcherding and Deacon (1972).
4. See, for example, Tullock's (1967) argument that cycles are unlikely in a large number setting, in a model that is a special case of Plott's (1967) majority rule equilibrium. These models are discussed at greater length in Holcombe (1985: Ch. 4).
5. See Plott and Levine (1978) for an illustration in a committee setting.
6. See Holcombe (1983) Chapters 3 and 4, for further discussion.
7. See also McKelvey (1979) and Hinich (1977). The intuition behind McKelvey's mathematical demonstration can be seen in a simpler graphical framework in Holcombe (1985), Chapter 4.
8. Riker (1980) abandoned hope of finding a stable political equilibrium analogous to general equilibrium in the market, while Tullock (1982) asked how the public sector could appear so stable in the face of these formal proofs of instability.
9. This is an example of a structure induced equilibrium, to use the terminology of Shepsle and

Weingast (1981), although they consider the notion in relation to legislative choice rather than in a general election. Slutsky (1977) also illustrates the point that if the voter choice set is limited so that it does not cover all points in the policy space, a majority rule equilibrium can exist. 10. See Holcombe (1985: 71–74) for some additional discussion along these lines.

## References

- Arrow, K.J. (1951). *Social choice and individual values*. New Haven: Yale University Press.
- Barlow, R. (1970). Efficiency aspects of local school finance. *Journal of Political Economy* 78 (5): 1028–1040.
- Bator, F.M. (1958). The anatomy of market failure. *Quarterly Journal of Economics* 72 (August): 351–379.
- Bergstrom, T., and Goodman, R. (1973). Private demand for public goods. *American Economic Review* 63 (June): 280–296.
- Borcherding, T.E., and Deacon, R.T. (1972). The demand for services of non-federal governments. *American Economic Review* 62 (December): 891–901.
- Bowen, H. (1943). The interpretation of voting in the allocation of economic resources. *Quarterly Journal of Economics* 58 (Fall): 27–48.
- Black, D. (1958). *The theory of committees and elections*. Cambridge: Cambridge University Press.
- Buchanan, J.M. (1975). Public finance and public choice. *National Tax Journal* 28 (December): 383–394.
- Downs, A. (1957). *An economic theory of democracy*. New York: Harper and Row.
- Hinich, M.J. (1977). Equilibrium in spatial voting: The median voter result is an artifact. *Journal of Economic Theory* 16 (2): 208–219.
- Holcombe, R.G. (1977). The Florida system: A Bowen equilibrium referendum process. *National Tax Journal* 30 (March): 77–84.
- Holcombe, R.G. (1980). An empirical test of the median voter model. *Economic Inquiry* 18 (2): 260–274.
- Holcombe, R.G. (1983). *Public finance and the political process*. Carbondale: Southern Illinois University Press.
- Holcombe, R.G. (1985). *An economic analysis of democracy*. Carbondale: Southern Illinois University Press.
- Inman, R.P. (1978). Testing political economy's 'as if' assumption: Is the median income voter really decisive? *Public Choice* 33 (4): 45–65.
- McEachern, W.A. (1978). Collective decision rules and local debt choice: A test of the median voter hypothesis. *National Tax Journal* 31 (2): 129–136.
- McKelvey, R.D. (1976). Intrasitivities in multi dimensional voting models and some implications for agenda control. *Journal of Economic Theory* 12 (3): 472–482.
- McKelvey, R.D. (1979). General conditions for global intrasitivities in formal voting models. *Econometrica* 47 (5): 1085–1112.
- Mueller, D.C. (1979). *Public Choice*. Cambridge: Cambridge University Press.
- Mueller, D.C. (1987). The 'Virginia School' and public choice. In R. Slesinger and G. Beeson (Eds.), *Economics 87/88*. Guilford, CT: Dushkin Publishing Group.
- Munley, V.G. (1984). Has the median voter found a ballot box he can control? *Economic Inquiry* 22 (3): 323–336.
- Niskanen, W.A. (1971). *Bureaucracy and representative government*. Chicago and New York: Aldine-Atherton.

- Plott, C.R. (1967). A notion of equilibrium and its possibility under majority rule. *American Economic Review* 62 (4): 787–806.
- Plott, C.R., and Levine, M.E. (1978). A model of agenda influence on committee decisions. *American Economic Review* 68 (1): 146–160.
- Pommerene, W.W. (1978). Institutional approaches to public expenditures: Empirical evidence from Swiss municipalities. *Journal of Public Economics* 9 (April): 163–201.
- Pommerene, W.W., and Frey, B.S. (1976). Two approaches to estimating public expenditures. *Public Finance Quarterly* 4 (4): 395–407.
- Riker, W.H. (1980). Implications from the disequilibrium of majority rule for the study of institutions. *American Political Science Review* 74 (2): 432–436.
- Romer, T., and Rosenthal, H. (1978). Political resource allocation, controlled agendas, and the status quo. *Public Choice* 33 (4): 27–43.
- Romer, T., and Rosenthal, H. (1979a). The elusive median voter. *Journal of Public Economics* 12 (2): 143–170.
- Romer, T., and Rosenthal, H. (1979b). Bureaucrats versus voters: On the political economy of resource allocation by direct democracy. *Quarterly Journal of Economics* 93 (4): 563–587.
- Romer, T., and Rosenthal, H. (1982). Median voters or budget maximizers: Evidence from school expenditure referenda. *Economic Inquiry* 26 (4): 556–578.
- Rowley, C.K. (1984). The relevance of the median voter theorem. *Journal of Institutional and Theoretical Economics* 140 (1): 104–126.
- Shepsle, K.A., and Weingast, B.R. (1981). Structure-induced equilibrium and legislative choice. *Public Choice* 37 (3): 503–519.
- Slutsky, S. (1977). A voting model for the allocation of public goods: Existence of an equilibrium. *Journal of Economic Theory* 14 (April): 299–325.
- Tullock, G. (1967). The general irrelevance of the general impossibility theorem. *Quarterly Journal of Economics* 81 (2): 256–270.
- Tullock, G. (1982). Why so much stability? *Public Choice* 37 (2): 189–202.

# DO VOTERS AFFECT OR ELECT POLICIES? EVIDENCE FROM THE U. S. HOUSE\*

DAVID S. LEE  
ENRICO MORETTI  
MATTHEW J. BUTLER

There are two fundamentally different views of the role of elections in policy formation. In one view, voters can *affect* candidates' policy choices: competition for votes induces politicians to move toward the center. In this view, elections have the effect of bringing about some degree of policy compromise. In the alternative view, voters merely *elect* policies: politicians cannot make credible promises to moderate their policies, and elections are merely a means to decide which one of two opposing policy views will be implemented. We assess which of these contrasting perspectives is more empirically relevant for the U. S. House. Focusing on elections decided by a narrow margin allows us to generate quasi-experimental estimates of the impact of a "randomized" change in electoral strength on subsequent representatives' roll-call voting records. We find that voters merely *elect* policies: the degree of electoral strength has no effect on a legislator's voting behavior. For example, a large *exogenous* increase in electoral strength for the Democratic party in a district does not result in shifting both parties' nominees to the left. Politicians' inability to credibly commit to a compromise appears to dominate any competition-induced convergence in policy.

## I. INTRODUCTION

How do voters influence government policies? An economist's answer is that they do so by compelling politicians to adopt "middle ground" platforms. Competition for votes can force even the most partisan Republicans and Democrats to moderate their policy choices. In the extreme case, competition may be so strong that it leads to "full policy convergence": opposing parties are forced to adopt identical policies [Downs 1957].<sup>1</sup> More realisti-

\* An earlier version of the paper, "Credibility and Policy Convergence: Evidence from U. S. Roll Call Voting Records," is online as NBER Working Paper No. 9315, October 2002. We thank David Card, John DiNardo, and Melvin Hinich for helpful discussions, and David Autor, Hongbin Cai, Anne Case, Dhammika Dharmapala, Justin Wolfers, and participants in workshops at the University of North Carolina—Chapel Hill, University of Texas—Austin, University of Chicago Department of Economics and Graduate School of Business, Princeton University, University of California—Los Angeles, Stanford University Graduate School of Business, University of California—Davis, University of California—Riverside, and University of California—Irvine, and two anonymous referees for comments and suggestions. We also thank James Snyder and Michael Ting for providing data for an earlier draft. This material is based upon work supported by the National Science Foundation under Grant No. 0214351.

1. Empirical studies indicate that Republican and Democratic legislators vote very differently, even when they share the same constituency. For example,

cally, competition leads to “partial policy convergence”: candidates do pursue more moderate policies, even if they are not forced to adopt identical platforms [Wittman 1983; Calvert 1985]. This less rigid and arguably more realistic understanding of Downs’ insight has become central to how economists think about political competition. Indeed, the so-called “Downsian paradigm” has remained the backbone of many models in political economy.

There is, however, a growing recognition of a serious shortcoming of this paradigm. In a recent survey of the literature, Besley and Case [2003] emphasize that the assumptions about politicians’ commitment and motivation in the Downsian paradigm “are unreasonable and outcomes are highly unrobust to deviations from them.” Downsian convergence depends on the assumption that elected politicians always implement the policies that they promised as candidates. But Alesina [1988] shows that when partisan politicians cannot credibly promise to implement more moderate policies, the result can be full policy *divergence*: the winning candidate, after obtaining office, simply pursues his most-preferred policy. In this case, voters fail to compel candidates to reach any kind of policy compromise.

What emerges, then, are two fundamentally different views of the role of elections in a representative democracy. On the one hand, when electoral promises are credible—as in a Downsian partial convergence—candidates seek middle ground policies, and general elections bring about some degree of policy “compromise.” On the other hand, when promises to enact moderate policies are not credible—as in full policy divergence—general elections are merely a means to decide which candidate’s preferred policy will be implemented. Which of these two competing views is empirically more relevant? This paper assesses the relative importance of the two contrasting perspectives in explaining how Representatives vote in the U. S. House.

As is apparent from Alesina’s [1988] analysis of the role of credibility, the two broad views have sharply different predictions for how a politician’s electoral strength influences her policy choices. When politicians have incentives to moderate their platforms—as in partial policy convergence—the relative electoral

---

Poole and Rosenthal [1984] show that senators from the same state but from different political parties have different voting records. This is inconsistent with Downs’ original model, in which candidates adopt identical positions—“complete policy convergence.” See also Snyder and Groseclose [2000] and Levitt [1996].

strength of the two parties matters. More specifically, when electoral support is high, a candidate can afford to vote in a relatively more partisan way if he is elected; a weaker candidate would be forced to choose a more moderate policy. An increase in electoral strength for the Democratic party in a district, for example, would cause both parties' nominees to shift to the "left." On the other hand, when voters do not believe promises of policy compromises—as in full policy divergence—the relative electoral strength of the two candidates is irrelevant, as politicians simply pursue their own personal policy views. That is, an increase in electoral strength for the Democratic party in a district leaves legislators' actions unchanged.

Therefore, an assessment of the relative importance of the two views requires estimating the effect of a candidate's electoral strength on subsequent roll-call voting records. To do so, we consider electoral races where a Democrat holds the seat—and hence an electoral advantage—and measure the roll-call voting records of the winners of these elections. We measure the extent to which they are more liberal than the voting records of winners of elections where the Republican had held the seat; i.e., where the Democrat was relatively weaker.<sup>2</sup> Of course, which of the two parties holds a district seat—and hence the electoral advantage—is clearly endogenously determined, influenced by the political leanings of the voters, the quality of candidates, resources available to the campaigns, and other unmeasured characteristics of the district and the candidates. A naive comparison that does not account for these differences between Democratic and Republican districts is likely to yield biased estimates. What is needed is an *exogenous* variation in who holds the seat—and hence greater electoral strength—in order to measure how politicians' actions respond to the odds of winning an election.

To isolate such exogenous variation, we exploit a quasi experiment embedded in the Congressional electoral system that generates essentially "random assignment" of which party holds a seat—and therefore which party holds the electoral advantage. In particular, we focus our analysis on the set of electoral races in

2. Our empirical strategy obviously accounts for the fact that Democrats are more liberal than Republican. That is, the roll-call behavior of a winner of an electoral race where a Democrat held the seat will tend to be more liberal simply because—due to the advantage of incumbency—the winner will more likely be a Democrat. This fact in itself would cause a difference in voting records, even if Representatives ignored electoral pressures and simply voted their own ideological position. It is easy to account for this factor—as we will show in Section II.

which the incumbent party had *barely* won the *previous* election (say by 0.01 percent of the vote). The key identifying assumption is that districts where the Democrats barely won are comparable—in all other ways—to districts where the Republicans barely won. We present empirical evidence that strongly supports this assumption: Democratic and Republican districts are in general very different, but among close elections, they are similar in every characteristic that we examine, including various demographic characteristics of the population, racial composition, size of the district, income levels, and geographical location. Our quasi experiment, then, addresses the endogeneity problem by isolating arguably independent and exogenous variation in candidates' electoral strength across Congressional districts.

Using this regression-discontinuity design and voting record data from the U. S. House (1946–1995), we find that the degree of electoral strength has no effect on a legislator's voting behavior.<sup>3</sup> Candidates with weak electoral support do not adopt more moderate positions than do stronger candidates, holding other factors constant. For example, a large exogenous increase in electoral strength for the Democratic party in a district does not result in shifting both parties' nominees to the left. This suggests that voters seem not to *affect* politicians' choices during general elections; instead, they appear to merely *elect* policies through choosing a legislator. That is, they do not *influence* policy through their Representatives' choices as much as they are implicitly *presented with* policy choices by different candidates.<sup>4</sup>

Our findings are consistent with the inability of opposing candidates to credibly commit to a policy compromise. It appears that the central prediction of the Downsian paradigm—that individual politicians' policy choices are constrained by voters' sentiments—has little empirical support, at least in the context of U. S. House general elections. Our findings provide some empirical justification for the notion that candidates confront a credi-

3. The voting score data include (1) the actual roll-call data from ICPSR Study 4 "United States Congressional Roll Call Voting Records," (2) Groseclose, Levitt, and Snyder's [1999] Inflation-Adjusted ADA Scores [Groseclose 2002], (3) McCarty, Poole, and Rosenthal's [1997] DW-NOMINATE scores [Poole 2002a], (4) Poole and Daniels' [1985] Interest Group Ratings [Poole 2002b], and (5) Poole's [1999] rank order data [Poole 2002c].

4. This leaves open the question of how candidates are selected. There are several models where candidates are endogenous. (See, for example, Persson and Tabellini [2000] for an introduction to this literature.) In this paper we take the candidates' ideologies as exogenous. We return to this point below.

bility problem. This notion has been explicitly adopted in recent theoretical analyses [Besley and Coate 1997, 1998].

It is important to recognize that our findings say little about whether members of the U. S. House generally represent their “constituencies.” Instead, our analysis focuses on the role of *general elections* in inducing candidates with different policy stances to move toward the center. Although we find a small effect of the pressures of a general election on candidates, this does not imply that election outcomes do not “represent” the desires of the electorate. First, and most obviously, voters still do choose between the two available policy platforms. Second, “representativeness” does not necessarily occur only through general elections. Pre-election channels (primary elections, for example) may also be important in inducing representativeness. Indeed, within each district, the Republican and Democratic nominees may, respectively, represent the “median” Republican and “median” Democratic voter.

The paper is organized as follows. Sections II and III provide background and motivation for our analysis, and describe our empirical strategy, first informally, and then within a formal conceptual framework. Section IV describes the context and the data, and Section V presents our empirical results. We relate our findings to the existing literature in Section VI. Section VII concludes.

## II. BACKGROUND AND CONCEPTUAL FRAMEWORK

### II.A. *Role of Credibility in Political Competition*

Voters can influence policy in two distinct ways. Competing political candidates have incentives to adopt positions that reflect the preferences of the electorate because doing so raises the chances they will win the election. That is, voters can *affect* the policy choices of politicians. Alternatively, voters always impact policy outcomes by selecting a leader among several candidates, who each may have already decided on a particular policy based on other reasons. In this way, voters may simply *elect* policies. Whether voters affect or elect policies depends on whether or not candidates are able to make credible promises to implement moderate policies.

A large class of models of political competition assumes that they can. The most well-known example is the simple “median

voter” model of political competition [Downs 1957]. Two candidates, who care only about winning office, compete for votes by taking a stance in a single dimensional policy space. Voters cast their vote based on these positions, and the equilibrium result is that the politicians carry out identical policies—the one most preferred by the “median voter.” In this extreme example, voters have a powerful *effect* on politicians’ choices, to the point where it is irrelevant *which* of the two candidates is ultimately elected.

A similar outcome results when opposing candidates care not only about winning the election, but also about the implemented policies themselves. Opposing parties may not choose *identical* positions, but in general electoral competition will compel them to choose policies more moderate than their most preferred choices [Wittman 1983; Calvert 1985]. The basic insight that voters *affect* candidates’ positions by inducing spatial competition is robust to various generalizations of the simple model utilized by Downs [Osborne 1995].

But it is much less robust to the assumption that candidates can commit to policy pronouncements, as emphasized in Besley and Case [2003]. When politicians have ideological preferences over policy outcomes, credibility becomes an issue. Specifically, Alesina [1988] points out that Downs’ equilibrium may fall apart if parties care about policies and there is no way to make binding precommitments to announced policies. After winning the election, what incentive does a legislator have to keep a promise of a more moderate policy? In a one-shot game, the only time-consistent equilibrium is that candidates carry out their *ex post* most-preferred policy. Electoral pressures do not at all compel opposing candidates to moderate their positions. Voters’ only role in affecting policy outcomes is to *elect* a politician, whose policy position is unaffected by electoral pressures.

In a repeated election framework, both policy convergence and divergence are possible, as politicians can establish credibility through building reputations. If voters and opposing parties believe that there are sufficiently high costs to deviating from moderate promises, it is possible to achieve some degree of *policy convergence* [Alesina 1988]. Voters *affect* policies because of candidates’ incentives to maintain a reputation. But if both parties and voters do not expect any compromise, the *fully divergent* outcome occurs in every election. Candi-

dates do not deviate from their ex post most-preferred policy, and voters only *elect* policies.<sup>5</sup>

The goal of this paper is to examine which phenomenon is more empirically relevant for describing roll-call voting patterns of U. S. House Representatives. Does the expectation of how voters will cast their ballot *affect* how legislators vote, or do voters simply *elect* a legislator among candidates with fixed policy positions? The answer to this question has important implications for understanding and modeling policy formation in a representative democracy.

If voters primarily affect politicians' decisions, then "centripetal" political forces generated by the broader voting population would largely outweigh any "centrifugal" forces that pull candidates' positions apart (e.g., party discipline, special interest groups). It would also imply that candidates are able to convince voters that they will compromise on policy, through the building of reputations or other mechanisms. The Downsian paradigm would then seem to be a reasonable, first-order description of policy formation as it relates to U. S. House elections.

On the other hand, if voters primarily elect policies, then "centrifugal" forces largely would dominate any Downsian convergence. It would then become more important, for example, to understand how a nominee, and the policies that she supports, is chosen by the party: primary elections could be more influential than general elections for policy formation. It would also provide an empirical basis for assuming that candidates face a serious credibility problem in their policy pronouncements. There is a growing recognition of the inadequacy of the Downsian paradigm on this point [Besley and Case 2003].

Existing studies have established that, controlling for constituency characteristics, Democratic representatives possess more liberal voting records than Republican members of Congress.<sup>6</sup> This constitutes strong evidence against the extreme case of *complete* policy convergence (e.g., the median voter theorem), but is too stringent a test of the more general notion of Downsian electoral competition. Therefore, to measure the relative impor-

5. It is also true that even if discount rates are sufficiently low, the fully divergent outcome still remains a subgame perfect equilibrium of the repeated election game.

6. The full convergence hypothesis has been tested, and rejected by many authors. For example, Poole and Rosenthal [1984] show that senators from the same state but from different political parties have different voting records. For a discussion of empirical regularities in the literature, see Snyder and Ting [2001a].

tance of competition-induced convergence, it is necessary to empirically distinguish between *partial convergence*, where voters affect politicians' policy choices—despite the undeniable party effect—and *complete divergence*, where voters merely elect policies. This is the goal of our study.

### *II.B. Identification Strategy*

We now describe the main difficulties of addressing this question, and how we confront them with our identification strategy. Here we will intentionally be less formal, in order to provide the intuition of our approach. A more rigorous exposition of our conceptual and econometric framework is presented in the next section. Throughout the discussion we assume a two-party political system.

The most straightforward way to determine whether voters primarily affect or elect policy choices is to simply compare candidates' most-preferred policies (hereinafter "bliss points") and the policies they would actually choose. If the voting records were more moderate than their bliss points, this would indicate that the expected voting behavior of the electorate factored into the candidates' decisions. If there were no difference between their choices and their bliss points, this would imply that voters merely influence the relative odds of which of the two candidates' policies is "elected." Unfortunately, such a comparison is impossible, since there are no reliable measures of candidates' bliss points.

In this paper we utilize a simple empirical test of whether voters primarily affect or elect policy choices, based on how Representatives' roll-call voting behavior is affected by exogenous changes in their electoral strength. The test is based on the predictions of Alesina's [1988] model of electoral competition. In the next section we formally develop the idea, but the intuition is very simple. If candidates are constrained by their constituents' preferences, we should observe that exogenous changes in their electoral strength have an impact on how they intend to vote if elected to Congress. On the other hand, if promises to adopt moderate policies are noncredible, then the electoral strength of a candidate should be irrelevant to how (s)he intends to vote.

Throughout the paper we use the following notation for the timing of elections.  $t$  and  $t + 1$  represent separate electoral cycles. For example, when  $t = 1992$ , it includes the 1992 campaign, the November 1992 election, and the 1993–1994 Congressional session. Similarly,  $t + 1$  would include the 1994 campaign,

the November 1994 election, and the 1995–1996 Congressional session.

Our strategy is based on the following thought experiment. Imagine that we could decide the outcome of Congressional electoral races in, say, 1992 with the flip of a coin (but we allow all subsequent elections to be determined in the usual way). This initial randomization guarantees that the group of districts where the Democrat won would be, in all other respects, similar to the newly Republican districts. For example, the two groups of districts would be similar in the ideological positions of the voters and candidates, the demographic characteristics, the resources that were available to the candidates, and so forth.

Because incumbents are known to possess an electoral advantage, the outcome of the 1992 race would impact what happens in the 1994 election. Democrats are likely to be in a relatively stronger electoral position where they are incumbents, and similarly for Republicans. The key point is that the random assignment of who wins in 1992 essentially generates random assignment in which party's nominee has greater electoral strength for the 1994 election. We could use this change in electoral strength to test the hypothesis of complete divergence against the alternative of partial convergence.

Specifically, we could examine the 1995–1996 voting “scores” of the winners of the 1994 elections where the Democrats had held the seat during the 1994 campaign, and compare them with the scores of winners of elections where a Republican held the seat. This difference would represent a valid causal effect of who holds the seat during the 1994 electoral races on 1995–1996 voting records. We call this the “overall effect,” and it is the sum of two components.

The first component would reflect that the 1995–1996 voting scores of the winners where a Democrat held the seat during the 1994 electoral race will tend to be more liberal simply because—due to the electoral advantage of holding the seat—the winner will more likely be a Democrat. And as we know, Democrats have more liberal voting scores. This first component reflects how voters *elect* policies: how they impact policy by simply altering the relative odds of which party's nominee is chosen. As we show more formally in the next section, this component can be directly estimated by answering the questions how much more likely is the winner to be a Democrat if the seat is already held by a

Democrat and what is the expected difference between how Republicans and Democrats vote, other things constant.

The remaining, second component would reflect how candidates might respond to an exogenous increase or decrease in the probability of winning the election in 1994. If legislators are pressured to keep their election promises, then a Democrat who is challenging an incumbent Republican in 1994 would be expected to have less liberal voting records in 1995–1996 (if elected) compared with an incumbent Democrat. After all, the challenger would be in a much weaker electoral position than the incumbent. This second component reflects how expected voting behavior *affects* the policy choices of candidates. It is computed by subtracting the first component from the overall effect.

The relative magnitudes of the two components indicate which equilibrium—full divergence or partial convergence—is relatively more important. If the “elect” component is dominant, it suggests full policy divergence: politicians simply vote their own policy views, unaffected by electoral pressures. If the “affect” component is important, it suggests partial policy convergence: policy choices are constrained by electoral pressure imparted by voters.

What allows us to perform this decomposition into the two components? The initial “random assignment” of who wins the 1992 election does. Without the random assignment, it would be difficult to distinguish between any of these effects and differences due to spurious reasons. After all, in the real world, the party that holds a district seat—and the electoral advantage—is clearly endogenously determined, influenced by the ideologies of the voters and candidates, and other unmeasured characteristics of the districts. A naive comparison that does not account for all these unobservable differences between Democratic and Republican districts is likely to yield biased estimates.

For example, Democratic legislators will have more liberal voting scores than Republicans (for simplicity, consider the period of the 1990s). But Democrats are also more likely to be elected in places like Massachusetts and than in places like Alabama. So it is not clear how much of this voting gap reflects the typical difference between Republican and Democratic nominees and how much of the gap reflects the typical difference between Representatives from Massachusetts and Alabama.

How do we generate the initial “flip of the coin” decision of who wins the 1992 election? We use a quasi experiment that is

embedded in the Congressional electoral system. Specifically, our empirical strategy focuses on elections that were decided by a very narrow margin in 1992, as revealed by the final vote tally. For example, we begin by examining elections that were decided by less than a 2 percent vote share. We argue that among these elections, it is virtually random which of the two parties won the seat [Lee 2003]. For the sake of exposition, we defer to a later section the discussion of why we believe this to be true, and the description of the empirical evidence that strongly supports this assumption. We have used 1992 and 1994 in this explanation of our empirical strategy. In practice, in our empirical analysis we use data for the period 1946–1995.

### III. THEORETICAL AND ECONOMETRIC FRAMEWORK

In this section we 1) formally define what it means to ask the question of whether voters primarily affect or elect policies, and 2) explain how our empirical strategy is able to distinguish between these two phenomena.

#### *III.A. Model*

We utilize the repeated election framework of Alesina [1988], adopting that study's modeling conventions and notation. Consider two parties,  $D$  (Democrats) and  $R$  (Republicans), in a particular Congressional district. The policy space is unidimensional, where party  $D$ 's and  $R$ 's per-period policy preferences are represented by quadratic loss functions,  $u(l) = -(1/2)(l - c)^2$  and  $v(l) = -(1/2)l^2$ , respectively, where  $l$  is the policy variable and  $c(>0)$ , and 0, are their respective bliss points. As in Alesina, the analysis makes no distinction between the "party" and an individual nominee, so that the "electoral strength" of the party in a district is equivalent to the "electoral strength" of the party's nominee in that district, during the election. Also, candidates'/ parties' bliss points are assumed to be exogenously determined.<sup>7</sup>

The timing of elections is as follows. Before election  $t$ , voters

7. This framework has little to say on the question of how candidates are selected. Alternative frameworks are possible and may generate different predictions. For example, the models proposed by Bernhardt and Ingberman [1985] and Banks and Kiewiet [1989] are quite different in spirit from the model used here. In those models, the challenger is at disadvantage because she cannot adopt the incumbent's position and is therefore forced to take a more extreme position. In equilibrium the low probability of defeating incumbent members of Congress deters potentially strong rivals from challenging them [Banks and Kiewiet 1989].

form expectations of the parties' policies, denoted  $x^e$  and  $y^e$ . At this point, the outcome of the election is uncertain to all agents in the model, with the probability of party  $D$  winning being  $P$ , which is "common knowledge."  $P(x^e, y^e)$  is a function of  $x^e$  and  $y^e$ , and by assumption, when  $x^e > y^e$ , then  $\partial P/\partial x^e, \partial P/\partial y^e < 0$ ; that is, more votes can be gained by moderating the policy position. If party  $D$  wins the election,  $x$  is implemented, and if party  $R$  wins,  $y$  is implemented. A rational expectations equilibrium is assumed throughout;  $x = x^e$ , and  $y = y^e$ . The game then repeats for period  $t + 1$ . Note that period  $t$  includes both the election and the subsequent Congressional session, and similarly for  $t + 1$ . For example, if  $t = 1992$ ,  $t$  refers to the November 1992 election and the roll-call votes  $RC_t$  in the 1993–1994 Congressional session;  $t + 1$  refers to the November 1994 election and the roll-call votes  $RC_{t+1}$  in the 1995–1996 session.

Alesina [1988] shows that the efficient frontier is given by  $x^* = y^* = \lambda c$ , where  $\lambda \in (0,1)$ . Because of the concavity preferences, both parties prefer a moderate policy with certainty to a fair bet. Three Nash equilibria are possible.

(a) Complete Convergence:  $x^* = y^* = \lambda^* c$ .

In this equilibrium, opposing parties agree to a moderate policy, by Nash bargaining on the efficient frontier. The "Folk Theorem" equilibrium is one where both parties "announce" the same, moderate policy, and the voters expect the moderate outcome, but as soon as a party deviates from the announced position, reputation is lost, and the game reverts to the uncooperative outcome,  $y^* = 0, x^* = c$ . As long as discount rates are sufficiently low, promises to adopt policy compromises are credible.

For our purposes, the key result is that  $dx^*/dP^* = dy^*/dP^* = (d\lambda^*/dP^*)c > 0$ , where  $P^*$  represents the underlying "popularity" of party  $D$ : the probability that party  $D$  would win at fixed policy positions,  $x^e = c$  and  $y^e = 0$ .<sup>8</sup> An increase in  $P^*$  represents an *exogenous* increase in the popularity of party  $D$ , which would boost party  $D$ 's "bargaining power" so that the equilibrium moves closer to her bliss point. This exogenous increase comes about from a "helicopter drop" of Democrats in the district, or campaign resources, or the advantage that comes from being the incumbent in the district. In this equilibrium, policy choices are implicitly

8.  $\lambda$  is used to characterize the entire efficient frontier.  $\lambda^*$ , on the other hand, denotes the Nash bargaining equilibrium.

constrained by voters. Thus, when  $dx^*/dP^*, dy^*/dP^* > 0$ , we say that voters *affect* candidates' policy choices.

Indeed, in this equilibrium—similar to Downs' original “median voter” model—voters exclusively affect policy choices, and do not elect policies at all: it is irrelevant for policy which party is actually elected.

(b) Partial Convergence:  $0 \leq y^* \leq x^* \leq c$ .

Is the result that voters *affect* policies— $dx^*/dP^*, dy^*/dP^* > 0$ —robust to minor deviations from the complete convergence equilibrium? We show that it is. This agrees with our intuition that voters can induce policy compromise, even if they cannot force them to adopt identical positions. It also agrees with our intuition that a rejection of complete convergence says little about the relative degree to which voters affect or elect policies. Rejecting complete convergence simply implies that  $y^* < x^*$ , but nothing about whether  $0 < y^*$  or  $x^* < c$ .

It is possible to extend Alesina's model to allow for parties to care about winning the seat, *per se*, in addition to caring about the policy outcome.<sup>9</sup> The result is that in general,  $0 \leq y^* \leq x^* \leq c$ , because there are values where  $x = y$  is not Pareto efficient. Both parties can be made better off by one party moving closer to its bliss point, because there is an explicit benefit to obtaining office. A detailed proof is available on request.

The important point, for our purposes, is that the comparative static  $dx^*/dP^*, dy^*/dP^* > 0$  is robust to this logical extension to the model. With an exogenously higher  $P^*$ , party  $D$  has a better “bargaining position” and therefore can compel the parties to agree on a position closer to party  $D$ 's bliss point.

(c) Complete Divergence:  $x^* = c, y^* = 0$ .

In this equilibrium, voters expect nothing else than the parties to carry out their bliss points if elected, and the parties do just this. This can arise if promises to implement policy compromises are not credible. In this case, an increase in  $P^*$  now does *nothing* to the equilibrium:  $dx^*/dP^* = dy^*/dP^* = 0$ . This is a “corner solution,” whereby an exogenous shock to  $P^*$  has no effect on candidates positions. Here, voters merely *elect* politicians' fixed policies.

9. Our extension should not be confused with that of Alesina and Spear [1988], in which parties agree to split the benefits of office. In our extension, they cannot split the benefits of office. This case should also not be confused with the partial convergent equilibria that can arise if discount rates are too low to support fully convergent equilibria. Alesina [1988] proves existence of these equilibria.

Among the above three equilibria, the full convergence equilibrium is not very realistic, and has already been empirically rejected by several authors. But a rejection of full convergence says little about whether politicians' behaviors are better characterized by partial convergence (voters can affect policy outcomes) or complete policy divergence (voters only elect policies). Distinguishing between these two equilibria is our goal. For this purpose, the key result of the theoretical framework is that differentiating between partial and complete divergence is equivalent to assessing whether  $dx^*/dP^*$ ,  $dy^*/dP^* > 0$  or  $dx^*/dP^*$ ,  $dy^*/dP^* = 0$ .

We assume that voters are forward-looking and have rational expectations. This implies that voting records  $RC_{t+1}$ —roll-call votes *after* the election—are on average equal to voters' expectations. It is important to note that this is *not* the same as assuming that candidates can make binding precommitments. Politicians always have the option of not carrying out their preelection policy pronouncements. But in Alesina's repeated game equilibrium, candidates do carry out their "announced" policies because of the need to maintain a reputation.<sup>10</sup>

### *III.B. Estimating Framework*

The above framework directly leads to our empirical strategy. Note first that the roll-call voting record  $RC_t$  of the representative in the district following the election  $t$  can be written as

$$(1) \quad RC_t = (1 - D_t)y_t + D_t x_t,$$

where  $D_t$  is the indicator variable for whether the Democrat won election  $t$ . A similar equation applies for  $RC_{t+1}$ . Simply put, only the winning candidate's intended policy is ultimately observable. In Appendix 1 we provide conditions under which the above expression can be transformed into

$$(2) \quad RC_t = \text{constant} + \pi_0 P_t^* + \pi_1 D_t + \varepsilon_t$$

$$(3) \quad RC_{t+1} = \text{constant} + \pi_0 P_{t+1}^* + \pi_1 D_{t+1} + \varepsilon_{t+1},$$

where  $P^*$  is the measure of the electoral strength of party  $D$ —the probability of a party  $D$  victory at fixed platforms  $c$  and 0—and  $\varepsilon$  reflects heterogeneity in bliss points across districts. This equa-

10. Of course, the equilibrium depends on candidates not discounting the future too much.

tion simply parameterizes the derivatives  $dx^*/dP^*$ ,  $dy^*/dP^*$  as  $\pi_0$ . It also allows an independent effect of party,  $\pi_1$ , which is reasonable given the existing evidence that party affiliation is an important determinant of roll-call voting records. In this equation, partial convergence (voters *affect* policy choices) implies that  $\pi_0 > 0$ . Full divergence (voters only *elect* policies) implies that  $\pi_0 = 0$ .

In general, we cannot observe  $P^*$ , so equation (2) cannot be directly estimated by OLS. But suppose that one could randomize  $D_t$ . Then  $D_t$  would be independent of  $\varepsilon_t$  and  $P_t^*$ . Also, if bliss points are exogenous—and hence are not influenced by who won the previous election—then  $D_t$  will have no impact on  $\varepsilon_{t+1}$ . It follows that

$$(4) \quad E[RC_{t+1}|D_t = 1] - E[RC_{t+1}|D_t = 0] = \pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}] + \pi_1[P_{t+1}^D - P_{t+1}^R] = \gamma$$

$$(5) \quad E[RC_t|D_t = 1] - E[RC_t|D_t = 0] = \pi_1$$

$$(6) \quad E[D_{t+1}|D_t = 1] - E[D_{t+1}|D_t = 0] = P_{t+1}^D - P_{t+1}^R,$$

where  $D$  and  $R$  superscripts denote which party held the seat—and hence held the electoral advantage. For example,  $P_{t+1}^D$  denotes the *equilibrium* probability of a Democrat victory in  $t + 1$  given that a Democrat held the seat during the campaign of  $t + 1$ ;  $P_{t+1}^{*R}$  represents the “electoral strength” of the Democrat during the campaign of  $t + 1$ , given that a Republican held the seat. Note that while we cannot estimate  $P_{t+1}^{*D}$  and  $P_{t+1}^{*R}$ , we can estimate the  $P_{t+1}^D$  and  $P_{t+1}^R$  from the data.<sup>11</sup>

These three equations form the basis of our empirical analysis. Equation (4) shows that the total effect  $\gamma$  of a Democratic victory in  $t$  on voting records  $RC_{t+1}$  is the sum of two components,  $\pi_1[P_{t+1}^D - P_{t+1}^R]$ , and the remainder,  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ . The first term is the “elect” component. The second term is the “affect” component. The equation shows that the overall effect  $\gamma$  can be estimated by the simple difference in voting scores  $RC_{t+1}$  between districts won by Democrats and Republicans in  $t$ .

The next two equations show how to estimate the “elect”

11. It is important to distinguish between  $P^*$  and  $P$ .  $P^*$  is a measure of the underlying “popularity” of a party, the probability that party  $D$  will win if parties  $D$  and  $R$  are expected to choose  $c$  and  $0$ , respectively. A change in  $P^*$  represents an exogenous change in popularity. On the other hand,  $P$  is the probability that party  $D$  will win, at whatever policies the parties are expected to choose.

component, which is the product of  $\pi_1$  and  $[P_{t+1}^D - P_{t+1}^R]$ .  $\pi_1$  is estimated by the difference in voting records  $RC_t$ .<sup>12</sup>  $P_{t+1}^D - P_{t+1}^R$  is estimated by the difference in the fraction of districts won by Democrats in  $t + 1$ .

The “affect” component,  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ , can be estimated by  $\gamma - \pi_1[P_{t+1}^D - P_{t+1}^R]$ . If voters merely “elect” policies (complete divergence), we should observe little change in the candidates’ intended policies following an exogenous increase in the probability of victory; that is,  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$  should be small. If voters not only choose politicians, but also affect their policy choices (partial convergence), candidates should move toward their bliss points in response to an exogenously higher probability of winning; that is,  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$  should be relatively large. This simple decomposition allows us to make quantitative statements about the relative importance of the “affect” and “elect” phenomena. We can compute what fraction of the total effect  $\gamma$  is explained by the “elect component”  $\pi_1[P_{t+1}^D - P_{t+1}^R]$ , and what fraction by the “affect component”  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ .

Note that the initial “random assignment” of  $D_t$  is crucial here. Without this, the estimated differences above would in general be biased for the quantities  $\gamma$ , and  $\pi_1$ .<sup>13</sup> As an example, without this random assignment, the simple difference in how Republicans and Democrats vote after election  $t$  would reflect both  $\pi_1$ , and that candidates are likely to have more liberal “bliss points” where Democrats hold the seat.

We argue that the examination of suitably “close” elections in period  $t$  isolates “as good as random assignment” in  $D_t$ . As would be expected from a valid regression-discontinuity design, among

12. As will be evident below, in principle, one could obtain an alternative estimate of  $\pi_1$ , by examining the difference in records  $RC_{t+1}$  among close elections in time  $t + 1$ . In practice, however, this makes little difference because we are pooling data from many years (i.e., the difference between estimating  $\pi_1$  from data 1946–1994 and estimating it using data from 1948–1994).

13. More formally, without random assignment of  $D_t$ , the three expressions would become

$$\begin{aligned} E[RC_{t+1}|D_t = 1] - E[RC_{t+1}|D_t = 0] &= \gamma + E[\varepsilon_{t+1}|D_t = 1] - E[\varepsilon_{t+1}|D_t = 0] \\ E[RC_t|D_t = 1] - E[RC_t|D_t = 0] &= \pi_0(E[P_t^*|D_t = 1] - E[P_t^*|D_t = 0]) \\ &\quad + \pi_1 + E[\varepsilon_t|D_t = 1] - E[\varepsilon_t|D_t = 0] \\ E[D_{t+1}|D_t = 1] - E[D_{t+1}|D_t = 0] &= P_{t+1}^D - P_{t+1}^R. \end{aligned}$$

It is clear, from the expressions above, that without random assignment, the parameter estimates  $\gamma$  and  $\pi_1$  would be biased.

elections decided by a very narrow margin, as long as there is some unpredictable component of the ultimate vote tally, who wins the election will be mostly determined by pure chance (e.g., unpredictable components of voter turnout on election day). This is shown more formally in Appendix 1.

By now it should be clear why it is necessary to examine the impact of who wins in  $t$  on  $RC_{t+1}$ , roll-call votes in period  $t + 1$ . The impact of who wins in  $t$  on  $RC_t$ , roll-call votes in  $t$ , only yields  $\pi_1$ . By estimating  $\pi_1$  alone, it is only possible to test complete convergence  $\pi_1 = 0$ . This extreme hypothesis has already been tested by several studies.<sup>14</sup> But  $\pi_1$  alone is not sufficient to say anything about the size of the “elect” phenomenon relative to the “affect” phenomenon. It is not sufficient for testing full divergence against partial convergence, and hence it is not sufficient for evaluating the Downsian perspective versus the alternative view that politicians face difficulties in credibly committing to policy compromises.

#### IV. ROLL-CALL VOTING RECORDS IN THE U. S. HOUSE

##### IV.A. Context

There are several reasons why the U. S. House of Representatives provides an ideal setting in which to empirically assess whether voters primarily affect or elect policies. First, the U. S. federal legislative body is virtually a two-party system, and policy convergence is frequently modeled in a two-party context. When there are more than two candidates, the basic insight of Downs' [1957] approach to policy convergence arguments becomes more complicated (see Osborne [1995]).

Second, it is well-known that Democrats and Republicans have different (and often directly opposing) policy positions. It is meaningful to ask whether electoral competition compels opposing parties' nominees to moderate their positions in the face of strong incentives to vote along party lines. If the U. S. House were a relatively nonpartisan environment (with “bliss points” relatively close together), the distinction between voters affecting or electing policies would be less important, and a test to distinguish between them less useful.

14. See, for example, Poole and Rosenthal [1984], Levitt [1996], and Snyder and Groseclose [2000].

Third, the U. S. House is arguably the most likely setting in which to observe policy convergence, if establishing reputations is important. U. S. House elections are held every two years, and there are no term limits (as opposed to gubernatorial and presidential elections), meaning that political careers can consist of several terms in office. Furthermore, political tenure in the House is often a stepping-stone to participating in electoral races for higher offices. For these reasons, it is plausible that candidates for the U. S. House have high discount factors, which would allow reputation to support convergent equilibria.

Finally, our empirical analysis focuses on Representatives' voting records. These votes are directly observable, and are part of the public record. In principle, voters can compare a legislator's record with their platforms and promises as candidates (and opponents can advertise any deviations during election campaigns). Convergent equilibria of the kind described in Alesina [1988] require that policy positions are perfectly observable by voters and that it can be determined whether politicians deviate from policy pronouncements.

#### *IV.B. Data Description*

We now discuss the choice of the dependent variable.<sup>15</sup> There are several alternative ways to measure Representatives' voting on legislation. A widely used measure is a voting score provided by the liberal political organization, Americans for Democratic Action (ADA). For each Congress, the ADA chooses about twenty high-profile roll-call votes, and creates an index that varies between 0 and 100 for each Representative of the House. Higher scores correspond to a more "liberal" voting record. Throughout the paper our preferred voting record index is the ADA score. Later, we show that our results are robust to many alternative interest groups scores and other voting record indices.

We utilize data on ADA scores for all Representatives in the U. S. House from 1946–1995, linked to election returns data during that period.<sup>16</sup> There is considerable variation in ADA scores *within* each party. For example, the distribution of ADA scores for Democrat and Republican Representatives in the three

15. All the data and the programs used in this paper are available at <http://www.econ.ucla.edu/moretti/papers.html>

16. To make the comparison across Congresses possible, we follow the literature and use "adjusted" ADA scores throughout the paper. This adjustment to the nominal ADA score, was devised by Groseclose, Levitt, and Snyder [1999].

most recent Congresses shows significant overlapping between the parties. It is not uncommon for Democrat representatives to vote more conservatively than Republican candidates, and vice versa.

One advantage of using ADA scores is that it is a widely used index in the literature. However, one limitation is that it includes only twenty votes per Congress, and the choice of what issues to include and what weight to assign to each issue is necessarily arbitrary. To assess how robust our results are to alternative measures of “liberalness” of roll-call votes, we have reestimated all our models using three alternative sets of voting record measures.

First, we use the DW-NOMINATE scores constructed by McCarty, Poole, and Rosenthal [1997]. Poole and Rosenthal [1985] developed the NOMINATE procedure to estimate a low-dimensional measure of political ideology in a complex multidimensional political world. NOMINATE is an attempt to estimate the *underlying* ideology that drives observed roll-call behavior by assigning legislators the ideological points that maximize the number of correctly predicted roll-call votes. The NOMINATE data have the advantage of including all roll-call votes, not an arbitrary subset of votes. It also ignores the Representative’s political party and the legislative issue in question, so it is arguably more exogenous than the ADA scores.<sup>17</sup>

Second, for each member and each Congress, we construct our own measure of loyalty to the party leadership using the individual vote tallies on every issue voted on in the House. For

17. Poole and Rosenthal [1997] note that a single dimension would be unlikely to capture the division between Northern and Southern Democrats during the Civil Rights Era. Therefore, the NOMINATE procedure estimates a two-dimensional measure of ideology where the first dimension captures party loyalty and can be thought of as a liberal to conservative scale, and the second dimension captures the issues of race that divided the Democrats until the mid-1970s. To remain consistent with our discussion of a single ideological dimension, we restrict our analysis to the first dimension during the period where the second dimension had little predictive power. Specifically, we restrict our DW-NOMINATE analysis to 1975 and beyond. However, we have reestimated our models including DW-NOMINATE data for the entire 1946–1995 period and obtained very similar results. For completeness, we have also reestimated our ADA models including only data for the 1975–1995 period and obtained very similar results. We use the DW-NOMINATE scores as opposed to the Poole and Rosenthal’s [1991] earlier D-NOMINATE scores because the DW data cover up through the 106th Congress while the D-NOMINATE data ends with the 99th Congress. McCarty, Poole, and Rosenthal [1997] note that the D-NOMINATE and DW-NOMINATE scores are highly correlated where both scores are available. See Poole and Rosenthal [1997] for a description of the NOMINATE procedure. Poole’s [1999] rank order data yield similar results.

this measure, we calculate the percent of a representative's votes that agree with the Democrat party leader.<sup>18</sup>

Third, we use ratings from interest groups other than the ADA. We include both liberal and conservative ratings from groups such as the American Civil Liberties Union, the League of Women Voters, the League of Conservation Voters, the American Federation of Government Employees, the American Federation of State, County, Municipal Employees, the American Federation of Teachers, the AFL-CIO Building and Construction, the United Auto Workers, the Conservative Coalition, the U. S. Chamber of Commerce, the American Conservative Union, the Christian Voters Victory Fund, the Christian Voice, Lower Federal Spending, and Taxation with Representation. Not all the ratings are available in all years, so sample sizes vary when using these alternative ratings.

As we show below, our results are remarkably stable across alternative measures of roll-call votes. This finding lends some credibility to the conclusion that our estimates are not driven by the unique characteristics of one particular measure. See the Data Appendix for a detailed discussion of our samples and data sources.

## V. EMPIRICAL RESULTS

In this section we present our empirical results. Subsection V.A presents our main results with a simple graphical analysis that illustrates that changes in electoral strength appears to affect future voting records entirely because it alters the relative odds of which party's nominee will be elected to the House. That is, candidates do not seem to change their intended policies in response to large exogenous shocks to electoral strength. This is followed by more formal estimates of the key parameters of interest:  $\gamma$ ,  $\pi_1[P_{t+1}^D - P_{t+1}^R]$ , and  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ . Subsection V.B provides evidence supporting our main identifying assumption—that among elections that turn out to be close, who wins is “as good as randomly assigned.” In subsection V.C, we show that our results do not change substantially when we utilize a number of alternative voting record indices. Finally, in subsection V.D, we examine the sensitivity of our results to a functional form as-

18. The results are nearly identical if one uses the party whip instead of the party leader.

sumption utilized in our base model. The analysis within this more general framework confirms our findings.

### V.A. Main Empirical Results

*Graphical Analysis.* As discussed above, the way to distinguish between full divergence and partial divergence is to analyze the effects on roll-call votes of an exogenous change in the probability of winning the election. The total effect of such an exogenous change on roll-call behavior ( $\gamma$ ) can be split into two components: the elect component ( $\pi_1[P_{t+1}^D - P_{t+1}^R]$ ) and the affect component ( $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ ). If voters merely “elect” policies (complete divergence), we should observe little change in the candidates’ intended policies following an exogenous increase in the probability of victory:  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$  should be small. If voters not only choose politicians, but also affect their policy choices (partial convergence), candidates should move toward their bliss points in response to an exogenously higher probability of winning:  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$  should be relatively important.

We begin with a simple graphical analysis of ADA scores. Figure I plots ADA scores at time  $t + 1$  against the Democrat vote share at time  $t$ . As an example, we are relating the ADA scores for the representative from, say, the first California district during the 1995–1996 Congressional session to the Democratic vote share observed in the 1992 election in that district. In practice, we use all pairs of adjacent years from 1946 to 1995, except for the pairs where we cannot link districts due to redistricting (pairs with years ending with “0” and “2”).

Throughout the paper the unit of observation is the district in a given year. But to give an overall picture of the data, each point in Figure I is an average of the ADA score in period  $t + 1$  within 0.01-wide intervals of the vote share at time  $t$ . The vertical line marks 50 percent of the two-party vote share. Districts to the right of the vertical line are districts won by Democrats in election  $t$ , districts to the left are districts won by Republicans in election  $t$ . The continuous line is the predicted  $ADA_{t+1}$  score from a regression that includes a fourth-order polynomial in vote share and a dummy for observations above the 50 percent threshold, and an interaction of the dummy and the polynomial. The dotted lines represent pointwise 95 percent confidence intervals of this approximation.

A striking feature of the figure is that ADA scores appear to

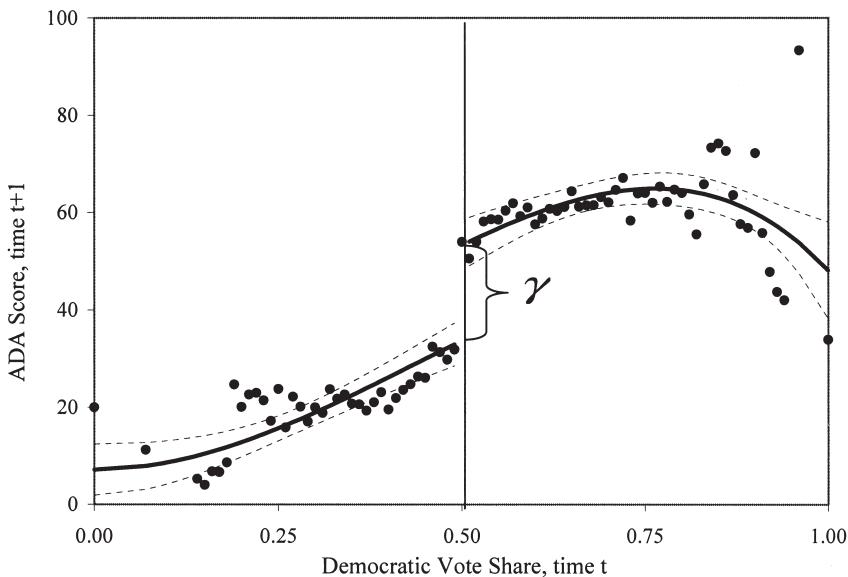


FIGURE I

Total Effect of Initial Win on Future ADA Scores:  $\gamma$

This figure plots ADA scores after the election at time  $t + 1$  against the Democrat vote share, time  $t$ . Each circle is the average ADA score within 0.01 intervals of the Democrat vote share. Solid lines are fitted values from fourth-order polynomial regressions on either side of the discontinuity. Dotted lines are pointwise 95 percent confidence intervals. The discontinuity gap estimates

$$\gamma = \underbrace{\pi_0(P_{t+1}^{SD} - P_{t+1}^{SR})}_{\text{"Affect"}} + \underbrace{\pi_1(P_{t+1}^{SD} - P_{t+1}^{SR})}_{\text{"Elect"}}$$

be a continuous and smooth function of vote shares everywhere, except at the threshold that determines party membership. There is a large discontinuous jump in ADA scores at the 50 percent threshold. Compare districts where the Democrat candidate barely lost in period  $t$  (for example, vote share is 49.5 percent), with districts where the Democrat candidate barely won (for example, vote share is 50.5 percent). If the regression discontinuity design is valid, the two groups of districts should appear ex ante similar in every respect—on average. The difference will be that in one group, the Democrats will be the incumbent for the next election ( $t + 1$ ), and in the other it will be the Republicans. Districts where the Democrats are the incumbent party for election  $t + 1$  elect representatives who have much higher ADA scores, compared with districts where the Republican candidate

barely won and became the incumbent—on average. The size of the jump appears to be fairly large, at around twenty ADA points.

What does this discontinuity mean? Formally, the gap is a credible estimate of the parameter  $\gamma$  in equation (4). Intuitively, it is unsurprising to observe some discontinuity. We know that party affiliation is an important determinant of roll-call behavior. We also know that if a Democrat (Republican) is elected in period  $t$ , a Democrat (Republican) is more likely to be elected in period  $t + 1$  in the same district, due to the incumbency advantage. The party effect, together with the electoral advantage of incumbency, suggests that we should expect to find a gap in Figure I. It is not surprising to observe that, for example, the 1995–1996 voting records are more liberal in the districts that were won by Democrats in 1992. The 1995–1996 representative, after all, is more likely be a Democrat. In Sections II and III we called this particular mechanism the “elect component,” and denoted it  $\pi_1[P_{t+1}^D - P_{t+1}^R]$ .

There is a second component that contributes to  $\gamma$ . If candidates are constrained by expected voters’ behavior, then a Democrat who is challenging an incumbent Republican (left side of the graph) would be expected to moderate his intended policies more, compared with an incumbent Democrat (right side of the graph). After all, the incumbent would be in a much stronger electoral position compared with the challenger. This is the other reason why voting scores should be more liberal where the Democrat is the incumbent, and hence why there should be a gap in Figure I. In Sections II and III we labeled this phenomenon “voters affecting policies,” and denoted it  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ .

The discontinuity  $\gamma$  illustrated in Figure I is equal to  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}] + \pi_1[P_{t+1}^D - P_{t+1}^R]$ . While it is not surprising to find that  $\gamma > 0$ , the real question is whether  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$  or  $\pi_1[P_{t+1}^D - P_{t+1}^R]$  dominates. Our empirical strategy is simple. We can directly estimate  $\pi_1[P_{t+1}^D - P_{t+1}^R]$ , by separately estimating  $\pi_1$ , the expected difference in voting between the two parties, as well as  $[P_{t+1}^D - P_{t+1}^R]$ , the electoral advantage to incumbency. We can subtract this from the total effect  $\gamma$  to determine the magnitude of the “affect component”  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ .

In Figure II we illustrate the two elements that make up the elect component,  $\pi_1$  and  $[P_{t+1}^D - P_{t+1}^R]$ . The top panel in Figure II plots ADA scores at time  $t$  against the Democrat vote share at time  $t$ . As in Figure I, average ADA scores appear to be a continuous and smooth function of vote shares everywhere, except at

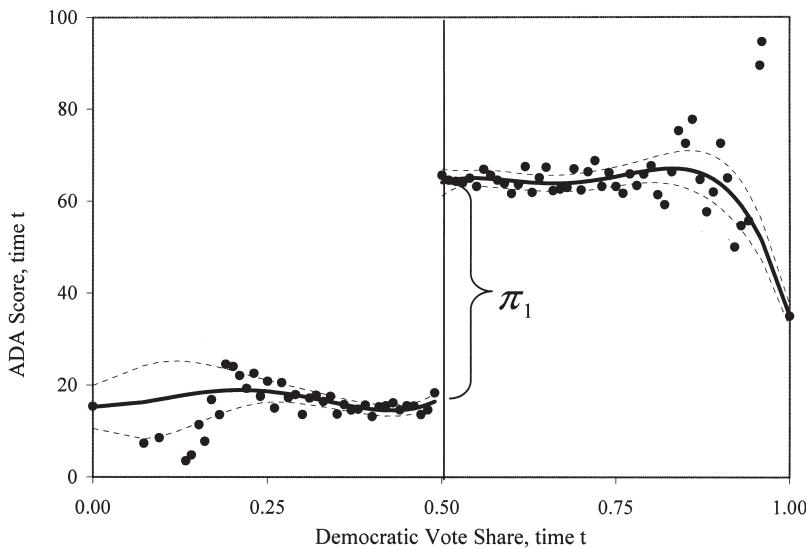


FIGURE IIa  
Effect of Party Affiliation:  $\pi_1$

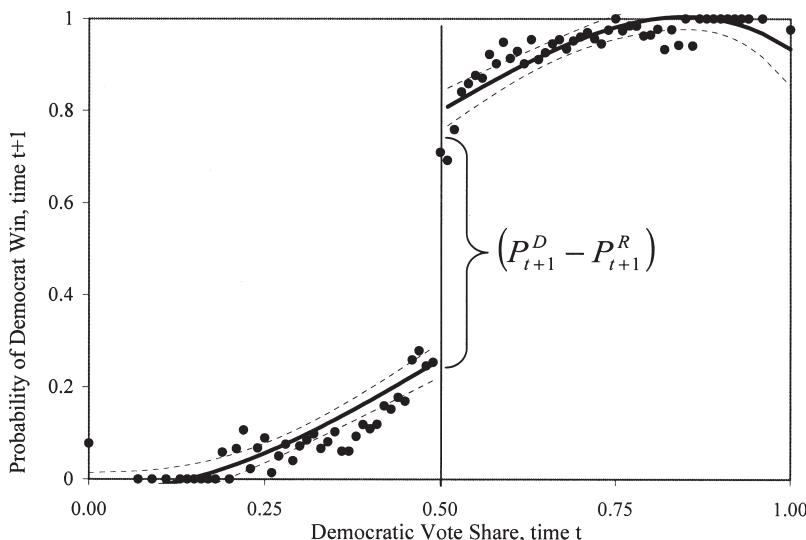


FIGURE IIb  
Effect of Initial Win on Winning Next Election:  $(P_{t+1}^D - P_{t+1}^R)$

Top panel plots ADA scores after the election at time  $t$  against the Democrat vote share, time  $t$ . Bottom panel plots probability of Democrat victory at  $t + 1$  against Democrat vote share, time  $t$ . See caption of Figure III for more details.

the threshold that determines party membership. The jump is a credible estimate of  $\pi_1$  in equation (5).

Compare a district where the Democrat candidate barely lost at time  $t$  (for example, vote share of 49.5 percent), with a district where the Democrat candidate barely won at time  $t$  (for example, vote share of 50.5 percent). Again, if the regression discontinuity design is valid and has generated random assignment of who wins in  $t$ , then the average voting records of Democrats who are *barely* elected will credibly represent, on average, how Democrats *would have* voted in the districts that were in actuality, barely won by Republicans (and vice versa). The observed difference in voting scores represents a credible estimate of the average policy differences between the two parties across districts—the direct influence of party affiliation on voting scores. The difference at the 50 percent threshold appears quite large, with a gap of about 45 points.

Finally, the bottom panel of Figure II plots estimates of the probability that the Democrat will win election  $t + 1$  for a given Democratic vote share at  $t$ . As in the previous cases, the figure shows a smooth function of vote shares everywhere, except at the threshold that determines which party won  $t$ . The size of the jump estimates [ $P_{t+1}^D - P_{t+1}^R$ ] in equation (6).<sup>19</sup>

The discontinuity around the 50 percent threshold indicates that, for example, districts which barely elected a Democrat in  $t$  are more likely to elect a Democrat in  $t + 1$ , consistent with a causal incumbency advantage. This is consistent with these districts experiencing exogenous increases in the probability of electing a Democrat (Republican) in 1994.

The total effect  $\gamma$ , given by the discontinuity in Figure I, appears to be about twenty ADA points. Figure IIa shows that the estimate of  $\pi_1$  is about 45 points, and Figure IIb shows that [ $P_{t+1}^D - P_{t+1}^R$ ] is around 0.5. The “elect component”  $\pi_1 [P_{t+1}^D - P_{t+1}^R]$  is thus approximately  $45 \times 0.5 = 22.5$ . The small difference (20–22.5) implies that the *entire* effect of an exogenous change in electoral strength on future ADA scores is not operating through how candidates’ policy choices respond to changes in the probability of winning. Instead, the effect is operating through simply changing the relative odds that a party will retain control over the seat. That is, this graphical analysis indicates that voters

19. This regression-discontinuity estimate of the incumbency advantage is documented in Lee [2001, 2003].

TABLE I  
RESULTS BASED ON ADA SCORES—CLOSE ELECTIONS SAMPLE

Variable	Total effect			Elect component	Affect component
	$\gamma$	$\pi_1$	$(P_{t+1}^D - P_{t+1}^R)$	$\pi_1[(P_{t+1}^D - P_{t+1}^R)]$	$\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$
	$ADA_{t+1}$	$ADA_t$	$DEM_{t+1}$	(col. (2)*col. (3))	(col. (1)) – (col. (4))
	(1)	(2)	(3)	(4)	(5)
Estimated gap	21.2 (1.9)	47.6 (1.3)	0.48 (0.02)		
				22.84 (2.2)	-1.64 (2.0)

Standard errors are in parentheses. The unit of observation is a district-congressional session. The sample includes only observations where the Democrat vote share at time  $t$  is strictly between 48 percent and 52 percent. The estimated gap is the difference in the average of the relevant variable for observations for which the Democrat vote share at time  $t$  is strictly between 50 percent and 52 percent and observations for which the Democrat vote share at time  $t$  is strictly between 48 percent and 50 percent. Time  $t$  and  $t + 1$  refer to congressional sessions.  $ADA_t$  is the adjusted ADA voting score. Higher ADA scores correspond to more liberal roll-call voting records. Sample size is 915.

primarily elect policies (full divergence) rather than affect policies (partial convergence).

Here we quantify our estimates more precisely. In the analysis that follows, we restrict our attention to “close elections”—where the Democrat vote share in time  $t$  is strictly between 48 and 52 percent. As Figures I and II show, the difference between barely elected Democrat and Republican districts among these elections will provide a reasonable approximation to the discontinuity gaps. There are 915 observations, where each observation is a district-year.<sup>20</sup>

Table I, column (1), reports the estimated total effect  $\gamma$ , the size of the jump in Figure I. Specifically, column (1) shows the difference in the average  $ADA_{t+1}$  for districts for which the Democrat vote share at time  $t$  is strictly between 50 percent and 52 percent and districts for which the Democrat vote share at time  $t$  is strictly between 48 percent and 50 percent. The estimated difference is 21.2.

In column (2) we estimate the coefficient  $\pi_1$ , which is equal to the size of the jump in Figure IIa. The estimate is the difference in the average  $ADA_t$  for districts for which the Democrat vote

20. In 68 percent of cases, the representative in period  $t + 1$  is the same as the representative in period  $t$ . The distribution of close elections is fairly uniform across the years. In a typical year there are about 40 close elections. The year with the smallest number is 1988, with twelve close elections. The year with the largest number is 1966, with 92 close elections.

share at time  $t$  is strictly between 50 percent and 52 percent and districts for which the Democrat vote share at time  $t$  is strictly between 48 percent and 50 percent. The estimated difference is 47.6.

In column (3) we estimate the quantity  $[P_{t+1}^D - P_{t+1}^R]$ , which is equal to the size of the discontinuity documented in Figure IIb. The estimated jump is 0.48. This indicates that if the Democrat (Republican) candidate wins a close election in a given district in, say, 1992, the Democrat (Republican) candidate in the same district has a 0.48 higher probability of winning in 1994. This is indeed consistent with the notion that the party that already holds a seat holds a substantial electoral advantage.

In column (4) we multiply the estimates in columns (2) and (3) to obtain an estimate of the elect component,  $\pi_1[P_{t+1}^D - P_{t+1}^R]$ . The product is 22.84, which is not statistically different from the estimate of  $\gamma$  in column (1). Because estimates of  $\gamma$  and  $\pi_1[P_{t+1}^D - P_{t+1}^R]$  are quite similar, we conclude that the “affect component” is quite small. In column (5) we subtract the estimate in column (4) from the estimate in column (1) to yield  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ . The difference is virtually zero.

What is the relative importance of the “elect component”  $\pi_1[P_{t+1}^D - P_{t+1}^R]$ , and the “affect component”  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ ? Our results indicate that  $\pi_1[P_{t+1}^D - P_{t+1}^R]$  overwhelmingly dominates  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ . Indeed, it entirely explains the overall effect  $\gamma$ . Below, we show that this is true not only on average, but also for every decade taken separately.

### *V.B. Tests for Quasi-Random Assignment*

It is important to note that our empirical test crucially relies on the assumption of random assignment of the winner in close elections in  $t$ . Specifically, the key identifying assumption in our analysis is that as one compares closer and closer elections, all *predetermined* characteristics of Republican and Democratic districts (including the district-specific bliss points) become more and more similar. If this assumption does not hold, our estimates are likely to be biased.

Intuitively, this assumption seems to make sense. While Republican and Democrat districts are likely to be very different in general, the difference should decline as we examine elections whereby “pure luck” is a more important determinant of who wins—in other words, elections that turn out to be won by a tiny

margin. In the Appendix we provide a formal discussion of this assumption. Here, we provide two pieces of empirical evidence to support this assumption.

First, if examining close elections truly provides random assignment, characteristics determined before time  $t$  should be the same on both sides of the 50 percent threshold—on average.<sup>21</sup> We find that as we compare closer and closer elections, Republican and Democrat districts do have similar observable characteristics. Consider, for example, geographical location. There are sizable geographical differences in the entire sample. Averaging over the entire time period, Democrats are significantly more likely to be elected in the South than in the North and the West. However, as we start restricting the sample to closer and closer elections, the geographical differences decrease. For elections that are within only two percentage points from the threshold, the differences are not statistically significant.

This is shown graphically in Figures III and IV, which plot average district characteristics against Democratic vote share. Other than geographical location, we consider the following pre-determined characteristics: real income, percentage with high school degree, percentage black, percentage eligible to vote, and size of the voting population. Generally, the figures indicate that the difference at the 50 percent threshold is small and statistically insignificant.

Table II illustrates the same point by quantifying the difference between Democrat and Republican districts for a larger set of characteristics. In particular, we examine all the characteristics shown in Figures III and IV, as well as the fraction of open seats, percent urban, percent manufacturing employment, and percent eligible to vote.<sup>22</sup> Column (1) includes the entire sample. Columns (2) to (5) include only districts with Democrat vote share between 25 percent and 75 percent, 40 percent and 60 percent, 45 percent and 55 percent, and 48 percent and 52 percent, respectively. The model in column (6) is equivalent to Figures III and IV, since it includes a fourth-order polynomial in Democrat vote

21. See Lee [2003] for the conditions under which RD designs can generate variation in the treatment that is as good as randomized.

22. Data on districts' characteristics in each election year are from the last available Census of Population. Because the census takes place every ten years, standard errors allow for clustering at the district-decade level.

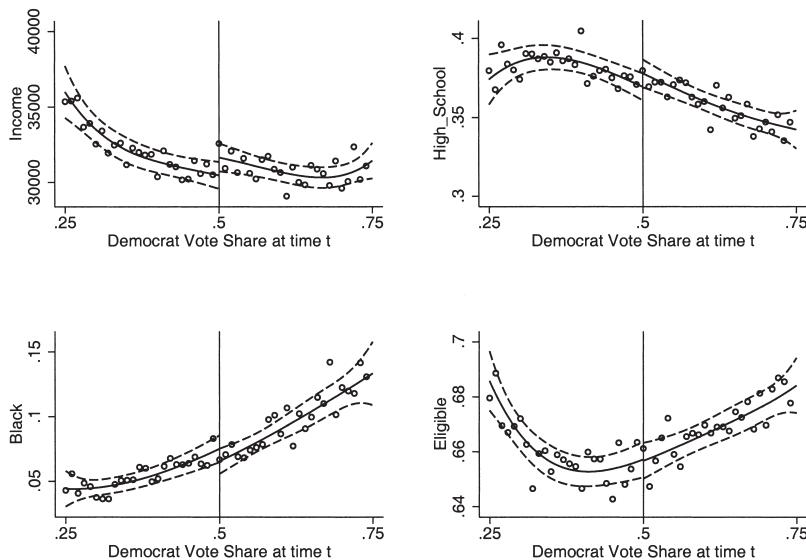


FIGURE III  
Similarity of Constituents' Characteristics in Bare Democrat and Republican Districts—Part 1

Panels refer to (from top left to bottom right) the following district characteristics: real income, percentage with high-school degree, percentage black, percentage eligible to vote. Circles represent the average characteristic within intervals of 0.01 in Democrat vote share. The continuous line represents the predicted values from a fourth-order polynomial in vote share fitted separately for points above and below the 50 percent threshold. The dotted line represents the 95 percent confidence interval.

share. The coefficient reported in column (6) is the predicted difference at 50 percent. The table confirms that, for many observable characteristics, there is no significant difference in a close neighborhood of 50 percent. One important exception is the percentage black, for which the magnitude of the discontinuity is statistically significant.<sup>23</sup>

As a consequence, estimates of the coefficients in Table I from regressions that include these covariates would be expected to produce similar results—as in a randomized experiment—since

23. This is due to few outliers in the outer part of the vote share range. When the polynomial is estimated including only districts with vote share between 25 percent and 75 percent, the coefficients becomes insignificant. The gap for percent urban and open seats, while not statistically significant at the 5 percent level, is significant at the 10 percent level.

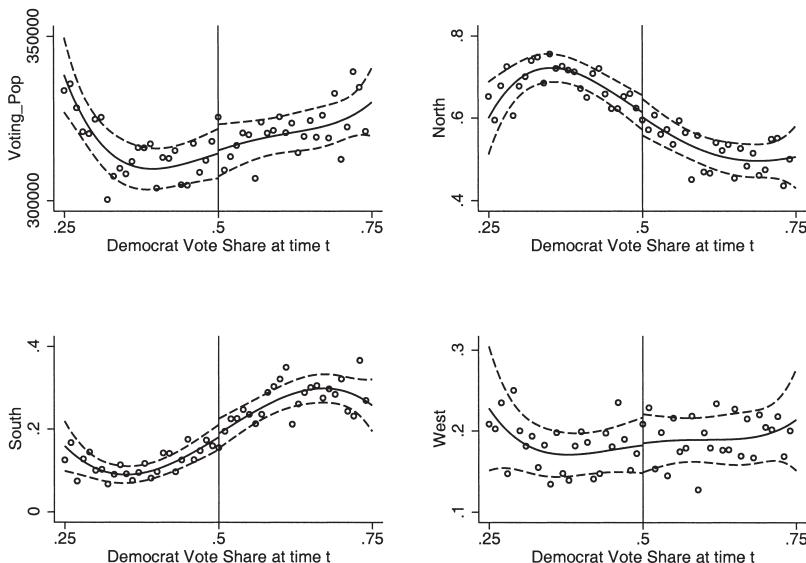


FIGURE IV  
Similarity of Constituents' Characteristics in Bare Democrat and Republican Districts—Part 2

Panels refer to (from top left to bottom right) the following district characteristics: voting population, North, South, West. Circles represent the average characteristic within intervals of 0.01 in Democrat vote share. The continuous line represents the predicted values from a fourth-order polynomial in vote share fitted separately for points above and below the 50 percent threshold. The dotted line represents the 95 percent confidence interval.

all predetermined characteristics appear to be orthogonal to  $D_t$ . We have reestimated all the models in Table I conditioning on all of the district characteristics in Table II, and found estimates that are virtually identical to the ones in Table I.

As a similar empirical test of our identifying assumption, in Figure V we plot the ADA scores from the Congressional sessions that preceded the determination of the Democratic two-party vote share in election  $t$ . Since these past scores have already been determined by the time of the election, it is yet another predetermined characteristic (just like demographic composition, income levels, etc.). If the RD design is valid, then we should observe no discontinuity in these lagged ADA scores—just as we would expect, in a randomized experiment, to see no systematic differences in any variables determined prior to the experiment. The

TABLE II  
DIFFERENCE IN DISTRICT CHARACTERISTICS BETWEEN DEMOCRAT AND REPUBLICAN  
DISTRICTS, BY DISTANCE FROM 50 PERCENT

	All (1)	+/- 25 (2)	+/- 10 (3)	+/- 5 (4)	+/- 2 (5)	Polynomial (6)
North	-0.211 (0.018)	-0.156 (0.019)	-0.096 (0.021)	-0.054 (0.024)	-0.059 (0.036)	-0.041 (0.045)
South	0.250 (0.015)	0.145 (0.014)	0.093 (0.016)	0.053 (0.019)	0.009 (0.028)	0.015 (0.036)
West	-0.031 (0.013)	-0.012 (0.015)	-0.036 (0.020)	-0.003 (0.017)	0.001 (0.020)	0.001 (0.036)
Log income	-0.086 (0.013)	-0.036 (0.012)	0.014 (0.014)	0.026 (0.017)	0.030 (0.026)	0.052 (0.033)
Percentage high-school grad.	-0.035 (0.003)	-0.024 (0.003)	-0.008 (0.004)	-0.001 (0.004)	0.001 (0.007)	0.008 (0.008)
Percentage urban	0.070 (0.011)	0.065 (0.011)	0.053 (0.012)	0.053 (0.014)	0.056 (0.023)	0.053 (0.028)
Percentage black	0.082 (0.005)	0.042 (0.004)	0.013 (0.004)	0.003 (0.005)	-0.003 (0.009)	-0.053 (0.013)
Manufacturing employment	-0.002 (0.001)	0.000 (0.001)	0.004 (0.002)	0.004 (0.002)	0.005 (0.004)	0.003 (0.005)
Total population	-1817.9 (3517.3)	3019.2 (3723.0)	4961.5 (4562.4)	3211.4 (5524.2)	8640.4 (8427.9)	2007.5 (10483.0)
Percentage eligible to vote	0.005 (0.002)	0.010 (0.002)	0.007 (0.003)	0.006 (0.004)	-0.003 (0.006)	-0.003 (0.007)
Open seats	0.070 (0.011)	0.065 (0.011)	0.053 (0.012)	0.053 (0.014)	0.056 (0.023)	0.053 (0.028)
Number of observations	13413	10229	4174	2072	910	13413

Standard errors are in parentheses. The unit of observation is a district-congressional session. Columns (1) to (5) report the difference in average district characteristics between Democrat and Republican districts. Column (1) includes the entire sample. Columns (2) to (5) include only districts with Democrat vote share between 25 percent and 75 percent, 40 percent and 60 percent, 45 percent and 55 percent, and 48 percent and 52 percent, respectively. The model in column (6) includes a fourth-order polynomial in Democrat vote share that enters separately for vote share above and below 50 percent. The coefficient reported in column (6) is the predicted difference at 50 percent. All standard errors account for district-decade clustering.

lack of discontinuity in the figure lends further credibility to our identifying assumption.<sup>24</sup>

Overall, the evidence strongly supports a valid regression discontinuity design. And as a consequence, it appears that among close elections, who wins appears virtually randomly assigned, which is the identifying assumption of our empirical strategy.

24. The estimated gap is 3.5 (5.6).

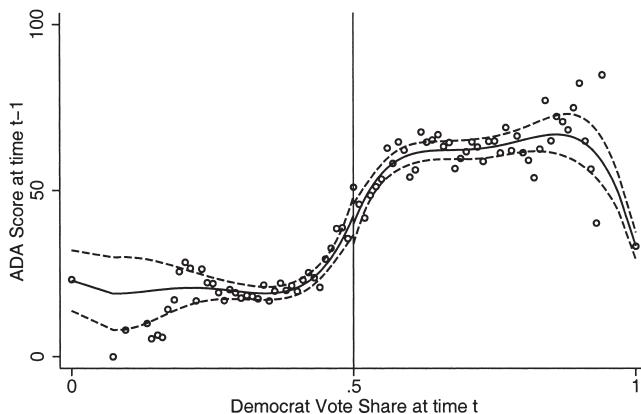


FIGURE V

Specification Test: Similarity of Historical Voting Patterns between Bare Democrat and Republican Districts

The panel plots one time lagged ADA scores against the Democrat vote share. Time  $t$  and  $t - 1$  refer to congressional sessions. Each point is the average lagged ADA score within intervals of 0.01 in Democrat vote share. The continuous line is from a fourth-order polynomial in vote share fitted separately for points above and below the 50 percent threshold. The dotted line is the 95 percent confidence interval.

### V.C. Sensitivity to Alternative Measures of Voting Records

Our results so far are based on a particular voting index, the ADA score. In this section we investigate whether our results generalize to other voting scores. We find that the findings do not change when we use alternative interest groups scores, or other summary measures of representatives' voting records.

Table III is analogous to Table I, but instead of using ADA scores, it is based on two alternative measures of roll-call voting. The top panel is based on McCarty, Poole, and Rosenthal's DW-NOMINATE scores. The bottom panel is based on the percent of individual roll-call votes cast that are in agreement with the Democrat party leader. All the qualitative results obtained using ADA scores (Table I) hold up using these measures. When we use the DW-NOMINATE scores,  $\gamma$  is  $-0.36$ , remarkably close to the corresponding estimate of  $\pi_1[P_{t+1}^D - P_{t+1}^R]$  in column (4), which is  $-0.34$ . The estimates are negative here because, unlike ADA scores, higher Nominate scores correspond to a more conservative voting record. When we use the measure "percent voting with the Democrat leader,"  $\gamma$  is  $0.13$ , almost indistinguishable from the

TABLE III  
RESULTS BASED ON NOMINATE SCORES AND ON PERCENT VOTED LIKE DEMOCRAT  
LEADERSHIP—CLOSE ELECTIONS SAMPLE

Variable	Total effect		$(P_{t+1}^D - P_{t+1}^R)$	$\pi_1(P_{t+1}^D - P_{t+1}^R)$ (col. (2)*col. (3))	Elect component	Affect component
	$\gamma$	$\pi_1$			$DEM_{t+1}$	$\pi_0(P_{t+1}^{RD} - P_{t+1}^{R^*})$
	$Z_{t+1}$	$Z_t$				(5)
(a) Results based on Nominate scores						
Estimated gap	–0.36 (0.03)	–0.58 (0.02)	0.62 (0.04)		–0.34 (0.04)	–0.02 (0.04)
(b) Results based on percent voted like Democrat leadership						
Estimated gap	0.13 (0.01)	0.29 (0.006)	0.46 (0.02)		0.13 (0.02)	0.00 (0.02)

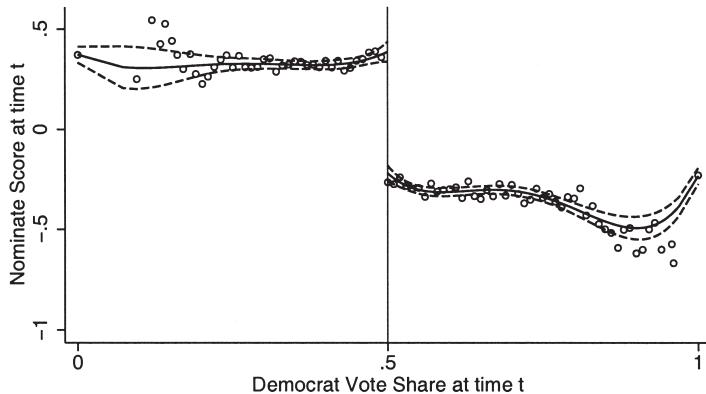
Standard errors are in parentheses. The unit of observation is a district-congressional session. The sample includes only observations where the Democrat vote share at time  $t$  is strictly between 48 percent and 52 percent. The estimated gap is the difference in the relevant variable for observations for which the Democrat vote share at time  $t$  is strictly between 50 percent and 52 percent and observations for which the Democrat vote share at time  $t$  is strictly between 48 percent and 50 percent. Time  $t$  and  $t + 1$  refer to congressional sessions. The top panel uses the DW-NOMINATE score constructed by McCarty, Poole, and Rosenthal. Higher Nominate scores correspond to more conservative roll-call voting records. The bottom panel uses the percent of a representative's votes that agree with the Democrat party leader. Sample size is 276 in top panel and 1010 in bottom panel.

estimate  $\pi_1[P_{t+1}^D - P_{t+1}^R]$  in column (4), which is 0.13. We show the graphical analysis for the estimate of  $\pi_1$  in Figure VI.

Our empirical findings are also not sensitive to the use of ratings from various liberal and conservative interest groups. Liberal interest groups include the American Civil Liberties Union, the League of Women Voters, the League of Conservation Voters, the American Federation of Government Employees, the American Federation of State, County, and Municipal Employees, the American Federation of Teachers, the AFL-CIO Building and Construction, and the United Auto Workers. Conservative groups include the Conservative Coalition, the U. S. Chamber of Commerce, the American Conservative Union, and the Christian Voice. All the ratings range from 0 to 100. For liberal groups, low ratings correspond to conservative roll-call votes, and high ratings correspond to liberal roll-call votes. For conservative groups the opposite is true.

These alternative ratings yield results that are qualitatively similar to our findings in Table I and III. Instead of presenting these results in a table format as we did in Table I and III, we present the main results in graphical form. We summarize our

### Nominate Scores



Percent Vote Equal to Democrat Party Leader

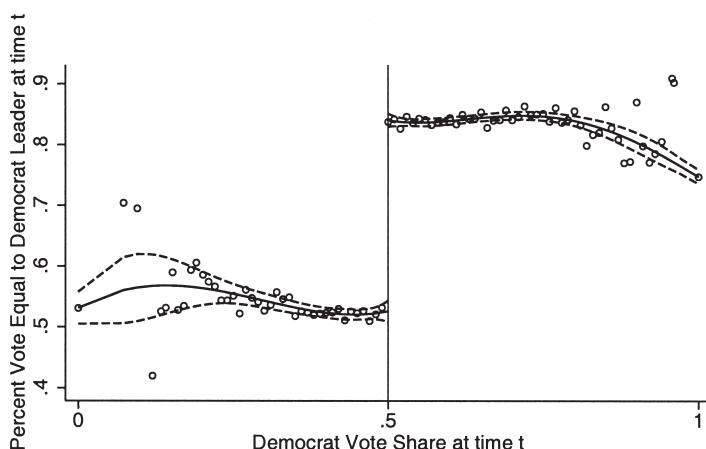


FIGURE VI

Nominate Scores, by Democrat Vote Share; and Percent Voted with Democrat Leader, by Democrat Vote Share

The top panel plots DW-Nominate scores at time  $t$  against the Democrat vote share at time  $t$ . Circles represent the average Nominate score within intervals of 0.01 in Democrat vote share. The bottom panel plots the fraction of a Representative's votes that agree with the Democrat party leader at time  $t$  against the Democrat vote share at time  $t$ . Circles represent the percent voted with Democrat leader within intervals of 0.01 in Democrat vote share. The continuous line is from a fourth-order polynomial fit in vote share fitted separately for points above and below the 50 percent threshold. The dotted line is the 95 percent confidence interval.

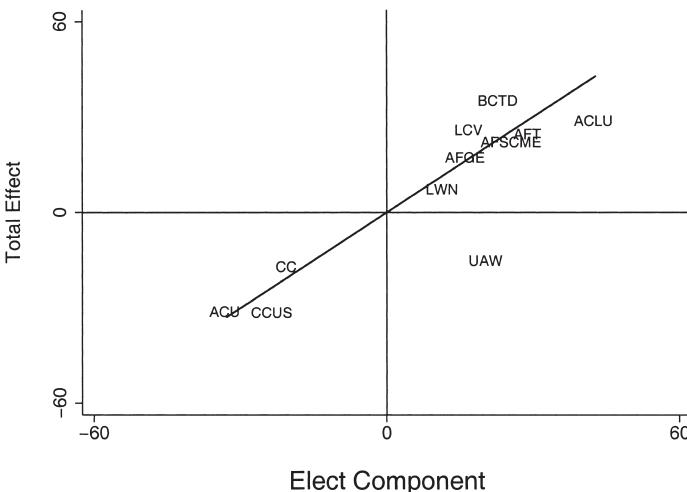


FIGURE VII  
Total Effect [ $\gamma$ ] Versus "Elect Policy" Component [ $\pi_1(P_{t+1}^D - P_{t+1}^R)$ ], for Alternative Interest Group Ratings

The figure plots estimates of the Total Effect [ $\gamma$ ] on the  $y$ -axis against estimates of the "Elect Policy" component [ $\pi_1(P_{t+1}^D - P_{t+1}^R)$ ] on the  $x$ -axis for eleven interest group ratings. The line is the  $45^\circ$  line. The closer the total effect is to the "elect component" effect for an interest group, the closer that group symbol is to the  $45^\circ$  line. All ratings are between 0 and 100. Higher ratings of liberal interest groups correspond to more liberal roll-call voting records. Higher ratings of conservative interest groups correspond to more conservative roll-call voting records.

ACLU is America Civil Liberties Union; LWV is League of Women Voters; LCV is League of Conservation Voters; AFGE is American Federal Government Employees; AFSCME is American Federation State, County, Municipal Employees; AFT is American Federation of Teachers; BCTD is AFL-CIO Building and Construction; UAW is United Auto Workers; CC is Conservative Coalition; CCUS is U. S. Chamber of Commerce; ACU is American Conservative Union.

results in Figure VII, where we plot our estimate of  $\gamma$  against our estimate of  $\pi_1[P_{t+1}^D - P_{t+1}^R]$  for each of these alternative interest group ratings. The diagonal is the  $45^\circ$  degree line. Most estimates are on the line or close to the line, indicating again that across a variety of different interest groups scores, the results are highly consistent with the full policy divergence hypothesis.<sup>25</sup>

Our qualitative findings seem insensitive to the choice of voting score. Representatives' policy positions, on a wide array of

25. Figures VIII to XI show that the relationship between ratings, and the democrat vote share shares the same general features as the relationships for ADA, DW-NOMINATE, and "Percent vote with Democrat Leader." In all cases, we find a large discontinuity.

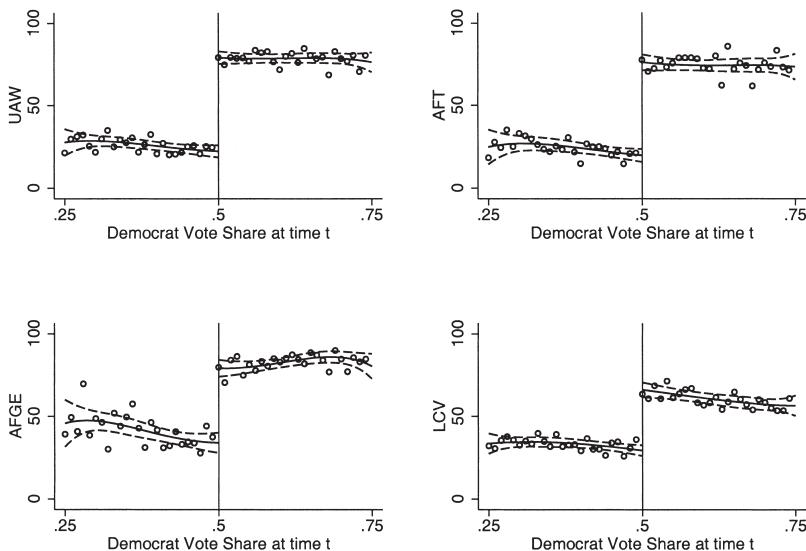


FIGURE VIII

## Liberal Interest Groups Ratings, by Democrat Vote Share—Part 1

The top panel on the left refers to ratings from United Auto Workers. The top panel on the right refers to ratings from American Federation of Teachers. The bottom panel on the left refers to ratings from American Federation of Government Employees. The bottom panel on the right refers to ratings from League of Conservative Voters. Circles represent average ratings within intervals of 0.01 in Democrat vote share. The continuous line is from a fourth-order polynomial in vote share fitted separately for points above and below the 50 percent threshold. The dotted line is the 95 percent confidence interval.

issues, do not seem to respond to exogenous changes in electoral strength. Voters appear to elect, rather than affect, candidates' platforms.

#### V.D. Heterogeneity

We now turn to the important issue of heterogeneity. Candidates' bliss points can be very different across districts or over time. For example, in any given year, a Democrat from Alabama is likely to have a bliss point that is quite different from a Democrat from Massachusetts. Our main results in subsection V.A are based on a model where candidates' positions can vary across districts and years. But one implicit functional form assumption that we adopted was that the *difference* in policy positions between Democrat and Republican candidates is constant across districts and over time. This assumption is violated if, for

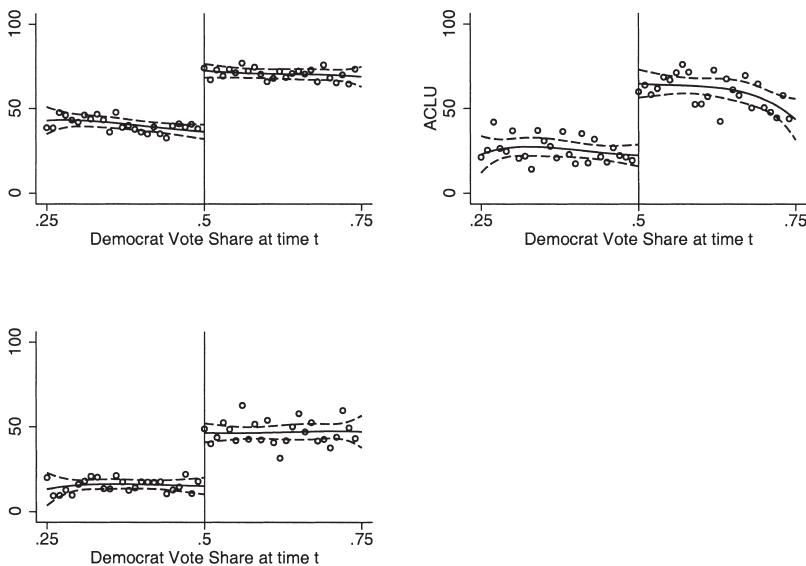


FIGURE IX

## Liberal Interest Groups Ratings, by Democrat Vote Share—Part 2

The top panel on the left refers to ratings from League of Women Voters. The middle panel on the right refers to ratings from American Civil Liberties Union. The bottom panel on the left refers to ratings from Taxation with Representation. Circles represent average ratings within intervals of 0.01 in Democrat vote share. The continuous line is from a fourth-order polynomial in vote share fitted separately for points above and below the 50 percent threshold. The dotted line is the 95 percent confidence interval.

example, the gap in intended policies between Democrats and Republicans from Alabama is different from the gap between Democrats and Republicans from Massachusetts.

We show, however, that our findings are robust to a more general framework that allows for virtually unrestricted heterogeneity in the gap between opposing candidates' policies across time and districts. The details can be found in Appendix 3.

Intuitively, we know that the total effect  $\gamma$  is partially driven by the impact of who wins election  $t$  on the composition of Democrats and Republicans in office after election  $t + 1$ . In the Appendix we show that it is possible, ex post, to identify the "marginal" districts that switched from Republican to Democrat in  $t + 1$ , because the Democrat won in  $t$ . By deleting these districts from the sample, we can examine the impact of an increase in electoral strength on policy positions, without the confounding effects of the compositional change (i.e., the "elect component").

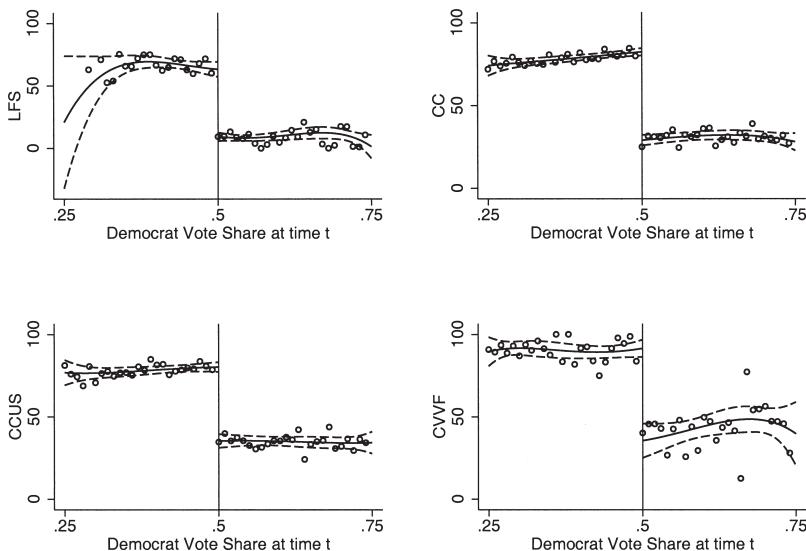


FIGURE X

## Conservative Interest Groups Ratings, by Democrat Vote Share—Part 1

The top panel on the left refers to ratings from the Lower Federal Spending Support Score. The top panel on the right refers to ratings from the Conservative Coalition. The bottom panel on the left refers to ratings from the Chamber of Commerce. The bottom panel on the right refers to ratings from the Christian Voters Victory Fund. Circles represent average ratings within intervals of 0.01 in Democrat vote share. The continuous line is from a fourth-order polynomial in vote share fitted separately for points above and below the 50 percent threshold. The dotted line is the 95 percent confidence interval.

Our findings using this more general framework indicate small estimates of  $\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$ . This implies that our finding that voters primarily elect policies is not an artifact of a functional form assumption of our basic specification.

We conclude this section considering a different type of heterogeneity: heterogeneity over time. In Table IV we replicate the estimates in Table I, presenting separate estimates by decade.<sup>26</sup> Column (2) shows that the discontinuity estimated by pooling all the years (Table II, column (2)) masks some variation in the discontinuity gap across states and years. This is not surprising, as the political science literature, for example, has noted that in the South, Democrats and Republicans are ideologically closer

26. Because there are few observations in 1940 and 1990, we merge 1940 with 1950, and 1980 with 1990.

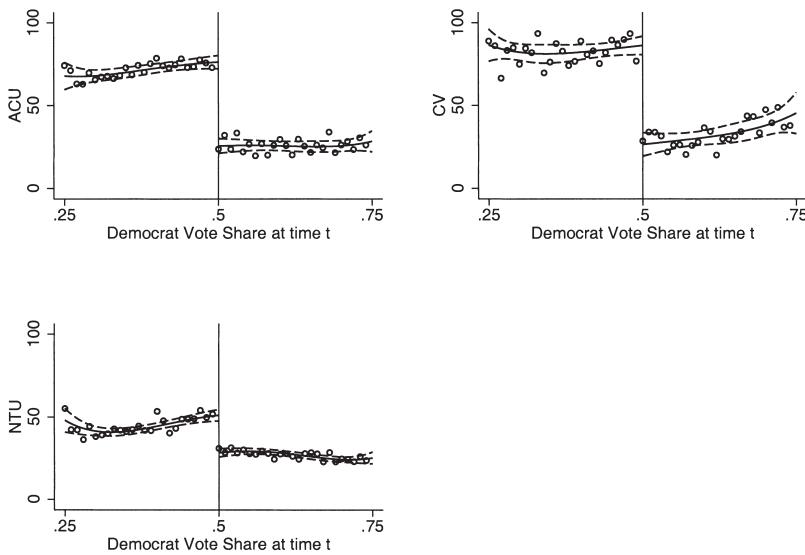


FIGURE XI

## Conservative Interest Groups Ratings, by Democrat Vote Share—Part 2

The top panel on the left refers to ratings from the American Conservative Union. The top panel on the right refers to ratings from the Christian Voice. The bottom panel on the left refers to ratings from the National Taxpayers' Union. Circles represent average ratings within intervals of 0.01 in Democrat vote share. The continuous line is from a fourth-order polynomial in vote share fitted separately for points above and below the 50 percent threshold. The dotted line is the 95 percent confidence interval.

than they are in the North. The estimated discontinuity is relatively smaller during the 1970s, and relatively larger during the 1990s. When we stratify by region and decade (not shown), the discontinuity is relatively smaller in the South in the 1950s and 1970s. Consistent with previous evidence, column (3) shows that the incumbency advantage has increased over time. The key findings of the table are in column (5). Notably, the relative importance of the “elect component” and the “affect component” remains fairly stable over time. Entries in column (5) indicate that our findings in Table I hold in each decade from 1940 to 1990.<sup>27</sup>

27. We have also estimated separate models for presidential election years and midterm election years. Our results are similar to the ones in Table I. For example, estimates similar to column (5) for presidential election years and midterm election years are, respectively, -0.1 (2.8) and -2.5 (3.0).

TABLE IV  
RESULTS BASED ON ADA SCORES, BY DECADE—CLOSE ELECTIONS SAMPLE

Variable	(1)		(2)		(3)		(4)		(5)	
	Total effect						Elect component		Affect component	
	$\gamma$	$\pi_1$	$(P_{t+1}^D - P_{t+1}^R)$	$\pi_1[(P_{t+1}^D - P_{t+1}^R)]$	$\pi_0[P_{t+1}^{*D} - P_{t+1}^{*R}]$	$(\text{col. (2)} * \text{col. (3)})$	$(\text{col. (1)}) - (\text{col. (4)})$			
1946–1958		14.2 (3.2)	41.7 (2.3)	0.41 (0.05)			17.0 (4.8)	−2.8 (4.0)		
1960–1968		23.5 (3.5)	49.5 (2.7)	0.51 (0.05)			25.2 (4.9)	−1.7 (4.1)		
1970–1978		11.5 (4.7)	46.6 (3.1)	0.40 (0.06)			18.6 (5.1)	−7.1 (5.1)		
1980–1996		46.8 (3.7)	56.6 (2.8)	0.76 (0.05)			43.0 (4.9)	3.8 (4.5)		

Standard errors are in parentheses. The unit of observation is a district-congressional session. The sample includes only observations where the Democrat vote share at time  $t$  is strictly between 48 percent and 52 percent. The estimated gap is the difference in the average of the relevant variable for observations for which the Democrat vote share at time  $t$  is strictly between 50 percent and 52 percent and observations for which the Democrat vote share at time  $t$  is strictly between 48 percent and 50 percent. Time  $t$  and  $t + 1$  refer to congressional sessions.  $ADA_t$  is the adjusted ADA voting score. Higher ADA scores correspond to more liberal roll-call voting records. Sample sizes are 322 in 1946–1958; 245 in 1960–1968; 183 in 1970–1978; 164 in 1980–1996.

## VI. RELATION TO PREVIOUS EMPIRICAL LITERATURE

A number of empirical studies have directly or indirectly examined the policy convergence issue.<sup>28</sup> Typically, the studies examine whether party affiliation matters for the observed voting records of the legislator. Most studies find evidence of this, which is strictly inconsistent with the *complete* policy convergence result. For example, Poole and Rosenthal [1984] show that senators from the same state belonging to different parties have significantly different voting records.

28. An example of early empirical work in this area is Miller and Stokes [1963]. The literature is too large to be summarized here. Other examples include, but are not limited to, Snyder and Ting [2001a], Fiorina [1999], Poole and Rosenthal [2001], Snyder and Ting [2001b], Lott and Davis [1992], Canes-Wrone, Brady, and Cogan [2002], Krehbiel [2000], Bender [1991], McArthur and Marks [1988], and McCarty, Poole, and Rosenthal [2000].

More recently, Snyder and Groseclose [2000] estimate the effect of party affiliation on roll-call votes using an identification strategy based on the assumption that on votes with 65 percent or more legislators on one side, legislators are not subject to party influence. Their main concern is that the correlation between party affiliation and voting may be spurious, possibly reflecting legislators' or constituency preferences. They conclude that in the majority of cases, party affiliation is a significant determinant of roll calls.<sup>29</sup>

We have argued here, however, that evidence against the (arguably unrealistic) outcome of *complete* convergence says little about the relative importance of the broader Downsian notion that electoral competition pressures politicians to moderate their positions. An appropriate assessment of this phenomenon rests on being able to empirically differentiate between *partial* convergence and complete policy *divergence*.

Other studies have noted a correlation between the electorate's and legislator's ideologies across constituencies and over time [Levitt 1996; Snyder and Ting 2001a]. This observation does *not* constitute evidence that candidates are being forced to deviate from their bliss points (i.e., partial convergence). Since bliss points are likely to vary across districts within the United States, this correlation could simply reflect that candidates' ideologies may be correlated with those of his fellow citizens in the district. These correlations suggest that the U. S. Congress may achieve a modest degree of "representation" across constituencies, in voting on legislation. Our findings suggest that whatever degree of representation that may arise, it appears not to occur through Downsian competition during a general election.

Our findings are generally consistent with existing evidence that candidates exhibit ideological rigidity. Many studies have shown that representatives do not change their positions during their careers, even when their constituency changes. For example, redistricting does not seem to make much of a difference.<sup>30</sup>

## VII. CONCLUSION

The arguably central notion of the Downsian paradigm is that the pressure of electoral competition constrains candidates'

29. See Krehbiel [2000] and McCarty, Poole, and Rosenthal [2000] for a different interpretation of the Snyder and Groseclose results.

30. See Poole and Rosenthal [1997] for an overview.

policies. This paper finds that voters do not *affect* policies in this way, but rather primarily *elect* policies. The empirical analysis finds that there is little evidence that members of the U. S. House alter their positions (in either direction) in response to a large exogenous change in the probability of winning the election. This result is unlikely to be spurious or a mere statistical artifact, because it is derived using a quasi experiment embedded in the Congressional electoral system that isolates arguably exogenous variation in the probability of winning the election.

The finding that voters primarily elect policies does not necessarily constitute evidence regarding the overall “representativeness” of members of the U. S. Congress. For example, one can define “representativeness” as the degree to which a politician represents the preferences of his/her core supporters. Our analysis would say nothing about this degree of representativeness. In loose terms, we have no way of knowing whether the Democrat’s (Republican’s) preferred policy exactly coincides with the most-preferred policy choice of the “median” Democratic (Republican) voter. It is possible that electoral competition in the U. S. primary system can work to force each party’s nominee to represent the median voter *within* each party. Instead, our analysis takes candidates’ preferred policies—representative or not—as given, and asks whether electoral competition in *general elections* compels politicians to deviate from those bliss points.

A leading explanation for our findings is that the difficulty of establishing credible commitments to moderate policies is an important real-world phenomenon, and dominates any Downsian convergence effect. This complete divergence hypothesis yields a very strong prediction—that exogenous changes in the relative popularity of the candidates has *no impact* on the candidates’ positions. Our quasi-experimental evidence is consistent with this bold prediction. Thus, at a minimum, there appears to be empirical support for the assumption that politicians are unable to overcome this credibility problem—an assumption that has been adopted in recent theoretical analyses of representative democracy (e.g., Besley and Coate [1997, 1998]).

Indeed, if one accepts this interpretation, it suggests the need for a more nuanced view than the Downsian perspective of how the American political system translates voters’ preferences into actual policy. When voters merely “elect” instead of “affect” policies, voters do not *influence* policy choices as much as they are *presented with* choices. Investigating how citizens can affect

which candidates, and hence which policies appear “on the ballot” seems to be a useful direction for future empirical research.

## APPENDIX 1

### 1. Derivation of Equation (2)

By definition, we have  $RC_t = y^*(k, \lambda^*, c) + (x^*(k, \lambda^*, c) - y^*(k, \lambda^*, c))D_t$ . Both  $x^*$  and  $y^*$  depend on  $k$ , the Nash bargaining point  $\lambda^* \in (0,1)$  and the bliss point  $c$ . Symmetry of the  $P$  function yields a symmetric efficient frontier (around the line  $y^* = c - x^*$ ). Consider that  $P_t^*$  can take on only two values,  $P_t^{*D}$  and  $P_t^{*R}$  (with  $P_t^{*D} > P_t^{*R}$ ), corresponding to which party was the incumbent party in election  $t$ . Assuming symmetry of the incumbency advantage, this implies  $P_t^{*D} = 1 - P_t^{*R}$ . Given the above symmetries, it follows that the Nash bargains are such that  $\lambda^*(P_t^{*D}) = 1 - \lambda^*(P_t^{*R})$ . Symmetry then implies that  $x^*(k, \lambda^*(P_t^{*D}), c) - y^*(k, \lambda^*(P_t^{*D}), c)$  will be equal to  $x^*(k, \lambda^*(P_t^{*R}), c) - y^*(k, \lambda^*(P_t^{*R}), c)$ , and this gap we can denote,  $\pi_1$ . Defining  $\pi_0 \equiv [y^*(k, \lambda^*(P_t^{*D}), c) - y^*(k, \lambda^*(P_t^{*R}), c)]/[P_t^{*D} - P_t^{*R}]$ , we obtain

$$RC_t = \text{constant} + \pi_0 P_t^* + \pi_1 D_t + \varepsilon_t,$$

where we have added a translational shift  $\varepsilon$  of the entire setup, to reflect heterogeneity across districts around the “average” district (the bliss points would be  $(\varepsilon_t, c + \varepsilon_t)$ ).

### 2. Random Assignment from Close Elections

Consider the vote share production function for all districts:

$$(7) \quad VS_t = vs(x^*(P_t^*, \varepsilon_t), y^*(P_t^*, \varepsilon_t), \varepsilon_t, e_t)$$

$$(8) \quad = \widehat{vs}(P_t^*, \varepsilon_t) + \hat{e}_t,$$

where  $\varepsilon_t$  and  $P_t^*$  vary across districts, and  $\varepsilon_t$  represents translational shifts in the parties’ bliss points.  $e_t$  is an unpredictable component of the vote share.  $\widehat{vs}_t$  is the forecast of the vote share, which depends on information at the time of the election, and  $\hat{e}_t$  is the forecast error.

By Bayes Rule, we have

$$(9) \quad \phi(P_t^*, \varepsilon_t | VS_t = v_0) = \frac{f(v_0 | P_t^*, \varepsilon_t) \cdot g(P_t^*, \varepsilon_t)}{h(v_0)},$$

where  $\phi(\cdot | \cdot)$  is the joint density of  $P_t^*$ ,  $\varepsilon_t$  conditional on  $VS_t$ .  $f(\cdot | \cdot)$  is the density of  $VS_t$  conditional on  $P_t^*$  and  $\varepsilon_t$ ,  $g(\cdot)$  is the marginal joint density of  $P_t^*$ ,  $\varepsilon_t$  and  $h(\cdot)$  is the marginal density of  $VS_t$ .

If  $\hat{e}_t$  has continuous density conditional on  $\widehat{vs}$ , then for a fixed  $P_t^*$ ,  $\varepsilon_t$ ,  $f(v_0 | P_t^*, \varepsilon_t)$  is continuous in  $v_0$ .  $h(v_0)$  is also continuous in  $v_0$ , which implies that  $\phi(P_t^*, \varepsilon_t | VS_t = v_0)$  is continuous in  $v_0$ . This implies that the distribution of  $P_t^*$ ,  $\varepsilon_t$  is very nearly the same between  $VS_t = 1/2 + \Delta$  and  $VS_t = 1/2 - \Delta$ , for  $\Delta$  small.<sup>31</sup>

## APPENDIX 2: DATA

The election data used in our analysis come from ICPSR Study 6311, “Elections to the United States House of Representatives, 1898–1992,” compiled by Gary King [1995]. ICPSR 6311 reports the state, district, incumbency status, and the total votes received by the Republican and Democrat candidates. ICPSR 6311 has the advantage that King distinguishes between “normal” elections that can be characterized as a Republican running against a Democrat, with no important independent candidates, and all other elections labeled as “exceptions.” We drop the 304 “exceptions” between 1946 and 1994 leaving 10,138 observations—roughly 422 observations per congressional session from 1946 to 1992. The 1994 election data were all obtained from the U. S. House of Representatives’ Office of the Clerk’s Web site.

The Americans for Democratic Action interest group ratings were taken from Timothy J. Groseclose’s Web site [2002]. As mentioned in the text, we use “adjusted” ADA scores to be able to compare ratings across sessions of congress. See Groseclose, Levitt, and Snyder [1999] for a description of the procedure used to calculate the “adjusted” ADA scores. We replicated the analysis in the paper with the “nominal” ADA scores and found no qualitative difference in our findings.

Congressional District Data for 1962–1970 were taken from ICPSR Study 10, “United States Congressional District Data Books.” Congressional District Data for 1972–1980 were taken from ICPSR Study 11, “United States Congressional District Data Book for the Ninety-Third Congress.” Congressional District

31. See DiNardo and Lee [2002] for a similar proof when  $P_{92}^*$  and  $\varepsilon$  have discrete support, and Lee [2003] for conditions under which an RD design can generate random assignment of treatment.

Data for 1982–1990 were taken from ICPSR Study 8710, “Census of Population and Housing, 1980 [United States]: STF3D Congressional District-level Extract.” Congressional District Data for 1992 were taken from the 1990 Congressional District level extract. All districts during a given decade were assigned the most recent census value. For example, 1974 congressional districts were assigned the 1970 census values.

Keith T. Poole’s [2002] Web site provided much of the data used in this paper. The non-ADA interest group ratings came from Poole and Daniels’ [1985] interest group rating data which contain 59 interest group ratings scores, each covering some or all of the years between 1959 and 1981. McCarty, Poole, and Rosenthal’s DW-NOMINATE scores and Poole’s rank ordering data were also found at Poole’s Web site. For a description of the NOMINATE procedure, see Poole and Rosenthal [1997], and for a specific description of the DW-NOMINATE scores see McCarty, Poole, and Rosenthal [1997] and for a description of the rank order procedure see Poole [1999].

James Snyder was kind enough to provide us with the Presidential Election returns by Congressional District from 1952 to 1996 for a previous draft of this paper.

#### APPENDIX 3: ADDITIONAL ESTIMATES UNDER UNRESTRICTED HETEROGENEITY

To see how it is possible to introduce heterogeneity in the difference between parties’ positions across districts and legislatures, we focus, as before, on districts that experience close elections at time  $t$  and assume that the winner in close elections is random.

Imagine that after election  $t + 1$ , *the econometrician* can divide this set of districts into three groups. The first group (which we call the *top group*) includes districts that were going to be Democrat in time  $t + 1$ , irrespective of the outcome at time  $t$ . This group includes the districts that were (exogenously) assigned a Republican representative at time  $t$  but become Democrat at time  $t + 1$ . But it also includes districts that were (exogenously) assigned a Democratic representative at time  $t$ , remained Democrat at time  $t + 1$ , and would have been Democrat at time  $t + 1$  even if they had been assigned a Republican winner at time  $t$ . In the terminology of Angrist, Imbens, and Rubin [1996], these districts are “always takers,” because irre-

spective of the quasi-random assignment of  $DEM_t$ , they are Democrat at time  $t + 1$ .

It is crucial to remember that even though *the econometrician* can identify the districts in this particular group, we maintain the assumption that all agents in the model *cannot*. If any party knew with certainty that it was going to win in  $t + 1$ , irrespective of the outcome of election  $t$ , there would be little incentive to moderate positions. We emphasize that we are not introducing a new assumption at this point. The assumption that electoral outcomes are not completely certain (particularly among close elections in  $t$ ) provides the basis for Alesina's theoretical framework, and the analysis throughout this paper.

In a parallel fashion, we can identify a "bottom group," which includes districts that, irrespective of the outcome at time  $t$ , become Republican at time  $t + 1$ . This group includes the districts that were (exogenously) assigned a Democrat representative at time  $t$  but are Republican at time  $t + 1$ . It also includes districts that were (exogenously) assigned a Republican representative at time  $t$ , remained Republican after time  $t + 1$ , and *would* have been Republican at time  $t + 1$  even if they had been assigned to be Democrat at time  $t$ . These districts could be called "never takers" [Angrist, Imbens, and Rubin 1996], because irrespective of the "random assignment" of  $DEM_t$ , they are Republican at time  $t + 1$ . Again, we can only identify these districts *ex post*, *from the point of view of the econometrician*.

The remaining districts make up the third and final group, which we call the *middle group*. In the terminology of Angrist, Imbens, and Rubin [1996] these districts are the "compliers."

We can think of the parameter  $\gamma$  that we estimated in our main result section as a weighted average of each group specific  $\gamma$ :

$$(10) \quad \gamma = Pr_1\gamma_1 + Pr_2\gamma_2 + Pr_3\gamma_3,$$

where  $\gamma_1$  is the  $\gamma$  for the top group,  $\gamma_2$  is for  $\gamma$  for the middle group, and  $\gamma_3$  is the  $\gamma$  for the bottom group; and  $Pr$  is the relative size of each group.

If we can identify the three groups, we can provide a test of complete divergence against partial convergence based on the size of the group specific  $\gamma$ s. In particular, we have sharp predictions on the size of  $\gamma_1$  and  $\gamma_3$  depending on whether complete divergence or partial convergence is true.

Under complete divergence, we expect the ADA scores of representatives from both the top and bottom group to be unaffected by who wins the election at time  $t$ . This is because, despite the underlying change in electoral strength, candidates are unable to credibly promise anything but their extreme bliss points. So it should not matter whether they had a strong or moderate chance of winning the election in  $t + 1$ .

By contrast, under partial convergence, a Democrat victory in  $t$  should cause ADA scores to rise for the top and bottom groups. A Democrat victory in  $t$  causes an increase in electoral strength for the Democrats for election  $t + 1$ . With greater “bargaining power” the compromise would move toward the Democrats’ bliss points.

Essentially, by dividing the districts into these three groups, we have eliminated the “elect component” when focusing only on the top and bottom groups. Among the top and bottom groups, the change in electoral strength will not change the composition of Democrats and Republicans in office after  $t + 1$ , because we have deleted the marginal districts (the middle group). Thus, both  $\gamma_1$  and  $\gamma_3$  are direct estimates of the “affect component”  $\pi_0(P_{t+1}^{*D} - P_{t+1}^{*R})$ .

How do we identify these three groups in practice? In most contexts, it would be impossible to identify these groups of districts. However, in our context, it is possible because we actually observe the Democratic vote share, which is the underlying index that perfectly determines who wins in  $t + 1$ .

To see how this is done, consider elections  $t$  and  $t + 1$ . The “vote production function” is  $VS_{t+1} = vs(x_{t+1}^*(P_{t+1}^*), y_{t+1}^*(P_{t+1}^*), e_{t+1})$ , where  $VS_{t+1}$  is the vote share for the Democrats.  $e$  is an unpredictable component of the vote share, which must be present if electoral outcomes are uncertain. Taking a linear approximation to this function yields (normalizing all coefficients to 1)  $VS_{t+1} = P_{t+1}^* + e_{t+1}$ .

Consider the districts where the Democrat defeated the incumbent Republican party in  $t + 1$ . These are the districts in which  $VS_{t+1} = P_{t+1}^{*R} + e_{t+1} > \frac{1}{2}$ . Next, consider the Democratic incumbents who succeeded in  $t + 1$  who have vote share  $VS_{t+1} = P_{t+1}^{*D} + e_{t+1} > \frac{1}{2}$ . To make this group comparable, we must select a subset of these districts. That is, we must choose  $VS_{t+1} > \frac{1}{2} + \bar{\theta}$ , where  $\bar{\theta} = (P_{t+1}^{*D} - P_{t+1}^{*R})$ . Doing this allows us to focus

on those districts such that  $P_{t+1}^{*R} + e_{t+1} > \frac{1}{2}$ . Intuitively, some of the districts where the Democratic incumbents won in  $t + 1$  would *not* have won if the district had been won by a Republican in  $t$ . This selection procedure eliminates these “marginal” districts.

Since  $D_t$  is essentially randomly assigned (among close elections), we have that  $E[RC_{t+1}|D_0 = 1, VS_{t+1} > \frac{1}{2} + \bar{\theta}] = E[x_{t+1}^*(P_{t+1}^{*D})|P_{t+1}^{*R} + e_{t+1} > \frac{1}{2}]$  and  $E[RC_{t+1}|D_0 = 0, VS_{t+1} > \frac{1}{2}] = E[x_{t+1}^*(P_{t+1}^{*R})|P_{t+1}^{*R} + e_{t+1} > \frac{1}{2}]$ . Therefore, the difference in these quantities is  $E[x_{t+1}^*(P_{t+1}^{*D}) - x_{t+1}^*(P_{t+1}^{*R})|D_t = 0, D_{t+1} = 1]$ . This is the response of Democrats’ policy choice to differing probabilities of winning  $t + 1$ ,  $P_{t+1}^{*D}$  and  $P_{t+1}^{*R}$ .

How do we practically choose this higher cutoff  $\bar{\theta}$ ? We observe the proportion of the  $D_t = 0$  group of districts that have  $P_{t+1}^{*R} + e_{t+1} > \frac{1}{2}$ . This proportion is  $\Lambda \equiv \Pr[D_{t+1} = 1|D_t = 0]$ . Now consider the  $D_t = 1$  group of districts. If  $D_t$  is randomly assigned, then we must choose  $\bar{\theta}$  so that  $\Pr[VS_{t+1} > \frac{1}{2} + \bar{\theta}|D_t = 1] = \Lambda$ . This can be done since  $VS_{t+1}$  is perfectly observable.

*Empirical Results with Heterogeneity.* We now describe how to empirically implement the theoretical framework described above. Our empirical strategy is similar to the one used to obtain our main results in subsection V.A. As with our basic results, we include in our analysis only districts that experienced close elections at time  $t$ . In these districts we assume again that the winner party at time  $t$  is virtually random. The main difference with our previous empirical strategy is that we now allow for a more unrestricted form of heterogeneity across districts and divide our sample of close elections in three groups.

We begin by showing in Panel A in Table V that in the sample of close elections, among districts that were won by Democrat in period  $t$ , 72.6 percent remain Democrat in  $t + 1$ , and 27.4 percent are won by Republicans. Among districts that were won by Republican in period  $t$ , 75.9 percent remain Republican in  $t + 1$ , and 24.1 percent are won by Democrat.

If challengers, who are electorally weaker, are responding to electoral pressures by moderating their positions, we should see that a successful Democrat challenger’s ADA score is less liberal than its comparison group, and that a successful Republican challengers’ ADA score is more liberal than its comparison group.

Take, for example, the top left corner of panel A. It represents

TABLE V  
ADDITIONAL ESTIMATES USING ADA SCORES—CLOSE ELECTIONS SAMPLE

	Districts, where $DEM_t = 0$		Districts, where $DEM_t = 1$	
	(1)	(2)	(3)	(4)
Percent $DEM_{t+1} = 1$	24.1		72.6	
Percent $DEM_{t+1} = 0$	75.9		27.4	
Total	100		100	
Panel B	Average $ADA_{t+1}$ in districts, where $DEM_t = 0$	Average $ADA_{t+1}$ in districts, where $DEM_t = 1$	Difference in average $ADA_{t+1}$ (col. (2)) – (col. (1))	Obs. (4)
	(1)	(2)	(3)	
Districts in top group	67.6 (3.0)	65.9 (2.9)	-1.7 (3.0)	224
Districts in middle group	18.6 (1.7)	66.1 (1.8)	47.4 (1.8)	441
Districts in bottom group	21.1 (2.2)	16.5 (2.1)	-4.6 (2.3)	250

Standard errors are in parentheses. The unit of observation is a district-congressional session. The sample includes only observations where the Democrat vote share at time  $t$  is strictly between 48 percent and 52 percent. Column (1) refers to observations for which the Democrat vote share at time  $t$  is strictly between 48 percent and 50 percent. Column (2) refers to observations for which the Democrat vote share at time  $t$  is strictly between 50 percent and 52 percent. Time  $t$  and  $t + 1$  refer to congressional sessions. Higher ADA scores correspond to more liberal roll-call voting records.

Democrats who were able to win the district even though it was held by a Republican. We would like to compare the voting score of those districts with the voting score of *otherwise comparable districts* where the Democrat party is the incumbent party, and the candidate from the Democrat party was elected at time  $t$ .

The empirical challenge is to identify the correct counterfactual. The methodology described above provides a solution and leads us to split the districts into three groups, as shown in Panel B. We empirically identify the three groups as follows. As before, we consider only the 915 observations that experience close elections at time  $t$ . We include in the top group of districts where a Republican (barely) won at time  $t$  but a Democrat regained the district at time  $t + 1$  as well as the districts with the largest Democrat vote shares at time  $t + 1$ , among the districts where a Democrat won at time  $t$ . Specifically, we included the top 24.1

percent districts with the largest Democrat vote share among the districts where a Democrat won at time  $t$ . There are 224 such districts.

Similarly, we include in the bottom group districts where a Democrat (barely) won at time  $t$  but a Republican regained the district at time  $t + 1$ , as well as districts with the largest Republican vote shares at time  $t + 1$ , among the districts where a Republican won at time  $t$ . Specifically, we included the 27.4 percent districts with the largest Republican vote share, since among all the districts where the Republican lost at time  $t$ , the Republican candidate won at time  $t + 1$  in 27.4 percent of the cases. There are 250 such districts. Finally, we include in the middle group the remaining districts. There are 441 such districts.

Because of the way we constructed the three groups, none of the top group districts in column (1) has a representative in period  $t + 1$  who is the same representative in period  $t$ . On the contrary, 91 percent of the top group districts in column (2) has a representative in period  $t + 1$  who is the same representative in period  $t$ . The corresponding figures for the bottom group are, respectively, 0 and 96 percent. The corresponding figures for the middle group are 90 percent and 89 percent.

Panel B shows the average voting scores, for each of the three groups, by the winning party at time  $t$ . In column (1) we report the average ADA at time  $t + 1$  for districts that were Republican at time  $t$  for each of the three groups. In column (2) we report the average ADA score at time  $t + 1$  for districts that were Democrat at time  $t$ . In column (3) we report the difference.

Consider the top and the bottom group. Consistent with the hypothesis of complete divergence, there is virtually no difference in the voting scores of the two sets of districts. Consider the middle group. As expected, there is a large difference in the voting scores of the two sets of districts.

Overall, the results in Table V lead us to conclude that our earlier results are robust to a more general framework that allows for unrestricted heterogeneity in the *differences* in policy positions across districts and over time. Finally, we have estimated similar models using alternative measures of roll-call behavior (DW-Nominate scores and percentage of vote consistent with Democrat leadership). In results available

from the authors, we find similar results: virtually no response of voting records to changes in electoral strength.

UNIVERSITY OF CALIFORNIA, BERKELEY, AND NATIONAL BUREAU OF ECONOMIC  
RESEARCH  
UNIVERSITY OF CALIFORNIA, BERKELEY, AND NATIONAL BUREAU OF ECONOMIC  
RESEARCH  
UNIVERSITY OF CALIFORNIA, BERKELEY

## REFERENCES

- Alesina, Alberto, "Credibility and Policy Convergence in a Two-Party System with Rational Voters," *American Economic Review*, LXXVIII (1988), 796–805.
- Alesina, Alberto, and Stephen E. Spear, "An Overlapping Generations Model of Electoral Competition," *Journal of Public Economics*, XXXVII (1988), 359–379.
- Angrist, Joshua, Guido Imbens, and Donald Rubin, "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, XCI (1996), 444–445.
- Banks, Jeffrey S., and D. Roderick Kiewiet, "Explaining Patterns of Candidate Competition in Congressional Elections," *American Journal of Political Science*, XXXIII (1989), 997–1015.
- Bernhardt, Daniel, and Daniel Ingberman, "Candidate Reputation and the Incumbency Effect," *Journal of Public Economics*, XXVII (1985), 47–67.
- Bender, Bruce A., "The Influence of Ideology on Congressional Voting," *Economic Inquiry*, XXIX (1991), 416–428.
- Besley, Timothy, and Stephen Coate, "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, CXII (1997), 85–114.
- Besley, Timothy, and Stephen Coate, "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis," *American Economic Review*, LXXXVIII (1998), 139–156.
- Besley, Timothy, and Anne Case, "Political Institutions and Policy Outcomes," *Journal of Economic Literature*, XLI (2003), 7–73.
- Calvert, Randall L., "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence," *American Journal of Political Science*, XXIX (1985), 69–95.
- Canes-Wrone, Brandace, David W. Brady, and John F. Cogan, "Out of Step, Out of Office: Electoral Accountability and House Members' Voting," *American Political Science Review*, XCVI (2002), 127–140.
- DiNardo, John, and David S. Lee, "The Impact of Unionization on Establishment Closure: A Regression Discontinuity Analysis of Representation Elections," NBER Working Paper No. 8993, June 2002.
- Downs, Anthony, *An Economic Theory of Democracy* (New York, NY: Harper and Row, 1957).
- Fiorina, Morris P., "Whatever Happened to the Median Voter?" paper prepared for the MIT conference on Parties and Congress, Cambridge, MA, October 2, 1999.
- Groseclose, Timothy J., ADA SCORES, 1947–1999 [Computer File]. <http://faculty-gsb.stanford.edu/groseclose/archive.htm> 12 February 2002.
- Inter-university Consortium for Political and Social Research, and Carroll McKibbin, ROSTER OF UNITED STATES CONGRESSIONAL OFFICEHOLDERS AND BIOGRAPHICAL CHARACTERISTICS OF MEMBERS OF THE UNITED STATES CONGRESS, 1789–1996: MERGED DATA [Computer file], 10th ICPSR ed., Ann Arbor, MI, Inter-university Consortium for Political and Social Research [producer and distributor], 1997.
- Inter-university Consortium for Political and Social Research, UNITED STATES CONGRESSIONAL ROLL CALL VOTING RECORDS, 1789–1990 [Parts 1–202] [Computer file], ICPSR ed., Ann Arbor, MI, Inter-university Consortium for Political and Social Research [producer and distributor], 2001.

- Inter-university Consortium for Political and Social Research, and Congressional Quarterly, Inc., UNITED STATES CONGRESSIONAL ROLL CALL VOTING RECORDS, 1789–1994 [Parts 203–206] [Computer file], ICPSR ed., Ann Arbor, MI, Inter-university Consortium for Political and Social Research/Washington, DC, Congressional Quarterly, Inc. [producers], 1997, Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1997.
- Inter-university Consortium for Political and Social Research, and Congressional Quarterly, Inc., UNITED STATES CONGRESSIONAL ROLL CALL VOTING RECORDS, 1789–1996 [Parts 207–208] [Computer file], ICPSR ed., Ann Arbor, MI, Inter-university Consortium for Political and Social Research/Washington, DC, Congressional Quarterly, Inc. [producers], 1998, Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1998.
- King, Gary, ELECTIONS TO THE UNITED STATES HOUSE OF REPRESENTATIVES, 1898–1992 [Computer file]. ICPSR version, Cambridge, MA, Gary King [producer], 1994, Ann Arbor, MI, Inter-university Consortium for Political and Social Research [distributor], 1995.
- Krehbiel, Keith, "The Coefficient of Party Influence," Stanford Graduate School of Business Research Paper No. 1646, 2000.
- Lee, David S., "The Electoral Advantage to Incumbency and Voters' Valuation of Politicians' Experience: A Regression Discontinuity Analysis of Close Elections," University of California, Berkeley Center for Labor Economics, Working Paper No. 31, April 2001.
- , "Randomized Experiments from Non-Random Selection in U. S. House Elections," University of California, Berkeley manuscript, September 2003.
- Levitt, Steven D., "How Do Senators Vote? Disentangling the Role of Voter Preferences, Party Affiliation, and Senator Ideology," *American Economic Review*, LXXXVI (1996), 425–441.
- Levitt, Steven D., Timothy Groseclose, and James M. Snyder, Jr., "Comparing Interest Group Scores across Time and Chambers: Adjusted ADA Scores for the U. S. Congress," *American Political Science Review*, XCIII (1999), 33–50.
- Lott, John, and Michael Davis, "A Critical Review and Extension of the Political Shirkung Literature," *Public Choice*, LXXIV (1992), 461–484.
- McArthur, John R., and Stephen V. Marks, "Constituent Interests versus Legislator Ideology: The Role of Political Opportunity Cost," *Economic Inquiry*, XXVI (1988), 461–470.
- McCarty, Nolan M., Keith T. Poole, and Howard Rosenthal, *Income Redistribution and the Realignment of American Politics* (Washington, DC: The AEI Press (1997).
- McCarty, Nolan, Keith T. Poole, and Howard Rosenthal, "The Hunt for Party Discipline in Congress," *American Political Science Review*, XCV (2000), 673–687.
- Miller, Warren E., and Donald E. Stokes, "Constituency Influence in Congress," *American Political Science Review*, LVII (1963), 45–56.
- Osborne, Martin J., "Spatial Models of Political Competition under Plurality Rule: A Survey of Some Explanations of the Number of Candidates and the Position they Take," *Canadian Journal of Economics*, XXVIII (1995), 261–301.
- Persson, Torsten, and Guido Tabellini, *Political Economics* (Cambridge, MA: MIT Press, 2000).
- Poole, Keith T., "Changing Minds? Not in Congress!" GSIA WP No. 1997-22, in *Changing Minds*, Michael Munger and Bryan Roberts, eds. (New York: Cambridge University Press, 1999).
- , DW-NOMINATE SCORES FOR HOUSE, 1st–106th SESSION OF CONGRESS [Computer File]. [http://voteview.uh.edu/default\\_nomdata.htm](http://voteview.uh.edu/default_nomdata.htm), 2002.
- , INTEREST GROUP RATINGS, 1959–1981 [Computer File]. <http://voteview.uh.edu/dwnominate.htm>, 2002.
- , RANK ORDER FOR THE HOUSE, 1947–1996 [Computer File]. [http://voteview.uh.edu/default\\_nomdata.htm](http://voteview.uh.edu/default_nomdata.htm), 2002.
- Poole, Keith T., and R. Steven Daniels, "Ideology, Party and Voting in the U. S. Congress 1959–1980," *American Political Science Review*, LXXIX (1985), 373–399.
- Poole, Keith T., and Howard Rosenthal, "The Polarization of American Politics," *Journal of Politics*, XLVI (1984), 102–131.

- Poole, Keith T., and Howard Rosenthal, "A Spatial Model for Roll Call Analysis," *American Journal of Political Science*, XXIX (1985), 331–347.
- Poole, Keith T., and Howard Rosenthal, "Patterns of Congressional Voting," *American Journal of Political Science*, XXXV (1991), 331–347.
- Poole, Keith T., and Howard Rosenthal, *Congress: A Political-Economic History of Roll Call Voting* (Oxford, UK: Oxford University Press, 1997).
- Poole, Keith T., and Howard Rosenthal, "D-Nominate after 10 Years," *Legislative Studies Quarterly*, XXVI (2001), 5–30.
- Snyder, James, and Timothy Groseclose, "Estimating Party Influence in Congressional Roll-Call Voting," *American Journal of Political Science*, XLIV (2000), 187–205.
- James Snyder, and Michael Ting, "Party Labels, Roll-Call Votes, and Elections," mimeo, 2001a.
- James Snyder, and Michael Ting, "An Information Rational for Political Parties," mimeo, 2001b.
- U. S. Department of Commerce, Bureau of the Census, CENSUS OF POPULATION AND HOUSING, 1980 [UNITED STATES]: SUMMARY TAPE FILE 3D CONGRESSIONAL DISTRICT-LEVEL EXTRACT [Computer file]. Milwaukee, WI: John McAdams, Marquette University [producer], 198?. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1987.
- , CENSUS OF POPULATION AND HOUSING, 1990 [UNITED STATES]: SUMMARY TAPE FILE 3D CONGRESSIONAL DISTRICTS OF THE UNITED STATES, 103RD CONGRESS [Computer File]. Washington, DC, 1993.
- , UNITED STATES CONGRESSIONAL DISTRICT DATA BOOKS, 1961–1965 [Computer file]. ICPSR ed., Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 196?
- , UNITED STATES CONGRESSIONAL DISTRICT DATA BOOK FOR THE NINETY-THIRD CONGRESS, 1973 [Computer file]. ICPSR ed., Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1999.
- Wittman, Donald, "Candidate Motivation: A Synthesis of Alternative Theories," *American Political Science Review*, LXXVII (1983), 142–157.



## CALL FOR NOMINATIONS

### \$300,000 Nemmers Prize in Economics

**Northwestern University** invites nominations for the Erwin Plein Nemmers Prize in Economics, to be awarded during the 2026–27 academic year. The prize pays the recipient \$300,000. Recipients of the Nemmers Prize present lectures, participate in department seminars, and engage with Northwestern faculty and students in other scholarly activities.

Details about the prize and the nomination process can be found at [nemmers.northwestern.edu](http://nemmers.northwestern.edu). Candidacy for the Nemmers Prize is open to those with careers of outstanding achievement in their disciplines as demonstrated by major contributions to new knowledge or the development of significant new modes of analysis. Individuals of all nationalities and institutional affiliations are eligible except current or recent members of the Northwestern University faculty and past recipients of the Nemmers or Nobel Prize. **Nominations will be accepted until January 14, 2026.**

*The Nemmers prizes are made possible by a generous gift to Northwestern University by the late Erwin Esser Nemmers and the late Frederic Esser Nemmers.*

**Northwestern**

Nemmers Prizes • Office of the Provost • Northwestern University • Evanston, Illinois 60208  
[nemmers.northwestern.edu](http://nemmers.northwestern.edu)



---

Political business cycles 40 years after Nordhaus

Author(s): Eric Dubois

Source: *Public Choice*, January 2016, Vol. 166, No. 1/2 (January 2016), pp. 235-259

Published by: Springer

Stable URL: <https://www.jstor.org/stable/10.2307/48719654>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*Springer* is collaborating with JSTOR to digitize, preserve and extend access to *Public Choice*

## Political business cycles 40 years after Nordhaus

Eric Dubois<sup>1</sup>

Received: 6 June 2015 / Accepted: 22 January 2016 / Published online: 1 February 2016  
© Springer Science+Business Media New York 2016

**Abstract** The aim of this article is to survey the huge literature that has emerged in the last four decades following Nordhaus's (*Rev Econ Stud* 42(2):169–190, 1975) publication on political business cycles (PBCs). I first propose some developments in history of thought to examine the context in which this ground-breaking contribution saw the light of the day. I also present a simplified version of Nordhaus's model to highlight his key results. I detail some early critiques of this model and the fields of investigations to which they gave birth. I then focus on the institutional context and examine its influence on PBCs, the actual research agenda. Finally, I derive some paths for future research.

**Keywords** Political business cycles · Politico-economic cycles · Electoral cycles · Opportunistic cycles · Conditional political business cycles

**JEL Classification** B22 · D72 · D78 · E32

### 1 Introduction<sup>1</sup>

Forty years ago, William Nordhaus was the first to formalize in an analytical framework the idea that the course of macroeconomic variables is influenced by purely political considerations (Nordhaus 1975). Until this ground-breaking contribution, in most

<sup>1</sup> This paper necessitated reading 530 articles or books. The complete list of references is available upon request. For convenience, I have disregarded unpublished papers (with a few exceptions) and papers not written in English. Despite all our efforts, I was not able to locate approximately 20 published articles or books related to the topic. In some contributions, the study of political business cycles is reduced to a simple electoral dummy introduced as a control variable among others and sometimes not related

---

✉ Eric Dubois  
eric.dubois@univ-paris1.fr

<sup>1</sup> CES, Université de Paris 1, Maison des Sciences Economiques, 106-112, boulevard de l'Hôpital, 75647 Paris Cedex 13, France

academic works, the government was considered in economic models to be a social planner, maximizing a social welfare function which coincides with the utility function of the representative agent in the economy. This behavior of a “benevolent dictator” in Knut Wicksell’s words is inherited from Tinbergen (1952, chapters 1 and 9) and Theil (1958, chapter 8). In Nordhaus’s (1975) theory, governments are driven by private interest and care only about their reelection prospects. They exploit the short-term Phillips curve and benefit from the naïve expectations of voters to attain their goal. As voters are concerned about unemployment, the incumbent improves the probability of being reelected by increasing the inflation rate so that the unemployment rate decreases just before the election. After the election, the government faces a high inflation rate and then implements austerity measures, leading to more unemployment. Unemployment and inflation are thus subject to cyclical fluctuations linked to the rhythm of elections and these fluctuations are called “political business cycles” (PBCs).<sup>2</sup> Without going into a detailed presentation of the history of thought, it is interesting to see how this article emerged and in particular to examine both the forerunners in contributions and the context of the era.

The first trace of PBCs can be found in Akerman (1947), who showed that between 1830 and 1945 in the United States, short-term economic cycles were linked to the four-year presidential election cycle. But his analysis was exclusively empirical and demonstrated no precise causal mechanism. Ten years later, Downs (1957a, 1957b), although not explicitly dealing with cycles in macroeconomic variables, proposed a hypothesis that would serve as a foundation for Nordhaus’s (1975) paper: politicians are driven by private interest. In Anthony Downs’s (1957b, p. 28) words: “We assume that they act solely in order to attain the income, prestige, and power which come from being in office”. This motivation gives rise to an operational objective for politicians: to maximize the number of votes in their favor and win the election. However, Downs says nothing about the macroeconomic variables the government has to manipulate to reach that goal. He simply states that “by means of economic and other actions, [the government] tries to manipulate both present and future utility pay-offs to voters in a way that will win their votes” (Downs 1957b, p. 176). According to Downs (1957a, p. 137), the government “is an entrepreneur

---

Footnote 1 continued

to Nordhaus (1975). I have probably missed some studies like these. Finally, I note that PBCs, explicitly related or not to Nordhaus (1975), have been investigated in somewhat exotic areas (compared to economics): Wasserman (1983) in relation to suicides, Nincic (1990) and Gaubatz (1991) in international relations, Shughart and Tollison (1985), Lagona and Padovano (2008) and Brechler and Geršl (2014) in legislative activity, Block and Vaaler (2004) in ratings published by agencies, Thies and Porche (2007) in agricultural producer protection, Ladewig (2008) in the housing market, Horgos and Zimmermann (2010) in the activity of interest groups, Kächelein et al. (2011) in electricity supply, and Wehner (2013) in legislative budgetary decisions. Unusual applications in the economic area include Paiva (1996) on prices in regulated industries, Castro and Veiga (2004) on the timing of stabilization programs, Dreher and Vaubel (2004) on credits from the IMF and IBRD, Özatay (2007) on public sector prices, Faye and Niehaus (2012) on foreign aid, Sukhtankar (2012) on prices paid to farmers for sugar cane, Lami et al. (2014) on household consumption spending, and Klien (2014) on water tariffs.

<sup>2</sup> Although it is the canonical expression, “political business cycles” can lead to possible confusion as it in fact refers to a particular type of cycle, namely opportunistic. An opportunistic cycle is the result of political manipulations *before* an election. In contrast, partisan cycles appear as a result of manipulation *after* the election. A generic term to designate both opportunistic and partisan cycles could be “politico-economic cycles” or “electoral cycles”. The first author to use the expression political business cycle even if not directly linked to elections was Kalecki (1943, p. 330). Michal Kalecki used this expression to qualify the inflation–unemployment cycle, which reflects fluctuations in the political power struggle between conflicting social classes. Moreover, he was rather silent on the government’s behavior. In his work, the government appears as a passive relay in class struggles, favoring the different classes alternately through stop and go policies.

selling policies for votes instead of products for money". In this view, there is no room for ideology. Politicians treat the possible political choices solely as a means of fulfilling their private objectives, objectives that can only be achieved if they are elected: "parties formulate policies in order to win elections, rather than win elections in order to formulate policies" (Downs 1957b, p. 28). Worried only about reelection, they forget the general interest; the benevolent dictator *à la* Wicksell mentioned earlier vanishes. The private interest hypothesis is certainly not new (see Smith 1776), but it is the first time that it is applied to politicians.<sup>3</sup>

In their two co-written papers, Bruno Frey and Lawrence Lau take up the Downsian hypothesis of self-interest to characterize the government in their model (Frey and Lau 1968; Lau and Frey 1971). Popularity is thus considered a form of pressure upon government that can lead to electoral defeat if neglected.<sup>4</sup>

In 1971, in a speech at a symposium on the appraisal of the Employment Act of 1946, Milton Friedman envisaged the possibility of PBCs under some restrictive assumptions:

...if, in the United States, we had highly precise knowledge of the short-term effects of monetary and fiscal policy, and the authorities were not bound by any rigid and well-enforced rule in the conduct of monetary and fiscal policy, then I would predict with great confidence a four-year cycle, with unemployment reaching its trough in years divisible by four (i.e., presidential election years) and inflation reaching its peak in the following year. (Friedman 1972, p. 196)

Nordhaus's (1975) paper thus comes after several others. The interesting point here is why this particular paper has received so much attention and has occluded most of the previous literature. To identify the reasons for this success, one has to go back to the beginning of the 1970s when William Nordhaus wrote his paper.

In 1972, Nordhaus wrote a first version of his paper and published it in the *Cowles Foundation Discussion Papers* series (Nordhaus 1972). At the Cowles Foundation, he met Gerald Kramer and Ray Fair, who had both shown that economic conditions influence election results (Kramer 1971; Fair 1975). These works are extremely important because they complete those of Downs: as economic outcomes influence the vote, the incumbent who wants to be reelected has an incentive to manipulate the economy in order to gain votes. To some extent, Nordhaus (1975) therefore echoes the works of Gerald Kramer and Ray Fair.

A second reason can be found in the fact that in the 1970s the work of Michal Kalecki received new attention and most of his books were then reprinted (see Boddy and Crotty 1975; Feiwel 1974). Nordhaus (1975) thus emerged in a period when politico-economic cycles were fashionable.

Third, the popularity gained by Nordhaus (1975) can be explained by the fact that he proposed an analytical model that was missing in Akerman (1947) or Downs (1957a, 1957b) and which probably appeared more appealing and trendy than that of Frey and Lau (1968) or Lau and Frey (1971), notably because of the use of the Phillips curve. He also provided an empirical analysis that only Akerman (1947) had conducted before him.

<sup>3</sup> All the same, one notes that the reelection motivation is present, for example, in Tocqueville (1835).

<sup>4</sup> However, in these papers (and in subsequent papers by the same authors and their coauthors), one should note that the concept of cycle is missing, the political pressure being permanent and not only during elections. I will disregard here the literature concerning this "satisficing" version of the PBC model, but this idea will later be taken up and combined in the traditional electoral cycle (see Sect. 3.4).

Fourth, concerning the economic context, William Nordhaus wrote in a period when the instability of macroeconomic variables was increasing, notably with regard to the inflation rate. These fluctuations, when not expected, are sources of uncertainty that penalize investment and undermine growth. Economic scholars were searching for the origins of this instability and Nordhaus (1975) provided an answer: the volatility of inflation comes from electoral manipulations.

Finally, at the beginning of the 1970s, the United States experienced a confidence crisis in its elected representatives that reached its climax in August 1974 with President Nixon's resignation after the Watergate scandal. A cynical vision of manipulator politicians was in the air and Nordhaus (1975) caught that atmosphere.<sup>5</sup>

Having addressed these considerations of the history of thought, this article aims to survey the existing literature on this topic. Of course, many good surveys already exist on PBCs.<sup>6</sup> The aim here is not only to update these studies, but also to complete them, as many aspects developed below are generally not tackled. In Sect. 2, a simplified version of Nordhaus's model, drawn from Artus (1993), is presented with key assumptions and key results. As this model, despite its appealing formalization, led to mixed empirical conclusions, scholars have put forward several critiques, which have opened up new research fields. These main developments are summarized in Sect. 3. Section 4 is devoted to exploration of a particular critique addressed to Nordhaus's model: the fact that the cycle could be conditional, especially conditional on the institutional context. Section 5 envisages some promising directions for future research.

## 2 Nordhaus's (1975) model: a simple presentation of the theory and conclusions of first tests

The economy is characterized by an expectation-augmented Phillips curve:

$$Y_j = \bar{Y} + \gamma(\pi_j - \pi_j^e), \quad (1)$$

where  $Y_j$  is the output level in period  $j$ ,  $\bar{Y}$  is the equilibrium output level,  $\pi_j$  is the inflation rate in period  $j$ ,  $\pi_j^e$  is the expected inflation rate for the period  $j$  and  $\gamma$  is a positive constant.

The government is penalized by both unemployment and inflation. It maximizes a vote function defined as follows:

$$V_j = -l(\tilde{Y} - Y_j)^2 - m\pi_j, \quad (2)$$

where  $\tilde{Y}$  is the target level of output and  $l$  and  $m$  are parameters measuring discontent with unemployment and inflation, respectively. Even if the government does not control it directly, the inflation rate is the action variable of the government. The result of the minimization of (2) can be interpreted as the inflation rate the government would like see implemented; it is the preferred inflation rate of the government.

<sup>5</sup> One might note that other articles or books dealing with the idea of PBCs, sometimes written independently of Nordhaus (1975), exist: Schumpeter (1939, in particular p. 391), Breton (1974, in particular p. 49), Ben-Porath (1975), Lindbeck (1976), Frey and Ramser (1976), Umstead (1977), Wagner (1977) and Lächler (1978).

<sup>6</sup> See de Haan (2013), de Haan and Klomp (2013), Drazen (2000, chapter 7; 2001), Franzese (2000, 2002), Franzese and Long-Jusko (2006), Olters (2001) and Shi and Svensson (2003) for the most recent references.

One supposes that there are two periods, noted  $t$  and  $t + 1$ , and that elections take place at the end of period  $t + 1$ .

The government minimizes the following intertemporal loss function:

$$V = \frac{V_t}{1 + \rho} + V_{t+1}, \quad (3)$$

where  $\rho$  is a “forgetting rate” which is supposed to be  $> 0$ . This kind of discount rate applied to the first period models the fact that individuals have a short memory. At the time of the elections, voters have partially forgotten the first period. In combining (1) and (2) and in setting  $\tilde{Y} - \bar{Y} = \alpha > 0$ , one obtains:

$$V_j = -l \left[ \alpha - \gamma(\pi_j - \pi_j^e) \right]^2 - m\pi_j. \quad (4)$$

and then:

$$V = \frac{-l[\alpha - \gamma(\pi_t - \pi_t^e)]^2 - m\pi_t}{1 + \rho} - l[\alpha - \gamma(\pi_{t+1} - \pi_{t+1}^e)]^2 - m\pi_{t+1}. \quad (5)$$

The key assumption of the model is the characterization of the individuals’ expectations. One supposes here that individuals form their expectations in a naïve way<sup>7</sup>:

$$\pi_j^e = \pi_{j-1}. \quad (6)$$

In incorporating (6) in (5) and minimizing with respect to  $\pi_t$  and  $\pi_{t+1}$ , one obtains the first-order conditions for the government program:

$$\pi_{t+1} = \pi_t + \frac{\alpha}{\gamma} - \frac{m}{2l\gamma^2} \quad (7)$$

$$\pi_t = \pi_{t-1} + \frac{\alpha}{\gamma} - \frac{m(2 + \rho)}{2l\gamma^2} \quad (8)$$

For  $\frac{m}{2l\gamma^2} < \frac{\alpha}{\gamma} < \frac{m(2 + \rho)}{2l\gamma^2}$ , one has the classic Nordhaus (1975) result:

$$\pi_t < \pi_{t-1} \text{ and } \pi_{t+1} > \pi_t. \quad (9)$$

A politico-economic cycle occurs. In the period after the elections, the actual inflation rate is less than expected. Output declines, which leads to higher unemployment (this will be forgotten in the next period). Post-electoral periods are therefore recession periods. In the period before the elections, the actual inflation rate is higher than expected. Output increases, which leads to lower unemployment. Pre-electoral periods are therefore expansion years. Voters are manipulated as they rely on the (low) inflation rate in the previous period to build their expectations.

The first econometric tests searching for PBCs in inflation and/or unemployment led to mixed conclusions. Allen et al. (1986), McGavin (1987), Haynes and Stone (1988, 1989, 1990), Keil (1988) and Findlay (1990) conclude in favor of Nordhaus’s theory, whereas McCallum (1978), Golden and Poterba (1980), Alt and Chrystal (1981), Beck (1982),

<sup>7</sup> Naïve rather than adaptive expectations, as in the Nordhaus model, are used here without loss of generality.

Alesina (1988, 1989) and Davidson et al. (1990) do not. This inconclusive literature has raised some questions about Nordhaus's model, which has been criticized on both theoretical and empirical grounds.

### 3 First critiques, new fields

I present here a list of the main critiques<sup>8</sup> addressed to Nordhaus's model and the developments that issued from these critiques.

#### 3.1 Heterogeneous preferences over inflation and unemployment: partisan cycles

In the Nordhaus model, when inflation increases around the election owing to political manipulations, one might think that it does not affect all voters in the same way as it may penalize those, for example, who are owners of capital. Nordhaus (1975, p. 173) foresees this possibility of a heterogeneous electorate when he states: "It is probably accurate to say that in the United States Republicans have consistently been more concerned about inflation and Democrats about unemployment". Kirschen et al. (1964) are the first to classify economic policy goals according to the ideology of political parties. They consider, for example, that full employment is the first priority of socialist or labor parties, whereas price stability is the chief objective of conservative parties. In Frey and Lau (1968), besides the reelection constraint, there is an internal pressure on the government, namely ideological satisfaction. The government therefore faces a trade-off as both vote maximization and ideological satisfaction are entered as arguments in the "felicity function" that the government seeks to maximize. This idea of a dual behavior would be taken up by Lau and Frey (1971) and empirically tested by several papers at the end of the 1970s and the beginning of the 1980s, the more frequently quoted papers being Frey and Schneider (1978a, 1978b). Finally, in 1977, Donald Wittman demonstrates analytically that parties with policy preferences as well as private interests do not converge (Wittman 1977; completed by Wittman 1983). This goes against Downs (1957b), who argues that parties in competition tend to move closer to a median platform in order to capture the maximum number of voters. Donald Wittman's works therefore provide a foundation for partisan cycles, developed by Hibbs (1977). According to Hibbs's theory—and unlike Nordhaus (1975)—voters have heterogeneous preferences and therefore parties have different ideological goals. The usual way of accounting for these differences between parties (and therefore governments) is to consider that they do not have the same ideal positions on the short-term Phillips curve. Specifically, conservative governments prefer lower inflation and therefore higher unemployment than liberal governments. The changeover of political

<sup>8</sup> For example, the specification of Nordhaus's (1975) vote function has been criticized (e.g., Frey and Ramser 1976; Paldam 1981). I do not develop this point here because the debate about vote function was older and the critique of Nordhaus (1975) did not give rise to a new strand of literature in this area. However, this issue is related to the effectiveness of PBCs. Do the PBCs pay off in terms of votes? Examining the impact of pre-election budget deficit on reelection probability, Brender and Drazen's (2008) conclusion is negative. Aidt et al. (2011) challenge this view and find that greater expenditures in the election year lead to greater vote differences between the incumbent and the main opponent (see also Veiga and Veiga 2007b, in this line). Moreover, some refinements of the Nordhaus model are not shown here because they have not yet given rise to further developments.

power between conservative and liberal governments leads to “partisan cycles”.<sup>9</sup> Partisan theory would later be developed and tested by, among many others, Chappell and Keech (1986), Alesina (1987, 1988, 1989), Alesina and Roubini (1992), Alesina et al. (1992, 1993, 1997) and Hibbs (1994).

### 3.2 Rational expectations: rational political business cycles

In the same year and the same month (April 1975) that William Nordhaus’s article was published, an article written by Thomas Sargent and Neil Wallace launched what was called “the rational expectations revolution” (Sargent and Wallace 1975). Some even spoke about “evangelism” (Price 1997, 408). The impact on PBC theory was so important that Alberto Alesina (1988, p. 15) would speak about a presumed “devastating effect”. Indeed, if people are able perfectly to anticipate government policy, there is no longer a politico-economic cycle in the Nordhaus (1975) sense. Minimizing the loss function leads the government to set an inflation rate equal to zero for both periods, thus making any expansive policy impossible; PBCs disappear.<sup>10</sup>

More than 10 years would pass by before scholars tried to reconcile Nordhaus’s (1975) model with rational expectations theory. The two main contributions here are Rogoff and Sibert (1988) and Rogoff (1990).<sup>11</sup> These articles no longer envisage an inflation/unemployment trade-off but a state budget problem. Moreover, temporary information asymmetry between voters and governments is substituted for the hypothesis of non-rational voters. Voters are rational but do not have complete information about the government’s competence; they observe it with a delay. Competence is defined as the capacity to provide public goods and transfers in an efficient way (i.e., by limiting waste). To appear competent, the government pursues an expansive budgetary policy while temporarily hiding the harmful consequences from the agents (the increase in the budget deficit). By doing so, the government creates the illusion of prosperity until the time when voters realize that they will have to “reimburse” by one means or another the deficit thus caused.

Generally, following a distinction first introduced by Persson and Tabellini (2000), these models are labelled adverse selection models because the government has an informational advantage over voters concerning its true competence. Lohmann (1998a) is the first to relax this assumption. Both voters and the government have the same information about the government’s competence. However, if voters no longer face an adverse selection problem, they face a moral hazard problem. The incumbent can indeed modulate his or her level of effort and appear better than he or she really is by using policy instruments unobservable to the public.

In taking advantage of information asymmetry and/or trying to appear as competent as possible before the election, politicians behave as in Nordhaus’s model. His conclusions

<sup>9</sup> One notes, however, that in partisan theory, the cycle is involuntary: the electoral calendar is a source of uncertainty and disrupts the agents who, not knowing which policy will be implemented after the elections, anticipate an inflation rate between the parties’ preferred rates. After the elections, expectations necessarily turn out to be false, which generates a cycle.

<sup>10</sup> Another way, as chosen by Nordhaus (1989), is to defend adaptive expectations by showing that people do not have rational expectations. Early critics of Nordhaus’s view of expectations include, among others, McCallum (1978), Keech (1980) and Paldam (1981). In these studies, voters’ sophistication and memory are also debated.

<sup>11</sup> Cukierman and Meltzer (1986) are also often quoted, but if they emphasize the role of information asymmetries, they do not deal explicitly with PBCs. Persson and Tabellini (1990) translate Rogoff (1990) into an inflation-unemployment setting.

hold even when voters are not short-sighted and naïve as long as they are poorly informed about the environment, the government's objectives and/or its ability to manage the economy.

### 3.3 Instruments versus outcomes

As one has seen above, two key variables in the Nordhaus model are the unemployment and inflation rates, outcome variables that are not actually under the direct control of the government. This could be one of the reasons why PBCs are not detected in these kinds of variables. Tufte (1978) is the first to address the question of the means of action available to the government to generate economic expansion before the elections. According to him, “the economic policy instruments involved must be easy to start-up quickly and must yield clear and immediate economic benefits to a large number of voters” (Tufte 1978, p. 9). An increase in real disposable income seems to be a good candidate. Indeed, through taxes and transfers, it can quickly be manipulated and with much greater ease and flexibility than unemployment, inflation or output growth. If a cycle actually exists, it has to be sought not in the outcomes of economic policy, but rather in the instruments of economic policy, which are by definition more directly controllable.

Empirical tests of the PBC in instruments are much more convincing so that nowadays, tests in relation to outcomes are scarce.<sup>12</sup> It is impossible to cite all the hundreds of articles that seek PBCs in budgetary and monetary variables.<sup>13</sup> The most complete studies in terms of the sample include Ames (1987), Remmer (1993), Neto and Borsani (2004), and Nieto-Parra and Santiso (2012) for Latin America, Block (2002) for Africa, Hallerberg et al. (2002) for East Europe, Alesina et al. (1997), Mink and de Haan (2006) and Klomp and de Haan (2013a, 2013c) for industrial democracies and Schuknecht (1996, 1999, 2000) for developing countries.

Of course, the underlying hypothesis is that the government has perfect control over policy instruments. This issue will be examined in Sect. 4.

### 3.4 The political and economic context

Many scholars have emphasized that the context of an election has an influence on the PBC. Clearly, if the political and economic conditions are considered good by the agents, success is certain for the incumbent government and there is no need to generate a PBC (Dinkel 1981).<sup>14</sup> Two main indicators have been privileged in assessing the electoral context and therefore measuring the risk of a government losing the election.

<sup>12</sup> One should note, however, that an old and persisting literature exists that highlights the presence of PBCs in stock markets. For early references, see Umstead (1977) and Allvine and O'Neill (1980); for more recent references, see Sturm (2013) and Kräussl et al. (2014).

<sup>13</sup> The legitimacy of tests in terms of instruments rather than outcomes is reinforced by the fact that instruments are the action variables in rational PBC models. Besides all of their properties, their use for political purposes is, moreover, grounded theoretically. One also notes that budgetary variables are privileged over monetary variables because of the conjunction of two elements: (1) for one-country studies, there is a need for a degree of freedom which implies resorting to local data; (2) monetary policy is centralized at the national (or supra-national) level and is therefore not an instrument available to local governments.

<sup>14</sup> Conversely, when the incumbent expects to lose the election, it may prefer to use the debt strategically instead of generating a PBC. In doing so, it can constrain the next government by reducing the latter's room for maneuver. In Boroohah and van der Ploeg (1983), the “kamikaze” government, knowing it will lose power, pursues a ruinous economic policy, penalizing its successor to the point that the latter will not be reelected, which puts the suicidal government back at the helm. While Alesina and Tabellini (1990) predict

The first indicator is the level of popularity. Schultz (1995) was the first to amend the traditional PBC dummy variable to take account of the popularity of the government, the latter being measured as the differential between the governing party and the main opposition party. The test showed that whereas the unconditional electoral dummy and permanent popularity lead (in Frey and Schneider's style) are not significant, the product of both is. This conclusion would be confirmed by Carlsen (1997) in relation to re-election probability, Price (1998) for popularity level, Alt and Rose (2007) for the level of approval, Efthyvoulou (2012) in terms of voting intention polls and Klomp and Haan (2013b) using a measure of popular protest.<sup>15</sup>

The second indicator is the level of electoral competitiveness. When an election is not competitive, the incumbent has no incentive to generate a PBC. By splitting their sample between high competition states and low competition states, Baber and Sen (1986) show that the PBC in the situation of debt is stronger in the former. Clingermayer and Wood (1995) seem to have been the first to substitute an index of electoral competition for the traditional PBC dummy. They conclude that this index has a significant impact on debt before the election: debt is higher in states with strong interparty competition in the year preceding an election. The presence of such an effect is found in Vergne (2009), Aidt et al. (2011) and Benito et al. (2013b), but is rejected by Chaudhuri and Dasgupta (2005) and Schneider (2010).<sup>16</sup>

The threat of losing an election is only one aspect of the context. Some authors have suggested that the identity of the party in power could have an impact on the PBC. Swank and Swank (1993) show that in the United States, Republican administrations aim to reduce tax rates when elections are approaching, whereas Democratic administrations are more concerned with inflation and unemployment. Partisan differences are found in Haynes and Stones (1990), Kneebone and McKenzie (2001), Krause (2005), Aidt et al.

---

Footnote 14 continued

an increase in debt whatever the party. Persson and Swensson (1989) highlight partisan differences according to the preferences of each party. The lowest spending party reduces taxes and in doing so, increases debt, thus forcing the next party to spend less. However, the highest spending party raises taxes, which creates a surplus and makes the next party's policy easier to implement. This theory was tested with success by Pettersson-Lidbom (2001). In Tabellini and Alesina (1990), as the preferences of the next majority concerning the composition of expenditures are uncertain, the current median voter prefers higher debt to be sure that the next majority will have sufficient scope to implement its desired policy. Finally, in Aghion and Bolton (1990), the conservative party accumulates deficits when in power and voters reelect it as they deem the opposition (liberal) party unable to fight against deficits.

<sup>15</sup> Popularity can be considered a proxy for economic conditions, which can also have a direct impact on PBC. In Efthyvoulou (2012), the magnitude of a PBC is weaker when non-economic voting (measured by replies to a survey about households' financial situation) is high. In Hanusch (2012), economic disturbances blur competence signals, which dampen PBCs. In Canes-Wrone and Park (2012), the traditional PBC is invisible because it is cancelled by another cycle in private investment which drops before the election owing to electoral uncertainty. In Alpanda and Honig (2010), the government is constrained in generating a PBC by its borrowing capacity. Finally, endowment in natural resources can play a role (Vergne 2009).

<sup>16</sup> This can be related to the growing literature on leviathan behavior (Brennan and Buchanan 1980), in which politicians are viewed as power-maximizing agents whose only purpose is to maximize the size of the public sector. In this literature, the electoral constraint modelled by the margin (the difference in vote shares or the distance from a 50 % vote share) affects budget outcomes. Ultimately, the empirical test looks very much like a classic PBC test in which the margin is substituted for the traditional electoral dummy. However, the interpretation is different. In the leviathan model, the margin is expected to be related positively to expenditure and taxes. The larger the margin, the greater the government's capacity to increase expenditures and taxes (Dubois et al. 2007; Solé-Ollé 2003, 2006). In the PBC model, the larger the margin, the lower the incentive to intervene to secure reelection and the lower the expenditures and the higher the tax.

(2011), Garcia-Sanchez et al. (2011) and Benito et al. (2013b), but not in Grier (1987) and Haynes and Stone (1989). Besides the identity of the party, the distance between parties' platforms (polarization) can also play a role. The larger the distance, the greater the utility loss from losing office and the greater the manipulation the incumbent is willing to engage into be reelected (Alt and Lassen 2006; Klomp and de Haan 2013c). Finally, other dimensions of the context include political alignment or affiliation (i.e., when local and national executives belong to the same party) as shown in Lema and Streb (2013) and Rumi (2014), the strength of political parties (Shelton 2014) and the strength of the incumbent government, a divided or fragmented government being less able to generate a PBC because of coordination costs (Alt and Rose 2007; Rose 2008; Klomp and de Haan 2013c; but see Geys 2007, for a challenging view).

## 4 Taking into account the institutional context: conditional political business cycles

I focus here on three features of the institutional context: the electoral calendar, the governance system, regime and rule, and the constraints on economic policy.<sup>17</sup>

### 4.1 The electoral calendar: length of terms and endogenous elections

In Nordhaus's (1975) model, the institutional context appears as a watermark throughout the lengths of office terms. For Nordhaus (1975), reducing the length of terms or the number of terms imposes a time constraint on the government, which is no longer able to generate a cycle (see, in a similar vein, Besley and Case 1995; Streb 1999; Grier and Grier 2000; Bordignon et al. 2003; Rose 2006, 2008; Alt and Rose 2007; Klein 2010; Aidt et al. 2011; Benito et al. 2013b; Sjahrir et al. 2013; Klein and Sakurai 2015).<sup>18</sup> For Amacher and Boyes (1978), reducing the frequency of elections reduces their cost. At the extreme, an infinite length of the electoral cycle could enable such a result (MacRae 1977). Lindbeck (1976) suggests choosing the date of elections randomly. In both cases, if the date of the next election is not known, the government cannot anticipate when it has to generate the cycle. However, these options are not compatible with a democratic system. On the optimal length of the electoral term, see also Chappell and Peel (1979), Ginsburgh and Michel (1983), Keech and Simon (1983), Baleiras and Santos (2000) and Ferris and Voia (2009).

<sup>17</sup> Several other institutional variables that may exert an impact on the magnitude of PBCs have been studied: the quality of institutions by Shi and Svensson (2006), government transparency by Akhmedov and Zhuravskaya (2004), corruption by Shi and Svensson (2006), globalization by Eftihyvoulou (2011), social capital measured by the level of generalized trust by Kouvaras (2013), direct versus indirect elections by Sjahrir et al. (2013) and the degree of decentralization (Vergne 2009; Shelton 2014). Among these miscellaneous issues, the special role of the media emerges. As one has seen, information is crucial in the formation of PBCs as informational asymmetries could be at their heart. The better informed the voters, the better their evaluation of government competence, which diminishes the incentives for the government to generate a PBC. Shi and Svensson (2006), Akhmedov and Zhuravskaya (2004), Alt and Rose (2007), Vergne (2009) and Shelton (2014) have found that voters' access to media or media freedom does indeed have an impact on the magnitude of a PBC.

<sup>18</sup> Some of these studies do not deal explicitly with term limits but investigate an issue that is quite close, i.e., what happens to the PBC when the incumbent is seeking reelection (on the understanding that it is allowed to do so).

Besides the considerations linked to term length, one of the hypotheses of the Nordhaus (1975) model is the exogeneity of the electoral calendar. Elections are held on dates set by law, known to everyone. Frey and Schneider (1978b) and Lächler (1982) were among the first to remark that this hypothesis is not valid for all countries, as in Japan and the United Kingdom, for example, elections are organized by the government on the date desired (as long as it falls before the constitutional or statutory “sell by” date, e.g., six years in the UK). Elections are thus labelled “endogenous”. This institutional feature has an important implication for Nordhaus’s (1975) theory: rather than generate an economic cycle at the moment of the elections, the government will call elections in an expansionary economic period to ensure its reelection. As Inoguchi (1979) puts it, they “surf” on the economic cycle. It is no longer the economic cycle that follows the political cycle but the reverse. This can be viewed as a means of reducing distortions as the government no longer has to manipulate the business cycle but only to choose the election date with care (Terrones 1989). Nothing is changed in empirical terms as an election should be observed during an economic expansion, whereas in the exogenous version, an economic expansion should be observed in the run-up to an election. Finally, one can imagine a combination of both in which the government generates an economic expansion and then chooses the election date in consequence.

Following Inoguchi (1979), numerous empirical studies have been undertaken in which the occurrence of elections is explained by, among other determinants, the economic conditions: Ito and Park (1988), Ito (1990, 1991), Kohno and Nishizawa (1990), Cargill and Hutchison (1991), Suzuki (1994) and Heckelman and Berument (1998) for Japan, Annett (1993) for Ireland, Chowdhury (1993) and Khemani (2004) for India, Balke (1990), Heckelman and Berument (1998) and Smith (2003, 2004) for the United Kingdom, Reid (1998), Dickson et al. (2013) and Voia and Ferris (2013) for Canada and Telatar (2003) and Karakas (2014) for Turkey.<sup>19</sup> In their panel study, Shi and Svensson (2006) reject a direct effect of the timing of the election date on the magnitude of the PBC observed for the surplus. References that distinguish between pre-determined elections and endogenous elections in the detection of PBCs include Brender and Drazen (2005), Vergne (2009), Katsimi and Sarantides (2012) and Ehrhart (2013).

Finally, it has been shown that a political business cycle in a particular country could generate cross-country spillovers (Tufte 1978; Thompson and Zuk 1983; Ito 1991; Dobson and Dufrene 1993; Foerster 1994; Foerster and Schmitz 1997; Katada and McKeown 1998; Sapir and Sekkat 1999, 2002; Nippani and Arize 2005). To eliminate such imported cycles, some scholars wonder about the opportunity to synchronize electoral calendars and to set a uniform election date for all countries belonging to the same geographic area (Sapir and Sekkat 1999; Breuss 2008).

## 4.2 System, regime and rule

Regarding the system, the impact of two dimensions on PBC has been studied: an authoritarian versus democratic system and new versus old democracy. In a quite intuitive way, in a country where the incumbent is the sole candidate, the probability of losing power is nil and there is therefore no incentive to generate a PBC. Block et al. (2003) find support for this hypothesis. Nevertheless, it has been shown that PBCs occur even in non-democratic systems: in Egypt (Blaydes 2011), in Mexico (Gonzalez 2002), in China (Guo

<sup>19</sup> See also the panel data studies of Alesina et al. (1997) and Palmer and Whitten (2000). It should be noted that besides these empirical studies, analytical models of the choice of election date have been proposed (see, among others, Smith 1996; Kayser 2005).

2009), in Malaysia (Pepinsky 2007) and in the Republic of Korea (Soh 1988). The idea is that even in authoritarian systems, political leaders have to be sufficiently popular to avoid contestation and removal.

Block et al. (2003) hypothesize that initial or founding multiparty elections would present both the greatest incentives for and the fewest constraints on electoral economic manipulation. On the one hand, through the entry of challengers, an election introduces uncertainty for the authoritarian incumbent, who is therefore more likely to generate a PBC. On the other hand, in founding elections, there is little chance of a free press or other institutional constraints. In a similar vein, according to Brender and Drazen (2005), PBCs are more pronounced in new democracies because of lack of experience with electoral politics, lack of information available to voters, or both. This hypothesis is confirmed by Brender and Drazen (2007) and Klomp and de Haan (2013b, c) but rejected by Hagen (2007) and Barberia and Avelino (2011). Other articles on this topic include Veiga and Pinho (2007), Efthyvoulou (2012) and Shelton (2014). Although the debate has focused on old versus new democracy issues in recent years, some studies show that democracy reduces PBCs, mostly because it is a proxy for other factors that may mitigate PBCs, such as governmental transparency or the media (Akhmedov and Zhuravskaya 2004).

Persson and Tabellini (2003) were the first to study the impact of the political regime (parliamentary versus presidential) and electoral rule (proportional versus majority) on PBCs (see also Chang 2008; Efthyvoulou 2012; Klomp and de Haan 2013c; Shelton 2014). They expect different patterns in PBCs because, according to them, majoritarian elections are associated with stronger individual accountability than proportional elections, in which politicians are more accountable collectively. Moreover, proportional electoral rules give politicians stronger incentives to seek support by endorsing broad policy programs, such as welfare state spending, whereas majoritarian electoral rules instead induce them to target spending to smaller geographic groups (marginal districts with more swing voters). As individual political accountability supplies greater incentives to manipulate than collective accountability, stronger electoral cycles are expected under presidential regimes as decisions are concentrated in the hands of a few decision makers. Linked to the voting rule, Aidt and Mooney (2014) use the natural experiment of a change in suffrage (taxpayer suffrage *versus* universal suffrage) to study how it affects the PBC. If the PBC is present under both rules, all fiscal variables are not affected in the same way. They observe a reduction in tax under taxpayer suffrage (as the right to vote is linked to the obligation to pay local taxes) and an increase in capital spending under universal suffrage (as all adult residents can vote, which creates a larger demand for public goods).

### 4.3 Constraints on economic policy

In Nordhaus's (1975) model, the action variable of the government is the inflation rate. If one adopts a monetarist view of inflation, separating the monetary authority from the political power (for example, through an independent central bank<sup>20</sup>) should be a means of eliminating PBCs.<sup>21</sup> Most of the empirical literature draws this conclusion (Clark and

<sup>20</sup> In fact, through all of the settings that reduce inflationary bias: implementation of a rule (Barro and Gordon 1983), contract between the government and the central bank (Walsh 1995; see al-Nowaihi and Levine 1998, for the impact of such a contract on PBCs) or the appointment of a conservative central banker (Rogoff 1985).

<sup>21</sup> One might ask why the government could choose to give up the monetary instrument which could be so important in generating PBCs and therefore helpful in reelection. First, monetary policy would be less useful because it is less adapted to targeted interventions toward groups of voters. Fiscal instruments are more

Reichert 1998; Hadri et al. 1998; Gärtner 1999; Clark and Hallerberg 2000; Leertouwer and Maier 2001; Hallerberg et al. 2002; Hiroi 2009; Alpanda and Honig 2009, 2010).

However, some case studies cast doubt on this general conclusion as in countries such as Germany or the United States, PBCs seem to survive even when the monetary policy is conducted by an “independent” central bank. Numerous studies have shown the presence of monetary PBCs in the United States (Grier 1987, 1989; Williams 1990; Carlsen 1997; Abrams and Iossifov 2006) and in Germany (Johnson and Siklos 1994; Vaubel 1997; Lohmann 1998b). How can the Federal Reserve Bank and the Bundesbank, reputed to be the two most independent central banks in the world generate a PBC? An answer lies in the distinction between “*de jure*” or “legal” independence, i.e., that warranted by the central bank status, and “*de facto*” or “actual” independence, i.e., that which can be derived from observation of the central bank behavior (see, in particular Cukierman 1992, on this distinction). The difference observed between these two types of independence could be the result of pressures directly exerted by the government (Abrams 2006; Abrams and Butkiewicz 2012), or indirectly through interventions in the media presenting the government’s side (Havrilesky 1988; Froyen et al. 1997; Maier et al. 2002). Government can also control the independent central bank through appointments to the board: see Havrilesky and Schweitzer (1990), Gildea (1990) and Havrilesky and Gildea (1992) for the United States, and Sieg (1997) and Lohmann (1998b) for Germany. Other scholars argue that PBCs detected in monetary instruments are the results of PBCs in fiscal instruments accommodated by the monetary authority (Laney and Willett 1983; Allen 1986; Beck 1987; Allen and McCrickard 1991).

Besides an independent central bank, other monetary institutions may have an impact on the magnitude of the PBC. Among these, the foreign exchange regime is one the most studied (Clark and Reichert 1998; Schuknecht 1999; Clark and Hallerberg 2000; Leertouwer and Maier 2001; Hallerberg et al. 2002; Hiroi 2009; Alpanda and Honig 2010; Shelton 2014). The argument is that a fixed exchange rate regime limits the autonomy of domestic monetary policy and therefore prevents monetary manipulations for electoral purposes. Indeed, in a fixed exchange rate regime, attention is focused on maintaining exchange rates, which may imply the use of monetary tools in a way that the government considers undesirable for reelection prospects (e.g., monetary contraction, upward adjustment of interest rates). A complementary implication is that in this case, government focuses on fiscal policy, so that a PBC in fiscal policy is more likely to happen in a fixed exchange rate regime.

Up to now, I have considered constraints on monetary instruments. Clearly, several constraints also limit the use of fiscal instruments. Various fiscal rules can reduce PBCs: balanced budget requirements or budget targets (Hallerberg and von Hagen 1998; Clark and Hallerberg 2000; Rose 2006; Alt and Rose 2007; Donahue and Warin 2007; Schneider 2010; Garcia-Sánchez et al. 2011; Benito et al. 2013a; Klomp and de Haan 2013c), delegation to a strong finance minister (Hallerberg and von Hagen 1998; Clark and

---

Footnote 21 continued

easily manipulated and have more immediate and clear impacts. Political gains from inflation are relatively blurred. Second, the adoption of an independent central bank could be a sign of good will towards the financial market to obtain more credibility and a better reputation. Credibility and reputation are crucial to obtain the confidence of the financial market and this confidence should help finance the public debt, which is also a tool for generating PBCs. Third, adopting an independent central bank could reduce volatility; volatility generates uncertainty and penalizes investment and therefore economic growth, which is the key indicator in the voters’ eyes.

Hallerberg 2000), checks and balances<sup>22</sup> (Chang 2008; Streb et al. 2009, 2012, 2013; Ho and Huang 2013; Shelton 2014) and transparency<sup>23</sup> (Alt and Lassen 2006; Alt and Rose 2007; Vicente et al. 2013; Shelton 2014).

Because of all of these constraints, it is increasingly difficult for governments to pursue all-around and indiscriminate fiscal policies for electoral purposes. A solution is to play with the distribution rather than with the amount of spending. This need is reinforced by the fact that targeted spending also yields more visible results. Numerous studies have tried to detect a PBC in targeted spending.<sup>24</sup> The specific components most studied are public employment (Spafford 1981; Levitt 1997; Katsimi 1998; Ames et al. 2005; Coelho et al. 2006; Gamez and Ibarra-Yunez 2007; Remmer 2007), public wages (Borjas 1984; Klein 2004; Gamez and Ibarra-Yunez 2007), defense (Nincic and Cusack 1979; Cusack and Ward 1981; Griffin et al. 1982a, 1982b; Mintz and Hicks 1984; Zuk and Woodbury 1986; Kamlet and Mowery 1987; Mintz and Ward 1988; Mintz 1988; Su et al. 1993), culture (Getzner 2002; 2004; Noonan 2007; Nogare and Galizzi 2011; Benito et al. 2013b) and infrastructure (Blais and Nadeau 1992; Chaudhuri and Dasgupta 2005; Goeminne and Smolders 2014). More global studies investigate PBCs in several spending components to discriminate those privileged by the incumbent: for example, van Dalen and Swank (1996), Kneebone and McKenzie (2001), Galli and Rossi (2002), Serritzlew (2005), Veiga and Veiga (2007a, 2007b), Vergne (2009), Drazen and Eslava (2010), Katsimi and Sarantides (2012), Enkelmann and Leibrecht (2013), Brender and Drazen (2013). Most of these studies conclude in favor of PBCs in most visible components. However, one has to be cautious about what “visible” means here. For example, even if both imply an increase in education spending, building a school is visible whereas raising teachers’ wages is not (except to teachers). There could be also a problem of accountability depending on which elections and which aspects of spending are considered. For example, building a hospital may not favor the reelection of a local incumbent because voters hold the national government responsible for such spending.

## 5 Paths for future research

Can one derive some promising directions for future research on PBCs? Looking at past surveys, this exercise, even if traditional, appears to be quite perilous. Indeed, with the notable exception of Chortareas (1999), no scholar has been an accurate visionary in this respect. In his conclusion, Chortareas (1999, p. 137) states “The institutional structure within which the PBC is typically studied is assumed given. Institutions change and how this change affects the PBC is not usually discussed. Such analysis may provide valuable insights”. Launching ourselves into the breach, I propose a research agenda for the next two decades. However, to be prudent, I limit myself to proposing a collection of slight improvements. First, although some (national) studies already exist (Lema and Streb 2013;

<sup>22</sup> Usually measured by an index of the political constraints facing the executive when implementing policy.

<sup>23</sup> Usually measured by an index coding several aspects of the budget process (e.g., budget cycle frequency, the existence of multi-year expenditure forecasts, the requirement of published performance measures). In a somewhat related approach, Mourao (2008) shows that fiscal illusion, in hiding the real fiscal situation, aggravates the PBC. Finally, financial transparency should not be mistaken for transparency understood as the absence of corruption, which can also influence the PBC (Shi and Svensson 2006, Klomp and de Haan 2013c).

<sup>24</sup> Similarly, some scholars have studied PBCs in targeted taxes (e.g., Karran 1985; Ashworth and Heyndels 2002; Chaudhuri and Dasgupta 2005; Ehrhart 2013).

Rumi 2014), more attention could be paid to partisan vertical affinities. Indeed, one might think that PBCs at the local level would be reinforced if the local government shared the same political options as the national government. Second, it could be interesting to examine if PBCs are stronger in the case of simultaneous elections (see Veiga and Veiga 2007a, for a first attempt in Portugal). It would be logical as the stakes are more important. The US case, with presidential and congressional elections held in the same year (compared to midterm elections), or the French case, with presidential and legislative elections held in the same year since 2002 (compared to past elections), could be good fields of investigation. Third, even after four decades, one still knows very little about the duration of electoral effects. How long does the influence of political manipulation last? One quarter? One year? There is essentially no answer to this question. There is also no reason to think that this duration is fixed and cannot change from one country to another, or within the same country from one election to another, and/or from one instrument to another. Scholars who believe in PBCs tend to “force” the tests and select the duration for which the electoral dummy is (the most) significant. For transparency purposes, I militate in favor of generalizing practice, possibly through a table in an annex, to show the significance of the electoral variable for different durations. Fourth, the personal characteristics of candidates can alter the shape of the PBC. The first insights are given by Brollo and Troiano (2013) and Alesina et al. (2015). Brollo and Troiano (2013) find that, when elections are approaching, even if there is no gender difference in transfers and health care services, male mayors hire relatively more temporary public employees than female mayors. Alesina et al. (2015) show that younger politicians are more likely to increase expenditures in pre-election years. We can note that these personal characteristics may also have an indirect impact on PBCs by influencing the probability of reelection and therefore the incentives to generate PBCs. Fifth, although most key economic variables have been studied, one might note that no PBC has been sought in relation to income or wealth inequality. If one disregards the fact that it is an outcome, this variable is interesting for at least two reasons. First, there is no consensus as to whether inequalities are good for the economy. Second, in some countries (e.g., France), inequalities are a strong political marker, which leads us to think that different parties will not generate PBCs with the same pattern. Sixth and finally, changes in the definition of the population of registered voters could influence the PBCs. For example, switching from a system of suffrage based on a tax threshold to universal suffrage (as in Aidt and Mooney 2014), allowing women to vote, or lowering the voting age, are natural experiments through which modifications in the PBC could be studied. Indeed, one might consider that these changes could modify the structure of spending by shifting the target according to the new voters’ characteristics.

**Acknowledgments** I thank William F. Shughart II and the anonymous reviewers for their helpful comments and suggestions.

## References

- Abrams, B. A. (2006). How Richard Nixon pressured Arthur Burns: evidence from the Nixon tapes. *Journal of Economic Perspectives*, 20(4), 177–188.
- Abrams, B. A., & Butkiewicz, J. L. (2012). The political business cycle: new evidence from the Nixon tapes. *Journal of Money, Credit and Banking*, 44(2–3), 385–399.
- Abrams, B. A., & Iossifov, P. (2006). Does the Fed contribute to a political business cycle? *Public Choice*, 129(3–4), 249–262.

- Aghion, P., & Bolton, P. (1990). Government domestic debt and the risk of default: a political-economic model of the strategic role of debt. In R. Dornbusch & M. Draghi (Eds.), *Public debt management: theory and history* (pp. 315–349). Cambridge, MA: Cambridge University Press.
- Aidt, T. S., & Mooney, G. (2014). Voting suffrage and the political budget cycle: evidence from the London metropolitan boroughs 1902–1937. *Journal of Public Economics*, 112, 53–71.
- Aidt, T. S., Veiga, F. J., & Veiga, L. G. (2011). Election results and opportunistic policies: a new test of the rational political business cycle model. *Public Choice*, 148(1–2), 21–44.
- Akerman, J. (1947). Political economic cycle. *Kyklos*, 1, 107–117.
- Akhmedov, A., & Zhuravskaya, E. (2004). Opportunistic political cycles: test in a young democracy setting. *Quarterly Journal of Economics*, 119(4), 1301–1338.
- Alesina, A. (1987). Macroeconomic policy in a two-party system as a repeated game. *Quarterly Journal of Economics*, 102(3), 651–678.
- Alesina, A. (1988). Macroeconomics and politics. In S. Fischer (Ed.), *NBER macroeconomics annual* (pp. 13–52). Cambridge, MA: MIT Press.
- Alesina, A. (1989). Politics and business cycles in industrial democracies. *Economic Policy*, No, 8, 57–99.
- Alesina, A., Cohen, G., & Roubini, N. (1992). Macroeconomic policy and elections in OECD democracies. *Economics and Politics*, 4(1), 1–30.
- Alesina, A., Cohen, G., & Roubini, N. (1993). Electoral business cycle in industrial democracies. *European Journal of Political Economy*, 9(1), 1–23.
- Alesina, A., & Roubini, N. (1992). Political cycles in OECD economies. *Review of Economic Studies*, 59(4), 663–688.
- Alesina, A., Roubini, N., & Cohen, G. (1997). *Political cycles and the macroeconomy*. Cambridge, MA: MIT Press.
- Alesina, A., & Tabellini, G. (1990). A political theory of fiscal deficits and government debt in a democracy. *Review of Economic Studies*, 57(3), 403–414.
- Alesina, A., Troiano, U., & Cassidy, T. (2015). Old and young politicians. Unpublished working paper.
- Allen, H. V. (1986). The Federal Reserve and the electoral cycle. *Journal of Money, Credit and Banking*, 18(1), 88–94.
- Allen, S. D., & McCrickard, D. L. (1991). The influence of elections on Federal Reserve behaviour. *Economics Letters*, 37(1), 51–55.
- Allen, S. D., Sulock, J. M., & Sabo, W. A. (1986). The political business cycle: how significant? *Public Finance Quarterly*, 14(1), 107–112.
- Allvine, F. C., & O'Neill, D. E. (1980). Stock market returns and the presidential election cycle. *Financial Analysts Journal*, 36(5), 49–56.
- al-Nowaihi, A., & Levine, P. (1998). Can political monetary cycles be avoided? *Journal of Monetary Economics*, 42(3), 525–545.
- Alpanda, S., & Honig, A. (2009). The impact of central bank independence on political monetary cycles in advanced and developing nations. *Journal of Money, Credit and Banking*, 41(7), 1365–1389.
- Alpanda, S., & Honig, A. (2010). Political monetary cycles and de facto ranking of central bank independence. *Journal of International Money and Finance*, 29(6), 1003–1023.
- Alt, J. E., & Lassen, D. D. (2006). Transparency, political polarization, and political budget cycles in OECD countries. *American Journal of Political Science*, 50(3), 530–550.
- Alt, J. E., & Rose, S. (2007). Context-conditioned political budget cycles. In C. Boix & S. C. Stokes (Eds.), *Oxford handbook of comparative politics* (pp. 845–867). New York, NY: Oxford University Press.
- Amacher, R. C., & Boyes, W. J. (1978). Cycles in senatorial voting behavior: implications for the optimal frequency of elections. *Public Choice*, 33(3), 5–13.
- Ames, B. (1987). *Political survival: politicians and public policy in Latin America*. Berkeley, CA: University of California Press.
- Ames, B., Hiroi, T., & Renno, L. (2005). The political economy of personnel expenditures: brazilian states, 1965–1994. *Brazilian Journal of Political Economy*, 25(1), 50–69.
- Annett, A. M. (1993). Elections and macroeconomic outcomes in Ireland, 1948–91. *Economic and Social Review*, 25(1), 21–47.
- Artus, P. (1993). *Théorie de la croissance et des fluctuations*. Paris: P.U.F.
- Ashworth, J., & Heyndels, B. (2002). Tax structure turbulence in OECD countries. *Public Choice*, 111(1–2), 347–376.
- Baber, W. R., & Sen, P. K. (1986). The political process and the use of debt financing by State governments. *Public Choice*, 48(3), 201–215.
- Baleiras, R. N., & Santos, V. (2000). Behavioral and institutional determinants of political business cycles. *Public Choice*, 104(1–2), 121–147.
- Balke, N. S. (1990). The rational timing of parliamentary elections. *Public Choice*, 65(3), 201–216.

- Barberia, L. G., & Avelino, G. (2011). Do political budget cycles differ in Latin American democracies? *Economia*, 11(2), 101–134.
- Barro, R. J., & Gordon, D. B. (1983). Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics*, 12(1), 101–121.
- Beck, N. (1982). Does there exist a political business cycle: a Box-Tiao analysis. *Public Choice*, 38(2), 205–209.
- Beck, N. (1987). Elections and the Fed: is there a political monetary cycle? *American Journal of Political Science*, 31(1), 194–216.
- Benito, B., Bastida, F., & Vicente, C. (2013a). Creating room for manoeuvre: a strategy to generate political budget cycles under fiscal rules. *Kyklos*, 66(4), 467–496.
- Benito, B., Bastida, F., & Vicente, C. (2013b). Municipal elections and cultural expenditure. *Journal of Cultural Economics*, 37(1), 3–32.
- Ben-Porath, Y. (1975). The years of plenty and the years of famine – A political business cycle? *Kyklos*, 28(2), 400–403.
- Besley, T., & Case, A. (1995). Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits. *Quarterly Journal of Economics*, 110(3), 769–798.
- Blais, A., & Nadeau, R. (1992). The electoral budget cycle. *Public Choice*, 74(4), 389–403.
- Blaydes, L. (2011). *Elections and distributive politics in Mubarak's Egypt*. Cambridge, UK: Cambridge University Press.
- Block, S. A. (2002). Political business cycles, democratization, and economic reform: The case of Africa. *Journal of Development Economics*, 67(1), 205–228.
- Block, S. A., Ferree, K. E., & Singh, S. (2003). Multiparty competition, founding elections and political business cycles in Africa. *Journal of African Economies*, 12(3), 444–468.
- Block, S. A., & Vaaler, P. M. (2004). The price of democracy: sovereign risk ratings, bond spreads and political business cycles in developing countries. *Journal of International Money and Finance*, 23(6), 917–946.
- Boddy, R., & Crotty, J. (1975). Class conflict and macro-policy: the political business cycle. *Review of Radical Political Economics*, 7(1), 1–19.
- Bordignon, M., Cerniglia, F., & Revelli, F. (2003). In search for yardstick competition: a spatial analysis of Italian municipality property tax setting. *Journal of Urban Economics*, 54(2), 199–217.
- Borjas, G. L. (1984). Electoral cycles and the earnings of federal bureaucrats. *Economic Inquiry*, 22(4), 447–459.
- Boroohah, V. K., & van der Ploeg, F. (1983). *Political aspects of the economy*. Cambridge, UK: Cambridge University Press.
- Brechler, J., & Geršl, A. (2014). Political legislation cycle in the Czech Republic. *Constitutional Political Economy*, 25(2), 137–153.
- Brender, A., & Drazen, A. (2005). Political budget cycles in new versus established democracies. *Journal of Monetary Economics*, 52(7), 1271–1295.
- Brender, A., & Drazen, A. (2007). Electoral fiscal policy in new, old, and fragile democracies. *Comparative Economic Studies*, 49(3), 446–466.
- Brender, A., & Drazen, A. (2008). How do budget deficits and economic growth affect reelection prospects? *American Economic Review*, 98(5), 2203–2220.
- Brender, A., & Drazen, A. (2013). Elections, leaders, and the composition of government spending. *Journal of Public Economics*, 97, 18–31.
- Brennan, G., & Buchanan, J. M. (1980). *The power to tax: analytical foundations of fiscal constitution*. New York: Cambridge University Press.
- Breton, A. (1974). *The economic theory of representative government*. Chicago: Aldine Publishing Company.
- Breuss, F. (2008). A common election day for the Euro zone? *Kyklos*, 61(1), 19–32.
- Brollo, F., & Troiano, U. (2013). What happens when a woman wins an election? Evidence from close races in Brazil. Unpublished working paper.
- Canes-Wrone, B., & Park, J. K. (2012). Electoral business cycles in OECD countries. *American Political Science Review*, 106(1), 103–122.
- Cargill, T. F., & Hutchison, M. M. (1991). Political business cycles with endogenous election timing: evidence from Japan. *Review of Economics and Statistics*, 73(4), 733–739.
- Carlsen, F. (1997). Opinion polls and political business cycles: theory and evidence for the United States. *Public Choice*, 92(3–4), 387–406.
- Castro, V., & Veiga, F. J. (2004). Political business cycles and inflation stabilization. *Economics Letters*, 83(1), 1–6.

- Chang, E. C. (2008). Electoral incentives and budgetary spending: rethinking the role of political institutions. *Journal of Politics*, 70(4), 1086–1097.
- Chappell, H. W., & Keech, W. R. (1986). Party differences in macroeconomic policies and outcomes. *American Economic Review*, 76(2), 71–74.
- Chappell, D., & Peel, D. A. (1979). On the political theory of the business cycle. *Economics Letters*, 2(4), 327–332.
- Chaudhuri, K., & Dasgupta, S. (2005). Political determinants of central governments' economic policies in India: empirical investigation. *Journal of International Development*, 17(7), 957–978.
- Chortareas, G. E. (1999). Political business cycles: theory, evidence and extension. In J. M. Mbaku & M. S. Kimenyi (Eds.), *Institutions and collective choice in developing countries: applications of the theory of public choice* (pp. 113–144). Aldershot: Ashgate.
- Chowdhury, A. R. (1993). Political surfing over economic waves: parliamentary election timing in India. *American Journal of Political Science*, 37(4), 1000–1118.
- Clark, W. R., & Hallberg, M. (2000). Mobile capital, domestic institutions, and electorally induced monetary and fiscal policy. *American Political Science Review*, 94(2), 323–346.
- Clark, W. R., & Reichert, U. N. (1998). International and domestic constraints on political business cycles in OECD economies. *International Organization*, 52(1), 87–120.
- Clinger Mayer, J. C., & Wood, B. D. (1995). Disentangling patterns of state debt financing. *American Political Science Review*, 89(1), 108–120.
- Coelho, C., Veiga, F. J., & Veiga, L. G. (2006). Political business cycles in local employment: evidence from Portugal. *Economics Letters*, 93(1), 82–87.
- Cukierman, A. (1992). *Central bank strategy, credibility and independence: theory and evidence*. Cambridge, MA: MIT Press.
- Cukierman, A., & Meltzer, A. H. (1986). A positive theory of discretionary policy, the costs of democratic government and the benefits of a constitution. *Economic Inquiry*, 24(3), 367–388.
- Cusack, T. R., & Ward, M. D. (1981). Military spending in the United States, Soviet Union, and the People's Republic of China. *Journal of Conflict Resolution*, 25(3), 429–469.
- Davidson, L. S., Fratianni, M., & von Hagen, J. (1990). Testing for political business cycles. *Journal of Policy Modeling*, 12(1), 35–59.
- de Haan, J. (2013). Democracy, elections and government budget deficits. *German Economic Review*, 15(1), 131–142.
- de Haan, J., & Klomp, J. (2013). Conditional political budget cycles: a review of recent evidence. *Public Choice*, 157(3–4), 387–410.
- de Tocqueville, A. (1835). *De la démocratie en Amérique*. Paris: C. Gosselin.
- Dickson, V., Farnsworth, M., & Zhang, J. (2013). Opportunism and election timing by Canadian provincial and federal governments. *Canadian Public Policy*, 39(1), 101–118.
- Dinkel, R. (1981). Political business cycles in Germany and the United States: some theoretical and empirical considerations. In D. A. Hibbs & H. Fassbender (Eds.), *Contemporary Political Economy* (pp. 209–230). Amsterdam: North-Holland.
- Dobson, J., & Dufrene, U. B. (1993). The impacts of U.S. presidential elections on international security-markets. *Global Finance Journal*, 4(1), 39–47.
- Donahue, K., & Warin, T. (2007). The stability and growth pact: a European answer to the political budget cycle? *Comparative European Politics*, 5(4), 423–440.
- Downs, A. (1957a). An economic theory of political action in democracy. *Journal of Political Economy*, 65(2), 135–150.
- Downs, A. (1957b). *An economic theory of democracy*. New York: Harper & Row.
- Drazen, A. (2000). The political business cycle after 25 years. In B. S. Bernanke & K. Rogoff (Eds.), *NBER macroeconomic annual* (pp. 75–137). Cambridge, MA: MIT Press.
- Drazen, A. (2001). *Political economy in macroeconomics*. Princeton, NJ: Princeton University Press.
- Drazen, A., & Eslava, M. (2010). Electoral manipulation via voter-friendly spending: theory and evidence. *Journal of Development Economics*, 92(1), 39–52.
- Dreher, A., & Vaubel, R. (2004). Do IMF and IBRD cause moral hazard and political business cycles? Evidence from panel data. *Open Economies Review*, 15(1), 5–22.
- Dubois, E., Leprince, M., & Paty, S. (2007). The effects of politics on local tax setting: evidence from France. *Urban Studies*, 44(8), 1603–1618.
- Efthyvoulou, G. (2011). Political cycles under external economic constraints: evidence from Cyprus. *Journal of Economics and Business*, 63(6), 638–662.
- Efthyvoulou, G. (2012). Political budget cycles in the European Union and the impact of political pressures. *Public Choice*, 153(3–4), 295–327.

- Ehrhart, H. (2013). Elections and the structure of taxation in developing countries. *Public Choice*, 156(1–2), 195–211.
- Enkelmann, S., & Leibrecht, M. (2013). Political expenditure cycles and election outcomes: evidence from disaggregation of public expenditures by economic functions. *Economics Letters*, 121(1), 128–132.
- Fair, R. C. (1975). On controlling the economy to win elections. Cowles Foundation Discussion Paper No 397.
- Faye, M., & Niehaus, P. (2012). Political aid cycles. *American Economic Review*, 102(7), 3516–3530.
- Feiwel, G. R. (1974). Reflection on Kalecki's theory of political business cycle. *Kyklos*, 27(1), 21–48.
- Ferris, J. S., & Voia, M. (2009). What determines the length of a typical Canadian parliamentary government? *Canadian Journal of Political Science/Revue Canadienne de Science Politique*, 42(4), 881–910.
- Findlay, D. W. (1990). The political business cycle and republican administrations: an empirical investigation. *Public Finance Quarterly*, 18(3), 328–338.
- Foerster, S. R. (1994). *Stock market performance and elections: made-in-Canada effects?* (pp. 39–42). Summer: Canadian Investment Review.
- Foerster, S. R., & Schmitz, J. J. (1997). The transmission of U.S. election cycles to international stock returns. *Journal of International Business Studies*, 28(1), 1–27.
- Franzese, R. J. (2000). Review of A. Alesina & N. Roubini with G. Cohen, *Political Cycles and the Macroeconomy*. *Journal of Policy Analysis Management*, 19(3), 501–509.
- Franzese, R. J. (2002). Electoral and partisan cycles in economic policies and outcomes. *Annual Review of Political Science*, 5, 369–421.
- Franzese, R. J., & Long-Jusko, K. (2006). Political-economic cycles. In B. R. Weingast & D. Wittman (Eds.), *The Oxford handbook of political economy* (pp. 545–564). Oxford: Oxford University Press.
- Frey, B. S., & Lau, L. J. (1968). Towards a mathematical model of government behaviour. *Zeitschrift Für Nationalökonomie*, 28, 355–380.
- Frey, B. S., & Ramser, H. J. (1976). The political business cycle: a comment. *Review of Economic Studies*, 43(3), 553–555.
- Frey, B. S., & Schneider, F. (1978a). An empirical study of politico-economic interaction in the US. *Review of Economics and Statistics*, 60(2), 174–183.
- Frey, B. S., & Schneider, F. (1978b). A politico-economic model of the United Kingdom. *Economic Journal*, 88(350), 243–253.
- Friedman, M. (1972). Monetary policy. *Proceedings of the American Philosophical Society*, 116(3), 183–196.
- Froyen, R. T., Havrilesky, T. W., & Waud, R. N. (1997). The asymmetric effects of political pressures on U.S. monetary policy. *Journal of Macroeconomics*, 19(3), 471–493.
- Galli, E., & Rossi, S. P. S. (2002). Political budget cycles: the case of the Western German länder. *Public Choice*, 110(3–4), 283–303.
- Gamez, C., & Ibarra-yunez, A. (2007). Political cycles and public expenditures at subnational level: the case of Mexico. *Public Administration*, 4(16), 17–25.
- Garcia-Sanchez, I.-M., Prado-Lorenzo, J. M. & Cuadrado-Ballesteros, B. (2011). Do progressive governments undertake different debt burdens? Partisan vs. electoral cycles. *Revista de Contabilidad*, 14(1), 29–57.
- Gärtner, M. (1999). The election cycle in the inflation bias: evidence from the G-7 countries. *European Journal of Political Economy*, 15(4), 705–725.
- Gaubatz, K. T. (1991). Election cycles and war. *Journal of Conflict Resolution*, 35(2), 212–244.
- Getzner, M. (2002). Determinants of public cultural expenditures: an exploratory time series analysis for Austria. *Journal of Cultural Economics*, 26(4), 287–306.
- Getzner, M. (2004). Cultural policies and fiscal federalism. *Public Finance and Management*, 4(1), 21–50.
- Geys, B. (2007). Government weakness and electoral cycles in local public debt: evidence from Flemish municipalities. *Local Government Studies*, 33(2), 237–251.
- Gildea, J. A. (1990). Explaining FOMC members' votes. In T. Mayer (Ed.), *The political economy of American monetary policy* (pp. 211–227). New York: Cambridge University Press.
- Ginsburgh, V., & Michel, P. (1983). Random timing of elections and the political business cycle. *Public Choice*, 40(2), 155–164.
- Goeminne, S., & Smolders, C. (2014). Politics and public infrastructure investments in local governments: empirical evidence from Flemish municipalities (1996–2009). *Local Government Studies*, 40(2), 182–202.
- Golden, D. G., & Poterba, J. M. (1980). The price of popularity: the political business cycle re-examined. *American Journal of Political Science*, 24(4), 696–714.
- Gonzalez, M. (2002). Do changes in democracy affect the political budget cycle? Evidence from Mexico. *Review of Development Economics*, 6(2), 204–224.

- Grier, K. B. (1987). Presidential elections and Federal Reserve policy: an empirical test. *Southern Economic Journal*, 54(2), 475–486.
- Grier, K. B. (1989). On the existence of a political monetary cycle. *American Journal of Political Science*, 33(2), 376–389.
- Grier, R. M., & Grier, K. B. (2000). Political cycles in nontraditional settings: theory and evidence from the case of Mexico. *Journal of Law and Economics*, 43(1), 239–264.
- Griffin, L. J., Devine, J. A., & Wallace, M. (1982a). Monopoly capital, organized labor, and military expenditures in the United States, 1949–1976. *American Journal of Sociology*, 88(S), S113–S153.
- Griffin, L. J., Devine, J. A., & Wallace, M. (1982b). The political economy of military spending: evidence from the United States. *Cambridge Journal of Economics*, 6(1), 1–14.
- Guo, G. (2009). China's local political budget cycles. *American Journal of Political Science*, 53(3), 621–632.
- Hadri, K., Lockwood, B., & Maloney, J. (1998). Does central bank independence smooth the political business cycle in inflation some OECD evidence. *Manchester School of Economic and Social Studies*, 66(4), 377–395.
- Hagen, T. (2007). Estimating the effect of parliamentary elections on primary budget deficits in OECD countries. *Economics Bulletin*, 8(8), 1–5.
- Hallerberg, M., Vinhas da Souza, L., & Clark, W. R. (2002). Political business cycles in EU accession countries. *European Union Politics*, 3(2), 231–250.
- Hallerberg, M., & von Hagen, J. (1998). Electoral institutions and the budget process. In K. Fukasaku & R. Hausmann (Eds.), *Democracy, decentralization and deficits in Latin America* (pp. 65–94). Paris: Organisation for Economic Cooperation and Development.
- Hanusch, M. (2012). Mooted signals: economic disturbances and political budget cycles. *Journal of Applied Economics*, 15(2), 189–212.
- Havrilesky, T. M. (1988). Monetary policy signalling from the Administration to the Federal Reserve. *Journal of Money, Credit and Banking*, 20(1), 82–101.
- Havrilesky, T. M., & Gildea, J. A. (1992). Reliable and unreliable partisan appointees to the board of governors. *Public Choice*, 73(4), 397–417.
- Havrilesky, T. M., & Schweitzer, R. (1990). A theory of FOMC dissent voting with evidence from the time series. In T. Mayer (Ed.), *The political economy of American monetary policy* (pp. 197–210). New York: Cambridge University Press.
- Haynes, S. E., & Stone, J. A. (1988). Does the political business cycle dominate US unemployment and inflation? Some new evidence. In T. D. Willett (Ed.), *Political business cycles: the political economy of money, inflation and unemployment* (pp. 276–293). Durham: Duke University Press.
- Haynes, S. E., & Stone, J. A. (1989). An integrated test for electoral cycles in the US economy. *Review of Economics and Statistics*, 71(3), 426–434.
- Haynes, S. E., & Stone, J. A. (1990). Political models of the business cycle should be revived. *Economic Inquiry*, 28(3), 442–465.
- Heckelman, J. C., & Berument, H. (1998). Political business cycles and endogenous elections. *Southern Economic Journal*, 64(4), 987–1000.
- Hibbs, D. A. (1977). Political parties and macroeconomic policy. *American Political Science Review*, 71(4), 1467–1487.
- Hibbs, D. A. (1994). The partisan model of macroeconomic cycles. *Economics and Politics*, 6(1), 1–23.
- Hiroi, T. (2009). Exchange rate regime, central bank independence, and political business cycles in Brazil. *Studies in Comparative International Development*, 44(1), 1–22.
- Ho, Y. H., & Huang, C. J. (2013). Presidential election, checks and balances, and allocation of public expenditures in Taiwan. *Journal of Economics and Management*, 9(1), 31–53.
- Horgos, D., & Zimmermann, K. W. (2010). It takes two to tango: lobbies and the political business cycle. *Kyklos*, 63(3), 383–399.
- Inoguchi, T. (1979). Political surfing over economic waves: a simple model of the Japanese political-economic system in comparative perspective. Unpublished working paper.
- Ito, T. (1990). The timing of elections and political business cycles in Japan. *Journal of Asian Economics*, 1(1), 135–156.
- Ito, T. (1991). International impacts on domestic political economy: a case of Japanese general elections. *Journal of International Money and Finance*, 10, S73–S89.
- Ito, T., & Park, J. H. (1988). Political business cycles in the parliamentary system. *Economics Letters*, 27(3), 233–238.
- Johnson, D. R., & Siklos, P. L. (1994). Political effects on central bank behaviour: some international evidence. In P. L. Siklos (Ed.), *Variety of monetary reforms: lessons and experiences on the road to monetary union* (pp. 133–163). Dordrecht: Kluwer Academic Publishers.

- Kächelein, H., Lami, E., & Imami, D. (2011). Election-related cycles in publicly supplied goods in Albania. *South-Eastern Europe Journal of Economics*, 9(1), 13–25.
- Kalecki, M. (1943). Political aspects of full employment. *Political Quarterly*, 14(4), 322–331.
- Kamlet, M. S., & Mowery, D. C. (1987). Influences on executive and congressional budgetary priorities, 1955–1981. *American Political Science Review*, 81(1), 155–178.
- Karakas, M. (2014). Political business cycles in Turkey: a monetary approach. *Electronic Journal of Social Sciences*, 13(48), 159–174.
- Karran, T. (1985). The determinants of taxation in Britain: an empirical test. *Journal of Public Policy*, 5(3), 365–386.
- Katada, S. N., & McKeown, T. J. (1998). Aid politics and electoral politics: Japan, 1970–1992. *International Studies Quarterly*, 42(3), 591–602.
- Katsimi, M. (1998). Explaining the size of the public sector. *Public Choice*, 96(1–2), 117–144.
- Katsimi, M., & Sarantides, V. (2012). Do elections affect the composition of fiscal policy in developed, established democracies? *Public Choice*, 151(1–2), 325–362.
- Kayser, M. A. (2005). Who surfs, who manipulates? The determinants of opportunistic election timing and electorally motivated economic intervention. *American Political Science Review*, 99(1), 17–27.
- Keech, W. R. (1980). Elections and macroeconomic policy optimization. *American Journal of Political Science*, 24(2), 345–367.
- Keech, W. R., & Simon, C. P. (1983). Inflation, unemployment, and electoral terms: when can reform of political institutions improve macroeconomic policy? In K. R. Monroe (Ed.), *The political process and economic change* (pp. 77–107). New York: Agathon Press.
- Keil, M. W. (1988). Is the political business cycle really dead? *Southern Economic Journal*, 55(1), 86–99.
- Khemani, S. (2004). Political cycles in a developing economy: effect of elections in the Indian states. *Journal of Development Economics*, 73(1), 125–154.
- Kirschen, E. S., Bénard, J., Besters, H., Blackaby, F., Eckstein, O., Faaland, J., et al. (1964). *Economic policy in our time*. Amsterdam: North-Holland.
- Klein, N. (2004). Political cycles and economic policy in Israel: 1980–1999. *Israel Economic Review*, 2(1), 55–67.
- Klein, F. A. (2010). Reelection incentives and political budget cycle: evidence from Brazil. *Revista de Administração Pública*, 44(2), 283–337.
- Klein, F. A., & Sakurai, S. N. (2015). Term limits and political budget cycles at the local level: evidence from a young democracy. *European Journal of Political Economy*, 37, 21–36.
- Klien, M. (2014). Tariff increases over the electoral cycle: a question of size and salience. *European Journal of Political Economy*, 36, 228–242.
- Klomp, J., & de Haan, J. (2013a). Political budget cycles and election outcomes. *Public Choice*, 157(1–2), 245–267.
- Klomp, J., & de Haan, J. (2013b). Popular protest and political budget cycles: a panel data analysis. *Economics Letters*, 120(3), 516–520.
- Klomp, J., & de Haan, J. (2013c). Do political budget cycles really exist? *Applied Economics*, 45(3), 329–341.
- Kneebone, R. D., & McKenzie, K. J. (2001). Electoral and partisan cycles in fiscal policy: an examination of Canadian provinces. *International Tax and Public Finance*, 8(5–6), 753–774.
- Kohno, M., & Nishizawa, Y. (1990). A study of the electoral business cycle in Japan: elections and government spending on public construction. *Comparative Politics*, 22(2), 151–166.
- Kouvalas, O. (2013). Political budget cycles revisited, the case for social capital. Unpublished working paper.
- Kramer, G. H. (1971). Short-term fluctuations in U.S. voting behavior, 1896–1964. *American Political Science Review*, 65(1), 131–143.
- Krause, G. A. (2005). Electoral incentives, political business cycles, and macroeconomic performance: empirical evidence from Postwar U.S. personal income growth. *British Journal of Political Science*, 35(1), 77–101.
- Kräussl, R., Lucas, A., Rijsbergen, D. R., van der Sluis, P. J., & Vrugt, E. B. (2014). Washington meets Wall Street: a closer examination of the presidential cycle puzzle. *Journal of International Money and Finance*, 43, 50–69.
- Lächler, U. (1978). The political business cycle: a complementary study. *Review of Economic Studies*, 45(2), 369–375.
- Lächler, U. (1982). On political business cycles with endogenous election dates. *Journal of Public Economics*, 17(1), 111–117.
- Ladewig, J. W. (2008). Housing starts and the political business cycle. *American Politics Research*, 36(5), 776–798.

- Lagona, F., & Padovano, F. (2008). The political legislation cycle. *Public Choice*, 134(3–4), 201–229.
- Lami, E., Kächelein, H., & Imami, D. (2014). A new view into political business cycles: household expenditures in Albania. *Acta Oeconomica*, 64, 201–224.
- Laney, L. O., & Willett, T. D. (1983). Presidential politics, budget deficits and monetary policy in the United States, 1960–1976. *Public Choice*, 40(1), 53–69.
- Lau, L. J., & Frey, B. S. (1971). Ideology, public approval, and government behavior. *Public Choice*, 10, 21–40.
- Leertouwer, E., & Maier, P. (2001). Who creates political business cycles: should central banks be blamed? *European Journal of Political Economy*, 17(3), 445–463.
- Lema, D., & Streb, J. M. (2013). Party alignment and political budget cycles: the Argentine provinces. Unpublished working paper.
- Levitt, S. D. (1997). Using electoral cycles in police hiring to estimate the effect of police on crime. *American Economic Review*, 87(3), 270–290.
- Lindbeck, A. (1976). Stabilization policy in open economies with endogenous politicians. *American Economic Review*, 66(2), 1–19.
- Lohmann, S. (1998a). Rationalizing the political business cycle: a workhorse model. *Economics and Politics*, 10(1), 1–17.
- Lohmann, S. (1998b). Federalism and central bank independence: the politics of German monetary policy, 1957–1992. *World Politics*, 50(3), 401–446.
- MacRae, D. (1977). A political model of the business cycle. *Journal of Political Economy*, 85(2), 239–265.
- Maier, P., Sturm, J.-E., & de Haan, J. (2002). Political pressure on the Bundesbank: an empirical investigation using the Havrilesky approach. *Journal of Macroeconomics*, 24(1), 103–123.
- McCallum, B. T. (1978). The political business cycle: an empirical test. *Southern Economic Journal*, 44(3), 504–515.
- McGavin, B. H. (1987). The political business cycle: a reexamination of some empirical evidence. *Quarterly Journal of Business and Economics*, 26(1), 36–49.
- Mink, M., & de Haan, J. (2006). Are there political budget cycles in the Euro area? *European Union Politics*, 7(2), 191–211.
- Mintz, A. (1988). Electoral cycles and defence spending: a comparison of Israel and the United States. *Comparative Political Studies*, 21(3), 368–381.
- Mintz, A., & Hicks, A. (1984). Military Keynesianism in the United States, 1949–1976: disaggregating military expenditures and their determination. *American Journal of Sociology*, 90(2), 411–417.
- Mintz, A., & Ward, M. D. (1988). The evolution of Israel's military expenditures: 1960–1983. *The Western Political Quarterly*, 41(3), 489–507.
- Mourao, P. R. (2008). Political budget cycles and fiscal illusion – A panel data study. In H. Marques, E. Soukiazis, & P. Cerqueira (Eds.), *Perspectives on integration and globalisation* (pp. 83–110). Münster: LIT Verlag.
- Neto, O. A., & Borsani, H. (2004). Presidents and cabinets: the political determinants of fiscal behavior in Latin America. *Studies in Comparative International Development*, 39(1), 3–27.
- Nieto-Parra, S., & Santiso, J. (2012). Revisiting political budget cycles in Latin America. In J. Santiso & J. Dayton-Johnson (Eds.), *The Oxford handbook of Latin American political economy* (pp. 557–584). Oxford: Oxford University Press.
- Nincic, M. (1990). U.S. Soviet policy and the electoral connection. *World Politics*, 42(3), 370–396.
- Nincic, M., & Cusack, T. R. (1979). The political economy of US military spending. *Journal of Peace Research*, 16(2), 101–115.
- Nippiani, S., & Arize, A. C. (2005). U.S. presidential election impact on Canadian and Mexican stock markets. *Journal of Economics and Finance*, 29(3), 271–279.
- Nogare, C. D., & Galizzi, M. M. (2011). The political economy of cultural spending: evidence from Italian cities. *Journal of Cultural Economics*, 35(3), 203–231.
- Noonan, D. (2007). Fiscal pressures, institutional context, and constituents: a dynamic model of states' art agency appropriations. *Journal of Cultural Economics*, 31(4), 293–310.
- Nordhaus, W. D. (1972). The political business cycle. Cowles Foundation Discussion Paper No 333.
- Nordhaus, W. D. (1975). The political business cycle. *Review of Economic Studies*, 42(2), 169–190.
- Nordhaus, W. D. (1989). Alternative approaches to the political business cycles. *Brookings Papers on Economic Activity*, No, 2, 1–49.
- Olters, J.-P. (2001). Modeling politics with economic tools: a critical survey of the literature. IMF Working Paper No. 01/10.
- Özatay, F. (2007). Public sector price controls and electoral cycles. *Applied Economics*, 39(4), 527–539.
- Paiva, C. A. (1996). Electoral price cycles in regulated industries. *World Development*, 24(10), 1673–1680.

- Paldam, M. (1981). An essay on the rationality of the economic policy: the test-case of the electoral cycle. *Public Choice*, 37(2), 287–305.
- Palmer, H. D., & Whitten, G. D. (2000). Government competence, economic performance and endogenous election dates. *Electoral Studies*, 19(2–3), 413–426.
- Pepinsky, T. (2007). Autocracy, elections, and fiscal policy: evidence from Malaysia. *Studies in Comparative International Development*, 42(1–2), 136–163.
- Persson, T., & Swenson, L. (1989). Why a stubborn conservative would run a deficit. *Quarterly Journal of Economics*, 104(2), 325–345.
- Persson, T., & Tabellini, G. (1990). *Macroeconomic policy, credibility and politics*. Chur, Switzerland: Hardwood Academic Publishers.
- Persson, T., & Tabellini, G. (2000). *Political economics: explaining economic policy*. Cambridge, MA: MIT Press.
- Persson, T., & Tabellini, G. (2003). *The economic effects of constitutions*. Cambridge, MA & London: MIT Press.
- Petterson-Lidbom, P. (2001). An empirical investigation of the strategic use of debt. *Journal of Political Economy*, 109(3), 570–582.
- Price, S. (1997). Political business cycles and macroeconomic credibility: a survey. *Public Choice*, 92(3–4), 407–427.
- Price, S. (1998). Comment on “The politics of the political business cycle”. *British Journal of Political Science*, 28(1), 201–209.
- Reid, B. G. (1998). Endogenous elections, electoral budget cycles and Canadian provincial governments. *Public Choice*, 97(1–2), 35–48.
- Remmer, K. L. (1993). The political economy of elections in Latin America, 1980–1991. *American Political Science Review*, 87(2), 393–407.
- Remmer, K. L. (2007). The political economy of patronage: expenditure patterns in the Argentine provinces, 1983–2003. *Journal of Politics*, 69(2), 363–377.
- Rogoff, K. (1985). The optimal degree of commitment to an intermediate monetary target. *Quarterly Journal of Economics*, 100(403), 1069–1189.
- Rogoff, K. (1990). Equilibrium political budget cycles. *American Economic Review*, 80(1), 21–36.
- Rogoff, K., & Sibert, A. (1988). Elections and macro-economic policy cycle. *Review of Economic Studies*, 55(1), 1–16.
- Rose, S. (2006). Do fiscal rules dampen the political business cycle? *Public Choice*, 128(3–4), 407–431.
- Rose, S. (2008). The political manipulation of US state rainy day funds under rules versus discretion. *State Politics & Policy Quarterly*, 8(2), 150–176.
- Rumi, C. (2014). National electoral cycles in transfers to subnational jurisdictions: evidence from Argentina. *Journal of Applied Economics*, 17(1), 161–178.
- Sapir, A., & Sekkat, K. (1999). Optimum electoral areas: should Europe adopt a single election day? *European Economic Review*, 43(8), 1595–1619.
- Sapir, A., & Sekkat, K. (2002). Political cycles, Fiscal deficits, and Output spillovers in Europe. *Public Choice*, 111(1–2), 195–205.
- Sargent, T. J., & Wallace, N. (1975). Rational expectations, the optimal monetary instrument and the optimal money supply rule. *Journal of Political Economy*, 83(2), 241–254.
- Schneider, C. J. (2010). Fighting with one hand tied behind the back: political budget cycles in the West German states. *Public Choice*, 142(1–2), 125–150.
- Schuknecht, L. (1996). Political business cycles and fiscal policies in developing countries. *Kyklos*, 49(2), 155–170.
- Schuknecht, L. (1999). Fiscal policy cycles and the exchange rate regime in developing countries. *European Journal of Political Economy*, 15(3), 569–580.
- Schuknecht, L. (2000). Fiscal policy cycles and public expenditure in developing countries. *Public Choice*, 102(1–2), 115–130.
- Schultz, K. A. (1995). The politics of political business cycle. *British Journal of Political Science*, 25(1), 79–100.
- Schumpeter, J. A. (1939). *Business cycles: a theoretical, historical, and statistical analysis*. New York: McGraw Hill.
- Serritzlew, S. (2005). Breaking budgets: an empirical examination of Danish municipalities. *Financial Accountability & Management*, 21(4), 413–435.
- Shelton, C. A. (2014). Legislative budget cycles. *Public Choice*, 159(1–2), 251–275.
- Shi, M., & Svensson, J. (2003). Political budget cycles: a review of recent developments. *Nordic Journal of Political Economy*, 29(1), 67–76.

- Shi, M., & Svensson, J. (2006). Political budget cycles: do they differ across countries and why? *Journal of Public Economics*, 90(8–9), 1367–1389.
- Shughart, W. F., I. I., & Tollison, R. D. (1985). Legislation and political business cycles. *Kyklos*, 38(1), 43–59.
- Sieg, G. (1997). A model of partisan central banks and opportunistic political cycles. *European Journal of Political Economy*, 13(3), 503–516.
- Sjahrir, B. S., Kis-Katos, K., & Schulze, G. G. (2013). Political budget cycles in Indonesia at the district level. *Economics Letters*, 120(2), 342–345.
- Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. New York: Random House.
- Smith, A. (1996). Endogenous election timing in majoritarian parliamentary systems. *Economics and Politics*, 8(2), 85–110.
- Smith, A. (2003). Election timing in majoritarian parliaments. *British Journal of Political Science*, 33(3), 397–418.
- Smith, A. (2004). *Election timing*. Cambridge, MA: Cambridge University Press.
- Soh, B. H. (1988). Political instability and economic fluctuations in the Republic of Korea. *Public Choice*, 57(3), 259–274.
- Solé Ollé, A. (2003). Electoral accountability and tax mimicking: the effects of electoral margins, coalition government, and ideology. *European Journal of Political Economy*, 19(4), 685–713.
- Solé Ollé, A. (2006). The effects of party competition on budget outcomes: empirical evidence from local governments in Spain. *Public Choice*, 126(1–2), 145–176.
- Spafford, D. (1981). Highway employment and provincial elections. *Canadian Journal of Political Science*, 14(1), 135–143.
- Streb, J. M. (1999). Reelection or term limits? The short and the long run view of economic policy. *Estudios de economía*, 26(2), 187–206.
- Streb, J. M., Lema, D., & Garofalo, P. (2012). Temporal aggregation in political budget cycles. *Economía*, 13(1), 39–78.
- Streb, J. M., Lema, D., & Garofalo, P. (2013). Electoral cycles in international reserves: evidence from Latin America and the OECD. *Revista di Economia*, 20(2), 27–60.
- Streb, J. M., Lema, D., & Torrens, G. (2009). Checks and balances on political budget cycles: cross-country evidence. *Kyklos*, 62(3), 426–447.
- Sturm, R. R. (2013). Economic policy and the presidential election cycle in stock returns. *Journal of Economics and Finance*, 37(2), 200–215.
- Su, T. T., Kamlet, M. S., & Mowery, D. C. (1993). Modeling US budgetary and fiscal policy outcomes: a disaggregated, systemwide perspective. *American Journal of Political Science*, 37(1), 213–245.
- Sukhtankar, S. (2012). Sweetening the deal? Political connections and sugar mills in India. *American Economic Journal: Applied Economics*, 4(3), 43–63.
- Suzuki, M. (1994). Evolutionary voter sophistication and political business cycle. *Public Choice*, 81(3–4), 241–261.
- Swank, O. H., & Swank, J. (1993). In search of the motives behind US fiscal macroeconomic policy. *Applied Economics*, 25(8), 1013–1022.
- Tabellini, G., & Alesina, A. (1990). Voting on the budget deficit. *American Economic Review*, 80(1), 37–49.
- Telatar, F. (2003). Political business cycles in the parliamentary systems: evidence from Turkey. *Emerging Markets Finance and Trade*, 39(4), 24–39.
- Terrones, M. (1989). Macroeconomic policy choices under alternative electoral structures. University of Western Ontario research report No 8905.
- Theil, H. (1958). *Economic forecasts and policy*. Amsterdam: North-Holland.
- Thies, C., & Porche, S. (2007). The political economy of agricultural protection. *The Journal of Politics*, 69(1), 116–127.
- Thompson, W. R., & Zuk, G. (1983). American elections and the international electoral-economic cycle: a test of the Tufte hypothesis. *American Journal of Political Science*, 27(3), 464–484.
- Tinbergen, J. (1952). *On the theory of economic policy*. Amsterdam: North-Holland.
- Tufte, E. R. (1978). *Political control of the economy*. Princeton: Princeton University Press.
- Umstead, D. A. (1977). Forecasting stock market prices. *Journal of Finance*, 32(2), 427–441.
- van Dalen, H. P., & Swank, O. H. (1996). Government spending cycles: ideological or opportunistic. *Public Choice*, 89(1–2), 183–200.
- Vaibel, R. (1997). The bureaucratic and partisan behavior of independent central bank: German and international evidence. *European Journal of Political Economy*, 13(2), 201–224.
- Veiga, L. G., & Pinho, M. M. (2007). The political economy of intergovernmental grants: evidence from a maturing democracy. *Public Choice*, 133(3–4), 457–477.

- Veiga, L. G., & Veiga, F. J. (2007a). Political business cycles at the municipal level. *Public Choice*, 131(1–2), 45–64.
- Veiga, L. G., & Veiga, F. J. (2007b). Does opportunism pay off? *Economics Letters*, 96(2), 177–182.
- Vergne, C. (2009). Democracy, elections and allocation of public expenditures in developing countries. *European Journal of Political Economy*, 25(1), 63–77.
- Vicente, C., Benito, B., & Bastida, F. (2013). Transparency and political budget cycles at municipal level. *Swiss Political Science Review*, 19(2), 139–156.
- Voia, M. C., & Ferris, J. S. (2013). Do business cycle peaks predict election calls in Canada? *European Journal of Political Economy*, 29, 102–118.
- Wagner, R. E. (1977). Economic manipulation for political profit. *Kyklos*, 30(3), 395–409.
- Walsh, C. E. (1995). Optimal contract for central banker. *American Economic Review*, 85(1), 150–167.
- Wasserman, A. M. (1983). Political business cycles, presidential elections, and suicide and mortality. *American Sociological Review*, 48(5), 711–720.
- Wehner, J. (2013). Electoral budget cycles in legislatures. *Legislative Studies Quarterly*, 38(4), 545–570.
- Williams, J. T. (1990). The political manipulation of macroeconomic policy. *American Political Science Review*, 84(3), 767–796.
- Wittman, D. (1977). Candidates with policy preferences: a dynamic model. *Journal of Economic Theory*, 14(1), 180–189.
- Wittman, D. (1983). Candidate motivation: a synthesis of alternative theories. *American Political Science Review*, 77(1), 142–157.
- Alt, J. E. & Chrystal, K. A. (1981). Public sector behaviour: the status of the political business cycle. In Currie, D. A., Nobay, A. R. & Peel, D. (Eds.), *Macroeconomic analysis: essays in macroeconomics and econometrics*. London: Croom Helm, PP. 353–382.
- Zuk, G., & Woodbury, N. R. (1986). U.S. defense spending, electoral cycles, and Soviet-American relations. *Journal of Conflict Resolution*, 30(3), 445–468.

# Political Cycles in OECD Economies

ALBERTO ALESINA

*Harvard University, NBER and CEPR*

and

NOURIEL ROUBINI

*Yale University, NBER and CEPR*

*First version received August 1990; final version accepted July 1992 (Eds.)*

This paper studies whether the dynamic behaviour of GNP growth, unemployment and inflation is systematically affected by the timing of elections and of changes of governments. The sample include the last three decades in 18 OECD economies. We explicitly test the implication of several models of political cycles, both of the "opportunistic" and of the "partisan" type. Also, we confront the implication of recent "rational" models with more traditional approaches. Our results can be summarized as follows: (a) The "political business cycle" hypothesis, as formulated in Nordhaus (1975) on output and unemployment is generally rejected by the data; (b) inflation tends to increase immediately after elections, perhaps as a result of pre-electoral expansionary monetary and fiscal policies; (c) we find evidence of temporary partisan differences in output and unemployment and of long-run partisan differences in the inflation rate as implied by the "rational partisan theory" by Alesina (1987); (d) we find virtually no evidence of permanent partisan differences in output growth and unemployment.

## 1. INTRODUCTION

Different models of political cycles emphasize either the "opportunistic" or the "partisan" incentives of policymakers. In "opportunistic" models, the policymakers maximize their popularity or their probability of re-election. In "partisan" models different political parties represent the interests of different constituencies and, when in office, follow policies which are favourable to their supporting groups; specifically, the left-wing parties are more concerned with the problem of unemployment, while the right-wing parties are relatively more willing to bear the costs of unemployment to reduce inflation.

This literature has developed in two clearly distinct phases. The first one, in the mid-seventies, is due to the influential work by Nordhaus (1975), and Lindbeck (1976) on "opportunistic" cycles and by Hibbs (1977) on "partisan" cycles. These papers share a "pre-rational expectations" model of the economy and are based upon an exploitable "Phillips curve". The "political business cycle" model of Nordhaus predicts pre-electoral high growth and low unemployment, increasing inflation around the election time and a post-electoral recession, regardless of the political orientation of the incumbent government. Hibbs' "partisan" model implies systematic and permanent differences in the inflation/unemployment combination chosen by different political parties.

Macroeconomists soon lost interest in this subject, because at that time the profession was developing (or fighting against), the "rational expectations revolution."<sup>1</sup>

1. Furthermore, Nordhaus' "political business cycle" model did not receive much empirical support. Soon after the publication of Nordhaus' paper McCallum (1978) and Paldam (1979) presented negative empirical results for the U.S. and OECD economies respectively. More favourable results are shown by Tufte (1978), but only on a sample including a few American elections.

The second phase took off in the mid-eighties as a branch of the game-theoretic approach to the positive theory of policy. Cukierman and Meltzer (1986), Rogoff and Sibert (1988), Rogoff (1990), and Persson and Tabellini (1990) propose rational "opportunistic" models; Alesina (1987) develops a rational partisan approach. These models depart from their predecessors in two important dimensions. First, the assumption of economic agents' rationality makes real economic activity less directly and predictably influenced by monetary policy. Second, voters' rationality implies that they cannot be systematically "fooled" in equilibrium.

This second generation of models has empirical implications which are quite different from those of the earlier literature: the assumption of rationality reduces the extent and the likelihood of regular political cycles, although it does not eliminate them. For example, in models with rational economic agents and voters, Nordhaus' type cycles are mitigated. Rather than regular multi-year cycles on output and unemployment, one should observe, according for instance to Rogoff and Sibert (1988), short-lived electoral cycles on monetary and fiscal policy instruments, but not necessarily on the level of economic activity. Alesina (1987, 1988b) shows that in a partisan model with nominal wage contracts and rational voters, permanently different inflation rates across parties may result only in temporary, post-election differences in growth and unemployment.

This paper addresses two questions. First, whether or not the dynamic behaviour of GNP growth, unemployment and inflation is systematically affected by the timing of elections and of changes of governments. Second, whether or not the second generation of rational models has provided useful insights to interpret the evidence. The paper suggests an affirmative answer to both these questions by examining 18 OECD democracies, in the last three decades.

More specifically our results can be summarized as follows: (a) The "political business cycle" hypothesis, as formulated in Nordhaus on growth and unemployment is generally rejected by the data. (Some favourable evidence can be found in only two countries); (b) inflation tends to increase immediately after elections, perhaps as a result of pre-electoral expansionary monetary and fiscal policies; this result yields support to the Rogoff and Sibert (1988) and Rogoff (1990) model of "political budget cycles"; (c) we find evidence of temporary partisan differences in output and unemployment and of long-run partisan differences in the inflation rate as implied by the "rational partisan theory" of Alesina (1987). This pattern appears rather unambiguously in countries with a pure two-party system, or with clearly identifiable "right" and "left" coalitions; (d) we find virtually no evidence of permanent partisan differences in output and unemployment. Indirectly, results (c) and (d) yield support to the positive model of inflation developed by Kydland-Prescott (1977) and Barro and Gordon (1983a, b).

The qualitative features of these results are consistent with the finding on the United States by Alesina and Sachs (1988), Alesina (1988a) and Chapell and Keech (1988). The advantage of a multi-country study is that, of course, one has many more degrees of freedom. Elections and changes of governments are relatively infrequent events. Thus, the researcher is left with very few observations and only one country is considered. This is why systematic multi-country studies are particularly useful in this area.

The paper is organized as follows. In the next section we highlight the empirical implications of several models of political cycles. Since several comprehensive reviews of the literature have recently appeared (Alesina (1988a), Nordhaus (1989), Persson and Tabellini (1990)) we sketch the various models very succinctly. In Section 3 we present regressions on a panel data set of all the countries in the sample. Section 4 briefly discusses the results of country by country regressions. Section 5 employs Hamilton's (1989) method of timing recessions and expansions. The last section concludes.

## 2. MODELS OF POLITICO-ECONOMIC CYCLES

### 2.1. *The “political business cycle” (Nordhaus (1975))*

The assumptions underlying Nordhaus’ “political business cycle” (henceforth PBC) can be characterized as follows:

A.1. The economy is described by a Phillips curve:

$$u_t = \bar{u} + \alpha u_{t-1} + \gamma(\pi_t - \pi_t^e) + \varepsilon_t; \quad 0 < \alpha < 1; \gamma < 0. \quad (1)$$

where  $u$  is unemployment;  $\bar{u}/(1-\alpha)$  is the steady state “natural” level of unemployment;  $\pi$  is inflation;  $\pi^e$  is expected inflation;  $\varepsilon$  is a random shock with zero mean;  $\alpha$ ,  $\gamma$  are parameters. The autoregressive term in (1) captures various sources of persistence. The “natural” level of unemployment is normalized at zero, with no loss of generality. By Okun’s law, the same model can be written in terms of output growth,  $y_t$ , instead of unemployment. In the empirical work which follows, we will consider both GNP growth and unemployment.

A.2. Inflation expectations are adaptive:

$$\pi_t^e = \pi_{t-1} + \lambda(\pi_{t-1}^e - \pi_{t-1}); \quad 0 < \lambda < 1. \quad (2)$$

A.3. Inflation is directly controlled by the policymakers.<sup>2</sup>

A.4. Politicians are “opportunistic”: they only care about holding office, and they do not have “partisan” objectives.

A.5. Voters are “retrospective”. They judge the incumbent’s performance based upon the state of the economy during the incumbent’s term of office, and heavily discount past observations.

A.6. The timing of elections is exogenously fixed.

Under these assumptions, Nordhaus derives the following testable implications: (i) every government follows the same policy; (ii) towards the end of his term of office, the incumbent stimulates the economy to take advantage of the “short run” more favourable Phillips curve; (iii) the rate of inflation increases around the election time as a result of the pre-electoral economic expansion; after the election, inflation is reduced with contractionary policies.<sup>3</sup>

This basic model has recently been developed by investigating the role of rationality.

### 2.2. *Rational political business cycle models*

Persson and Tabellini (1990) propose a simple model which summarizes the basic insights of this approach, due to Rogoff and Sibert (1988). Assumptions A.1, A.3, A.4, and A.6 as in Nordhaus’ model are retained. Assumption A.2 is replaced by:

A.2'.  $\pi_t^e = E(\pi_t / I_{t-1})$ : rational expectations.

A.2".  $I_{t-1}$  includes all the relevant information except the level of “competence” of different policymakers.

Assumption A.5 is substituted by:

A.5'. Voters choose the candidate which is rationally expected to deliver the highest utility, if elected.

A.5". There are no differences in voters’ utility functions.

2. To be precise, Nordhaus (1975) assumes that policymakers control aggregate demand and, indirectly, inflation. This difference is inessential.

3. Nordhaus’ (1975) model predicts that inflation should increase *before* the election. However, given time lags between the effects of aggregate demand policies on output and inflation, one can build a model in which inflation increases *after*, rather than *before* the election. (See Lindbeck (1976)).

$E(\cdot)$  is the expectation operator and  $I_{t-1}$  is the information set of the voters at time  $(t-1)$  when expectations are formed. A.2") implies an asymmetry of information between the policymakers and the voters: the former know their own competence, but the latter do not.<sup>4</sup> Policymakers' "competence" is defined as their ability of keeping unemployment low (and GNP growth high) with a relatively low level of inflation.<sup>5</sup>

By taking advantage of this informational asymmetry, and by trying to appear as competent as possible before elections, the politicians behave in a way leading to a Nordhaus' type PBC. However, given voters' rationality and awareness of the politicians' incentives, the latter are limited in their "opportunistic" behaviour. Thus, the resulting cycles are more short-lived and less regular than in Nordhaus' model.

The original proponents of the "competence" model, i.e. Rogoff and Sibert (1988) and Rogoff (1990), consider a budget problem, rather than an inflation/unemployment trade-off, but with identical assumptions about the distribution of information. These papers have empirical implications on opportunistic cycles on monetary and fiscal variables, rather than on unemployment and output. In fact, the model by Rogoff and Sibert (1988) makes predictions on the inflation rate similar to those of the Nordhaus model, but does not imply any correlation between elections and GNP growth or unemployment.

### 2.3. *The "partisan theory"* (Hibbs (1977, 1987))

A strong version of the "partisan theory" (henceforth PT) based upon a non-rational expectation mechanism, adopts assumptions A.1, A.2, A.3 and A.6. Assumptions A.4 and A.5 are substituted by:

A.4'. Politicians are "partisan," in the sense that different parties maximize different objective functions. Left-wing parties attribute a higher cost to unemployment relative to inflation than right-wing parties.

A.5". Each voter is aware of the partisan difference and votes for the party which offers the policy closer to his most preferred outcome.

The assumption of partisanship is justified by the distributional consequences of unemployment. In periods of high unemployment, low growth and low inflation the relative share of income of the upper middle class, increases and the other way around, as shown by Hibbs (1987).

Thus, this model implies that different parties choose different points on the Phillips curve: output growth and inflation should be permanently higher and unemployment permanently lower when the left is in office than with right wing governments.<sup>6</sup>

### 2.4. "Rational partisan theory" (Alesina (1987))

Alesina (1987) and (1988b) suggests a "rational partisan theory" (henceforth RPT). This model adopts Assumption A.1, A.2', A.3, A.4', A.5" and A.6. The objective functions of the two parties can be written as:

$$W^i = \sum_{t=0}^T \delta^t [ -(\pi_t - c^i)^2 - b^i (u_t - K^i)^2 ], \quad 0 < \delta < 1; \quad (3)$$

4. In Cukierman and Meltzer (1986) the asymmetry of information is related to the knowledge of the realization of a random shock to the economy.

5. Formally, the degree of competence is modelled by adding a term in equation (1) which changes over time and is known by the policymaker but becomes known to the voters with a lag. For an explicit test of this model on U.S. data see Alesina, Londregan and Rosenthal (1990).

6. Nordhaus' PBC and Hibbs' PT can coexist. If one assumes that politicians are both opportunistic and partisan and voters are retrospective as implied by A.5, one obtains a "weaker" form of PT which incorporates elements of pre-electoral opportunistic behaviour. See Frey and Schneider (1978) and Nordhaus (1989).

where  $i = L, R$  identifies the "left" and the "right" parties. The difference between the two parties can be summarized by at least one of these three sets of inequalities:

$$c^L > c^R \geq 0; \quad b^L > b^R \geq 0; \quad K^L < K^R \leq \frac{\bar{u}}{1-\alpha}. \quad (4)$$

The last double inequality implies the time-inconsistency problem in monetary policy pointed out by Kydland-Prescott (1977) and Barro and Gordon (1983a, b). Since at least one of the two parties targets a level of output growth which is above the natural rate (normalized at zero), it introduces an "inflation bias" because of the lack of pre-commitments in monetary policy. Thus, a test of the RPT is indirectly a test for this specification of policymakers' objective functions.

This model generates a political cycle if we assume that uncontingent labour contracts are signed at discrete intervals (which do not coincide with the political terms of office) and that electoral outcomes are uncertain because of shocks to voters' preferences or to voters' participation rates in elections. The basic idea of the model is that, given the sluggishness in wage adjustments, changes in the inflation rate associated with changes in government create temporary deviations of real economic activity from its natural level.

More specifically, the following testable implications can be derived from the model: (i) at the beginning of a right-wing (left-wing) government output growth is below (above) its natural level and unemployment is above (below); (ii) after expectations, prices and wages adjust, output and unemployment return to their natural level; after this adjustment period, the level of economic activity should be independent of the party in office; (iii) the rate of inflation should remain higher throughout the term of a left-wing government; note that this occurs even if  $c^L = c^R$  in (3), as long as  $K^L < K^R$  or  $b^L > b^R$ . That is, the time consistent (but sub-optimal) inflation rate remains higher for left-wing parties even after the level of economic activity returns to its natural level.<sup>7</sup>

## 2.5. Previous empirical results

Most of the empirical studies on political cycles use post-war United States data. The evidence in favour of the RPT is relatively strong; evidence of "opportunistic" PBC is found for certain policy instruments (particularly government transfers) for limited sub-samples: for recent surveys of this empirical literature see Alesina (1988a) and Nordhaus (1989).

Multi-country studies are more scarce. Alt (1985) formally tests for partisan patterns in unemployment in twelve OECD democracies and finds evidence quite consistent with this approach. Paldam (1979) finds very weak evidence (if any at all) of Nordhaus' political business cycle on output and unemployment using a sample of seventeen OECD countries. The same author (1989a, b) reports stronger evidence of partisan effects using annual data. Alesina (1989) provides some qualitative tests with annual data using the same sample of countries; his results suggest that the RPT is broadly consistent with the evidence while the same paper does not find clear evidence of PBC on growth and unemployment. Alvarez, Garrett and Lange (1991) suggest that the degree of success of "partisan policies" may depend upon the characteristics of labour market institutions and of unions' behaviour. On the contrary, Sheffrin (1989) finds inconclusive results for

7. With the exception of Lindbeck (1976), very little attempt has been devoted to build political business cycle models (opportunistic and/or partisan) for small open economies. For such economies the exchange rate regime would greatly influence the options available to the politicians.

the RPT. However, his definition of "unexpected change" of governments is questionable. Sheffrin disregards the fact that in several countries the same party or coalition was elected repeatedly with no electoral uncertainty. (See Alesina (1991)).

The contribution of the present paper is that, unlike its predecessors, it considers all the different theories in a unified framework. Furthermore, unlike the recent work by Alesina, Paldam and Sheffrin, we use quarterly data rather than annual data and make use of different and more robust statistical tests. The use of quarterly data is important since the precise timing of cyclical fluctuations in relation to elections is crucial for the theories.

### 3. PANEL REGRESSIONS

#### 3.1. Data

We consider all the OECD countries which have been democracies in the sample period considered, which is 1960 to 1987. The extent of the sample is limited by availability of quarterly data; in fact, for some countries not all the series are available even for this period. The countries included are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Japan, Ireland, Italy, the Netherlands, New Zealand, Norway, Sweden, Switzerland, United Kingdom, and the United States.

The economic data are quarterly observations on inflation, output growth, and unemployment. Inflation is defined as the yearly rate of change of the CPI from IMF, IFS. Output growth is obtained as the rate of change of real GNP (or GDP), also from IMF, IFS. For unemployment, we use the total standardized unemployment rate from OECD. More details on country specific data issues can be found in Table A-1 in Appendix. The political data are election dates, the dates of changes of governments, and the political orientation of various governments. Dates of regime changes and elections do not always coincide in parliamentary systems in which changes of coalitions take place not only after elections. This information is summarized in Table A-2. Sources for these political data are Alt (1985) and Banks (1987). The identification of changes of political orientation of governments is usually unambiguous. Whenever ambiguities occurred in the case of coalition governments, we followed Alt's and Banks' conventions. It should be noted the countries for which positive results for the partisan theory are found, are those in which there are no ambiguities about the classification of government political orientation.

#### 3.2. Specification of empirical tests

The most direct way of testing the various theories is to run the following panel regressions of time-series cross-section data, for instance on output growth:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_n y_{t-n} + \alpha_{n+1} PDUM_t + \varepsilon_t \quad (5)$$

$y_t$  is the stacked vector of time-series data on output growth for the countries in the sample and PDUM is a political dummy which captures the implications of the different theories. The autoregressive specification for the dependent variable is chosen as the "best" using standard techniques. Similar regressions have been performed by McCallum (1978), Hibbs (1987), Alesina and Sachs (1988) and Alesina (1988a) on U.S. data. These tests are based upon the assumption that output growth and unemployment are generated

by a covariance-stationary stochastic process that can be expressed in autoregressive form as in (5).

Since the sample includes open economies (most of which are "small"), we must control for the effect of the world economy on domestic economies, for two reasons. First, the "partisan" or opportunistic goals of the politicians are likely to be defined, in small open economies, in relation to the rest of the world. Second, regardless of the governments' goals, international trade and financial linkages make OECD economies highly interdependent.

We have followed three approaches to capture these effects. The first one is to redefine each country's variable as a difference between the actual variable and a proxy for the OECD average of the same variable. The second one is to add as a regressor in equation (5) a proxy for a world or OECD average. The third one is to add time dummies in the regression. Our results concerning the relative performance of various political models are insensitive to the procedure used. As an indicator of an OECD average of each economic variable we consider the average of the seven largest economies in our sample, which are the U.S.A., Japan, Germany, France, the U.K., Italy, and Canada, weighted by each country's share of GNP over the total.<sup>8</sup>

In the remainder of this section we present results of panel regressions on the different political theories of the business cycle. We make use of a fixed-effect model with constant slopes. By doing so we take into account differences in long-term growth rates, unemployment, and inflation across countries but we assume that the other parameters of the model are constant and equal across countries.<sup>9</sup>

### 3.3. *The "rational partisan theory" (RPT)*

The political dummy used is:

$$\text{DRPTN} = \begin{cases} +1 & \text{in the } N \text{ quarters starting with that of} \\ & \text{a change of government toward the right} \\ -1 & \text{in the } N \text{ quarters starting with that of} \\ & \text{a change of government toward the left} \\ 0 & \text{otherwise.} \end{cases}$$

We tested the cases of  $N = 4, 6, 8$ . This choice of number of quarters is consistent with a wage contract model in which contracts have an average length of 1 or 2 years.

Note that the variable DRPTN assumes values different from zero only following actual changes of governments, but not after every election if the same government is reappointed. According to the RPT theory, inflation surprises and thus output fluctuations may occur even if an incumbent is re-appointed unexpectedly (Alesina, 1987). However, for long periods of time in many countries in the sample certain parties repeatedly won elections with virtually no political uncertainty. Furthermore, in countries with

8. In the regressions for the seven countries included in the creation of the proxy for the OECD averages, we have used different proxies which exclude the country in the left-hand side of the regression.

9. A priori, the correct dynamic specification of the model could differ across countries but in country by country regressions we found that the same AR specification is the best for almost all the countries in the sample. However, even if the same AR specification applies to each country, the estimates of the coefficients on the dynamic part of the model could differ across countries and suggest the use of a variable-slopes and variable-intercepts model. Given the loss of degrees of freedom involved, this procedure was not adopted. Country specific results are discussed in Section 4, and more extensively in the working paper version of this article.

endogenous timing of elections, which are the large majority of the sample (see Table A-2), in every period there is at least "some" probability that an election is called and that a change of government may occur. In addition, in parliamentary systems sometimes government changes occur in the middle of a term, with no elections. Rather than trying to estimate the degree of political uncertainty in every period, which would be rather difficult, we have chosen to estimate a somewhat weaker form of RPT, testing for temporary effects on real variables after actual *changes* of governments. An additional reason for doing so, is that several macroeconomic models in the "neo-Keynesian" tradition, imply that not only unexpected, but also expected aggregate demand policy may have some real effects. Thus, according to these approaches, stronger effects should be found after actual changes of governments, with actual changes of policies, relative to the case of reappointment of the same government.

Column (1) of Table 1 reports the result of the dynamic panel OLS regressions for the entire sample of countries and the time period for which data are available.<sup>10</sup> Japan and Switzerland are not included since they had no political change in the sample.<sup>11</sup> The dependent variable  $y$  is the rate of GNP growth defined as:

$$y = \left[ \frac{X_t - X_{t-4}}{X_{t-4}} \cdot 100 \right];$$

where  $X_t$  = level of real GNP in quarter  $t$ . The regressors are self explanatory:  $yw_t$  is the world growth average (described above); the AR (2) specification has been chosen as the "best" using standard techniques; the remaining regressors are country dummies.

The political dummy DRPT 6 has the correct sign and is statistically significant at the 1% confidence level: a change in government to the right (left) leads to a transitory fall (increase) in output growth. The one quarter lag in the political dummy is consistent with a reasonable interval between change of regime (in quarter  $t$ ) and change of policy (in period  $t+1$ ). The regressions with DRPT 4 and DRPT 8 (available upon request) yield analogous results: the pattern of the coefficients suggests that partisan effects are observable from about the second to the eighth/ninth quarters after the election. These results are consistent with findings on United States data by Alesina and Sachs (1988) and Alesina (1988a).<sup>12</sup>

In column 2 of the same table we present the result of the same regression for a subset of countries which have either a "pure" two-party system or at least more clearly identifiable "left" and "right" coalitions. There are U.S., U.K., France, Germany, Australia, New Zealand, Sweden, and Canada. The other countries in the sample have more fragmented political systems with governments formed with large coalitions of parties (often centre-left) which sometimes are short-lived and unstable. For obvious reasons, the second group of countries is less likely to exhibit regular partisan cycles. In the

10. It is known that in dynamic fixed-effects panel models the correlation between the error term and the lagged dependent variables might lead to biased and inconsistent estimates of the parameters (Hsiao (1986)). The problem is serious in panel sets where the number of agents ( $N$ ) considered is large but the number of time series observations ( $T$ ) is small. In that case, the maximum likelihood estimator of the dynamic model is inconsistent even if the number of agents becomes very large (Anderson and Hsiao (1982) and Nickell (1981)). The solution to this problem is to use instrumental-variable methods such as those suggested by Bhargava and Sargan (1983) and Pakes and Griliches (1984). Our panel data set, however, does not suffer of the above problem because of the use of a long time series (usually 112 data points). In the case where the time period  $T$  is large, the parameter estimates of the standard fixed-effects dynamic model are consistent (Hsiao (1986)).

11. Our results are unaffected by the inclusion of these two countries.

12. Analogous results on the political dummy available are obtained by using as a dependent variable the difference between domestic and world growth.

TABLE 1  
*Rational partisan theory*  
*Dependent variable: rate of growth of output (Y)*

Independent variables	(1) Coefficient ( <i>t</i> -statistics)	(2) Coefficient ( <i>t</i> -statistics)	(3) Coefficient ( <i>t</i> -statistics)	(4) Coefficient ( <i>t</i> -statistics)
Constant	0.130 (0.55)	-0.180 (-0.81)	0.500 (0.80)	1.282 (3.84)
Y(-1)	0.712 (28.1)	0.610 (17.27)	0.689 (25.59)	0.543 (15.04)
Y(-2)	-0.062 (-2.55)	-0.01 (-0.28)	-0.065 (-2.46)	-0.009 (-0.24)
YW	0.353 (11.82)	0.305 (9.33)	—	—
DRPT 6(-1)	-0.41 (-3.48)	-0.62 (-4.42)	-0.412 (-3.43)	-0.573 (-4.09)
U.S.A.	-0.45 (-1.49)	0.21 (0.81)	-0.376 (-1.19)	0.250 (1.10)
U.K.	-0.63 (-2.08)	-0.02 (-0.07)	-0.703 (-2.21)	-0.118 (-0.53)
Germany	-0.36 (-1.19)	0.28 (1.09)	-0.400 (-1.25)	0.187 (0.81)
France	-0.14 (-0.45)	0.49 (1.81)	-0.205 (-0.60)	0.440 (1.80)
Canada	0.12 (0.40)	0.80 (3.06)	0.073 (0.23)	0.798 (3.44)
Italy	0.04 (0.15)	—	0.000 (0.00)	—
Sweden	-0.56 (-1.63)	—	-0.600 (-1.64)	-0.079 (-0.30)
Belgium	-0.42 (-1.40)	—	-0.443 (-1.38)	—
Austria	0.14 (0.48)	—	0.161 (0.51)	—
Norway	-0.03 (-0.10)	—	0.001 (0.00)	—
Finland	0.05 (0.15)	—	0.005 (0.14)	—
Ireland	0.46 (1.13)	—	0.621 (1.40)	—
Australia	0.03 (-0.10)	0.007 (2.67)	0.000 (0.02)	0.658 (3.44)
New Zealand	-0.48 (-1.59)	0.12 (0.48)	-0.512 (-1.60)	—
Denmark	-0.39 (-1.29)	—	-0.399 (-1.24)	—
R <sup>2</sup>	0.61	0.60	0.63	0.64
S.E.	2.24	1.61	2.23	1.60

second regression, in fact, the coefficients on the political dummy are much larger in absolute value and even more precisely estimated.

The values of the coefficients in the second column of Table 1 imply that about eighteen months after a change of regime toward the right (left) the rate of growth of GNP is about 1.3% below (above) "normal". Thus, the difference in the rate of growth between the beginning of a left-wing government and the beginning of a right-wing government reaches a peak of about 2.6%.

Columns (3) and (4) of Table 1 report the same regressions in which we added time dummies, instead of the world growth variable. The 27 time dummies, one for each year

of our sample, are not reported in the Table. Column (3) includes all the countries; column (4) the same subset of countries of column (2). The coefficient on the DRPT dummy variable is virtually unchanged, relative to columns (1) and (2). Generally, we always found that our results on the political variables are totally unaffected by substituting the "world variable" with time dummies. Therefore, in what follows we report only the result using the world variable. The corresponding results obtained using time dummies are available upon request.

In Table 2 the dependent variable is the difference ( $U_t^{DIF}$ ) between the domestic unemployment rate, ( $U_t$ ) and the "OECD unemployment rate,"  $UW_t$ , defined analogously to the average GNP growth. In evaluating results on employment one has to be cautious because of problems of hysteresis (see, for instance, Blanchard and Summers (1986)). By taking the difference of domestic unemployment from a world weighted average, unit

TABLE 2  
*Test of rational partisan theory*  
Dependent variable:  $U_t^{DIF}$

Variable	(1) Coefficient ( <i>t</i> -statistic)	(2) Coefficient ( <i>t</i> -statistic)
Constant	0.152 (3.37)	0.101 (2.92)
$U^{DIF}(-1)$	1.284 (42.25)	1.332 (39.35)
$U^{DIF}(-2)$	-0.300 (-11.14)	-0.359 (-10.49)
DRPT 6(-2)	0.063 (3.11)	0.086 (3.20)
Australia	-0.126 (-2.12)	-0.082 (-1.67)
Austria	-0.218 (-3.31)	
Belgium	-0.030 (-0.54)	
Canada	-0.128 (-2.30)	-0.051 (-1.18)
Denmark	-0.093 (-1.51)	
Finland	-0.164 (-2.80)	
France	-0.110 (-1.83)	-0.06 (-1.25)
Germany	-0.140 (-2.31)	-0.116 (-2.17)
Ireland	0.071 (1.05)	
Italy	-0.072 (-1.30)	
Norway	-0.216 (-3.13)	
Sweden	-0.208 (-3.32)	-0.198 (-3.32)
U.K.	-0.132 (-2.34)	-0.082 (-1.80)
U.S.A.	-0.138 (-2.48)	-0.061 (-1.38)
R <sup>2</sup>	0.99	0.98
S.E.	0.35	0.31

roots problems are somewhat mitigated, but certainly not eliminated. Table 2 shows results which are quite consistent with those on GNP growth. The political dummy is significant at the 1% level and the fit improves when the sample is restricted to seven bi-partisan countries (note that New Zealand is missing from these regressions because of lack of quarterly unemployment data). The dummy DRPT6 is lagged two quarters to capture the slow response of unemployment to policy changes relative to output. In any case, analogous results (available upon request) are obtain if this variable is lagged only one quarter or when DRPT 4 and DRPT 8 are used. The values of the coefficients in the second column of Table 2 imply that about six quarters after a change of regime toward the right (left) the unemployment rate is about 1.5 percentage points above (below) normal.<sup>13</sup>

Let us now turn to inflation. The theory implies that one should observe permanent differences across governments on the inflation rate. Thus, we have defined a political dummy, RADM, as follows:

$$\text{RADM} = \begin{cases} +1 & \text{if a right-wing government is in office, including} \\ & \text{the quarter of the change of government} \\ -1 & \text{if a left-wing government is in office including} \\ & \text{the quarter of the change of government.} \end{cases}$$

In Table 3 the dependent variable is domestic inflation ( $\pi$ ) defined as the rate of change of CPI:

$$\pi_t = \left[ \frac{\text{CPI}_t - \text{CPI}_{t-4}}{\text{CPI}_{t-4}} \cdot 100 \right].$$

The variable for world inflation ( $\pi_W$ ) is defined analogously to the world output growth.<sup>14</sup>

In the first regression, which includes the entire sample of countries, the sign of the coefficient on RADM (-1) is correct and it is marginally insignificant at the 10% level ( $t = -1.65$ ). The second regression includes only the eight "bi-partisan" countries: here the coefficient on RADM (-1) is larger and significant at the 5% level.

The value of the coefficients in the second regressions imply a difference in the steady-state inflation rate between the two regimes of about 1.4%. This relatively low value reflects the fact that our sample includes the sixties, with a low and stable inflation and countries, such as Germany, with a low inflation rate throughout the sample period. We have run the same regressions of Table 3 for the post-fixed exchange rates regimes, from 1972 to 1987. In these regressions (available upon request) the coefficient on the RADM dummy is more precisely estimated and implies (in the sample of 8 "bi-partisan" countries) a difference in the inflation rate across political regimes of about 2.5 percent.

In fact, we have tested whether all the regressions of Table 1, 2 and 3 improve in the post-1971 period, since in the fixed exchange rate period (1960–1971 in our sample) the macroeconomic policies of each countries were more constrained and integrated. All the  $t$ -statistics on the political dummies improve and the value of the coefficients increase in absolute value in the post-1971 regressions. However, the problem in pursuing this comparison, pre- and post-1971, is that there are very few changes of regimes in the pre-71 period (see Table A-1); in many countries there are *no* changes of regimes in the

13. The variable  $U_t^{\text{DIF}}$  shows a high level of persistence. Thus, even a "temporary" policy shock has rather persistent effects.

14. One could add oil prices to the equation but the inclusion of the world inflation variable already proxies for this role of world-wide oil shocks.

TABLE 3  
*Rational partisan theory*  
*Dependent variable:  $\pi$*

Variable	(1) Coefficient ( <i>t</i> -statistic)	(2) Coefficient ( <i>t</i> -statistic)
Constant	-0.075 (-0.63)	0.593 (4.80)
$\pi(-1)$	1.085 (45.54)	1.210 (35.22)
$\pi(-2)$	-0.136 (-3.92)	-0.272 (-5.15)
$\pi(-3)$	-0.097 (-4.34)	-0.074 (-2.30)
$\pi W$	0.146 (13.15)	0.127 (9.35)
RADM (-1)	-0.05 (-1.65)	-0.084 (-2.17)
Australia	0.329 (2.15)	-0.307 (-2.27)
Austria	-0.064 (-0.42)	
Belgium	0.027 (0.18)	
Canada	0.070 (0.45)	-0.580 (-3.97)
Denmark	0.352 (2.27)	
Finland	0.41 (2.65)	
France	0.333 (2.18)	-0.293 (-2.19)
Germany	-0.255 (-1.66)	-0.853 (-5.68)
Ireland	0.66 (4.22)	
Italy	0.665 (4.13)	
New Zealand	0.66 (4.27)	
Norway	0.299 (1.94)	
Sweden	0.264 (1.69)	-0.405 (-2.83)
U.K.	0.516 (3.34)	-0.14 (-1.07)
U.S.A.	-0.041 (-0.27)	-0.65 (-4.55)
$R^2$	0.94	0.95
S.E.	1.13	0.98

sixties. Thus, the political dummies in the pre-1971 regression are very imprecisely estimated and hard to compare with the post-1971 sample.<sup>15</sup>

15. We also tested whether the dynamic process of inflation has changed moving from the fixed to the flexible rate system. The regressions of Table 3 were computed allowing the coefficients on the lagged dependent variable to be different before and after 1972. The results (available upon request) confirm that inflation is significantly more persistent in the post-1972 period with flexible rates. However, our results concerning the statistical significance of the variable RADM remain unchanged, even when we allow for a structural break in 1972. These results are available.

Finally, it is worth noting that in the inflation regression several of the coefficients on the country dummies are statistically significant, indicating, as it is well known, that different countries have had substantially different average inflation rates in the sample period considered here. An often cited explanation for these country differences is the degree of Central Bank independence (Alesina (1989), Grilli, Masciandaro and Tabellini (1991), Alesina and Summers (1992)). More independent Central Banks appear to have been associated with lower average inflation rates.<sup>16</sup>

In summary, these results are quite favourable to the RPT. The implication of this hypothesis is not rejected on both the level of economic activity (growth and unemployment) and inflation, particularly for a subset of countries with more clearly identifiable government changes from left to right and vice versa.<sup>17</sup>

### 3.4. "Partisan theory" with permanent effects

Hibbs' PT implies permanent differences in output and unemployment in addition to permanent differences in inflation across governments. Thus, one way of comparing the Hibbs' PT with the RPT is to run the same regressions of Tables 1 and 2 using the "permanent" partisan dummy RADM rather than the "transitory" political dummy DRPTN. The results are shown in Tables 4 and 5: all the coefficients on the political dummy are insignificant, even though with the right sign. In these tables the fixed effects coefficients are not reported since they are very similar to those of Tables 1 and 2. Additional regressions with alternative lag structures (for instance lagging RADM more than one quarter) yield no support for the theory.

TABLE 4  
*Partisan theory (Hibbs)*  
*Dependent variable: Y*

Variable*	(1) Coefficient (t-statistic)	(2) Coefficient (t-statistic)
Constant	0.12 (0.51)	-0.17 (-0.79)
Y(-1)	0.720 (28.47)	0.629 (17.76)
Y(-2)	-0.061 (-2.53)	-0.01 (-0.26)
YW	0.349 (11.62)	0.289 (8.77)
RADM (-1)	-0.03 (-0.54)	-0.02 (-0.35)
R <sup>2</sup>	0.61	0.59
S.E.	2.25	1.61

\* The estimated regression includes country fixed effects that are not reported in the table.

16. We thought about adding as a regressor in the inflation equation one of the index of Central Bank independence; (see for instance Alesina and Summers (1992)). However, such index assign numerical values to different countries and these numbers do not vary over time. Therefore, the country dummies which are already included in the regression capture the same effect.

17. The significance of the coefficients of the political dummies in Tables 1, 2 and 3 is not due to the predominant influence of any single country. If one drops any of the 16 countries and retains the other 15, the coefficients on the political dummies remain significant.

TABLE 5  
*Partisan theory (Hibbs)*  
*Dependent variable:  $U^{DIF}$*

Variable	(1) Coefficient (t-statistics)	(2) Coefficient (t-statistics)
Constant	0.14 (3.41)	0.016 (0.62)
$U^{DIF}(-1)$	1.29 (49.0)	1.43 (41.1)
$U^{DIF}(+1)$	-0.20 (-11.3)	-0.45 (-12.7)
RADM (-1)	0.009 (1.00)	0.0009 (-0.09)
$R^2$	0.98	0.98
S.E.	0.33	0.24

\* The estimated regression includes country fixed effects that are not reported in the table.

An additional test confirmed our results. We defined a new dummy variable DPRTNX, which is the "complement" of the DRTPN variable; that is, it takes the value of 1 during right-wing governments *after* the first  $N$  quarters, and -1 *after* the first  $N$  quarters of left-wing governments. We added this new variable in our panel regressions of Tables 1 and 2. The coefficient on this variable has the opposite sign to the DRPTN dummies and is statistically insignificant. The coefficients on the DRPTN dummies, instead, remain statistically highly significant. This test confirms that the effects of changes of governments on growth and unemployment are transitory.

The results on the RPT and PT viewed together, indirectly provide some empirical support to the inflation-bias model of Kydland and Prescott (1977) and Barro and Gordon (1983a, b). In fact, our regressions show that a permanent difference in inflation rate is associated with temporary deviations of output and unemployment from trend. Thus, the governments that are more concerned about growth and unemployment relative to inflation, after a temporary initial expansion, are caught in the sub-optimal equilibrium with an inflation bias. In fact, inflation remains high even though the level of economic activity returns to its "natural" value. This is precisely the feature of the sub-optimal time-consistent equilibrium.

### 3.5. The "political business cycle"

Nordhaus' (1975) PBC model can be tested on growth and unemployment by constructing a political dummy of the following form:

$$NRDN = \begin{cases} 1 & \text{in the } (N-1) \text{ quarters preceding an election} \\ & \text{and in the election quarter} \\ 0 & \text{otherwise.} \end{cases}$$

We have chosen  $N = 4, 6$  and  $8$ . A relatively short pre-electoral output expansion is consistent with this theory, which views the electorate as short-sighted, Nordhaus (1975, 1989). Furthermore, since in many countries in the sample several elections occur in less

than four year intervals, a longer specification of the pre-electoral period seems unreasonable.

Tables 6 and 7 report the results on output and unemployment for the 18 countries in the sample, using NRD 6. (The fixed-effect coefficients are not reported.) In both tables the coefficients of NRD 6 are insignificant; in the growth regression the coefficient has the opposite sign from the theory prediction. Several alternative specifications with NRD 4 and NRD 8, using the difference of domestic growth from the world as the dependent variable and alternative lag structures, yield no support for the theory. In fact, the coefficient on the political dummy has the "wrong" sign in the majority of the regressions.

We also tested whether the NRD dummy approaches statistical significance, when partisan effects are held constant. Regressions including both the DRPT and the NRD

TABLE 6  
*Test for political business cycle theory*  
Dependent variable: Y

Variable*	(1) Coefficient (t-statistic)	(2) Coefficient (t-statistic)
Constant	0.12 (0.49)	-0.19 (-0.81)
Y (-1)	0.732 (29.49)	0.631 (17.27)
Y (-2)	-0.059 (-2.48)	-0.015 (-0.43)
YW	0.344 (12.02)	0.280 (8.47)
NRD 6	-0.09 (-0.78)	0.06 (0.49)
R <sup>2</sup>	0.65	0.60
S.E.	2.25	1.73

\* The estimated regression includes country fixed effects that are not reported in the table.

TABLE 7  
*Political business cycle theory*  
Dependent variable: U<sup>DIF</sup>

Variable*	(1) Coefficient (t-statistic)	(2) Coefficient (t-statistic)
Constant	0.166 (3.75)	0.020 (0.68)
U <sup>DIF</sup> (-1)	1.323 (51.02)	1.433 (39.67)
U <sup>DIF</sup> (-2)	-0.336 (-12.81)	-0.446 (-12.08)
NRD 6	-0.011 (-0.64)	-0.001 (-0.63)
R <sup>2</sup>	0.99	0.98
S.E.	0.32	0.24

\* The estimated regression includes country fixed effects that are not reported in the table.

dummies were run, with no support for the PBC, while the DRPT dummy remained statistically significant (results are available).<sup>18</sup>

The PBC not only as formulated in Nordhaus (1975) but also, with caveats discussed above, in the "rational" models by Rogoff and Sibert (1988) and Persson and Tabellini (1990) implies an increase of the inflation rate around elections. Furthermore, governments may prefer to raise prices under their direct control after, rather than before elections, thus directly contributing to a post-electoral upward jump in inflation. We have tested this implication in Table 8, where the dummy ELE is defined as follows:

$$\text{ELE} = \begin{cases} 1 & \text{in the 4 quarters following an election, and in} \\ & \text{the election quarter.} \\ 0 & \text{otherwise.} \end{cases}$$

The dummy ELE is significant at the one per cent level. Additional regressions (available upon request) confirm that the upward jump in inflation does not occur before the election, but only in the election quarter and lasts three to five quarters.<sup>19</sup> If confirmed by direct findings on policy instruments, this result suggests that around elections monetary and fiscal policy instruments may be manipulated, even though these policies do not seem to affect real economic activity, as implied by Rogoff and Sibert (1988) and Rogoff (1990). Alesina, Cohen and Roubini (1992) present evidence on monetary and fiscal policy instruments which is consistent with this hypothesis.

Up to this point, the different theories have been tested separately, that is by including only one political dummy variable in each regression. Our results were also confirmed

TABLE 8  
*Political business cycle*  
Dependent variable:  $\pi$

Variable	Coefficient (t-statistic)
Constant	-0.131 (-1.08)
$\pi(-1)$	1.078 (46.90)
$\pi(-2)$	-0.113 (-3.36)
$\pi(-3)$	-0.113 (-5.23)
$\pi W$	0.141 (13.09)
ELE	0.263 (4.61)
$R^2$	0.93
S.E.	1.14

\* The estimated regression includes country fixed effects that are not reported in the table.

18. An even more extreme version of this model of voters' myopia would imply that they ignore the influence of the world economy on their countries' performance and thus politicians simply attempt to expand their economies, regardless of the world economy. This hypothesis can be tested running the same regressions of Tables 6, 7 without correcting for the effect of the world economy. The results (available upon request) show no support for the PBC.

19. The statistical significance of the coefficient on ELE is unaffected by allowing for a structural break of the inflation process in 1972. See also footnote 15.

when we ran a general nesting model. Specifically, we estimated regressions on growth and unemployment in which all three dummy variables, DRPT, RADM and NRD were included. Only the DRPT variable was significant. We also calculated an *F*-test comparing the unrestricted model with all the three political variables and the restricted model with only the DRPT: for both samples of countries we could not reject the restricted model, at very high levels of significance.

As far as inflation is concerned, we tested whether the two dummies RADM and ELE remain jointly significant when used as regressors in the same equation. For the smaller group of eight countries we reject the hypotheses that either one and both variables are zero. *F*-tests reject models in which either one or both variables are excluded. For the complete sample of countries, our *F*-tests continue to reject the hypothesis that the coefficient on ELE is zero, but we cannot reject, at standard levels of confidence, the hypothesis that the coefficient on RADM is zero. These different results on the two samples of countries are consistent with the evidence presented in Tables 3 and 8 above.<sup>20</sup>

#### 4. COUNTRY RESULTS

In this section we summarize the results obtained by performing country by country regressions with the same specification of the panel regressions. These results are presented more extensively in the working paper version of this article (NBER Working Paper no. 3478).

We begin with the RPT. Growth and unemployment regressions were run for each country using six dummies DRPTN ( $-J$ ) with  $N = 4, 6, 8$  and  $J = 1, 2$ . Countries may differ with regard to the time delay in implementing a new policy after a regime change or with regard to how persistent the transitory increase in output will be after the policy change. This is why the most appropriate specification for the DRPTN dummy may vary across countries. For inflation we run regressions using the dummy RADM ( $-J$ )  $J = 1, \dots, 5$ . Longer lags for inflation, relative to the growth regression, can be easily explained by the lag between output and inflation movements following changes in macroeconomic policies. Our results can be summarized in three points:<sup>21</sup>

(1) In seven countries, Australia, Denmark, Germany, France, New Zealand, the U.S., and the U.K., all the regressions on growth, inflation and unemployment show evidence favourable to the RPT, although not all the coefficients on the political variables are significant at the usual confidence levels (5 or 10%) in every regression. The results on the U.K. are greatly strengthened if the sample is restricted to the post-fixed rates period.<sup>22</sup>

(2) In seven other countries, Austria, Belgium, Finland, Ireland, the Netherlands, Norway, Sweden, the coefficients on the political dummies exhibit the sign predicted by the theory sometimes approaching statistical significance, in either the growth and/or the unemployment regressions. No significant results were found in the inflation regressions. For example, Sweden has a very strongly significant coefficient on DRPT 6 in the growth equation.

20. All the results of these *F*-tests are available from the authors. Also, there is no difference in the results of these tests regardless of whether we use time dummies or "world variables".

21. All the regressions from which the following results are derived are displayed in the NBER Working Paper version of this paper.

22. The significant difference between the pre- and post-1971 results for the U.K. is explained primarily by the observation of the Labour government elected in October 1964. This government, constrained by a commitment not to devalue the pound, could not pursue expansionary policies.

(3) Canada and Italy show no significant coefficients in any regressions. The case of Canada, however, is explained by the almost perfect correlation between the U.S. and Canadian business cycle. In fact the U.S. political dummies are statistically quite significant (five percent confidence level) in the Canadian equations! Thus, it is not clear whether for the purpose of this paper Canada really provides an independent observation.

In summary, six of the eight countries with more clearly identifiable left-right governments (that is the U.S., Germany, France, the U.K., Australia, and New Zealand) plus Denmark exhibit evidence of RPT effects.<sup>23</sup> All the parliamentary systems with large coalition governments show little sign of RPT, particularly on inflation.<sup>24</sup>

We find that the implications for growth and unemployment of the PT with permanent effects are rejected in every country except for Germany and for the borderline case of Sweden. All the other countries clearly reject the theory; in several cases the sign of the coefficient on the political dummy is opposite to the theory prediction.

In order to test the implication of the PBC model, we run the unemployment and growth regressions trying both the NRD 4 and NRD 6 dummies. In four countries, Germany, Japan, U.K. and New Zealand the coefficient on at least one of the NRD's is significant. In Australia and France the coefficients have the sign inconsistent with the theory and are statistically significant. In all the other countries the coefficients are insignificant.

Finally, we performed the PBC regressions on inflation, using the dummy ELE as for the panel regressions of Table 8. Several countries, such as Denmark, France, Germany, Italy, and New Zealand show significant (ten percent or better) post-electoral upward jumps in the inflation rate. In several other countries (e.g. Japan, Norway, and U.K.) the sign is correct but the *t*-statistic does not reach a significant level, although is above 1.

In summary, in only two countries, Germany and New Zealand, both the level of economic activity and the inflation rate follow the predictions of Nordhaus' PBC model. The results for New Zealand are not too unexpected, given that, until recently this country had one of the least independent Central Banks. On the contrary, the case of Germany, with a Central Bank with a strong reputation for independence, appears somewhat surprising.<sup>25</sup>

## 5. TESTS OF THE RPT MODEL USING HAMILTON'S MODEL

One of the strongest and most interesting results which we have highlighted thus far, is that downturns and upsurges in growth tend to follow changes of governments, as predicted by the RPT. In this section we pursue this observation further by deriving direct measures of the dating of the business cycle in different countries and study their relation to the dates of government changes. From a conceptual point of view, the recent

23. Our results regarding the RPT theory on output growth and unemployment are confirmed by another set of regressions in which we used a distributed lag of the variable representing the changes in partisan regime, instead of the DRPT variable. For the countries with a significant DRPT effect we find that an *F*-test on the distributed lag variable rejects the null hypothesis that the sum of all coefficients is equal to zero.

24. A referee has noted that in this group of eight countries we have several "large" economies, while many of the countries with coalition governments are small and very open economies and has suggested that it may be the size and the degree of openness of the economy which affect the government's ability to implement partisan macroeconomic policies.

25. It is interesting to note that in his original paper, Nordhaus (1975) had found support for his theory precisely on these two countries!

literature on unit roots and GNP has offered a number of alternative approaches to the problem of distinguishing between trend and cyclical components of output. Most of the literature<sup>26</sup> is based on the assumption that GNP growth is characterized by a linear stationary process.

TABLE 9

*Test of the RPT model using Hamilton's filter for the business cycle  
Dependent variable: probability of being in a low (relative to average OECD) growth state. Panel regressions on seven bi-partisan countries*

Variable	Estimated coefficient	t-statistic
Constant	0.09	4.74
PROBS ( <i>t</i> - 1)	0.82	38.43
DRPT 6	0.046	3.10
United States	-0.08	-3.26
Germany	0.05	2.33
France	0.002	0.01
Australia	-0.024	-1.05
Canada	-0.020	-0.891
Sweden	-0.029	-1.14
<i>R</i> <sup>2</sup> = 0.80		
S.E. = 0.167		

TABLE 10

*Test of the RPT model using Hamilton's filter for the business cycle  
Dependent variable: probability of being in a low (relative to average OECD) growth state. Time series regressions on seven bi-partisan countries*

Country	Constant	Lagged dep. variable	DRPTN	R2	D.W.
United States	0.02 (1.55)	0.73 (12.0)	0.08 <sup>a</sup> (3.07)	0.67	2.08
Germany	0.14 (3.31)	0.83 (16.9)	0.06 <sup>a</sup> (1.59)	0.80	1.60
France	0.04 (1.65)	0.92 (23.9)	0.09 <sup>b</sup> (2.10)	0.88	1.83
Australia	0.11 (3.04)	0.71 (10.4)	0.10 <sup>c</sup> (1.83)	0.53	1.98
Canada	0.05 (2.01)	0.85 (16.7)	0.04 <sup>b</sup> (0.71)	0.73	1.93
Sweden	0.13 (3.80)	0.62 (6.66)	0.05 <sup>c</sup> (2.04)	0.51	1.96
United Kingdom	0.20 (4.68)	0.61 (7.73)	0.02 <sup>b</sup> (0.81)	0.37	1.95
United Kingdom	0.21 (4.52)	0.56 (6.19)	0.06 <sup>a</sup> (1.87)	0.41	1.93

<sup>a</sup> DRPT 6.<sup>b</sup> DRPT 4.<sup>c</sup> DRPT 8.<sup>d</sup> 1967-1987 Sample.

26. See Nelson and Plosser (1982), Campbell and Mankiw (1987), Watson (1986), Clark (1987), King, Plosser, Stock and Watson (1987).

Hamilton (1989) studies the implications of specifying the first differences of log GNP as a non-linear stationary process. His idea is to view the economy as characterized by two states, a high-growth (expansion) state and a low-growth (recession) state and model the switch between these two states as being governed by a Markov process. One of the bi-products of the estimation of the model is a non-linear filter that delivers optimal estimates of the dating of the business cycle based on past observations on output. In particular, for each quarter the filter provides an estimate of the probability that the economy is in a recession (or a boom) given the information available in the data. Given Hamilton's success in characterizing the U.S. business cycles, using his filter, we used the same statistical approach to derive estimates of the dating of the business cycle for other OECD countries in our sample. After having done that, we tested the relation between these estimates of the business cycle and the changes of governments.

Given the positive evidence in favour of the RPT model for countries with a political system close to a two-party structure, we have considered only these countries.<sup>27</sup> To control for the effects of the world business cycle on the growth rate of the various economies, in our maximum likelihood estimates of Hamilton's model we use  $y^{DIF}$  (defined as the difference between country  $i$  growth rate and the average growth rate of the major OECD countries) as the two-state variable to be explained.

Once we have obtained an estimate of the dating of the business cycle, we perform a regression of the estimated probability (PROBS) of being in a low-growth state (relative to the OECD average) on a constant, the dummy for the RPT model (DRPTN) and the first lag of the dependent variable. The latter is introduced to capture the observed persistence of the probability of being in a particular state of the world. The basic regressions is:

$$\text{PROBS}_t = \alpha_0 + \alpha_1 \text{PROBS}_{t-1} + \alpha_2 \text{DRPTN}_{t-1} + \varepsilon_t. \quad (6)$$

Table 9 reports the results of fixed-effects panel regressions of equation (6) above for 7 countries with a political structure close enough to a two-party system, as discussed above.<sup>28</sup> The political variable (DRPT 6), used to capture the effects of the RPT model, has the correct sign and is statistically significant at the 1% confidence level.

In Table 10 we report the results of separate time series regressions for each of the seven countries considered. The coefficient on the RPT dummy (DRPT) is significant at the 10% confidence level or better in five of the seven countries: United States, Germany, France, Australia and Sweden. The DRPT coefficient for the United Kingdom is significant (at the 10% level) only if we start the sample in 1970 (as discussed in Section 4). The seventh country, Canada does not show statistical significance for the RPT variable, as discussed in the previous section.

## 6. CONCLUSIONS

The most interesting result of this paper is that the more recent models of political cycles significantly outperform their predecessors. The rational partisan model by Alesina (1987) and the rational "opportunistic" model by Rogoff and Sibert (1988) are consistent with the overall pattern of results for several countries.

27. Given our country results, Hamilton's tests on the other countries are not likely to support the theory.  
 28. New Zealand is excluded because of lack of quarterly data.

The main findings of this paper can be summarized as follows:

(1) With the exception of two countries (Germany and New Zealand), we found no evidence of a systematic opportunistic cycle of the Nordhaus type either for output or unemployment.<sup>29</sup>

(2) The data show an electoral cycle on the inflation rate, consistent with the models of budget cycles of Rogoff and Sibert (1988).

(3) The implications of the "rational partisan theory" are consistent with the empirical evidence particularly for a subset of countries with a bi-partisan system or with clearly identifiable movements from left to right and vice-versa. This theory is less applicable, and in fact tends to fail, in countries with large coalition governments with frequent government collapses.

(4) The "partisan theory" with permanent effects on output and unemployment is generally rejected.

Thus, a political cycle which seems to appear fairly consistently in several countries is the following: left-wing governments expand the economy when elected; for a while (about 2 years) they succeed, then inflation expectations adjust and the economy returns to its natural rate of growth. At this point, left-wing governments are trapped into the time-consistent equilibrium with an inflation bias à la Barro and Gordon (1983b). Note that, when left-wing governments approach the new election in this high inflation, they may try to reduce the latter, particularly if inflation is perceived as the main economic problem of the time (Lindbeck (1976)). When right-wing governments are elected they fight inflation, causing a recession or a growth slowdown. Later in their term, the economy goes back at its natural rate of growth and inflation remains low.

Two explanations can account for the relatively little evidence of a Nordhaus-type opportunistic cycle on growth and unemployment: first, a "rational" electorate imposes a limit on this behaviour; an excessive attempt to pursue opportunistic policies may be perceived as counterproductive by policymakers. Second, it may be quite difficult to create expansions precisely timed before elections.

However, the results on post-electoral inflation increase may signal the occurrence of pre-electoral opportunistic budget policies. In fact, Alesina (1989) and Alesina, Cohen and Roubini (1992) show that budget deficits and money growth tend to increase in election years in several OECD democracies. Similar evidence on budget cycles in the U.S. is also discussed in Tufte (1978), Alesina (1988a), and Nordhaus (1989). Pre-electoral fiscal "favours" to key constituencies may be electorally very useful and easy to implement, relative to an attempt to increase the rate of growth of GNP. These opportunistic monetary and fiscal policies can very well co-exist with the partisan cycles found in the data. Even "partisan" politicians prefer to be in office, rather than out; by being in office they can implement their desired goals. Thus, they may engage in short term pre-electoral opportunistic policies if the latter enhance their chances for re-election.

29. Following Ito (1990), in the Working Paper version of this article, we also tested for an opportunistic model with "endogenous timing of elections." We checked whether the probability that early elections are called is affected by the state of the economy. We confirmed Ito's results on Japan, but we did not find supporting evidence for this hypothesis in any other country in our sample.

## APPENDIX

TABLE A-1

*Description of data*

**Inflation:** Inflation is obtained as:  $\pi_t = [(P_t - P_{t-4})/P_{t-4}] \times 100$  where  $P_t$  is the Consumer Price Index in quarter  $t$ . For all countries the sample is 1960:1, 1987:4, and CPI is taken from line 64 of IMF-IFS.

**Output and Unemployment:** Country-by-country sample and sources.

**Note:** Countries which use other measures of GDP do so because real quarterly GDP is not available.

Australia	GDP – real quarterly GDP from OECDMEI (1960:1–1987:4). Unemployment – unemployment rate – adjusted – OECDMEI (1965:1–1987:4).
Austria	GDP – real quarterly GDP from IMF-IFS (1960:1–1987:4). Unemployment – unemployment rate – total – adjusted – OECDMEI (1969:1–1986:4).
Belgium	GDP – quarterly Industrial Production from OECDMEI (1960:1–1987:4). Unemployment – unemployment rate – total insured – adjusted – OECDMEI (1960:1–1987:4).
Canada	GDP – real quarterly GDP from IMF-IFS (1960:1–1987:4). Unemployment – unemployment rate – total – adjusted – OECDMEI (1960:1–1987:4).
Denmark	GDP – real ANNUAL GDP from IMF-IFS (1960:1–1987:4) (converted into quarterly data by assuming that quarter-to-quarter annual change corresponds to year-to-year change). Unemployment – unemployment rate – registered – unemployed – adjusted – OECDMEI (1970:1–1987:4).
Finland	GDP – real quarterly GDP from IMF-IFS (1970:1–1987:4). Unemployment – unemployment rate – total – adjusted – OECDMEI (1960:1–1987:4).
France	GDP – real quarterly GDP from IMF-IFS (1965:1–1987:4). Unemployment – unemployment rate – total – adjusted – OECDMEI (1967:1–1987:4).
Germany	GDP – real quarterly GDP from IMF-IFS (1960:1–1987:4). Unemployment – unemployment rate – adjusted – OECDMEI (1965:1–1987:4).
Ireland	GDP – quarterly Industrial Production from OECDMEI (1975:1–1986:4). Unemployment – unemployment rate – adjusted – OECDMEI (1975:1–1987:4).
Italy	GDP – real quarterly GDP from IMF-IFS (1960:1–1987:4). Unemployment – unemployment rate – adjusted – OECDMEI (1960:1–1987:4).
Japan	GDP – real quarterly GDP from IMF-IFS (1960:1–1987:4). Unemployment – unemployment rate – adjusted – OECDMEI (1965:1–1987:4).
Netherlands	GDP – quarterly Industrial Production from OECDMEI (1960:1–1987:4). Unemployment – unemployment rate – registered unemployed – OECDMEI (1971:1–1987:4), no adjusted available.
New Zealand	GDP – real ANNUAL GDP from IMF-IFS (1960:1–1987:4) (converted into quarterly data by assuming that quarter-to-quarter annual change corresponds to year-to-year change). Unemployment – not available.
Norway	GDP – real ANNUAL GDP from IMF-IFS (1960:1–1987:4) (converted into quarterly data by assuming that quarter-to-quarter annual change corresponds to year-to-year change). Unemployment – unemployment rate – adjusted – OECDMEI (1972:1–1987:4).
Sweden	GDP – real quarterly GDP from IMF-IFS (1969:1–1987:4). Unemployment – unemployment rate – total insured – adjusted – OECDMEI (1969:1–1983:4).
Switzerland	GDP – real quarterly GDP form IMF-IFS (1967:1–1986:4). Unemployment – ratio of total unemployed to labor force – adjusted – OECDMEI (1974:4–1987:3).
U.K.	GDP – real quarterly GDP form IMF-IFS (1960:1–1987:4). Unemployment – unemployment rate – registered – civilian – adjusted – OECDMEI (1960:1–1987:4).
U.S.A.	GDP – real quarterly GDP from IMF-IFS (1960:1–1987:4). Unemployment – unemployment rate – total – adjusted – OECDMEI (1960:1–1987:4).

TABLE A-2  
*Election and regime change*  
E = Election; CH L = Change Left; Ch R = Change Right

AUSTRALIA: Endogenous Timing, 3 Yrs				AUSTRIA: Endogenous Timing, 4 Yrs			
1961:4	E	RIGHT	a	1959:2	E	RIGHT	c
1963:4	E			1962:4	E	E	
1966:4	E			1966:1	E	CH R	
1969:4	E			1970:1	E	CH L	
1972:4	E	CH L		1971:4	E		(*)
1974:2	E		(*) <sup>b</sup>	1975:4	E		
1975:4	E	CH R		1979:2	E		
1977:4	E			1983:2	E	CH R	c
1980:4	E			1986:4	E	CH R	
1983:1	E	CH L					
1984:4	E		(*)				
1987:3	E						
BELGIUM: Endogenous Timing, 4 Yrs				CANADA: Endogenous Timing, 5 Yrs			
1961:1	E	RIGHT		1962:2	E	RIGHT	
1965:2	E			1963:2	E	CH L	(*)
1968:1	E	CH L		1965:4	E		
1971:4	E			1968:2	E		
1973:1		CH R		1972:4	E		
1974:1	E			1974:3	E		(*)
1977:2	E	CH L		1979:2	E	CH R	
1978:4	E		(*)	1980:1	E	CH L	(*)
1981:4	E	CH R		1984:3	E	CH R	
1985:4	E	CH L					
1987:4	E						
DENMARK: Endogenous Timing, 4 Yrs				FINLAND: Endogenous Timing, 4 Yrs			
1960:4	E	LEFT		1962:1	E	LEFT	
1964:3	E			1963:4		CH R	
1966:4	E			1966:1	E	CH L	
1968:1	E	CH R	(*)	1970:1	E		
1971:3	E	CH L		1972:1	E		
1973:4	E	CH R		1975:3	E	CH R	
1975:1	E	CH L	(*)	1977:2	E	CH L	
1977:1	E			1979:1	E		
1979:4	E			1983:1	E	CH R	
1981:4	E			1987:1	E	CH R	
1982:3		CH R					
1984:1	E						
1987:3	E						
FRANCE: Endogenous Timing, 5 Yrs				GERMANY: Endogenous Timing, 4 Yrs			
1962:4	E	RIGHT		1961:3	E	RIGHT	
1967:1	E			1965:3	E		
1968:2	E		(*)	1966:4		CH L	c
1973:1	E			1969:3	E	CH L	
1978:1	E			1972:4	E	CH R	
1981:2	E	CH L		1976:4	E		
1984:3		CH R		1980:4	E		
1986:1	E	CH R		1982:4		CH R	
				1983:1	E		
				1987:1	E		

## IRELAND: Endogenous Timing, 5 Yrs

1961:4	E	RIGHT
1965:2	E	
1969:2	E	
1973:1	E	CH L
1977:2	E	CH R
1981:2	E	CH L
1982:1	E	CH R
1982:4	E	CH L
1987:1	E	(*)

## ITALY: Endogenous Timing, 5 Yrs

1962:4		RIGHT
1963:2	E	CH L
1968:2	E	
1972:2	E	
1974:4		CH R
1976:2	E	CH L
1979:2	E	
1983:2	E	
1987:2	E	

## JAPAN: Endogenous Timing, 4 Yrs

1960:4	E	RIGHT
1963:4	E	
1967:1	E	
1969:4	E	
1972:4	E	
1976:4	E	
1979:4	E	
1980:2	E	(*)
1983:4	E	
1986:3	E	

## NETHERLANDS: Endogenous Timing, 4 Yrs

1959:1	E	RIGHT
1963:2	E	
1965:2		CH L
1967:1	E	CH R
1971:1	E	
1972:4	E	
1973:2		CH L
1977:2	E	
1977:4		CH R
1981:2	E	CH L
1982:3	E	CH R
1986:2	E	

## NEW ZEALAND: Endogenous Timing, 3 Yrs

1960:4	E	RIGHT
1963:4	E	
1966:4	E	
1969:4	E	
1972:4	E	CH L
1975:4	E	CH R
1978:4	E	
1981:4	E	
1984:3	E	CH L
1987:3	E	

## NORWAY: Exogenous Timing, 4 Yrs

1961:3	E	LEFT
1965:3	E	CH R
1969:3	E	
1971:4		CH L
1972:4		CH R
1973:3	E	CH L
1977:3	E	
1981:3	E	CH R
1985:3	E	
1986:2		CH L

SWEDEN: Exogenous Timing, 3 Yrs  
since late 60s, constitutional reform

1960:3	E	LEFT
1964:3	E	
1968:3	E	
1970:3	E	
1973:3	E	
1976:3	E	CH R
1979:3	E	
1982:3	E	CH L
1985:3	E	

## SWITZERLAND: Exogenous Timing, 4 Yrs

1959:4	E	RIGHT
1963:4	E	
1967:4	E	
1971:4	E	
1975:4	E	
1979:4	E	
1983:4	E	
1987:4	E	

## U.K.: Endogenous Timing, 5 Yrs

## U.S.A.: Exogenous Timing, 4 Yrs

					RIGHT
1959:4	E	RIGHT		1960:4	CH L
1964:4	E	CH L		1964:4	E
1966:1	E		(*)	1968:4	CH R
1970:2	E	CH R		1972:4	E
1974:1	E			1976:4	CH L
1974:3	E	CH L	(*)	1980:4	CH R
1979:2	E	CH R		1984:4	E
1983:2	E				
1987:2	E				

a RIGHT or LEFT indicates the type of government in power at the beginning of the sample which is 1959:1. We also indicate for each country whether elections dates are endogenous or exogenous and the official number of years between two elections.

b Elections denoted with an asterisk "\*" are not included in tests of the political business cycle theory because they are too close (less than two years) to previous elections. They are however included in tests of the opportunistic endogenous election model.

c Both Germany and Austria had grand coalitions of Left and Right parties. Thus, a finer administration variable was used in the the RPT inflation and partisan (Hibbs) regressions. This also explains the occurrence of a rightward shift from an already central Right leaning party.

Source: Election Dates are obtained from Banks (1989); dates of changes of government and their classification of "Right" and "Left" are obtained from Alt (1985) and Banks (1989).

**Acknowledgements.** Prepared for the conference on Economic Policy in Political Equilibrium, Stockholm, June 14-16, 1990. We are especially grateful to our discussant Assar Lindbeck and to many other conference participants for very useful comments. James Hamilton generously shared his computer programs. Mathias Dewatripont, Benjamin Friedman, Sule Ozler, Torsten Persson, Dani Rodrik, James Stock, Philippe Weil, three anonymous referees and participants of seminars at Carnegie-Mellon, Chicago, Harvard, M.I.T., NBER, North Carolina, Trade Union Institute in Stockholm, and NBER Summer Institute provided very helpful comments on earlier drafts. Gerald Cohen was our excellent research assistant. The first draft of this paper was written while Alesina was an Olin Fellow at the NBER; he gratefully acknowledges financial support from the Olin and Sloan Foundations.

## REFERENCES

- ALESINA, A. (1987), "Macroeconomic Policy in a Two-Party System as a Repeated Game", *Quarterly Journal of Economics*, **102**, 651-678.
- ALESINA, A. (1988a), "Macroeconomics and Politics", *NBER Macroeconomics Annual*, (Cambridge: M.I.T. Press) 13-52.
- ALESINA, A. (1988b), "Credibility and Policy Convergence in a Two-Party System with Rational Voters", *American Economic Review*, **78**, 796-806.
- ALESINA, A. (1989), "Politics and Business Cycles in Industrial Democracies", *Economic Policy*, **8**, 55-87.
- ALESINA, A. (1990), "Evaluating Rational Partisan Business Cycle Theory: A Response", *Economics and Politics*, **3**, 63-72.
- ALESINA, A., COHEN, G. and ROUBINI, N. (1992), "Macroeconomic Policy and Elections in OECD Democracies", *Economics and Politics*, **5**, 1-30.
- ALESINA, A., LONDREGAN, J. and ROSENTHAL, H. (1990), "A Political-Economy Model of the United States" (unpublished).
- ALESINA, A. and SACHS, J. (1988), "Political Parties and the Business Cycle in the United States, 1948-1984", *Journal of Money, Credit and Banking*, **20**, 63-82.
- ALESINA, A. and SUMMERS, L. (1992), "Central Bank Independence and Macroeconomic Performance: Some Comparative Evidence", *Journal of Money, Credit and Banking* (forthcoming).
- ALT, J. (1985), "Political Parties, World Demand, and Unemployment: Domestic and International Sources of Economic Activity", *American Political Science Review*, **79**, 1016-1040.
- ALVAREZ, M., GARRET, G. and LANGE, P. (1989), "Government Partisanship, Labor Organizations and Macroeconomic Performance", *American Political Science Review*, **83**, 539-556.
- ANDERSON, T. W. and HSIAO, C. (1982), "Formulation and Estimation of Dynamic Models Using Panel Data", *Journal of Econometrics*, **18**, 47-82.

- BANKS, A. (1987) *Political Handbook of the World*.
- BARRO, R. and GORDON, D. (1983a), "Rules, Discretion, and Reputation in a Model of Monetary Policy", *Journal of Monetary Economics*, 12, 101-122.
- BARRO, R. and GORDON, D. (1983b), "A Positive Theory of Monetary Policy in a Natural Rate Model", *Journal of Political Economy*, 31, 589-610.
- BHARGAVA, A. and SARGAN, J. D. (1983), "Estimating Dynamic Random Effects Models from Panel Data Covering Short Time Periods", *Econometrica*, 51, 1635-1659.
- BLANCHARD, O. and SUMMERS, L. (1986), "Hysteresis and the European Unemployment Problem", *NBER Macroeconomics Annual 1986*.
- CAMPBELL, J. Y. and MANKIW, N. G. (1987), "Permanent and Transitory Components in Macroeconomic Fluctuations", *American Economic Review (Papers and Proceedings)*, 77, 111-117.
- CLARK, P. K. (1987), "The Cyclical Component of U.S. Economic Activity", *Quarterly Journal of Economics*, 102, 797-814.
- FREY, B. and SCHNEIDER, F. (1978), "An Empirical Study of Politico-Economic Interaction in the United States", *The Review of Economics and Statistics*, 60, 174-183.
- GRILLI, V., MASCIANDARO, D. and TABELLINI, G. (1991), "Political and Monetary Institutions and Public Financial Policies in the Industrial Countries", *Economic Policy*, 13, 341-392.
- HAMILTON, J. (1989), "A New Approach to the Economic Analysis of Non-Stationary Time Series and the Business Cycle", *Econometrica*, 57, 357-384.
- HIBBS, D. (1977), "Political Parties and Macroeconomic Policy", *The American Political Science Review*, 71, 1467-1487.
- HIBBS, D. (1987) *The American Political Economy* (Cambridge, MA: Harvard University Press).
- HSIAO, C. (1986) *Analysis of Panel Data* (Cambridge: Cambridge University Press).
- ITO, T. (1990), "The Timing of Elections and Political Business Cycles in Japan", *Journal of Asian Economics*, 1, 135-146.
- KING, R., PLOSSER, C., STOCK, J. and WATSON, M. (1987), "Stochastic Trends and Economic Fluctuations" (NBER Working Paper No. 2229).
- KYDLAND, F. and PRESCOTT, E. (1977), "Rules Rather Than Discretion: The Inconsistency of Optimal Plans", *Journal of Political Economy*, 85, 473-490.
- LINDBECK, A. (1976), "Stabilization Policies in Open Economies with Endogenous Politicians", *American Economic Review (Papers and Proceedings)*, 66, 1-19.
- McCALLUM, B. (1978), "The Political Business Cycle: An Empirical Test", *Southern Economic Journal*, 44, 504-515.
- NELSON, C. R. and PLOSSER, C. I. (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications", *Journal of Monetary Economics*, 10, 139-162.
- NICKELL, S. (1981), "Biases in Dynamic Models with Fixed Effects", *Econometrica*, 49, 1399-1416.
- NORDHAUS, W. (1975), "The Political Business Cycle", *Review of Economic Studies*, 42, 169-190.
- NORDHAUS, W. (1989), "Alternative Models to Political Business Cycles", *Brookings Papers on Economic Activity* 1989, No. 2.
- PALDAM, M. (1979), "Is There An Electoral Cycle?", *Scandinavian Journal of Economics*, 0, 000-000.
- PALDAM, M. (1989a), "Politics Matter After All: Testing Alesina's Theory of RE Partisan Cycles" (Aarhus University Working Paper).
- PALDAM, M. (1989b), "Politics Matter After All: Testing Hibbs' Theory of Partisan Cycles" (Aarhus University Working Paper).
- PAKES, A. and GRILICHES, Z. (1984), "Estimating Distributed Lags in Short Panels with an Application to the Specification of Depreciation Patterns and Capital Stock Constructs", *Review of Economic Studies*, 51, 243-262.
- PERSSON, T. and TABELLINI, G. (1990) *Macroeconomic Policy, Credibility and Politics* (London: Harwood Academic Publishers).
- ROGOFF, K. (1990), "Equilibrium Political Budget Cycles", *American Economic Review*, 80, 21-36.
- ROGOFF, K. and SIBERT, A. (1988), "Equilibrium Political Business Cycles", *Review of Economic Studies*, 55, 1-16.
- SHEFFRIN, S. (1989), "Evaluating Rational Partisan Business Cycle Theory", *Economics and Politics*, 1, 239-259.
- TERRONES, M. (1987), "Macroeconomic Policy Cycles under Alternative Electoral Structures: A Signalling Approach" (unpublished).
- TUFT, E. (1978) *Political Control of the Economy* (Princeton, N.J.: Princeton University Press).
- WATSON, M. W. (1986), "Univariate Detrending Methods with Stochastic Trends", *Journal of Monetary Economics*, 18, 49-75.

## The simple analytics of slack-maximizing bureaucracy

PAUL GARY WYCKOFF

*School of Public and Environmental Affairs, Indiana University, Bloomington, IN 47405*

Submitted 31 March 1989; accepted 22 May 1989

**Abstract.** Using a utility-based graphical model of bureaucratic choice, this paper develops four empirical predictions from the theory of slack-maximizing bureaucracy. These predictions are compared to those resulting from the Niskanen budget-maximizing model of bureaucracy. Slack-maximizing and budget-maximizing bureaucracies are similar in their response to changes in cost and in their generation of "flypaper effects", but they differ in their responses to matching and lump-sum grants.

### 1. Introduction

Since the publication of *Bureaucracy and Representative Government* (1971), analysts have debated Niskanen's simple assumption of budget-maximization. This controversy has arisen because the bureaucrat confronts his legislative sponsor with an all-or-nothing choice, and all-or-none demand curves are always elastic (see Patinkin, 1963, and Section 2 below). Therefore, to maximize the size of his budget, a bureaucrat would have to reduce the price charged to the sponsor to the level of his costs, eliminating productive inefficiency. This means that there would be no discretionary funds available for increased staff, perquisites of office, high salaries, etc., over and above what is minimally required to produce the services of the bureaucracy (Migue and Belanger, 1974).

As a result of this facet of Niskanen's model, a number of writers have proposed "slack-maximizing" models of bureaucracy (see Migue and Belanger, 1974; Breton and Wintrobe, 1975; and Gonzalez and Mehay, 1985). In these models, the bureaucrat tries to maximize the difference between his revenue and the minimum cost of production. Following Cyert and March (1963), I will call this difference "organizational slack", although Migue and Belanger call it "discretionary profit" and Orzechowski (1977) uses the term "fiscal residuum". This slack can then be used to purchase whatever non-productive expenditures the bureaucrat desires. The slack-maximizing approach also incorporates the models of Williamson (1964) and others, in which the bureaucrat wishes to maximize his staff. Under Williamson's assumption, the bureaucrat maximizes organizational slack and turns the whole of this slack into extra personnel.

Niskanen has carefully detailed the empirical consequences of budget-

maximizing bureaucracy on the expenditure patterns of the sponsor, but no similar analysis has been done for slack-maximizing bureaucracy. As a result, it is impossible for empirical analysts to tell if a slack-maximizing bureaucracy is operating, or whether a particular bureaucracy is budget- or slack-maximizing. This paper tries to fill this gap. Using a simple graphical model of expenditures, the four major empirical effects of slack-maximization are derived. For comparison purposes, equivalent results for budget-maximizing bureaucracy are also detailed. I conclude that there is surprisingly little empirical difference between slack-maximizing and budget-maximizing bureaucracies, so that in many circumstances it will not be necessary to pin down the bureaucrat's motives to predict his behavior well. There are, however, two empirical differences that can be used to discriminate between slack-maximizing and budget-maximizing bureaucracies if the analyst is interested in that question.

## 2. Equilibrium conditions

One major difficulty with previous graphical models of bureaucracy is their reliance on the sponsor's ordinary demand curve for analysis. The bureaucrat's position is taken to be similar to that of an ordinary private monopolist, in which consumers can choose the amount of a good they wish to purchase at any given price. The difficulty, of course, as pointed out by Niskanen, is that the sponsor trades a budget for a level of output without the opportunity for marginal changes in the level of consumption. Because of an information advantage and/or agenda control, the bureaucrat presents the sponsor with an all-or-nothing choice. This choice, therefore, is fundamentally different from that presented by an ordinary monopolist.

Most analysts have recognized the nature of this all-or-nothing choice, and have tried to incorporate this fact into their ordinary-demand-curve based analysis. But the ordinary demand curve is fundamentally inappropriate for this work, because it is based on the assumption that consumers are able to make marginal changes. This means that it is impossible to derive all of the empirical consequences of bureaucracy using ordinary demand curves. The alternative is to use all-or-none demand curves, or as is done below, use the sponsor's indifference curves as the primary units of analysis. McGuire (1973) pioneered the use of indifference curves to predict the responses of bureaucracy to changes in its environment.

Figure 1 shows the equilibrium conditions for the slack-maximizing and budget-maximizing bureaus. Let the term "cost" represent the true minimum resources necessary to produce the public good, which will be lower than the price charged to the sponsor in the case of slack-maximizing bureaucracies.

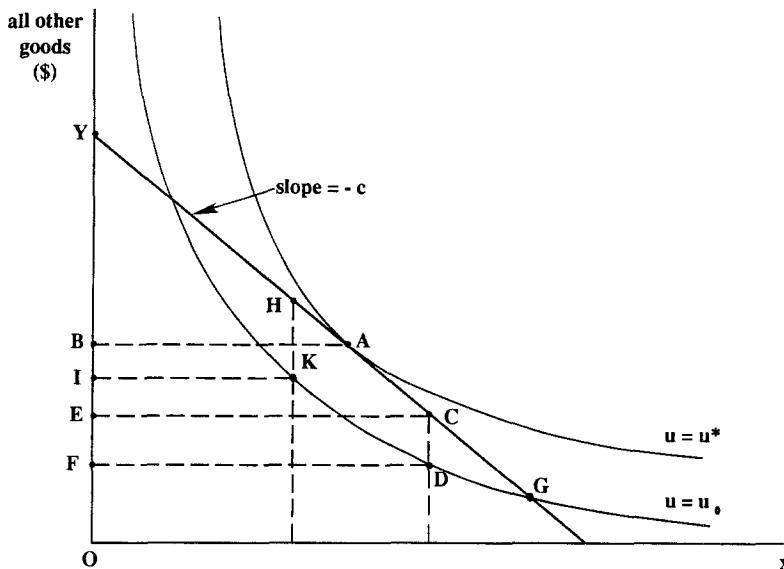


Figure 1. Equilibrium conditions.

Further, let  $c$  be the (constant) per-unit cost of the good. The sponsor is assumed to divide his income between the good produced by the bureaucracy ( $x$ ) and all other goods, with the latter measured in dollars. Point A describes the equilibrium conditions under competitive provision of  $x$  (for example, when the sponsor can choose from several private firms that produce the good). The sponsor finds a tangency between his budget line (which has a slope of  $-c$  and an all other goods intercept equal to his income  $Y$ ) and his highest possible indifference curve  $u = u^*$ . Because all other goods is measured in dollars and  $OY$  represents the sponsor's income,  $OB$  must represent the sponsor's spending on all other goods and (by the budget constraint)  $BY$  must represent his spending on  $x$ .

Point D is a representative point available to the bureaucrat. The bureaucrat uses his information advantage and/or agenda control to confront the sponsor with an all-or-nothing choice, forcing the equilibrium point to move in two dimensions. First, the bureaucrat can force the sponsor to buy more or less of the good than he might like (moving him along his budget constraint); and second, he can charge the sponsor more than the true cost of the good (thus moving the budget point off the budget constraint). (The bureau's ability to generate organization slack forces us to be more precise in our use of the term "budget constraint"; in this context, we mean the set of all bundles of  $x$  and all other goods which would exhaust the sponsor's budget if the sponsor were charged only for the true cost of producing  $x$ .) The bureaucrat's budget and slack expansion is limited, however, by the fact that the sponsor must achieve

a utility level equal to  $u_0$ , the level he obtains without bureaucratic provision of the good. In some cases, the alternative to bureaucratic provision is doing without the good entirely; in other cases, this will mean using private good substitutes like private security guards and volunteer fire departments. In Figure 1, I have assumed that private substitutes are available, so that the no bureaucracy option involves positive  $x$  output and all other goods consumption less than  $Y$ ; if this were not the case, the  $u_0$  indifference curve would cut through the all other goods axis at the  $Y$  intercept.

At point D, the following geometry establishes that the cost of producing the good is given by distance YE:

$$\begin{aligned} (\text{distance YE})/(\text{distance EC}) &= \text{slope of budget line} \\ &= c \end{aligned} \tag{1}$$

Hence

$$\begin{aligned} \text{distance YE} &= c(\text{distance EC}) = cx \\ &= \text{total cost of } x \end{aligned} \tag{2}$$

The bureaucrat's budget is represented by distance YF, leaving organizational slack of EF. To achieve utility level  $u_0$  at the given level of  $x$ , the sponsor must retain OF to spend on all other goods.

Point G is the budget-maximizing point. To maximize his budget (the vertical distance between  $Y$  and a new point F), the bureaucrat moves down and to the right along indifference curve  $u_0$ . However, the bureaucrat is limited by the constraint that his revenues at least cover his costs. Bureaucratic budget expansion therefore stops at point G because beyond that point the costs of production would exceed the proposed budget. Note that point G involves a larger budget and greater output than point A, the sponsor's preferred output, but no organizational slack.

Point K is the slack-maximizing bureaucratic point. Under slack-maximization, the bureaucrat seeks a point on the indifference curve that maximizes the difference between revenues and the cost of producing  $x$ . Geometrically, the bureaucrat seeks that level of output which maximizes the vertical distance between the budget line and indifference curve  $u = u_0$ . This must occur where the slope of the indifference curve is equal to the slope of the budget line; otherwise, a small movement to the left or right would increase organizational slack. Note that point K involves a larger budget (YI) than the sponsor would choose on his own, but smaller output and total costs than point A.

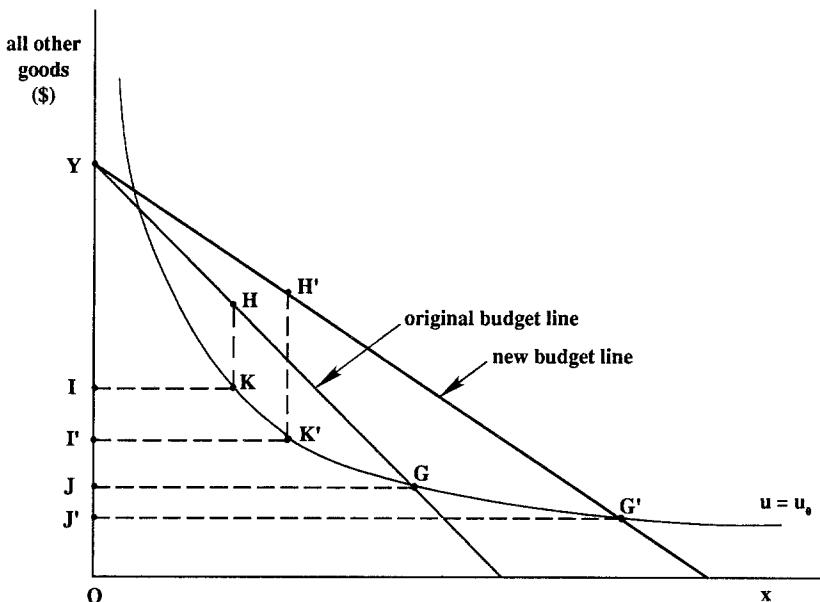


Figure 2. Elasticity of demand.

### 3. Elasticity of demand

Figure 2 shows that, for both the slack-maximizing and budget-maximizing cases, the estimated demand curve will always be cost elastic. More precisely, if we measure the elasticity of demand by comparing expenditure levels of the bureau at different levels of cost, we will always conclude that the demand curve is elastic.

In Figure 2, the cost of  $x$  falls, forming a new budget line. The budget-maximizing point shifts from  $G$  to  $G'$ , and the budget of the bureau increases from  $YJ$  to  $YJ'$ . Since a fall in cost leads to an increase in expenditure, the underlying demand curve must be elastic. Similarly, the slack-maximizing point goes from  $K$  to  $K'$ , reflecting increased output. (This must be the case, because the point on indifference curve  $u = u_0$  which has the same slope as the new, flatter budget line must be down and to the right from the corresponding point under the old budget line.) The budget of the slack-maximizing bureau rises from  $YI$  to  $YI'$ . Again, since a fall in cost leads to an increase in expenditures, the underlying demand curve must be elastic.

For intuition behind these results, we return to the equilibrium conditions of Section 2. The budget-maximizing bureaucrat obeys no standard marginal conditions; his constraint is simply that revenue must cover costs. More precisely, in the initial equilibrium, the amount the sponsor can be forced to pay

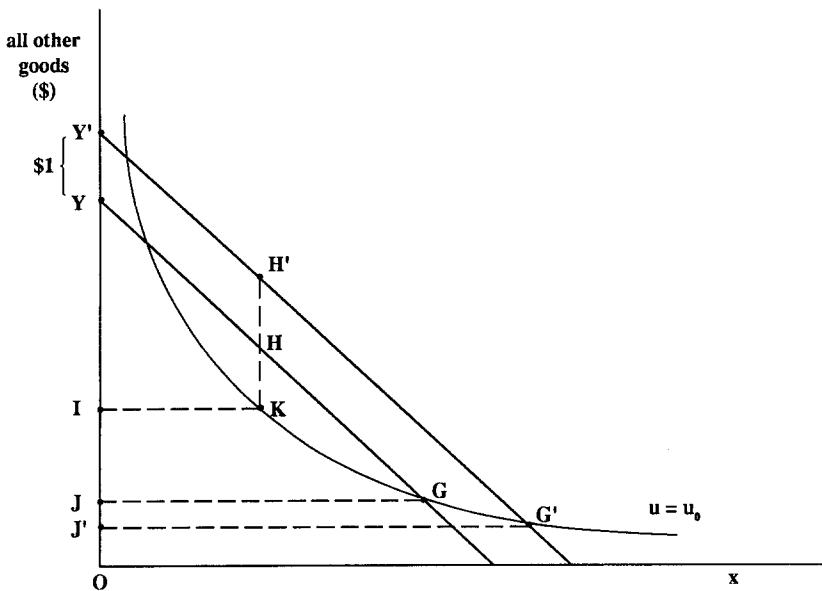


Figure 3. Lump-sum grant effects.

for additional units of  $x$ , even under the threat of losing bureau output entirely, is less than the cost of those additional units. As the cost of the units declines, however, the bureaucrat now finds that the revenue from the sponsor for those additional units does cover their costs, and both output and the budget expand.

For the slack-maximizer, the intuition is to remember that the only income left to the sponsor is the amount necessary to buy enough of all other goods to get him to utility level  $u_0$ . When the cost of  $x$  falls, the bureaucracy buys more of  $x$  and the sponsor therefore needs less of all other goods. His expenditure on all other goods therefore falls, and given the same income, the amount spent by the bureaucracy increases.

#### 4. Lump-sum grant effects

Figure 3 illustrates one empirical difference between slack-maximizing and budget-maximizing bureaus: their responses to lump-sum grants. Suppose, following Niskanen, that the bureau is of “mixed” type, receiving revenue from both the government (in the form of a grant) and from service recipients. Alternatively, following Beck (1981) or Wyckoff (1988a), suppose that the bureau’s sponsor is a state or local government which receives intergovernmental grants from higher levels of government. In either case, the exogenous lump-sum

grant sets off a comparative static change which reveals the nature of the bureau.

In Figure 3, the lump-sum grant has been normalized to a value of \$1 for clarity. The budget line shifts out parallel to the old line, with the vertical distance between the lines equal to \$1. The budget-maximizing bureaucrat shifts its equilibrium from point G to point G', increasing expenditures from YJ to Y'J'. Since YY' is equal to one dollar, the total increase in spending must exceed one dollar by the amount JJ'. The central result, then, is that a one-dollar increase in lump-sum aid increases expenditures by more than one dollar.

Conceptually, it is easiest to think of the sponsor initially turning over the entire one dollar grant to the bureaucrat. The budget-maximizing bureaucrat, since he is productively efficient, uses that one dollar to increase output by  $1/c$  units. But that additional output increases the sponsor's utility above the initial level  $u_0$ . This additional utility creates the opportunity for the bureaucrat to expand the budget even further, forcing the sponsor to give up more of all other goods in exchange for more x. While x continues to have positive utility for the sponsor, he would prefer not to pay for additions to the already excessive levels of x, so this action reduces his utility back to  $u_0$ , restoring equilibrium.

By contrast, one dollar of lump-sum aid increases the expenditure of a slack-maximizing bureau by *exactly* one dollar. In Figure 3, the lump-sum aid does not change points K or I at all. Originally, the indifference curve was parallel to the budget constraint at point K, and the parallel shift in the budget line maintains this relationship after the lump-sum aid increase. The expenditure on x therefore increases from YI to Y'I, an increase of one dollar.

Once again it is useful to think of the bureaucrat receiving a budget increase of one dollar as a preliminary allocation of the lump-sum grant. In this case, the slack-maximizing bureaucrat simply transforms that increase into increased staff, higher salaries, better working conditions, and so on, with no increase in output. In contrast with the budget-maximizing case, however, there is no tendency for the system to move from this preliminary point. The sponsor still continues to enjoy just as much x and all other goods as before, so he continues with utility level  $u_0$ . Ordinarily, the bureaucrat continually seeks out ways alter his budget in the hope of getting the sponsor to utility level  $u_0$  at lower total expense, thus leaving more resources for bureaucratic slack. However, under the lump-sum grant, the relative costs of x and all other goods haven't changed, so there is no incentive for further changes in the budget.

## 5. Lump-sum vs. matching grants

Figure 4 illustrates another difference between budget-maximizing and slack-maximizing bureaucrats: on a dollar-for-dollar basis, budget-maximizers show

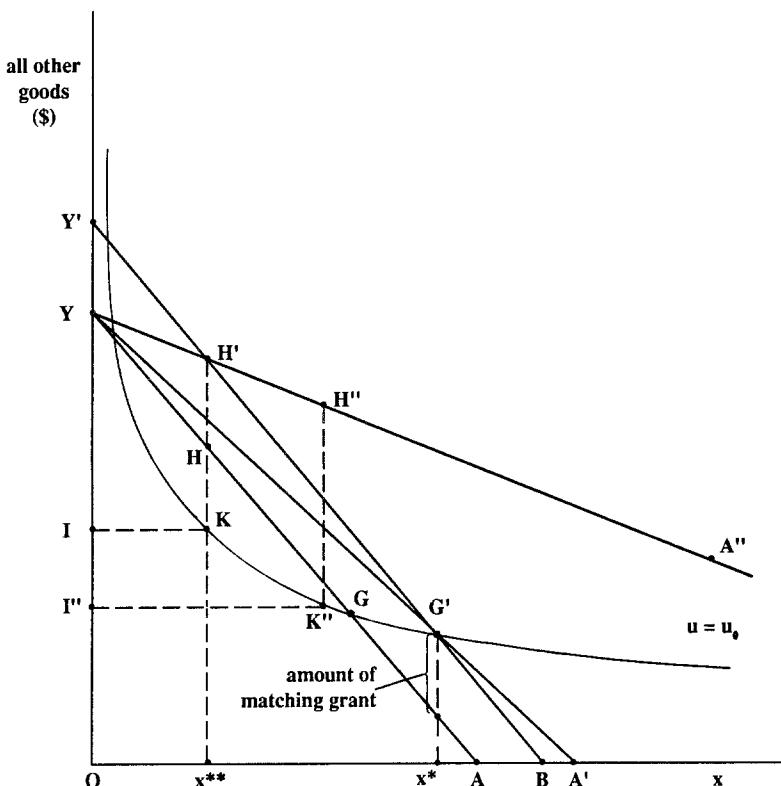


Figure 4. Lump-sum vs. matching grants.

exactly the same reaction to matching grants as lump-sum grants, while slack-maximizers do not. These results are applicable to both mixed bureaus, who might receive their grant from the government in either a lump-sum or per-unit form, and state or local government bureaus, who face both matching and lump-sum grants from higher levels of government.

Picking up from Figure 3, Figure 4 shows the effects of a lump-sum grant once again. The budget-maximizing bureau moves from point G on the original budget line YA to point G' on the new budget line Y'B. Now, suppose that we remove this lump-sum grant and replace it with an equivalent matching grant. Since the amount of the matching grant is endogenous, the term "equivalent" must be carefully defined. In the literature on intergovernmental grants, a matching grant is usually said to be equivalent to a lump-sum grant if, at the level of  $x$  formerly chosen by the jurisdiction, the community would get the same amount of aid under the matching grant as under the lump-sum grant. In Figure 4, this demands that at the level of  $x$  represented by  $x^*$ , the vertical distance between the original and matching-grant-aided budget lines

(which represents the amount of the aid) must be the same as under the lump-sum grant. Such a matching grant scheme is represented by the budget line  $YA'$ .

Contrary to the case without bureaucratic provision, equilibrium under budget-maximizing bureaus is exactly the same under the two kinds of grants. Both grant systems lead to an equilibrium at point  $G'$ . This is because the budget-maximizing bureaucrat is constrained only by the fact that revenues must equal costs. Therefore, all grants that increase revenues by the same amount have exactly the same effect. The budget-maximizing bureaucrat obeys no conventional marginal conditions, so changes in the marginal cost of the good have no effect except in their impact on revenue.

By contrast, the slack-maximizing bureaucrat is very sensitive to marginal changes in cost. Recall that, under a lump-sum grant, the bureaucrat maintains equilibrium at point  $K$  and simply increases slack by the amount of the grant. Once again, let us remove the lump-sum grant and construct an equivalent matching grant. This is a matching grant which, at the level of  $x$  represented by  $x^{**}$ , gives the sponsor the same aid (and hence the same vertical distance between the original and aided budget lines) as under the lump-sum grant. Such a budget line is  $YA''$ . Under the matching grant, however, the slack-maximizer would not continue to choose point  $K$ ; instead, the new budget line and the indifference curve would be parallel at a point like  $K''$ . For the slack-maximizer, then, lump-sum and matching grants do *not* have the same dollar-for-dollar impacts. Matching grants have a price as well as an income effect; the lower cost of  $x$  changes the cost-minimizing mix of  $x$  and all other goods required to bring the sponsor to utility level  $u_0$ , so the change in the budget is different than under the lump-sum grant.

## **6. Lump-sum grants vs. income increases**

Finally, I turn to a result which is peculiar to the intergovernmental aid case: the comparison between the effects of lump-sum grants and increases in sponsor income. Empirical studies commonly show that lump-sum aid increases cause more growth in expenditure than increases in the sponsor's income that result in the same the budget line. This result has come to be known as "the flypaper effect", because public sector money tends to stick in the public sector and be spent there, while private income tends to stick in the private sector. Wyckoff (1988a) contains a theoretical explanation of this effect for budget-maximizing bureaus and an empirical test of this explanation; here I extend the theory to slack-maximizing bureaus.

The root of the explanation is to notice that, while lump-sum aid and income increases have the same effect on the budget line, they do not have the same

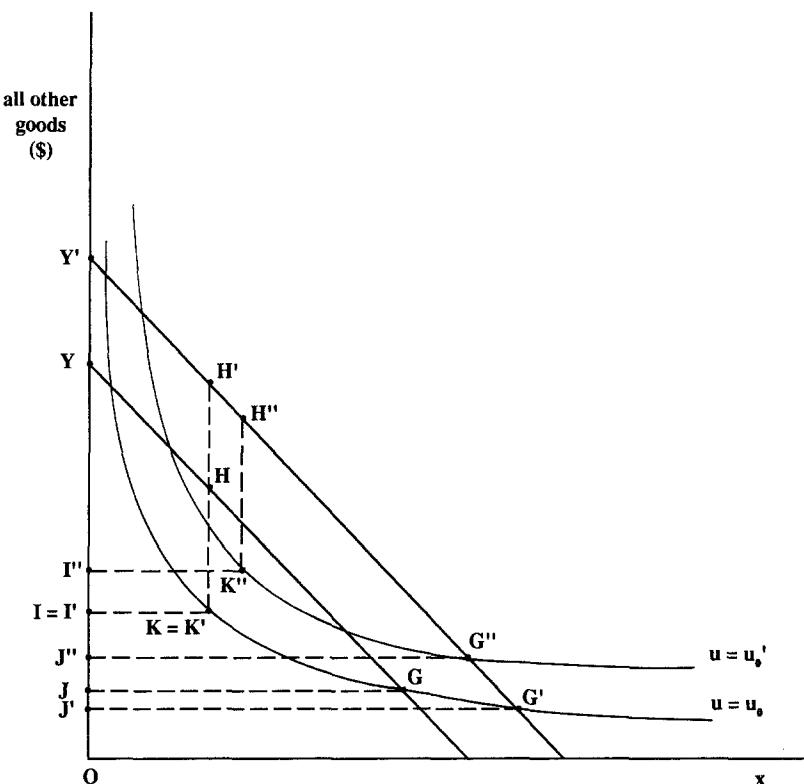


Figure 5. Lump-sum grants vs. income increases.

effect on the sponsor's utility level without bureaucratic provision,  $u_0$ . Income increases improve the sponsor's utility without bureaucratic provision, pushing up  $u_0$ , while lump-sum aid does not. This causes an asymmetry between lump-sum aid and income increases.

In Figure 5, a lump-sum aid increase moves the budget-maximizing equilibrium from G to G'. An income increase has the same effect on the budget line, but also increases the sponsor's reversion level of utility from  $u_0$  to  $u_0'$ . For the citizens of a state or local government, the next best alternative to bureaucratic provision in one jurisdiction is to move to another community. With an income increase, the extra income will be useful in that alternative location, so  $u_0$  must increase. On the other hand, if a citizen moves out of the jurisdiction, increases in lump-sum aid in his former community have no effect on him, so  $u_0$  is unaffected by lump-sum aid. If the sponsor's utility function is responsive to those of the citizens in the jurisdiction, then, there will be a difference in  $u_0$  under the two cases. Therefore, the equilibrium under income increases will be like G'' rather than G'. Instead of increasing from YJ to Y'J', there will be a smaller increase from YJ to Y'J''.

*Table 1.* Summary of results

Type of provision	Cost elasticity of demand	Effect of \$1 lump-sum grant	Lump-sum vs. matching grants	Lump-sum grants vs. increase in sponsor's income
Competitive supply	No restriction	Less than \$1 increase in expenditure*	No equivalence between two types of grants	Identical effects on spending
Budget-maximizing bureaucracy	Cost elastic	Greater than \$1 increase in expenditure	If grant amounts are the same, effects are identical	Lump-sum grants stimulate more spending
Slack-maximizing bureaucracy	Cost elastic	\$1 increase in expenditure	No equivalence between two types of grants	Lump-sum grants stimulate more spending

\* assuming that the sponsor's income elasticity of demand is less than infinite.

Similar results occur in the slack-maximizing case. A lump-sum grant leaves the slack-maximizer at point K. An income increase, however, because it affects the sponsor's reversion level, moves equilibrium to a point like K''. Rather than increasing from YI to Y'I, expenditures increase only from YI to Y'I''.

## 7. Conclusions

Table 1 summarizes the results of this analysis. Slack-maximizing and budget-maximizing bureaucracies are similar in their price elasticities and their generation of flypaper effects, but they differ in their responses to lump-sum grants and their relative responses to lump-sum and matching grants.

The important question for the empirical investigator, of course, is whether these differences matter for his research. The answer depends upon the comparative static changes he can observe, and upon his purposes. If the analyst only has data for a federal bureau which receives no lump-sum or matching grants, carefully specifying his model of bureaucracy is pointless, since the only comparative static tool he has — changes in cost — displays similar effects in both kinds of bureaus. If, on the other hand, the analyst is explicitly interested in determining whether bureaus are budget- or slack-maximizing, he should try to obtain data on bureaus receiving lump-sum and/or matching grants, be-

cause at present that is the only empirical way to distinguish between these two types of organizations.

The question of the observable differences between these two kinds of bureaucracies is further complicated by the ability of the sponsor to use "monitoring devices" to offset the information advantage of the bureau. As pointed out by Breton and Wintrobe (1975), in the real world the sponsor can employ techniques like strict accounting and budgeting controls, inspections of the bureau's facilities, public hearings on the bureau's performance, and the use of "watchdog" organizations like GAO, CBO, and (at the local level) professional city managers. These monitoring devices, by offsetting to some extent the information advantage of the bureau, result in a hybrid situation, between the competitive and bureaucratic extremes.

The presence of such monitoring devices might seem to obviate the need for the theoretical models developed above, models which are based on complete domination of the budget process by the bureaucrat. Even in this situation, however, it is extremely important to detail the characteristics of the bureaucratic extremes in order to identify the separate strands which make up the intermediate cases. The traits of the bureaucratic extremes are likely to remain, in muted form, when monitoring devices are incorporated into the model. For example, in a recent paper I noted that estimates of demand for local public goods that are generated from surveys of voters' preferences indicate a very low price elasticity and minimal impact of lump-sum aid on spending. If monitoring creates a situation in between the bureaucratic and competitive extremes, then, we ought to observe elasticities of demand from studies of actual spending data which are *not* greater than one in absolute value, but which *are* larger than the survey estimates. Similarly, the effect of lump-sum aid on expenditure might not be as large as in the bureaucratic models, but it ought to be larger than estimated by surveys of voter attitudes. The empirical work in Wyckoff (1988b) confirms that these are exactly what we observe in studies of local education expenditures.

## References

- Beck, J.H. (1981). Budget-maximizing bureaucracy and the effects of state aid on school expenditures. *Public Finance Quarterly* 9.2 (April): 159–182.
- Breton, A. and Wintrobe, R. (1975). The equilibrium size of a budget maximizing bureau: A note on Niskanen's theory of bureaucracy. *Journal of Political Economy* 83.1 (February): 195–207.
- Cyert, R.M. and March, J.G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice-Hall.
- Gonzalez, R.A. and Mehay, S. (1985). Bureaucracy and the divisibility of local public output. *Public Choice* 45(1): 89–101.
- McGuire, M. (1973). Notes on grants-in-aid and economic interactions among governments. *Canadian Journal of Economics* 6.2 (May): 207–221.

- Miguel, J.-L. and Belanger, G. (1974). Toward a general theory of managerial discretion. *Public Choice* 17 (Spring): 27–43.
- Niskanen, W.A. (1971). *Bureaucracy and representative government*. Chicago: Aldine-Atherton.
- Orzechowski, W. (1977). Economic models of bureaucracy: Survey, extensions, and evidence. In T.E. Borcherding (Ed.), *Budgets and bureaucrats: The sources of government growth*, 229–259. Durham, NC: Duke University Press.
- Patinkin, D. (1963). Demand curves and consumer's surplus. In C.F. Christ et al., *Measurement in economics: Studies in mathematical economics and econometrics in honor of Yehuda Grunfeld*, 83–112. Stanford, CA: Stanford University Press.
- Williamson, O. (1964). *The economics of discretionary behavior: Managerial objectives in a theory of the firm*. Englewood Cliffs, NJ: Prentice-Hall.
- Wyckoff, P.G. (1988a). A bureaucratic theory of flypaper effects. *Journal of Urban Economics* 23.1 (January): 115–129.
- Wyckoff, P.G. (1988b). *Testing bureaucratic influence on local school expenditures by comparing survey and expenditure data*. Unpublished paper, 1988.

## ***Bureaucrats, Budgets and the Growth of the State: Reconstructing an Instrumental Model***

PATRICK DUNLEAVY\*

This article forms part of a longer-term project dealing with the impact of public choice theories in political science.<sup>1</sup> The focus here is on economic models of bureaucracy, which despite their increasing theoretical significance and influence on practical politics have heretofore been little analysed, except by their exponents. I have argued elsewhere<sup>2</sup> that amongst existing public choice accounts there are two seminal works, Antony Downs's pluralist treatment in *Inside Bureaucracy* and William Niskanen's new right thesis in *Bureaucracy and Representative Government*.<sup>3</sup> The central innovation of economic approaches is their stress on rational officials' attachment to budget maximization strategies. In Downs's case this is a finite maximand limited by bureaucrats' conservatism and other motivations. But in Niskanen's case budget maximization is an open-ended process, constrained only by external limits on agencies' abilities to push up their budgets. None the less, despite their disparate approaches and conclusions, both these books share four failings common to almost all other public choice work in the field:

- (1) They operate with vague and ill-defined definitions of bureaucrats' utility functions.
- (2) They assume that all bureaucracies are hierarchical line agencies.

\* Department of Government, London School of Economics and Political Science. Hugh Ward, Spencer Zifcak, Brendan O'Leary, George Jones and other LSE colleagues, and members of the LSE Graduate Seminar on Public Policy in 1982-4, helped me to clarify these ideas. As a paper, this article was given in a more extended form to Politics staff seminars at the Universities of Bristol, Warwick, Glasgow and Edinburgh, and to the Public Sector Budgeting Conference at the University of Manchester, 18 May 1984. I would like to thank numerous individual commentators from all these sessions whose suggestions I have gratefully embodied here.

<sup>1</sup> See Patrick Dunleavy and Hugh Ward, 'Exogenous Voter Preferences and Parties with State Power: Some Internal Problems of Economic Theories of Party Competition', *British Journal of Political Science*, xi (1981), 350-80.

<sup>2</sup> P. Dunleavy, 'Bureaucrats, Budgets and the Growth of the State: Part I, Existing Public Choice Approaches', unpublished paper, available from the author at LSE.

<sup>3</sup> A. Downs, *Inside Bureaucracy* (New York: Little, Brown, 1967); W. Niskanen, *Bureaucracy and Representative Government* (New York: Aldine-Atherton, 1973). Other less important work in the genre includes: L. von Mises, *Bureaucracy* (New Haven: Yale University Press, 1944); G. Tullock, *The Politics of Bureaucracy* (Washington, D.C.: Public Affairs Press, 1965); J. Buchanan et al., *The Economics of Politics* (London: Institute for Economic Affairs, 1978); A. Breton, *The Economic Theory of Representative Government* (Chicago: Aldine, 1974), Chap. 9; J. Migue and G. Berlanger, 'Towards a General Theory of Managerial Discretion', *Public Choice*, xvii (1974), 27-43; T. E. Borcherding, ed., *Budgets and Bureaucrats* (Durham, North Carolina: Duke University Press, 1977).

- (3) They have no valid way of accounting for variations in bureaucratic motivations within a public-choice methodology.
- (4) They treat systems of bureaus as if they behaved in the same way as single bureaus.

Starting from the core assumptions of public choice models, this article outlines an alternative analysis of bureaucratic behaviour, one in which budget maximization is a more remote or unlikely influence upon agency policies. The framework for this reconstruction includes the following assumptions (which I believe capture the essentials of an economic approach to the subject). Agency policies are set by bureaucrats interacting with their sponsor body. Apart from general information from citizens about bureau behaviour, the sponsor largely depends upon government agencies for information about the costs and value of producing in given ranges of output. Bureaus negotiate an annual budget with their sponsor for a whole block of output. Bureaucrats are essentially instrumental, maximizing their personal utilities when making official decisions. By this I mean that they are concerned exclusively with the satisfaction of self-regarding, relatively hard-edged preferences. A bureau's aggregate policy behaviour is set by some combination of individual decisions made by its officials, although the result of this process may be an outcome desired by no bureau member. Within broad limits, officials' influence on bureau policy is always extensively rank structured, with those near the top of bureaus being most influential.

This article develops six main propositions in opposition to existing public-choice models:

- (1) Collective action problems exist within bureaucracies and have an important influence upon overall bureau behaviour.
- (2) Bureaucrats' utilities are normally associated with only a part of the overall budget under an agency's control.
- (3) There are sharp differences between different kinds of agencies in the extent to which officials' welfare is positively associated with budget increments.
- (4) Bureaucrats in policy positions will only maximize budgets up to an internal optimal level, defined by the intersection of curves showing their discounted marginal utility pay-offs from budget increments and the marginal costs they incur in advocating further expenditure.
- (5) Utility-maximizing bureaucrats are empirically more likely to be orientated towards the intrinsic character of their work tasks than to pecuniary or near-pecuniary considerations. Their strategies for improving work-related utilities hence focus not on budget maximization but on reshaping their bureaus to bring them into a closer conformity with an ideal form.
- (6) If bureaucrats maximize budgets then state growth should have produced a progressive expansion of large line bureaucracies – in fact, a rare pattern in liberal democracies. But if bureaucrats pursue bureau-shaping strategies, this would create a much more fragmented state, dominated

by elite central departments without line responsibilities. A great deal of state growth fits this pattern.

### I. COLLECTIVE ACTION PROBLEMS

One of the more critical hidden assumptions of existing public choice theories of bureaucracy is that government agencies are extremely hierarchical bodies. In Niskanen's case the assumption is that agencies with a separate identifiable budget are run *completely* by their top official, who is in his terms the only 'bureaucrat'. Everyone else is reduced to the status of 'employee'.<sup>4</sup> Little wonder then that Niskanen's account predicts over-supply behaviour, for in many ways budget maximization produces benefits which are private goods for the hegemonic top official. As soon as we relax this assumption to look at situations closer to actual bureaucracies, it is clear that officials may confront difficulties in organizing to achieve common objectives. In British central government, for example, departments may have up to eighty staff at Assistant Secretary level or above, which would be generally acknowledged as an influential policy-making rank.<sup>5</sup> We clearly need to move away from extreme hierarchic assumptions about how agencies are run. We should also recognize a point given little emphasis in existing accounts, namely that officials pursuing their own interests confront a wide range of options, among which a (collective) strategy of budget maximization is only one.

To take these points in reverse order, budget-maximization in an agency were no one official has complete hegemony increasingly takes on the character of a collective rather than an individual good. It lies at the 'public good' end of a spectrum of utility maximizing strategies which is shown in Figure 1.

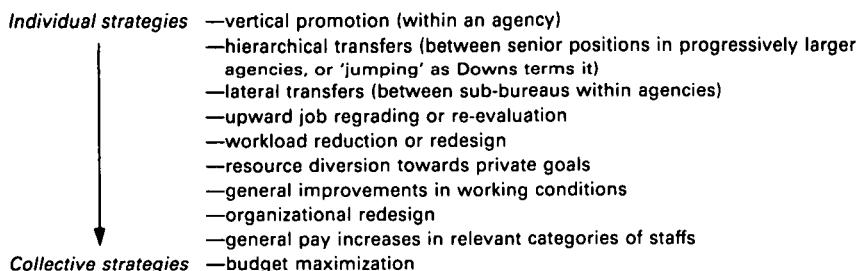


Fig. 1. The spectrum of utility-maximizing strategies

<sup>4</sup> Niskanen, *Bureaucracy and Representative Government*, p. 22. In his more popular exposition, *Bureaucracy: Servant or Master?* (London: Institute for Economic Affairs, 1973), Niskanen admits (p. 11) that he sometimes uses the term in this specialized sense, and sometimes just to mean 'civil servant'. Readers must presumably guess which is which, for the distinction is buried thereafter.

<sup>5</sup> M. Lee, 'Whitehall and Retrenchment', in C. Hood and M. Wright, eds, *Big Government in Hard Times* (London: Martin Robertson, 1981), p. 41, shows that in 1980 there were over 800 people in the top three ranks, spread across forty-seven departments. For more detailed figures, see A. Dunsire and C. Hood, *Bureaumetrics* (Farnborough, Hants: Gower, 1980).

Individual utility-maximizing strategies are obviously the most direct ways in which a given bureaucrat can increase her own welfare, because the effort expended in pursuit of them will (if successful) generate a pay-off which does not need to be shared with others. By contrast, the hallmark of collective strategies is that the connection between the achievement of a given public-goods outcome and the individual bureaucrat's welfare becomes more remote. In particular, a given budget increase for an agency will have rather indeterminate implications for most of its staff and officials, even within policy-making ranks. Consequently we should expect bureaucrats to put their efforts primarily into individual utility-maximizing strategies, implying that their energies will only be displaced into the pursuit of collective goods as other options are foreclosed or become fully exploited. It is commonly argued in occupational sociology, for example, that manual workers put more emphasis on collective forms of pay bargaining than non-manual staffs do because for them individual strategies of promotion, career advancement, or renegotiation of their personal terms of work with employers offer relatively little prospect of significantly improving their situation.<sup>6</sup> Similarly we might expect that within a government agency the opportunities for individual welfare maximization will be fewest in the bottom ranks of bureaucrats. Hence collective strategies for improving officials' welfare may be resorted to more readily. But in the policy-making ranks at the top of bureaus, the scope for exploiting individual strategies will be much greater and the propensity to resort to collective strategies correspondingly reduced.

It could be objected; however, that in real life budget increments are by no means a diffuse or shared benefit. Instead of dealing with the abstract marginal increment to the bureau's global budget discussed by public choice theory, we should perhaps focus attention upon much more specific or 'tagged' funding increases. But once we move away from a whole-bureau perspective focusing on the overall budget, not only does any model necessarily become much more complex, but it also becomes highly improbable that officials in general will see their interests as advanced by all (tagged) budget increases. Instead they will favour only those which will positively affect their own position. They are quite likely to be indifferent towards, or to oppose, budget increments going to other sections of their own organization. Officials in stagnant or slow-growing sections of an agency could easily face severe welfare losses because of budget increases in more dynamic sections. Not only could they experience 'relative deprivation', but in more concrete terms they are likely to see the balance of influence and prestige within the agency shift towards the growth areas at their expense. Consequently an account of how officials behave with tagged rather than global budget increases needs to be constructed in terms of the emergence of minimum winning coalitions for certain types of budgetary expansion rather than

<sup>6</sup> See, for example, Colin Crouch, *Trade Unions: The Logic of Collective Action* (London: Fontana, 1982), pp. 67-74. Crouch's analysis of trade unions resembles at some points the approach used here.

others. It also requires a developed apparatus for describing the degree of sectionalization in bureaus.<sup>7</sup> For these reasons we shall stick to an analysis of generalized budget increases or decreases while recognizing explicitly that this is a simplifying assumption which should ideally be superseded at a later stage by a more complex account.

The second point which needs to be recognized is that although bureaus are rank-structured environments there is very little likelihood that they are ever completely dominated by one individual or even a small leadership group with cohesive interests. As Downs concisely put it: 'The concept of bureaus as monolithic structures is largely a myth'.<sup>8</sup> Hence the realization of collective benefits for bureau members is likely to require concerted action by a number of officials which may be quite large (especially in hostile or turbulent environmental conditions). To see how the rank-structured nature of bureaucracy changes the collective action problems involved, consider the basic equation which an individual official must confront in deciding whether or not to press for a (generalized) budget increase:

- the net utility derived from a marginal budget increment (i.e. the benefits received after allowing for any costs associated with budgetary growth)
- discounted by the probability that the individual official's advocacy will be decisive in securing the budget increase
- must be greater than
- the costs of advocating the budget increment.<sup>9</sup>

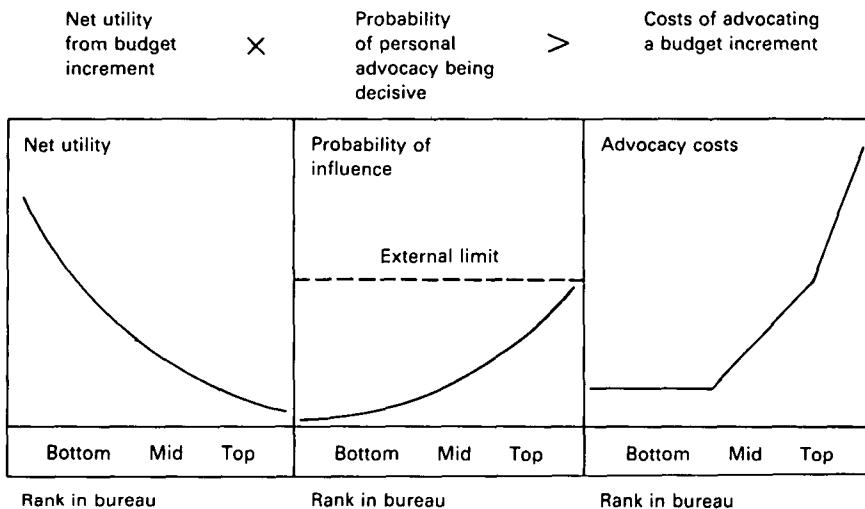
Each of these terms will be influenced by the rank of the official in question, roughly along the lines shown by the charts in Figure 2 where rank is shown along the horizontal dimension in terms of 'top', 'middle' and 'bottom' positions.

The utility pay offs from generalized budgetary increments are likely to be distributed more or less inversely to rank positions. I assume that those at the bottom gain most, while those at the top gain least. In fact, really large benefits may be concentrated on peripheral staffs – those with no job security, people acting as contractual consultants, spin-off staffs or those on part-time contracts. Certainly these are the groups with most to lose from any budget

<sup>7</sup> For interesting comments on the interchangeability of coalition theory and collective goods models, see R. Abrams, *Foundations of Political Analysis* (New York: Columbia University Press, 1980), pp. 329–46.

<sup>8</sup> Downs, *Inside Bureaucracy*, p. 133. Downs continues: 'If bureaus were really monolithic, control over nearly all their activities would be concentrated in the hands of their topmost officials. However, these officials must always delegate some of their powers to subordinates'.

<sup>9</sup> This equation in the same general terms was, of course, specified by A. Downs, *An Economic Theory of Democracy* (New York: Harper and Row, 1957) as the calculus facing citizens deciding whether to vote in a democracy, and applied to the analysis of why people join interest groups by Mancur Olson, Jr. in *The Logic of Collective Action* (Cambridge, Mass: Harvard University Press, 1965). In each case the format is that the net utility gained from action, discounted by the probability of the action influencing the eventual outcome, must exceed the (non-discounted) costs incurred in taking the action.



*Fig. 2. The distribution of net utilities, influence probabilities and advocacy costs within a hypothetical bureau*

reduction. In most bureaucratic organizations, peripheral staffs are a significant (if largely unstudied) component. For example, in the United States where federal government employment is subject to strict manpower limits, estimates suggest that there are between 5 and 25 per cent extra staff employed in this way, who never show up in returns to Congress.<sup>10</sup>

The probability that an individual's advocacy of a budget increment will be decisive obviously increases dramatically with rank. In particular, part-time, hived-off, and lower rank officials may have virtually negligible individual impact upon their agency's policy. But within the policy-making ranks we should expect to see the probability of influence rising sharply with rank position, reaching a threshold of maximum influence which will be environmentally determined. Where the agency confronts a favourable environment the probability of securing its advocated policy may approximate 1. But this is likely to be an exceptional situation, contrary to Niskanen's claim that it is the

<sup>10</sup> See J. Bennet and M. Johnston, *The Political Economy of Federal Government Growth* (College Station, Texas: Center for Education and Research in Free Enterprise, Texas A & M University, 1980), pp. 38-41. The problem with widely varying estimates here is in distinguishing between contracted-in staffs employed by the agency directly, and the much more massive phenomenon of contracted-out services being delivered by employees of firms or other agencies. We are concerned here with contracting in. Of course, contracted-out work often includes substantial staffing implications as part of larger contracts for products or service delivery. But since contracted-out staffs remain directly employed by other bodies, and since they often cannot even be sensibly construed as falling within the public sector, this kind of indirect employment is better analysed using the 'contract agency' classification of bureau situations in the typology developed below.

normal position facing bureaus.<sup>11</sup> Where an agency confronts a hostile or turbulent environment then even the most influential official in setting its internal policies will have a relatively low chance of seeing agency policies approved by its sponsor. If the distribution of influence is as shown here then we can expect changes in environmental hostility or favourability towards the agency to have a major impact upon top bureaucrats' behaviour, by bringing dramatic changes in the way they discount their utility pay-offs from budget increments. But such external changes will have much less effect upon the attitudes of lower staffs, whose probable influence is small even in a favourable environment.<sup>12</sup>

Finally, we should expect the costs of advocating a budget increment to be significantly rank-structured as well. Bottom rank personnel will have few opportunities to influence bureau policy directly. Their contribution may be confined quite largely to supporting their union or other collective organization in making a case for more spending. But in the middle ranks of the organization, the opportunities for actively promoting budgetary expansion are greater and the personal costs of such advocacy begin to rise steeply with rank. For top officials a budgetary increase (over and above normal or technical incremental adjustments for inflation, etc.) typically involves preparing special papers, attending difficult meetings, cultivating external allies and contacts, responding to sponsor criticisms or investigations, and justifying the bureau's case in public. The empirical literature on budget-making strongly suggests that the vast bulk of all budget negotiations and controversy concentrates on marginal increments or decrements to a largely unanalysed 'base budget'.<sup>13</sup> So the advocacy costs associated with departures from the base budget seem likely to be substantial in more senior ranks.

If the distribution of utilities, influence probabilities and advocacy costs is along the lines set out here it is possible to detect a central paradox that is likely to confront bureaucracies. While budget maximization would probably produce overall benefits for bureau members, it is also likely that officials will tend to free ride on advocating higher funding, for different reasons. Those at the bottom of the rank structure stand to gain most from an increased budget, but they discount these gains by the very low probability that their personal advocacy of any increase will be decisive. Hence even though their advocacy costs are small, it is rather unlikely that low ranking officials with permanent or secure jobs will feel it is worth their while to press for increased

<sup>11</sup> R. Goodin, 'Rational Bureaucrats and Rational Politicians in Washington and Whitehall', *Public Administration*, LX (1982), 23–41.

<sup>12</sup> Top bureaucrats are also likely to be much better informed about changes in the external environment than lower ranking officials, so that we should expect to see their reactions adapting much more quickly and accurately to external changes. In contrast, lower officials' behaviour may well perpetuate anachronistic attitudes into a new period, remaining conservative in their behaviour in environments favourable for organizational growth, but also carrying over a previously defined growth orientation into an era of cutback management.

<sup>13</sup> See P. Jackson, *The Political Economy of Bureaucracy* (Oxford: Phillip Allen, 1982), Chap. 5 *passim* for a good review.

spending. Officials at the top of the bureau by contrast have a significant probability of influencing outcomes, but they stand to gain least from budgetary expansion and confront high advocacy costs in exercising their influence. They are particularly unlikely to incur these costs in a hostile or turbulent environment where the overall likelihood of the agency being successful is low or unpredictable.

Of course, much will depend on the cardinal values which we might be able to ascribe to the different components given in the equation above. By itself the analysis given here does not demonstrate that a collective action problem will *necessarily* confront budget maximizing bureaucrats.<sup>14</sup> But if we flesh out this account with some intuitively realistic values or examples, it seems likely that collective action problems will be the norm rather than the exception – especially when we consider that rational bureaucrats will only become involved in collective (rather than individual) attempts to improve their welfare as a last resort.

## 2. BUDGET MAXIMIZATION

Pushing up the budget as a possible goal of utility-seeking bureaucrats is a more complicated concept than economic models have so far acknowledged. I noted above that attention is focused on generalized budgetary expansion rather than tagged increases. But there are still important distinctions which need to be made within this general focus between different kinds of budget.<sup>15</sup> We may distinguish at least three possible meanings of 'budget'.

- (1) An agency's *programme budget* (PB) consists of all the expenditure over which it exercises supervision or control, even if large parts of this total are passed on to other public sector agencies for final implementation. For example, the programme budget of the Department of Health and Social Security (DHSS) in Britain consists of all public expenditure on health care, personal social services and social security – even though health care funds are passed on to NHS authorities and personal social services funds go to local authorities rather than being spent by DHSS staff.
- (2) An agency's *bureau budget* (BB) consists of those parts of its programme budget for which it is completely or solely responsible to the sponsor body. Thus the BB is the funding which goes into programmes directly controlled by the bureau's own decisions. This implies that so long as the funding stays inside the public sector, policy is being implemented by the bureau's own staff, without the involvement of separate or subordinate

<sup>14</sup> I would like to thank Jeremy Waldron of Edinburgh University for helping to clarify the issues involved here.

<sup>15</sup> Both Downs and Niskanen shift between these meanings without apparently perceiving their different implications. Niskanen in particular explains agency behaviour quite largely in terms where 'budget' equals 'bureau budget'. But in his vaguer accounts of how budget maximization creates state growth, 'budget' normally equals 'programme budget'.

agencies. For example, the bureau budget of the DHSS covers its major direct function, the social security system, plus the administrative costs involved in its supervising the health care and personal social service systems.

- (3) An agency's *core budget* (CB) consists of those parts of its bureau budget which are spent on maintaining its own operations, rather than going outside the agency as contracts to private firms, or as services or transfer payments to citizens, clients or external groups who are the beneficiaries of the agency's activities. For example, the DHSS core budget includes only those revenue costs which are spent on its own staff, their accommodation and their day-to-day activities. It includes the money spent on administering the social security system. But it excludes the money which is actually transferred by this system to pensioners, people on supplementary benefits, or those receiving other state aid.

A diagrammatic classification of the most frequently occurring elements in public expenditure is shown in Figure 3. The classification suggested is reasonably straightforward. Capital spending is included in the bureau budget but not the core budget since contracts for this are assumed to go to private construction firms or other contractors. But contracts for revenue items associated with the basic administration of the bureau (for example, rents paid to private landlords or purchasing of office equipment) are included in the core budget.

The point of making these distinctions between types of budget is to bring out the variable extent to which the personal utilities of bureau members are likely to be involved in the expansion of different elements of the overall budget. Table 1 shows a rough and ready list of the reasons why budget maximization has been seen by public choice writers as improving bureaucrats' welfare. Codifying some of the comments made in section 1 about the salience of different utility considerations for top, middle and low rank bureaucrats, an interesting picture emerges.<sup>16</sup> The most general and the most basic utility gains from increasing budgets are all associated with the core budget. And these are also the utility gains which have most importance for bottom and middle ranking officials. By contrast the more general and diffuse utility gains from budgetary expansion, and those with most importance for top ranking officials, are primarily linked to the bureau budget. In this group of pay-offs only facilitating non-conflictual management of the bureau can be regarded as associated primarily with the core budget. And only an agency's

<sup>16</sup> For the sake of argument I have assumed in constructing Table 1 that all these net utility aspects show gains with budget increases. In practice, it seems unlikely that this will be the case. For example, budgetary expansion may keep the peace inside agencies but simultaneously increase bureau inertia. Similarly, increasing bureau budgets may allow more agency patronage over contractors; but enhanced pressure for outputs may also increase an agency's dependence on external firms. P. Dunleavy, *The Politics of Mass Housing in Britain, 1945-75* (Oxford: Clarendon Press, 1981) documents exactly this kind of power-dependency switchback between public housing agencies and construction corporations in Britain's post-war housing policy.

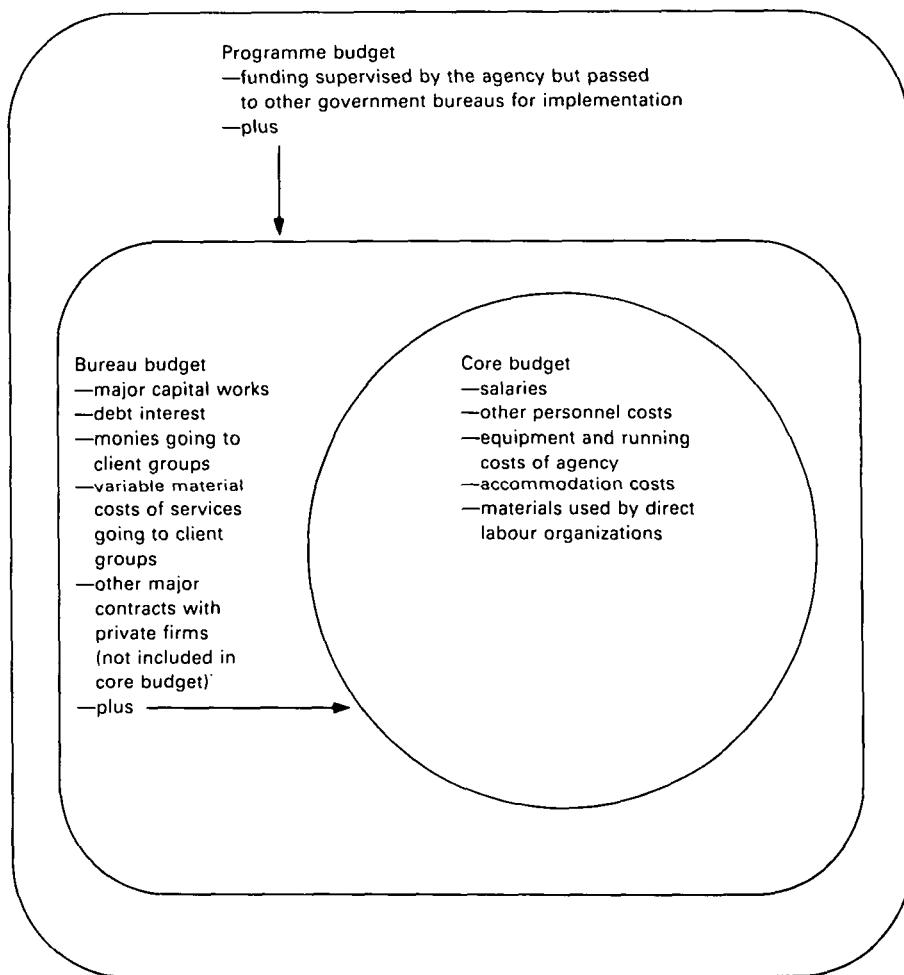


Fig. 3. Components of programme, bureau and core budgets

patronage power is positively and directly associated with the expansion of the programme budget. An enlarged programme budget will be useful also in allowing top management to build up slack resources to respond to crises or unexpected calls for action. But since the programme budget may have to be largely allocated to other agencies, top officials gain maximum flexibility from retaining slack under their own direct control, in the bureau budget.

Turning to the obverse side of the utility gains from budgetary expansion, namely, the costs associated with advocating increased funding, it is apparent that there is an important asymmetry.<sup>17</sup> While the benefits of budgetary expansion are associated mainly with the core budget or the bureau

<sup>17</sup> Advocacy costs of promoting budget maximization (on the right-hand side of bureaucrats' choice equation) need to be clearly distinguished from potential disutilities of budgetary expansion (which are incorporated into the net utility calculation on the left-hand side of the equation).

TABLE I *Welfare Gains for Bureaucrats from Budget Maximization*

Type of welfare gain	Associated budget	Salience for ranks		
		Bottom	Middle	Top
Improving job security	Core budget	++	+	0
Expanding career prospects	Core	+	++	0
Increased demand for skills and labour	Core	+	++	0
Triggering upward regrading	Core	0	++	+
Reducing conflict in bureau management	Core budget	0	+	++
Boosting bureau prestige	Bureau budget	0	+	++
Improved relations with clients or contractors	Bureau budget	0	+	++
'Slack' creation to cope in crises	Bureau/programme budgets	0	+	++
Increased patronage powers	Programme budget	0	0	++

Key: high salience ++; medium salience +; low salience 0.

budget, the costs of advocacy (in terms of time spent, effort and resources required, increased visibility to the sponsor and the public, and level of external criticism received) are all likely to be closely associated with the programme budget. If education costs in local education authorities rise sharply, the burden of defending their record in Cabinet and Parliament falls quite largely on the Department of Education and Science, especially in the annual round of negotiations justifying public expenditure increases. And within a given level of environmental hostility, a rise in an agency's programme budget for which it was not responsible has the effect of squeezing out its chances of pushing through increases in its own core budget or bureau budget. Similarly it is on programme budget performance that sponsoring bodies and citizens in general base their judgements about the overall funding levels appropriate in particular policy areas. It is little use to the DES that its own operations are efficiently run if the legislature judges what level of funding to put into education, and into the DES's own core budget or bureau budget, by reference to how well schools are doing their job. Hence there is a key asymmetry in the equation given in Figure 2, namely that net utility gains attach primarily to core budgets or bureau budgets, while advocacy costs are always attached to the programme budget of an agency.

The implications of this imbalance for individual bureaucrats' behaviour will clearly depend crucially on the extent to which programme, bureau and core budgets are differentiated from each other in particular agencies.

### 3. VARIATIONS IN BUREAUCRATIC BEHAVIOUR PATTERNS

Agencies behave in different ways not because officials have different types of personality (as Downs argued), but because the structure of utility gains and

advocacy costs associated with budgetary expansion varies systematically with different kinds of agency. We can distinguish the following sorts of agencies:

*Control agencies* (CLAs) allocate budgets to and supervise the activities of other public sector organizations while having few or no major responsibilities for implementation or service delivery of their own. The Department of Education and Science is a fairly pure example of this type.

*Regulatory agencies* (RAs) control or supervise the operations of other agencies, private sector firms, or the general public in some respect, using licensing systems, reporting controls, performance standards or some other system of regulation. RAs do not directly produce material outputs but rather police other firms' and agencies' production activities. RA 'outputs' are thus rather intangible 'goods' involving the non-commission of offences and the non-occurrence of unacceptable outcomes. The Alkali Inspectorate is a fair example of this type. RAs also often use subsidies in their efforts to control or direct private sector activities into a preferred pattern of development.

*Transfer agencies* (TAs) administer general transfer payments or subsidies to individuals or to client interest groups. The social security system and the Ministry of Agriculture farm support functions are both examples of this sort of agency.

*Contracts agencies* (CTAs) primarily allocate work on a contract basis to private sector firms or to commercially-run public sector agencies such as public corporations. Contract agencies' own activity is largely confined to conducting research and development into projects, drawing up specifications, inviting tenders and supervising contracts, with the actual production of physical outputs carried out by the contractors. Defence procurement is a good example of a CTA function.

*Delivery agencies* (DAs) directly undertake the production of goods and services and their delivery to citizens. Hence implementation is carried out directly by DA employees – no doubt often working in complex networks of sub-bureaus, but with a clear line of authority or responsibility from top bureau officials to those at the grass roots. Local authority schools provision in the United Kingdom is a good example of a DA set-up.

This typology is not exhaustive, since it makes no mention of taxing agencies or those supplying commercial services.<sup>18</sup> But the types distinguished here are very common and cover the bulk of public service organizations in most liberal democracies. The usefulness of the typology rests in large part on its close connection with the distinctions made in Section 2 between programme, bureau and core budgets. Four criteria seem especially useful in distinguishing these agency types from each other:

<sup>18</sup> Taxing agencies might be regarded as a sub-type of regulatory agencies, but they are distinctive in raising revenues as a result of their activities, making it difficult to apply simple demand/cost analysis to them. Mixed agency types with a commercial role similarly involve complex modelling, although Niskanen drops a few hints about them.

- the overall scale or size of an agency's programme budget, which obviously needs to be defined relative to the spatial scale of its operations and the scale of rival or parallel agencies;
- the importance of its bureau budget in the overall programme budget (the BB/PB ratio);
- the importance of the agency's core budget in its programme budget (the CB/PB ratio); and
- the importance of the agency's core budget in its bureau budget (the CB/BB ratio).

If we also graph the growth of the bureau budget and of the core budget in cash terms against the growth of the programme budget in cash terms for these agency types, we obtain the plots and table entries shown in Figure 4.

The impact of these differences in PB/BB/CB relationships between types of agencies can also be considered using the diagrams from Section 1 above.

*Net utility gains* from marginal programme budget increases show the greatest variation across agencies. The lowest curve is probably that for transfer agencies, since if the PB increase has no impact on the CB, it is hard to see how even low rank bureaucrats derive much benefit from it. For example, if pensions go up in cash terms the DHSS simply increases the figures printed by computers on millions of giro cheques. Even if the programme and bureau budget increase very substantially, there is no apparent reason why the DHSS core budget cannot remain unchanged. Control agencies similarly have low and flattish net utility curves across ranks since the bulk of any PB increase will go to other agencies. Contract agencies' curves are higher since a fairly fixed but low proportion of a PB increase is spent on core administration costs.<sup>19</sup> Delivery and regulatory agencies have the highest net utility curves, turning up most sharply at the bottom rank position – since an increased PB translates directly into an increased CB, with little leakage outside the agency, and with maximum benefits for rank-and-file employees in terms of improved job security, better career prospects and increased demand for their services.

*Probabilities of influence* show less change, certainly for the more hierarchical agencies (DAs, RAs, CTAs and TAs). But control agencies may have much lower curves because they are more dependent on other public agencies in securing a given PB increase. And they may have flatter curves because middle ranking officials can have more influence over control agency policy than in larger line bureaucracies.

*Advocacy costs* of a PB increase steeply amongst top officials in all agencies, but again control agencies may be distinguished by higher advocacy costs amongst middle ranking officials because of their greater opportunities for influence.

<sup>19</sup> In British local government, for example, administration cost ratios of 15–25 per cent of the programme budget are regularly applied where council staffs do agency work for other bodies (such as police or water authorities).

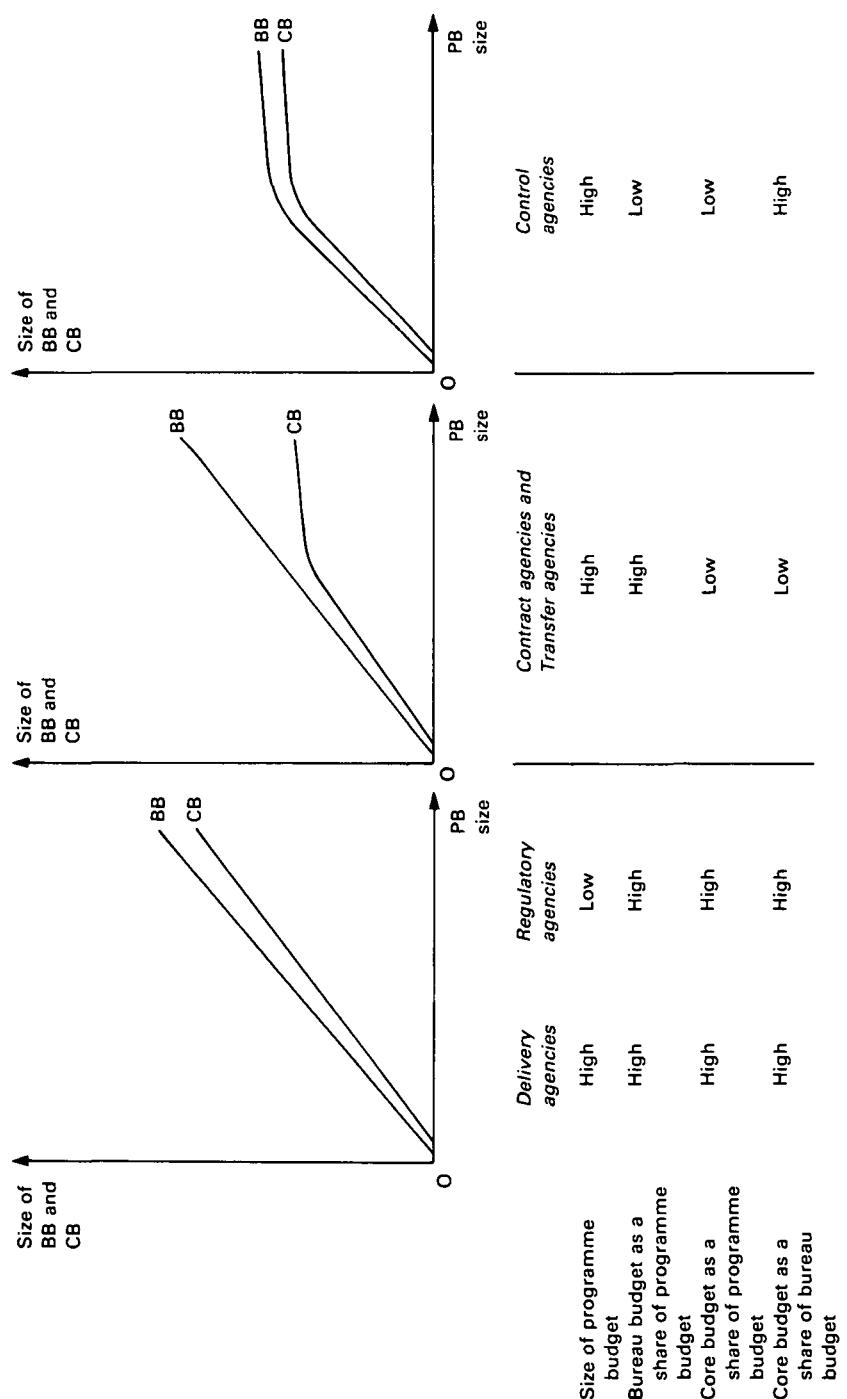


Fig. 4. The inter-relationship of programme, bureau and core budgets in five types of agencies

So far we have skated around an important definitional question, namely, what is to count as an 'agency' for the purposes of separating organizations into agency types. Clearly for the purposes of analysis we will normally need to decompose supposedly unified organizations into components consisting of different agency types. The categorization set out here is one which applies to agency *roles* rather than to whole institutions. Agency roles cannot be defined in organizational terms but only by reference to analytically delineated policy fields. Divisions within organizations between different component roles will be most common in national government organizations or other agencies which control complex administrative systems. For example, the DHSS has three agency roles: as a control agency for the personal social services supervising local authorities who deliver services on the ground; as a control agency for the National Health Service, where the delivery agencies are ultimately District Health Authorities, whose funding is routed via an additional tier of Regional Health Authorities which exercise more detailed control functions; and as a transfer agency directly administering delivery of pensions and other benefits. Similarly the Defence Department in Britain has DA, CLA and CTA roles.<sup>20</sup>

The final dimension of variations in bureaucratic behaviour concerns changes over time. So far we have assumed that the pattern of BB/PB relationships is fairly constant once an initial bureau-consolidating phase of growth has taken place. But for DAs, CTAs (and possibly also RAs and TAs), the *growth of the bureau* may run into severe constraints after a time. For example, there are severe size constraints on the growth of central government agencies imposed by:

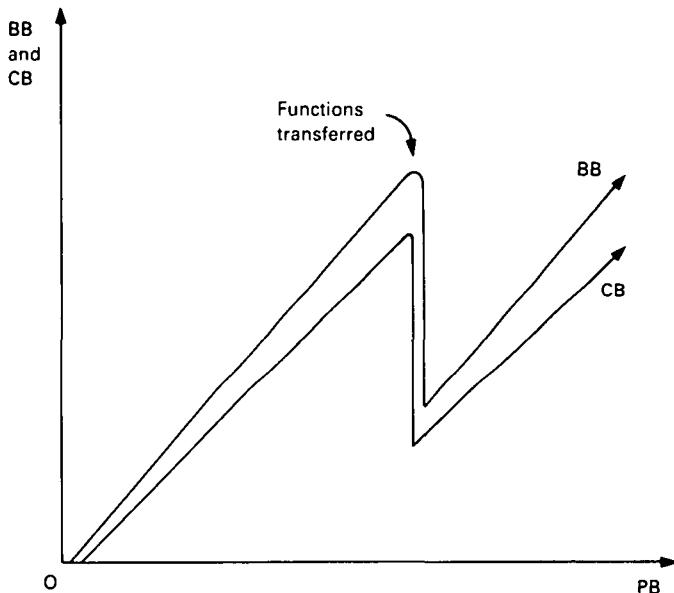
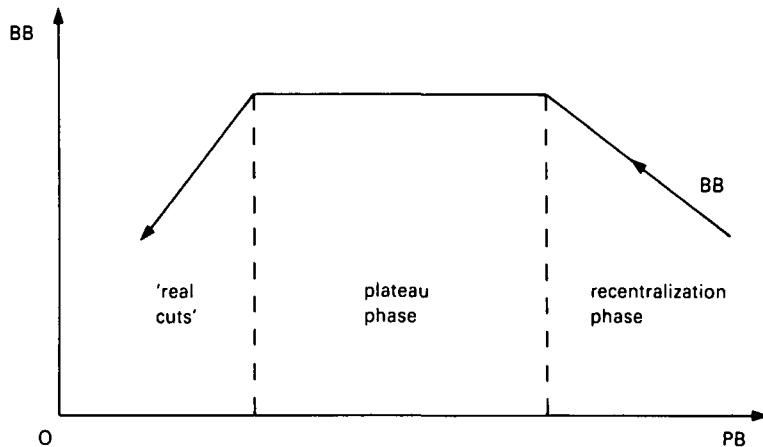
- the funnelling of management into political-administrative bottlenecks, such as the minister-permanent secretary relationship in the United Kingdom;<sup>21</sup>
- the spreading of top management and political attention over too wide an area, too thinly;<sup>22</sup>
- the accumulation of inertia in very large departments.

So it is perfectly feasible that the expansion of the bureau coming up against such constraints could trigger a period in which the bureau loses some of its existing functions to other (rival) departments or to quasi-government agencies set up to hive-off less salient functions. Either threat could produce a zig-zag growth curve for the bureau budget graphed against the programme budget, perhaps after a period of declining BB rate of growth (see Figure 5a).

<sup>20</sup> For a more detailed discussion see Dunsire and Hood, *Bureaumetrics*. Dunsire and Hood's painstaking empirical work was clearly restricted in value by their decision to focus on an institutional category ('the department'), rather than to develop statistics organized around more theoretically constituted variables.

<sup>21</sup> On which see Peter Kellner and Crowther-Hunt, *The Civil Servants* (London: Macdonald, 1980), pp. 174–238.

<sup>22</sup> A factor alleged to underlie the (partial) failure of giant departments in Whitehall. See C. Pollitt, *Manipulating the Machine* (London: Allen & Unwin, 1983).

(a) *Zig-zag bureau and core budget curves when functions are lost or hived off*(b) *Re-expansion of the bureau budget following cutbacks in the programme budget: the recentralization curve**Fig. 5. 'Zig-zag' and recentralizing bureau budget curves*

This sort of curve implies that there will be periods in the agency's evolution when rational officials will not advocate increases in the programme budget, because to do so could produce a quantum reduction in the bureau budget to which their utilities are more closely linked.

The BB/PB relationship also raises problems when we come to consider *declining budgets* in control agencies or contract agencies. In both cases dwindling programme budgets can be associated with a period of rising bureau budgets as the control agency recentralizes powers or functions from subordinate agencies, or as the contract agency brings back in-house functions it has previously sent out to outside professionals (Figure 5b).<sup>23</sup> (Even line agencies which use outside consultants or part-time staff may show this pattern.) After a while, the recentralization phase could be succeeded by cutbacks striking home within the central department itself. But there may well be extensive periods when officials in CLAs or CTAs welcome cutbacks in the programme budget as a means of increasing or stabilizing their bureau budgets, to which their utilities are linked.

#### 4. BUREAUCRATS' OPTIMAL BUDGET LEVELS

Part of the appeal of public-choice models of bureaucracy has clearly been the apparent correspondence between their account of budget-maximizing officials and 'common-sense' experiences of how bureaucracies operate. The bureaucratic empire-builder is not an off-beat theoretical construct, but seems to be an important and pervasive 'everyday' image of government officials as well. To define an alternative bureaucratic welfare maximand therefore entails uncovering a more plausible set of official objectives, one which can explain empirical phenomena with more precision or over a wider range of situations than the budget-maximization hypothesis. I argue in this section that even if bureaucrats are budget-maximizers, they are none the less always constrained by an internal benefit/cost calculation. Hence in contrast to Niskanen's picture of an open-ended budget maximizer constrained only by sponsor funding restrictions, bureaucrats pursue an optimal programme budget level with a finite limit. Section 5 sets out an alternative conception of what bureaucrats want (the 'bureau-shaping' model), and relates it to

<sup>23</sup> Dunleavy, *The Politics of Mass Housing in Britain* shows, for example, that in 1966–67, at the height of the public housing boom, local authority architects designed just half of the dwellings involved, private architects 30 per cent, and contractors' architects 20 per cent. By 1973, when the public housing programme was half the size, local authority architects' share increased to 75 per cent. In addition, the work tasks which public architects did on housing changed dramatically. In one authority (Birmingham) they were doing *solely* contract drafting and supervision (plus landscaping) in the late 1960s, whereas by the mid-1970s the council architects had regained control of designing most of their department's housing. This kind of effect will exist wherever public agency staffing does not expand fully to accommodate workloads in 'boom' periods, so that the character of their work tasks tends to stretch to accommodate the variation. This phenomenon implies that staff utilities at middle or lower levels may rise quite sharply in periods of limited budgetary reductions when workloads are reinternalized.

anecdotal evidence about their preferences. And Section 6 argues that the budget-maximizing model cannot account for the institutional form of modern state growth which the bureau-shaping model can easily explain.

We can make more explicit some of the earlier hints about the conditions under which budget maximization may or may not be a significant influence upon individual bureaucratic behaviour by looking at mappings of the discounted net utilities of a programme budget increment against the costs of advocating such an increment. These graphs need to be drawn for a particular rank of bureaucrat and a particular type of agency. But in practice we can distinguish three basic types of plot (Figure 6).

Discounted marginal utility curves are influenced mainly by:

- (i) *probability of influence*: hence top bureaucrats' curves are quite far away from the origin while bottom bureaucrats' curves are much closer in and shallower; and
- (ii) *the CB/PB and BB/PB ratios*: hence top bureaucrats in delivery, contract and regulatory agencies have fairly elastic slopes, but those in control and transfer agencies have steeply declining curves.

Marginal advocacy cost curves are influenced mainly by:

- (i) *the size of the existing PB*, set against the sorts of functions which the agency is carrying out, since with constant functions successive budgetary increments become progressively more costly to obtain;
- (ii) *external hostility* to an agency getting a PB increment: at the point where an agency is unlikely to get a further increase however hard it pushes, the cost curve (for top bureaucrats especially) becomes vertical;
- (iii) *changes in external hostility*: if it increases or decreases then the cost curve shifts up to the right or down to the left respectively; and
- (iv) *rank*: cost curves are higher and rise more steeply further up the rank hierarchy.

Budget maximization in these graphs implies that the bureaucrat will advocate expansion in the agency's activities if the actual PB of the bureau is to the left of her equilibrium PB position, but will do nothing if it is to the right of that position, switching attention instead to other individual or collective strategies for improving her welfare. (Note that the alternative to budget maximization is inaction, not advocacy of budgetary reductions.)

Because of the varying shapes of the discounted marginal utility and marginal cost curves, the same sort of changes in the external environment will have different impacts upon bureaucratic behaviour. For top bureaucrats in delivery, contract or regulatory agencies a shift of cost curves to the left will probably trigger large decreases in the equilibrium level of PB for that official, since increased external hostility produces sharp adjustments of the cost curve back along a utility curve that slopes fairly gently. But in control or transfer agencies at the same rank a large-scale adjustment in the marginal advocacy cost curve may not affect bureaucratic behaviour much because the DMU

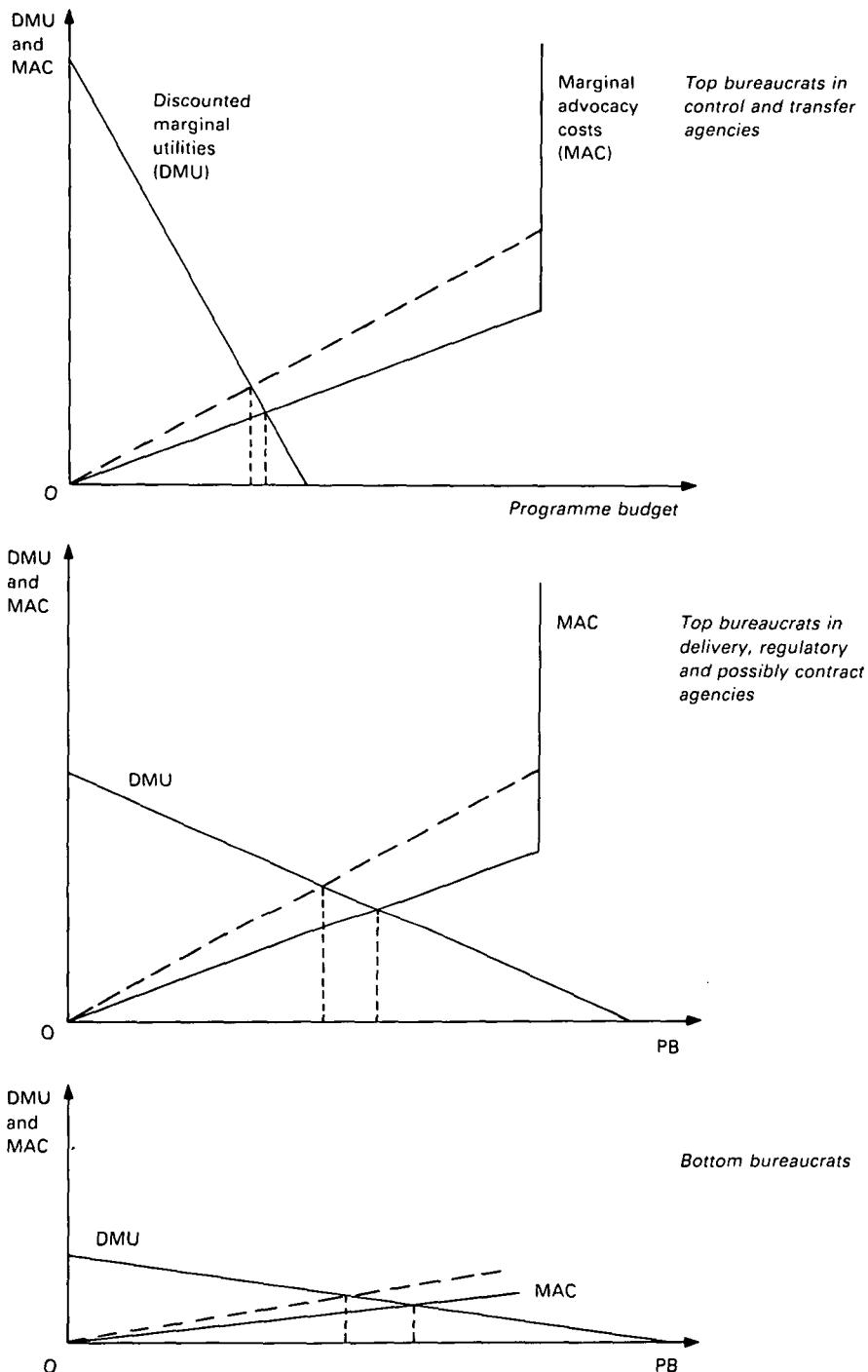


Fig. 6. The choice problem for bureaucrats in deciding whether to advocate budgetary increases

curve is so inelastic. For bottom bureaucrats, the same external shifts which manifest themselves as marked changes in the top bureaucrats' cost curves will produce only small increases in the shallow slope of cost curves. But since their discounted utility curves typically are also shallow, even a small shift may produce a sharp fall in their equilibrium PB level.

The internal determinants of the equilibrium PB point funnel through changes in the discounted net marginal utility curve. In particular, top bureaucrats may be in a position powerfully to increase the utility they derive from a given level of PB by internally reorganizing their agency. For example, if personnel ceilings are being enforced then the automation or computerization of clerical tasks can free manpower quotas to be redeployed at policy-making levels (for instance, by hiring more policy analysts).<sup>24</sup> Most accounts in the public-choice literature suggest that officials will simply push the DMU curve progressively further outwards from the origin, creating a succession of PB equilibrium points which follow the marginal cost curve outwards in the sequence shown in Figure 7a (*A*, *B*, *C*, etc.). Only if the expansion of the DMU curve is so great or the marginal cost curve is so restrictive that the equilibrium points lie along the vertical section of the marginal cost curve will there be any stasis in the level of PB which bureaucrats seek.

But it is actually rather unlikely that major internal reorganizations will simply inflate the DMU curve along its existing shape. Far more commonly we could expect the DMU curve to change shape, in particular to swing in a clockwise direction so as to maximize the benefits which senior bureaucrats derive from lower levels of the programme budget, while reducing their dependence on high levels of PB in order to improve their welfare (Figure 7b). Especially by hiving off responsibility for implementing public policies to other agencies or private sector bodies, top bureaucrats could improve their own welfare (for example, by shedding troublesome direct managerial responsibilities and gaining increased staff and time resources for intellectually more attractive tasks such as planning and guidance). Assume that exactly this kind of change underlies the shifts from DMU<sub>1</sub> to DMU<sub>2</sub>, shown here. If the marginal cost curve cuts these utility curves in their lower reaches as shown then the equilibrium budget level declines from *A* to *B*. Only if the marginal cost curve cuts the DMU curves in their upper reaches (i.e. if the external constraints on budgetary expansion are particularly restrictive) will the equilibrium programme budget position push steadily outwards. As in the

<sup>24</sup> Bennett and Johnson, *The Political Economy of Federal Government Growth*, pp. 30-7 demonstrate that the proportion of American federal employees in the bottom four of eighteen General Schedule grades declined from 40 per cent in 1959 to 21 per cent in 1978, while the proportion in the top five ranks grew from 6.5 per cent to 14 per cent in the same period. They compare this with the relatively static total of federal employees (up from 2.24 million in 1959 to 2.48 million twenty years later, and down as a proportion of the labour force from 3.2 to 2.4 per cent over the period), concluding that: 'a massive shift in General Schedule grades has occurred towards policy-making and program administration levels' (p. 38).

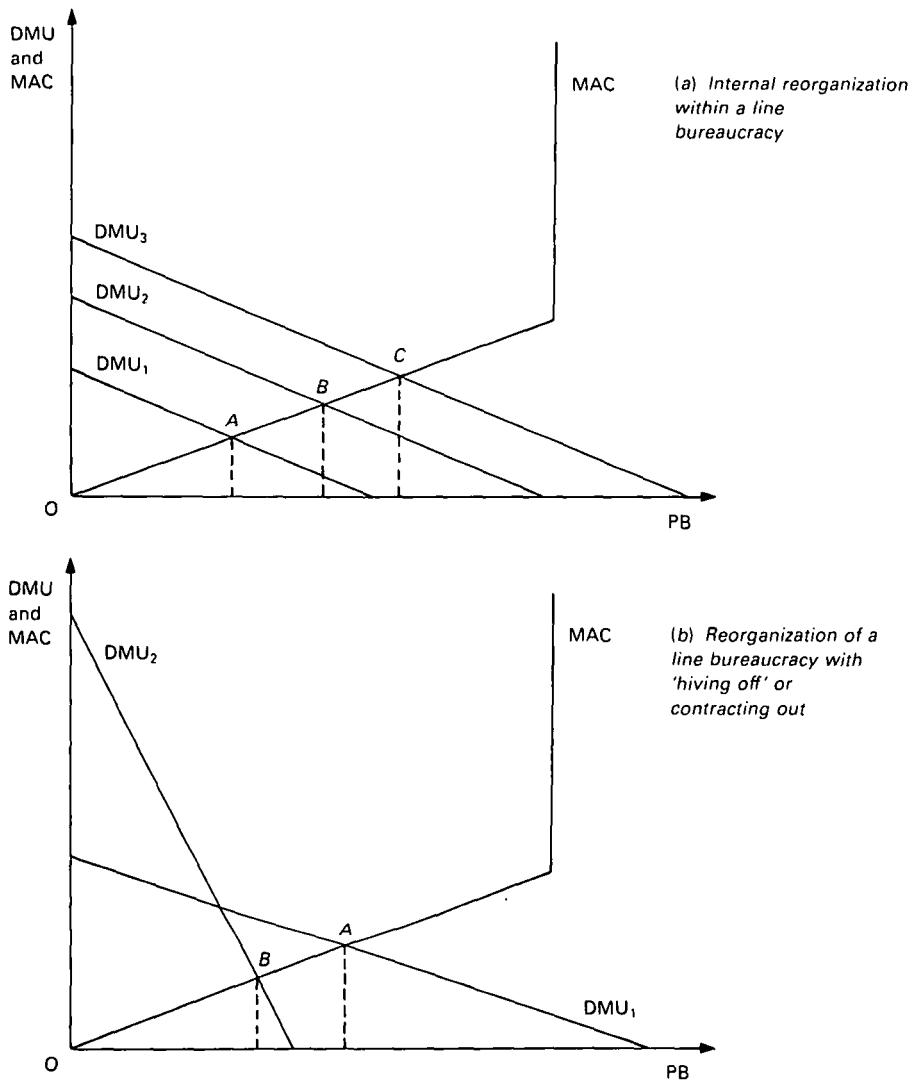


Fig. 7. The impact of bureau reorganizations on top bureaucrats' optimal budget equilibria

earlier example, if the DMU and marginal advocacy cost curves intersect in the vertical section of the MAC curve the equilibrium PB position will again be static.

A particularly interesting variant of the phenomenon of shifting DMU curves occurs where a delivery agency has been progressively transformed into a control agency. This sort of change will shift the DMU radically clockwise at the same time changing its shape from an elastic to an inelastic curve. *Ceteris paribus*, therefore, the optimum programme budget level for

top bureaucrats in the new control agency will be lower than that advocated while they still had direct line responsibilities. It should be clear that a prerequisite for an internal reorganization of this kind is that the area under the DMU curve but above the MAC curve should be greater with a control agency arrangement than under delivery agency arrangements – as is clearly the case here.

##### 5. THE BUREAU-SHAPING OBJECTIVE

Welfare-maximizing officials in policy-making ranks should emphasize collective strategies different from budget-maximization for three motivational reasons. First, there is a general presumption in the existing public-choice literature, most organization theory, and in Sections 1 and 2 above, that senior managers put less stress than lower ranking bureau members on the pecuniary or near-pecuniary components of their utility function. Instead there is a progressive shift of emphasis towards non-pecuniary utilities with increasing rank – a primary concern with status, prestige, patronage and influence, rather than with income, job security, or perks.

Secondly, the distinctive characteristics of public service employment place severe limits on the extent to which the pecuniary utilities of officials can be increased by such collective strategies as budget-maximization or the use of managerial discretion to divert resources to improving their personal welfare. Typically government officials' salaries are constrained within restrictive and standardized upper salary ceilings, lacking any counterpart to the very large and individualized 'prizes' paid out as salaries to key executives in private corporations. For example, the increase in responsibility from a single secondary school headteacher, to Director of Education in a county council (with many such schools), to Permanent Secretary at the DES (with a brief covering all schools and many other functions besides) boosts public-service salaries in Britain by only a few thousand pounds. In addition, it is common to find stringent controls on perks in the public sector, with relatively few company cars or fringe benefits, strict rules against officials having additional employment or business activities, standardized provision of accommodation and equipment, scarce support staff rationed out by standardized formulae, and personnel often administered centrally to minimize the creation of slack by individual managers. Compared with their counterparts in private companies, senior government officials are likely to find that budget-maximization is a remarkably frustrating activity in terms of direct near-pecuniary utility pay-offs.<sup>25</sup> Similarly blanket limits on staff numbers, centralized audit systems, and life-time career paths to senior positions, are all common features of public service employment systems which tend to reduce officials' ability to pursue their individual interests in pecuniary terms. All these traits are built

<sup>25</sup> See J. Gretton and A. Harrison, *How Much are Public Servants Worth?* (Oxford: Blackwell, 1982).

TABLE 2 Positive and Negative Values Ascribable to Bureaucrats

Positively valued	Negatively valued
<i>'Staff' functions</i>	<i>'Line' functions</i>
Individually innovative work	Routine work
Longer time horizons	Short time horizons
Broad scope of concerns	Narrow scope of concerns
Developmental rhythm	Repetitive rhythm
High level of managerial discretion	Low level of managerial discretion
Low level of public visibility	High level of grass-roots public visibility
<i>Collegial atmosphere</i>	<i>Corporate atmosphere</i>
Small-sized work unit	Large-sized work units
Restricted hierarchy and predominance of elite personnel	Extended hierarchy and predominance of non-elite personnel
Co-operative work patterns	Work patterns characterized by coercion and resistance
Congenial personal relations	Conflictual personal relations
<i>Central location</i>	<i>Peripheral location</i>
Proximate to the political power centres	Remote from political contacts
Metropolitan (capital city) location	Provincial location
Conferring high status social contacts	Remote from high status contacts

into public organizations with the aim of achieving exactly the kind of effects which they tend to produce. They are deliberate rather than accidental features, designed to displace senior officials' energies and efforts into work and policy-related aspects of their careers rather than into feathering personal nests. While senior officials' pay and conditions are normally maintained at levels sufficient to preserve their pre-existing position in the occupational class structure, they are also calculated to sift out from promotion people anxious to maximize pecuniary utilities.

Thirdly, it follows from the points above that a realistic individual-level model of why people enter career paths leading to senior positions in public agencies (or of why people temporarily transfer into such positions from the private sector in American-style systems) is likely to emphasize non-pecuniary elements in their utilities, especially aspects of their welfare which are related to the intrinsic characteristics of the work involved. Without trying to posit some other-regarding or ideological commitment by officials to their bureau or its 'mission' – i.e. without breaching the utility-maximizing instrumentalism which is the core of an economic approach – we can none the less acknowledge ample scope for the non-pecuniary elements of their utilities to be influential. In particular, Table 2 shows a list of pro- and anti-values which could plausibly be ascribed to self-regarding bureaucrats pursuing their own welfare. Clearly there is a pecuniary parameter in such officials' concerns

– a level of income and of near-money benefits which they will seek to achieve as a condition of the pursuit of other utilities. But this may be a constraint which is surmounted relatively easily and thereafter not very influential positively or negatively in structuring individual behaviour. It is certainly not an element of their utilities which officials seek to *maximize*. There are likely to be sharp differences in the perceived welfare of officials who share comparable salaries but are located in different positions on the dimensions included in Table 2 – between, say, similarly paid officials in the Cabinet Office and the DHSS central benefits office in Newcastle. These perceived differences translate directly into variable prestige, public recognition and influence. This model not only offers an instrumental explanation of why people enter public sector employment at policy-rank level, but also offers a plausible account of how officials typically pursue their careers. The key strategies individuals adopt in pursuit of objectives suggested by Table 2 are individual ones, searching for career paths which leave them favourably placed to reach an appropriate rank in a suitable sort of agency. Their success at this individual level will be far and away the most important influence upon their overall welfare (just as the success of people whose utilities are pecuniary is determined primarily by their individual strategies rather than the pursuit of collective goods such as budget maximization).

But if bureaucrats can maximize their welfare within a pecuniary constraint primarily by pursuing at an individual level the pro-values set out here, is there any *collective goods strategy* which they can pursue towards the same sort of objectives? How would this analogous strategy operate and what whole-bureau goals can be derived in line with it? In particular, how could such a strategy be as continuously present as budget-maximization, where the annual budgetary cycle can permit individual efforts to be fairly consistently deployed in boosting the budget? I suggest that rational officials in decision-making ranks characteristically adopt a *bureau-shaping strategy* designed to bring their bureau into a progressively closer approximation to 'staff' (rather than 'line') functions, a collegial atmosphere and a central location. They maximize this objective within a continuous bureau budget constraint but one which varies with the character and size of the agency. At each stage of this process officials seek to achieve a satisfactory level of budget, but this level in turn is set by their previous success in enhancing the bureau's conformity with the pro-values set out above. One of the most important consequences of successful bureau-shaping activities is a reduction in the size of the budget constraint over time – that is, the progressive unlinking of top officials' utilities from dependence on a high absolute level of programme or bureau budget – as the successful bureau takes on more of the small, central, elite character which crystallizes bureaucrats' pro-values.

There are four key means of pursuing bureau-shaping strategies:

- (i) *Major internal reorganizations* of the bureau, essentially those which increase its degree of conformity with an elite policy-making ideal. Thus

new acquisitions of functions will be concentrated at the policy-making level, while existing routine functions will be 'hived-off' to quasi-governmental agencies, 'hived-in' to separately designated departmental agencies or accountable management units, or contracted-out to private firms. In general where existing functions inconsistent with the bureau's ideal image cannot be coped with in one of these ways, they will be shunted into well-defined enclaves which need to be involved as little as possible with senior management. Often geographical separation is a key means of achieving this result.

- (ii) *Redefinition of relationships with external 'partners'* (such as subordinate public agencies, contractors, regulatees or client interest groups) can be important whether or not it is associated with an internal reorganization of the bureau. Agencies centrally involved with external organizations continuously seek to adjust their relationship with them so as to cut down on routine workloads but to maximize their agency's policy control. Hence they promote hands-off, auto-pilot controls for run-of-the-mill matters but increased discretionary involvement in policy-relevant issues. A shift towards a more corporatist style of relationship is frequently associated with this kind of change.<sup>26</sup> The bureau also tries to minimize its dependence upon external organizations, as the inter-organizational literature argues.<sup>27</sup> A high-density managerial or control workload can be a liability for a bureau if external or subordinate organizations refuse to co-operate. Eliminating such a potent source of inconvenience and stress, and replacing such arrangements with a more robust and insulating control apparatus is usually a priority. Bureaus also seek to extend the scope of their patronage of external bodies, but only where this can be achieved in line with their preferred image.
- (iii) *Competition with other bureaus* at the same level of government can be associated with both the preceding mechanisms. Bureaus always defend the *scope* of responsibilities involved in the existing programme budget, although they may be only weakly committed to defending given programme budget *levels*. Bureaus are by no means simple-minded imperialists. They compete with their rivals for programme tasks and policy areas which fit in with their ideal bureau form (especially those tasks with a high proportion of policy-ranking staff, which command useful resources and confer prestige or influence, and which tend to increase the average level of managerial discretion within the bureau). But bureaus may want to export troublesome and costly low-grade tasks to rivals, especially where doing so carries no major implications for a reduced programme budget.
- (iv) *Transformation of internal work practices* can bring major benefits to officials in policy-making ranks by increasing the interest of their work

<sup>26</sup> A. Cawson, *Corporatism and Social Welfare* (London: Heinemann, 1982) applies this idea to the social policy field in the United Kingdom.

<sup>27</sup> K. Hant and F. W. Scharpf, eds, *Interorganizational Policy-making* (London: Sage, 1978).

tasks, lengthening the time horizons used in decision making, and extending their discretionary ability to control policy. A shift towards more sophisticated management and policy analysis systems insulates the agency from criticism by rival bureaus, external partners or the sponsor body. It also tends to shift the balance of bureau personnel towards more high-level, skilled or professional staffs, improving existing bureau members' status and work content, as well as their career advancement prospects. Given the restrictive manpower ceilings applied in many public service personnel systems, a pre-condition for such changes may be the contracting out, reduction, computerization or automation of routine work tasks so that the staffing allocations involved can be redeployed in ways which confer more fruitful pay-offs for senior officials.<sup>28</sup> Bureau policy staffs also tend to promote more accountable management for routine enclave areas or lower-level staffs, but emphasize collegial decision making and diffused responsibility among policy-rank officials.<sup>29</sup>

It should be clear from this account that there is no *a priori* reason to think that bureau-shaping activities are less prevalent than budget maximization, or that the scope for individual officials to contribute to bureau-shaping strategies is any less than their ability to push up budgets. Like budget maximization, the pursuit of a bureau-shaping strategy requires collective action, especially by top (and perhaps, also, middle) ranking officials. But bureau-shaping has a much more important and visible connection with these officials' welfare than does generalized budgetary expansion *per se*.

The bureau-shaping model seems to fit closely with a large but disorganized stockpile of anecdotal data about how bureaucrats see themselves and about what they say they are trying to achieve. A strategy of maximizing a bureau's conformity to an ideal, high status organizational pattern, within a budget constraint contingent on the existing bureau configuration, certainly seems consistent with phenomena remarked upon by a wide range of approaches to administrative behaviour. Bureau-shaping activity appears to be every bit as commonplace and as frequently pointed out as are tendencies to budget maximization.

## 6. THE FORM OF STATE GROWTH

The main empirical evidence for this conception of bureaucratic motivations is the characteristic pattern of modern state growth. On the Niskanen model

<sup>28</sup> For example, in the United States federal government overall establishment numbers were frozen from 1977 to 1980; in the United Kingdom civil service manpower was first frozen in 1979, and then programmed to decrease by 7 per cent over three years (thanks primarily to the shedding of low-level line functions and to limited privatization). See G. K. Fry, 'The Development of the Thatcher Government's "Grand Strategy" for the Civil Service', *Public Administration*, LXII (1984), 322-36.

<sup>29</sup> Kellner and Crowther-Hunt, *The Civil Servants*, Chap. 12.

we should be able to take this pattern as indicative of government officials' 'revealed preferences'. Yet there is a basic contradiction between the over-supply hypothesis and the way in which governmental growth has taken place. If bureaucrats maximize their budgets, we should expect to see state growth taking the form of the accretion of high budget functions by large line agencies. In particular we should have witnessed an expansion of already existing large agencies, or certainly no diminution in their scale of activities or overall importance. And lastly, we should be confident in predicting a continuing centralization of functions and personnel within national government agencies, rather than at subordinate levels of the state.<sup>30</sup> None of these logical extrapolations from Niskanen's model can be sustained. Instead the evidence for each of them is markedly unfavourable. State growth has overwhelmingly been achieved by setting up decentralized networks of many smaller agencies.<sup>31</sup> Existing large line agencies have tended to lose functions or to be broken-up into their components. And new governmental growth has overwhelmingly been concentrated outside central (or federal) government, in sub-national governments, local authorities, spatially decentralized systems of quasi-governmental agencies, or in single-function agencies at the national level. Within central government agencies themselves there has been a marked shift towards non-executant status for the departments which remain. Civil service numbers have remained static or declined, while central departments' functions have concentrated on higher level managerial tasks.<sup>32</sup> Most individual central agencies have become smaller or stayed the same size, and almost all have lost 'line' functions and progressively assumed more of an exclusively elite character. All of these trends are exactly what the account of bureaucratic motivation given here would predict.

There are really only two options open to devotees of budget-maximization models in explaining these tendencies. Either they must argue that budget-maximizing officials have been most strikingly ineffective in influencing the *institutional form* of state expansion, while continuing to argue that they have been crucially important in influencing the *extent* of state expansion. Or they must conclude that senior officials' transparent lack of influence on the form of state expansion also reflects their supernumerary role in determining the extent of expansion. Neither conclusion seems particularly palatable, the first

<sup>30</sup> The criticism made here against the budget-maximizing thesis also applies with particular force to the aggregate-level explanation put forward by G. Tullock, 'Dynamic Hypotheses on Bureaucracy', *Public Choice*, xvii (1974), 128–32.

<sup>31</sup> For discussions of this 'Balkanization of the policy process' see P. Self, *Administrative Theories and Politics* (London: Allen & Unwin, 1977), p. 279; P. Dunleavy, 'The Limits to Local Government', in M. Boddy and C. Fudge, eds, *Local Socialism?* (London: Macmillan, 1984), and 'Quasi-Governmental Sector Professionalism', in A. Barker, ed., *Quangos in Britain* (London: Macmillan, 1982).

<sup>32</sup> C. Hood, 'Keeping the Centre Small: Explanations of Agency Type', *Political Studies*, xxvi (1978), 30–46; P. Dunleavy and R. Rhodes, 'Beyond Whitehall', in H. Drucker, P. Dunleavy, A. Gamble and G. Peele, eds, *Developments in British Politics* (London: Macmillan, 1983), pp. 104–13; Bennet and Johnson, *The Political Economy of Federal Government Growth*, *passim*.

because it seems logically inconsistent except under rather odd conditions, the second because it is implausible to suppose that large bureaucratic agencies have played so uninfluential a role in the growth of the state. The simplest and most direct way out of the conundrum is to drop the assumption that bureaucrats maximize budgets in favour of the view that they maximize the fit between their bureau and an optimal form of bureau design, within a contingently determined bureau budget constraint. Hence officials pursue a bureau shaping strategy rather than a policy of continuous budgetary expansion. The bureau shaping model views the form of state growth as a key reflection of officials' instrumental preferences, not as some puzzling anomaly falling outside the scope of effective theoretical explanation.

#### CONCLUSIONS AND DISCUSSION

Public choice models of bureaucracy which predict open-ended budget maximization are badly flawed internally. Bureaucrats typically do not embark on collective action modes of improving their welfare unless they have exhausted individual welfare-boosting strategies. If they do choose to try and increase budgets, rational officials typically confront familiar collective-action problems. In particular, although lower ranking bureaucrats have most to gain from budgetary expansion, they will know that the attainment of increments is almost completely insensitive to their individual advocacy, so that even though their advocacy costs are small, campaigning for budgetary expansion is unlikely to advance their individual utility. Higher ranking officials are aware that the attainment of budgetary growth will be much more sensitive to their personal contribution, but typically they have much less to gain from increments and will confront substantial advocacy costs in seeking to push through increases in the agency's base budget.

Budget maximization is anyway an ambiguous concept, since utility pay-offs are primarily associated with growth in the agency's core or bureau budgets, while advocacy costs are associated with the programme budget. There are additionally quite major differences between agency types in the extent to which officials associate their welfare with the growth of the programme budget. In delivery agencies (the classic line bureaucracies dwelt on in public choice models) the connection is close and positive. But in control agencies it is typically remote and variable.

Top bureaucrats' motivations can be modelled in terms of discounted marginal utilities and marginal advocacy costs, whose interaction identifies an optimal budget position which is an equilibrium point. Changes in the ways in which agencies are organized may shift this equilibrium point outwards over time – if the change involves simply reorganizing the way in which an agency carries out a fixed role. Or they may cause it to become smaller – if the change involves reshaping the agency into a different role, especially hiving off line functions to create a central control agency. This last example is especially relevant because it has been a predominant trend in the development of liberal democratic systems of public administration.

Finally, if we consider the characteristics of public service employment systems, it seems likely that the welfare of higher ranking bureaucrats is closely bound up with the intrinsic characteristics of their work. Rational bureaucrats would therefore concentrate on developing 'bureau-shaping' strategies designed primarily to bring their agency into line with an ideal configuration conferring high status and agreeable work tasks, within a budgetary constraint contingent on the existing and potential shape of the agency's activities. This hypothesis closely fits the tendency for state growth to create a deconcentrated network of multiple agencies with a non-executant central core, rather than the expansion of national line bureaucracies along the pattern predicted by the budget-maximization thesis.

Neither the pattern of argument nor the conclusions reached here should be read as an endorsement of the appropriateness of a public-choice approach for the analysis of political and administrative phenomena such as bureaucracy. This article is partly a pragmatic, *ad hominem* argument exploring the mileage that might be gained from reconstructing a public-choice model before discarding it, rather than seeking to rehabilitate it. There are certain heuristic advantages in developing an argument in opposition to the budget-maximizing hypothesis from within the terminology and using some of the analytical apparatus of 'soft' economic models. Chief of these is the closeness of these approaches to a structuralist mode of analysis, despite the avowedly individualistic beliefs of their exponents.<sup>33</sup> In contrast to conventional public administration with its stress on describing actors' behaviour in terms of the motivations and intentions of biographically realistic individuals, economic models offer a partial account of how actors operate as bearers-of-roles-in-organizations. As reconstructed here, a public-choice model can be read as an analysis of the situational logic facing any instrumental actor in a given role, with no necessary connection to the voluntaristic and dispositional style of explanation used by writers such as Downs. Lastly, of course, despite the characteristic ideological baggage of their exponents, public-choice models highlight the prevalence of 'emergent effects' in structures or systems of interaction.<sup>34</sup> Admittedly the level of structural explanation aimed at here is inter-positional and inter-organizational, rather than the macro-level accounts most often associated with radical approaches in the social sciences.<sup>35</sup> But there are good reasons to suppose that bureaucratic phe-

<sup>33</sup> Jackson, *The Political Economy of Bureaucracy*, p. 87, notes: 'In the neo-classical system agents are treated as if they are mindless automata who respond in a fully programmed fashion to external stimuli such as price and quantity signals'. For comparison, see J. Piaget, *Structuralism* (London: Routledge and Kegan Paul, 1962), which remains the best general introduction to the approach.

<sup>34</sup> R. Boudon, 'Undesired Consequences and Types of Systems of Interdependence', in P. Blau and R. Merton, eds, *Continuities in Structuralist Inquiry* (London: Sage, 1981). See also R. Boudon, *Effets pervers et ordre social* (Paris: Presses Universitaires de France, 1977).

<sup>35</sup> C. Warriner, 'Levels in the Study of Social Structure', defines inter-positional and inter-organizational levels of analysis, while Wolfe V. Heydebrand, 'Marxist Structuralism', discusses radical approaches, both in Blau and Merton, eds, *Continuities in Structuralist Inquiry*.

nomena are not simply reducible to explanation in terms of macro-social forces.<sup>36</sup> The internal organization and operations of the state apparatus instead have a specificity and significance of their own. If we are to develop an applied structuralist framework for analysing the logic of these particular social processes then the avenues indicated here seem worth exploring further.

<sup>36</sup> P. Dunleavy, 'Is There a Radical Approach to Public Administration?', *Public Administration*, LX (1982), 215–33.

---

An Essay on Fiscal Federalism

Author(s): Wallace E. Oates

Source: *Journal of Economic Literature*, Sep., 1999, Vol. 37, No. 3 (Sep., 1999), pp. 1120-1149

Published by: American Economic Association

Stable URL: <https://www.jstor.org/stable/2564874>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



is collaborating with JSTOR to digitize, preserve and extend access to *Journal of Economic Literature*

JSTOR

# An Essay on Fiscal Federalism

WALLACE E. OATES<sup>1</sup>

## 1. *Introduction*

FISCAL DECENTRALIZATION is in vogue. Both in the industrialized and in the developing world, nations are turning to devolution to improve the performance of their public sectors. In the United States, the central government has turned back significant portions of federal authority to the states for a wide range of major programs, including welfare, Medicaid, legal services, housing, and job training. The hope is that state and local governments, being closer to the people, will be more responsive to the particular preferences of their constituencies and will be able to find new and better ways to provide these services. In the United Kingdom, both Scotland and Wales have opted under the Blair government for their own regional parliaments. And in Italy the movement toward decentralization has gone so far as to encompass a serious proposal for the separation of the nation into two independent countries. In the developing world, we likewise see widespread interest in fiscal decentralization with the objective of breaking the grip of central

planning that, in the view of many, has failed to bring these nations onto a path of self-sustaining growth.

But the proper goal of restructuring the public sector cannot simply be decentralization. The public sector in nearly all countries consists of several different levels. The basic issue is one of aligning responsibilities and fiscal instruments with the proper levels of government. As Alexis de Tocqueville observed more than a century ago, "The federal system was created with the intention of combining the different advantages which result from the magnitude and the littleness of nations" (1980, v. I, p. 163). But to realize these "different advantages," we need to understand which functions and instruments are best centralized and which are best placed in the sphere of decentralized levels of government. This is the subject matter of fiscal federalism. As a subfield of public finance, fiscal federalism addresses the vertical structure of the public sector. It explores, both in normative and positive terms, the roles of the different levels of government and the ways in which they relate to one another through such instruments as intergovernmental grants.<sup>2</sup>

<sup>1</sup> Professor of Economics, University of Maryland, and University Fellow, Resources for the Future. I am most grateful for a host of helpful comments on an earlier draft from Robert Inman, Ronald McKinnon, Daniel Rubinfeld, Robert Schwab, John Wallis, Barry Weingast, and three anonymous referees; for research assistance from Tugrul Gurgur; and for the splendid editorial guidance of John Pencavel and John McMillan.

<sup>2</sup> This economic use of the term "federalism" is somewhat different from its standard use in political science, where it refers to a political system with a constitution that guarantees some range of autonomy and power to both central and

My purpose in this essay is not to provide a comprehensive survey of fiscal federalism. I begin with a brief review and some reflections on the traditional theory of fiscal federalism: the assignment of functions to levels of government, the welfare gains from fiscal decentralization, and the use of fiscal instruments. I then turn to some of the new directions in recent work in the field and explore a series of current topics: laboratory federalism, interjurisdictional competition and environmental federalism, the political economy of fiscal federalism, market-preserving federalism, and fiscal decentralization in the developing and transitional economies. Some of this research is expanding the scope of the traditional analyses in important and interesting ways. This will provide an opportunity both to comment on this new work and to suggest some potentially fruitful avenues for further research.

## *2. The Basic Theory of Fiscal Federalism: Some Comments*

The traditional theory of fiscal federalism lays out a general normative framework for the assignment of functions to different levels of government and the appropriate fiscal instruments for carrying out these functions (e.g., Richard Musgrave 1959; Oates 1972). At the most general level, this theory contends that the central government

---

decentralized levels of government. For an economist, nearly all public sectors are more or less federal in the sense of having different levels of government that provide public services and have some scope for *de facto* decision-making authority (irrespective of the formal constitution). In retrospect, it seems to me that the choice of the term "fiscal federalism" was probably an unfortunate one, since it suggests a narrow concern with budgetary matters. The subject of fiscal federalism, as I suggest above, encompasses much more, namely the whole range of issues relating to the vertical structure of the public sector.

should have the basic responsibility for the macroeconomic stabilization function and for income redistribution in the form of assistance to the poor. In both cases, the basic argument stems from some fundamental constraints on lower level governments. In the absence of monetary and exchange-rate prerogatives and with highly open economies that cannot contain much of the expansionary impact of fiscal stimuli, provincial, state, and local governments simply have very limited means for traditional macroeconomic control of their economies. Similarly, the mobility of economic units can seriously constrain attempts to redistribute income. An aggressive local program for the support of low-income households, for example, is likely to induce an influx of the poor and encourage an exodus of those with higher income who must bear the tax burden.<sup>3</sup> In addition to these functions, the central government must provide certain "national" public goods (like national defense) that provide services to the entire population of the country.

Decentralized levels of government have their *raison d'être* in the provision of goods and services whose consumption is limited to their own jurisdictions. By tailoring outputs of such goods and services to the particular preferences and circumstances of their

<sup>3</sup> It is straightforward to show that a system of decentralized poor relief is characterized by a garden-variety externality that results in suboptimal levels of support for the poor. More specifically, increases in support payments in one jurisdiction confer external benefits in the form of a reduced number of poor households elsewhere. On this, see Charles Brown and Oates (1985). There is, moreover, evidence for the U.S. that state-level decisions on levels of welfare support are interdependent; Luz Amparo Saavedra (1998), among others, finds that states have responded to decreases (increases) in benefit levels in other states by reducing (raising) their own benefits to welfare recipients. For an excellent survey of this whole issue, see Jan Brueckner (1998).

constituencies, decentralized provision increases economic welfare above that which results from the more uniform levels of such services that are likely under national provision. The basic point here is simply that the efficient level of output of a "local" public good (i.e., that for which the sum of residents' marginal benefits equals marginal cost) is likely to vary across jurisdictions as a result of both differences in preferences and cost differentials. To maximize overall social welfare thus requires that local outputs vary accordingly.

These precepts, however, should be regarded more as general "guidelines" than firm "principles." As has been pointed out in the literature, there is certainly some limited scope for decentralized macroeconomic efforts (Edward Gramlich 1987) and for assistance to the poor. In particular, there is a theoretical case for some poor relief at local levels (Mark Pauly 1973), and the fact is that state and local governments undertake a significant amount of redistributive activity.<sup>4</sup>

Moreover, this prescription is a quite general one. It does not offer a precise delineation of the specific goods and services to be provided at each level of government. And indeed the spatial pattern of consumption of certain goods and services like education and health is open to some debate. As a result, we find in cross-country comparisons some divergence in just what is considered,

<sup>4</sup> However, Martin Feldstein and Marian Vailant Wrobel (1998) present some recent evidence suggesting that state government attempts to redistribute income are largely unsuccessful. They find that progressive state income taxes in the U.S. have had little impact on the net-of-tax relative wage rates of skilled versus nonskilled workers. Their claim is that the mobility of workers across state borders undoes efforts at redistribution—and does so very quickly. The result is no redistribution, only deadweight losses from inefficient locational decisions.

say, "local" in its incidence. The specific pattern of goods and services provided by different levels of government will thus differ to some extent in time and place.<sup>5</sup> This is to be expected. Nonetheless, there remains much to be said for the basic principle of fiscal decentralization: the presumption that the provision of public services should be located at the lowest level of government encompassing, in a spatial sense, the relevant benefits and costs.<sup>6</sup>

Let me offer three observations on the general theory. First, the foundations of the Decentralization Theorem need some elaboration. The theorem is itself a straightforward normative proposition that states simply that ". . . in the absence of cost-savings from the centralized provision of a [local public] good and of interjurisdictional externalities, the level of welfare will always be at least as high (and typically higher) if Pareto-efficient levels of consumption are provided in each jurisdiction than if any single, uniform level of consumption is maintained across all jurisdictions" (Oates 1972, p. 54). The theorem thus establishes, on grounds of economic efficiency, a presumption in favor of the decentralized provision of public goods with localized effects. While the proposition may seem almost trivially obvious, it is of some interest both in terms of setting forth the conditions needed for its validity and, with

<sup>5</sup> For two useful treatments of the assignment of specific public services to the appropriate level of government, see Anwar Shah (1994, ch. 1) and Ronald McKinnon and Thomas Neehyba (1997).

<sup>6</sup> In Europe, proponents of fiscal decentralization refer to the "principle of subsidiarity." The precept here is that public policy and its implementation should be assigned to the lowest level of government with the capacity to achieve the objectives. This principle has been formally adopted as part of the Maastricht Treaty for European Union. Its intellectual roots, interestingly, are found in twentieth-century Catholic social philosophy. On this see Robert Inman and Daniel Rubinfeld (forthcoming).

some further analysis, for providing some insights into the determinants of the magnitude of the welfare gains from fiscal decentralization (Oates 1998).

But there is more to the story. The presumption in favor of decentralized finance is established by simply assuming that centralized provision will entail a uniform level of output across all jurisdictions. In a setting of perfect information, it would obviously be possible for a benevolent central planner to prescribe the set of differentiated local outputs that maximizes overall social welfare; there would be no need for fiscal decentralization (although one might wish to describe such an outcome as decentralized in spirit!). The response to this observation has been two-fold. First, one can realistically introduce some basic imperfections (or asymmetries) in information. More specifically, individual local governments are presumably much closer to the people and geography of their respective jurisdictions; they possess knowledge of both local preferences and cost conditions that a central agency is unlikely to have. And, second, there are typically political pressures (or perhaps even constitutional constraints) that limit the capacity of central governments to provide higher levels of public services in some jurisdictions than others. These constraints tend to require a certain degree of uniformity in central directives. There are thus important informational and political constraints that are likely to prevent central programs from generating an optimal pattern of local outputs.

My second observation concerns the magnitude of the welfare gains from fiscal decentralization. We can, in principle, measure the gains from the decentralized provision of public goods relative to a more uniform, centrally determined level of output. The theory

suggests that the magnitude of these gains depends both on the extent of the heterogeneity in demands across jurisdictions and any interjurisdictional differences in costs. In particular, we find that the potential gains from decentralization stemming from interjurisdictional differences in demand vary inversely with the price elasticity of demand. If the costs of provision are the same across jurisdictions, but demands differ, then the extent of the welfare loss from a centrally imposed, uniform level of output increases, other things equal, with the price inelasticity of demand.<sup>7</sup> There is a large body of econometric evidence that finds that the demand for local public goods is typically highly price inelastic. This suggests that the potential welfare gains from decentralized finance may well be quite large.<sup>8</sup>

Pursuing this point into the realm of positive economics, we might expect the magnitude of the potential gains from fiscal decentralization to have some explanatory power. Where these gains are large, we would expect to find that the public sector is more decentralized. In exploring this issue some years ago, I found some (perhaps vague) evidence in its support: in a sample of countries, the fiscal share of the central government varied inversely with an

<sup>7</sup> In tax analysis, we are accustomed to a quite different result: the deadweight loss varies directly with the price elasticity of demand. Here it is just the reverse, since the distortion takes place on the quantity, rather than the price, axis. But interestingly, if the source of the difference in efficient local outputs is cost differentials, then the gains from fiscal decentralization bear the opposite relationship to the case where their source is differences in levels of demand: these gains then vary directly with the price elasticity of demand (Oates 1998).

<sup>8</sup> For surveys of this econometric literature, see Rubinfeld (1987) and Oates (1996a). For an attempt actually to measure the welfare gains from decentralization, see David Bradford and Oates (1974); they find large gains.

index of "sectionalism," a measure of the extent to which people in geographical subareas of a country identify "self-consciously and distinctively with that area" (Oates 1972, pp. 207–208). More recently, Koleman Strumpf and Felix Oberholzer-Gee (1998), in a more sharply focused study of states and counties in the United States, find that the decision to allow counties a local option to legalize the consumption of alcoholic beverages depends significantly on a measure of the heterogeneity in preferences across counties within each state. There is, I think, some interesting work to be done in exploring the extent to which the potential gains from decentralization can explain the observed variation in actual governmental structure and policies.<sup>9</sup>

Third, I sense a widespread impression, suggested in some of the literature, that the gains from decentralization have their source in the famous Tiebout model (Charles Tiebout 1956). In this model, highly mobile households "vote with their feet": they choose as a jurisdiction of residence that locality that provides the fiscal package best suited to their tastes. In the limiting case, the Tiebout solution does indeed generate a first-best outcome that mimics the outcome in a competitive market. But the gains from decentralization, although typically enhanced by such mobility, are by no means wholly dependent upon them.<sup>10</sup> In fact, if

<sup>9</sup> Another interesting case is the setting of federal standards for safe drinking water. After mandating a set of standards for the quality of drinking water to be met in all jurisdictions in the Safe Drinking Water Act of 1974, the federal government has backed off and now allows a range of exceptions in recognition of the large interjurisdictional differences in per-capita costs of meeting the standards (U.S. Congressional Budget Office 1997).

<sup>10</sup> In certain settings, mobility can itself be a source of distorted outcomes. See, for example, the seminal paper by Frank Flatters, Vernon Henderson, and Peter Mieszkowski (1974).

there were absolutely nothing mobile—households, factors, or whatever—there would still exist, in general, gains from decentralization. The point here is simply that even in the absence of mobility, the efficient level of output of a "local" public good, as determined by the Samuelson condition that the sum of the marginal rates of substitution equals marginal cost, will typically vary from one jurisdiction to another. To take one example, the efficient level of air quality in Los Angeles is surely much different from that in, say, Chicago.

This point is of importance, because the Tiebout model is often viewed as a peculiarly U.S. construction. The relatively footloose households that it envisions, responding to such things as local schools and taxes, seem to characterize the U.S. much better than, say, most European countries. As a result, observers outside the U.S. tend to believe that this strand of the theory of local finance is of limited relevance in their settings. While there may well be some truth to this, it most emphatically does not follow that there are no longer any significant welfare gains from the decentralized provision of public goods.

### 3. *Fiscal Instruments in a Federal System*

To carry out their functions, the various levels of government require specific fiscal instruments. On the revenue side, governments will typically have access to tax and debt instruments. But in a federal system there is a further method for allocating funds among the different levels of the public sector: intergovernmental grants. One level of government may generate tax revenues in excess of its expenditures and then transfer the surplus to another level of government to finance part of the latter's budget. I want to review and

comment briefly on the use of these fiscal instruments in a federal fiscal system.

### 3.1 Taxation in a Federal System

The determination of the vertical structure of taxes is known in the literature as the "tax-assignment problem" (Charles McLure 1983). And the basic issue here is the normative question: Which taxes are best suited for use at the different levels of government? The question is typically posed in a setting in which there exists a nation state with a central government, where there is little or no mobility across national borders; at decentralized levels, in contrast, economic agents, goods, and resources have significant mobility across jurisdictional boundaries with the extent of this mobility increasing at successively lower levels of government. "Local" government, for analytical purposes, may sometimes be characterized as operating in a setting in which economic units can move costlessly among jurisdictions.

The difference in the mobility of taxed units at the central and decentralized levels has important implications for the design of the vertical structure of taxation. Taxes, as we know, can be the source of distortions in resource allocation, as buyers shift their purchases away from taxed goods. In a spatial setting, such distortions take the form of locational inefficiencies, as taxed units (or owners of taxed items) seek out jurisdictions where they can obtain relatively favorable tax treatment. High excise taxes in one jurisdiction, for example, may lead purchasers to bear unproductive travel costs in order to purchase the taxed items in jurisdictions with lower tax rates.

Such examples can suggest the conclusion that decentralized levels of government should avoid the taxation of

highly mobile economic units (be they households, capital, or final goods). But this in itself is not correct. The real implication is that decentralized levels of government should avoid *nonbenefit* taxes on mobile units. Or, more accurately, the analysis shows that on efficiency grounds decentralized governments *should* tax mobile economic units with benefit levies (Oates and Robert Schwab 1991; Oates 1996b). Such economic units, in short, should pay for the benefits that they receive from the public services that local governments provide to them.

The most well-known case of this is the earlier-discussed Tiebout model in which local jurisdictions use benefit taxes that effectively communicate to households the cost of consuming different levels of local public goods; this results in an efficient pattern of consumption of these goods. But this is true not only for households. If local governments provide local inputs that increase the productivity of capital employed in their jurisdictions, then they should levy benefit taxes on capital in order to provide the set of signals needed for the efficient deployment of capital across localities (Oates and Schwab 1991). In sum, efficiency requires not only that decentralized jurisdictions refrain from nonbenefit taxation of mobile economic units, but that they actively engage in benefit taxation where the public sector provides services to these units.

The public sector must for various reasons rely to a substantial extent on nonbenefit taxes. Redistributive programs that provide assistance to the poor, for example, simply transfer income. But, as noted earlier, such programs are not well suited to use at decentralized levels of government, where the mobility of economic units across local boundaries can undermine the

workings of such programs. It is for this reason that the literature suggests that nonbenefit taxes, to the extent they are needed, are best employed by higher levels of government.

But provincial, state, and local governments do, in fact, make use of some such levies.<sup>11</sup> In a seminal treatment of this issue making use of an optimal taxation framework, Roger Gordon (1983) has explored the ramifications of the decentralized use of a wide range of nonbenefit taxes. And Gordon finds several forms of potential distortion that result from an individual jurisdiction's ignoring the effects of its fiscal decisions elsewhere in the system; these include inefficiencies involving, for example, the "exporting" of tax burdens, external congestion effects, and impacts on levels of revenues in other jurisdictions, as well as certain equity issues associated with a generally regressive pattern of tax incidence.<sup>12</sup>

The analysis suggests, moreover, some guidelines for the use of such taxes. A reliance on *resident-based* taxes rather than *source-based* taxes, for example, can lessen tax-induced distortions by reducing the scope for tax-exporting (Inman and Rubinfeld 1996; McKinnon and Nechyba 1997).<sup>13</sup> The

<sup>11</sup> There is a lively and important debate in the local finance literature over whether or not local property taxation, as employed in the U.S., constitutes benefit taxation. Bruce Hamilton (1975, 1976) and William Fischel (1992) make the case that local property taxes combined with local zoning ordinances produce what is effectively a system of benefit taxation. Peter Mieszkowski and George Zodrow (1989) take the opposite view.

<sup>12</sup> See Inman and Rubinfeld (1996) for an excellent restatement and extension of the Gordon analysis. David Wildasin (1998a) provides a valuable survey of the various implications of factor mobility both for economic efficiency and for the redistributive impact of public policy.

<sup>13</sup> Resident-based taxes (also called "destination-based taxes") are levies on factors of production (such as land, labor, and capital) based on the owner's residence and on goods and services based on the residence of the consumer. In contrast,

analysis, moreover, establishes a presumption for the taxation of relatively immobile economic units. A particularly attractive tax base is unimproved land, since a tax on a factor or good in perfectly inelastic supply will not be the source of any locational inefficiencies. Such taxes (and any associated benefits from spending programs) will simply be capitalized into local land values. Thus, fiscally hard-pressed city governments have at their disposal a tax base that cannot escape them through mobility. There is some evidence in this regard that the city of Pittsburgh, which has used a graded property tax under which land is taxed at five times the rate on structures, has experienced an expansion in building activity that might not have been forthcoming in the presence of a higher tax on mobile capital (Oates and Schwab 1997).

### 3.2 Intergovernmental Grants and Revenue Sharing

Intergovernmental grants constitute a distinctive and important policy instrument in fiscal federalism that can serve a number of different functions. The literature emphasizes three potential roles for such grants: the internalization of spillover benefits to other jurisdictions, fiscal equalization across jurisdictions, and an improved overall tax system.

Grants can take either of two general forms. They can be "conditional grants" that place any of various kinds of restrictions on their use by the recipient. Or they can be "unconditional," that is,

---

source-based taxes (or "origin taxes") involve taxing factors where they are employed and goods and services where they are purchased. Under resident-based taxation, governments have much less capacity to export the incidence of their taxes onto economic units elsewhere. Source-based taxes, however, are often easier to administer and, in certain forms, tend to be more commonly used by state and local governments.

lump-sum transfers to be used in any way the recipient wishes. The theory prescribes that conditional grants in the form of matching grants (under which the grantor finances a specified share of the recipient's expenditure) be employed where the provision of local services generates benefits for residents of other jurisdictions. The rationale here is simply the usual Pigouvian one for subsidies that induce individuals (in this case policy-makers or the electorate) to incorporate spillover benefits into their decision-making calculus. The magnitude of the matching shares, in such instances, should reflect the extent of the spillovers.<sup>14</sup>

In contrast, unconditional grants are typically the appropriate vehicle for purposes of fiscal equalization. The purpose of these grants is to channel funds from relatively wealthy jurisdictions to poorer ones. Such transfers are often based on an equalization formula that measures the "fiscal need" and "fiscal capacity" of each province, state, or locality. These formulae result in a disproportionate share of the transfers going to those jurisdictions with the greatest fiscal need and the least fiscal capacity.<sup>15</sup>

Although widely used, equalizing intergovernmental grants are by no means

<sup>14</sup> Matching grants (possibly negative) can, in principle, also serve to correct some of the distortions associated with the decentralized use of nonbenefit taxes (Gordon 1983).

<sup>15</sup> Fiscal equalization can also make use of matching grants. If the objective of the equalization program is to equalize taxable capacity, the granting government may choose to supplement the revenue base of fiscally poorer jurisdictions by matching any revenues they collect by some specified percentage. Such a measure has the potential of allowing all jurisdictions to raise the same tax revenues per capita for a given tax rate (irrespective of the actual size of their tax base). This form of fiscal equalization is sometimes called "power-equalization" and has gotten some attention in the U.S. for state programs to achieve various equity goals—most notably in the area of school finance (e.g., Feldstein 1976; and Nechyba 1996).

a necessary feature of fiscal federalism (Dan Usher 1995; Robin Broaday 1996). Economists normally think of redistributive measures from rich to poor as those that transfer income from high-to low-income individuals. Intergovernmental equalizing transfers require a somewhat different justification based on social values.<sup>16</sup> In practice, such equalizing grants play a major role in many countries: in the fiscal systems of Australia, Canada, and Germany, for example, there are substantial transfers of income from wealthy provinces or states to poorer ones. In the United States, in contrast, equalizing grants from the federal to state governments have never amounted to much. Intergovernmental grants in the U.S. typically address specific functions or programs, but usually do not accomplish much in the way of fiscal equalization. At the levels of the states, however, there are many such programs under which states provide equalizing grants to local jurisdictions—notably school districts.

Fiscal equalization is a contentious issue from an efficiency perspective. Some observers see such grants as playing an important role in allowing poorer jurisdictions to compete effectively with fiscally stronger ones. This view holds that, in the absence of such grants, fiscally favored jurisdictions can exploit their position to promote continued economic growth, some of which comes

<sup>16</sup> The issue here is that from the perspective of redistributing income from rich to poor, equalizing intergovernmental grants are bound to have some perverse effects. For such grants, although transferring income from wealthy to poor *on average*, will inevitably result in some income transfers from poor individuals who reside in wealthy jurisdictions to rich persons in generally poor areas. In this sense, such equalizing measures are not as effective as programs that redistribute income from rich to poor *individuals*. But a society may well wish, for other reasons, to provide additional support for the provision of local public services (such as schools) in relatively low-income areas (e.g., Inman and Rubinfeld 1979).

at the expense of poorer ones. Fiscal equalization, from this perspective, helps to create a more level playing field for interjurisdictional competition.<sup>17</sup>

But the case is not entirely persuasive. Others have argued that fiscal equalization can stand in the way of needed regional adjustments that promote development in poorer regions. McKinnon (1997a), for example, contends that in the United States, the economic resurgence of the South following World War II resulted from relatively low levels of wages and other costs. It was this attraction of low wages and costs that ultimately induced economic movement to the South, bringing with it a new prosperity. Fiscal equalization, from this perspective, may actually hold back the development of poorer areas by impeding the needed interregional flow of resources (both emigration and immigration) in response to cost differentials.

But the primary justification for fiscal equalization must be on equity grounds. And it is as a redistributive issue that it continues to occupy a central place on the political stage. In some cases, as in Canada, it may provide the glue necessary to hold the federation together. In other instances, like Italy, it may become a divisive force, where regions, weary of large and longstanding transfers of funds to poorer areas, actually seek a dissolution of the union. Fiscal equalization is a complex economic and political issue.

The third potential role for intergovernmental grants is to sustain a more equitable and efficient overall tax system. For reasons we have discussed,

<sup>17</sup> As Broadway and Flatters (1982) have shown, equalizing grants may be required to offset distorting locational incentives where some jurisdictions offer pecuniary fiscal advantages to potential residents resulting, for example, from large, taxable natural resource endowments.

centrally administered, nonbenefit taxes with a single rate applying to the national tax base will not generate the sorts of locational inefficiencies associated with varying rates across decentralized jurisdictions. Moreover, central taxes can be more progressive, again without establishing fiscal incentives for relocation. There is, in fact, considerable evidence to indicate that state and local systems of taxes are typically more regressive than central taxation (e.g., Howard Chernick 1992). There is thus some force in an argument for "revenue sharing" under which the central government effectively serves as a tax-collecting agent for decentralized levels of government.<sup>18</sup> The central government then transfers funds, in a presumably unconditional form, to provinces, states, and/or localities. It is certainly possible, where the polity wishes, to build equalizing elements into these transfers. While there is here a real case for the use of intergovernmental grants, a most important qualification is that such a system of grants must not be too large in the sense of undermining fiscal discipline at lower levels of government (more on this later).

The prescriptive theory of intergovernmental grants thus leads to a vision of a system in which there exists a set of open-ended matching grants, where the matching rates reflect the extent of benefit spillovers across jurisdictional boundaries, and a set of unconditional grants for revenue sharing and, perhaps, equalization purposes. Such a conception has, however, only modest

<sup>18</sup> This argument has even more force where, as in some developing countries and emerging democracies, provincial and local governments simply lack the capacity for effective tax administration. In this setting, central transfers and/or the piggybacking of supplementary rates on top of centrally administered taxes may be the only realistic options. See, for example, Inman (forthcoming).

explanatory power. We do, in fact, find federal matching programs that have supported a number of state and local activities with spillover effects, including, for example, grants for interstate highway construction. However, on closer examination, important anomalies appear. These grants are often closed, rather than open, ended. They thus do not provide incentives for expansion at the margin. Moreover, the federal matching shares are typically much larger than justifiable by any plausible level of spillover benefits. More generally, in a careful study of the intergovernmental grant system, Inman (1988) concludes that the economic theory of intergovernmental grants does not provide a very satisfactory explanation of the structure of U.S. grant programs; he finds that a political model can do a much better job of explaining U.S. grant programs.<sup>19</sup>

Some years ago, David Bradford and I (1971a,b) tried to lay the foundations for a positive theory of the response to intergovernmental grants by setting forth a framework in which the budgetary decisions of the recipients of such grants are treated explicitly in a collective-choice setting. In short, we treated these grants, not as grants to an individual decision-maker, but rather as grants to polities that make budgetary decisions by some collective algorithm (such as simple majority rule). This exercise produced some intriguing equivalence theorems. For example, it is straightfor-

ward to show that a lump-sum grant to a group of people is fully equivalent in all its effects, both allocative and distributive, to a set of grants directly to the individuals in the group. Moreover, this result applies to an important class of collective-choice procedures, encompassing several of the major models employed in the public-finance literature. These theorems, known as the "veil hypothesis," thus imply that a grant to a community is fully equivalent to a central tax rebate to the individuals in the community; intergovernmental grants, according to this view, are simply a "veil" for a federal tax cut.

The difficulty is that this hypothesis has not fared well in empirical testing. It implies that the budgetary response to an intergovernmental transfer should be (roughly) the same as the response to an equal increase in private income in the community. But empirical studies of the response to grants have rejected this equivalence time and again. Such studies invariably find that state and local government spending is much more responsive to increases in intergovernmental receipts than it is to increases in the community's private income. And this has come to be known as the "flypaper effect"—money sticks where it hits. While this finding may not be all that surprising, it is not so easy to reconcile with models of rational choice, for it suggests that the same budget constraint gives rise to different choices depending on what form the increment to the budget takes. There is now a large literature that tries in a variety of ways (some quite ingenious) to explain the flypaper effect.<sup>20</sup> James Hines and Richard Thaler (1995) have suggested recently that this is just one of a more general class of cases where having

<sup>19</sup> As Inman and Rubinfeld (1996) point out, the prescriptive theory of grants presumes a central planner or political process that "will select socially preferred policies" (p. 325). However, the public-choice literature makes clear the potential of central-government political mechanisms to make inefficient choices concerning policies that affect various groups differently. In addition, a grant-distributing agency may have its own objectives; for an excellent study of how such objectives can influence the pattern of grants, see Chernick (1979).

<sup>20</sup> For surveys and interpretations of this literature, see Gramlich (1977), Ronald Fisher (1982), Oates (1994), and Hines and Thaler (1995).

money on hand (e.g., from grants) has a much different effect on spending behavior than where the money must be raised (e.g., by taxation).

Much of the early empirical work on the expenditure response to intergovernmental grants studied the period from the 1950's through the 1970's, when these grants exhibited a continuing path of expansion. As a result, much of the interest focused on the budgetary response to increases in grants. However, in more recent times, efforts at fiscal retrenchment and devolution have led to large cuts in a wide range of federal grant programs. And this has raised the interesting and important question of whether the response to cuts in grants is similar in sign and magnitude to the response to increases in these grants. Gramlich (1987), for example, observed that during this period of retrenchment, state and local governments responded to the cutbacks in grants by picking up much of the slack: they increased their own taxes and replaced in large part the lost grant funds so as to maintain levels of existing programs. If Gramlich is right, then we should observe a basic asymmetry in response: the spending of recipients should be more responsive to increases in grant monies than to decreases in these revenues. This issue is of some importance if we are to understand the budgetary implications of the ongoing process of fiscal decentralization. In the first study of this issue, William Stine (1994), examining the response of county governments in Pennsylvania, found just the opposite of Gramlich's prediction: his estimates imply that these county governments not only failed to replace lost grant revenues, but that they reduced their spending from own-revenues on these programs as well, giving rise to a "super-flypaper effect." There are, however, some tricky

and troublesome issues of measurement and interpretation in the Stine study. Subsequently, using national aggregate data on the state and local government sector, Shama Gamkhar and I (1996) were unable to reject the hypothesis that the expenditure response to increases and decreases in intergovernmental grants has the same absolute value per dollar of grants. Our findings are thus consistent with the proposition that the flypaper effect operates symmetrically in both directions. But much clearly remains to be done on this issue.

#### *4. A Note on Jurisdictional Boundaries*

The treatment to this point has implicitly taken as given a pattern of boundaries that divide the nation-state into a set of jurisdictions for decentralized governance. The existence and magnitude of spillover effects from localized public policies clearly depend on the geographical extent of the relevant jurisdiction. One way to deal with such spillovers is to increase the size of the jurisdiction, thereby internalizing all the benefits and costs. The problem, of course, is that such an extension may involve welfare losses from the reduced capacity to differentiate local outputs. There is clearly some kind of tradeoff here between internalizing spillover benefits (and costs) and allowing local differentiation.

In practice, much of the problem stems from a set of existing boundaries that are largely historically and culturally determined and that may make little sense in terms of the economic and geographical realities. Consider, for example, the United States. Suppose that we were to begin with a *tabula rasa*, a completely undefined set of boundaries for states and localities. And we set for ourselves the task of laying out both a rational set of levels of government and

borders for the jurisdictions at each level of government. One thing seems clear: such a system of jurisdictions would bear little resemblance to our existing map. The states, in particular, are quite poorly designed to deal with the provision of certain important public goods, notably environmental resources. To take one example, rivers were used historically (for understandable reasons) to mark off one state from another. But from the perspective of effective management of a public good, this is the worst sort of border. It means that two independent and autonomous jurisdictions are making decisions that affect the public good whose output they jointly share. It seems clear that it would make much more sense to place such resources within a single jurisdiction. My own surmise is that a much more rational map would probably entail (1) some fairly sizeable regional governments that extend over watersheds, air sheds, and other environmental resources; (2) metropolitan governments that encompass center cities and the suburbs that house many city workers; and (3) smaller local governments that allow groups of residents to determine services of relevance mainly to themselves.

But political realities being what they are, we can expect to continue our collective life with much the same map in place. There does, however, remain some flexibility in terms of creating useful compacts or associations of jurisdictions to deal with particular issues. The management of the Chesapeake Bay, for example, is in important organizational ways now the joint enterprise of the relevant states (Delaware, Maryland, Pennsylvania, and Virginia), and Washington, D.C., with an important role also played by the federal government. Likewise, the recognition that the management of ground-level ozone involves

pollutants that travel long distances across the midwestern and northeastern parts of the United States has led, under congressional legislation in 1990, to the formation of an Ozone Transport Region (OTR) for the coordination of efforts to manage air quality in eleven eastern states and the District of Columbia. Such regional organizations can be seen as the outcome of a kind of Coasian process in which interjurisdictional externalities are addressed through negotiation and coordinated decision-making. The history of such enterprises, however, attests to their difficulty. The fascinating study by Bruce Ackerman et al. (1974), for example, of the attempt to create a "model regional agency" in the form of the Delaware River Basin Commission reveals all the complexities and perverse incentives that can bedevil such joint enterprises. Nevertheless, such coordination does, in principle, offer an important avenue for addressing such interjurisdictional concerns.

##### *5. Laboratory Federalism and Welfare Reform*

It seems ironical in the light of the preceding treatment of principles (or guidelines) for fiscal federalism to find that welfare reform is in the vanguard of U.S. moves toward fiscal decentralization. The analysis suggests that the threat of mobility of both low and high income households will result in decentralized policies that provide too little assistance to the poor (sometimes described as a "race to the bottom"). Nevertheless, the decision has been made to shift the primary responsibility for poor relief back to the states. Under measures signed into law in 1996, the federal government has replaced the longstanding federal entitlement programs, which came with both detailed rules and generous matching grants to

the states, by a system of block grants with few strings attached. The states now have broad scope to determine both the form and levels of assistance under their programs to assist poor households.<sup>21</sup>

How are we to understand this reform? Does it represent an outright rejection of the economic principles of fiscal federalism? My answer is a qualified no. There exists widespread recognition of, and concern with, the likely shortcomings of a decentralized system of poor relief. Policy makers are well aware of the threat of strategic cuts in state levels of welfare support. But, as I read it, we have decided to live with this threat in order to seek out superior policy alternatives. And this brings us to another dimension of fiscal federalism: laboratory federalism.

In a setting of imperfect information with learning-by-doing, there are potential gains from experimentation with a variety of policies for addressing social and economic problems. And a federal system may offer some real opportunities for encouraging such experimentation and thereby promoting "technical progress" in public policy. This point was made long ago by James Bryce (1888) who, in his insightful study of the U.S. system of government, observed that "Federalism enables a people to try experiments which could not safely be tried in a large centralized country" (Vol. I, p. 353). Better known is a later statement by Justice Louis Brandeis, who wrote in 1932 that

There must be power in the States and the Nation to remould, through experimentation, our economic practices and institutions to meet changing social and economic needs

<sup>21</sup> For an excellent and recent review of this whole debate in a historical context, see Therese McGuire (1997). Rebecca Blank (1997) provides a concise and insightful treatment of the new welfare legislation and its potential implications.

... It is one of the happy incidents of the federal system that a single courageous State may, if its citizens choose, serve as a laboratory; and try novel social and economic experiments without risk to the rest of the country. (Osborne 1988)

It is my sense that this is the primary thrust behind the current welfare reform. There exists much disappointment and dissatisfaction with the operation and results under the traditional federal welfare programs. But we really don't have a clear sense of how to restructure them to achieve our societal goals of providing needed relief and, at the same time, establishing an effective set of incentives to move people off welfare and into jobs. The recent legislation that transfers the responsibility for these programs back to the states represents, I believe, a recognition of the failure of existing programs and an attempt to make use of the states as "laboratories" to try to find out what sorts of programs can work.<sup>22</sup>

There are, in fact, a number of important and intriguing examples of policies whose advent was at the state or local level and that later became fixtures of federal policy. Unemployment insurance, for example, was a state-level policy before the federal government made it effectively mandatory on a national scale in the 1930s. More recently, in the area of environmental policy, the experience in a number of states with their own forms of Emissions Trading was an important prelude to the adoption, in the 1990 Clean Air Act Amendments, of a national trading program in sulfur allowances to address the problem of acid rain. Without this experience in a number of states, I seriously doubt that policy-makers would have been willing to introduce such a new and unfamiliar policy measure as tradable emissions rights on a national

<sup>22</sup> For a concurring view, see Craig Volden (1997).

scale. More generally, since the dawn of the nation, programs successfully developed at the state level have often provided models for subsequent federal programs.

States, of course, may learn from others so that the diffusion of successful policy innovations may be horizontal as well as vertical. Both forms of diffusion have been the subject of study by a number of political scientists. Virginia Gray (1973) and Everett Rogers (1983), for example, have found that the cumulative distribution of states by date of adoption takes the S-curve shape, familiar from the study of the spread of other forms of innovation. Others, like Jack Walker (1969), James Lutz (1987), David Huff et al. (1988), and David Nice (1994), have explored the geographical and other determinants of the pattern of adoptions by states. Empirical studies of vertical diffusion are less numerous. Thomas Anton (1989), Keith Boeckelman (1992), and Michael Sparer and Lawrence Brown (1996) have examined the extent to which federal measures draw on the experience of the states. Some of this literature is relatively skeptical of the link. Sparer and Brown, for example, argue that (at least for health care) "These laboratory adoptions and adaptations are probably more the exception than the rule" (p. 196).

What are we to make of all this? A little reflection suggests first that there is nothing in principle to prevent the central government from undertaking limited experiments without committing the nation to an untested and risky policy measure. Indeed, there have been a number of such social experiments with, for example, income-maintenance and housing-allowance programs that have generated valuable information about how programs work and the response of participants to various values of the key parameters. We don't neces-

sarily need states as the "laboratories" for experiments. At the same time, one might suspect that relatively independent efforts in a large number of states will generate a wider variety of approaches to public policy than a set of centrally designed experiments.

A basic problem here is that there has been little in the way of a real theory of laboratory federalism to organize our thought and to guide empirical studies. However, the beginnings of some theory are emerging, and they are quite illuminating. Susan Rose-Ackerman (1980) and, more recently, Strumpf (1997) have taken two quite different formal approaches to policy innovation in a federal system. One insight emerging from their analyses is an important, if familiar and unsurprising, one. There exists a basic "information externality" in that states that adopt new and experimental policies generate valuable information for others. And this creates a standard sort of incentive for free-riding. From this perspective, we might expect too little experimentation and policy innovation in a highly decentralized public sector. Indeed, as Strumpf shows, it is unclear whether a centralized or decentralized outcome will result in more policy innovation.<sup>23</sup>

The underprovision of experimentation at state and local levels can be addressed through a system of subsidies to encourage these activities. And this raises another point regarding existing welfare reform in the U.S. Under earlier programs, federal aid took a matching form such that the federal government effectively shared the costs and risks of new state-level programs. But

<sup>23</sup> The Rose-Ackerman and Strumpf analyses, incidentally, also produce a number of subtle and more surprising results. Strumpf finds, for example, that a state with a higher expected return from experimentation can have a lower propensity to experiment.

under the new welfare reform measures, matching aid has been replaced by block grants. This in itself serves to reduce incentives for experimentation. There are some conflicting incentives here. On the one hand, the new legislation gives the states broader scope for experimentation, but it places the full cost of any new measures on the state with no sharing from the center. The net outcome on the amount of experimentation is thus *a priori* unclear.

More generally, we need a lot more work on the implications of fiscal decentralization for both the amount and kinds of policy experimentation and innovation. As I have suggested, there are some clear and important cases where innovation and experimentation at state and local levels have led to new policy measures that have had broad national application. But it is much less clear how we are to understand this experience in terms of the overall effectiveness of a federal system in policy innovation.

#### *6. Interjurisdictional Competition and Environmental Federalism: A Challenge to the Basic View*

The preceding sections have set forth an economic conception of a federal system. It is one in which the central government plays the major role in macroeconomic stabilization policies, takes the lead in redistributive measures for support for the poor, and provides a set of national public goods. Decentralized levels of government focus their efforts on providing public goods whose consumption is limited primarily to their own constituencies. In this way, they can adapt outputs of such services to the particular tastes, costs, and other circumstances that characterize their own jurisdictions.

The general idea of decentralizing the provision of public services to the jurisdictions of concern has been widely

recognized. It manifests itself clearly on both sides of the Atlantic. We see it in Europe under the nomenclature of the "principle of subsidiarity," where it is explicitly enshrined in the Maastricht Treaty as a fundamental principle for European union. In the U.S., it often appears more informally as an aversion to the "one size fits all" approach.

Somewhat paradoxically, however, this view is the subject of a widespread and fundamental challenge both at the theoretical and policy levels. The source of this challenge is the claim that interjurisdictional competition among decentralized levels of government introduces serious allocative distortions. In their eagerness to promote economic development with the creation of new jobs (so the argument goes), state and local officials tend to hold down tax rates and, consequently, outputs of public services so as to reduce the costs for existing and prospective business enterprise. This results in a "race to the bottom" with suboptimal outputs of public services.<sup>24</sup>

This argument has a substantial history. Some thirty years ago, for example, George Break (1967) made the case for the detrimental effects of interjurisdictional competition:

The trouble is that state and local governments have been engaged for some time in an increasingly active competition among themselves for new business . . . In such an environment government officials do not lightly propose increases in their own tax rates that go much beyond those prevailing in nearby states or in any area with similar natural attractions for industry . . . Active tax competition, in short, tends to produce either a generally low level of state-local tax effort or a state-local tax structure with strong regressive elements. (Break 1967, pp. 23–24).

<sup>24</sup> Competition may also take place between different levels of government. On such "vertical competition" (as well as horizontal competition), see Albert Breton (1998).

Fear of losing local business and jobs thus leads to suboptimal levels of state and local public goods. Such competition can involve regulatory as well as purely fiscal policies. John Cumberland (1979, 1981) has extended the Break argument to encompass the setting of standards for local environmental quality. In the Break spirit, Cumberland contends that state and local governments engage in "destructive interregional competition." In order to attract new business and create jobs, public officials compete by reducing local environmental standards to lower the costs of pollution control for firms that locate within their borders. In this instance, interjurisdictional competition leads to excessive environmental degradation. The implication of the Cumberland view is that national standards for environmental quality are needed to prevent the excessive levels of pollution forthcoming under state and local standard setting.

More recently, Alice Rivlin (1992) has echoed these views in her "rethinking of U.S. federalism." Although advocating an extensive devolution of public-sector responsibilities to state and local government, Rivlin sees it as almost axiomatic that competition among the states results in inadequate levels of public services. Her remedy is a system of shared taxes under which the revenues from a new national value-added tax would be shared among the states. This, she argues, would free the states so that they would not have "to worry so much about losing businesses to neighboring states with lower tax rates" (p. 142).

This line of argument has proved quite powerful in the policy arena. There are strong forces for the "harmonization" of fiscal and environmental measures in Europe that draw heavily on this proposition. Likewise, the case for the "race to the bottom" has pro-

vided basic support for the centralization of environmental management in the United States.

What I want to stress here is the fundamental character of this challenge to the basic model of fiscal federalism. The claim is that the decentralized provision of public services is basically flawed; in the words of one recent U.S. observer, we need centralization in order to "Save the States from Themselves" (Peter Enrich 1996).<sup>25</sup>

But is this claim in fact true? This turns out to be a very complicated question both in theoretical and empirical terms. There is now a substantial theoretical literature that addresses this issue. In one set of papers, my colleague Robert Schwab and I have developed a series of models that explore the conditions under which horizontal competition among governments is efficiency-enhancing (Oates and Schwab 1988, 1991, 1996). It turns out that it is straightforward to develop an analogue to perfect competition in the private sector. In such a setting, governments compete with one another for a mobile capital stock that both generates income for local residents and provides a tax base for them—and such competition leads local officials to adopt efficient levels of outputs of public goods and tax rates. In these models, the invisible hand works in much the same way as in the private sector to channel policy decisions in individual jurisdictions into an efficient outcome from a national perspective.

These models, moreover, are quite rich in terms of the variety of policy instruments. Public officials provide not

<sup>25</sup> There is, incidentally, a very extensive, interesting, and lively debate on this matter among legal scholars. Recent issues of the law journals are full of papers on interjurisdictional competition and its consequences. See, for example, Richard Revesz (1992) and Daniel Esty (1996).

only outputs for local residents, but public inputs that enhance the productivity of locally employed capital, and environmental regulations that impose costs on local business and improve local environmental quality. They finance these public outputs with a set of taxes on local residents and capital. And there is no race to the bottom here. Instead, jurisdictions find it in their own interest to charge benefit taxes that lead to efficient decisions in both the public and private sectors.<sup>26</sup>

The problem is that these models make some strong assumptions. Let me note three of them here: jurisdictions behave as price-takers in national or international capital markets; public officials seek in their decisions to maximize the welfare of their constituencies; and these officials have access to the needed fiscal and regulatory policy instruments to carry out their programs efficiently. It is not hard to show (or surprising to find) that violations of any of these conditions can lead to distorted outcomes. Suppose, for example, that local policy makers are Niskanen-type agents that seek to maximize, not the well-being of their constituencies, but rather the size of the local public budget. It is then straightforward to show that they will set excessively lax environmental standards in order to encourage a larger inflow of capital so as to enlarge the local tax base (Oates and Schwab 1988).

The Oates-Schwab models provide a

<sup>26</sup> I should emphasize here that all public outputs (including environmental quality) are entirely local in these models; there are no spillover effects into other jurisdictions. The analysis, incidentally, extends not only to fiscal instruments, but regulatory ones as well (such as environmental standards). The analysis of "regulatory federalism" is, in principle, analogous to that of fiscal federalism. The same general principles concerning decentralization apply to fiscal and regulatory instruments.

kind of baseline from which one can introduce a range of quite plausible and realistic modifications that can be the source of allocative distortions. A large number of papers explore outcomes either where jurisdictions are sufficiently large to have some influence over the price of capital or where local governments are restricted in their access to policy instruments and must, for example, tax business and household capital at the same rate. Many of these papers employ game-theoretic approaches in which there is strategic interaction among the jurisdictions (Wildasin 1988). In such settings, we find that outcomes can easily occur that involve suboptimal levels of public outputs.<sup>27</sup>

The theoretical literature thus generates some diverse findings on this issue. There seem to be some basic efficiency-enhancing aspects of interjurisdictional competition, but there are clearly a range of "imperfections" that can be the source of allocative distortions. The real issue here is the magnitude of these distortions. Are we dealing with minor deviations from efficient outcomes—or does such competition produce major welfare losses? The pure theory can't help us much in answering this question. Moreover, some of the terminology is not very helpful. In particular, the description of interjurisdictional competition as involving a "race to the bottom" seems quite misleading. Such a descriptive image may well be an effective rhetorical device: it conjures up a vision of one jurisdiction cutting its tax rates and lowering its environmental standards, only to be outdone by a neighboring jurisdiction, in a process that leads to a downward spiral to the "bottom" (suggesting a very bad

<sup>27</sup> See John Wilson (1996) for an excellent survey of this literature.

outcome indeed). However, the models that generate these results are nothing of the sort. They are often game-theoretic models that produce Nash equilibria with suboptimal public outputs as the outcome. What matters here is the extent of the suboptimality. And the race-to-the-bottom terminology tends to obscure this issue.

Unfortunately, we do not have many empirical studies to bring to bear on this matter. There is a substantial descriptive literature addressing economic competition among state and local governments in the U.S., with some interesting findings (Timothy Bartik 1991). But this body of work really does not shed much light on the normative question of whether such competition is efficiency-enhancing or not (Paul Courant 1994). In an interesting study that is of relevance, Anne Case, James Hines, and Harvey Rosen (1993) find evidence of strategic interaction in state-level fiscal policies. Using a similar methodology, Jan Brueckner (1998) finds empirical support for policy interdependence in the adoption of growth-control measures by local governments in California. But at this juncture, I think it is fair to say that the jury is still out on this matter. The welfare implications of interjurisdictional competition remain the subject of a lively ongoing debate with a real need for further empirical work to supplement the large theoretical literature. In my own view, the existing work is not sufficient to make a compelling case for the abandonment of (or basic amendment to) the principle of fiscal decentralization. The case remains strong, it seems to me, for leaving "local matters in local hands." Moreover, as we shall see shortly, there is another literature that takes a very different (and unambiguously positive) view of the role of interjurisdictional competition.

## *7. Fiscal Federalism: Expanding the Scope of the Analysis*

The normative framework for most of the literature in fiscal federalism (and for my treatment in this essay as well) consists of the traditional principles of welfare economics. From this perspective, institutions are evaluated in terms of their impact on efficiency in resource allocation and the distribution of income. However, the choice of a system of governance involves other values as well: the extent of political participation, the protection of individual rights, and the development of various civic virtues. Political theorists throughout the ages have explored the ways in which different political systems address these various objectives of the polity. In addition, the vertical structure of government may have important implications for the way in which the public sector functions and its impact on the operation of a system of markets. In this section, I want to explore some of the new (and older) literature that addresses some broader implications of fiscal federalism.

### *7.1 Economic and Political Objectives in a Federal System*

The first issue involves extending the conceptual horizon to encompass additional political objectives. What might this add to our more narrowly focused economic view of fiscal federalism? Inman and Rubinfeld, in one strand of their important new work on fiscal federalism, have (and are) exploring this issue in an attempt to redefine and extend the analytical framework to encompass some of these additional political and constitutional dimensions of public-sector structure.

The approach of Inman and Rubinfeld (1997a,b,c) incorporates explicitly certain political goals into a more extended

objective function. In such a setting, we find ourselves examining tradeoffs between such goals as economic efficiency and political participation. In one such illustration, they present a "federalism frontier" in which (over the relevant range) increased political participation comes at the expense of economic efficiency (1997a, p. 1230).

The basic presumption here is that more decentralized political systems are conducive to increased citizen impact on political outcomes and political participation. The evidence on this issue, in truth, is somewhat mixed, but overall it suggests on balance "that both citizen influence and effort increase as the size of government declines" (1997a, p. 1215). The basic political objectives thus strengthen the case for increased decentralization; they point to a system that is more decentralized than one chosen simply on the grounds of an exercise in economic optimization.

While this is suggestive at a general level, it raises the more difficult question of how one addresses these tradeoffs in the actual design of fiscal institutions. How, for example, can we define and measure in a meaningful way the marginal rate of substitution between economic efficiency and political participation and incorporate this into the design of a political system? To approach this question in a substantive way requires the study of more specific issues. And here Inman and Rubinfeld (1997a) provide a provocative beginning with a careful study of "anti-trust state-action doctrine." This involves an intriguing series of Supreme Court decisions in which state programs, that—had they been designed and introduced by producers themselves, would have constituted a violation of anti-trust laws—were upheld on the basis of state legislative sovereignty. Although the history of this doctrine is a complicated one, it is in-

teresting that the Court has seen fit to set aside, in certain instances, the presumed economic consequences of certain state regulations in favor of decentralized political choices, so long as they "were decided by an open, participatory political process, as evidenced by state legislative involvement" (1997a, p. 1252).

It seems unlikely that we can ever hope to quantify such tradeoffs in a formally satisfying way. But the Inman-Rubinfeld work does suggest that careful analysis can certainly help to clarify the nature of the tradeoffs involved in the vertical design of the political system and allow economics to play a broader role in the debate. It is interesting, moreover, that the political objectives seem, on the whole, to strengthen the case for fiscal decentralization.

## 7.2 Public-Sector Institutions: *Market-Preserving Federalism*

An alternative approach to federalism, related to the "new institutional economics," sees political decentralization in terms of its capacity to sustain a productive and growing market economy. From this perspective, Barry Weingast (1995), Ronald McKinnon (1997a), and their colleagues have explored the institutional structure of a system that promises to provide a stable framework for a market system (see also McKinnon and Nechyba 1997 and Qian and Weingast 1997). Weingast's point of departure is a "fundamental political dilemma of an economic system," namely that "a government strong enough to protect property rights and enforce contracts is also strong enough to confiscate the wealth of its citizens" (1995, p. 1).<sup>28</sup>

The attraction of federalism for Weingast is its potential for providing a

<sup>28</sup> However, as Martin McGuire and Mancur Olson (1996) have shown, even a self-aggrandizing autocrat (if secure) has powerful incentives for supporting an economically efficient system.

political system that can support an efficient system of markets. In a provocative treatment, Weingast lays out a set of three conditions for a federal system that characterize what he calls "market-preserving federalism." These conditions require that (1) decentralized governments have the primary regulatory responsibility over the economy; (2) the system constitutes a common market in which there are no barriers to trade; and (3) decentralized governments face "hard budget constraints." By this last condition, Weingast means that lower-level governments have neither the capacity to create money nor access to unlimited credit. And it implies further that the central government does not stand ready to bail them out in instances of fiscal distress.

Weingast goes on to argue in historical terms that eighteenth century England and the United States in the nineteenth century were effectively such systems of market-preserving federalism, and that this fostered in important and fundamental ways the process of economic growth. It proved critical, argues Weingast, to the industrial revolution in England and supported a system of "thriving markets" in the United States throughout the nineteenth century.

McKinnon (1997a) has explored in more detail the importance of Weingast's last condition of a hard budget constraint. Crucial to this view is the separation of monetary and fiscal powers. In a federal system, if the central government controls the common currency, then lower-level governments will be limited to fiscal instruments and will not have access to the "soft" option of monetized debt. As McKinnon points out, state and local governments in the United States engage in extensive debt finance for capital projects. This makes good economic sense in terms of spreading the payments for long-lived

capital projects over their useful life. But they have no recourse to public sources for funding this debt; they operate in private credit markets just like private borrowers. These markets themselves, through the determination of credit ratings and other forms of monitoring fiscal performance, create an environment in which the fiscal authorities must behave in responsible ways.<sup>29</sup> These markets, by creating a hard budget constraint in terms of debt finance, have imposed a very useful discipline on decentralized fiscal behavior.<sup>30</sup>

More generally, a hard budget constraint implies that decentralized governments must place a basic reliance on their own sources of revenues. They must not be overly dependent on transfers from above. I discussed in an earlier section the potential role for intergovernmental grants, but Weingast and McKinnon (as well as others) remind us of the important discipline that stems from self-financing. It is especially important that intergovernmental grants not be expansible in the sense that recipients can turn to the grant system to bail them out of fiscal difficulties (Wildasin 1998b). In particular, public authorities need to fund their own expenditures at the margin.<sup>31</sup>

The institutional perspective reminds us that there is more to the design of a

<sup>29</sup> James Poterba and Kim Rueben (1997), for example, have found that those states with tighter anti-deficit rules, and more restrictive limitations on the authority of the state legislature to issue debt, pay lower rates of interest on their bonds.

<sup>30</sup> McKinnon (1997b) has gone on to argue that much of the impetus for European Monetary Union has as its source a collectively imposed budgetary retrenchment. His interesting argument is that European decision makers, realizing that they cannot achieve fiscal stability with continued access to monetary powers, are seeking through EMU to create the hard budget constraints that are the prerequisite for responsible fiscal management.

<sup>31</sup> This is subject to the qualification that matching grants may be needed to internalize interjurisdictional spillover benefits.

federal fiscal system than just the allocation of functions to the appropriate levels of government. In addition, we need sets of formal and informal institutions that embody the rights sorts of incentives for public decision makers (Olson 1990). These rules or procedures must make the costs of public programs as fully visible as their benefits in ways that make public officials accountable for their decisions (Shah 1998).

The treatment of fiscal structure in this section is not unrelated to Geoffrey Brennan and James Buchanan's (1980) view of fiscal decentralization as a mechanism for controlling the size of the public sector. Drawing by analogy on the conventional theory of monopoly in the private sector, they envision the government sector as a monolithic agent, a "Leviathan," that seeks its own aggrandizement through maximizing the extraction of tax revenues from the economy. From this perspective, the design of the constitution and associated institutions has as a major objective the placing of a set of constraints that limits Leviathan's access to tax and other fiscal instruments. Fiscal decentralization can, in their view, play a most important role in constraining public sector growth. Competition among decentralized governments for mobile economic units greatly limits the capacity of Leviathan to channel resources into the public sector. As Brennan and Buchanan put it, competition among governments in the context of the "interjurisdictional mobility of persons in pursuit of 'fiscal gains' can offer partial or possibly complete substitutes for explicit fiscal constraints on the taxing power" (1980, p. 184).<sup>32</sup>

<sup>32</sup> In a more formal treatment of this matter, Dennis Epple and Allan Zelenitz (1981) have shown that while competition among jurisdictions can constrain government rent-seeking behavior, it cannot altogether eliminate it.

The Brennan-Buchanan view suggests the hypothesis that the overall size of the public sector "should be smaller, *ceteris paribus*, the greater the extent to which taxes and expenditures are decentralized" (1980, p. 185). The evidence on this hypothesis is, however, at best mixed. For example, I was unable to find any systematic relationship between public-sector size and the extent of fiscal decentralization (Oates 1985). However, some later and more disaggregated studies have found some tendencies of this kind (See Oates 1989 for a survey of this work.).

More generally, there is not much evidence on the relationship between fiscal decentralization and economic performance. But there is some. Jeff Huther and Anwar Shah (1996) at the World Bank have assembled a large and diverse set of indices for eighty nations. These indices encompass a wide variety of measures of economic and political structure and performance: quality of governance, political freedom, political stability, debt-to-GNP ratios, measures of income, the degree of equality in the distribution of income, and many more. In examining the statistical associations among these various indices, they find in nearly every case a statistically significant and positive correlation between increased decentralization and improved performance (either in political or economic terms). There are obvious and important qualifications here. Such associations do not prove causation. In particular, the degree of fiscal decentralization is itself the outcome of a complex of political and economic forces. Nonetheless, the initial results are suggestive and invite further exploration. Elsewhere, Sang-Loh Kim (1995) in an intriguing econometric study making use of an international panel data set, has estimated a Barro-type growth model. In addition to the

usual explanatory variables, he included a measure of fiscal decentralization that, in most of his estimated equations, has a significant and positive partial association with the rate of economic growth. Kim's findings thus support Shah's contention that fiscal decentralization enhances economic performance—in this case, more rapid economic growth. In contrast, Heng-fu Zou and his colleagues have found a negative relationship between economic growth and fiscal decentralization in two studies, one examining a sample of forty-six countries over the period 1970–89 (Davoodi and Zou 1998) and the other a study of the growth of provinces in China (Zhang and Zou 1998). Much obviously remains to be done at the empirical level in order to give us a better sense of the relationship of fiscal decentralization to economic and political performance.

There is also much more to do at the conceptual level. While Weingast's initial forays into market-preserving federalism are certainly provocative, they raise at least as many questions as they answer. It is fair, I think, to characterize the analysis as fairly "loose" at this stage. For example, are Weingast's conditions for market-preserving federalism to be regarded as necessary or sufficient (or both) for an effective political foundation for a private market economy? Jonathan Rodden and Susan Rose-Ackerman (1997) have raised a number of probing questions concerning the Weingast analysis. There is clearly much to chew on here. The next step, it seems to me, is to attempt to formalize these relationships more explicitly so as to get a better sense of how different political and budgetary institutions influence the functioning of a market system.

Finally, it is impossible to leave this section without noting an obvious irony

that has no doubt occurred to the reader. In the earlier section on interjurisdictional competition, the central concern was that such competition leads to too little in the way of public outputs. There it was argued that competition for new firms and jobs may lead to public budgets that are too small, and to overly lax environmental standards. In contrast, the thrust of this section has been on the beneficial effects of competition as a disciplining force that restrains the tendencies in the public sector towards excessive spending and other forms of fiscal misbehavior. One's view of the role of intergovernmental competition clearly depends on how one views the operation of the public sector more generally!

#### *8. Fiscal Decentralization and Economic Development*

When examining international cross-sectional data on intergovernmental structure, one is immediately struck by the sharp contrast in the extent of fiscal decentralization in the industrialized and developing countries. In a study of my own involving a group of forty-three countries (Oates 1985), the sample statistics revealed an average share of central-government spending in total public expenditure of 65 percent in the subsample of eighteen industrialized countries, as contrasted to 89 percent in the subsample of twenty-five developing nations. In terms of total public revenues, the central-government share for this same subsample of developing countries was over 90 percent!

Although there are real concerns with the accuracy of some of these fiscal data (Richard Bird 1986), the general presumption that the developing countries are characterized by relatively high degrees of fiscal centralization seems firmly grounded. And this, moreover, is

not something new. Writing over forty years ago, Alison Martin and W. Arthur Lewis (1956) noted that "the weakness of local government in relation to central government is one of the most striking phenomena of under-developed countries" (p. 231).

What are we to make of this? Some observers attribute the poor economic performance of many of the developing countries in large measure to the failure of central planning and make a strong case for the devolution of fiscal responsibilities. But the issue is clearly more complicated than this. In particular, the question arises as to whether fiscal decentralization is a cause or a result of economic development. Roy Bahl and Johannes Linn (1992), for example, argue that as economies grow and mature, economic gains from fiscal decentralization emerge. As they put it, "Decentralization more likely comes with the achievement of a higher stage of economic development" (p. 391); the "threshold level of economic development" at which fiscal decentralization becomes attractive "appears to be quite high" (p. 393). From this perspective, it is economic development that comes first; fiscal decentralization then follows. But not all would agree. More generally, it seems to me, we must regard intergovernmental structure as part of a larger political and economic system that both influences and is determined by the interplay of a variety of political and economic forces. It may well be that fiscal decentralization itself has a real contribution to make to improved economic and political performance at different stages of development.

To gain further insight into this issue, we might turn to the historical experience of the industrialized countries and examine the course of fiscal decentralization through extended periods of economic growth. This, in fact, does not

prove to be very helpful. If we look at the United States, for example, we find that in the late nineteenth century the public sector was both very small and highly decentralized. At the turn of the century, the public sector accounted for only about 8 percent of GNP in the U.S., while the central-government share of total public expenditure was around 30–35 percent. By 1955, the central-government share of public spending had roughly doubled from one-third to two-thirds.<sup>33</sup> The fiscal records of other industrialized nations like Great Britain reveal roughly similar patterns.

The point is that the trend over this period of economic growth was *not* one of increasing fiscal decentralization; it was just the reverse! It is worth noting, however, that these centralizing tendencies seem to have played out around the middle of the century. For most of the industrialized countries, fiscal centralization ratios appear to have peaked in the decade of the 1950's, and since that time, they have actually declined slightly in most cases (Oates 1978; Werner Pommerehne 1977). What typically seems to be taking place is a complicated process of intergovernmental evolution. We see efforts at devolution in a number of OECD countries accompanied, at the same time, by the emergence of a new top layer of government in the European Community.

But all this may not have much relevance for the developing nations. This is because they have a very different starting point for the growth process. As Diana Conyers (1990) stresses, "Most less developed countries inherited relatively centralized systems of governments from their colonial powers, and in the first years of independence there

<sup>33</sup> See John Wallis and Oates (1997) for a description and analysis of the evolution of American federalism in the twentieth century.

was often a tendency to maintain—if not strengthen—central control and centralized systems of planning, in order to encourage a sense of national unity and reinforce the new government and its policies” (p. 16). Thus, many of these countries entered upon nationhood with highly centralized government sectors; they have not undergone anything like the process of public-sector evolution experienced in the industrialized countries.

The implication of all this is that the potential of fiscal decentralization for improving economic and political performance must be evaluated in terms of the specific circumstances that characterize the current state of a developing nation. There remains, in my view and that of some others (Shah 1994), a strong case on traditional grounds for a significant degree of decentralization in public-sector decision-making in the developing nations. This case, as we have discussed, rests both on the potential economic gains from adapting levels of public outputs to specific regional or local conditions and on the political appeal of increased participation in governance. The economic case has been made formally in purely static terms (as noted earlier in the treatment of the Decentralization Theorem), but it may well have some validity in a dynamic setting of economic growth. Development policies that are sensitive to particular regional or local needs for infrastructure and even human capital are likely to be more effective in promoting economic growth than are centrally determined policies that largely ignore these geographical differences. There exists, incidentally, no formal theory of fiscal decentralization and economic growth; it might be useful to set out such a theory, for a framework that incorporates jurisdiction-specific investment programs might provide some insights into the parameters on which

improved growth performance depends.<sup>34</sup>

The prescriptive literature on fiscal structure for the developing countries harks back directly to several of the points made in the preceding sections. In particular, there is a heavy emphasis on reliance on own finance in order to create hard budget constraints. This can have special relevance in the developing-country context, where decentralized governments often have very limited access to their own major sources of tax and other revenues and are heavily dependent on transfers from above. In some instances, provincial or state governments may even have access to the public banking system to absorb their debt issues. This predictably leads to large budgetary deficits and both fiscal and monetary instability.

This literature makes reference to the problem of “vertical imbalance,” meaning a disparity between different levels of government in their expenditure commitments and their access to revenues. Although the concept suffers from certain ambiguities, it does focus attention on the important issue of the widespread inadequacy of revenue sources at decentralized levels of government. The often heavy reliance of provincial, state, and local governments on transfers from above undercuts incentives for responsible fiscal decision-making; fiscal decisions become outcomes of politically driven negotiations between central and “local” authorities, not the result of weighing benefits and costs of prospective public programs.

The case for establishing adequate

<sup>34</sup> Some observers, like Remy Prud'homme (1997), argue that the case for fiscal decentralization has been much exaggerated. Prud'homme claims that many of the premises of the fiscal federalism vision are typically not satisfied in the developing-country setting; decentralized government bodies, he argues, are frequently unresponsive to the needs of their constituencies and manifest widespread corruption.

and effective tax systems at decentralized levels of government is one of the critical issues of fiscal federalism in the developing world. And it is a truly challenging problem (Bahl and Linn 1992; Bird 1992). The earlier section dealing with the tax-assignment problem set forth some of the properties of "good" taxes at decentralized levels of government. But provincial and local governments in developing countries often face serious obstacles to the use of these tax bases. The scope, for example, for using local property taxes is circumscribed in many instances by the absence of the requisite institutions for tax administration. As Bahl and Linn (1992) point out, there is typically more potential for such taxes in urban than in rural areas in most developing countries. The obstacles are real, but there are ongoing and extensive efforts to build up the administrative capacity for more effective revenue systems.

Fiscal reform efforts in the developing world thus must focus on (1) Restructuring systems of intergovernmental grants, in some instances to reduce the extent of financing that they provide to decentralized levels of government, and, more generally, to remove the perverse incentives that they often embody for fiscal behavior on the part of recipients; (2) Redesigning revenue systems so as to provide decentralized levels of government a much expanded access to own-revenues to finance their budgets and thereby reduce their dependence on transfers from above; and (3) Reviewing the use and restrictions on debt finance to ensure that debt issues are not a ready way to finance deficits on the current account. All three of these avenues of reform contribute in important ways to the establishment of a hard budget constraint, but one that permits decentralized levels of government to do their job. Finally, running

through all these dimensions of fiscal reform is the crucial attention to fiscal decision-making institutions and procedures themselves to introduce mechanisms that provide incentives for public officials to act in the public interest; this means largely, as Shah (1998) stresses, establishing channels for accountability.<sup>35</sup> In the interim, provincial and local governments cannot be left to fend entirely for themselves; depending on the specific circumstances, there will often be a need for significant transfers from the center, especially to impoverished jurisdictions. But the general direction of needed reform seems clear.

The ongoing efforts to decentralize the public sectors of former socialist states encounter much the same set of issues. But the problems are in some ways even more complicated, inasmuch as the process of decentralization is going on alongside a process of privatization; the complicated and sometimes chaotic transition from a command economy to a market system does not provide a stable environment within which to restructure the public sector. Nevertheless, a comprehensive process of fiscal decentralization is underway in much of Central and Eastern Europe, and it involves the same issues of defining the fiscal responsibilities of the different levels of government and introducing the fiscal instruments and procedures needed both to support emerging private markets and to deliver needed public services (Bird, Ebel, and Wallich 1995).

#### *9. Some Concluding Observations*

The evolution of the vertical structure of the public sector continues in

<sup>35</sup> See Govinda Rao (1998) for an illuminating treatment in the Indian context of the wide range of mechanisms (or "subterranean transfers" as he calls them) through which central government subsidizes the states.

interesting and novel ways. As I noted earlier, the first half of the twentieth century was characterized by a strong trend toward increased fiscal centralization. Indeed, some acute political observers in the nineteenth century forecast this trend. Tocqueville, writing in the first half of the nineteenth century, predicted that "in the democratic ages which are opening upon us . . . centralization will be the natural government" (1945, Vol. II, p. 313). And nearer the end of the century, Lord Bryce reiterated this forecast (at least for the U.S.). After reviewing both the "centrifugal" and "centripetal" forces at work in American government, Bryce concluded that while the centrifugal forces were "likely, as far as we can see, to prove transitory . . . the centripetal forces are permanent and secular forces, working from age to age" (1901, Vol. II, p. 844). Bryce then proceeded to forecast that ". . . the importance of the States will decline as the majesty and authority of the National government increase" (1901, Vol. II, p. 844). Later, Edward McWhinney (1965) went on to generalize all this to what he calls "Bryce's Law," the proposition that ". . . federalism is simply a transitory step on the way to governmental unity" (p. 105).

But such forecasts have not been borne out. The second half of the twentieth century has seen the extent of centralization in most of the industrialized countries reach some sort of peak with a modest swing back in the direction of devolution of public sector activity. There are, as Bryce suggests, important forces working in both directions, and one can expect the net effect to move in different directions as nations evolve over time.

What does seem to be taking place is a growing complexity and specialization in the vertical structure of the public

sector. Recent decades have seen the creation of special districts to provide particular public services and the formation of metropolitan area governments to bring center cities and their suburbs into a single jurisdiction (again for purposes of addressing specific needs such as transportation and housing). It is especially striking to witness in the European Community the moves toward devolution in many member countries, while, at the same time, the Community develops a set of supranational institutions for governance and economic management. Other countries, like South Africa and the former socialist states, are struggling with their own sets of pressing issues in their attempts to find effective mechanisms for political and fiscal decentralization.

While the existing literature in fiscal federalism can provide some general guidance on these issues, my sense is that most of us working in the field feel more than a little uneasy when proffering advice on many of the decisions that must be made on vertical fiscal and political structure. We have much to learn!

#### REFERENCES

- Ackerman, Bruce et al. 1974. *The Uncertain Search for Environmental Quality*. NY: Free Press.
- Anton, Thomas J. 1989. *American Federalism and Public Policy*. Philadelphia: Temple U. Press.
- Bahl, Roy W. and Johannes F. Linn. 1992. *Urban Public Finance in Developing Countries*. Oxford: Oxford U. Press.
- Bartik, Timothy J. 1991. *Who Benefits from State and Local Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute.
- Bird, Richard M. 1986. "On Measuring Fiscal Centralization and Fiscal Balance in Federal States," *Gov. Pol.*, 4, pp. 384-404.
- . 1992. *Tax Policy and Economic Development*. Baltimore: Johns Hopkins Press.
- Bird, Richard M.; Robert D. Ebel, and Christine I. Wallich, eds. 1995. *Decentralization of the Socialist State*. Washington, DC: World Bank.
- Blank, Rebecca. 1997. "Policy Watch: The 1996 Welfare Reform," *J. Econ. Perspect.*, 11:1, pp. 169-78.

- Broadway, Robin. 1996. "Review of 'The Uneasy Case for Equalization Payments' by Dan Usher," *Nat. Tax J.*, 49:4, pp. 677-86.
- Broadway, Robin and Frank R. Flatters. 1982. "Efficiency and Equalization Payments in a Federal System of Government: A Synthesis and Extension of Recent Results," *Can. J. Econ.*, 15:4, pp. 613-33.
- Boeckelman, Keith. 1992. "The Influence of States on Federal Policy Adoptions," *Pol. Stud. J.*, 20:3, pp. 365-75.
- Bradford, David F. and Wallace E. Oates. 1971a. "Towards a Predictive Theory of Intergovernmental Grants," *Amer. Econ. Rev.*, 61:2, pp. 440-48.
- \_\_\_\_\_. 1971b. "The Analysis of Revenue Sharing in a New Approach to Collective Fiscal Decisions," *Quart. J. Econ.*, 85:3, pp. 416-39.
- \_\_\_\_\_. 1974. "Suburban Exploitation of Central Cities and Governmental Structure," in *Redistribution Through Public Choice*. Harold Hochman and George Peterson, eds. NY: Columbia U. Press, pp. 43-90.
- Break, George F. 1967. *Intergovernmental Fiscal Relations in the United States*. Washington, DC: Brookings Institution.
- Brennan, Geoffrey and James Buchanan. 1980. *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. Cambridge: Cambridge U. Press.
- Breton, Albert. 1998. *Competitive Governments: An Economic Theory of Politics and Public Finance*. Cambridge: Cambridge U. Press.
- Brown, Charles C. and Wallace E. Oates. 1987. "Assistance to the Poor in a Federal System," *J. Pub. Econ.*, 32, pp. 307-30.
- Brueckner, Jan E. 1998a. "Testing for Strategic Interaction Among Local Governments: The Case of Growth Controls," *J. Urban Econ.*, 44, pp. 438-67.
- \_\_\_\_\_. 1998b. "Welfare Reform and the Race to the Bottom: Theory and Evidence," Institute of Gov. and Public Affairs, U. Illinois, WP 64.
- Bryce, James. 1901. *The American Commonwealth*. London: Macmillan [first published 1888].
- Case, Anne; James R. Hines, Jr., and Harvey S. Rosen. 1993. "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States," *J. Pub. Econ.*, 52, pp. 285-307.
- Chernick, Howard. 1979. "An Economic Model of the Distribution of Project Grants," in *Fiscal Federalism and Grants-in-Aid*. P. Mieszkowski and W. Oakland, eds. Washington DC: Urban Institute, pp. 81-103.
- \_\_\_\_\_. 1992. "A Model of the Distributional Incidence of State and Local Taxes," *Pub. Fin. Quart.*, Oct., pp. 572-85.
- Conyers, Diana. 1990. "Centralization and Development Planning: A Comparative Perspective," in *Decentralizing for Participatory Planning*. P. de Valk and K. Wekwete, eds. Aldershot: Avebury.
- Courant, Paul N. 1994. "How Would You Know a Good Economic Policy If You Tripped Over One?" *Nat. Tax J.*, 47, pp. 863-81.
- Cumberland, John H. 1979. "Interregional Pollution Spillovers and Consistency of Environmental Policy," in *Regional Environmental Policy: The Economic Issues*. H. Siebert et al., eds. NY: NYU Press, pp. 255-81.
- \_\_\_\_\_. 1981. "Efficiency and Equity in Interregional Environmental Management," *Rev. Regional Stud.*, 2, pp. 1-9.
- Davoodi, Hamid and Heng-fu Zou. 1998. "Fiscal Decentralization and Economic Growth," *J. Urban Econ.*, 43, pp. 244-57.
- Enrich, Peter D. 1996. "Saving the States from Themselves: Commerce Clause Constraints on State Tax Incentives for Business," *Harvard Law Rev.*, 110, pp. 378-461.
- Apple, Dennis and Allan Zelenitz. 1981. "The Implications of Competition Among Jurisdictions: Does Tiebout Need Politics?" *J. Polit. Econ.*, 89, pp. 1197-217.
- Esty, Daniel C. 1996. "Revitalizing Environmental Federalism," *Mich. Law Rev.*, 95, pp. 570-653.
- Feldstein, Martin S. 1975. "Wealth Neutrality and Local Choice in Public Education," *Amer. Econ. Rev.*, 65:1, pp. 75-89.
- Feldstein, Martin and Marian Vaillant Wrobel. 1998. "Can State Taxes Redistribute Income?" *J. Pub. Econ.*, 68:3, pp. 369-96.
- Fischel, William. 1992. "Property Taxation and the Tiebout Model: Evidence for the Benefit View from Zoning and Voting," *J. Econ. Lit.*, 30:1, pp. 171-77.
- Fisher, Ronald C. 1982. "Income and Grant Effects on Local Expenditure: The Flypaper Effect and Other Difficulties," *J. Urban Econ.*, 12, pp. 324-45.
- Flatters, Frank; Vernon Henderson, and Peter Mieszkowski. 1974. "Public Goods, Efficiency, and Regional Fiscal Equalization," *J. Pub. Econ.*, 3, pp. 99-112.
- Gamkhar, Shama and Wallace Oates. 1966. "Asymmetries in the Response to Increases and Decreases in Intergovernmental Grants: Some Empirical Findings," *Nat. Tax J.*, 49, pp. 501-12.
- Gordon, Roger. 1983. "An Optimal Tax Approach to Fiscal Federalism," *Quart. J. Econ.*, 97, pp. 567-86.
- Gramlich, Edward M. 1977. "Intergovernmental Grants: A Review of the Empirical Literature," in *The Political Economy of Fiscal Federalism*. Wallace Oates, ed. Lexington, MA: Heath-Lexington, pp. 219-39.
- \_\_\_\_\_. 1987. "Federalism and Federal Deficit Reduction," *Nat. Tax J.*, 40, pp. 299-313.
- Gray, Virginia. 1973. "Innovation in the States: A Diffusion Study," *Amer. Polit. Sci. Rev.*, 67, pp. 1174-85.
- Hamilton, Bruce W. 1975. "Zoning and Property Taxation in a System of Local Governments," *Urban Stud.*, 12:2, pp. 205-11.
- \_\_\_\_\_. 1976. "Capitalization of Intrajurisdictional Differences in Local Tax Prices," *Amer. Econ. Rev.*, 66:5, pp. 743-53.

- Hines, James R. Jr., and Richard H. Thaler. 1995. "The Flypaper Effect," *J. Econ. Perspect.*, 9, pp. 217-26.
- Huff, David L.; James M. Lutz, and Rajendra Srivastava. 1988. "A Geographical Analysis of the Innovativeness of States," *Econ. Geogr.*, 64, pp. 137-46.
- Huther, Jeff and Anwar Shah. 1996. "A Simple Measure of Good Governance and Its Application to the Debate on the Appropriate Level of Fiscal Decentralization." Washington, DC: World Bank.
- Inman, Robert P. 1988. "Federal Assistance and Local Services in the United States: The Evolution of a New Federalist Fiscal Order," in *Fiscal Federalism*. Harvey Rosen, ed. Chicago: U. Chicago Press, pp. 33-74.
- \_\_\_\_\_. Forthcoming. "On Designing Intergovernmental Transfers With an Application in the New South Africa," in *Environmental and Public Economics*. Arvind Panagariya, Paul Portney, and Robert Schwab, eds. Cheltenham, UK: Edward Elgar.
- Inman, Robert P., and Daniel L. Rubinfeld. 1979. "The Judicial Pursuit of Local Fiscal Equity," *Harvard Law Rev.*, 92, pp. 1662-750.
- \_\_\_\_\_. 1996. "Designing Tax Policy in Federalist Economies: An Overview," *J. Pub. Econ.*, 60:3, pp. 307-34.
- \_\_\_\_\_. 1997a. "Making Sense of the Antitrust State-Action Doctrine: Balancing Political Participation and Economic Efficiency in Regulatory Federalism," *Texas Law Rev.*, 75, pp. 1203-99.
- \_\_\_\_\_. 1997b. "The Political Economy of Federalism," in *Perspectives on Public Choice: A Handbook*. D. Mueller, ed. Cambridge: Cambridge U. Press, pp. 73-105.
- \_\_\_\_\_. 1997c. "Rethinking Federalism," *J. Econ. Perspect.*, 11:4, pp. 43-64.
- \_\_\_\_\_. Forthcoming. "Subsidiarity and the European Union," in *The New Palgrave Dictionary of Economics and Law*.
- Kim, Sang Loh. 1995. *Fiscal Decentralization, Fiscal Structure, and Economic Performance: Three Empirical Studies*. Unpublished Ph.D. Dissertation, U. Maryland.
- Lutz, James M. 1987. "Regional Leadership Patterns in the Diffusion of Public Policies," *Amer. Polit. Quart.*, 15, pp. 387-98.
- Martin, Alison and W. Arthur Lewis. 1956. "Patterns of Public Revenue and Expenditure," *Manchester Sch. Econ. Soc. Stud.*, 24, pp. 203-44.
- McGuire, Martin C. and Mancur Olson, Jr. 1996. "The Economics of Autocracy and Majority Rule: The Invisible Hand and the Use of Force," *J. Econ. Lit.*, 34, pp. 72-96.
- McGuire, Therese J. 1997. "Intergovernmental Fiscal Relations and Social Welfare Policy," in *Intergovernmental Fiscal Relations*. Ronald C. Fisher, ed. Boston: Kluwer, pp. 173-98.
- McKinnon, Ronald I. 1997a. "Market-Preserving Fiscal Federalism in the American Monetary Union," in *Macroeconomic Dimensions of Public Finance: Essays in Honor of Vito Tanzi*. Mario Blejer and Teresa Ter-Minassian, eds. London: Routledge, pp. 73-93.
- \_\_\_\_\_. 1997b. "EMU as a Device for Collective Fiscal Retrenchment," *Amer. Econ. Rev.*, 87, pp. 227-29.
- McKinnon, Ronald I. and Thomas Neehyba. 1997. "Competition in Federal Systems: The Role of Political and Financial Constraints," in *The New Federalism: Can the States be Trusted?* John Ferejohn and Barry Weingast, eds. Stanford: Hoover Institution Press, pp. 3-61.
- McLure, Charles E., Jr. ed. 1983. *Tax Assignment in Federal Countries*. Canberra: Australian National U.
- McWhinney, Edward. 1965. *Comparative Federalism*. 2nd ed. Toronto: U. Toronto Press.
- Mieszkowski, Peter, and George R. Zodrow. 1989. "Taxation and the Tiebout Model: The Differential Effects of Head Taxes, Taxes on Land, Rents, and Property Taxes," *J. Econ. Lit.*, 27:3, pp. 1098-146.
- Musgrave, Richard M. 1959. *The Theory of Public Finance*. NY: McGraw-Hill.
- Neehyba, Thomas. 1996. "A Computable General Equilibrium Model of Intergovernmental Aid," *J. Pub. Econ.*, 62, pp. 363-99.
- Nice, David C. 1994. *Policy Innovation in State Government*. IA: Iowa U. Press.
- Oates, Wallace E. 1972. *Fiscal Federalism*. NY: Harcourt Brace Jovanovich.
- \_\_\_\_\_. 1978. "The Changing Structure of Intergovernmental Fiscal Relations," in *Secular Trends of the Public Sector*. H. Recktenwald, ed. Paris: Editions Cujas, pp. 151-60.
- \_\_\_\_\_. 1985. "Searching for Leviathan: An Empirical Study," *Amer. Econ. Rev.*, 75, pp. 748-57.
- \_\_\_\_\_. 1989. "Searching for Leviathan: A Reply and Some Further Reflections," *Amer. Econ. Rev.*, 79, pp. 578-83.
- \_\_\_\_\_. 1994. "Federalism and Government Finance," in *Modern Public Finance*. J. Quigley and E. Smolensky, eds. Cambridge, MA: Harvard U. Press, pp. 126-51.
- \_\_\_\_\_. 1996a. "Estimating the Demand for Public Goods: The Collective Choice and Contingent Valuation Approaches," in *The Contingent Valuation of Environmental Resources*. D. Bjornstad and J. Kahn, eds. Aldershot: Edward Elgar, pp. 211-30.
- \_\_\_\_\_. 1996b. "Taxation in a Federal System: The Tax-Assignment Problem," *Pub. Econ. Rev.*, 1, pp. 35-60.
- \_\_\_\_\_. 1998. "On the Welfare Gains from Fiscal Decentralization," U. Maryland Econ. Dept. WP 98-05.
- Oates, Wallace E. and Robert M. Schwab. 1988. "Economic Competition Among Jurisdictions: Efficiency-Enhancing or Distortion-Inducing?" *J. Pub. Econ.*, 35, pp. 333-54.
- \_\_\_\_\_. 1991. "The Allocative and Distributive Implications of Local Fiscal Competition," in

- Competition Among States and Local Governments.* D. Kenyon and J. Kincaid, eds. Washington, DC: Urban Institute, pp. 127–45.
- . 1996. "The Theory of Regulatory Federalism: The Case of Environmental Management," in *The Economics of Environmental Regulation*. Wallace Oates. Aldershot: Edward Elgar, pp. 319–31.
- . 1997. "The Impact of Urban Land Taxation: The Pittsburgh Experience," *Nat. Tax J.*, 50, pp. 1–21.
- Olson, Mancur Jr. 1990. "The IRIS Idea." IRIS, U. Maryland.
- Osborne, David. 1988. *Laboratories of Democracy*. Boston: Harvard Business School Press.
- Pauly, Mark V. 1973. "Income Redistribution as a Local Public Good," *J. Pub. Econ.*, 2, pp. 35–58.
- Pommerehne, Werner W. 1977. "Quantitative Aspects of Federalism: A Study of Six Countries," in *The Political Economy of Fiscal Federalism*. Wallace Oates, ed. Lexington, MA: Heath-Lexington, pp. 275–355.
- Poterba, James M. and Kim S. Rueben. 1997. "State Fiscal Institutions and the U.S. Municipal Bond Market," NBER WP 6237.
- Prud'homme, Remy. 1995. "On the Dangers of Decentralization," *World Bank Res. Observer*, pp. 201–10.
- Qian, Yingyi, and Barry R. Weingast. 1997. "Federalism as a Commitment to Preserving Market Incentives," *J. Econ. Perspect.*, 11:4, pp. 83–92.
- Rao, M. Govinda. 1998. "Invisible Transfers in Indian Federalism," Australian National U. Econ. Dept. WP 98/1.
- Revesz, Richard L. 1992. "Rehabilitating Interstate Competition: Rethinking the 'Race-to-the-Bottom' Rationale for Federal Environmental Regulation," *NYU Law Rev.*, 67, pp. 1210–54.
- Rivlin, Alice. 1992. *Reviving the American Dream*. Washington, DC: The Brookings Institution.
- Rodden, Jonathan, and Susan Rose-Ackerman. 1997. "Does Federalism Preserve Markets?" *U. Virginia Law Rev.*, 83, pp. 1521–72.
- Rogers, Everett. 1983. *Diffusion of Innovations*. 3rd ed. NY: Free Press.
- Rose-Ackerman, Susan. 1980. "Risk Taking and Reelection: Does Federalism Promote Innovation?" *J. Legal Stud.*, 9, pp. 593–616.
- Rubinfeld, Daniel L. 1987. "The Economics of the Local Public Sector," in *Handbook of Public Economics*, Vol. II. Alan Auerbach and Martin Feldstein, eds. Amsterdam: North-Holland, pp. 571–645.
- Saavedra, Luz Amparo. 1998. "A Model of Welfare Competition with Evidence from AFDC," Inst. Govt. and Public Affairs, U. Illinois, WP 63.
- Shah, Anwar. 1994. *The Reform of Intergovernmental Fiscal Relations in Developing and Emerging Market Economies*. Washington, DC: World Bank.
- . 1998. "Fostering Fiscally Responsive and Accountable Governance: Lessons from Decentralization," in *Evaluation and Development: The Institutional Dimension*. Robert Picciotto and Eduardo Wiesner, eds. Washington, DC: World Bank, pp. 83–96.
- Sparer, Michael S., and Lawrence D. Brown. 1996. "States and the Health Care Crisis: The Limits and Lessons of Laboratory Federalism," in *Health Policy, Federalism, and the American States*. Robert F. Rich and William D. White, eds. Washington, DC: Urban Institute Press, pp. 181–202.
- Stine, William F. 1994. "Is Local Government Revenue Response to Federal Aid Symmetrical?" *Nat. Tax J.*, 47, pp. 799–816.
- Strumpf, Koleman S. 1997. "Does Fiscal Decentralization Increase Policy Innovation?" unpublished paper.
- Strumpf, Koleman S. and Felix Oberholzer-Gee. 1998. "Endogenous Institutions and Policy Decentralization: Local Liquor Control from 1934–1970"; unpublished paper.
- Tiebout, Charles. 1956. "A Pure Theory of Local Expenditures," *J. Polit. Econ.*, 64, pp. 416–24.
- Tocqueville, Alexis de. 1945. *Democracy in America*. NY: Vintage Books Random House; first published 1838.
- US Congressional Budget Office. 1997. *Federalism and Environmental Protection: Case Studies for Drinking Water and Ground-Level Ozone*. Washington, DC: US Govt. Printing Office.
- Usher, Dan. 1995. *The Uneasy Case for Equalization Payments*. Vancouver, BC: The Fraser Institute.
- Volden, Craig. 1997. "Entrusting the States with Welfare Reform," in *The New Federalism: Can the States Be Trusted?* John Ferejohn and Barry Weingast, eds. Stanford: Hoover Institution Press, pp. 65–96.
- Walker, Jack L. 1969. "The Diffusion of Innovation Among the American States," *Amer. Polit. Sci. Rev.*, 63, pp. 880–99.
- Wallis, John J. and Wallace E. Oates. 1998. "The Impact of the New Deal on American Federalism," in *The Defining Moment: The Great Depression and the American Economy in the 20th Century*. M. Bordo, C. Goldin, and E. White, eds. Chicago: U. Chicago Press, pp. 155–80.
- Weingast, Barry R. 1995. "The Economic Role of Political Institutions: Market-Preserving Federalism and Economic Development," *J. Law Econ. Org.*, 11, pp. 1–31.
- Wildasin, David E. 1988. "Nash Equilibria in Models of Fiscal Competition," *J. Pub. Econ.*, 35, pp. 229–40.
- . 1998a. "Factor Mobility and Redistributive Policy: Local and International Perspectives," in *Public Finance in a Changing World*. Peter B. Sorensen, ed. London: MacMillan, pp. 151–92.
- . 1998b. "Externalities and Bailouts: Hard and Soft Budget Constraints in Intergovernmental Fiscal Relations," World Bank Policy Research WP 1843.

- Wilson, John Douglas. 1996. "Capital Mobility and Environmental Standards: Is There a Theoretical Basis for a Race to the Bottom?" in *Fair Trade and Harmonization: Prerequisites for Free Trade?* Vol. I. Jagdish Bhagwati and Robert Hudec, eds. Cambridge, MA: MIT Press, pp. 393–427.
- Zhang, Tao and Heng-fu Zou. 1998. "Fiscal Decentralization, Public Spending, and Economic Growth in China," *J. Pub. Econ.*, 67, pp. 221–40.

A Pure Theory of Local Expenditures

Author(s): Charles M. Tiebout

Source: *Journal of Political Economy*, Oct., 1956, Vol. 64, No. 5 (Oct., 1956), pp. 416-424

Published by: The University of Chicago Press

Stable URL: <https://www.jstor.org/stable/1826343>

#### REFERENCES

~~Linked references are available on JSTOR for this article:~~

~~[https://www.jstor.org/stable/1826343?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/1826343?seq=1&cid=pdf-reference#references_tab_contents)~~

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Journal of Political Economy*

JSTOR

# A PURE THEORY OF LOCAL EXPENDITURES<sup>1</sup>

CHARLES M. TIEBOUT

Northwestern University

ONE of the most important recent developments in the area of "applied economic theory" has been the work of Musgrave and Samuelson in public finance theory.<sup>2</sup> The two writers agree on what is probably the major point under investigation, namely, that no "market type" solution exists to determine the level of expenditures on public goods. Seemingly, we are faced with the problem of having a rather large portion of our national income allocated in a "non-optimal" way when compared with the private sector.

This discussion will show that the Musgrave-Samuelson analysis, which is valid for federal expenditures, need not apply to local expenditures. The plan of the discussion is first to restate the assumptions made by Musgrave and Samuelson and the central problems with which they deal. After looking at a key difference between the federal versus local cases, I shall present a simple model. This model yields a solution for the level of expenditures for local public

goods which reflects the preferences of the population more adequately than they can be reflected at the national level. The assumptions of the model will then be relaxed to see what implications are involved. Finally, policy considerations will be discussed.

## THE THEORETICAL ISSUE

Samuelson has defined public goods as "*collective consumption goods* ( $X_n + 1, \dots, X_n + n$ ) which all enjoy in common in the sense that each individual's consumption of such a good leads to no subtraction from any other individual's consumption of that good, so that  $X_n + j = X_n^i + j$  simultaneously for each and every  $i$ th individual and each collective good."<sup>3</sup> While definitions are a matter of choice, it is worth noting that "consumption" has a much broader meaning here than in the usual sense of the term. Not only does it imply that the act of consumption by one person does not diminish the opportunities for consumption by another but it also allows this consumption to be in another form. For example, while the residents of a new government housing project are made better off, benefits also accrue to other residents of the community in the form of the external economies of slum clearance.<sup>4</sup> Thus many goods that appear to lack the attributes of public goods may

<sup>1</sup> I am grateful for the comments of my colleagues Karl de Schweinitz, Robert Eisner, and Robert Strotz, and those of Martin Bailey, of the University of Chicago.

<sup>2</sup> Richard A. Musgrave, "The Voluntary Exchange Theory of Public Economy," *Quarterly Journal of Economics*, LII (February, 1939), 213-17; "A Multiple Theory of the Budget," paper read at the Econometric Society annual meeting (December, 1955); and his forthcoming book, *The Theory of Public Economy*; Paul A. Samuelson, "The Pure Theory of Public Expenditures," *Review of Economics and Statistics*, XXXVI, No. 4 (November, 1954), 387-89, and "Diagrammatic Exposition of a Pure Theory of Public Expenditures," *ibid.*, XXXVII, No. 4 (November, 1955), 350-56.

<sup>3</sup> "The Pure Theory . . .," *op. cit.*, p. 387.

<sup>4</sup> Samuelson allows for this when he states that "one man's circus may be another man's poison," referring, of course, to public goods ("Diagrammatic Exposition . . .," *op. cit.*, p. 351).

properly be considered public if consumption is defined to include these external economies.<sup>5</sup>

A definition alternative to Samuelson's might be simply that a public good is one which should be produced, but for which there is no feasible method of charging the consumers. This is less elegant, but has the advantage that it allows for the objections of Enke and Margolis.<sup>6</sup> This definition, unfortunately, does not remove any of the problems faced by Musgrave and Samuelson.

The core problem with which both Musgrave and Samuelson deal concerns

<sup>5</sup> There seems to be a problem connected with the external-economies aspect of public goods. Surely a radio broadcast, like national defense, has the attribute that A's enjoyment leaves B no worse off; yet this does not imply that broadcasting should, in a normative sense, be a public good (the arbitrary manner in which the level of radio programs is determined aside). The difference between defense and broadcasting is subtle but important. In both cases there is a problem of determining the optimal level of outputs and the corresponding level of benefits taxes. In the broadcasting case, however, A may be quite willing to pay more taxes than B, even if both have the same "ability to pay" (assuming that the benefits are determinate). Defense is another question. Here A is not content that B should pay less. A makes the *social judgment* that B's preference *should* be the same. A's preference, expressed as an annual defense expenditure such as \$42.7 billion and representing the majority view, thus determines the level of defense. Here the A's may feel that the B's *should pay* the same amount of benefits tax.

If it is argued that this case is typical of public goods, then, once the level is somehow set, the voluntary exchange approach and the benefit theory associated with it do not make sense. If the preceding analysis is correct, we are now back in the area of equity in terms of ability to pay.

<sup>6</sup> They argue that, for most of the goods supplied by governments, increased use by some consumer-voters leaves less available for other consumer-voters. Crowded highways and schools, as contrasted with national defense, may be cited as examples (see Stephen Enke, "More on the Misuse of Mathematics in Economics: A Rejoinder," *Review of Economics and Statistics*, XXXVII [May, 1955], 131-33; and Julius Margolis, "A Comment on the Pure Theory of Public Expenditure," *Review of Economics and Statistics*, XXXVII [November, 1955], 247-49).

the mechanism by which consumer-voters register their preferences for public goods. The consumer is, in a sense, surrounded by a government whose objective it is to ascertain his wants for public goods and tax him accordingly. To use Alchian's term, the government's revenue-expenditure pattern for goods and services is expected to "adapt to" consumers' preferences.<sup>7</sup> Both Musgrave and Samuelson have shown that, in the vertically additive nature of voluntary demand curves, this problem has only a conceptual solution. If all consumer-voters could somehow be forced to reveal their true preferences for public goods, then the amount of such goods to be produced and the appropriate benefits tax could be determined.<sup>8</sup> As things now stand, there is no mechanism to force the consumer-voter to state his true preferences; in fact, the "rational" consumer will underestimate his preferences and hope to enjoy the goods while avoiding the tax.

The current method of solving this problem operates, unsatisfactorily, through the political mechanism. The expenditure wants of a "typical voter" are somehow pictured. This objective on the expenditure side is then combined with an ability-to-pay principle on the revenue side, giving us our current budget. Yet in terms of a satisfactory theory of public finance, it would be desirable (1) to force the voter to reveal his preferences; (2) to be able to satisfy them in

<sup>7</sup> Armen A. Alchian, "Uncertainty, Evolution, and Economic Theory," *Journal of Political Economy*, LVIII (June, 1950), 211-21.

<sup>8</sup> The term "benefits tax" is used in contrast to the concept of taxation based on the "ability to pay," which really reduces to a notion that there is some "proper" distribution of income. Conceptually, this issue is separate from the problem of providing public goods and services (see Musgrave, "A Multiple Theory . . .," *op. cit.*).

the same sense that a private goods market does; and (3) to tax him accordingly. The question arises whether there is any set of social institutions by which this goal can be approximated.

#### LOCAL EXPENDITURES

Musgrave and Samuelson implicitly assume that expenditures are handled at the central government level. However, the provision of such governmental services as police and fire protection, education, hospitals, and courts does not necessarily involve federal activity.<sup>9</sup> Many of these goods are provided by local governments. It is worthwhile to look briefly at the magnitude of these expenditures.<sup>10</sup>

Historically, local expenditures have exceeded those of the federal government. The thirties were the first peace-time years in which federal expenditures began to pull away from local expenditures. Even during the fiscal year 1954, federal expenditures on *goods and services exclusive of defense* amounted only to some 15 billions of dollars, while local expenditures during this same period amounted to some 17 billions of dollars. There is no need to quibble over which comparisons are relevant. The important point is that the often-neglected local expenditures are significant and, when viewed in terms of expenditures on goods and services only, take on even more significance. Hence an important question arises whether at this level of govern-

<sup>9</sup> The discussion that follows applies to local governments. It will be apparent as the argument proceeds that it also applies, with less force, to state governments.

<sup>10</sup> A question does arise as to just what are the proper expenditures to consider. Following Musgrave, I shall consider only expenditures on goods or services (his Branch I expenditures). Thus interest on the federal debt is not included. At the local level interest payments might be included, since they are considered payments for services currently used, such as those provided by roads and schools.

ment any mechanism operates to insure that expenditures on these public goods approximate the proper level.

Consider for a moment the case of the city resident about to move to the suburbs. What variables will influence his choice of a municipality? If he has children, a high level of expenditures on schools may be important. Another person may prefer a community with a municipal golf course. The availability and quality of such facilities and services as beaches, parks, police protection, roads, and parking facilities will enter into the decision-making process. Of course, non-economic variables will also be considered, but this is of no concern at this point.

The consumer-voter may be viewed as picking that community which best satisfies his preference pattern for public goods. This is a major difference between central and local provision of public goods. At the central level the preferences of the consumer-voter are given, and the government tries to adjust to the pattern of these preferences, whereas at the local level various governments have their revenue and expenditure patterns more or less set.<sup>11</sup> Given these revenue and expenditure patterns, the consumer-voter moves to that community whose local government best satisfies his set of preferences. The greater the number of communities and the greater the variance among them, the closer the consumer will come to fully realizing his preference position.<sup>12</sup>

<sup>11</sup> This is an assumption about reality. In the extreme model that follows the patterns are assumed to be absolutely fixed.

<sup>12</sup> This is also true of many non-economic variables. Not only is the consumer-voter concerned with economic patterns, but he desires, for example, to associate with "nice" people. Again, the greater the number of communities, the closer he will come to satisfying his total preference function, which includes non-economic variables.

## A LOCAL GOVERNMENT MODEL

The implications of the preceding argument may be shown by postulating an extreme model. Here the following assumptions are made:

1. Consumer-voters are fully mobile and will move to that community where their preference patterns, which are set, are best satisfied.
2. Consumer-voters are assumed to have full knowledge of differences among revenue and expenditure patterns and to react to these differences.
3. There are a large number of communities in which the consumer-voters may choose to live.
4. Restrictions due to employment opportunities are not considered. It may be assumed that all persons are living on dividend income.
5. The public services supplied exhibit no external economies or diseconomies between communities.

Assumptions 6 and 7 to follow are less familiar and require brief explanations:

6. For every pattern of community services set by, say, a city manager who follows the preferences of the older residents of the community, there is an optimal community size. This optimum is defined in terms of the number of residents for which this bundle of services can be produced at the lowest average cost. This, of course, is closely analogous to the low point of a firm's average cost curve. Such a cost function implies that some factor or resource is fixed. If this were not so, there would be no logical reason to limit community size, given the preference patterns. In the same sense that the average cost curve has a minimum for one firm but can be reproduced by another there is seemingly no reason why a duplicate community cannot exist. The assumption that some factor is fixed

explains why it is not possible for the community in question to double its size by growth. The factor may be the limited land area of a suburban community, combined with a set of zoning laws against apartment buildings. It may be the local beach, whose capacity is limited. Anything of this nature will provide a restraint.

In order to see how this restraint works, let us consider the beach problem. Suppose the preference patterns of the community are such that the optimum size population is 13,000. Within this set of preferences there is a certain demand per family for beach space. This demand is such that at 13,000 population a 500-yard beach is required. If the actual length of the beach is, say, 600 yards, then it is not possible to realize this preference pattern with twice the optimum population, since there would be too little beach space by 400 yards.

The assumption of a fixed factor is necessary, as will be shown later, in order to get a determinate number of communities. It also has the advantage of introducing a realistic restraint into the model.

7. The last assumption is that communities below the optimum size seek to attract new residents to lower average costs. Those above optimum size do just the opposite. Those at an optimum try to keep their populations constant.

This assumption needs to be amplified. Clearly, communities below the optimum size, through chambers of commerce or other agencies, seek to attract new residents. This is best exemplified by the housing developments in some suburban areas, such as Park Forest in the Chicago area and Levittown in the New York area, which need to reach an optimum size. The same is true of communities that try to attract manufacturing indus-

tries by setting up certain facilities and getting an optimum number of firms to move into the industrially zoned area.

The case of the city that is too large and tries to get rid of residents is more difficult to imagine. No alderman in his right political mind would ever admit that the city is too big. Nevertheless, economic forces are at work to push people out of it. Every resident who moves to the suburbs to find better schools, more parks, and so forth, is reacting, in part, against the pattern the city has to offer.

The case of the community which is at the optimum size and tries to remain so is not hard to visualize. Again proper zoning laws, implicit agreements among realtors, and the like are sufficient to keep the population stable.

Except when this system is in equilibrium, there will be a subset of consumer-voters who are discontented with the patterns of their community. Another set will be satisfied. Given the assumption about mobility and the other assumptions listed previously, movement will take place out of the communities of greater than optimal size into the communities of less than optimal size. The consumer-voter moves to the community that satisfies his preference pattern.

The act of moving or failing to move is crucial. Moving or failing to move replaces the usual market test of willingness to buy a good and reveals the consumer-voter's demand for public goods. Thus each locality has a revenue and expenditure pattern that reflects the desires of its residents. The next step is to see what this implies for the allocation of public goods at the local level.

Each city manager now has a certain demand for  $n$  local public goods. In supplying these goods, he and  $m - 1$  other city managers may be considered as go-

ing to a national market and bidding for the appropriate units of service of each kind: so many units of police for the  $i$ th community; twice that number for the  $j$ th community; and so on. The demand on the public goods market for each of the  $n$  commodities will be the sum of the demands of the  $m$  communities. In the limit, as shown in a less realistic model to be developed later, this total demand will approximate the demand that represents the true preferences of the consumer-voters—that is, the demand they would reveal, if they were forced, somehow, to state their true preferences.<sup>13</sup> In this model there is no attempt on the part of local governments to "adapt to" the preferences of consumer-voters. Instead, those local governments that attract the optimum number of residents may be viewed as being "adopted by" the economic system.<sup>14</sup>

#### A COMPARISON MODEL

It is interesting to contrast the results of the preceding model with those of an even more severe model in order to see how these results differ from the normal market result. It is convenient to look at this severe model by developing its private-market counterpart. First assume that there are no public goods, only private ones. The preferences for these goods can be expressed as one of  $n$  patterns. Let a law be passed that all persons living in any one of the communities shall spend their money in the particular pattern described for that community by law. Given our earlier assumptions 1 through 5, it follows that, if the consum-

<sup>13</sup> The word "approximate" is used in recognition of the limitations of this model, and of the more severe model to be developed shortly, with respect to the cost of mobility. This issue will be discussed later.

<sup>14</sup> See Alchian, *op. cit.*

ers move to the community whose law happens to fit their preference pattern, they will be at their optimum. The  $n$  communities, in turn, will then send their buyers to market to purchase the goods for the consumer-voters in their community. Since this is simply a lumping together of all similar tastes for the purpose of making joint purchases, the allocation of resources will be the same as it would be if normal market forces operated. This conceptual experiment is the equivalent of substituting the city manager for the broker or middleman.

Now turn the argument around and consider only public goods. Assume with Musgrave that the costs of additional services are constant.<sup>15</sup> Further, assume that a doubling of the population means doubling the amount of services required. Let the number of communities be infinite and let each announce a different pattern of expenditures on public goods. Define an empty community as one that fails to satisfy anybody's preference pattern. Given these assumptions, including the earlier assumptions 1 through 5, the consumer-voters will move to that community which *exactly* satisfies their preferences. This must be true, since a one-person community is allowed. The sum of the demands of the  $n$  communities reflects the demand for local public services. In this model the demand is exactly the same as it would be if it were determined by normal market forces.

However, this severe model does not make much sense. The number of communities is indeterminate. There is no reason why the number of communities will not be equal to the population, since each voter can find the one that exactly fits his preferences. Unless some socio-logical variable is introduced, this may

reduce the solution of the problem of allocating public goods to the trite one of making each person his own municipal government. Hence this model is not even a first approximation of reality. It is presented to show the assumptions needed in a model of local government expenditures, which yields the same optimal allocation that a private market would.

#### THE LOCAL GOVERNMENT MODEL RE-EXAMINED

The first model, described by the first five assumptions together with assumptions 6 and 7, falls short of this optimum. An example will serve to show why this is the case.

Let us return to the community with the 500-yard beach. By assumption, its optimum population was set at 13,000, given its preference patterns. Suppose that some people in addition to the optimal 13,000 would choose this community if it were available. Since they cannot move into this area, they must accept the next best substitute.<sup>16</sup> If a perfect substitute is found, no problem exists. If one is not found, then the failure to reach the optimal preference position and the substitution of a lower position becomes a matter of degree. In so far as there are a number of communities with similar revenue and expenditure patterns, the solution will approximate the ideal "market" solution.

Two related points need to be mentioned to show the allocative results of this model: (1) changes in the costs of one of the public services will cause changes in the quantity produced; (2) the

<sup>15</sup> In the constant cost model with an infinite number of communities this problem does not arise, since the number of beaches can be doubled or a person can find another community that is a duplicate of his now filled first choice.

<sup>16</sup> Musgrave, "Voluntary Exchange . . .," *op. cit.*

costs of moving from community to community should be recognized. Both points can be illustrated in one example.

Suppose lifeguards throughout the country organize and succeed in raising their wages. Total taxes in communities with beaches will rise. Now residents who are largely indifferent to beaches will be forced to make a decision. Is the saving of this added tax worth the cost of moving to a community with little or no beach? Obviously, this decision depends on many factors, among which the availability of and proximity to a suitable substitute community is important. If enough people leave communities with beaches and move to communities without beaches, the total amount of lifeguard services used will fall. These models then, unlike their private-market counterpart, have mobility as a cost of registering demand. The higher this cost, *ceteris paribus*, the less optimal the allocation of resources.

This distinction should not be blown out of proportion. Actually, the cost of registering demand comes through the introduction of space into the economy. Yet space affects the allocation not only of resources supplied by local governments but of those supplied by the private market as well. Every time available resources or production techniques change, a new location becomes optimal for the firm. Indeed, the very concept of the shopping trip shows that the consumer does pay a cost to register his demand for private goods. In fact, Koopmans has stated that the nature of the assignment problem is such that in a space economy with transport costs there is *no* general equilibrium solution as set by market forces.<sup>17</sup>

<sup>17</sup> Tjalling Koopmans, "Mathematical Ground-work of Economic Optimization Theories," paper read at the annual meeting of the Econometric Society (December, 1954).

Thus the problems stated by this model are not unique; they have their counterpart in the private market. We are maximizing within the framework of the resources available. If production functions show constant returns to scale with generally diminishing factor returns, and if indifference curves are regularly convex, an optimal solution is possible. On the production side it is assumed that communities are forced to keep production costs at a minimum either through the efficiency of city managers or through competition from other communities.<sup>18</sup> Given this, on the demand side we may note with Samuelson that "each individual, in seeking as a competitive buyer to get to the highest level of indifference subject to given prices and *tax*, would be led as if by an Invisible Hand to the grand solution of the social maximum position."<sup>19</sup> Just as the consumer may be visualized as walking to a private market place to buy his goods, the prices of which are set, we place him in the position of walking to a community where the prices (taxes) of community services are set. Both trips take the consumer to market. There is no way in which the consumer can avoid revealing his preferences in a spatial economy. Spatial mobility provides the local public-goods counterpart to the private market's shopping trip.

<sup>18</sup> In this model and in reality, the city manager or elected official who is not able to keep his costs (taxes) low compared with those of similar communities will find himself out of a job. As an institutional observation, it may well be that city managers are under greater pressure to minimize costs than their private-market counterparts—firm managers. This follows from (1) the reluctance of the public to pay taxes and, what may be more important, (2) the fact that the costs of competitors—other communities—are a matter of public record and may easily be compared.

<sup>19</sup> "The Pure Theory . . .," *op. cit.*, p. 388. (Italics mine.)

## EXTERNAL ECONOMIES AND MOBILITY

Relaxing assumption 5 has some interesting implications. There are obvious external economies and diseconomies between communities. My community is better off if its neighbor sprays trees to prevent Dutch elm disease. On the other hand, my community is worse off if the neighboring community has inadequate law enforcement.

In cases in which the external economies and diseconomies are of sufficient importance, some form of integration may be indicated.<sup>20</sup> Not all aspects of law enforcement are adequately handled at the local level. The function of the sheriff, state police, and the FBI—as contrasted with the local police—may be cited as resulting from a need for integration. In real life the diseconomies are minimized in so far as communities reflecting the same socioeconomic preferences are contiguous. Suburban agglomerations such as Westchester, the North Shore, and the Main Line are, in part, evidence of these external economies and diseconomies.

Assumptions 1 and 2 should be checked against reality. Consumer-voters do not have perfect knowledge and set preferences, nor are they perfectly mobile. The question is how do people actually react in choosing a community. There has been very little empirical study of the motivations of people in choosing a community. Such studies as have been undertaken seem to indicate a surprising awareness of differing revenue and expenditure patterns.<sup>21</sup> The general disdain with which proposals to integrate municipalities are met seems to

<sup>20</sup> I am grateful to Stanley Long and Donald Markwalder for suggesting this point.

<sup>21</sup> See Wendell Bell, "Familism and Suburbanization: One Test of the Choice Hypothesis," a paper read at the annual meeting of the American Sociological Society, Washington, D.C., August, 1955. Forthcoming in *Rural Sociology*, December, 1956.

reflect, in part, the fear that local revenue-expenditure patterns will be lost as communities are merged into a metropolitan area.

## POLICY IMPLICATIONS

The preceding analysis has policy implications for municipal integration, provision for mobility, and set local revenue and expenditure patterns. These implications are worth brief consideration.

On the usual economic welfare grounds, municipal integration is justified only if more of any service is forthcoming at the same total cost and without reduction of any other service. A general reduction of costs along with a reduction in one or more of the services provided cannot be justified on economic grounds unless the social welfare function is known. For example, those who argue for a metropolitan police force instead of local police cannot prove their case on purely economic grounds.<sup>22</sup> If one of the communities were to receive less police protection after integration than it received before, integration could be objected to as a violation of consumers' choice.

Policies that promote residential mobility and increase the knowledge of the consumer-voter will improve the allocation of government expenditures in the same sense that mobility among jobs and knowledge relevant to the location of industry and labor improve the allocation of private resources.

Finally, we may raise the normative question whether local governments *should*, to the extent possible, have a fixed revenue-expenditure pattern. In a large, dynamic metropolis this may be

<sup>22</sup> For example, in Cook County—the Chicago area—Sheriff Joseph Lohman argues for such a metropolitan police force.

impossible. Perhaps it could more appropriately be considered by rural and suburban communities.

#### CONCLUSION

It is useful in closing to restate the problem as Samuelson sees it:

*However, no decentralized pricing system can serve to determine optimally these levels of collective consumption.* Other kinds of "voting" or "signaling" would have to be tried. . . . Of course utopian voting and signaling schemes can be imagined. . . . The failure of market catallactics in no way denies the following truth: given sufficient knowledge the optimal decisions can always be found by scanning over all the attainable states of the world and selecting the one which according to the postulated ethical welfare function is best. The solution "exists"; the problem is how to "find" it.<sup>23</sup>

It is the contention of this article that, for a substantial portion of collective or public goods, this problem *does have a*

conceptual solution. If consumer-voters are fully mobile, the appropriate local governments, whose revenue-expenditure patterns are set, are adopted by the consumer-voters. While the solution may not be perfect because of institutional rigidities, this does not invalidate its importance. The solution, like a general equilibrium solution for a private spatial economy, is the best that can be obtained given preferences and resource endowments.

Those who are tempted to compare this model with the competitive private model may be disappointed. Those who compare the reality described by this model with the reality of the competitive model—given the degree of monopoly, friction, and so forth—*may* find that local government represents a sector where the allocation of public goods (as a reflection of the preferences of the population) need not take a back seat to the private sector.

<sup>23</sup> "The Pure Theory . . . ,"*op. cit.*, pp. 388-89.

89033508631



b89033508631a

# THE LOGIC OF COLLECTIVE ACTION

Public Goods  
and the  
Theory of Groups

PAMELA OLIVER  
SOCIOLOGY DEPT.

MANCUR OLSON

Harvard University Press  
Cambridge • Massachusetts  
London • England

Mancur Olson  
The Logic of Collective Action

## Introduction

It is often taken for granted, at least where economic objectives are involved, that groups of individuals with common interests usually attempt to further those common interests. Groups of individuals with common interests are expected to act on behalf of their common interests much as single individuals are often expected to act on behalf of their personal interests. This opinion about group behavior is frequently found not only in popular discussions but also in scholarly writings. Many economists of diverse methodological and ideological traditions have implicitly or explicitly accepted it. This view has, for example, been important in many theories of labor unions, in Marxian theories of class action, in concepts of "counter-vailing power," and in various discussions of economic institutions. It has, in addition, occupied a prominent place in political science, at least in the United States, where the study of pressure groups has been dominated by a celebrated "group theory" based on the idea that groups will act when necessary to further their common or group goals. Finally, it has played a significant role in many well-known sociological studies.

The view that groups act to serve their interests presumably is based upon the assumption that the individuals in groups act out of self-interest. If the individuals in a group altruistically disregarded their personal welfare, it would not be very likely that collectively they would seek some selfish common or group objective. Such altruism, is, however, considered exceptional, and self-interested behavior is usually thought to be the rule, at least when economic issues are at stake; no one is surprised when individual businessmen seek higher profits, when individual workers seek higher wages, or when individual consumers seek lower prices. The idea that groups tend to act in support of their group interests is supposed to follow logically from this widely accepted premise of rational, self-interested behavior. In other words, if the members of some group have a common interest or objective, and if they would all be better off if that objective were achieved, it has been thought to follow logically that the individuals in that group would, if they were rational and self-interested, act to achieve that objective.

But it is *not* in fact true that the idea that groups will act in their

self-interest follows logically from the premise of rational and self-interested behavior. It does *not* follow, because all of the individuals in a group would gain if they achieved their group objective, that they would act to achieve that objective, even if they were all rational and self-interested. Indeed, unless the number of individuals in a group is quite small, or unless there is coercion or some other special device to make individuals act in their common interest, *rational, self-interested individuals will not act to achieve their common or group interests.* In other words, even if all of the individuals in a large group are rational and self-interested, and would gain if, as a group, they acted to achieve their common interest or objective, they will still not voluntarily act to achieve that common or group interest. The notion that groups of individuals will act to achieve their common or group interests, far from being a logical implication of the assumption that the individuals in a group will rationally further their individual interests, is in fact inconsistent with that assumption. This inconsistency will be explained in the following chapter.

If the members of a large group rationally seek to maximize their personal welfare, they will *not* act to advance their common or group objectives unless there is coercion to force them to do so, or unless some separate incentive, distinct from the achievement of the common or group interest, is offered to the members of the group individually on the condition that they help bear the costs or burdens involved in the achievement of the group objectives. Nor will such large groups form organizations to further their common goals in the absence of the coercion or the separate incentives just mentioned. These points hold true even when there is unanimous agreement in a group about the common good and the methods of achieving it.

The widespread view, common throughout the social sciences, that groups tend to further their interests, is accordingly unjustified, at least when it is based, as it usually is, on the (sometimes implicit) assumption that groups act in their self-interest because individuals do. There is paradoxically the logical possibility that groups composed of either altruistic individuals or irrational individuals may sometimes act in their common or group interests. But, as later, empirical parts of this study will attempt to show, this logical possibility is usually of no practical importance. Thus the customary view that groups of individuals with common interests tend to further those common interests appears to have little if any merit.

None of the statements made above fully applies to small groups, for the situation in small groups is much more complicated. In small groups there may very well be some voluntary action in support of the common purposes of the individuals in the group, but in most cases this action will cease before it reaches the optimal level for the members of the group as a whole. In the sharing of the costs of efforts to achieve a common goal in small groups, there is however a surprising tendency for the "exploitation" of the *great by the small.*

The proofs of all of the logical statements that have been made above are contained in Chapter I, which develops a logical or theoretical explanation of certain aspects of group and organizational behavior. Chapter II examines the implications of this analysis for groups of different size, and illustrates the conclusion that in many cases small groups are more efficient and viable than large ones. Chapter III considers the implications of the argument for labor unions, and draws the conclusion that some form of compulsory membership is, in most circumstances, indispensable to union survival. The fourth chapter uses the approach developed in this study to examine Marx's theory of social classes and to analyze the theories of the state developed by some other economists. The fifth analyzes the "group theory" used by many political scientists in the light of the logic elaborated in this study, and argues that that theory as usually understood is logically inconsistent. The final chapter develops a new theory of pressure groups which is consistent with the logical relationships outlined in the first chapter, and which suggests that the membership and power of large pressure-group organizations does not derive from their lobbying achievements, but is rather a by-product of their other activities.

Though I am an economist, and the tools of analysis used in this book are drawn from economic theory, the conclusions of the study are as relevant to the sociologist and the political scientist as they are to the economist. I have, therefore, avoided using the diagrammatic-mathematical language of economics whenever feasible. Unfortunately, many noneconomists will find one or two brief parts of the first chapter expressed in an obscure and uncongenial way, but all of the rest of the book should be perfectly clear, whatever the reader's disciplinary background.

# I

## A Theory of Groups and Organizations

### A. THE PURPOSE OF ORGANIZATION

Since most (though by no means all) of the action taken by or on behalf of groups of individuals is taken through organizations, it will be helpful to consider organizations in a general or theoretical way.<sup>1</sup> The logical place to begin any systematic study of organizations is with their purpose. But there are all types and shapes and sizes of organizations, even of economic organizations, and there is then some question whether there is any single purpose that would be characteristic of organizations generally. One purpose that is nonetheless characteristic of most organizations, and surely of practically all organizations with an important economic aspect, is the furtherance of the interests of their members. That would seem obvious, at least from the economist's perspective. To be sure, some organizations may out of ignorance fail to further their members' interests, and others may be enticed into serving only the ends of the leadership.<sup>2</sup>

1. Economists have for the most part neglected to develop theories of organizations, but there are a few works from an economic point of view on the subject. See, for example, three papers by Jacob Marschak, "Elements for a Theory of Teams," *Management Science*, I (January 1955), 127-137, "Towards an Economic Theory of Organization and Information," in *Decision Processes*, ed. R. M. Thrall, C. H. Combs, and R. L. Davis (New York: John Wiley, 1954), pp. 187-220, and "Efficient and Viable Organization Forms," in *Modern Organization Theory*, ed. Mason Haire (New York: John Wiley, 1959), pp. 307-320; two papers by R. Radner, "Application of Linear Programming to Team Decision Problems," *Management Science*, V (January 1959), 143-150, and "Team Decision Problems," *Annals of Mathematical Statistics*, XXXIII (September 1962), 857-881; C. B. McGuire, "Some Team Models of a Sales Organization," *Management Science*, VII (January 1961), 101-130; Oskar Morgenstern, *Prolegomena to a Theory of Organization* (Santa Monica, Calif.: RAND Research Memorandum 734, 1951); James G. March and Herbert A. Simon, *Organizations* (New York: John Wiley, 1958); Kenneth Boulding, *The Organizational Revolution* (New York: Harper, 1953).

2. Max Weber called attention to the case where an organization continues to exist for some time after it has become meaningless because some official is making a living out of it. See his *Theory of Social and Economic Organization*, trans. Talcott Parsons and A. M. Henderson (New York: Oxford University Press, 1947), p. 318.

But organizations often perish if they do nothing to further the interests of their members, and this factor must severely limit the number of organizations that fail to serve their members.

The idea that organizations or associations exist to further the interests of their members is hardly novel, nor peculiar to economics; it goes back at least to Aristotle, who wrote, "Men journey together with a view to particular advantage, and by way of providing some particular thing needed for the purposes of life, and similarly the political association seems to have come together originally, and to continue in existence, for the sake of the *general* advantages it brings."<sup>3</sup> More recently Professor Leon Festinger, a social psychologist, pointed out that "the attraction of group membership is not so much in sheer belonging, but rather in attaining something by means of this membership."<sup>4</sup> The late Harold Laski, a political scientist, took it for granted that "associations exist to fulfill purposes which a group of men have in common."<sup>5</sup>

The kinds of organizations that are the focus of this study are expected to further the interests of their members.<sup>6</sup> Labor unions are expected to strive for higher wages and better working conditions for their members; farm organizations are expected to strive for favorable legislation for their members; cartels are expected to strive for higher prices for participating firms; the corporation is expected to further the interests of its stockholders;<sup>7</sup> and the state is expected

3. *Ethics* viii.9.1160a.

4. Leon Festinger, "Group Attraction and Membership," in *Group Dynamics*, ed. Dorwin Cartwright and Alvin Zander (Evanston, Ill.: Row, Peterson, 1953), p. 93.

5. *A Grammar of Politics*, 4th ed. (London: George Allen & Unwin, 1939), p. 67.

6. Philanthropic and religious organizations are not necessarily expected to serve only the interests of their members; such organizations have other purposes that are considered more important, however much their members "need" to belong, or are improved or helped by belonging. But the complexity of such organizations need not be debated at length here, because this study will focus on organizations with a significant economic aspect. The emphasis here will have something in common with what Max Weber called the "associative group"; he called a group associative if "the orientation of social action with it rests on a rationally motivated agreement." Weber contrasted his "associative group" with the "communal group" which was centered on personal affection, erotic relationships, etc., like the family. (See Weber, pp. 136-139, and Grace Coyle, *Social Process in Organized Groups*, New York: Richard Smith, Inc., 1930, pp. 7-9.) The logic of the theory developed here can be extended to cover communal, religious, and philanthropic organizations, but the theory is not particularly useful in studying such groups. See my pp. 61n17, 159-162.

7. That is, its members. This study does not follow the terminological usage of those organization theorists who describe employees as "members" of the organization for which they work. Here it is more convenient to follow the language of everyday

to further the common interests of its citizens (though in this nationalistic age the state often has interests and ambitions apart from those of its citizens).

Notice that the interests that all of these diverse types of organizations are expected to further are for the most part *common* interests: the union members' common interest in higher wages, the farmers' common interest in favorable legislation, the cartel members' common interest in higher prices, the stockholders' common interest in higher dividends and stock prices, the citizens' common interest in good government. It is not an accident that the diverse types of organizations listed are all supposed to work primarily for the *common* interests of their members. Purely personal or individual interests can be advanced, and usually advanced most efficiently, by individual, unorganized action. There is obviously no purpose in having an organization when individual, unorganized action can serve the interests of the individual as well as or better than an organization; there would, for example, be no point in forming an organization simply to play solitaire. But when a number of individuals have a common or collective interest—when they share a single purpose or objective—individual, unorganized action (as we shall soon see) will either not be able to advance that common interest at all, or will not be able to advance that interest adequately. Organizations can therefore perform a function when there are common or group interests, and though organizations often also serve purely personal, individual interests, their characteristic and primary function is to advance the common interests of groups of individuals.

The assumption that organizations typically exist to further the common interests of groups of people is implicit in most of the literature about organizations, and two of the writers already cited make this assumption explicit: Harold Laski emphasized that organizations exist to achieve purposes or interests which "a group of men have in common," and Aristotle apparently had a similar notion in mind when he argued that political associations are created and maintained because of the "general advantages" they bring. R. M.

usage instead, and to distinguish the members of, say, a union from the employees of that union. Similarly, the members of the union will be considered employees of the corporation for which they work, whereas the members of the corporation are the common stockholders.

MacIver also made this point explicitly when he said that "every organization presupposes an interest which its members all share."<sup>8</sup>

Even when unorganized groups are discussed, at least in treatments of "pressure groups" and "group theory," the word "group" is used in such a way that it means "a number of individuals with a common interest." It would of course be reasonable to label even a number of people selected at random (and thus without any common interest or unifying characteristic) as a "group"; but most discussions of group behavior seem to deal mainly with groups that do have common interests. As Arthur Bentley, the founder of the "group theory" of modern political science, put it, "there is no group without its interest."<sup>9</sup> The social psychologist Raymond Cattell was equally explicit, and stated that "every group has its interest."<sup>10</sup> This is also the way the word "group" will be used here.

Just as those who belong to an organization or a group can be presumed to have a common interest,<sup>11</sup> so they obviously also have purely individual interests, different from those of the others in the organization or group. All of the members of a labor union, for example, have a common interest in higher wages, but at the same time each worker has a unique interest in his personal income, which depends not only on the rate of wages but also on the length of time that he works.

8. R. M. MacIver, "Interests," *Encyclopaedia of the Social Sciences*, VII (New York: Macmillan, 1932), 147.

9. Arthur Bentley, *The Process of Government* (Evanston, Ill.: Principia Press, 1949), p. 211. David B. Truman takes a similar approach; see his *The Governmental Process* (New York: Alfred A. Knopf, 1958), pp. 33-35. See also Sidney Verba, *Small Groups and Political Behavior* (Princeton, N.J.: Princeton University Press, 1961), pp. 12-13.

10. Raymond Cattell, "Concepts and Methods in the Measurement of Group Syntality," in *Small Groups*, ed. A. Paul Hare, Edgard F. Borgatta, and Robert F. Bales (New York: Alfred A. Knopf, 1955), p. 115.

11. Any organization or group will of course usually be divided into subgroups or factions that are opposed to one another. This fact does not weaken the assumption made here that organizations exist to serve the common interests of members, for the assumption does not imply that intragroup conflict is neglected. The opposing groups within an organization ordinarily have some interest in common (if not, why would they maintain the organization?), and the members of any subgroup or faction also have a separate common interest of their own. They will indeed often have a common purpose in defeating some other subgroup or faction. The approach used here does not neglect the conflict within groups and organizations, then, because it considers each organization as a unit only to the extent that it does in fact attempt to serve a common interest, and considers the various subgroups as the relevant units with common interests to analyze the factional strife.

#### B. PUBLIC GOODS AND LARGE GROUPS

The combination of individual interests and common interests in an organization suggests an analogy with a competitive market.<sup>12</sup> The firms in a perfectly competitive industry, for example, have a common interest in a higher price for the industry's product. Since a uniform price must prevail in such a market, a firm cannot expect a higher price for itself unless all of the other firms in the industry also have this higher price. But a firm in a competitive market also has an interest in selling as much as it can, until the cost of producing another unit exceeds the price of that unit. In this there is no common interest; each firm's interest is directly opposed to that of every other firm, for the more other firms sell, the lower the price and income for any given firm. In short, while all firms have a common interest in a higher price, they have antagonistic interests where output is concerned. This can be illustrated with a simple supply-and-demand model. For the sake of a simple argument, assume that a perfectly competitive industry is momentarily in a disequilibrium position, with price exceeding marginal cost for all firms at their present output. Suppose, too, that all of the adjustments will be made by the firms already in the industry rather than by new entrants, and that the industry is on an inelastic portion of its demand curve. Since price exceeds marginal cost for all firms, output will increase. But as all firms increase production, the price falls; indeed, since the industry demand curve is by assumption inelastic, the total revenue of the industry will decline. Apparently each firm finds that with price exceeding marginal cost, it pays to increase its output, but the result is that each firm gets a smaller profit. Some economists in an earlier day may have questioned this result,<sup>13</sup> but the fact that profit-maximizing firms in a perfectly competitive industry can act contrary to their interests as a group is now widely understood and accepted.<sup>14</sup> A group of profit-maximizing firms can act to reduce their aggregate profits because in perfect competition each firm is, by definition, so small that it can ignore the effect of its output on price. Each firm finds it to its advantage to increase output to the point where mar-

12. See J. M. Clark, *The Economics of Overhead Costs* (Chicago: University of Chicago Press, 1923), p. 417, and Frank H. Knight, *Risk, Uncertainty and Profit* (Boston: Houghton Mifflin, 1921), p. 193.

13. Edward H. Chamberlin, *Monopolistic Competition*, 6th ed. (Cambridge, Mass.: Harvard University Press, 1950), p. 4.

ginal cost equals price and to ignore the effects of its extra output on the position of the industry. It is true that the net result is that all firms are worse off, but this does not mean that every firm has not maximized its profits. If a firm, foreseeing the fall in price resulting from the increase in industry output, were to restrict its own output, it would lose more than ever, for its price would fall quite as much in any case and it would have a smaller output as well. A firm in a perfectly competitive market gets only a small part of the benefit (or a small share of the industry's extra revenue) resulting from a reduction in that firm's output.

For these reasons it is now generally understood that if the firms in an industry are maximizing profits, the profits for the industry as a whole will be less than they might otherwise be.<sup>14</sup> And almost everyone would agree that this theoretical conclusion fits the facts for markets characterized by pure competition. The important point is that this is true because, though all the firms have a common interest in a higher price for the industry's product, it is in the interest of each firm that the other firms pay the cost—in terms of the necessary reduction in output—needed to obtain a higher price.

About the only thing that keeps prices from falling in accordance with the process just described in perfectly competitive markets is outside intervention. Government price supports, tariffs, cartel agreements, and the like may keep the firms in a competitive market from acting contrary to their interests. Such aid or intervention is quite common. It is then important to ask how it comes about. How does a competitive industry obtain government assistance in maintaining the price of its product?

Consider a hypothetical, competitive industry, and suppose that most of the producers in that industry desire a tariff, a price-support program, or some other government intervention to increase the price for their product. To obtain any such assistance from the government, the producers in this industry will presumably have to organize a lobbying organization; they will have to become an active pressure group.<sup>15</sup> This lobbying organization may have to conduct a con-

14. For a fuller discussion of this question see Mancur Olson, Jr., and David McFarland, "The Restoration of Pure Monopoly and the Concept of the Industry," *Quarterly Journal of Economics*, LXXVI (November 1962), 613–631.

15. Robert Michels contends in his classic study that "democracy is inconceivable without organization," and that "the principle of organization is an absolutely essential condition for the political struggle of the masses." See his *Political Parties*,

siderable campaign. If significant resistance is encountered, a great amount of money will be required.<sup>16</sup> Public relations experts will be needed to influence the newspapers, and some advertising may be necessary. Professional organizers will probably be needed to organize "spontaneous grass roots" meetings among the distressed producers in the industry, and to get those in the industry to write letters to their congressmen.<sup>17</sup> The campaign for the government assistance will take the time of some of the producers in the industry, as well as their money.

There is a striking parallel between the problem the perfectly competitive industry faces as it strives to obtain government assistance, and the problem it faces in the marketplace when the firms increase output and bring about a fall in price. *Just as it was not rational for a particular producer to restrict his output in order that there might be a higher price for the product of his industry, so it would not be rational for him to sacrifice his time and money to support a lobbying organization to obtain government assistance for the industry. In neither case would it be in the interest of the individual producer to assume any of the costs himself. A lobbying organization, or indeed a labor union or any other organization, working in the interest of a large group of firms or workers in some industry, would get no assistance from the rational, self-interested individuals in that industry.* This would be true even if everyone in the industry were absolutely convinced that the proposed program was in their interest (though in fact some might think otherwise and make the organization's task yet more difficult).<sup>18</sup>

Although the lobbying organization is only one example of the logical analogy between the organization and the market, it is of

trans. Eden and Cedar Paul (New York: Dover Publications, 1959), pp. 21–22. See also Robert A. Brady, *Business as a System of Power* (New York: Columbia University Press, 1943), p. 193.

16. Alexander Heard, *The Costs of Democracy* (Chapel Hill: University of North Carolina Press, 1960), especially note 1, pp. 95–96. For example, in 1947 the National Association of Manufacturers spent over \$4.6 million, and over a somewhat longer period the American Medical Association spent as much on a campaign against compulsory health insurance.

17. "If the full truth were ever known . . . lobbying, in all its ramifications, would prove to be a billion dollar industry." U.S. Congress, House, Select Committee on Lobbying Activities, *Report*, 81st Cong., 2nd Sess. (1950), as quoted in the *Congressional Quarterly Almanac*, 81st Cong., 2nd Sess., VI, 764–765.

18. For a logically possible but practically meaningless exception to the conclusion of this paragraph, see footnote 68 in this chapter.

some practical importance. There are many powerful and well-financed lobbies with mass support in existence now, but these lobbying organizations do not get that support because of their legislative achievements. The most powerful lobbying organizations now obtain their funds and their following for other reasons, as later parts of this study will show.

Some critics may argue that the rational person will, indeed, support a large organization, like a lobbying organization, that works in his interest, because he knows that if he does not, others will not do so either, and then the organization will fail, and he will be without the benefit that the organization could have provided. This argument shows the need for the analogy with the perfectly competitive market. For it would be quite as reasonable to argue that prices will never fall below the levels a monopoly would have charged in a perfectly competitive market, because if one firm increased its output, other firms would also, and the price would fall; but each firm could foresee this, so it would not start a chain of price-destroying increases in output. In fact, it does not work out this way in a competitive market; nor in a large organization. When the number of firms involved is large, no one will notice the effect on price if one firm increases its output, and so no one will change his plans because of it. Similarly, in a large organization, the loss of one dues payer will not noticeably increase the burden for any other one dues payer, and so a rational person would not believe that if he were to withdraw from an organization he would drive others to do so.

The foregoing argument must at the least have some relevance to economic organizations that are mainly means through which individuals attempt to obtain the same things they obtain through their activities in the market. Labor unions, for example, are organizations through which workers strive to get the same things they get with their individual efforts in the market—higher wages, better working conditions, and the like. It would be strange indeed if the workers did not confront some of the same problems in the union that they meet in the market, since their efforts in both places have some of the same purposes.

However similar the purposes may be, critics may object that attitudes in organizations are not at all like those in markets. In organizations, an emotional or ideological element is often also involved. Does this make the argument offered here practically irrelevant?

A most important type of organization—the national state—will serve to test this objection. Patriotism is probably the strongest non-economic motive for organizational allegiance in modern times. This age is sometimes called the age of nationalism. Many nations draw additional strength and unity from some powerful ideology, such as democracy or communism, as well as from a common religion, language, or cultural inheritance. The state not only has many such powerful sources of support; it also is very important economically. Almost any government is economically beneficial to its citizens, in that the law and order it provides is a prerequisite of all civilized economic activity. But despite the force of patriotism, the appeal of the national ideology, the bond of a common culture, and the indispensability of the system of law and order, no major state in modern history has been able to support itself through voluntary dues or contributions. Philanthropic contributions are not even a significant source of revenue for most countries. Taxes, *compulsory* payments by definition, are needed. Indeed, as the old saying indicates, their necessity is as certain as death itself.

If the state, with all of the emotional resources at its command, cannot finance its most basic and vital activities without resort to compulsion, it would seem that large private organizations might also have difficulty in getting the individuals in the groups whose interests they attempt to advance to make the necessary contributions voluntarily.<sup>10</sup>

**The reason the state cannot survive on voluntary dues or payments,**

19. Sociologists as well as economists have observed that ideological motives alone are not sufficient to bring forth the continuing effort of large masses of people. Max Weber provides a notable example:

"All economic activity in a market economy is undertaken and carried through by individuals for their own ideal or material interests. This is naturally just as true when economic activity is oriented to the patterns of order of corporate groups . . .

"Even if an economic system were organized on a socialistic basis, there would be no fundamental difference in this respect . . . The structure of interests and the relevant situation might change; there would be other means of pursuing interests, but this fundamental factor would remain just as relevant as before. It is of course true that economic action which is oriented on purely ideological grounds to the interest of others does exist. But it is even more certain that the mass of men do not act in this way, and it is an induction from experience that they cannot do so and never will . . .

"In a market economy the interest in the maximization of income is necessarily the driving force of all economic activity." (Weber, pp. 319-320.)

Talcott Parsons and Neil Smelser go even further in postulating that "performance" throughout society is proportional to the "rewards" and "sanctions" involved. See their *Economy and Society* (Glencoe, Ill.: Free Press, 1954), pp. 50-69.

but must rely on taxation, is that the most fundamental services a nation-state provides are, in one important respect,<sup>20</sup> like the higher price in a competitive market: they must be available to everyone if they are available to anyone. The basic and most elementary goods or services provided by government, like defense and police protection, and the system of law and order generally, are such that they go to everyone or practically everyone in the nation. It would obviously not be feasible, if indeed it were possible, to deny the protection provided by the military services, the police, and the courts to those who did not voluntarily pay their share of the costs of government, and taxation is accordingly necessary. The common or collective benefits provided by governments are usually called "public goods" by economists, and the concept of public goods is one of the oldest and most important ideas in the study of public finance. A common, collective, or public good is here defined as any good such that, if any person  $X_i$  in a group  $X_1, \dots, X_4, \dots, X_n$  consumes it, it cannot feasibly be withheld from the others in that group.<sup>21</sup> In

20. See, however, section E of this chapter, on "exclusive" and "inclusive" groups.

21. This simple definition focuses upon two points that are important in the present context. The first point is that most collective goods can only be defined with respect to some specific group. One collective good goes to one group of people, another collective good to another group; one may benefit the whole world, another only two specific people. Moreover, some goods are collective goods to those in one group and at the same time private goods to those in another, because some individuals can be kept from consuming them and others can't. Take for example the parade that is a collective good to all those who live in tall buildings overlooking the parade route, but which appears to be a private good to those who can see it only by buying tickets for a seat in the stands along the way. The second point is that once the relevant group has been defined, the definition used here, like Musgrave's, distinguishes collective good in terms of infeasibility of excluding potential consumers of the good. This approach is used because collective goods produced by organizations of all kinds seem to be such that exclusion is normally not feasible. To be sure, for some collective goods it is physically possible to practice exclusion. But, as Head has shown, it is not necessary that exclusion be technically impossible; it is only necessary that it be infeasible or uneconomic. Head has also shown most clearly that nonexcludability is only one of two basic elements in the traditional understanding of public goods. The other, he points out, is "jointness of supply." A good has "jointness" if making it available to one individual means that it can be easily or freely supplied to others as well. The polar case of jointness would be Samuelson's pure public good, which is a good such that additional consumption of it by one individual does not diminish the amount available to others. By the definition used here, jointness is not a necessary attribute of a public good. As later parts of this chapter will show, at least one type of collective good considered here exhibits no jointness whatever, and few if any would have the degree of jointness needed to qualify as pure public goods. Nonetheless, most of the collective goods to be studied here do display a large measure of jointness. On the definition and importance of public goods, see John G. Head,

other words, those who do not purchase or pay for any of the public or collective good cannot be excluded or kept from sharing in the consumption of the good, as they can where noncollective goods are concerned.

Students of public finance have, however, neglected the fact that *the achievement of any common goal or the satisfaction of any common interest means that a public or collective good has been provided for that group.*<sup>22</sup> The very fact that a goal or purpose is *common* to a group means that no one in the group is excluded from the benefit or satisfaction brought about by its achievement. As the opening paragraphs of this chapter indicated, almost all groups and organizations have the purpose of serving the common interests of their members. As R. M. MacIver puts it, "Persons . . . have common interests in the degree to which they participate in a cause . . . which indivisibly embraces them all."<sup>23</sup> It is of the essence of an organization that it provides an inseparable, generalized benefit. It follows that the provision of public or collective goods is the fundamental function of organizations generally. A state is first of all an organization that provides public goods for its members, the citizens; and other types of organizations similarly provide collective goods for their members.

And just as a state cannot support itself by voluntary contributions, or by selling its basic services on the market, neither can other large organizations support themselves without providing some sanction,

"Public Goods and Public Policy," *Public Finance*, vol. XVII, no. 3 (1962), 197-219; Richard Musgrave, *The Theory of Public Finance* (New York: McGraw-Hill, 1959); Paul A. Samuelson, "The Pure Theory of Public Expenditure," "Diagrammatic Exposition of A Theory of Public Expenditure," and "Aspects of Public Expenditure Theories," in *Review of Economics and Statistics*, XXXVI (November 1954), 387-390, XXXVII (November 1955), 350-356, and XL (November 1958), 332-338. For somewhat different opinions about the usefulness of the concept of public goods, see Julius Margolis, "A Comment on the Pure Theory of Public Expenditure," *Review of Economics and Statistics*, XXXVII (November 1955), 347-349, and Gerhard Colm, "Theory of Public Expenditures," *Annals of the American Academy of Political and Social Science*, CLXXXIII (January 1936), 1-11.

22. There is no necessity that a public good to one group in a society is necessarily in the interest of the society as a whole. Just as a tariff could be a public good to the industry that sought it, so the removal of the tariff could be a public good to those who consumed the industry's product. This is equally true when the public-good concept is applied only to governments; for a military expenditure, or a tariff, or an immigration restriction that is a public good to one country could be a "public bad" to another country, and harmful to world society as a whole.

23. R. M. MacIver in *Encyclopaedia of the Social Sciences*, VII, 147.

or some attraction distinct from the public good itself, that will lead individuals to help bear the burdens of maintaining the organization. The individual member of the typical large organization is in a position analogous to that of the firm in a perfectly competitive market, or the taxpayer in the state: his own efforts will not have a noticeable effect on the situation of his organization, and he can enjoy any improvements brought about by others whether or not he has worked in support of his organization.

There is no suggestion here that states or other organizations provide *only* public or collective goods. Governments often provide noncollective goods like electric power, for example, and they usually sell such goods on the market much as private firms would do. Moreover, as later parts of this study will argue, large organizations that are not able to make membership compulsory *must also* provide some noncollective goods in order to give potential members an incentive to join. Still, collective goods are the characteristic organizational goods, for ordinary noncollective goods can always be provided by individual action, and only where common purposes or collective goods are concerned is organization or group action ever indispensable.<sup>24</sup>

### C. THE TRADITIONAL THEORY OF GROUPS

There is a traditional theory of group behavior that implicitly assumes that private groups and associations operate according to principles entirely different from those that govern the relationships among firms in the marketplace or between taxpayers and the state. This "group theory" appears to be one of the principal concerns of many political scientists in the United States, as well as a major preoccupation of many sociologists and social psychologists.<sup>25</sup> This traditional theory of groups, like most other theories, has been developed by different writers with varying views, and there is accordingly an inevitable injustice in any attempt to give a common

24. It does not, however, follow that organized or coordinated group action is always necessary to obtain a collective good. See section D of this chapter, "Small Groups."

25. For a discussion of the importance of "groups" of various sorts and sizes for the theory of politics, see Verba, *Small Groups and Political Behavior*; Truman, *Governmental Process*; and Bentley, *Process of Government*. For examples of the type of research and theory about groups in social psychology and sociology, see *Group Dynamics*, ed. Cartwright and Zander, and *Small Groups*, ed. Hare, Borgatta, and Bales.

treatment to these different views. Still, the various exponents of the traditional understanding of groups do have a common relationship to the approach developed in the present study. It is therefore appropriate to speak here in a loose way of a single traditional theory, provided that a distinction is drawn between the two basic variants of this theory: the casual variant and the formal variant.

In its most casual form, the traditional view is that private organizations and groups are ubiquitous, and that this ubiquity is due to a fundamental human propensity to form and join associations. As the famous Italian political philosopher Gaetano Mosca puts it, men have an "instinct" for "herding together and fighting with other herds." This "instinct" also "underlies the formation of all the divisions and subdivisions . . . that arise within a given society and occasion moral and, sometimes, physical conflicts."<sup>26</sup> Aristotle may have had some similar gregarious faculty in mind when he said that man was by nature a political animal.<sup>27</sup> The ubiquitous and inevitable character of group affiliation was emphasized in Germany by Georg Simmel, in one of the classics of sociological literature,<sup>28</sup> and in America by Arthur Bentley, in one of the best-known works on political science.<sup>29</sup> This universal joining tendency or propensity is often thought to have reached its highest intensity in the United States.<sup>30</sup>

The formal variant of the traditional view also emphasizes the universality of groups, but does not begin with any "instinct" or "tendency" to join groups. Instead it attempts to explain the associations and group affiliations of the present day as an aspect of the evolution of modern, industrial societies out of the "primitive" societies that preceded them. It begins with the fact that "primary groups"<sup>31</sup>—groups so small that each of the members has face-to-face

26. *The Ruling Class* (New York: McGraw-Hill, 1939), p. 163.

27. *Politics* i.2.9.1253a. Many others have also emphasized the human propensity towards groups; see Coyle, *Social Process in Organized Groups*; Robert Lowie, *Social Organization* (New York: Rinehart & Co., 1948); Truman, especially pp. 14-43.

28. Georg Simmel, *Conflict and the Web of Group Affiliations*, trans. Kurt Wolff and Reinhard Bendix (Glencoe, Ill.: Free Press, 1950).

29. Bentley, *Process of Government*.

30. Alexis de Tocqueville, *Democracy in America* (New York: New American Library, 1956), p. 198; James Bryce, *The American Commonwealth*, 4th ed. (New York: Macmillan, 1910), pp. 281-282; Charles A. Beard and Mary R. Beard, *The Rise of American Civilization*, rev. ed. (New York: Macmillan, 1949), pp. 761-762; and Daniel Bell, *The End of Ideology* (Glencoe, Ill.: Free Press, 1960), esp. p. 30.

31. Charles H. Cooley, *Social Organization* (New York: Charles Scribner's Sons,

relationships with the others—like family and kinship groups are predominant in primitive societies. As Talcott Parsons contends, "it is well-known that in many primitive societies there is a sense in which kinship 'dominates' the social structure; there are few concrete structures in which participation is independent of kinship status."<sup>32</sup> Only small family or kinship type units represent the interests of the individual. R. M. MacIver describes it this way in the *Encyclopaedia of the Social Sciences*: "Under more simple conditions of society the social expression of interests was mainly through caste or class groups, age groups, kin groups, neighborhood groups, and other unorganized or loosely organized solidarities."<sup>33</sup> Under "primitive" conditions the small, family-type units account for all or almost all human "interaction."

But, these social theorists contend, as society develops, there is structural differentiation: new associations emerge to take on some of the functions that the family had previously undertaken. "As the social functions performed by the family institution in our society have declined, some of these secondary groups, such as labor unions, have achieved a rate of interaction that equals or surpasses that of certain of the primary groups."<sup>34</sup> In Parsons' words, "It is clear that in the more 'advanced' societies a far greater part is played by non-kinship structures like states, churches, the larger business firms, universities and professional societies . . . The process by which non-kinship units become of prime importance in the social structure inevitably entails 'loss of function' on the part of some or even all of the kinship units."<sup>35</sup> If this is true, and if, as MacIver claims, "the most marked structural distinction between a primitive society and a civilized society is the paucity of specific associations in the one

1909), p. 23; George C. Homans, *The Human Group* (New York: Harcourt, Brace, 1950), p. 1; Verba, pp. 11-16.

32. Talcott Parsons and Robert F. Bales, *Family* (Glencoe, Ill.: Free Press, 1955), p. 9; see also Talcott Parsons, Robert F. Bales, and Edward A. Shils, *Working Papers in the Theory of Action* (Glencoe, Ill.: Free Press, 1953).

33. MacIver in *Encyclopaedia of the Social Sciences*, VII, 144-148, esp. 147. See also Truman, p. 25.

34. Truman, pp. 35-36; see also Eliot Chapple and Carlton Coon, *Principles of Anthropology* (New York: Henry Holt, 1942), pp. 443-462.

35. Parsons and Bales, p. 9. See also Bernard Barber, "Participation and Mass Apathy in Associations," in *Studies in Leadership*, ed. Alvin W. Gouldner (New York: Harper, 1950), pp. 477-505, and Neil J. Smelser, *Social Change in the Industrial Revolution* (London: Routledge & Kegan Paul, 1959).

and their multiplicity in the other,"<sup>36</sup> then it would seem that the large association in the modern society is in some sense an equivalent of the small group in the primitive society, and that the large, modern association and the small, primitive group must be explained in terms of the same fundamental source or cause.<sup>37</sup>

What then is the fundamental source which accounts alike for the small primary groups in primitive societies and the large voluntary association of modern times? This the advocates of the formal variant of the theory have left implicit and unclear. It could be the supposed "instinct" or "tendency" to form and join associations, which is the hallmark of the casual variant of the traditional view; this predilection for forming and joining groups would then manifest itself in small family and kinship groups in primitive societies and in large voluntary associations in modern societies. This interpretation would however probably be unfair to many of the theorists who subscribe to the formal variant of the traditional theory, for many of them doubtless would not subscribe to any theory of "instincts" or "propensities." They are no doubt aware that no explanation whatever is offered when the membership of associations or groups is said to be due to an "instinct" to belong; this merely adds a word, not an explanation. Any human action can be ascribed to an instinct or propensity for that kind of action, but this adds nothing to our knowledge. If instincts or propensities to join groups are ruled out as meaningless, what then could be the source of the ubiquitous groups and associations, large and small, posited by the traditional theory? Probably some of the traditional theorists were thinking in "functional" terms—that is of the functions that groups or associations of different types and sizes can perform. In primitive societies small primary groups prevailed because they were best suited (or at

36. MacIver in *Encyclopaedia of the Social Sciences*, VII, 144-148, esp. 147. See also Louis Wirth, "Urbanism as a Way of Life," *American Journal of Sociology*, XLIV (July 1938), 20; Walter Firey, "Coalition and Schism in a Regional Conservation Program," *Human Organization*, XV (Winter 1957), 17-20; Herbert Goldhamer, "Social Clubs," in *Development of Collective Enterprise*, ed. Seba Eldridge (Lawrence: University of Kansas Press, 1943), p. 163.

37. For a different interpretation of the voluntary association see Oliver Garneau, *The Political Life of the American Medical Association* (Cambridge, Mass.: Harvard University Press, 1941), p. 3: "With the advent of political intervention and control, particularly over the economy, it became evident that the formation of policy could not be confined to ballot or legislature. To fill the gap the voluntary group was resorted to, not only by the individual who felt himself alone, but by the government which felt itself ignorant."

least sufficient) to perform certain functions for the people of these societies; in modern societies, by contrast, large associations are supposed to predominate because in modern conditions they alone are capable of performing (or are better able to perform) certain useful functions for the people of these societies. The large voluntary association, for example, could then be explained by the fact that it performed a function—that is, satisfied a demand, furthered an interest, or met a need—for some large number of people that small groups could not perform (or perform so well) in modern circumstances. This demand or interest provides an incentive for the formation and maintenance of the voluntary association.

It is characteristic of the traditional theory in all its forms that it assumes that participation in voluntary associations is virtually universal, and that small groups and large organizations tend to attract members for the same reasons. The casual variant of the theory assumed a propensity to belong to groups without drawing any distinctions between groups of different size. Though the more sophisticated variant may be credited with drawing a distinction between those functions that can best be served by small groups and those that can best be served by large associations, it nonetheless assumes that, when there is a need for a large association, a large association will tend to emerge and attract members, just as a small group will when there is a need for a small group. Thus in so far as the traditional theory draws any distinction at all between small and large groups, it is apparently with respect to the scale of the functions they perform, not the extent they succeed in performing these functions or their capacity to attract members. It assumes that small and large groups differ in degree, but not in kind.

But is this true? Is it really the case that small, primary groups and large associations attract members in the same way, that they are about equally effective in performing their functions, or that they differ only in size but not in their basic character? This traditional theory is called into question by the empirical research which shows that the average person does *not* in fact typically belong to large voluntary associations and that the allegation that the typical American is a "joiner" is largely a myth.<sup>38</sup> It is therefore worth

38. Murray Hausknecht, *The Joiners—A Sociological Description of Voluntary Association Membership in the United States* (New York: Bedminster Press, 1962); Mirra Komarovsky, "The Voluntary Associations of Urban Dwellers," *American*

asking if it is really true that there is no relation between the size of a group and its coherence, or effectiveness, or appeal to potential members; and whether there is any relation between the size of a group and the individual incentives to contribute toward the achievement of group goals. These are questions which must be answered before the traditional theory of groups can be properly assessed. What needs to be known, in the words of the German sociologist Georg Simmel, is "the bearing which the number of sociated individuals has upon the form of social life."<sup>39</sup>

One obstacle, it would seem, to any argument that large and small groups operate according to fundamentally different principles, is the fact, emphasized earlier, that any group or organization, large or small, works for some collective benefit that by its very nature will benefit all of the members of the group in question. Though all of the members of the group therefore have a common interest in obtaining this collective benefit, they have no common interest in paying the cost of providing that collective good. Each would prefer that the others pay the entire cost, and ordinarily would get any benefit provided whether he had borne part of the cost or not. If this is a fundamental characteristic of all groups or organizations with an economic purpose, it would seem unlikely that large organizations would be much different from small ones, and unlikely that there is any more reason that a collective service would be provided for a small group than a large one. Still, one cannot help but feel intuitively that sufficiently small groups would sometimes provide themselves with public goods.

This question cannot be answered satisfactorily without a study of the costs and benefits of alternative courses of action open to individuals in groups of different sizes. The next section of this chapter contains such a study. The nature of this question is such that some of the tools of economic analysis must be used. The following section contains a small amount of mathematics which, though extremely rudimentary, might naturally still be unclear to readers who have never studied that subject. Some points in the following section,

*Sociological Review*, XI (December 1946), 686–698; Floyd Dotson, "Patterns of Voluntary Membership Among Working Class Families," *American Sociological Review*, XVI (October 1951), 687; John C. Scott, Jr., "Membership and Participation in Voluntary Associations," *American Sociological Review*, XXII (June 1957), 315.

39. Georg Simmel, *The Sociology of Georg Simmel*, trans. Kurt H. Wolff (Glencoe, Ill.: Free Press [1950]), p. 87.

moreover, refer to oligopolistic groups in the marketplace, and the references to oligopoly may interest only the economist. Accordingly, some of the highlights of the following section are explained in an intuitively plausible, though loose and imprecise, way in the "non-technical summary" of section D, for the convenience of those who might wish to skip the bulk of the following section.

#### D. SMALL GROUPS

The difficulty of analyzing the relationship between group size and the behavior of the individual in the group is due partly to the fact that each individual in a group may place a different value upon the collective good wanted by his group. Each group wanting a collective good, moreover, faces a different cost function. One thing that will hold true in every case, however, is that the total cost function will be rising, for collective goods are surely like non-collective goods in that the more of the good taken, the higher total costs will be. It will, no doubt, also be true in virtually all cases that there will be significant initial or fixed costs. Sometimes a group must set up a formal organization before it can obtain a collective good, and the cost of establishing an organization entails that the first unit of a collective good obtained will be relatively expensive. And even when no organization or coordination is required, the lumpiness or other technical characteristics of the public goods themselves will ensure that the first unit of a collective good will be disproportionately expensive. Any organization will surely also find that as its demands increase beyond a certain point, and come to be regarded as "excessive," the resistance and the cost of additional units of the collective good rise disproportionately. In short, cost ( $C$ ) will be a function of the rate or level ( $T$ ) at which the collective good is obtained ( $C = f(T)$ ), and the average cost curves will have the conventional U shape.

One point is immediately evident. If there is some quantity of a collective good that can be obtained at a cost sufficiently low in relation to its benefit that some one person in the relevant group would gain from providing that good all by himself, then there is some presumption that the collective good will be provided. The total gain would then be so large in relation to the total cost that some one individual's share would exceed the total cost.

An individual will get some share of the total gain to the group,

a share that depends upon the number in the group and upon how much the individual will benefit from that good in relation to the others in the group. The total gain to the group will depend upon the rate or level at which the collective good is obtained ( $T$ ), and the "size" of the group ( $S_g$ ), which depends not only upon the number of individuals in the group, but also on the value of a unit of the collective good to each individual in the group. This could be illustrated most simply by considering a group of property owners lobbying for a property tax rebate. The total gain to the group would depend upon the "size" ( $S_g$ ) of the group, that is, the total assessed valuation of all the group property, and the rate or level ( $T$ ) of tax rebate per dollar of assessed valuation of property. The gain to an individual member of the group would depend upon the "fraction" ( $F_i$ ) of the group gain he got.

The group gain ( $S_g T$ ) could also be called  $V_g$ , for "value" to the group, and the gain to the individual  $V_i$ , for "value" to the individual. The "fraction" ( $F_i$ ) would then equal  $V_i/V_g$ , and the gain to the individual would be  $F_i S_g T$ . The advantage ( $A_i$ ) that any individual  $i$  would get from obtaining any amount of the collective or group good would be the gain to the individual ( $V_i$ ) minus the cost ( $C$ ).

What a group does will depend on what the individuals in that group do, and what the individuals do depends on the relative advantages to them of alternative courses of action. So the first thing to do, now that the relevant variables have been isolated, is to consider the individual gain or loss from buying different amounts of the collective good. This will depend on the way the advantage to the individual ( $A_i = V_i - C$ ) changes with changes in  $T$ , that is, on

$$dA_i/dT = dV_i/dT - dC/dT.$$

For a maximum,  $dA_i/dT = 0$ .<sup>40</sup> Since  $V_i = F_i S_g T$ , and  $F_i$  and  $S_g$  are, for now, assumed constant,<sup>41</sup>

$$\begin{aligned} d(F_i S_g T)/dT - dC/dT &= 0 \\ F_i S_g - dC/dT &= 0. \end{aligned}$$

40. The second-order conditions for a maximum must also be satisfied; that is,  $d^2 A_i/dT^2 < 0$ .

41. In cases where  $F_i$  and  $S_g$  are not constant, the maximum is given when:

$$\begin{aligned} d(F_i S_g T)/dT - dC/dT &= 0 \\ F_i S_g + F_i T(dS_g/dT) + S_g T(dF_i/dT) - dC/dT &= 0. \end{aligned}$$

This indicates the amount of the collective good that an individual acting independently would buy, if he were to buy any. This result can be given a general, common-sense meaning. Since the optimum point is found when

$$dA_i/dT = dV_i/dT - dC/dT = 0$$

and since  $dV_i/dT = F_i(dV_g/dT)$

$$\begin{aligned} F_i(dV_g/dT) - dC/dT &= 0 \\ F_i(dV_g/dT) &= dC/dT. \end{aligned}$$

This means that the optimal amount of a collective good for an individual to obtain, if he should obtain any, is found when the rate of gain to the group, multiplied by the fraction of the group gain the individual gets, equals the rate of increase of the total cost of the collective good. In other words, the rate of gain to the group ( $dV_g/dT$ ) must exceed the rate of increase in cost ( $dC/dT$ ) by the same multiple that the group gain exceeds the gain to the individual concerned ( $1/F_i = V_g/V_i$ ).<sup>42</sup>

But what matters most is *not* how much of the collective good will be provided if some is provided, but rather whether *any* of the collective good will be provided. And it is clear that, at the optimum point for the individual acting independently, the collective or group good will presumably be provided if  $F_i > C/V_i$ .

For if

$$\begin{aligned} F_i &> C/V_i \\ V_i/V_g &> C/V_g \end{aligned}$$

then

$$V_i > C.$$

Thus, if  $F_i > C/V_g$ , the gain to an individual from seeing that the collective good is provided will exceed the cost. This means there is a presumption that the collective good will be provided if the cost of the collective good is, at the optimal point for any individual in the group, so small in relation to the gain of the group as a whole

42. The same point could be made by focusing attention on the individual's cost and benefit functions alone, and neglecting the gains to the group. But this would divert attention from the main purpose of the analysis, which is studying the relation between the size of the group and the likelihood that it will be provided with a collective good.

from that collective good, that the total gain exceeds the total cost by as much as or more than the gain to the group exceeds the gain to the individual.

In summary, then, the rule is that there is a presumption that a collective good will be provided if, when the gains to the group from the collective good are increasing at  $1/F_i$  times the rate of increase in the total cost of providing that good (that is, when  $dV_g/dT = 1/F_i(dC/dT)$ ), the total benefit to the group is a larger multiple of the cost of that good than the gains to the group are of the gains to the individual in question (that is,  $V_g/C > V_i/V_i$ ).

The degree of generality of the basic idea in the foregoing model can be illustrated by applying it to a group of firms in a market. Consider an industry producing a homogeneous product, and assume that the firms in the industry *independently* seek to maximize profits. For simplicity, suppose also that marginal costs of production are zero. In order to avoid adding any new notational symbols, and to bring out the applicability of the foregoing analysis, assume that  $T$  now stands for price, that  $S_g$  now stands for the physical volume of the group's or industry's sales, and  $S_i$  for the size or physical volume of the sales of firm  $i$ .  $F_i$  still indicates the "fraction" of the total accounted for by the individual firm or member of the group. It indicates now the fraction of the total group or industry sales going to firm  $i$  at any given moment:  $F_i = S_i/S_g$ . The price,  $T$ , will affect the amount sold by the industry to an extent given by the elasticity of demand,  $E$ . The elasticity  $E = -T/S_g(dS_g/dT)$ , and from this a convenient expression for the slope of the demand curve,  $dS_g/dT$ , follows:  $dS_g/dT = -ES_g/T$ . With no production costs, the optimum output for a firm will be given when:

$$\begin{aligned} dA_i/dT &= d(S_i T)/dT = 0 \\ S_i + T(dS_i/dT) &= 0 \\ F_i S_g + T(dS_i/dT) &= 0. \end{aligned}$$

Here, where it is assumed that the firm acts independently, i.e., expects no reaction from other firms,  $dS_i = dS_g$ , so

$$F_i S_g + T(dS_g/dT) = 0$$

and since  $dS_g/dT = -ES_g/T$ ,

$$\begin{aligned} F_i S_g - T(ES_g/T) &= 0 \\ S_g(F_i - E) &= 0. \end{aligned}$$

This can happen only when  $F_i = E$ . Only when the elasticity of demand for the industry is less than or equal to the fraction of the industry's output supplied by a particular firm will that firm have any incentive to restrict its output. A firm that is deciding whether or not to restrict its output in order to bring about a higher price will measure the cost or loss of the foregone output against the gains it gets from the "collective good"—the higher price. The elasticity of demand is a measure of this. If  $F_i$  is equal to  $E$  it means that the elasticity of demand for the industry is the same as the proportion of the output of the industry shared by the firm in question; if the elasticity of demand is, say, 1/4, it means that a 1 per cent reduction in output will bring a 4 per cent increase in price, which makes it obvious that if a given firm has one fourth of the total industry output it should stop increasing, or restrict, its own output. If there were, say, a thousand firms of equal size in an industry, the elasticity of demand for the industry's product would have to be 1/1000 or less before there would be any restriction of output. Thus there are no profits in equilibrium in any industry with a really large number of firms. A profit-maximizing firm will start restricting its output, that is, will start acting in a way consistent with the interests of the industry as a whole, when the rate at which the gain to the group increases, as more  $T$  (a higher price) is provided, is  $1/F_i$  times as great as the rate at which the total cost of output restriction increases. This is the same criterion for group-oriented behavior used in the more general case explained earlier.

This analysis of a market is identical with that offered by Cournot.<sup>43</sup> This should not be surprising, for Cournot's theory is essentially a special case of a more general theory of the relationship between the interests of the member of a group and of the interests of the group as a whole. The Cournot theory can be regarded as a special case of the analysis developed here. The Cournot solution thus boils down to the common-sense statement that a firm will act to keep up the price of the product its industry sells only when the total cost of keeping up the price is not more than its share of the industry's gain from the higher price. The Cournot theory is, like the analysis of group action outside the market, a theory that asks

43. Augustin Cournot, *Researches into the Mathematical Principles of the Theory of Wealth*, trans. Nathaniel T. Bacon (New York: Macmillan, 1897), especially chap. vii, pp. 79-90.

when it is in the interest of an individual unit in a group to act in the interest of the group as a whole.

The Cournot case is in one respect simpler than the group situation outside the marketplace that is the main concern of this study. When a group seeks an ordinary collective good, rather than a higher price through output restriction, it finds, as the opening paragraph of this section argued, that the first unit of the collective good obtained will be more expensive per unit than some subsequent units of the good. This is because of the lumpiness and other technical characteristics of collective goods, and because it may sometimes be necessary to create an organization to obtain the collective good. This calls to attention the fact that there are two distinct questions that an individual in a nonmarket group must consider. One is whether the total benefit he would get from providing some amount of the collective good would exceed the total cost of that amount of the good. The other question is how much of the collective good he should provide, if some should be provided, and the answer here depends of course on the relationship between marginal, rather than total, costs and benefits.

There are similarly also two distinct questions that must be answered about the group as a whole. It is not enough to know whether a small group will provide itself with a collective good; it is also necessary to determine whether the amount of the collective good that a small group will obtain, if it obtains any, will tend to be Pareto-optimal for the group as a whole. That is, will the group gain be maximized? The optimal amount of a collective good for a group as a whole to obtain, if it should obtain any, would be given when the gain to the group was increasing at the same rate as the cost of the collective good, i.e., when  $dV_g/dT = dC/dT$ . Since, as shown earlier, each individual in the group would have an incentive to provide more of the collective good until  $F_i(dV_g/dT = dC/dT)$ , and since  $\Sigma F_i = 1$ , it would at first glance appear that the sum of what the individual members acting independently would provide would add up to the group optimum. It would also seem that each individual in the group would then bear a fraction,  $F_i$ , of the total burden or cost, so that the burden of providing the public good would be shared in the "right" way in the sense that the cost would be shared in the same proportion as the benefits.

But this is not so. Normally, the provision of the collective good will be strikingly suboptimal and the distribution of the burden will

be highly arbitrary. This is because the amount of the collective good that the individual obtains for himself will automatically also go to others. It follows from the very definition of a collective good that an individual cannot exclude the others in the group from the benefits of that amount of the public good that he provides for himself.<sup>44</sup> This means that no one in the group will have an incentive independently to provide any of the collective good once the amount that would be purchased by the individual in the group with the largest  $F_i$  was available. This suggests that, just as there is a tendency for large groups to fail to provide themselves with any collective good at all, so *there is a tendency in small groups toward a suboptimal provision of collective goods*. The suboptimality will be the more serious the smaller the  $F_i$  of the "largest" individual in the group. Since the larger the number in the group, other things equal, the smaller the  $F_i$ 's will be, the more individuals in the group, the more serious the suboptimality will be. Clearly then groups with larger numbers of members will generally perform less efficiently than groups with smaller numbers of members.

It is not, however, sufficient to consider only the number of individuals or units in a group, for the  $F_i$  of any member of the group will depend not only on how many members there are in the group, but also on the "size" ( $S_i$ ) of the individual member, that is, the extent to which he will be benefited by a given level of provision of the collective good. An owner of vast estates will save more from a given reduction in property taxes than the man with only a modest cottage, and other things equal will have a larger  $F_i$ .<sup>45</sup> A group com-

44. In the rest of this section it is convenient and helpful to assume that every member of the group receives the same amount of the public good. This is in fact the case whenever the collective good is a "pure public good" in Samuelson's sense. This assumption is, however, more stringent than is usually necessary. A public good may be consumed in unequal amounts by different individuals, yet be a full public good in the sense that one individual's consumption does not in any way diminish that of another. And even when additional consumption by one individual does lead to marginal reductions in the amount available to others, the qualitative conclusions that there will be suboptimality and disproportionate burden sharing still hold.

45. Differences in size can also have some importance in market situations. The large firm in a market will get a larger fraction of the total benefit from any higher price than a small firm, and will therefore have more incentive to restrict output. This suggests that the competition of a few large firms among the many small ones, contrary to some opinions, can lead to a serious misallocation of resources. For a different view on this subject, see Willard D. Arant, "The Competition of the Few among the Many," *Quarterly Journal of Economics*, LXX (August 1956), 327-345.

posed of members of unequal  $S_i$ , and, therefore, unequal  $F_i$ , will show less of a tendency toward suboptimality (and be more likely to provide itself with some amount of a collective good) than an otherwise identical group composed of members of equal size.

Since no one has an incentive to provide any more of the collective good, once the member with the largest  $F_i$  has obtained the amount he wants, it is also true that the distribution of the burden of providing the public good in a small group will *not* be in proportion to the benefits conferred by the collective good. The member with the largest  $F_i$  will bear a disproportionate share of the burden.<sup>46</sup> Where small groups with common interests are concerned, then, *there is a systematic tendency for "exploitation"<sup>47</sup> of the great by the small!*

The behavior of small groups interested in collective goods can sometimes be quite complex—much more complex than the preceding paragraphs would suggest.<sup>48</sup> There are certain institutional

46. The discussion in the text is much too brief and simple to do full justice even to some of the most common situations. In what is perhaps the most common case, where the collective good is *not* a money payment to each member of some group, and not something that each individual in the group can sell for money, the individuals in the group must compare the additional cost of another unit of the collective good with the additional "utility" they would get from an additional unit of that good. They could not, as the argument in the text assumes, merely compare a money cost with a money return, and indifference curves would accordingly also have to be used in the analysis. The marginal rate of substitution would be affected not only by the fact that the taste for additional units of the collective good would diminish as more of the good was consumed, but also by the income effects. The income effects would lead a group member that had sacrificed a disproportionate amount of his income to obtain the public good to value his income more highly than he would have done had he got the collective good free from others in the group. Conversely, those who had not borne any of the burden of providing the collective good they enjoyed would find their real incomes greater, and unless the collective good were an inferior good, this gain in real income would strengthen their demand for the collective good. These income effects would tend to keep the largest member of the group from bearing *all* of the burden of the collective good (as he would in the much too simple case considered in the text). I am thankful to Richard Zeckhauser for bringing the importance of income effects in this context to my attention.

47. The moral overtones of the word "exploitation" are unfortunate; no general moral conclusions can follow from a purely logical analysis. Since the word "exploitation" is, however, commonly used to describe situations where there is a disproportion between the benefits and sacrifices of different people, it would be pedantic to use a different word here.

48. For one thing, the argument in the text assumes independent behavior, and thus neglects the strategic interaction or bargaining that is possible in small groups. As later parts of this chapter will show, strategic interaction is usually much less important in nonmarket groups seeking collective goods than it is among groups of firms in the marketplace. And even when there is bargaining, it will often be

arrangements and behavioral assumptions that will not always lead to the suboptimality and disproportionality that the preceding paragraphs have described. Any adequate analysis of the tendency toward suboptimal provision of collective goods, and toward disproportionate sharing of the burdens of providing them, would be too long to fit comfortably into this study, which is concerned mainly with large groups, and brings in small groups mainly for purposes of comparison and contrast. The problem of small groups seeking collective goods is of some importance, both theoretically<sup>49</sup> and practically, and has not been adequately treated in the literature. It will accordingly be analyzed in more detail in forthcoming articles. The Nontechnical Summary of this section will list a few of the specific cases that this approach to small groups and organizations can be used to study.

The necessary conditions for the optimal provision of a collective good, through the voluntary and independent action of the members of a group, can, however, be stated very simply. The marginal cost of additional units of the collective good must be shared in exactly the same proportion as the additional benefits. Only if this is done will each member find that his own marginal costs and benefits are

---

the case that there will be a disparity of bargaining power that will lead to about the same results as are described in the text. When a group member with a large  $F_i$  bargains with a member with a small  $F_i$ , all he can do is threaten the smaller member by saying, in effect, "If you do not provide more of the collective good, I will provide less myself, and you will then be worse off than you are now." But when the large member restricts his purchase of the public good, he will suffer more than the smaller member, simply because his  $F_i$  is greater. The large member's threat is thus not apt to be credible. Another factor that works in the same direction is that the maximum amount of collective good provision that a successful bargain can extract from the small member is less than the amount a successful bargain can bring forth from the large member. This means that the large member may not gain enough even from successful bargaining to justify the risks or other costs of bargaining, while the small member by contrast finds that the gain from a successful bargain is large in relation to his costs of bargaining. The bargaining problem is of course more complex than this, but it is nonetheless clear that bargaining will usually lead toward the same results as the forces explained in the text.

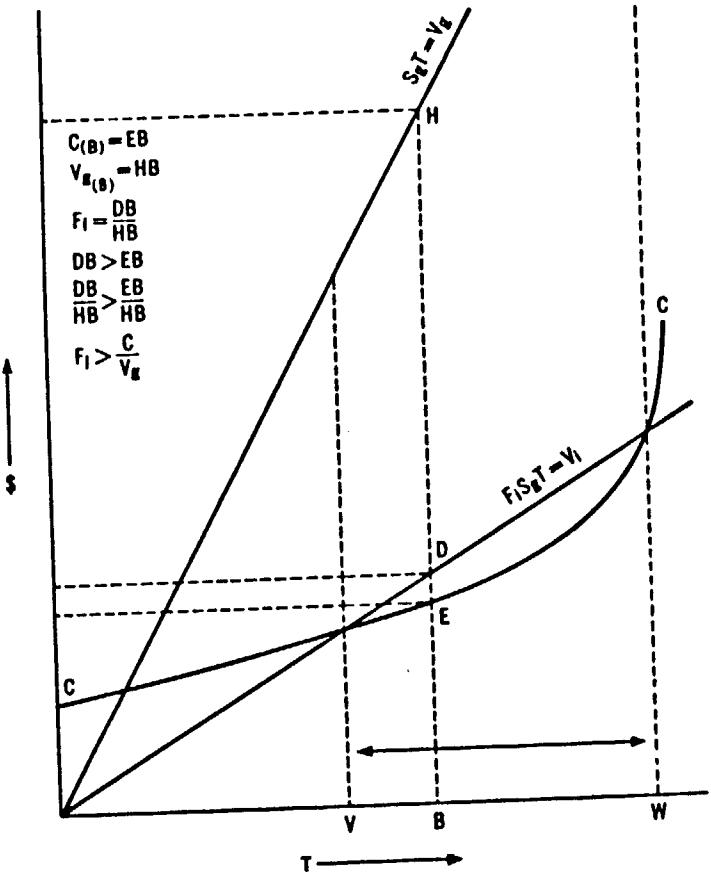
49. Erik Lindahl's famous "voluntary theory of public exchange" can, I believe, usefully be amended and expanded with the aid of the analysis adumbrated in the text. I am thankful to Richard Musgrave for bringing to my attention the fact that Lindahl's theory and the approach used in this study must be closely related. He sees this relationship in a different way, however. For analyses of Lindahl's theory see Richard Musgrave, "The Voluntary Exchange Theory of Public Economy," *Quarterly Journal of Economics*, LIII (February 1939), 213-237; Leif Johansen, "Some Notes on the Lindahl Theory of Determination of Public Expenditures," *International Economic Review*, IV (September 1963), 346-358; John G. Head, "Lindahl's Theory of the Budget," *Finanzarchiv*, XXIII (October 1964), 421-454.

equal at the same time that the total marginal cost equals the total or aggregate marginal benefit. If marginal costs are shared in any other way, the amount of collective good provided will be *suboptimal*.<sup>50</sup> It might seem at first glance that if some cost allocations lead to a suboptimal provision of a collective good, then some other cost allocations would lead to a supraoptimal supply of that good; but this is not so. In any group in which participation is voluntary, the member or members whose shares of the marginal cost exceed their shares of the additional benefits will stop contributing to the achievement of the collective good *before* the group optimum has been reached. And there is no conceivable cost-sharing arrangement in which some member does not have a marginal cost greater than his share of the marginal benefit, except the one in which every member of the group shares marginal costs in exactly the same proportion in which he shares incremental benefits.<sup>51</sup>

50. There is an illustration of this point in many farm tenancy agreements, where the landlord and tenant often share the produce of the crop in some prearranged proportion. The farm's output can then be regarded as a public good to the landlord and tenant. Often the tenant will provide *all* of the labor, machinery, and fertilizer, and the landlord will maintain *all* of the buildings, drainage, ditches, etc. As some agricultural economists have rightly pointed out, such arrangements are inefficient, for the tenant will use labor, machinery, and fertilizer only up to the point where the marginal cost of these factors of production equals the marginal return from his share of the crop. Similarly, the landlord will provide a suboptimal amount of the factors he provides. The only way in which this suboptimal provision of the factors can be prevented in a share-tenancy is by having the landlord and tenant share the costs of each of the (variable) factors of production in the same proportion in which they share the output. Perhaps this built-in inefficiency in most share-tenancy agreements helps account for the observation that in many areas where farmers do not own the land they farm, land reform is necessary to increase agricultural efficiency. See Earl O. Heady and E. W. Kehlberg, *Effect of Share and Cash Renting on Farming Efficiency* (Iowa Agricultural Experiment Station Bulletin 386), and Earl O. Heady, *Economics of Agricultural Production and Resource Use* (New York: Prentice-Hall, 1952), esp. pp. 592 and 620.

51. A similar argument could sometimes be used to help explain the common observation that there is "public squalor" midst "private splendor," that is, a suboptimal supply of public goods. Such an argument would be relevant at least in those situations where proposed Pareto-optimal public expenditures benefit a group of people smaller than the group that would be taxed to pay for these expenditures. The point that even Pareto-optimal public expenditures usually benefit groups of people smaller than the group taxed to pay for these expenditures was suggested to me by Julius Margolis' useful paper on "The Structure of Government and Public Investment," in *American Economic Review: Papers and Proceedings*, LIV (May 1964), 236-247. See my "Discussion" of Margolis' paper (and others) in the same issue of the *American Economic Review* (pp. 250-251) for a suggestion of a way in which a model of the kind developed in this study can be used to explain private

Though there is a tendency for even the smallest groups to provide suboptimal amounts of a collective good (unless they arrange marginal cost-sharing of the kind just described), the more important point to remember is that some sufficiently small groups can pro-



affluence and public squalor. It is interesting that John Head (*Finanzarchiv*, XXIII, 453-454) and Leif Johansen (*International Economic Review*, IV, 353), though they started out at different points from mine and used instead Lindahl's approach, still had arrived at conclusions on this point that are not altogether different from mine. For interesting arguments that point to forces that could lead to supra-optimal levels of government expenditure, see two other papers in the issue of the *American Economic Review* cited above, namely "Fiscal Institutions and Efficiency in Collective Outlay" (pp. 227-235) by James M. Buchanan, and "Divergencies between Individual and Total Costs within Government" (pp. 243-249) by Roland N. McKean.

vide themselves with some amount of a collective good through the voluntary and rational action of one or more of their members. In this they are distinguished from really large groups. There are two things to determine in finding out whether there is any presumption that a given group will voluntarily provide itself with a collective good. First, the optimal amount of the collective good for each *individual* to buy, if he is to buy any, must be discovered; this is given when  $F_i(dV_i/dT) = dC/dT$ .<sup>52</sup> Second, it must be determined whether any member or members of the group would find at that individual optimum that the benefit to the group from the collective good exceeded the total cost by more than it exceeded the member's own benefit from that collective good; that is, whether  $F_i > C/V_g$ . The argument may be stated yet more simply by saying that, *if at any level of purchase of the collective good, the gain to the group exceeds the total cost by more than it exceeds the gain to any individual, then there is a presumption that the collective good will be provided, for then the gain to the individual exceeds the total cost of providing the collective good to the group*. This is illustrated in the accompanying figure, where an individual would presumably be better off for having provided the collective good, whether he provided amount  $V$  or amount  $W$  or any amount in between. If any amount of the collective good between  $V$  and  $W$  is obtained, even if it is not the optimal amount for the individual,  $F_i$  will exceed  $C/V_g$ .

#### *Nontechnical summary of Section D*

The technical part of this section has shown that certain small groups can provide themselves with collective goods without relying on coercion or any positive inducements apart from the collective good itself.<sup>53</sup> This is because in some small groups each of the mem-

52. If  $F_i$  is not a constant, this individual optimum is given when:  

$$F_i(dV_g/dT) + V_g(dF_i/dT) = dC/dT$$

53. I am indebted to Professor John Rawls of the Department of Philosophy at Harvard University for reminding me of the fact that the philosopher David Hume sensed that small groups could achieve common purposes but large groups could not. Hume's argument is however somewhat different from my own. In *A Treatise of Human Nature*, Everyman edition (London: J. M. Dent, 1952), II, 239, Hume wrote: "There is no quality in human nature which causes more fatal errors in our conduct, than that which leads us to prefer whatever is present to the distant and remote, and makes us desire objects more according to their situation than their intrinsic value. Two neighbours may agree to drain a meadow, which they possess

bers, or at least one of them, will find that his personal gain from having the collective good exceeds the total cost of providing some amount of that collective good; there are members who would be better off if the collective good were provided, even if they had to pay the entire cost of providing it themselves, than they would be if it were not provided. In such situations there is a presumption that the collective good will be provided. Such a situation will exist only when the benefit to the group from having the collective good exceeds the total cost by more than it exceeds the gain to one or more individuals in the group. Thus, in a very small group, where each member gets a substantial proportion of the total gain simply because there are few others in the group, a collective good can often be provided by the voluntary, self-interested action of the members of the group. In smaller groups marked by considerable degrees of inequality—that is, in groups of members of unequal "size" or extent of interest in the collective good—there is the greatest likelihood that a collective good will be provided; for the greater the interest in the collective good of any single member, the greater the likelihood that that member will get such a significant proportion of the total benefit from the collective good that he will gain from seeing that the good is provided, even if he has to pay all of the cost himself.

Even in the smallest groups, however, the collective good will not ordinarily be provided on an optimal scale. That is to say, the members of the group will not provide as much of the good as it would be in their common interest to provide. Only certain special

---

in common: because it is easy for them to know each other's mind; and each must perceive, that the immediate consequence of his failing in his part, is the abandoning of the whole project. But it is very difficult, and indeed impossible, that a thousand persons should agree in any such action; it being difficult for them to concert so complicated a design, and still more difficult for them to execute it; while each seeks a pretext to free himself of the trouble and expense, and would lay the whole burden on others. Political society easily remedies both these inconveniences. Magistrates find an immediate interest in the interest of any considerable part of their subjects. They need consult nobody but themselves to form any scheme for promoting that interest. And as the failure of any one piece in the execution is connected, though not immediately, with the failure of the whole, they prevent that failure, because they find no interest in it, either immediate or remote. Thus, bridges are built, harbours opened, ramparts raised, canals formed, fleets equipped, and armies disciplined, everywhere, by the care of government, which, though composed of men subject to all human infirmities, becomes, by one of the finest and most subtle inventions imaginable, a composition which is in some measure exempted from all these infirmities."

institutional arrangements will give the individual members an incentive to purchase the amounts of the collective good that would add up to the amount that would be in the best interest of the group as a whole. This tendency toward suboptimality is due to the fact that a collective good is, by definition, such that other individuals in the group cannot be kept from consuming it once any individual in the group has provided it for himself. Since an individual member thus gets only part of the benefit of any expenditure he makes to obtain more of the collective good, he will discontinue his purchase of the collective good before the optimal amount for the group as a whole has been obtained. In addition, the amounts of the collective good that a member of the group receives free from other members will further reduce his incentive to provide more of that good at his own expense. Accordingly, *the larger the group, the farther it will fall short of providing an optimal amount of a collective good.*

This suboptimality or inefficiency will be somewhat less serious in groups composed of members of greatly different size or interest in the collective good. In such unequal groups, on the other hand, there is a tendency toward an arbitrary sharing of the burden of providing the collective good. The largest member, the member who would on his own provide the largest amount of the collective good, bears a disproportionate share of the burden of providing the collective good. The smaller member by definition gets a smaller fraction of the benefit of any amount of the collective good he provides than a larger member, and therefore has less incentive to provide additional amounts of the collective good. Once a smaller member has the amount of the collective good he gets free from the largest member, he has more than he would have purchased for himself, and has no incentive to obtain any of the collective good at his own expense. In small groups with common interests there is accordingly *a surprising tendency for the "exploitation" of the great by the small.*

The argument that small groups providing themselves with collective goods tend to provide suboptimal quantities of these goods, and that the burdens of providing them are borne in an arbitrary and disproportionate way, does not hold in all logically possible situations. Certain institutional or procedural arrangements can lead to different outcomes. The subject cannot be analyzed adequately in any brief discussion. For this reason, and because the main focus of this book is on large groups, many of the complexities of small-group

behavior have been neglected in this study. An argument of the kind just outlined could, however, fit some important practical situations rather well, and may serve the purpose of suggesting that a more detailed analysis of the kind outlined above could help to explain the apparent tendency for large countries to bear disproportionate shares of the burdens of multinational organizations, like the United Nations and NATO, and could help to explain some of the popularity of neutralism among smaller countries. Such an analysis would also tend to explain the continual complaints that international organizations and alliances are not given adequate (optimal) amounts of resources.<sup>54</sup> It would also suggest that neighboring local governments in metropolitan areas that provide collective goods (like commuter roads and education) that benefit individuals in two or more local government jurisdictions would tend to provide inadequate amounts of these services, and that the largest local government (e.g., the one representing the central city) would bear disproportionate shares of the burdens of providing them.<sup>55</sup> An analysis of the foregoing type might, finally, provide some additional insight into the phenomenon of price leadership, and particularly the possible disadvantages involved in being the largest firm in an industry.

The most important single point about small groups in the present context, however, is that they may very well be able to provide themselves with a collective good simply because of the attraction of the collective good to the individual members. In this, small groups differ from larger ones. The larger a group is, the farther it will fall short of obtaining an optimal supply of any collective good, and the less likely that it will act to obtain even a minimal amount of such a good. In short, the larger the group, the less it will further its common interests.

#### E. "EXCLUSIVE" AND "INCLUSIVE" GROUPS

The movement in and out of the group must no longer be ignored. This is an important matter; for industries or market groups differ

54. Some of the complexities of behavior in small groups are treated in Mancur Olson, Jr., and Richard Zeckhauser, "An Economic Theory of Alliances," *Review of Economics and Statistics*, XLVIII (August 1966), 266-279, and in "Collective Goods, Comparative Advantage, and Alliance Efficiency," in *Issues of Defense Economics* (A Conference of the Universities-National Bureau-Committee for Economic Research), Roland McKean, ed., (New York: National Bureau of Economic Research, 1967), pp. 25-48. [Footnote added in 1970.]

55. I am indebted to Alan Williams of York University in England, whose study of local government brought the importance of these sorts of spillovers among local governments to my attention.

fundamentally from nonmarket groups in their attitudes toward movement in and out of the group. The firm in an industry wants to keep new firms from coming in to share the market and wants as many as possible of those firms already in the industry to leave the industry. It wants the group of firms in the industry to shrink until there is preferably only one firm in the group: its ideal is a monopoly. Thus the firms in a given market are competitors or rivals. In nonmarket groups or organizations seeking a collective good the opposite is true. Usually the larger the number available to share the benefits and costs the better. An increase in the size of the group does not bring competition to anyone, but may lead to lower costs for those already in the group. The truth of this view is evident from everyday observation. Whereas firms in a market lament any increase in competition, associations that supply collective goods in nonmarket situations almost always welcome new members. Indeed, such organizations sometimes attempt to make membership compulsory.

Why is there this difference between the market and nonmarket groups which previous sections of this chapter have shown to have striking similarities? If the businessman in the market, and the member of the lobbying organization, are alike in that each of them finds that the benefits of any effort made to achieve group goals would accrue mostly to other members of the group, then why are they so much different where entry and exit from the group are concerned? The answer is that in a market situation the "collective good"—the higher price—is such that if one firm sells more at that price, other firms must sell less, so that the benefit it provides is fixed in supply; but in nonmarket situations the benefit from a collective good is not fixed in supply. Only so many units of a product can be sold in any given market without driving down the price, but any number of people can join a lobbying organization without necessarily reducing the benefits for others.<sup>56</sup> Usually in a market situation what one firm captures another firm cannot obtain; essentially in a nonmarket situation what one consumes another may also enjoy. If a firm in a market situation prospers, it becomes a more formidable rival; but if an individual in a nonmarket group prospers, he may

56. In a social club that gives members status because it is "exclusive," the collective good in question is like a suprarevenue price in a market, not like the normal nonmarket situation. If the top "400" were to become the top "4000," the benefits to the entrants would be offset by the losses of old members, who would have traded an exalted social connection for one that might be only respectable.

well then have an incentive to pay a larger share of the cost of the collective good.

Because of the fixed and thus limited amount of the benefit that can be derived from the "collective good"—the higher price—in the market situation, which leads the members of a market group to attempt to reduce the size of their group, this sort of collective good will here be called an "exclusive collective good."<sup>57</sup> Because the supply of collective goods in nonmarket situations, by contrast, automatically expands when the group expands, this sort of public good should be called an "inclusive collective good."<sup>58</sup>

57. This usage of the idea of the collective good is, to be sure, in some respects over-broad, in that the collective-good concept is not needed to analyze market behavior; other theories are usually better for that purpose. But it is helpful in this particular context to treat a supracompetitive price as a special type of collective good. It is a useful expositional technique for bringing out parallels and contrasts between market and nonmarket situations with respect to the relationships between individual interests and group-oriented action. I hope that in the following pages it will also offer some insight into organizations that have functions both inside and outside the market, and into the extent of bargaining in market and nonmarket groups.

58. There are some interesting parallels between my concepts of "exclusive" and "inclusive" collective goods and some recent work by other economists. There is, first, a relationship between these concepts and John Head's previously cited article on "Public Goods and Public Policy" (*Public Finance*, XVII, 197-219). I did not understand all of the implications of my discussion of inclusive and collective goods until I had read all of Head's article. As I now see it, these concepts can be explained in terms of his distinction between the two defining characteristics of the traditional public good: infeasibility of exclusion and jointness of supply. My exclusive collective good is then a good such that, at least within some given group, exclusion is not feasible, but at the same time such that there is no jointness of supply whatever, so that the members of the group hope that others can be kept out of the group. My inclusive collective good is also such that exclusion is infeasible, at least within some given group, but it is however also characterized by at least some considerable degree of jointness in supply, and this accounts for the fact that additional members can enjoy the good with little or no reduction in the consumption of the old members.

There is, second, a relationship between my inclusive-exclusive distinction and a paper by James M. Buchanan entitled "An Economic Theory of Clubs" (mime.). Buchanan's paper assumes that exclusion is possible, but that a (severely limited) degree of jointness in supply exists, and shows that on these assumptions the optimal number of users of a given public good is normally finite, will vary from case to case, and may sometimes be quite small. Buchanan's approach and my own are related in that both of us ask how the interests of a member of a group enjoying a collective good will be affected by increases or decreases in the number of people who consume the good. Both of us have been working on this problem independently, and until very recently in ignorance of each other's interest in this particular question. Buchanan generously suggests that I may have asked this question earlier than he did, but whereas I have barely touched upon the question merely to facilitate other parts of my general argument, he has developed an interesting and general model which shows the relevance of this question to a wide range of policy problems.

Whether a group behaves exclusively or inclusively, therefore, depends upon the nature of the objective the group seeks, not on any characteristics of the membership. Indeed, the same collection of firms or individuals might be an exclusive group in one context and an inclusive group in another. The firms in an industry would be an exclusive group when they sought a higher price in their industry by restricting output, but they would be an inclusive group, and would enlist all the support they could get, when they sought lower taxes, or a tariff, or any other change in government policy. The point that the exclusiveness or inclusiveness of a group depends on the objective at issue, rather than on any traits of the membership, is important, since many organizations operate both in the market to raise prices by restricting output, and also in the political and social systems to further other common interests. It might be interesting, if space permitted, to study such groups with the aid of the distinction between exclusive and inclusive collective goods. The logic of this distinction suggests that such groups would have ambivalent attitudes toward new entrants. And in fact they do. Labor unions, for example, sometimes advocate the "solidarity of the working class" and demand the closed shop, yet set up apprenticeship rules that limit new "working class" entrants into particular labor markets. Indeed, this ambivalence is a fundamental factor with which any adequate analysis of what unions seek to maximize must deal.<sup>59</sup>

A further difference between inclusive and exclusive groups is evident when formally organized, or even informally coordinated,

59. There is some uncertainty about what unions maximize. It is sometimes thought that unions do not maximize wage rates, since higher wages reduce the quantity of labor demanded by the employer and thereby also union membership. This reduction in membership is in turn contrary to the institutional interests of the union and harmful to the power and prestige of the union leaders. Yet some unions, such as the United Mine Workers, have in fact raised wages to a point they conceded would reduce employment in their industry. One possible explanation is that unions seek inclusive collective goods from government, as well as higher wages in the market. In this nonmarket capacity each union has an interest in acquiring new members, outside its "own" industry or craft as well as inside it. Higher wages do not hinder the expansion of a union in other industries or skill categories. Indeed, the higher the wages a union wins in any given labor market the greater the prestige of its leaders and the greater its appeal to workers in other labor markets, thus facilitating the growth of union membership outside its original clientele. This is something a union may be happy to do because this will help it fulfill its political, lobbying function. Interestingly, the CIO, and the catch-all District 50 of the UMW, may possibly have allowed the influence of John L. Lewis and the UMW to expand at some times when union wage levels limited employment in coal mining. I am thankful to one of my former students, John Beard, for stimulating ideas on this point.

behavior is attempted. When there is organized or coordinated effort in an inclusive group, as many as can be persuaded to help will be included in that effort.<sup>60</sup> Yet it will not (except in marginal cases, where the collective good is only just worth its cost) be essential that every individual in the group participate in the organization or agreement. In essence this is because the nonparticipant normally does not take the benefits of an inclusive good away from those who do cooperate. An inclusive collective good is by definition such that the benefit a noncooperator receives is not matched by corresponding losses to those who do cooperate.<sup>61</sup>

When a group seeks an exclusive collective good through an agreement or organization of the firms in the market—that is, if there is explicit or even tacit collusion in the market—the situation is much different. In such a case, though the hope is that the number of firms in the industry will be as small as possible, it is paradoxically almost always essential that there be 100 per cent participation of those who

60. Riker's interesting argument, in *The Theory of Political Coalitions*, that there will be a tendency toward minimum winning coalitions in many political contexts, does not in any way weaken the conclusion here that inclusive groups try to increase their membership. Nor does it weaken any of the conclusions in this book, for Riker's argument is relevant only to zero-sum situations, and no such situations are analyzed in this book. Any group seeking an inclusive collective good would not be in a zero-sum situation, since the benefit by definition increases in amount as more join the group, and as more of the collective good is provided. Even groups seeking exclusive collective goods do not fit Riker's model, for though the amount that can be sold at any given price is fixed, the amount the price will be raised and thus the gain to the group are variable. It is unfortunate that Riker's otherwise stimulating and useful book considers some phenomena, like military alliances, for which his zero-sum assumption is most inappropriate. See William H. Riker, *The Theory of Political Coalitions* (New Haven, Conn.: Yale University Press, 1962).

61. If the collective good were a "pure public good" in Samuelson's sense, the benefit the noncooperator receives would not only not lead to a corresponding loss to those who did cooperate; it would not lead to any loss whatever for them. The pure-public-good assumption seems, however, to be unnecessarily stringent for present purposes. It would surely often be true that after some point, additional consumers of a collective good would, however slightly, reduce the amount available to others. The argument in the text therefore does not require that inclusive collective goods be pure public goods. When an inclusive collective good is not a pure public good, however, those in the group enjoying the good would not welcome additional members who failed to pay adequate dues. Dues would not be adequate unless they were at least equal in value to the reduction in the consumption of the old members entailed by the consumption of the new entrant. As long as any significant degree of "jointness in supply" remains, however, the gains to new entrants will exceed the dues payment needed to ensure that the old members will be adequately compensated for any curtailment in their own consumption, so the group will remain truly "inclusive."

remain in the group. In essence this is because even one nonparticipant can usually take all the benefits brought about by the action of the collusive firms for himself. Unless the costs of the nonparticipating firm rise too rapidly with increases in output,<sup>62</sup> it can continually expand its output to take advantage of the higher price brought about by the collusive action until the collusive firms, if they foolishly continue to maintain the higher price, have reduced their output to zero, all for the benefit of the nonparticipating firm. The non-participating firm can deprive the collusive firms of all the benefits of their collusion because the benefit of any given suprareactive price is fixed in amount; so whatever he takes the collusive firms lose. There is then an all-or-none quality about exclusive groups, in that there must often be either 100 per cent participation or else no collusion at all. This need for 100 per cent participation has the same effects in an industry that a constitutional provision that all decisions must be unanimous has in a voting system. Whenever unanimous participation is required, any single holdout has extraordinary bargaining power; he may be able to demand for himself most of the gain that would come from any group-oriented action.<sup>63</sup> Moreover, any one in the group can attempt to be a holdout, and demand a greater share of the gain in return for his (indispensable) support. This incentive to holdouts makes any group-oriented action less likely than it would otherwise be. It also implies that each member has a great incentive to bargain; he may gain all by a good bargain, or lose all in a bad one. This means much more bargaining is likely in any situation where 100 per cent participation is required than when some smaller percentage can undertake group-oriented activity.

It follows that the relationship among individuals in inclusive and

62. If marginal costs rise very steeply, and accordingly no firm has an incentive to increase output greatly in response to the higher price, a single holdout need not be fatal to a collusive agreement. But a holdout will still be costly, for he will tend to gain more from the collusion than a firm that colludes, and whatever he gains the collusive firms lose.

63. On the implications of a unanimity requirement, see the important book by James M. Buchanan and Gordon Tullock, *The Calculus of Consent: Logical Foundations of Constitutional Democracy* (Ann Arbor: University of Michigan Press, 1962), especially chap. viii, pp. 96-116. I believe that some complications in their useful and provocative study could be cleared up with the aid of some of the ideas developed in the present study; see for example my review of their book in the *American Economic Review*, LII (December 1962), 1217-1218.

exclusive groups usually is quite different, whenever groups are so small one member's action has a perceptible effect on any other member, so that individual relationships matter. The firms in the exclusive group want as few others in the group as possible, and therefore each firm warily watches other firms for fear they will attempt to drive it out of the industry. Each firm must, before it takes any action, consider whether it will provoke a "price war" or "cut-throat competition." This means that each firm in an exclusive group must be sensitive to the other firms in the group, and consider the reactions they may have to any action of its own. At the same time, any group-oriented behavior in an exclusive group will usually require 100 per cent participation, so each firm in an industry is not only a rival of every other firm, but also an indispensable collaborator in any collusive action. Therefore, whenever any collusion, however tacit, is in question, each firm in the industry may consider bargaining or holding out for a larger share of the gains. The firm that can best guess what reaction other firms will have to each move of its own will have a considerable advantage in this bargaining. This fact, together with the desire of the firms in an industry to keep the number in that industry as small as possible, makes each of the firms in any industry with a small number of firms very anxious about the reactions other firms will have to any action it takes. In other words, both the desire to limit the size of the group, and the usual need for 100 per cent participation in any kind of collusion, increase the intensity and complexity of oligopolistic reactions. The conclusion that industries with small numbers of firms will be characterized by oligopolistic interaction with mutual dependence recognized is of course familiar to every economist.

It is not however generally understood that in inclusive groups, even small ones, on the other hand, bargaining or strategic interaction is evidently much less common and important. This is partly because there is no desire to eliminate anyone from the inclusive group. It is also partly because nothing like unanimous participation is normally required, so that individuals in the inclusive group are not so likely to try to be holdouts in order to get a larger share of the gain. This tends to reduce the amount of bargaining (and also makes group-oriented action more likely). Though the problem is extremely complex, and some of the tools needed to determine exactly how much bargaining there will be in a given situation do not now exist, it nonetheless seems very likely that there is much less strategic

interaction in inclusive groups, and that the hypothesis of independent behavior will frequently describe members of these groups reasonably well.

#### F. A TAXONOMY OF GROUPS

To be sure, there can also be many instances in inclusive or non-market groups in which individual members do take into account the reactions of other members to their actions when they decide what action to take—that is, instances in which there is the strategic interaction among members characteristic of oligopolistic industries in which mutual dependence is recognized. In groups of one size range at least, such strategic interaction must be relatively important. That is the size range where the group is not so small that one individual would find it profitable to purchase some of the collective good himself, but where the number in the group is nonetheless sufficiently small that each member's attempts or lack of attempts to obtain the collective good would bring about noticeable differences in the welfare of some, or all, of the others in the group. This can best be understood by assuming for a moment that an inclusive collective good is already being provided in such a group through a formal organization, and then asking what would happen if one member of the group were to cease paying his share of the cost of the good. If, in a reasonably small organization, a particular person stops paying for the collective good he enjoys, the costs will rise noticeably for each of the others in the group; accordingly, they may then refuse to continue making their contributions, and the collective good may no longer be provided. However, the first person could realize that this might be the result of his refusal to pay anything for the collective good, and that he would be worse off when the collective good is not provided than when it was provided and he met part of the cost. Accordingly he might continue making a contribution toward the purchase of the collective good. He might; or he might not. As in oligopoly in a market situation, the result is indeterminate. The rational member of such a group faces a strategic problem and while the Theory of Games and other types of analyses might prove very helpful, there seems to be no way at present of getting a general, valid, and determinate solution at the level of abstraction of this chapter.<sup>64</sup>

64. It is of incidental interest here to note also that oligopoly in the marketplace is in some respects akin to logrolling in the organization. If the "majority" that vari-

What is the range of this indeterminateness? In a small group in which a member gets such a large fraction of the total benefit that he would be better off if he paid the entire cost himself, rather than go without the good, there is some presumption that the collective good will be provided. In a group in which no one member got such a large benefit from the collective good that he had an interest in providing it even if he had to pay all of the cost, but in which the individual was still so important in terms of the whole group that his contribution or lack of contribution to the group objective had a noticeable effect on the costs or benefits of others in the group, the result is indeterminate.<sup>65</sup> By contrast, in a large group in which no single individual's contribution makes a perceptible difference to the group as a whole, or the burden or benefit of any single member of the group, it is certain that a collective good will *not* be provided unless there is coercion or some outside inducements that will lead the members of the large group to act in their common interest.<sup>66</sup>

---

ous interests in a legislature need is viewed as a collective good—something that a particular interest cannot obtain unless other interests also share it—then the parallel is quite close. The cost each special-interest legislator would like to avoid is the passage of the legislation desired by the other special-interest legislators, for if these interests gain from their legislation, often others, including his own constituents, may lose. But unless he is willing to vote for the legislation desired by the others, the particular special-interest legislator in question will not be able to get his own legislation passed. So his goal would be to work out a coalition with other special-interest legislators in which they would vote for exactly the legislation he wanted, and he in turn would give them as little in return as possible, by insisting that they moderate their legislative demands. But since every potential logroller has this same strategy, the result is indeterminate: the logs may be rolled or they may not. Every one of the interests will be better off if the logrolling is done than if it is not, but as individual interests strive for better legislative bargains the result of the competing strategies may be that no agreement is reached. This is quite similar to the situation oligopolistic groups are in, as they all desire a higher price and will all gain if they restrict output to get it, but they may not be able to agree on market shares.

65. The result is clearly indeterminate when  $F_i$  is less than  $C/V$ , at every point and it is also true that the group is not so large that no one member's actions have a noticeable effect.

66. One friendly critic has suggested that even a large pre-existing organization could continue providing a collective good simply by conducting a kind of plebiscite among its members, with the understanding that if there were not a unanimous or nearly unanimous pledge to contribute toward providing the collective good, this good would no longer be provided. This argument, if I understand it correctly, is mistaken. In such a situation, an individual would know that if others provided the collective good he would get the benefits whether he made any contribution or not. He would therefore have no incentive to make a pledge unless a completely unanimous set of pledges was required, or for some other reason his one pledge would decide whether or not the good would be provided. But if a pledge were required

The last distinction, between the group so large it definitely cannot provide itself with a collective good, and the oligopoly-sized group which may provide itself with a collective good, is particularly important. It depends upon whether any two or more members of the group have a perceptible interdependence, that is, on whether the contribution or lack of contribution of any one individual in the group will have a perceptible effect on the burden or benefit of any other individual or individuals in the group. Whether a group will have the possibility of providing itself with a collective good without coercion or outside inducements therefore depends to a striking degree upon the number of individuals in the group, since the larger the group, the less the likelihood that the contribution of any one will be perceptible. It is not, however, strictly accurate to say that it depends solely on the number of individuals in the group. The relation between the size of the group and the significance of an individual member cannot be defined quite that simply. A group which has members with highly unequal degrees of interest in a collective good, and which wants a collective good that is (at some level of provision) extremely valuable in relation to its cost, will be more apt to provide itself with a collective good than other groups with the same number of members. The same situation prevails in the study of market structure, where again the number of firms an industry can have and still remain oligopolistic (and have the possibility of supracompetitive returns) varies somewhat from case to case. The standard for determining whether a group will have the capacity to act, without coercion or outside inducements, in its group interest is (as it should be) the same for market and non-market groups: it depends on whether the individual actions of any one or more members in a group are noticeable to any other individuals in the group.<sup>67</sup> This is most obviously, but not exclusively, a function of the number in the group.

---

of every single member, or if for any other reason any one member could decide whether or not the group would get a collective good, this one member could deprive all of the others in the group of great gains. He would therefore be in a position to bargain for bribes. But since any other members of the group might gain just as much from the same holdout strategy, there is no likelihood that the collective good would be provided. See again Buchanan and Tullock, pp. 96-116.

67. The noticeability of the actions of a single member of a group may be influenced by the arrangements the group itself sets up. A previously organized group, for example, might ensure that the contributions or lack of contributions of any member of the group, and the effect of each such member's course on the burden and benefit for others, would be advertised, thus ensuring that the group effort

It is now possible to specify when either informal coordination or formal organization will be necessary to obtain a collective good. The smallest type of group—the group in which one or more members get such a large fraction of the total benefit that they find it worthwhile to see that the collective good is provided, even if they have to pay the entire cost—may get along without any group agreement or organization. A group agreement might be set up to spread the costs more widely or to step up the level of provision of the collective good. But since there is an incentive for unilateral and individual action to obtain the collective good, neither a formal organization nor even an informal group agreement is indispensable to obtain a collective good. In any group larger than this, on the other hand, no collective good can be obtained without some group agreement, coordination, or organization. In the intermediate or oligopoly-sized group, where two or more members must act simultaneously before a collective good can be obtained, there must be at least tacit coordination or organization. Moreover, the larger a group is, the more agreement and organization it will need. The larger the group, the greater the number that will usually have to be included in the group agreement or organization. It may not be necessary that the entire group be organized, since some subset of the whole group may be able to provide the collective good. But to establish a group agreement or organization will nonetheless always tend to be more difficult the larger the size of the group, for the larger the group the more difficult it will be to locate and organize even a subset of the group, and those in the subset will have an incentive to continue bargaining with the others in the group until the burden is widely shared, thereby adding to the expense of bargaining. In short, costs of organization are an increasing function of the number of individuals in the group. (Though the more

would not collapse from imperfect knowledge. I therefore define "noticeability" in terms of the degree of knowledge, and the institutional arrangements, that actually exist in any given group, instead of assuming a "natural noticeability" unaffected by any group advertising or other arrangements. This point, along with many other valuable comments, has been brought to my attention by Professor Jerome Rothenberg, who does, however, make much more of a group's assumed capacity to create "artificial noticeability" than I would want to do. I know of no practical example of a group or organization that has done much of anything, apart from improve information, to enhance the noticeability of an individual's actions in striving for a collective good.

members in the group the greater the total costs of organization, the costs of organization per person need not rise, for there are surely economies of scale in organization.) In certain cases a group will already be organized for some other purpose, and then these costs of organization are already being met. In such a case a group's capacity to provide itself with a collective good will be explained in part by whatever it was that originally enabled it to organize and maintain itself. This brings attention back again to the costs of organization and shows that these costs cannot be left out of the model, except for the smallest type of group in which unilateral action can provide a collective good. The costs of organization must be clearly distinguished from the type of cost that has previously been considered. The cost functions considered before involved only the direct resource costs of obtaining various levels of provision of a collective good. When there is no pre-existing organization of a group, and when the direct resource costs of a collective good it wants are more than any single individual could profitably bear, additional costs must be incurred to obtain an agreement about how the burden will be shared and to coordinate or organize the effort to obtain the collective good. These are the costs of communication among group members, the costs of any bargaining among them, and the costs of creating, staffing, and maintaining any formal group organization.

A group cannot get infinitesimally small quantities of a formal organization, or even of an informal group agreement; a group with a given number of members must have a certain minimal amount of organization or agreement if it is to have any at all. Thus there are significant initial or minimal costs of organization for each group. Any group that must organize to obtain a collective good, then, will find that it has a certain minimum organization cost that must be met, however little of the collective good it obtains. The greater the number in the group, the greater these minimal costs will be. When this minimal organizational cost is added to the other initial or minimal costs of a collective good, which arise from its previously mentioned technical characteristics, it is evident that the cost of the first unit of a collective good will be quite high in relation to the cost of some subsequent units. However immense the benefits of a collective good, the higher the absolute total costs of getting any

amount of that good, the less likely it is that even a minimal amount of that good could be obtained without coercion or separate, outside incentives.

This means that there are now three separate but cumulative factors that keep larger groups from furthering their own interests. First, the larger the group, the smaller the fraction of the total group benefit any person acting in the group interest receives, and the less adequate the reward for any group-oriented action, and the farther the group falls short of getting an optimal supply of the collective good, even if it should get some. Second, since the larger the group, the smaller the share of the total benefit going to any individual, or to any (absolutely) small subset of members of the group, the less the likelihood that any small subset of the group, much less any single individual, will gain enough from getting the collective good to bear the burden of providing even a small amount of it; in other words, the larger the group the smaller the likelihood of oligopolistic interaction that might help obtain the good. Third, the larger the number of members in the group the greater the organization costs, and thus the higher the hurdle that must be jumped before any of the collective good at all can be obtained. For these reasons, the larger the group the farther it will fall short of providing an optimal supply of a collective good, and very large groups normally will not, in the absence of coercion or separate, outside incentives, provide themselves with even minimal amounts of a collective good.<sup>68</sup>

68. There is one logically conceivable, but surely empirically trivial, case in which a large group could be provided with a very small amount of a collective good without coercion or outside incentives. If some very small group enjoyed a collective good so inexpensive that any one of the members would benefit by making sure that it was provided, even if he had to pay all of the cost, and if millions of people then entered the group, with the cost of the good nonetheless remaining constant, the large group could be provided with a little of this collective good. This is because by hypothesis in this example the costs have remained unchanged, so that one person still has an incentive to see that the good is provided. Even in such a case as this, however, it would still not be quite right to say that the large group was acting in its group interest, since the output of the collective good would be incredibly suboptimal. The optimal level of provision of the public good would increase each time an individual entered the group, since the unit cost of the collective good by hypothesis is constant, while the benefit from an additional unit of it increases with every entrant. Yet the original provider would have no incentive to provide more as the group expanded, unless he formed an organization to share costs with the others in this (now large) group. But that would entail incurring the considerable costs of a large organization, and there would be no way these costs could be covered through the voluntary and rational action of the individuals in the group. Thus,

Now that all sizes of groups have been considered, it is possible to develop the classification of groups that is needed. In an article that was originally part of this study, but which has been published elsewhere,<sup>69</sup> this writer and his co-author argued that the concept of the group or industry can be given a precise theoretical meaning, and should be used, along with the concept of pure monopoly, in the study of market structure. In that article the situation in which there was only one firm in the industry was called pure monopoly. The situation where the firms are so few that the actions of one firm would have a noticeable effect on some one other firm or group of firms was called oligopoly; and the situation where no one firm had any noticeable effect on any other firm was called "atomistic competition." The category of atomistic competition was subdivided into pure competition and monopolistic competition within the large group, and oligopoly was also divided into two subdivisions according as the product was homogeneous or differentiated.

For inclusive or nonmarket groups the categories must be slightly different. The analog to pure monopoly (or pure monopsony) is obviously the single individual outside the market seeking some non-collective good, some good without external economies or diseconomies. In the size range that corresponds to oligopoly in market groups, there are two separate types of nonmarket groups: "privileged" groups and "intermediate" groups. A "privileged" group is a

if the total benefit from a collective good exceeded its costs by the thousandfold or millionfold, it is logically possible that a large group could provide itself with some amount of that collective good, but the level of provision of the collective good in such a case would be only a minute fraction of the optimal level. It is not easy to think of practical examples of groups that would fit this description, but one possible example is discussed on page 161, note 94. It would be easy to rule out even any such exceptional cases, however, simply by defining all groups that could provide themselves with some amount of a collective good as "small groups" (or by giving them other names), while putting all groups that could not provide themselves with a collective good in another class. But this easy route must be rejected, for that would make this part of the theory tautologous and thus incapable of refutation. Therefore the approach here has been to make the (surely reasonable) empirical hypothesis that the total costs of the collective goods wanted by large groups are large enough to exceed the value of the small fraction of the total benefit that an individual in a large group would get, so that he will not provide the good. There may be exceptions to this, as to any other empirical statement, and thus there may be instances in which large groups could provide themselves with (at most minute amounts of) collective goods through the voluntary and rational action of one of their members.

69. Olson and McFarland (note 14 above).

group such that each of its members, or at least some one of them, has an incentive to see that the collective good is provided, even if he has to bear the full burden of providing it himself. In such a group there is a presumption<sup>70</sup> that the collective good will be obtained, and it may be obtained without any group organization or coordination whatever. An "intermediate" group is a group in which no single member gets a share of the benefit sufficient to give him an incentive to provide the good himself, but which does not have so many members that no one member will notice whether any other member is or is not helping to provide the collective good. In such a group a collective good may, or equally well may not, be obtained, but no collective good may ever be obtained without some group coordination or organization.<sup>71</sup> The analog to atomistic competition in the nonmarket situation is the very large group, which will here be called the "latent" group. It is distinguished by the fact that, if one member does or does not help provide the collective good, no other one member will be significantly affected and therefore none has any reason to react. Thus an individual in a "latent" group, by definition, cannot make a noticeable contribution to any group effort, and since no one in the group will react if he makes no contribution, he has no incentive to contribute. Accordingly, large or "latent" groups have no incentive to act to obtain a collective good because, however valuable the collective good might be to the group as a whole, it does not offer the individual any incentive to pay dues

70. It is conceivable that a "privileged" group might not provide itself with a collective good, since there might be bargaining within the group and this bargaining might be unsuccessful. Imagine a privileged group in which every member of the group would gain so much from the collective good that he would be better off if he paid the full cost of providing the collective good than he would be if the good were not provided. It is still conceivable that each member of the group, knowing that each of the others would also be better off if they provided the good alone than they would be if no collective good were obtained, would refuse to contribute anything toward obtaining the collective good. Each could refuse to help provide the collective good on the mistaken assumption that the others would provide it without him. It does not seem very likely that all of the members of the group would go on making this mistake permanently, however.

71. "The character of the numerically intermediate structure, therefore, can be explained as a mixture of both: so that each of the features of both the small and large group appears in the intermediate group, as a fragmentary trait, now emerging, now disappearing or becoming latent. Thus, the intermediate structures objectively share the essential character of the smaller and larger structures—partially or alternately. This explains the subjective uncertainty regarding the decision to which of the two they belong." (Simmel, *Sociology of Georg Simmel*, pp. 116-117.)

to any organization working in the latent group's interest, or to bear in any other way any of the costs of the necessary collective action.

Only a *separate and "selective" incentive* will stimulate a rational individual in a latent group to act in a group-oriented way. In such circumstances group action can be obtained only through an incentive that operates, not indiscriminately, like the collective good, upon the group as a whole, but rather *selectively* toward the individuals in the group. The incentive must be "selective" so that those who do not join the organization working for the group's interest, or in other ways contribute to the attainment of the group's interest, can be treated differently from those who do. These "selective incentives" can be either negative or positive, in that they can either coerce by punishing those who fail to bear an allocated share of the costs of the group action, or they can be positive inducements offered to those who act in the group interest.<sup>72</sup> A latent group that has been led to act in its group interest, either because of coercion of the individuals in the group or because of positive rewards to those individuals, will here be called a "mobilized" latent group.<sup>73</sup> Large groups are thus called "latent" groups because they have a latent power or capacity for action, but that potential power can be realized or "mobilized" only with the aid of "selective incentives."

The chances for group-oriented action are indeed different in each of the categories just explained. In some cases one may have some expectation that the collective or public good will be provided; in other cases one may be assured that (unless there are selective incentives) it will not; and still other cases could just as easily go either

72. Coercion is here defined to be a punishment that leaves an individual on a lower indifference curve than he would have been on had he borne his allocated share of the cost of the collective good and not been coerced. A positive inducement is defined to be any reward that leaves an individual who pays his allocated share of the cost of a collective good and receives the reward, on a higher indifference curve than he would have been had he borne none of the cost of the collective good and lost the reward. In other words, selective incentives are defined to be greater in value, in terms of each individual's preferences, than each individual's share of the cost of the collective good. Sanctions and inducements of smaller value will not be sufficient to mobilize a latent group. On some of the problems of distinguishing and defining coercion and positive incentives see Alfred Kuhn, *The Study of Society: A Unified Approach* (Homewood, Ill.: Richard D. Irwin, Inc. and the Dorsey Press, Inc., 1963), pp. 365-370.

73. Deutsch has also used the term "mobilization" in a somewhat similar context, but his use of the word is not the same. See Karl Deutsch, "Social Mobilization and Political Development," *American Political Science Review*, LV (September 1961), 493-514.

way. In any event, size is one of the determining factors in deciding whether or not it is possible that the voluntary, rational pursuit of individual interest will bring forth group-oriented behavior. Small groups will further their common interests better than large groups.

The question asked earlier in this chapter can now be answered. It now seems that small groups are not only quantitatively, but also qualitatively, different from large groups, and that the existence of large associations cannot be explained in terms of the same factors that explain the existence of small groups.

## THE WELFARE COSTS OF TARIFFS, MONOPOLIES, AND THEFT

GORDON TULLOCK

RICE UNIVERSITY

In recent years a considerable number of studies have been published that purport to measure the welfare costs of monopolies and tariffs.<sup>1</sup> The results have uniformly shown very small costs for practices that economists normally deplore. This led 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."<sup>2</sup> Judging 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."<sup>3</sup>

It is my purpose to take the other route suggested by Mundell and demonstrate that the "tools on which these studies are founded" produce an underestimation of the welfare costs of tariffs and monopolies. The classical economists were not concerning themselves with trifles when they argued against tariffs, and the Department of Justice is not dealing with a minuscule problem in its attacks on monopoly.

### STATICS

The present method for measuring these costs was pioneered by Professor Harberger.<sup>4</sup> Let us, therefore, begin with a very simple use of his diagram to analyze a tariff. Figure 1 shows a commodity that can be produced

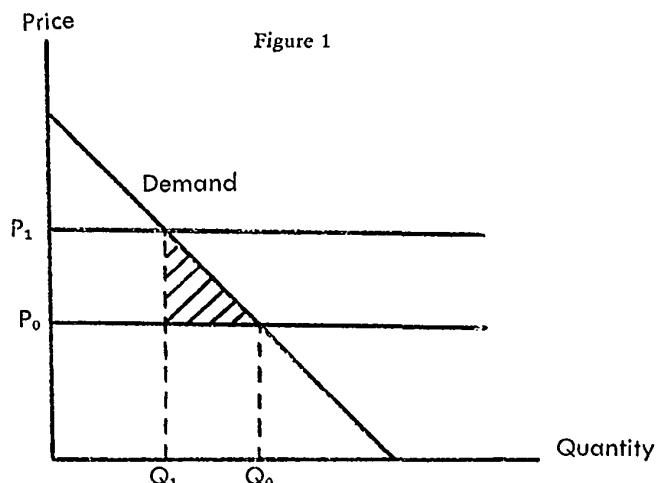
<sup>1</sup>These studies are conveniently listed with a useful table of the welfare losses computed in each in Harvey Leibenstein, "Allocative Efficiency vs. 'X-Efficiency,'" *Am. Econ. Rev.*, June 1966, 56, 392-415.

<sup>2</sup>R. A. Mundell, Review of L. H. Janssen, *Free Trade, Protection and Customs Union*, *Am. Econ. Rev.*, June 1962, 52, 622.

<sup>3</sup>*Op. cit.*, p. 392. In this article Leibenstein consistently uses the phrase "allocative efficiency" to refer solely to the absence of tariffs and monopolies.

<sup>4</sup>A. C. Harberger, "Using the Resources at Hand More Effectively," *Am. Econ. Rev.*, May 1959, 49, 134-46. 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 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. (*Cf.* Leibenstein, *op. cit.*)

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, etc., who do the actual collection and coast guards who prevent smuggling. Further, customs brokers are normally hired by the shipper to expedite the movement of their goods through customs.<sup>5</sup> Normally we pay little attention to collections 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.

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 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, building tunnels, for example, which go nowhere. Now the social

<sup>5</sup>Strictly speaking, the customs brokerage should be added on to the tax thus producing a larger welfare triangle.

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.<sup>6</sup> 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 its social cost cannot be measured simply by the welfare triangle. Figure 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.<sup>7</sup> 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 they had been employed in other industries.<sup>8</sup> 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. 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 the production of 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.

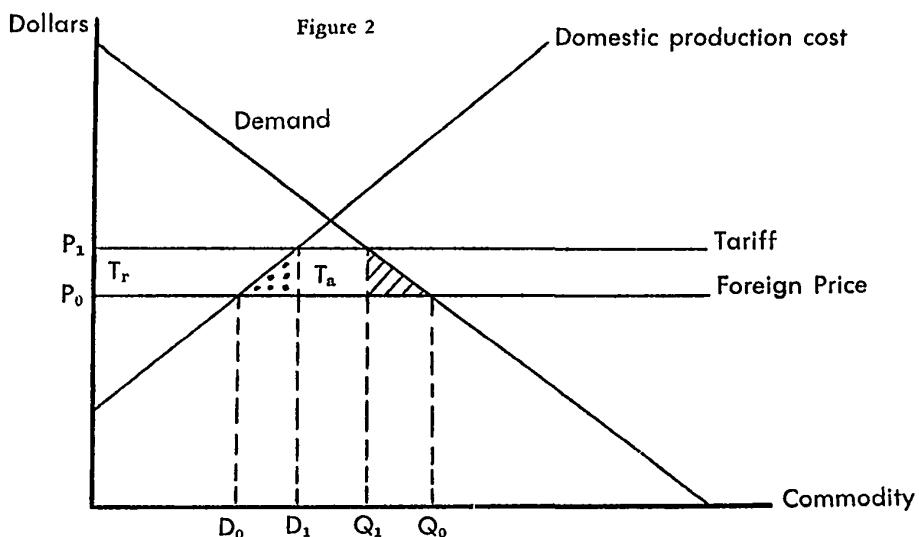
<sup>6</sup>The government action might slightly increase the rents on the resources used to build the tunnel, and thus the owners of specialized resources might benefit slightly, but clearly this is a very trivial effect.

<sup>7</sup>Tibor Scitovsky, *Economic Theory and Western European Integration*, Stanford 1958.

<sup>8</sup>There might be sizable but temporary rents to the firstcomers when the industry was first established.

TULLOCK: WELFARE COSTS

227



Let us, however, consider the situation in which there is some domestic production before the imposition of a tariff. Figure 2 shows a commodity part of the consumption 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 which would be 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 resources equal to the area of the trapezoid  $T_r$ .<sup>9</sup> 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.

<sup>9</sup>See J. Wemelsfelder, "The Short Term Effect of the Lowering of Import Duties in Germany," *Econ. Jour.*, March 1960, 70, 94-104.

DYNAMICS: THE COST OF TRANSFERS

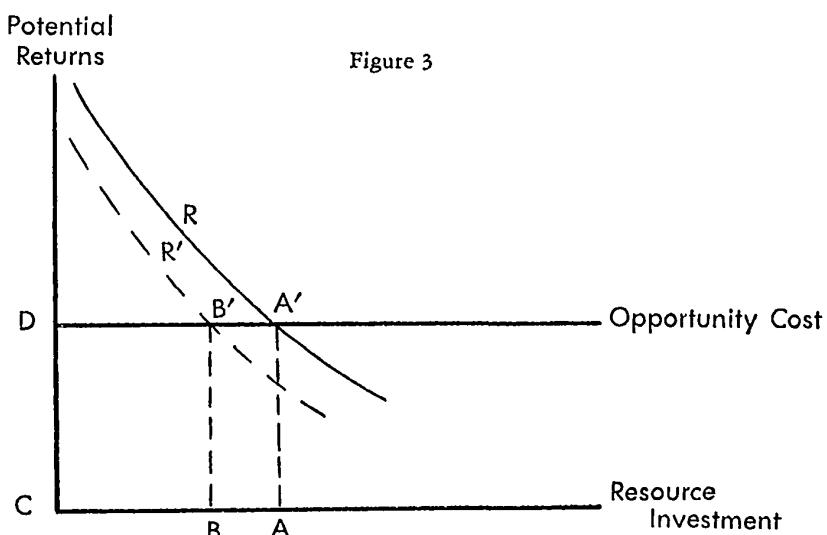
The trapezoid  $T_r$ , however, would appear to be a pure transfer, and hence not to be included in the computation of the cost of the tariff. Strictly speaking this is so, but looking at the matter dynamically, there is another social cost involved, and its magnitude is a function of the size of this 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 producing 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 attempts 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 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 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 important costs, I should like to take a detour through the economics of theft.<sup>10</sup> 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 at all, 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 and/or unimportant problem any more than our laws against monopoly.

<sup>10</sup>The economics of illegal activities is an underdeveloped area, but Harold Demsetz discusses the subject briefly in "The Exchange and Enforcement of Property Rights," *Jour. of Law and Econ.*, October 1964, 7, 11-26. J. Randolph Norsworthy's Doctoral Dissertation, *A Theory of Tax Evasion and Collection*, 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 which I began four years ago and which has languished in draft form for almost all of those four years.

Figure 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 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 thieves 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. On the horizontal axis are shown the resources invested in loss minimizing activities.<sup>11</sup> 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

<sup>11</sup>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.

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 your 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 only involves 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 Figure 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. 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, etc. The new returns are shown by  $R'$  on Figure 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.<sup>12</sup> 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 protective expenditures. Clearly

<sup>12</sup>It may be suggested that society should not be interested in the saving of 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.

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 which could be paid for from the cost of leaving them off would be less effective in preventing theft than they are.<sup>13</sup> The total social cost of theft 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, our major subject is monopoly. To return to Figure 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 a prize of this size dangling before our eyes, 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.<sup>14</sup> 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 to either break the monopoly or 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 industry 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

<sup>13</sup>James Buchanan and Gordon Tullock, "Public and Private Interaction Under Reciprocal Externality," in *The Public Economy of Urban Communities*, Julius Margolis, Ed., Washington, D.C. 1964, pp. 52-73.

<sup>14</sup>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 the likelihood of getting it. Thus they change the discount rate, rather than the payoff.

tariff treatment. In present day United States 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 a 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.<sup>15</sup> 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 most 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 function 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.

<sup>15</sup>This helpful analogy was suggested to me by Dr. William Niskanen.



---

## The Theory of Economic Regulation

Author(s): George J. Stigler

Source: *The Bell Journal of Economics and Management Science*, Vol. 2, No. 1 (Spring, 1971), pp. 3-21

Published by: RAND Corporation

Stable URL: <https://www.jstor.org/stable/3003160>

Accessed: 08-01-2026 13:06 UTC

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



RAND Corporation is collaborating with JSTOR to digitize, preserve and extend access to *The Bell Journal of Economics and Management Science*

# The theory of economic regulation

George J. Stigler

The University of Chicago

*The potential uses of public resources and powers to improve the economic status of economic groups (such as industries and occupations) are analyzed to provide a scheme of the demand for regulation. The characteristics of the political process which allow relatively small groups to obtain such regulation is then sketched to provide elements of a theory of supply of regulation. A variety of empirical evidence and illustration is also presented.*

■ The state—the machinery and power of the state—is a potential resource or threat to every industry in the society. With its power to prohibit or compel, to take or give money, the state can and does selectively help or hurt a vast number of industries. That political juggernaut, the petroleum industry, is an immense consumer of political benefits, and simultaneously the underwriters of marine insurance have their more modest repast. The central tasks of the theory of economic regulation are to explain who will receive the benefits or burdens of regulation, what form regulation will take, and the effects of regulation upon the allocation of resources.

Regulation may be actively sought by an industry, or it may be thrust upon it. A central thesis of this paper is that, as a rule, regulation is acquired by the industry and is designed and operated primarily for its benefit. There are regulations whose net effects upon the regulated industry are undeniably onerous; a simple example is the differentially heavy taxation of the industry's product (whiskey, playing cards). These onerous regulations, however, are exceptional and can be explained by the same theory that explains beneficial (we may call it "acquired") regulation.

Two main alternative views of the regulation of industry are widely held. The first is that regulation is instituted primarily for the protection and benefit of the public at large or some large subclass of the public. In this view, the regulations which injure the public—as when the oil import quotas increase the cost of petroleum products to America by \$5 billion or more a year—are costs of some social goal (here, national defense) or, occasionally, perversions of the regulatory philosophy. The second view is essentially that the political process defies rational explanation: "politics" is an imponderable, a constantly and unpredictably shifting mixture of forces of the most diverse nature, comprehending acts of great moral virtue (the emancipation of slaves) and of the most vulgar venality (the congressman feathering his own nest).

---

The author obtained the B.B.A. degree from the University of Washington, the M.B.A. degree from Northwestern, and the Ph.D. degree from the University of Chicago. He is presently Charles R. Walgreen Distinguished Service Professor of American Institutions at the University of Chicago, and has published numerous articles and texts in the field of economics. Dr. Stigler is Vice Chairman of the Securities Investor Protective Commission.

THEORY OF  
REGULATION / 3

Let us consider a problem posed by the oil import quota system: why does not the powerful industry which obtained this expensive program instead choose direct cash subsidies from the public treasury? The "protection of the public" theory of regulation must say that the choice of import quotas is dictated by the concern of the federal government for an adequate domestic supply of petroleum in the event of war—a remark calculated to elicit uproarious laughter at the Petroleum Club. Such laughter aside, if national defense were the goal of the quotas, a tariff would be a more economical instrument of policy: it would retain the profits of exclusion for the treasury. The non-rationalist view would explain the policy by the inability of consumers to measure the cost to them of the import quotas, and hence their willingness to pay \$5 billion in higher prices rather than the \$2.5 billion in cash that would be equally attractive to the industry. Our profit-maximizing theory says that the explanation lies in a different direction: the present members of the refining industries would have to share a cash subsidy with all new entrants into the refining industry.<sup>1</sup> Only when the elasticity of supply of an industry is small will the industry prefer cash to controls over entry or output.

This question, why does an industry solicit the coercive powers of the state rather than its cash, is offered only to illustrate the approach of the present paper. We assume that political systems are rationally devised and rationally employed, which is to say that they are appropriate instruments for the fulfillment of desires of members of the society. This is not to say that the state will serve any person's concept of the public interest: indeed the problem of regulation is the problem of discovering when and why an industry (or other group of like-minded people) is able to use the state for its purposes, or is singled out by the state to be used for alien purposes.

## **1. What benefits can a state provide to an industry?**

■ The state has one basic resource which in pure principle is not shared with even the mightiest of its citizens: the power to coerce. The state can seize money by the only method which is permitted by the laws of a civilized society, by taxation. The state can ordain the physical movements of resources and the economic decisions of households and firms without their consent. These powers provide the possibilities for the utilization of the state by an industry to increase its profitability. The main policies which an industry (or occupation) may seek of the state are four.

The most obvious contribution that a group may seek of the government is a direct subsidy of money. The domestic airlines received "air mail" subsidies (even if they did not carry mail) of \$1.5 billion through 1968. The merchant marine has received construction and operation subsidies reaching almost \$3 billion since World War II. The education industry has long shown a masterful skill in obtaining public funds: for example, universities and colleges have received federal funds exceeding \$3 billion annually in recent years, as well as subsidized loans for dormitories and other construction. The veterans of wars have often received direct cash bonuses.

---

<sup>1</sup> The domestic producers of petroleum, who also benefit from the import quota, would find a tariff or cash payment to domestic producers equally attractive. If their interests alone were consulted, import quotas would be auctioned off instead of being given away.

We have already sketched the main explanation for the fact that an industry with power to obtain governmental favors usually does not use this power to get money: unless the list of beneficiaries can be limited by an acceptable device, whatever amount of subsidies the industry can obtain will be dissipated among a growing number of rivals. The airlines quickly moved away from competitive bidding for air mail contracts to avoid this problem.<sup>2</sup> On the other hand, the premier universities have not devised a method of excluding other claimants for research funds, and in the long run they will receive much-reduced shares of federal research monies.

The second major public resource commonly sought by an industry is control over entry by new rivals. There is considerable, not to say excessive, discussion in economic literature of the rise of peculiar price policies (limit prices), vertical integration, and similar devices to retard the rate of entry of new firms into oligopolistic industries. Such devices are vastly less efficacious (economical) than the certificate of convenience and necessity (which includes, of course, the import and production quotas of the oil and tobacco industries).

The diligence with which the power of control over entry will be exercised by a regulatory body is already well known. The Civil Aeronautics Board has not allowed a single new trunk line to be launched since it was created in 1938. The power to insure new banks has been used by the Federal Deposit Insurance Corporation to reduce the rate of entry into commercial banking by 60 percent.<sup>3</sup> The interstate motor carrier history is in some respects even more striking, because no even ostensibly respectable case for restriction on entry can be developed on grounds of scale economies (which are in turn adduced to limit entry for safety or economy of operation). The number of federally licensed common carriers is shown in Figure 1: the immense growth of the freight hauled by trucking common carriers has been associated with a steady secular decline of numbers of such carriers. The number of applications for new certificates has been in excess of 5000 annually in recent years: a rigorous proof that hope springs eternal in an aspiring trucker's breast.

We propose the general hypothesis: every industry or occupation that has enough political power to utilize the state will seek to control entry. In addition, the regulatory policy will often be so fashioned as to retard the rate of growth of new firms. For example, no new savings and loan company may pay a dividend rate higher than that prevailing in the community in its endeavors to attract deposits.<sup>4</sup> The power to limit selling expenses of mutual funds, which is soon to be conferred upon the Securities and Exchange Commission, will serve to limit the growth of small mutual funds and hence reduce the sales costs of large funds.

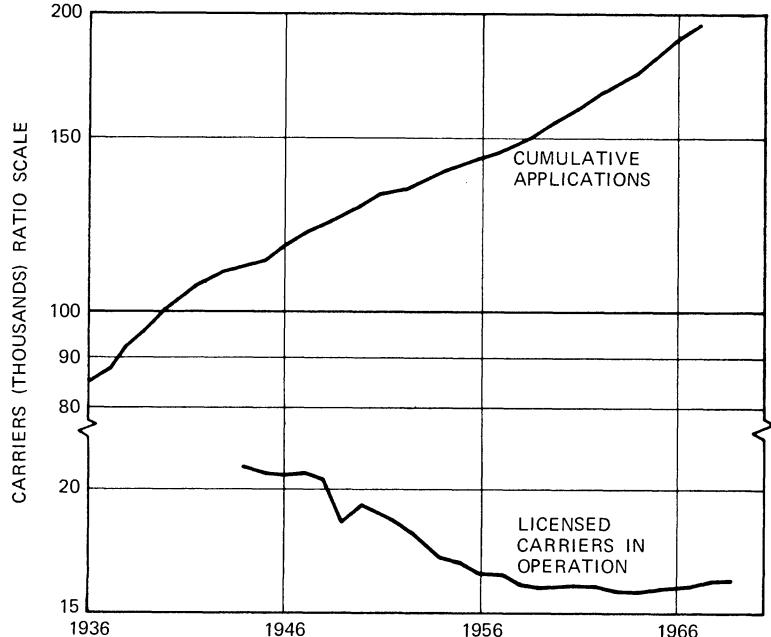
One variant of the control of entry is the protective tariff (and the corresponding barriers which have been raised to interstate movements of goods and people). The benefits of protection to an industry, one might think, will usually be dissipated by the entry of new domestic producers, and the question naturally arises: Why does the industry not also seek domestic entry controls? In a few industries

<sup>2</sup> See [7], pp. 60 ff.

<sup>3</sup> See [10].

<sup>4</sup> The Federal Home Loan Bank Board is the regulatory body. It also controls the amount of advertising and other areas of competition.

FIGURE 1  
CERTIFICATES FOR INTERSTATE MOTOR CARRIERS



SOURCE: TABLE 5

(petroleum) the domestic controls have been obtained, but not in most. The tariff will be effective if there is a specialized domestic resource necessary to the industry; oil-producing lands is an example. Even if an industry has only durable specialized resources, it will gain if its contraction is slowed by a tariff.

A third general set of powers of the state which will be sought by the industry are those which affect substitutes and complements. Crudely put, the butter producers wish to suppress margarine and encourage the production of bread. The airline industry actively supports the federal subsidies to airports; the building trade unions have opposed labor-saving materials through building codes. We shall examine shortly a specific case of inter-industry competition in transportation.

The fourth class of public policies sought by an industry is directed to price-fixing. Even the industry that has achieved entry control will often want price controls administered by a body with coercive powers. If the number of firms in the regulated industry is even moderately large, price discrimination will be difficult to maintain in the absence of public support. The prohibition of interest on demand deposits, which is probably effective in preventing interest payments to most non-business depositors, is a case in point. Where there are no diseconomies of large scale for the individual firm (e.g., a motor trucking firm can add trucks under a given license as common carrier), price control is essential to achieve more than competitive rates of return.

**Limitations upon political benefits.** These various political boons are not obtained by the industry in a pure profit-maximizing form. The political process erects certain limitations upon the exercise of cartel policies by an industry. These limitations are of three sorts.

TABLE 1

IMPORT QUOTAS OF REFINERIES AS PERCENT  
OF DAILY INPUT OF PETROLEUM  
(DISTRICTS I - IV, JULY 1, 1959 - DEC. 31, 1959)

SIZE OF REFINERY (THOUSANDS OF BARRELS)	PERCENT QUOTA
0-10	11.4
10-20	10.4
20-30	9.5
30-60	8.5
60-100	7.6
100-150	6.6
150-200	5.7
200-300	4.7
300 AND OVER	3.8

SOURCE: HEARING, SELECT COMMITTEE ON SMALL BUSINESS, U. S. CONGRESS,  
88th CONG., 2nd SESS., AUG. 10 AND 11, 1964, [12] P. 121.

First, the distribution of control of the industry among the firms in the industry is changed. In an unregulated industry each firm's influence upon price and output is proportional to its share of industry output (at least in a simple arithmetic sense of direct capacity to change output). The political decisions take account also of the political strength of the various firms, so small firms have a larger influence than they would possess in an unregulated industry. Thus, when quotas are given to firms, the small firms will almost always receive larger quotas than cost-minimizing practices would allow. The original quotas under the oil import quota system will illustrate this practice (Table 1). The smallest refiners were given a quota of 11.4 percent of their daily consumption of oil, and the percentage dropped as refinery size rose.<sup>5</sup> The pattern of regressive benefits is characteristic of public controls in industries with numerous firms.

Second, the procedural safeguards required of public processes are costly. The delays which are dictated by both law and bureaucratic thoughts of self-survival can be large: Robert Gerwig found the price of gas sold in interstate commerce to be 5 to 6 percent higher than in intrastate commerce because of the administrative costs (including delay) of Federal Power Commission reviews [5].

Finally, the political process automatically admits powerful outsiders to the industry's councils. It is well known that the allocation of television channels among communities does not maximize industry revenue but reflects pressures to serve many smaller communities. The abandonment of an unprofitable rail line is an even more notorious area of outsider participation.

These limitations are predictable, and they must all enter into the calculus of the profitability of regulation of an industry.

- **An illustrative analysis.** The recourse to the regulatory process is of course more specific and more complex than the foregoing sketch

<sup>5</sup> The largest refineries were restricted to 75.7 percent of their historical quota under the earlier voluntary import quota plan.

suggests. The defensive power of various other industries which are affected by the proposed regulation must also be taken into account. An analysis of one aspect of the regulation of motor trucking will illustrate these complications. At this stage we are concerned only with the correspondence between regulations and economic interests; later we shall consider the political process by which regulation is achieved.

The motor trucking industry operated almost exclusively within cities before 1925, in good part because neither powerful trucks nor good roads were available for long-distance freight movements. As these deficiencies were gradually remedied, the share of trucks in intercity freight movements began to rise, and by 1930 it was estimated to be 4 percent of ton-miles of intercity freight. The railroad industry took early cognizance of this emerging competitor, and one of the methods by which trucking was combatted was state regulation.

By the early 1930's all states regulated the dimensions and weight of trucks. The weight limitations were a much more pervasive control over trucking than the licensing of common carriers because even the trucks exempt from entry regulation are subject to the limitations on dimensions and capacity. The weight regulations in the early 1930's are reproduced in the appendix (Table 6). Sometimes the participation of railroads in the regulatory process was incontrovertible: Texas and Louisiana placed a 7000-pound payload limit on trucks serving (and hence competing with) two or more railroad stations, and a 14,000-pound limit on trucks serving only one station (hence, not competing with it).

We seek to determine the pattern of weight limits on trucks that would emerge in response to the economic interests of the concerned parties. The main considerations appear to be the following:

- (1) Heavy trucks would be allowed in states with a substantial number of trucks on farms: the powerful agricultural interests would insist upon this. The 1930 Census reports nearly one million trucks on farms. One variable in our study will be, for each state, trucks per 1000 of agricultural population.<sup>6</sup>
- (2) Railroads found the truck an effective and rapidly triumphing competitor in the shorter hauls and hauls of less than carload traffic, but much less effective in the carload and longer-haul traffic. Our second variable for each state is, therefore, length of average railroad haul.<sup>7</sup> The longer the average rail haul is, the less the railroads will be opposed to trucks.
- (3) The public at large would be concerned by the potential damage done to the highway system by heavy trucks. The better the state highway system, the heavier the trucks that would be permitted. The percentage of each state's highways that had a high type surface is the third variable. Of course good highways are more likely to exist where the potential contribution of trucks to a state's economy is greater, so the causation may be looked at from either direction.

<sup>6</sup> The ratio of trucks to total population would measure the product of (1) the importance of trucks to farmers, and (2) the importance of farmers in the state. For reasons given later, we prefer to emphasize (1).

<sup>7</sup> This is known for each railroad, and we assume that (1) the average holds within each state, and (2) two or more railroads in a state may be combined on the basis of mileage. Obviously both assumptions are at best fair approximations.

We have two measures of weight limits on trucks, one for 4-wheel trucks ( $X_1$ ) and one for 6-wheel trucks ( $X_2$ ). We may then calculate two equations,

$$X_1 \text{ (or } X_2) = a + bX_3 + cX_4 + dX_5,$$

where

$X_3$  = trucks per 1000 agricultural labor force, 1930 ,

$X_4$  = average length of railroad haul of freight traffic, 1930,

$X_5$  = percentage of state roads with high-quality surface, 1930.

(All variables are fully defined and their state values given in Table 7 on page 20.)

The three explanatory variables are statistically significant, and each works in the expected direction. The regulations on weight were less onerous; the larger the truck population in farming, the less competitive the trucks were to railroads (i.e., the longer the rail hauls), and the better the highway system (see Table 2).

□ The foregoing analysis is concerned with what may be termed the industrial demand for governmental powers. Not every industry will have a significant demand for public assistance (other than money!), meaning the prospect of a substantial increase in the present value of the enterprises even if the governmental services could be obtained gratis (and of course they have costs to which we soon turn). In some economic activities entry of new rivals is extremely difficult to control—consider the enforcement problem in restricting the supply of domestic servants. In some industries the substitute products cannot be efficiently controlled—consider the competition offered to bus lines by private car-pooling. Price fixing is not feasible where every

TABLE 2

REGRESSION ANALYSIS OF STATE WEIGHT LIMITS ON TRUCKS  
(T VALUES UNDER REGRESSION COEFFICIENTS)

DEPENDENT VARIABLE	N	CONSTANT	$X_3$	$X_4$	$X_5$	R <sup>2</sup>
$X_1$	48	12.28 (4.87)	0.0336 (3.99)	0.0287 (2.77)	0.2641 (3.04)	0.502
$X_2$	46	10.34 (1.57)	0.0437 (2.01)	0.0788 (2.97)	0.2528 (1.15)	0.243

$X_1$  = WEIGHT LIMIT ON 4-WHEEL TRUCKS (THOUSANDS OF POUNDS), 1932-33  
 $X_2$  = WEIGHT LIMIT ON 6-WHEEL TRUCKS (THOUSANDS OF POUNDS), 1932-33  
 $X_3$  = TRUCKS ON FARMS PER 1,000 AGRICULTURAL LABOR FORCE, 1930  
 $X_4$  = AVERAGE LENGTH OF RAILROAD HAUL OF FREIGHT (MILES), 1930  
 $X_5$  = PERCENT OF STATE HIGHWAYS WITH HIGH-TYPE SURFACE,  
DEC. 31, 1930

SOURCES:  $X_1$  AND  $X_2$ : THE MOTOR TRUCK RED BOOK AND DIRECTORY [11], 1934 EDITION, P. 85-102, AND U.S. DEPT. OF AGRIC., BUR. OF PUBLIC ROADS, DEC. 1932 [13].

$X_3$ : CENSUS OF AGRICULTURE, 1930, VOL. IV, [14].

$X_4$ : A.A.R.R., BUR. OF RAILWAY ECONOMICS, RAILWAY MILEAGE BY STATES, DEC. 31, 1930 [1] AND U.S.I.C.C., STATISTICS OF RAILWAYS IN THE U.S., 1930 [18].

$X_5$ : STATISTICAL ABSTRACT OF THE U.S., 1932 [16].

unit of the product has a different quality and price, as in the market for used automobiles. In general, however, most industries will have a positive demand price (schedule) for the services of government.

## 2. The costs of obtaining legislation

■ When an industry receives a grant of power from the state, the benefit to the industry will fall short of the damage to the rest of the community. Even if there were no deadweight losses from acquired regulation, however, one might expect a democratic society to reject such industry requests unless the industry controlled a majority of the votes.<sup>8</sup> A direct and informed vote on oil import quotas would reject the scheme. (If it did not, our theory of rational political processes would be contradicted.) To explain why many industries are able to employ the political machinery to their own ends, we must examine the nature of the political process in a democracy.

A consumer chooses between rail and air travel, for example, by voting with his pocketbook: he patronizes on a given day that mode of transportation he prefers. A similar form of economic voting occurs with decisions on where to work or where to invest one's capital. The market accumulates these economic votes, predicts their future course, and invests accordingly.

Because the political decision is coercive, the decision process is fundamentally different from that of the market. If the public is asked to make a decision between two transportation media comparable to the individual's decision on how to travel—say, whether airlines or railroads should receive a federal subsidy—the decision must be abided by everyone, travellers and non-travellers, travellers this year and travellers next year. This compelled universality of political decisions makes for two differences between democratic political decision processes and market processes.

(1) The decisions must be made simultaneously by a large number of persons (or their representatives): the political process demands simultaneity of decision. If *A* were to vote on the referendum today, *B* tomorrow, *C* the day after, and so on, the accumulation of a majority decision would be both expensive and suspect. (*A* might wish to cast a different vote now than last month.)

The condition of simultaneity imposes a major burden upon the political decision process. It makes voting on specific issues prohibitively expensive: it is a significant cost even to engage in the transaction of buying a plane ticket when I wish to travel; it would be stupendously expensive to me to engage in the physically similar transaction of voting (i.e., patronizing a polling place) whenever a number of my fellow citizens desired to register their views on railroads versus airplanes. To cope with this condition of simultaneity, the voters must employ representatives with wide discretion and must eschew direct expressions of marginal changes in preferences. This characteristic also implies that the political decision does not predict voter desires and make preparations to fulfill them in advance of their realization.

---

<sup>8</sup> If the deadweight loss (of consumer and producer surplus) is taken into account, even if the oil industry were in the majority it would not obtain the legislation if there were available some method of compensation (such as sale of votes) by which the larger damage of the minority could be expressed effectively against the lesser gains of the majority.

(2) The democratic decision process must involve “all” the community, not simply those who are directly concerned with a decision. In a private market, the non-traveller never votes on rail versus plane travel, while the huge shipper casts many votes each day. The political decision process cannot exclude the uninterested voter: the abuses of any exclusion except self-exclusion are obvious. Hence, the political process does not allow participation in proportion to interest and knowledge. In a measure, this difficulty is moderated by other political activities besides voting which do allow a more effective vote to interested parties: persuasion, employment of skilled legislative representatives, etc. Nevertheless, the political system does not offer good incentives like those in private markets to the acquisition of knowledge. If I consume ten times as much of public service A (streets) as of B (schools), I do not have incentives to acquire corresponding amounts of knowledge about the public provision of these services.<sup>9</sup>

These characteristics of the political process can be modified by having numerous levels of government (so I have somewhat more incentive to learn about local schools than about the whole state school system) and by selective use of direct decision (bond referenda). The chief method of coping with the characteristics, however, is to employ more or less full-time representatives organized in (disciplined by) firms which are called political parties or machines.

The representative and his party are rewarded for their discovery and fulfillment of the political desires of their constituency by success in election and the perquisites of office. If the representative could confidently await reelection whenever he voted against an economic policy that injured the society, he would assuredly do so. Unfortunately virtue does not always command so high a price. If the representative denies ten large industries their special subsidies of money or governmental power, they will dedicate themselves to the election of a more complaisant successor: the stakes are that important. This does not mean that every large industry can get what it wants or all that it wants: it does mean that the representative and his party must find a coalition of voter interests more durable than the anti-industry side of every industry policy proposal. A representative cannot win or keep office with the support of the sum of those who are opposed to: oil import quotas, farm subsidies, airport subsidies, hospital subsidies, unnecessary navy shipyards, an inequitable public housing program, and rural electrification subsidies.

The political decision process has as its dominant characteristic infrequent, universal (in principle) participation, as we have noted: political decisions must be infrequent and they must be global. The voter's expenditure to learn the merits of individual policy proposals and to express his preferences (by individual and group representation as well as by voting) are determined by expected costs and returns, just as they are in the private marketplace. The costs of comprehensive information are higher in the political arena because information must be sought on many issues of little or no direct concern to the individual, and accordingly he will know little about most matters before the legislature. The expressions of preferences in voting will be less precise than the expressions of preferences in the

<sup>9</sup> See [2].

marketplace because many uninformed people will be voting and affecting the decision.<sup>10</sup>

The channels of political decision-making can thus be described as gross or filtered or noisy. If everyone has a negligible preference for policy A over B, the preference will not be discovered or acted upon. If voter group X wants a policy that injures non-X by a small amount, it will not pay non-X to discover this and act against the policy. The system is calculated to implement all strongly felt preferences of majorities and many strongly felt preferences of minorities but to disregard the lesser preferences of majorities and minorities. The filtering or grossness will be reduced by any reduction in the cost to the citizen of acquiring information and expressing desires and by any increase in the probability that his vote will influence policy.

The industry which seeks political power must go to the appropriate seller, the political party. The political party has costs of operation, costs of maintaining an organization and competing in elections. These costs of the political process are viewed excessively narrowly in the literature on the financing of elections: elections are to the political process what merchandizing is to the process of producing a commodity, only an essential final step. The party maintains its organization and electoral appeal by the performance of costly services to the voter at all times, not just before elections. Part of the costs of services and organization are borne by putting a part of the party's workers on the public payroll. An opposition party, however, is usually essential insurance for the voters to discipline the party in power, and the opposition party's costs are not fully met by public funds.

The industry which seeks regulation must be prepared to pay with the two things a party needs: votes and resources. The resources may be provided by campaign contributions, contributed services (the businessman heads a fund-raising committee), and more indirect methods such as the employment of party workers. The votes in support of the measure are rallied, and the votes in opposition are dispersed, by expensive programs to educate (or uneducate) members of the industry and of other concerned industries.

These costs of legislation probably increase with the size of the industry seeking the legislation. Larger industries seek programs which cost the society more and arouse more opposition from substantially affected groups. The tasks of persuasion, both within and without the industry, also increase with its size. The fixed size of the political "market," however, probably makes the cost of obtaining legislation increase less rapidly than industry size. The smallest industries are therefore effectively precluded from the political process unless they have some special advantage such as geographical concentration in a sparsely settled political subdivision.

If a political party has in effect a monopoly control over the governmental machine, one might expect that it could collect most of the benefits of regulation for itself. Political parties, however, are

---

<sup>10</sup> There is an organizational problem in any decision in which more than one vote is cast. If because of economies of scale it requires a thousand customers to buy a product before it can be produced, this thousand votes has to be assembled by some entrepreneur. Unlike the political scene, however, there is no need to obtain the consent of the remainder of the community, because they will bear no part of the cost.

perhaps an ideal illustration of Demsetz' theory of natural monopoly [4]. If one party becomes extortionate (or badly mistaken in its reading of effective desires), it is possible to elect another party which will provide the governmental services at a price more closely proportioned to costs of the party. If entry into politics is effectively controlled, we should expect one-party dominance to lead that party to solicit requests for protective legislation but to exact a higher price for the legislation.

The internal structure of the political party, and the manner in which the perquisites of office are distributed among its members, offer fascinating areas for study in this context. The elective officials are at the pinnacle of the political system—there is no substitute for the ability to hold the public offices. I conjecture that much of the compensation to the legislative leaders takes the form of extra-political payments. Why are so many politicians lawyers?—because everyone employs lawyers, so the congressman's firm is a suitable avenue of compensation, whereas a physician would have to be given bribes rather than patronage. Most enterprises patronize insurance companies and banks, so we may expect that legislators commonly have financial affiliations with such enterprises.

The financing of industry-wide activities such as the pursuit of legislation raises the usual problem of the free rider.<sup>11</sup> We do not possess a satisfactory theory of group behavior—indeed this theory is the theory of oligopoly with one addition: in the very large number industry (e.g., agriculture) the political party itself will undertake the entrepreneurial role in providing favorable legislation. We can go no further than the infirmities of oligopoly theory allow, which is to say, we can make only plausible conjectures such as that the more concentrated the industry, the more resources it can invest in the campaign for legislation.

□ **Occupational licensing.** The licensing of occupations is a possible use of the political process to improve the economic circumstances of a group. The license is an effective barrier to entry because occupational practice without the license is a criminal offense. Since much occupational licensing is performed at the state level, the area provides an opportunity to search for the characteristics of an occupation which give it political power.

Although there are serious data limitations, we may investigate several characteristics of an occupation which should influence its ability to secure political power:

(1) *The size of the occupation.* Quite simply, the larger the occupation, the more votes it has. (Under some circumstances, therefore, one would wish to exclude non-citizens from the measure of size.)

(2) *The per capita income of the occupation.* The income of the occupation is the product of its number and average income, so this variable and the preceding will reflect the total income of the occupation. The income of the occupation is presumably an index of the probable rewards of successful political action: in the absence of specific knowledge of supply and demand functions, we expect

---

<sup>11</sup> The theory that the lobbying organization avoids the “free-rider” problem by selling useful services was proposed by Thomas G. Moore [8] and elaborated by Mancur Olson [9]. The theory has not been tested empirically.

licensing to increase each occupation's equilibrium income by roughly the same proportion. In a more sophisticated version, one would predict that the less the elasticity of demand for the occupation's services, the more profitable licensing would be. One could also view the income of the occupation as a source of funds for political action, but if we view political action as an investment this is relevant only with capital-market imperfections.<sup>12</sup>

The average income of occupational members is an appropriate variable in comparisons among occupations, but it is inappropriate to comparisons of one occupation in various states because real income will be approximately equal (in the absence of regulation) in each state.

(3) *The concentration of the occupation in large cities.* When the occupation organizes a campaign to obtain favorable legislation, it incurs expenses in the solicitation of support, and these are higher for a diffused occupation than a concentrated one. The solicitation of support is complicated by the free-rider problem in that individual members cannot be excluded from the benefits of legislation even if they have not shared the costs of receiving it. If most of the occupation is concentrated in a few large centers, these problems (we suspect) are much reduced in intensity: regulation may even begin at the local governmental level. We shall use an orthodox geographical concentration measure: the share of the occupation of the state in cities over 100,000 (or 50,000 in 1900 and earlier).

(4) *The presence of a cohesive opposition to licensing.* If an occupation deals with the public at large, the costs which licensing imposes upon any one customer or industry will be small and it will not be economic for that customer or industry to combat the drive for licensure. If the injured group finds it feasible and profitable to act jointly, however, it will oppose the effort to get licensure, and (by increasing its cost) weaken, delay, or prevent the legislation. The same attributes—numbers of voters, wealth, and ease of organization—which favor an occupation in the political arena, of course, favor also any adversary group. Thus, a small occupation employed by only one industry which has few employers will have difficulty in getting licensure; whereas a large occupation serving everyone will encounter no organized opposition.

An introductory statistical analysis of the licensing of select occupations by states is summarized in Table 3. In each occupation the dependent variable for each state is the year of first regulation of entry into the occupation. The two independent variables are

- (1) the ratio of the occupation to the total labor force of the state in the census year nearest to the median year of regulation,
- (2) the fraction of the occupation found in cities over 100,000 (over 50,000 in 1890 and 1900) in that same year.

---

<sup>12</sup> Let  $n$  = the number of members of the profession and  $y$  = average income. We expect political capacity to be in proportion to  $(ny)$  so far as benefits go, but to reflect also the direct value of votes, so the capacity becomes proportional to  $(n^a y)$  with  $a > 1$ .

TABLE 3

INITIAL YEAR OF REGULATION AS A FUNCTION OF  
RELATIVE SIZE OF OCCUPATION AND DEGREE OF URBANIZATION

OCCUPATION	NUMBER OF STATES LICENSING	MEDIAN CENSUS YEAR OF LICENSING	REGRESSION COEFFICIENTS (AND T-VALUES)		R <sup>2</sup>
			SIZE OF OCCUPATION (RELATIVE TO LABOR FORCE)	URBANIZATION (SHARE OF OCCUPA- TION IN CITIES OVER 100,000*)	
BEAUTICIANS	48	1930	-4.03 (2.50)	5.90 (1.24)	0.125
ARCHITECTS	47	1930	-24.06 (2.15)	-6.29 (0.84)	0.184
BARBERS	46	1930	-1.31 (0.51)	-26.10 (2.37)	0.146
LAWYERS	29	1890	-0.26 (0.08)	-65.78 (1.70)	0.102
PHYSICIANS	43	1890	0.64 (0.65)	-23.80 (2.69)	0.165
EMBALMERS	37	1910	3.32 (0.36)	-4.24 (0.44)	0.007
REGISTERED NURSES	48	1910	-2.08 (2.28)	-3.36 (1.06)	0.176
DENTISTS	48	1900	2.51 (0.44)	-22.94 (2.19)	0.103
VETERINARIANS	40	1910	-10.69 (1.94)	-37.16 (4.20)	0.329
CHIROPRACTORS	48	1930	-17.70 (1.54)	11.69 (1.25)	0.079
PHARMACISTS	48	1900	-4.19 (1.50)	-6.84 (0.80)	0.082

SOURCES: THE COUNCIL OF STATE GOVERNMENTS, "OCCUPATIONAL LICENSING LEGISLATION IN THE STATES", 1952 [3], AND U.S. CENSUS OF POPULATION [15], VARIOUS YEARS.

\* 50,000 IN 1890 AND 1900.

We expect these variables to be negatively associated with year of licensure, and each of the nine statistically significant regression coefficients is of the expected sign.

The results are not robust, however: the multiple correlation coefficients are small, and over half of the regression coefficients are not significant (and in these cases often of inappropriate sign). Urbanization is more strongly associated than size of occupation with licensure.<sup>13</sup> The crudity of the data may be a large source of these disappointments: we measure, for example, the characteristics of the barbers in each state in 1930, but 14 states were licensing barbers by 1910. If the states which licensed barbering before 1910 had relatively more barbers, or more highly urbanized barbers, the predictions

<sup>13</sup> We may pool the occupations and assign dummy variables for each occupation; the regression coefficients then are:

$$\begin{aligned} \text{size of occupation relative to labor force: } & -0.450 (t = 0.59) \\ \text{urbanization} & : -12.133 (t = 4.00). \end{aligned}$$

Thus urbanization is highly significant, while size of occupation is not significant.

would be improved. The absence of data for years between censuses and before 1890 led us to make only the cruder analysis.<sup>14</sup>

In general, the larger occupations were licensed in earlier years.<sup>15</sup> Veterinarians are the only occupation in this sample who have a well-defined set of customers, namely livestock farmers, and licensing was later in those states with large numbers of livestock relative to rural population. The within-occupation analyses offer some support for the economic theory of the supply of legislation.

A comparison of different occupations allows us to examine several other variables. The first is income, already discussed above. The second is the size of the market. Just as it is impossible to organize an effective labor union in only one part of an integrated market, so it is impossible to regulate only one part of the market. Consider an occupation—junior business executives will do—which has a national market with high mobility of labor and significant mobility of employers. If the executives of one state were to organize, their scope for effective influence would be very small. If salaries were raised above the competitive level, employers would often recruit elsewhere so the demand elasticity would be very high.<sup>16</sup> The third variable is stability of occupational membership: the longer the members are in the occupation, the greater their financial gain from control of entry. Our regrettably crude measure of this variable is based upon the number of members aged 35–44 in 1950 and aged 45–54 in 1960: the closer these numbers are, the more stable the membership of the occupation. The data for the various occupations are given in Table 4.

The comparison of licensed and unlicensed occupations is consistently in keeping with our expectations:

- (1) the licensed occupations have higher incomes (also before licensing, one may assume),
- (2) the membership of the licensed occupations is more stable (but the difference is negligible in our crude measure),
- (3) the licensed occupations are less often employed by business enterprises (who have incentives to oppose licensing),
- (4) all occupations in national markets (college teachers, engineers, scientists, accountants) are unlicensed or only partially licensed.

---

<sup>14</sup> A more precise analysis might take the form of a regression analysis such as:  
Year of licensure = constant

$$+b_1 \text{ (year of critical size of occupation)} \\ +b_2 \text{ (year of critical urbanization of occupation)},$$

where the critical size and urbanization were defined as the mean size and mean urbanization in the year of licensure.

<sup>15</sup> Lawyers, physicians, and pharmacists were all relatively large occupations by 1900, and nurses also by 1910. The only large occupation to be licensed later was barbers; the only small occupation to be licensed early was embalmers.

<sup>16</sup> The regulation of business in a partial market will also generally produce very high supply elasticities within a market: if the price of the product (or service) is raised, the pressure of excluded supply is very difficult to resist. Some occupations are forced to reciprocity in licensing, and the geographical dispersion of earnings in licensed occupations, one would predict, is not appreciably different than in unlicensed occupations with equal employer mobility. Many puzzles are posed by the interesting analysis of Arlene S. Holen in [6], pp. 492–98.

TABLE 4

CHARACTERISTICS OF LICENSED AND UNLICENSED  
PROFESSIONAL OCCUPATIONS, 1960

OCCUPATION	MEDIAN AGE (YEARS)	MEDIAN EDUCATION (YEARS)	MEDIAN EARNINGS (50-52 WKS.)	INSTABILITY OF MEMBERSHIP*	PERCENT NOT SELF-EMPLOYED	PERCENT IN CITIES OVER 50,000	PERCENT OF LABOR FORCE
LICENSED:							
ARCHITECTS	41.7	16.8	\$ 9,090	0.012	57.8%	44.1%	0.045%
CHIROPRACTORS	46.5	16.4	6,360	0.053	5.8	30.8	0.020
DENTISTS	45.9	17.3	12,200	0.016	9.4	34.5	0.128
EMBALMERS	43.5	13.4	5,990	0.130	52.8	30.2	0.055
LAWYERS	45.3	17.4	10,800	0.041	35.8	43.1	0.308
PROF. NURSES	39.1	13.2	3,850	0.291	91.0	40.6	0.868
OPTOMETRISTS	41.6	17.0	8,480	0.249	17.5	34.5	0.024
PHARMACISTS	44.9	16.2	7,230	0.119	62.3	40.0	0.136
PHYSICIANS	42.8	17.5	14,200	0.015	35.0	44.7	0.339
VETERINARIANS	39.2	17.4	9,210	0.169	29.5	14.4	0.023
AVERAGE	43.0	16.3	8,741	0.109	39.7	35.7	0.195
PARTIALLY LICENSED:							
ACCOUNTANTS	40.4	14.9	6,450	0.052	88.1	43.5	0.698
ENGINEERS	38.3	16.2	8,490	0.023	96.8	31.6	1.279
ELEM. SCHOOL TEACHERS	43.1	16.5	4,710	(a)	99.1	18.8	1.482
AVERAGE	40.6	15.9	6,550	0.117(b)	94.7	34.6	1.153
UNLICENSED:							
ARTISTS	38.0	14.2	5,920	0.103	77.3	45.7	0.154
CLERGYMEN	43.3	17.0	4,120	0.039	89.0	27.2	0.295
COLLEGE TEACHERS	40.3	17.4	7,500	0.085	99.2	36.0	0.261
DRAFTSMEN	31.2	12.9	5,990	0.098	98.6	40.8	0.322
REPORTERS & EDITORS	39.4	15.5	6,120	0.138	93.9	43.3	0.151
MUSICIANS	40.2	14.8	3,240	0.081	65.5	37.7	0.289
NATURAL SCIENTISTS	35.9	16.8	7,490	0.264	96.3	32.7	0.221
AVERAGE	38.3	15.5	5,768	0.115	88.5	37.6	0.242

(\*) 1-R, WHERE R = RATIO: 1960 AGE 45-54 TO 1950 AGE 35-44.

(a) NOT AVAILABLE SEPARATELY; TEACHERS N.E.C. (INCL. SECONDARY SCHOOL AND OTHER) = 0.276

(b) INCLUDES FIGURE FOR TEACHERS N.E.C. IN NOTE (a)

SOURCE: U.S. CENSUS OF POPULATION, [15], 1960.

The size and urbanization of the three groups, however, are unrelated to licensing. The inter-occupational comparison therefore provides a modicum of additional support for our theory of regulation.

■ The idealistic view of public regulation is deeply imbedded in professional economic thought. So many economists, for example, have denounced the ICC for its pro-railroad policies that this has become a cliché of the literature. This criticism seems to me exactly as appropriate as a criticism of the Great Atlantic and Pacific Tea Company for selling groceries, or as a criticism of a politician for currying popular support. The fundamental vice of such criticism is that it misdirects attention: it suggests that the way to get an ICC which is not subservient to the carriers is to preach to the commissioners or to the people who appoint the commissioners. The only way to get a different commission would be to change the political

### 3. Conclusion

support for the Commission, and reward commissioners on a basis unrelated to their services to the carriers.

Until the basic logic of political life is developed, reformers will be ill-equipped to use the state for their reforms, and victims of the pervasive use of the state's support of special groups will be helpless to protect themselves. Economists should quickly establish the license to practice on the rational theory of political behavior.

## Appendix

TABLE 5

COMMON, CONTRACT AND PASSENGER MOTOR CARRIERS, 1935-1969<sup>1</sup>

YEAR ENDING	CUMULATIVE APPLICATIONS			OPERATING CARRIERS	
	GRAND- FATHER	NEW	TOTAL	APPROVED APPLICATIONS <sup>3</sup>	NUMBER IN OPERATION <sup>2</sup>
OCT. 1936	82,827	1,696	84,523	—	—
1937	83,107	3,921	87,028	1,114	—
1938	85,646	6,694	92,340	20,398	—
1939	86,298	9,636	95,934	23,494	—
1940	87,367	12,965	100,332	25,575	—
1941	88,064	16,325	104,389	26,296	—
1942	88,702	18,977	107,679	26,683	—
1943	89,157	20,007	109,164	27,531	—
1944	89,511	21,324	110,835	27,177	21,044
1945	89,518	22,829	112,347	—	20,788
1946	89,529	26,392	115,921	—	20,632
1947	89,552	29,604	119,156	—	20,665
1948	89,563	32,678	122,241	—	20,373
1949	89,567	35,635	125,202	—	18,459
1950	89,573	38,666	128,239	—	19,200
1951	89,574	41,889	131,463	—	18,843
1952	(89,574) <sup>4</sup>	44,297	133,870	—	18,408
1953	"	46,619	136,192	—	17,869
1954	"	49,146	138,719	—	17,080
1955	"	51,720	141,293	—	16,836
JUNE 1956	"	53,640	143,213	—	16,486
1957	"	56,804	146,377	—	16,316
1958	"	60,278	149,851	—	16,065
1959	"	64,171	153,744	—	15,923
1960	"	69,205	158,778	—	15,936
1961	"	72,877	162,450	—	15,967
1962	"	76,986	166,559	—	15,884
1963	"	81,443	171,016	—	15,739
1964	"	86,711	176,284	—	15,732
1965	"	93,064	182,637	—	15,755
1966	"	101,745	191,318	—	15,933
1967	"	106,647	196,220	—	16,003
1968	"	(6)	(6)	—	16,230 <sup>5</sup>
1969	"	(6)	(6)	—	16,318 <sup>5</sup>

SOURCE: U.S. INTERSTATE COMMERCE COMMISSION ANNUAL REPORTS [17].

1 EXCLUDING BROKERS AND WITHIN-STATE CARRIERS.

2 PROPERTY CARRIERS WERE THE FOLLOWING PERCENTAGES OF ALL OPERATING CARRIERS: 1944-93.4%; 1950-92.4%; 1960-93.0%; 1966-93.4%.

3 ESTIMATED.

4 NOT AVAILABLE; ASSUMED TO BE APPROXIMATELY CONSTANT.

5 1968 AND 1969 FIGURES ARE FOR NUMBER OF CARRIERS REQUIRED TO FILE ANNUAL REPORTS.

6 NOT AVAILABLE COMPARABLE TO PREVIOUS YEARS; APPLICATIONS FOR PERMANENT AUTHORITY DISPOSED OF (I.E., FROM NEW AND PENDING FILES) 1967-69 ARE AS FOLLOWS: 1967-7,049; 1968-5,724; 1969-5,186.

TABLE 6

WEIGHT LIMITS ON TRUCKS, 1932-33\*, BY STATES (BASIC DATA FOR TABLE 2).

STATE	MAXIMUM WEIGHT (IN LBS.)		STATE	MAXIMUM WEIGHT (IN LBS.)	
	4-WHEEL <sup>1</sup>	6-WHEEL <sup>2</sup>		4-WHEEL <sup>1</sup>	6-WHEEL <sup>2</sup>
ALABAMA	20,000	32,000	NEBRASKA	24,000	40,000
ARIZONA	22,000	34,000	NEVADA	25,000	38,000
ARKANSAS	22,200	37,000	NEW HAMPSHIRE	20,000	20,000
CALIFORNIA	22,000	34,000	NEW JERSEY	30,000	30,000
COLORADO	30,000	40,000	NEW MEXICO	27,000	45,000
CONNECTICUT	32,000	40,000	NEW YORK	33,600	44,000
DELAWARE	26,000	38,000	NO. CAROLINA	20,000	20,000
FLORIDA	20,000	20,000	NO. DAKOTA	24,000	48,000
GEORGIA	22,000	39,600	OHIO	24,000	24,000
IDAHO	24,000	40,000	OKLAHOMA	20,000	20,000
ILLINOIS	24,000	40,000	OREGON	25,500	42,500
INDIANA	24,000	40,000	PENNSYLVANIA	26,000	36,000
IOWA	24,000	40,000	RHODE ISLAND	28,000	40,000
KANSAS	24,000	34,000	SO. CAROLINA	20,000	25,000
KENTUCKY	18,000	18,000	SO. DAKOTA	20,000	20,000
LOUISIANA	13,400	N. A.	TENNESSEE	20,000	20,000
MAINE	18,000	27,000	TEXAS	13,500	N. A.
MARYLAND	25,000	40,000	UTAH	26,000	34,000
MASSACHUSETTS	30,000	30,000	VERMONT	20,000	20,000
MICHIGAN	27,000	45,000	VIRGINIA	24,000	35,000
MINNESOTA	27,000	42,000	WASHINGTON	24,000	34,000
MISSISSIPPI	18,000	22,000	WEST VA.	24,000	40,000
MISSOURI	24,000	24,000	WISCONSIN	24,000	36,000
MONTANA	24,000	34,000	WYOMING	27,000	30,000

\* RED BOOK [11] FIGURES ARE REPORTED (P. 89) AS "BASED ON THE STATE'S INTERPRETATIONS OF THEIR LAWS [1933] AND ON PHYSICAL LIMITATIONS OF VEHICLE DESIGN AND TIRE CAPACITY." PUBLIC ROADS [13] FIGURES ARE REPORTED (P. 167) AS "AN ABSTRACT OF STATE LAWS, INCLUDING LEGISLATION PASSED IN 1932."

1. 4-WHEEL: THE SMALLEST OF THE FOLLOWING 3 FIGURES WAS USED:

- (A) MAXIMUM GROSS WEIGHT (AS GIVEN IN RED BOOK, P. 90-91).
- (B) MAXIMUM AXLE WEIGHT (AS GIVEN IN RED BOOK, P. 90-91), MULTIPLIED BY 1.5 (SEE RED BOOK, P. 89).
- (C) MAXIMUM GROSS WEIGHT (AS GIVEN IN RED BOOK, P. 93).

EXCEPTIONS: TEXAS AND LOUISIANA—SEE RED BOOK, P. 91.

2. 6-WHEEL: MAXIMUM GROSS WEIGHT AS GIVEN IN PUBLIC ROADS, P. 167. THESE FIGURES AGREE IN MOST CASES WITH THOSE SHOWN IN RED BOOK, P. 93, AND WITH PUBLIC ROADS MAXIMUM AXLE WEIGHTS MULTIPLIED BY 2.5 (SEE RED BOOK, P. 93). TEXAS AND LOUISIANA ARE EXCLUDED AS DATA ARE NOT AVAILABLE TO CONVERT FROM PAYLOAD TO GROSS WEIGHT LIMITS.

TABLE 7

INDEPENDENT VARIABLES  
(BASIC DATA FOR TABLE 2 – CONT'D)

STATE	TRUCKS ON FARMS PER 1,000 AGRICULTURAL LABOR FORCE	AVERAGE LENGTH OF RAILROAD HAUL OF FREIGHT (MILES)	PERCENT OF STATE HIGHWAYS WITH HIGH-TYPE SURFACE
ALABAMA	26.05	189.4	1.57
ARIZONA	79.74	282.2	2.60
ARKANSAS	28.62	233.1	1.72
CALIFORNIA	123.40	264.6	13.10
COLORADO	159.50	244.7	0.58
CONNECTICUT	173.80	132.6	7.98
DELAWARE	173.20	202.7	21.40
FLORIDA	91.41	184.1	8.22
GEORGIA	32.07	165.7	1.60
IDAHO	95.89	243.6	0.73
ILLINOIS	114.70	207.9	9.85
INDIANA	120.20	202.8	6.90
IOWA	98.73	233.3	3.39
KANSAS	146.70	281.5	0.94
KENTUCKY	20.05	227.5	1.81
LOUISIANA	31.27	201.0	1.94
MAINE	209.30	120.4	1.87
MARYLAND	134.20	184.1	12.90
MASSACHUSETTS	172.20	144.7	17.70
MICHIGAN	148.40	168.0	6.68
MINNESOTA	120.40	225.6	1.44
MISSISSIPPI	29.62	164.9	1.14
MISSOURI	54.28	229.7	2.91
MONTANA	183.80	266.5	0.09
NEBRASKA	132.10	266.9	0.41
NEVADA	139.40	273.2	0.39
NEW HAMPSHIRE	205.40	129.0	3.42
NEW JERSEY	230.20	137.6	23.30
NEW MEXICO	90.46	279.0	0.18
NEW YORK	220.50	163.3	21.50
NO. CAROLINA	37.12	171.5	8.61
NO. DAKOTA	126.40	255.1	0.01
OHIO	125.80	194.2	11.20
OKLAHOMA	78.18	223.3	1.42
OREGON	118.90	246.2	3.35
PENNSYLVANIA	187.60	166.5	9.78
RHODE ISLAND	193.30	131.0	20.40
SO. CAROLINA	20.21	169.8	2.82
SO. DAKOTA	113.40	216.6	0.04
TENNESSEE	23.98	191.9	3.97
UTAH	101.70	235.7	1.69
VERMONT	132.20	109.7	2.26
VIRGINIA	71.88	229.8	2.86
WASHINGTON	180.90	254.4	4.21
WEST VIRGINIA	62.88	218.7	8.13
WISCONSIN	178.60	195.7	4.57
WYOMING	133.40	286.7	0.08

- (1) AVERAGE LENGTH OF RR HAUL OF (REVENUE) FREIGHT = AVERAGE DISTANCE IN MILES EACH TON IS CARRIED = RATIO OF NUMBER OF TON-MILES TO NUMBER OF TONS CARRIED. FOR EACH STATE, AVERAGE LENGTH OF HAUL WAS OBTAINED BY WEIGHTING AVERAGE LENGTH OF HAUL OF EACH COMPANY BY THE NUMBER OF MILES OF LINE OPERATED BY THAT COMPANY IN THE STATE (ALL FOR CLASS I RR'S).
- (2) PERCENTAGE OF STATE ROADS WITH HIGH-QUALITY SURFACE: WHERE HIGH-QUALITY (HIGH-TYPE) SURFACE CONSISTS OF BITUMINOUS MACADAM, BITUMINOUS CONCRETE, SHEET ASPHALT, PORTLAND CEMENT CONCRETE, AND BLOCK PAVEMENTS. ALL STATE RURAL ROADS, BOTH LOCAL AND STATE HIGHWAYS SYSTEMS, ARE INCLUDED.

## References

1. ASSOCIATION OF AMERICAN RAILROADS, BUREAU OF RAILWAY ECONOMICS. *Railway Mileage by States*. Washington, D. C.: December 31, 1930.
2. BECKER, G. S. "Competition and Democracy." *Journal of Law and Economics*, October 1958.
3. THE COUNCIL OF STATE GOVERNMENTS. "Occupational Licensing Legislation in the States." 1952.
4. DEMSETZ, H., "Why Regulate Utilities?" *Journal of Law and Economics*, April 1968.
5. GERWIG, R. W. "Natural Gas Production: A Study of Costs of Regulation." *Journal of Law and Economics*, October 1962, pp. 69-92.
6. HOLEN, A. S. "Effects of Professional Licensing Arrangements on Interstate Labor Mobility and Resource Allocation." *Journal of Political Economy*, Vol. 73 (1915), pp. 492-98.
7. KEYES, L. S. *Federal Control of Entry into Air Transportation*. Cambridge, Mass.: Harvard University Press, 1951.
8. MOORE, T. G. "The Purpose of Licensing." *Journal of Law and Economics*, October 1961.
9. OLSON, M. *The Logic of Collective Action*. Cambridge, Mass.: Harvard University Press, 1965.
10. PELTZMAN, S. "Entry in Commercial Banking." *Journal of Law and Economics*, October 1965.
11. *The Motor Truck Red Book and Directory*, 1934 Edition, pp. 85-102.
12. U. S. CONGRESS, SELECT COMMITTEE ON SMALL BUSINESS. *Hearings*, 88th Congress, 2nd Session, August 10 and 11, 1964.
13. U. S. DEPARTMENT OF AGRICULTURE, BUREAU OF PUBLIC ROADS. *Public Roads*. Washington, D. C.: U. S. Government Printing Office, December 1932.
14. U. S. DEPARTMENT OF COMMERCE, BUREAU OF THE CENSUS. *United States Census of Agriculture, 1930*, Vol. 4. Washington, D. C.: U. S. Government Printing Office, 1930.
15. ———. *United States Census of Population*. Washington, D. C.: U. S. Government Printing Office, appropriate years.
16. ———, BUREAU OF FOREIGN AND DOMESTIC COMMERCE. *Statistical Abstract of the U. S.*, 1932. Washington, D. C.: U. S. Government Printing Office, 1932.
17. U. S. INTERSTATE COMMERCE COMMISSION. *Annual Report*. Washington, D. C.: U. S. Government Printing Office, appropriate years.
18. ———. *Statistics of Railways in the United States, 1930*. Washington, D. C.: U. S. Government Printing Office, 1930.

---

## Libertarian Paternalism

Author(s): Richard H. Thaler and Cass R. Sunstein

Source: *The American Economic Review*, May, 2003, Vol. 93, No. 2, Papers and Proceedings of the One Hundred Fifteenth Annual Meeting of the American Economic Association, Washington, DC, January 3-5, 2003 (May, 2003), pp. 175-179

Published by: American Economic Association

Stable URL: <https://www.jstor.org/stable/3132220>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*

# BEHAVIORAL ECONOMICS, PUBLIC POLICY, AND PATERNALISM<sup>†</sup>

## Libertarian Paternalism

By RICHARD H. THALER AND CASS R. SUNSTEIN\*

Many economists are libertarians and consider the term “paternalistic” to be derogatory. Most would think that the phrase libertarian paternalism is an oxymoron. The modest goal of this essay is to encourage economists to rethink their views on paternalism. We believe that the anti-paternalistic fervor expressed by many economists is based on a combination of a false assumption and at least two misconceptions. The false assumption is that people always (usually?) make choices that are in their best interest. This claim is either tautological, and therefore uninteresting, or testable. We claim that it is testable and false—indeed, obviously false.

The first misconception is that there are viable alternatives to paternalism. In many situations, some organization or agent *must* make a choice that will affect the choices of some other people. The point applies to both private and public actors. Consider the problem facing the director of a company cafeteria who discovers that the order in which food is arranged influences the choices people make. To simplify, consider three alternative strategies: (1) she could make choices that she thinks would make the customers best off; (2) she could make choices at random; or (3) she could maliciously choose those items that she thinks would make the customers as obese as possible. Option 1

appears to be paternalistic, which it is, but would anyone advocate options 2 or 3?

The second misconception is that paternalism always involves coercion. As the cafeteria example illustrates, the choice of which order to present food items does not coerce anyone to do anything, yet one might prefer some orders to others on paternalistic grounds. Would many object to putting the fruit before the desserts at an elementary school cafeteria if the outcome were to increase the consumption ratio of apples to Twinkies? Is this question fundamentally different if the customers are adults? If no coercion is involved, we think that some types of paternalism should be acceptable to even the most ardent libertarian. We call such actions *libertarian paternalism*.

In our understanding, a policy counts as “paternalistic” if it is selected with the goal of influencing the choices of affected parties in a way that will make those parties better off. We intend “better off” to be measured as objectively as possible, and we clearly do not always equate revealed preference with welfare. That is, we emphasize the possibility that in some cases individuals make inferior choices, choices that they would change if they had complete information, unlimited cognitive abilities, and no lack of willpower. Once it is understood that some organizational decisions are inevitable, that a form of paternalism cannot be avoided, and that the alternatives to paternalism (such as choosing options to make people sick, obese, or generally worse off) are unattractive, we can abandon the less interesting question of whether to be paternalistic or not and turn to the more constructive question of how to choose among paternalistic options.<sup>1</sup>

<sup>†</sup> *Discussants:* Robert E. Hall, Stanford University; Gary Becker, University of Chicago; Robert Barro, Harvard University.

\* Thaler: Department of Economics and Behavioral Science, University of Chicago Graduate School of Business, 1101 E. 58th Street, Chicago, IL 60637 (e-mail: thaler@gsb.uchicago.edu); Sunstein: Law School and Department of Political Science, University of Chicago, 1111 E. 60th Street, Chicago, IL 60637 (e-mail: csunstei@midway.uchicago.edu). We are grateful to Richard Epstein, Russ Fuller, Owen Lamont, Eric Posner, Richard Posner, and David Wilcox for valuable comments on a previous draft.

<sup>1</sup> Readers interested in this topic should also consult Colin Camerer et al. (2001) for an illuminating discussion of related issues. That paper shares with the papers in this

### I. Are Choices Rational?

The presumption that individual choices should be free from interference is usually based on the assumption that people do a good job of making choices, or at least that they do a far better job than third parties could do. As far as we can tell, there is little empirical support for this claim. Research by psychologists and economists over the past three decades has raised questions about the rationality of the judgments and decisions that individuals make. People do not exhibit rational expectations, fail to make forecasts that are consistent with Bayes' rule, use heuristics that lead them to make systematic blunders, exhibit preference reversals (that is, they prefer A to B and B to A) and make different choices depending on the wording of the problem (for many examples, see the two recent collections of papers by Daniel Kahneman and Amos Tversky [2000] and by Thomas Gilovich et al. [2002]). Furthermore, in the context of intertemporal choice, people exhibit dynamic inconsistency, valuing present consumption much more than future consumption. In other words, people have self-control problems (see the other papers in this session [James Choi et al., 2003b; Ted O'Donoghue and Matthew Rabin, 2003] for details and references).

Many economists are skeptical of some of these findings, thinking that people may do a better job of choosing in the "real world" than they do in the laboratory. However, studies of actual choices for high stakes reveal many of the same problems. For example, the Surgeon General reports that 61 percent of Americans are either overweight or obese. Given the adverse effects obesity has on health, it is hard to claim that Americans are eating optimal diets.

Another illustration comes from the domain of savings behavior. Shlomo Benartzi and Thaler (2002) investigate how much investors like the portfolios they have selected in their defined-contribution savings plans. Employees volunteered to share their portfolio choices with the investigators (by bringing a copy of their

---

session the common goal of devising policies that help some agents who are making some mistake, while minimizing the costs imposed on others.

most recent statement to the lab). They were then shown the probability distributions of expected retirement income for three investment portfolios just labeled A, B, and C. Unbeknownst to the subjects, the three portfolios were their own and portfolios mimicking the average and median choices of their fellow employees. The distributions of expected returns were computed using the software of Financial Engines, the financial information company founded by William Sharpe. On average, the subjects rated the average portfolio equally with their own portfolio, and they judged the median portfolio to be significantly more attractive than their own. Indeed, only 20 percent of the subjects preferred their own portfolio to the median portfolio. Apparently, people do not gain much by choosing investment portfolios for themselves.

### II. Is Paternalism Inevitable?

As the cafeteria line example discussed above illustrates, planners are forced to make some design choices. A simple and important example is the selection of a "default option" to determine what happens if an agent fails to choose for himself. In a fully rational world such design choices would have little effect (at least in high-stakes situations) because agents would simply choose the best option for them regardless of the default. However, numerous experiments illustrate that there is a very strong "status quo" bias (see William Samuelson and Richard Zeckhauser, 1988; Kahneman et al., 1991). The existing arrangement, whether set out by private institutions or by government, tends to stick.

One illustration of this phenomenon comes from studies of automatic enrollment in 401(k) employee savings plans. Most 401(k) plans use an opt-in design. When employees first become eligible to participate in the 401(k) plan, they receive some plan information and an enrollment form that must be completed in order to join. Under the alternative of automatic enrollment, employees receive the same information but are told that unless they opt out, they will be enrolled in the plan (with some default options for savings rates and asset allocation). In companies that offer a "match" (the employer matches the employee's contributions accord-

ing to some formula, often a 50-percent match up to some cap), most employees eventually do join the plan, but enrollments occur much sooner under automatic enrollment. For example, Brigitte Madrian and Dennis Shea (2001) found that initial enrollments jumped from 49 percent to 86 percent, and Choi et al. (2002) find similar results for other companies.

Should the adoption of automatic enrollment be considered paternalistic? And, if so, should it therefore be seen as a kind of officious meddling with employee preferences? We answer these questions yes and no respectively. If the employer thinks (correctly, we believe) that most employees would prefer to join the 401(k) plan if they took the time to think about it and did not lose the enrollment form, then by choosing automatic enrollment they are acting paternalistically. They are attempting to steer employees' choices in directions that will promote employees' welfare. But since no one is forced to do anything, we think this steering should be considered unobjectionable to libertarians. The employer must choose some set of rules, and either plan affects employees' choices. No law of nature says that, in the absence of an affirmative election by employees, zero percent of earnings will go into a retirement plan. Because both plans alter choices, neither one can be said, more than the other, to count as a form of objectionable meddling.

Quick-minded readers might be tempted to think that there is a way out of this dilemma. Employers could avoid choosing a default if they *required* employees to make a choice, either in or out. But some thought reveals that this is not at all a way out of the dilemma; rather, it is simply another option among many that the employer can elect. In fact, Choi et al. (2003a) find that this rule increases enrollments (relative to the opt-in rule) though by not as much as automatic enrollment. Furthermore, the very requirement that employees make a choice has a paternalistic element. Many employees do not want to have to make a choice (and would choose not to have to do so). Should employers really force them to choose?

Why, exactly, does the setting of defaults have such large effects? With respect to savings, the designated default plan apparently carries a certain legitimacy for many employees,

perhaps because it seems to have resulted from some conscious thought about what makes most sense for most people. But there is a separate explanation, involving inertia. For any employee, a change from any status quo entails time and effort, and many people seem to prefer to avoid both of these, especially if they are prone to procrastination. When default rules are "sticky" and affect choices as a result, inertia might be the major reason.

For present purposes, the choice among these various explanations does not much matter. The point is only that paternalism, in the form of effects on individual choices, is often unavoidable. When paternalism seems absent, it is usually because the starting point appears so natural and obvious that its preference-shaping effects are invisible to most observers. But those effects are nonetheless there. Of course it is usually good not to block choices, and we do not mean to defend non-libertarian paternalism here. But in an important respect, the anti-paternalistic position is incoherent.

### III. Beyond the Inevitable (but Still Libertarian)

The inevitability of paternalism is most clear when the planner has to choose default rules. It is reasonable to ask whether the planner should go beyond the inevitable. Take the cafeteria example discussed above. Putting the fruit before the desserts is a fairly mild intervention. A more intrusive step would be to place the desserts in another location altogether, so that diners have to get up and get a dessert after they have finished the rest of their meal. This step raises the transactions costs of eating dessert, and according to a standard economic analysis the proposal is unattractive: it seems to make dessert-eaters worse off and no one better off. But once self-control costs are incorporated, we can see that some diners would prefer this arrangement, namely, those who would eat a dessert if it were put in front of them but would resist temptation if given a little help. To the extent that the dessert location is not hard to find, and no choice is forbidden, this approach meets libertarian muster.

In the domain of employee savings, Thaler and Benartzi (2003) have proposed a method of increasing contributions to 401(k) plans that also meets the libertarian test. Under this plan,

called Save More Tomorrow, employees are invited to sign up for a program in which their contributions to the savings plan are increased annually whenever they get a raise. Once employees join the plan, they stay in until they opt out or reach the maximum savings rate in the plan. In the first company to use this plan, the employees who joined increased their savings rates from 3.5 percent to 11.6 percent in a little over two years (three raises). Very few of the employees who join the plan drop out. This is successful libertarian paternalism in action.

#### **IV. How to Choose: The Toolbox of the Libertarian Paternalist**

How should sensible planners (a category we mean to include anyone who must design plans for others, from human-resource directors to bureaucrats to kings) choose among possible systems, given that some choice is necessary? We suggest two approaches to this problem.

If feasible, a comparison of possible rules should be done using a form of cost-benefit analysis. The goal of a cost-benefit study would be to measure the full ramifications of any design choice. To illustrate, take the example of automatic enrollment. Under automatic enrollment some employees will join the plan who otherwise would not. Presumably, some are made better off (especially if there is an employer match), but some may be made worse off (e.g., those who are highly liquidity-constrained). If the issue were just enrollment, we would guess that the gains would exceed the losses. We base this guess partly on revealed choices. Most employees do join the plan eventually, and very few who are automatically enrolled opt out when they figure out what has happened to them. We also judge that the costs of having too little saved up for retirement are typically greater than the costs of having saved too much.

In many cases, however, the planner will be unable to make a direct inquiry into welfare, either because too little information is available or because the costs of doing the analysis are not warranted. The committed anti-paternalist might say, in such cases, that people should simply be permitted to choose as they see fit. We hope that we have said enough to show why this response is unhelpful. What people choose

often depends on the starting point, and hence the starting point cannot be selected by asking what people choose. In these circumstances, the libertarian paternalist would seek indirect proxies for welfare: methods that test whether one or another approach is welfare-promoting without relying on unreliable guesswork about that question. We suggest three possible methods.

First, the libertarian paternalist might select the approach *that the majority would choose if explicit choices were required and revealed*. Useful though it is, this market-mimicking approach raises its own problems. Perhaps the majority's choices would be insufficiently informed. Perhaps those choices, in fact, would not promote the majority's welfare. At least as a presumption, however, it makes sense to follow those choices, if the planner knows what they would be. A deeper problem is that the majority's choices might themselves be a function of the starting point or the default rule. If so, the problem of circularity dooms the market-mimicking approach. But in some cases, at least, the majority is likely to go one way or the other regardless of the starting point; and to that extent, the market-mimicking strategy seems quite workable.

Second, the libertarian paternalist might select the approach *that would force people to make their choices explicit*. This approach might be chosen if the market-mimicking strategy fails, either because of the circularity problem or because the planner does not know which approach would in fact be chosen by the majority. We have seen the possibility of forced choices in the context of retirement plans; it would be easy to find other examples. Here too, however, there is a risk that the choices that are actually elicited will be inadequately informed or will not promote welfare. In the case of retirement plans, for example, forced choices have been found to produce higher participation rates than requiring opt-ins, but lower rates than requiring opt-outs. If it is likely that automatic enrollment is welfare-promoting, perhaps automatic enrollment should be preferred over forced choices. The only suggestion is that, where the social planner is unsure how to handle the welfare question, he might devise a strategy that requires people to choose.

Third, the libertarian paternalist might select the approach *that minimizes the number of opt-*

*outs.* For example, very few employees opt out of the 401(k) plan when they are automatically enrolled, though many opt in under the standard enrollment procedure. This is an *ex post* inquiry into people's preferences, in contrast to the *ex ante* approach favored by the market-mimicking strategy. With those numbers, there is reason to think that automatic enrollment is better, if only because more people are sufficiently satisfied to leave it in place.

## V. Conclusion

Our goal here has been to defend libertarian paternalism, an approach that preserves freedom of choice but that authorizes both private and public institutions to steer people in directions that will promote their welfare. Some kind of paternalism is likely whenever such institutions set out arrangements that will prevail unless people affirmatively choose otherwise. In these circumstances, the goal should be to avoid random, arbitrary, or harmful effects and to produce a situation that is likely to promote people's welfare, suitably defined.

## REFERENCES

- Benartzi, Shlomo and Thaler, Richard H.** "How Much Is Investor Autonomy Worth?" *Journal of Finance*, August 2002, 57(4), pp. 1593–1616.
- Camerer, Colin; Issacharoff, Samuel; Loewenstein, George; O'Donoghue, Ted and Rabin, Matthew.** "Regulation for Conservatives: Behavioral Economics and the Case for Asymmetric Paternalism." Working paper, California Institute of Technology, 2001.
- Choi, James; Laibson, David; Madrian, Brigitte C. and Metrick, Andrew.** "Defined Contribution Pensions: Plan Rules, Participant Decisions, and the Path of Least Resistance," in James M. Poterba, ed., *Tax policy and the economy*, Vol. 16. Cambridge, MA: MIT Press, 2002, pp. 67–113.
- \_\_\_\_\_. "Benign Paternalism and Active Decisions: A Natural Experiment." Working paper, University of Chicago, Graduate School of Business, 2003a.
- \_\_\_\_\_. "Optimal Defaults." *American Economic Review*, May 2003b (*Papers and Proceedings*), 93(2), pp. 180–85.
- Gilovich, Thomas; Griffin, Dale and Kahneman, Daniel**, eds. *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, U.K.: Cambridge University Press, 2002.
- Kahneman, Daniel; Knetsch, Jack L. and Thaler, Richard H.** "The Endowment Effect, Loss Aversion, and Status Quo Bias." *Journal of Economic Perspectives*, Winter 1991, 5(1), pp. 193–206.
- Kahneman, Daniel and Tversky, Amos**, eds. *Choices, values, and frames*. Cambridge, U.K.: Cambridge University Press, 2000.
- Madrian, Brigitte and Shea, Dennis.** "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics*, November 2001, 116(4), pp. 1149–87.
- O'Donoghue, Ted and Rabin, Matthew.** "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes." *American Economic Review*, May 2003 (*Papers and Proceedings*), 93(2), pp. 186–91.
- Samuelson, William and Zeckhauser, Richard J.** "Status Quo Bias in Decision Making." *Journal of Risk and Uncertainty*, March 1988, 1(1), pp. 7–59.
- Thaler, Richard H. and Benartzi, Shlomo.** "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy*, 2003 (forthcoming).

NBER WORKING PAPER SERIES

PSYCHOLOGY AND ECONOMICS:  
EVIDENCE FROM THE FIELD

Stefano DellaVigna

Working Paper 13420  
<http://www.nber.org/papers/w13420>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
September 2007

I would like to thank Roger Gordon (the editor), two anonymous referees, Dan Acland, Malcolm Baker, Brad Barber, Nicholas Barberis, Saurabh Bhargava, Colin Camerer, David Card, Raj Chetty, James Choi, Sanjit Dhami, Constanca Esteves, Ernst Fehr, Shane Frederick, Drew Fudenberg, David Hirshleifer, Eric Johnson, Lawrence F. Katz, Georg Kirchsteiger, Jeffrey Kling, Howard Kunreuther, David Laibson, Erzo F.P. Luttmer, Rosario Macera, Ulrike Malmendier, MichelAndre Marechal, John Morgan, Ted O'Donoghue, Ignacio Palacios-Huerta, Joshua Palmer, Vikram Pathania, Matthew Rabin, Ricardo Reis, Uri Simonsohn, Rani Spiegler, Bjarne Steffen, Justin Sydnor, Richard Thaler, Jeremy Tobacman, Michael Urbancic, Ebonya Washington, Kathryn Zeiler, and Jonathan Zinman for useful comments and suggestions. Thomas Barrios and Charles Lin provided excellent research assistance. I also want to thank the students of my class in Psychology and Economics who over the years helped shape the ideas in this paper. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2007 by Stefano DellaVigna. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Psychology and Economics: Evidence from the Field  
Stefano DellaVigna  
NBER Working Paper No. 13420  
September 2007  
JEL No. A1,C91,C93,D00,D64,D91,G1,M3

## **ABSTRACT**

The research in Psychology and Economics (a.k.a. Behavioral Economics) suggests that individuals deviate from the standard model in three respects: (i) non-standard preferences; (ii) non-standard beliefs; and (iii) non-standard decision-making. In this paper, I survey the empirical evidence from the field on these three classes of deviations. The evidence covers a number of applications, from consumption to finance, from crime to voting, from giving to labor supply. In the class of non-standard preferences, I discuss time preferences (self-control problems), risk preferences (reference dependence), and social preferences. On non-standard beliefs, I present evidence on overconfidence, on the law of small numbers, and on projection bias. Regarding non-standard decision-making, I cover limited attention, menu effects, persuasion and social pressure, and emotions. I also present evidence on how rational actors -- firms, employers, CEOs, investors, and politicians -- respond to the non-standard behavior described in the survey. I then summarize five common empirical methodologies used in Psychology and Economics. Finally, I briefly discuss under what conditions experience and market interactions limit the impact of the non-standard features.

Stefano DellaVigna  
UC, Berkeley  
Department of Economics  
549 Evans Hall #3880  
Berkeley, CA 94720-3880  
and NBER  
[sdellavi@econ.berkeley.edu](mailto:sdellavi@econ.berkeley.edu)

# 1 Introduction

The core theory used in economics builds on a simple but powerful model of behavior. Individuals make choices so as to maximize a utility function, using the information available, and processing this information appropriately. Individuals' preferences are assumed to be time-consistent and independent of the framing of the decision.

Many attempts to test these assumptions through laboratory experiments in both the psychology and the economics literature raise serious questions, though. In the laboratory, individuals are time-inconsistent (Thaler, 1981), show a concern for the welfare of others (Charness and Rabin 2002, Fehr and Gächter 2000), and exhibit an attitude toward risk that depends on framing and reference points (Kahneman and Tversky, 1979). They violate rational expectations, for example by overestimating their own skills (Camerer and Lovallo, 1999) and overprojecting from the current state (Read and van Leeuwen, 1998). They use heuristics to solve complex problems (Gabaix, Laibson, Moloche, and Weinberg, 2006) and are affected by transient emotions in their decisions (Loewenstein and Lerner, 2003).

Unclear from these experiments, though, is how much these deviations from the standard theory in the laboratory affect economic decisions in the field. In markets people hone their behavioral rules to match the incentives they face and sort into favorable economic settings (Levitt and List, fs2007). This is likely to limit the impact of deviations from the standard model in markets. However, other forces are likely to increase the impact. Important economic decisions such as the choice of retirement savings or a house purchase are taken seldom, with limited scope for feedback. In addition, firms often have incentives to accentuate the deviations of consumers to profit from them (DellaVigna and Malmendier, 2004).

The objective of this paper is to summarize a growing list of recent papers that document aspects of behavior in market settings that also deviate from the forecasts of the standard theory. This research area is known as Psychology and Economics (or Behavioral Economics). The evidence suggests deviations from the standard theory in each step of the decision-making process: 1) non-standard preferences, 2) incorrect beliefs, and 3) systematic biases in decision-making. For each of these three steps, I present an example of the laboratory evidence, introduce a simple model if available, and summarize the strength and weaknesses of the field evidence. Since the focus of the paper is on the field evidence, I do not survey the laboratory evidence or the theoretical literature.

To fix ideas, consider the following stylized version of the standard model, modified from Rabin (2002a). Individual  $i$  at time  $t = 0$  maximizes expected utility subject to a probability

distribution  $p(s)$  of the states of the world  $s \in S$ :

$$\max_{x_i^t \in X_i} \sum_{t=0}^{\infty} \delta^t \sum_{s_t \in S_t} p(s_t) U(x_i^t | s_t). \quad (1)$$

The utility function  $U(x|s)$  is defined over the payoff  $x_i^t$  of player  $i$  and future utility is discounted with a (time-consistent) discount factor  $\delta$ .

The first class of deviations from the standard model in (1) is non-standard preferences, discussed in Section 2. I focus on three dimensions: time preferences, risk preferences, and social preferences. With respect to time preferences, the findings on self-control problems, for example in retirement savings, challenge the assumption of a time-consistent discount factor  $\delta$ . With respect to risk preferences, the evidence such as on insurance decisions suggests that the utility function  $U(x_i|s)$  depends on a reference point  $r$ : the utility function becomes  $U(x_i|r, s)$ . With respect to social preferences, the evidence, for example on charitable giving, suggests that the utility function depends also on the payoff of other people  $x_{-i}$ : the utility is  $U(x_i, x_{-i}|s)$ . The research on non-standard preferences constitutes the bulk of the empirical research in Psychology and Economics.

The second class of deviations from the standard model in (1) is non-standard beliefs  $\tilde{p}(s) \neq p(s)$ , reviewed in Section 3. Systematic overconfidence about own ability can help explain managerial behavior of CEOs. Non-Bayesian forecasting rationalizes ‘gambler’s fallacy’ behavior in lotteries and overinference from past stock returns. The overprojection of current tastes on future tastes can explain aspects of the purchase of seasonal items.

The third class of deviations from the standard model is non-standard decision-making, discussed in Section 4. For given utility  $U(x|s)$  and beliefs  $p(s)$ , individuals resort to heuristics (Tversky and Kahneman, 1974) instead of solving the complex maximization problem (1). They simplify a complex decision by being inattentive to less salient features of a problem, from asset allocation to purchase decisions. They use sub-optimal heuristics when choosing from a menu of options  $X_i$ , such as for savings plans or loan terms. They are also subject to social pressure and persuasion, for example in their workplace performance and in voting decisions. Finally, they are affected by emotions, as in the case of investment decisions.

While I organize the deviations in three separate classes, the three types of deviations are often related. For example, persuasion leads to a different decision through the change in beliefs that it induces.

Are these deviations large enough to matter for our theories of how markets and institutions work? A key test for Psychology and Economics is whether it helps to understand markets and institutions. In Section 5, I provide evidence on how rational actors respond to these behavioral anomalies. In particular, I discuss the response of firms, employers, managers, investors, and politicians. These agents appear to have changed their own behavior in ways that would be puzzling given the standard theory but that are consistent with utility-maximizing responses

to the documented behavioral anomalies.

Following the summary of the evidence, in Section 6 I discuss the pros and cons of the five types of evidence used in Psychology and Economics: (i) Menu Choice; (ii) Natural Experiments; (iii) Field Experiments; (iv) Correlational Studies; and (v) Structural Identification.

Given this evidence, I expect that the documented deviations from the standard model will be increasingly incorporated in economic models. Indeed, features such as time inconsistency and reference dependence have become common assumptions. In the concluding Section, I present final remarks on why these deviations matter also in the field and discuss directions for future research in Psychology and Economics.

This overview differs from other surveys of Psychology and Economics (Rabin, 1998; Rabin, 2002a; Mullainathan and Thaler, 2001; Camerer, 2005) because it focuses on empirical research using non-laboratory data. A number of caveats are in order. First, this paper, being organized by psychological principles, does not provide an overview by field of application; the interested reader can consult as a starting point the book chapters in Diamond and Vartiainen (2007). Second, the emphasis of the paper is on (relatively) detailed summaries of a small number of papers for each deviation. As such, the survey provides a selective coverage of the field evidence, though it strives to cover all the important deviations.<sup>1</sup> Finally, this overview undersamples empirical studies in Marketing and provides a partial coverage of the research in Behavioral Finance, probably the most developed application of Psychology and Economics, for which a comprehensive survey of the empirical findings is available (Barberis and Thaler, 2004).

## 2 Non-standard Preferences

### 2.1 Self-Control Problems

The standard model (1) assumes a discount factor  $\delta$  between any two time periods that is independent of when the utility is evaluated. This assumption implies time consistency, that is, the decision maker has the same preferences about future plans at different points in time.<sup>2</sup>

**Laboratory Experiments.** Experiments on intertemporal choice, summarized in Loewenstein and Prelec (1992) and Frederick, Loewenstein, and O'Donoghue (2002), have cast doubt on this assumption. This evidence suggests that discounting is steeper in the immediate future than in the further future. For example, the median subject in Thaler (1981) is indifferent between \$15 now and \$20 in one month (for an annual discount rate of 345 percent) and between

---

<sup>1</sup>This overview does not discuss deviations from the standard model that are widely documented in experiments but not in the field, such as will-power exhaustion and the availability heuristics.

<sup>2</sup>Strictly speaking, the standard model merely assumes time consistency, not a constant discount factor  $\delta$ . Still, most of the evidence in this Section—the adoption of costly commitments or behavior that differs from the plans—directly violates time consistency and hence also this more general version of the standard model.

\$15 now and \$100 in ten years (for an annual discount rate of 19 percent).<sup>3</sup> The preference for immediate gratification captured in these studies appears to have identifiable neural underpinnings. Intertemporal decisions involving payoffs in the present activate different neural systems than decisions involving only payoffs in future periods (McClure et al., 2004).

Intertemporal preferences with these features capture *self-control problems*. When evaluating outcomes in the distant future, individuals are patient and make plans to exercise, stop smoking, and look for a better job. As the future gets near, the discounting gets steep, and the individuals engage in binge eating, light another (last) cigarette, and stay put on their job. Preferences with these features therefore induce time inconsistency.

**Model.** Laibson (1997) and O'Donoghue and Rabin (1999a) formalized these preferences using  $(\beta, \delta)$  preferences<sup>4</sup>, building on Strotz (1956), Phelps and Pollak (1968), and Akerlof (1991). Labelling as  $u_t$  the per-period utility, the overall utility at time  $t$ ,  $U_t$ , is

$$U_t = u_t + \beta\delta u_{t+1} + \beta\delta^2 u_{t+2} + \beta\delta^3 u_{t+3} + \dots$$

The only difference from the standard model (with  $\delta$  as the discount factor) is the parameter  $\beta \leq 1$ , capturing the self-control problems. For  $\beta < 1$ , the discounting between the present and the future is higher than between any future time periods, capturing the main finding of the experiments. For  $\beta = 1$ , this reduces to the standard model.

A second key element in this model is the modelling of expectations about future time preferences. O'Donoghue and Rabin (2001) allow the agent to be partially naive (that is, overconfident) about the future self-control problems. A partially naive  $(\beta, \delta)$  agent expects in the future period  $t+s$  to have the utility function

$$\hat{U}_{t+s} = u_{t+s} + \hat{\beta}\delta u_{t+s+1} + \hat{\beta}\delta^2 u_{t+s+2} + \hat{\beta}\delta^3 u_{t+s+3} + \dots$$

with  $\hat{\beta} \geq \beta$ . The agent may be sophisticated about the self-control problem ( $\hat{\beta} = \beta$ ), fully naive ( $\hat{\beta} = 1$ ), or somewhere in between. This model, therefore, combines self-control problems with a form of overconfidence, naiveté about future self-control.

Other models have been proposed to capture self-control problems, including axiomatic models that emphasize preferences over choice sets (Gul and Pesendorfer, 2001) and models of the conflict between two systems, a planner and a doer (Shefrin and Thaler, 1981 and

---

<sup>3</sup>The laboratory experiments on time preferences face at least three issues: (i) most experiments are over hypothetical choices, including Thaler (1981); (ii) in the experiments with real payments, issues of credibility regarding the future payments can induce seeming present bias; (iii) the discounting should apply to consumption units, rather than to money (in theory, over monetary outcomes, only the interest rate should matter). While none of the experiments fully addresses all three issues, the consistency of the evidence suggests that the phenomenon is genuine.

<sup>4</sup>These preferences are also labelled quasi-hyperbolic preferences, to distinguish them from (pure) hyperbolic preferences, and present-biased preferences.

Fudenberg and Levine, 2006, among others). For lack of space, and since most applied work has referred to the  $(\beta, \delta)$  model, we refer only to this latter model in what follows.

As an example of how the  $(\beta, \delta)$  model operates, consider a good with immediate payoff (relative to a comparison activity)  $b_1$  at  $t = 1$  and delayed payoff  $b_2$  at  $t = 2$ . An investment good, like exercising or searching for a job, has the features  $b_1 < 0$  and  $b_2 > 0$ : the good requires effort at present and delivers happiness tomorrow. Conversely, a leisure good, like consumption of tempting food or watching TV, has the features  $b_1 > 0$  and  $b_2 < 0$ : it provides an immediate reward, at a future cost.

How often does the agent *want* to consume, from an ex ante perspective? If the agent could set consumption one period in advance, at  $t = 0$ , she would consume if  $\beta\delta b_1 + \beta\delta^2 b_2 \geq 0$ , or

$$b_1 + \delta b_2 \geq 0. \quad (2)$$

(Notice that  $\beta$  cancels out, since all payoffs are in the future)

How much does the agent *actually* consume at  $t = 1$ ? The agent consumes if

$$b_1 + \beta\delta b_2 \geq 0. \quad (3)$$

Compared to the desired, optimal consumption, therefore, a  $(\beta, \delta)$  agent consumes too little investment good ( $b_2 > 0$ ) and too much leisure good ( $b_2 < 0$ ). This is the self-control problem in action. In response, a sophisticated agent looks for commitment devices to increase the consumption of investment goods and to reduce the consumption of leisure goods.

Finally, how much does the agent *expect* to consume? The agent expects to consume in the future if

$$b_1 + \hat{\beta}\delta b_2 \geq 0, \quad (4)$$

with  $\hat{\beta} \geq \beta$ . Compared to the actual consumption in (3), the agent overestimates the consumption of the investment good ( $b_2 > 0$ ) and underestimates the consumption of the leisure good ( $b_2 < 0$ ). Naiveté therefore leads to mispredictions of future usage.

I now present evidence on the consumption of investment goods (exercise and homeworks) and leisure goods (credit card take-up and life-cycle savings) that can be interpreted in light of this simple model.

**Exercise.** DellaVigna and Malmendier (2006) use data from three US health clubs offering a choice between a monthly contract  $X_M$  with lump-sum fee  $L$  of approximately \$80 per month and no payment per visit, and a pay-per-visit contract  $X_p$  with fee  $p$  of \$10. Denote by  $E(x_M)|_{X_M}$  the expected number of monthly visits under the monthly contract  $X_M$ . Under the standard model, individuals choosing the monthly contract must believe that  $pE(x_M)|_{X_M} \geq L$ , or  $L/E(x_M)|_{X_M} \leq p$ : the price per expected attendances under the monthly contract should be lower than the fee under payment-per-usage. Otherwise, the individual should have chosen the pay-per-usage treatment. DellaVigna and Malmendier (2006), however, find that health

club users that choose the monthly contract  $X_M$  attend only 4.8 times per month. These users pay \$17 per visit even though they could pay \$10 per visit, a puzzle for the standard model. A model with partially naive  $(\beta, \delta)$  members suggests two explanations for this finding. The users may be purchasing a commitment device to exercise more: the monthly membership reduces the marginal cost of a visit from \$10 to \$0, and helps to align actual attendance in (3) with desired attendance in (2). Alternatively, these agents may be overestimating their future health club attendance, as in (4). Direct survey evidence on expectation of attendance and evidence on contract renewal are most consistent with the latter interpretation.<sup>5</sup>

**Homeworks and Deadlines.** Ariely and Wertenbroch (2002) present evidence on homework completion and deadlines. The subjects are 51 professionals enrolled in a section of a semester-long executive education class at Sloan (MIT), with three homeworks as a requirement. At the beginning of the semester, they set binding deadlines (with a cost of lower grades for delay) for each of the homeworks. According to the standard model, they should set deadlines for the last day of the semester: there is no benefit to setting early deadlines, since the students do not receive feedback on the homeworks, and there is a cost of lower flexibility. (A maximization without constraints is always preferable to one with constraints.) According to a model of self-control, instead, the deadlines provide a useful commitment device. Since homework completion is an investment good ( $b_2 > 0$ ), individuals spend less time on it than they wish to ex ante (compare equations (2) and (3)). A deadline forces the future self to spend more time on the assignment. The results support the self-control model: 68 percent of the deadlines are set for weeks prior to the last week, indicating a demand for commitment.<sup>6</sup>

This result leaves open two issues. First, do the self-set deadlines improve performance relative to a setting with no deadlines? Second, is the deadline setting optimal? If the individuals are partially naive about the self-control, they will under-estimate the demand for commitment (equation (4)). In a second (laboratory) experiment, Ariely and Wertenbroch (2002) address both issues. Sixty students complete three proofreading assignments within 21 days. The control group can turn in each assignment at any time within the 21 days, a first treatment group can choose three deadlines (as in the class-room setting described above), and a second treatment group faces equal-spaced deadlines. The first result is that self-set deadlines indeed improve performance: the first treatment group does significantly better than the control group, detecting 50 percent more errors (on average, 105 versus 70) and earning substantially more as a result (on average, \$13 versus \$5). The second result is that the deadline setting is not optimal: the group with equal-spaced deadlines does significantly better than the other groups, on average detecting 130 errors and earning \$20. This provides evidence of

---

<sup>5</sup>In Section 5, I discuss how the contracts offered by health club companies are consistent with the assumption of naive  $(\beta, \delta)$  consumers (DellaVigna and Malmendier, 2004).

<sup>6</sup>Ariely and Wertenbroch (2002) also compare the performance in this section to the performance in another section with equal-spaced deadlines, with results similar to the ones described below. However, the students are not randomly assigned to the two sections.

partial naiveté about the self-control problems.

**Credit Card Take-up.** Ausubel (1999) provides evidence on credit card usage using a large-scale field experiment run by a credit card company. The company mailed randomized credit card offers, varying both the pre-teaser and the post-teaser interest rates. For example, compared to an offer of 6.9% interest rate for six months and 16% thereafter (the control group), the treatment group ‘Pre’ received a lower pre-teaser rate (4.9% followed by 16%); the treatment group ‘Post’, instead, received a lower post-teaser rate (6.9% followed by 14%). For each offer, Ausubel (1999) observes the response rate and 21 months of history of borrowing for the individuals that take the card. Across these offers, the average balance borrowed in the first 6 months is about \$2,000, while the average balance in the subsequent 15 months is about \$1,000.<sup>7</sup> Given these borrowing rates, the standard theory predicts that the increase in response rate for treatment ‘Post’ (relative to the control group) should be at least as large as for treatment ‘Pre’: neglecting compounded interest,  $15/12 * 2\% * \$1000$  is larger than  $6/12 * 2\% * \$2,000$  (the comparison would only be more favorable for the ‘Post’ treatment if we could observe the balances past 21 months). Instead, the increase in take-up rate for the ‘Pre’ treatment (386 people out of 100,000) is 2.5 times larger than the increase for the ‘Post’ treatment (154 people out of 100,000). Individuals over-respond to the pre-teaser interest rate. Ausubel’s interpretation of this result is that individuals (naively) believe that they will not borrow much on a credit card, past the teaser period. These findings are consistent with underestimation of future consumption for leisure goods, as in (4).

**Life-Cycle Savings.** The  $(\beta, \delta)$  model of self-control can also help explain puzzling features of life-cycle accumulation, historically the first application of these models. Building on Laibson (1997) and Angeletos et al. (2001), Laibson, Repetto and Tobacman (2006) estimate a fully-specified model of life-cycle accumulation with liquid and illiquid saving. They show that the  $(\beta, \delta)$  model can reconcile two facts: high credit card borrowing (11.7 percent of annual income) and substantial illiquid wealth accumulation (216 percent of annual income for the median consumer of age 50-59).<sup>8</sup> Standard models have a hard time explaining both facts, since credit card borrowing implies high impatience, which is at odds with substantial wealth accumulation. The model with self-control problems predicts high spending on liquid assets, but also a high demand for illiquid assets, which work as commitment devices.

Ashraf, Karlan, and Yin (2005) document directly the demand for illiquid savings as a commitment device, and its effect. They offer an account with a commitment device to 842 randomly determined households in the Philippines with a pre-existent bank account. Access to funds in these accounts is constrained to reaching a self-specified savings goal or a self-

---

<sup>7</sup>Of course, the differences in interest rates will affect the borrowing directly, through incentive and selection effects. However, these differences are small enough in the data that we can, to a first approximation, neglect them in these calculations.

<sup>8</sup>The figures (from Laibson et al., 2006) refer to high-school graduates.

specified time period. A control group of 466 households from the same sample is offered a verbal encouragement to save but with no commitment. The results reveal a sizeable demand for commitment, and an impact of commitment on savings. In the treatment group, 202 of 842 households take up the commitment savings product. In this group, savings in the bank after six months are 5.6 percentage points more likely to increase, compared to the control group that received a pure encouragement.<sup>9</sup> The difference is statistically significant. The comparison includes individuals in the treatment group that do not take up the commitment savings product; the treatment-on-the-treated estimate is larger by a factor of 842/202. Benartzi and Thaler (2004), described in Section 5 below, provide evidence of substantial demand for commitment devices in retirement savings in the US.

**Default Effects in 401(k)s.** The evidence on default effects is the final set of findings bearing on self-control problems.<sup>10</sup> Madrian and Shea (2001) consider the effect on the contribution rates in 401(k)s of a change in default. Before the change, the default is non-participation in retirement savings; after the change, the default is participation at a 3% rate in a money market fund. In both cases, employees can override the default with a phone call or by filing a form; also, in both cases, contributions receive a 50 percent match up to 6% of compensation. Madrian and Shea (2001) find that the change in default has a very large impact: one year after joining the company, the participation rate in 401(k)s is 86% for the treatment group and 49% for the control group.

Choi et al. (2004) show that these findings generalize to six companies in different industries with remarkably similar effect sizes. This finding is not limited to retirement choices in the U.S.. Cronqvist and Thaler (2004) examine the choice of retirement funds in Sweden after the privatization of social security in the year 2000. They find that 43.3 percent of new participants choose the default plan, despite the fact that the government encouraged individual choice, and despite the availability of 456 plans. Three years later, after the end of the advertisement campaign encouraging individual choice, the proportion choosing the default plan increased to 91.6 percent. Overall, the finding of large default effects is one of the most robust results in the applied economics literature of the last ten years.<sup>11</sup>

What explains the large default effect for retirement savings? Transaction costs alone are unlikely to explain default effects. Employees can change their retirement decisions at any time using the phone or a written form. Such small transaction costs are dwarfed by the tax advantages of 401(k) investments, particularly in light of the 50 percent match (up to 6% of compensation) in place at the Madrian and Shea (2001) company. At a mean compensation of about \$40,000, the match provides a yearly benefit of \$1,200, assuming a discount rate equal

---

<sup>9</sup>These figures refer to the total bank balance across all accounts for a household, that is, they are not due to switches of savings from an ordinary account to the account with commitment device.

<sup>10</sup>Samuelson and Zeckhauser (1988) is an early paper documenting default effects.

<sup>11</sup>Default effects matter in other decisions, such as contractual choice in health-clubs (DellaVigna and Malmendier, 2006), organ donation (Abadie and Gay, 2006), and car insurance plan choice (Johnson et al, 1993).

to the interest rate. It is hard to imagine transaction costs of this size.

O'Donoghue and Rabin (1999b and 2001) show that naive  $(\beta, \delta)$  agents can display a large default effect even with small transaction costs.<sup>12</sup> Consider a naive  $(\beta, \delta)$  agent that has to decide when to undertake a decision with immediate disutility from transaction costs  $b_1 < 0$  and delayed benefit  $b_2 > 0$ , such as enrolling in retirement savings. This agent would rather postpone this activity, given the self-control problems, as in equation (3). Moreover, this agent is (incorrectly) convinced that if she does not do the activity today, she'll do it tomorrow, as in (4). This agent postpones the activity day-after-day, ending up never doing it. O'Donoghue and Rabin (2001) show that, in the presence of naiveté, even a small degree of self-control problems can generate (infinite) procrastination. O'Donoghue and Rabin (1999b) presents calibrations for the case of retirement savings in a deterministic set-up. DellaVigna and Malmendier (2006) allow for stochastic transaction costs and show that naive  $(\beta, \delta)$  agents accumulate substantial delays in a costly activity (in their case, cancelling a health club membership). O'Donoghue and Rabin (2001) also show that, unlike naive agents, sophisticated  $(\beta, \delta)$  agents do not exhibit large default effects for reasonable parameter values. While these agents would like to postpone activities with immediate costs, they realize that doing an activity now is better than postponing it for a long time.

If procrastination of a financial transaction is indeed responsible for the default effects in Madrian and Shea (2001) and in Choi et al. (2004), we should expect that, if individuals were forced to make an active choice at enrollment, they would display their true preferences for savings. In this case, they bear the transaction cost whether they invest or not, and hence investing does not have an immediate cost, i.e.,  $b_1 = 0$ . In this situation, the short-run self does not desire to postpone the choice. Choi et al. (2005) analyze a company that required its employees to choose the retirement savings at enrollment. Under this Active Decision plan, 80% of workers enrolled in a 401(k) within one year of joining the company. Later, this company switched to a no-investment default, and the one-year enrollment rate declined to 50%. Requiring workers to choose, therefore, produces an enrollment rate that is only slightly lower than under the automatic enrollment in Madrian and Shea (2001).<sup>13</sup>

**Welfare.** These studies have welfare and policy implications. They suggest that savings rates for retirement in the US may be low due to a combination of procrastination and defaults set to no savings. The  $(\beta, \delta)$  model implies that the individuals are likely to be happier with defaults set to higher savings rates. A change in policy with defaults set to automatic enrollment is an example of cautious paternalism (Camerer et al., 2003), in that it would help substantially individuals with self-control problems and inflict little or no harm on individuals without self-control problems. These individuals can switch to a different savings rate for a

---

<sup>12</sup>Inattention and limited memory about 401(k) investment are other possible explanations.

<sup>13</sup>The effect of the Active Decision may also be due to a deadline effect for naive  $(\beta, \delta)$  employees, who know that the next occasion to enroll will not be until several months later.

low transaction cost. In Section 5, we present the results of a plan with automatic enrollment and other features designed to increase savings (Benartzi and Thaler, 2004). An alternative design could be based on the requirement to make an active choice, as in Choi et al. (2005). Social Security is a commitment device to save, albeit one that consumers cannot opt out of, and that thus can hurt consumers with no self-control problems.

**Summary.** A model of self-control problems with partial naiveté can rationalize a number of findings that are puzzling to the standard exponential model: (i) excessive preference for membership contracts in health clubs; (ii) positive effect of deadlines on homework grades and preference for deadlines; (iii) near-neglect of post-teaser interest rates in credit-card take-up; (iv) liquid debt and illiquid saving in life-cycle accumulation; (v) demand for illiquid savings as commitment devices; (vi) default effects in retirement savings and in other settings.

The partially-naive  $(\beta, \delta)$  model, therefore, does a good job of explaining qualitative patterns across a variety of settings involving self-control. A frontier of this research agenda is to establish whether one model can fit these different facts not just qualitatively, but also quantitatively. A few papers have estimated values for the time preference parameters. Laibson, Repetto, and Tobacman (2006) estimate annual time preference parameters ( $\beta = .70, \delta = .96$ ) on life-cycle accumulation data. Pášerman (forthcoming), building on DellaVigna and Pášerman (2005), uses job search data to estimate<sup>14</sup> ( $\beta = .40, \delta = .99$ ) for low-wage workers and ( $\beta = .89, \delta = .99$ ) for high-wage workers. Both papers assume sophistication.

## 2.2 Reference Dependence

The simplest version of the standard model as in (1) assumes that individuals maximize a global utility function over lifetime consumption  $U(x|s)$ .

**Laboratory Experiments.** A set of experiments on attitude toward risk call into question the assumption of a global utility function. An example (using hypothetical questions) from Kahneman and Tversky (1979) illustrates the point. A group of 70 subjects is asked to consider the situation: “In addition to whatever you own, you have been given 1,000. You are now asked to choose between A: (1,000, .50), and B: (500).” A different group of 68 subjects is asked to consider: “In addition to whatever you own, you have been given 2,000. You are now asked to choose between C: (-1,000, .50), and D: (-500).” The allocations A and C are identical, and so are B and D. However, in the first group only 16 percent of the subjects choose A, in contrast with 69 percent of subjects choosing C in the second group. Clearly, framing matters.

Choices in lotteries with real payoffs display similar violation of the standard theory. In Fehr and Goette (2007), 27 out of 42 subjects prefer 0 Swiss Franks for sure to the lottery  $(-5, p = .5; 8, p = .5)$ . Under the standard model, this implies an unreasonably high level of

---

<sup>14</sup>In Pášerman (2006), the model is estimated at the weekly level, so the  $\beta$  parameter refers to the one-week discounting. The  $\delta$  parameter is the annualized equivalent.

risk aversion (Rabin, 2000). A subject that made this choice for all wealth levels would also reject the lottery  $(-31, p = .5; \infty, p = .5)$ , which offers an infinite payout with probability .5.

**Model.** Kahneman and Tversky (1979), in the second most cited article in economics since 1970 (Kim, Morse, and Zingales, 2006), propose a reference-dependent model of utility that, unlike the standard model, can fit most of the experimental evidence on lottery choice. According to prospect theory, subjects evaluate a lottery  $(y, p; z, 1 - p)$  as follows:  $\pi(p)v(y - r) + \pi(1 - p)v(z - r)$ . Prospect theory is characterized by: (i) *Reference Dependence*. The value function  $v$  is defined over differences from a reference point  $r$ , instead of over the overall wealth; (ii) *Loss Aversion*. The value function  $v(x)$  has a kink at the reference point and is steeper for losses ( $x < 0$ ) than for gains ( $x > 0$ ); (iii) *Diminishing Sensitivity*. The value function  $v$  is concave over gains and convex over losses; (iv) *Probability weighting*. The decision-maker transforms the probabilities with a probability-weighting function  $\pi(p)$  that overweights small probabilities and underweights large probabilities.

The four features of prospect theory are designed to capture the evidence on risk-taking, including risk-aversion over gains, risk-seeking over losses, and contemporaneous preference for insurance and gambling. It can also capture framing effects as in the example above. Lottery A is evaluated as  $\pi(.5)v(1,000)$  and hence, given the concavity of  $v(x)$  for positive  $x$  and given  $\pi(.5) \approx .5$ , is inferior to lottery B, valued  $v(500)$ . Conversely, lottery C is evaluated as  $\pi(.5)v(-1,000)$  and, given the convexity of  $v(x)$  for negative  $x$ , is preferred to lottery D.

The large majority of the follow-up literature, however, adopts a simplified version of prospect theory incorporating only features (i) and (ii). The subjects maximize  $\sum_i p_i v(x_i|r)$ , where  $v(x|r)$  is defined as

$$v(x|r) = \begin{cases} x - r & \text{if } x \geq r; \\ \lambda(x - r) & \text{if } x < r, \end{cases} \quad (5)$$

where  $\lambda > 1$  denotes the loss aversion parameter. Prospect theory, even in the simplified version of expression (5), can explain the aversion to small risk exhibited experimentally. A prospect-theoretic subject evaluates the lottery  $(-5,.5; 8,.5)$  as  $.5\lambda * (-5) + .5 * 8 = 4 - 2.5\lambda$ . This subject prefers the status-quo for  $\lambda > 8/5$ . (The experimental evidence from Tversky and Kahneman (1992) suggests  $\lambda \approx 2.25$ ). I present a number of applications to economic phenomena, including ones not involving risk (such as the endowment effect and labor supply).

**Endowment Effect.** A finding consistent with prospect theory and inconsistent with the standard model is the so-called endowment effect, an asymmetry in willingness to pay (WTP) and willingness to accept (WTA). In the laboratory, Kahneman, Knetsch, and Thaler (1990) randomly allocate mugs to one group of experimental subjects. They then use an incentive-compatible procedure to elicit the WTA for subjects that received the mug, and the WTP for subjects that were not allocated the mug. According to the standard theory, the two valuations should on average be the same. The median WTA of \$5.75, however, is twice as large as the median WTP of \$2.25. Since theoretically wealth effects could explain this discrepancy, in a

different experiment Kahneman, Knetsch and Thaler introduce choosers, alongside buyers and sellers. Choosers, who are not endowed with a mug, choose between a mug and a sum of money; the experimenters elicit the price that induces indifference. Their choice is formally identical to the choice of the sellers (except for the fact that the choosers are not endowed with the mug); hence, according to the standard theory, the sum of money that makes them indifferent should correspond to the WTA of sellers. Instead, in this experiment the median WTA for sellers is \$7.12, while the price for choosers is \$3.12 (and the WTP for buyers is \$2.87). The asymmetry between WTA and WTP has implications such as low volume of trades in markets and inconsistencies in the elicitation of contingent valuations in environmental decisions.

The endowment effect is predicted by a reference-dependent utility function with loss-aversion  $\lambda > 1$ , as long as the subjects do not exhibit loss aversion with respect to money. Assume that the utility of the subjects is  $u(1)$  if they received a mug, and  $u(0)$  otherwise, with  $u(1) > u(0)$ . Consider subjects with a piece-wise linear utility function (5), where the reference point  $r$  depends on whether the subjects were assigned a mug. Subjects with the mug have reference point  $r = 1$  and assign utility  $u(1) - u(1) = 0$  to keeping the mug and utility  $\lambda[u(0) - u(1)] + p_{WTA}$  to selling the mug for the sum  $p_{WTA}$ . Subjects without the mug have reference point  $r = 0$  and assign value  $u(1) - u(0) - p_{WTP}$  to getting the mug at price  $p_{WTP}$  and utility  $u(0) - u(0) = 0$  to keeping the status quo. The prices that make both groups of subjects indifferent between having and not having the mug are

$$p_{WTA} = \lambda [u(1) - u(0)] \text{ and } p_{WTP} = u(1) - u(0),$$

hence  $p_{WTA} = \lambda p_{WTP}$ . A loss-aversion parameter  $\lambda = 5.75/2.25$  fits the evidence in Kahneman et al. (1990). Notice that choosers choose a mug if  $u(1) - u(0) \geq p_C$ , and hence  $p_C = p_{WTP}$  with referent-dependent preferences, approximately as observed.

Plott and Zeiler (2004) criticize this set of experiments on the ground that the endowment effect may be due to lack of experience of subjects. They elicit the WTP and WTA for a mug after extensive training and practice rounds, in 2 of 3 sessions including 14 rounds of trading of lotteries (for which no endowment effect is expected). In contrast to Kahneman et al. (1990), they find no evidence of the endowment effect for mugs, with a median WTA of \$5.00 and a median WTP of \$6.00. This result suggests that the endowment effect does not appear in economic settings where subjects are highly experienced and where they get repeated feedback. Of course, several important economic decisions, such as buying or selling a house, involve only limited experience and feedback.

List (2003 and 2004) provide field evidence consistent with this hypothesis for participants of a sports card fair. By selection, these subjects have at least some experience with sport cards, but some subjects are substantially more experienced than others. List (2003) randomly assigns sports memorabilia A or B as compensation for filling out a questionnaire. After the questionnaire is filled out, the participants are asked whether they would like to switch their

assigned memorabilia for the other one. Since the objects are chosen to be of comparable value, the standard model predicts trade about 50 percent of the time. Instead, subjects with low trading experience switch only 6.8 percent of the time, displaying a strong form of the endowment effect. Unlike inexperienced subjects, instead, subjects with high trading experience switch 46.7 percent of the time, displaying no endowment effect. The difference between the two groups is not due to the fact that inexperienced traders are approximately indifferent between the two memorabilia, and hence willing to stick to the status quo. In another treatment eliciting WTA and WTP, the WTA is substantially larger than the WTP for inexperienced subjects (18.53 versus 3.32), but not for experienced subjects (8.15 versus 6.27). Next, List (2003) attempts to test whether the difference between the two groups is due to self-selection of subjects without the endowment effect among the frequent traders, or is a causal effect of trading experience on the endowment effect. In a follow-up study performed months later, the endowment effect decreases in the trading experience accumulated in the intervening months, supporting the latter interpretation. Finally, and most surprisingly, List (2004) shows that the more experienced card traders also display substantially less endowment effect with respect to other goods, such as chocolates and mugs.

Overall, the evidence suggests that the endowment effect is a feature of trading behavior that market experience tempers.<sup>15</sup> This evidence leaves open (at least) two interpretations. One interpretation is that experience with the market leads individuals to become aware of their loss aversion, and counteract it: experience mitigates loss aversion. Another interpretation is that experience does not affect loss aversion, but it impacts the reference-point formation. Assume that experienced traders expect to trade the object that they are assigned with probability .5, independent of which group they are assigned to. As in Kőszegi and Rabin (2006), we model subjects as having a stochastic reference point,  $r = 1$  with probability .5 and  $r = 0$  otherwise. For individuals assigned the good, the (expected) value of keeping the good is  $.5 * [u(1) - u(0)] + .5 * [u(1) - u(1)] = .5 * [u(1) - u(0)]$ ; the (expected) value of selling the good  $.5 * [u(0) - u(0) + pWTA] + .5 * [\lambda(u(0) - u(1)) + pWTA] = .5 * [\lambda(u(0) - u(1))] + pWTA$ . This implies  $pWTA = .5(1 + \lambda)[u(1) - u(0)]$ . It is easy to show with similar calculation that

$$pWTP = .5(1 + \lambda)[u(1) - u(0)] = pWTA.$$

If experienced subjects have rational expectations about their reference point (Kőszegi and Rabin, 2006), they exhibit no endowment effect, even if they are loss-averse. The follow-up literature should consider carefully the determination of the reference point.

**Labor Supply.** As a second application, we consider the response of labor supply to wage fluctuations. This response, in general, reflects a complex combination of income and substitution effects (Card, 1994). Here, we consider a simple case in which income effects can, to a first approximation, be neglected. I consider jobs in which workers decide the labor supply

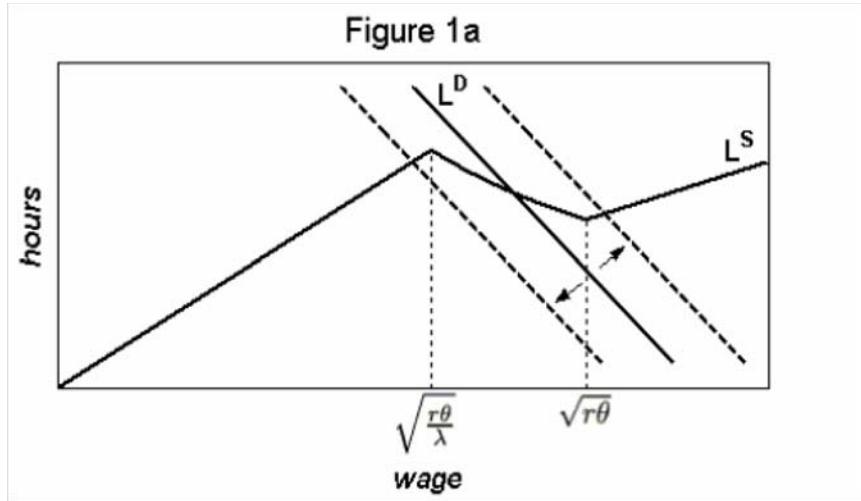
---

<sup>15</sup>In the Conclusion, I discuss further the role of experience.

daily, and in which the realization of the daily wage is idiosyncratic. Taxi drivers, for example, decide every day whether to drive for the whole shift or end earlier; the effective wage varies from day-to-day as the result of demand shifters such as weather and conventions. For these occupations, the income effect from (uncorrelated) changes in the daily wage is negligible, and we can neglect it by assuming a quasi-linear model. Assume that, each day, workers maximize the utility function  $U(Y) - \theta h^2/2$ , where the daily earning  $Y$  equals  $hw$ ,  $h$  is the number of hours worked,  $w$  is the daily wage, and  $\theta h^2/2$  is the (convex) cost of effort.

Following the simplified prospect theory formulation in (5), we assume that the utility function  $U(Y)$  equals  $(Y - r)$  for  $Y \geq r$ , and  $\lambda(Y - r)$  otherwise, where  $r$  is a target daily earning. Reference-dependent workers ( $\lambda > 1$ ) are loss-averse with respect to missing the daily target earning. For  $\lambda = 1$ , this model reduces to the standard model with risk-neutral workers.

In the standard model ( $\lambda = 1$ ), workers maximize  $wh - \theta h^2/2$ , yielding an upward-sloping labor supply curve  $h^* = w/\theta$ . As the wage increases, so do the hours supplied, in accordance to the substitution effect between leisure and consumption. A reference-dependent worker ( $\lambda > 1$ ), instead, exhibits a non-monotonic labor supply function (Figure 1a). For a low wage ( $w < \sqrt{r\theta/\lambda}$ ), the worker has not yet achieved the target earnings, and an increase in wage leads to an increase in hours worked ( $h^* = \lambda w/\theta$ ), as in the standard model. For a high wage ( $w > \sqrt{r\theta}$ ), the worker earns more than the target, and the labor supply is similarly upward-sloping, albeit flatter ( $h^* = w/\theta$ ). For intermediate levels of the wage ( $\sqrt{r\theta/\lambda} < w < \sqrt{r\theta}$ ), instead, the worker is content to earn exactly the daily target  $r$ . Any additional dollar earned makes it easier to reach the target and leads to reductions in the number of hours worked ( $h^* = r/w$ ); this generates a locally downward-sloping labor supply function.



Camerer, Babcock, Loewenstein, and Thaler (1997) use three data sets of hours worked and daily earnings for New York cab drivers to test whether the labor supply function is upward-sloping, as the standard theory above implies, or downward-sloping. Denote by  $Y_{i,t}$  and  $h_{i,t}$

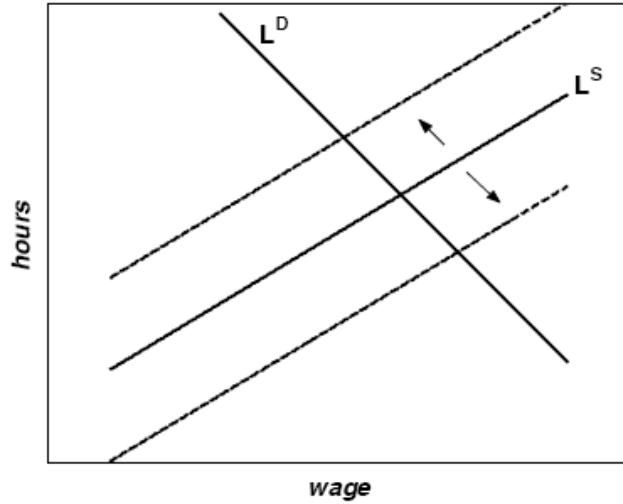
the daily earnings and the hours worked on day  $t$  by driver  $i$ . Camerer et al. (1997) estimate the OLS labor-supply equation

$$\log(h_{i,t}) = \alpha + \beta \log(Y_{i,t}/h_{i,t}) + \Gamma X_{i,t} + \varepsilon_{i,t}. \quad (6)$$

Increases in the daily wage, computed as  $Y_{i,t}/h_{i,t}$ , lead to decreases in the number of hours worked  $h_{i,t}$  with elasticities  $\hat{\beta} = -.186$  (s.e. .129),  $-.618$  (s.e. .051) and  $-.355$  (s.e. .051). The authors conclude that the data reject the standard model which predicts a positive elasticity, and support a reference-dependent model with daily earnings as the reference point. As Figure 1a shows, though, the labor supply function is not necessarily downward-sloping for target earners, and it is almost certainly not log-linear, unlike in specification (6). Nevertheless, the finding of a negative elasticity is consistent with reference-dependent preferences for shifts in labor demand corresponding to a wage in the interval  $\sqrt{\theta r/\lambda} < w < \sqrt{\theta r}$ .

Specification (6) is open to two main criticisms. First, a negative elasticity  $\hat{\beta}$  is expected if the daily fluctuations in wages for cab drivers are due to shifters of labor supply (like rain that make driving less pleasant), rather than shifters of labor demand. As Figure 1b illustrates, if labor supply shifts across days, the resulting equilibrium points plot out a downward-sloping curve even if the labor supply function is upward-sloping. Camerer et al. (1997) use interviews of cab drivers to argue that the factors affecting the wage are unlikely to change the marginal cost of driving; however, in the absence of an instrument for labor supply, this objection is a concern. Second, specification (6) suffers from division bias, which biases downward the estimate of  $\beta$ . Since the daily wage is computed as the ratio of daily earnings and hours worked, and since hours worked is the left-hand-side variable in (6), any measurement error in  $h_{i,t}$  induces a mechanical downward bias in  $\hat{\beta}$ . Camerer et al. (1997) address this objection by instrumenting the daily wage of worker  $i$  by the summary statistics of the daily wage of the other workers on the same shift. The estimates of  $\beta$  are still negative, though noisier.

Figure 1b



Farber (2005) uses a different data set of 584 trip sheets for 21 New York cab drivers and estimates a hazard model that does not suffer from division bias. For any trip  $t$  within a day, Farber (2005) estimates the probability of stopping as a function of the number of hours worked  $h_{i,t}$  and the daily cumulative earnings to that point,  $Y_{i,t}$ :

$$Stop_{i,t} = \Phi(\alpha + \beta_Y Y_{i,t} + \beta_h h_{i,t} + \Gamma X_{i,t}),$$

where  $\Phi$  is the c.d.f. of a standardized normal distribution. The standard theory predicts that  $\beta_Y$  should be zero (since earnings are not highly correlated within a day), while reference dependence predicts that  $\beta_Y$  should be positive. Farber (2005) finds that  $\beta_Y$  is positive ( $\hat{\beta}_Y = .015$ ), but not significantly so. While the author cannot reject the standard model, the point estimates are not negligible: a ten percent increase in  $Y_{i,t}$  (about \$15) is predicted to increase the probability of stopping by  $15 * .015 = .225$  percentage points, a 1.6 percent increase relative to the average of 14 percentage points. This corresponds to an elasticity between earnings and stopping of .16. These findings do not contradict prospect theory, since Farber (2005) does not test the hypothesis that cab drivers have reference-dependent preferences (Failing to reject the null is different from rejecting the alternative hypothesis of prospect theory, especially in light of the positive point estimates). In a more recent paper, Farber (2006) addresses this issue and tests, using the same data set, a simple model of labor supply which explicitly allows for reference-dependent preferences with a stochastic reference point. The findings provide weak evidence of reference dependence: the estimated model implies a loss-aversion coefficient  $\lambda$  significantly larger than zero. At the same time, however, the estimated variation across days in the reference daily earning is large enough that reference dependence loses predictive power.

Given the lack of an instrument for daily wage fluctuations, the evidence on the labor supply of taxi drivers is unlikely to settle the debate on reference dependence and labor supply. Fehr

and Goette (2007) provide new evidence using a field experiment on the labor supply of bike messengers. Like taxi drivers, bike messengers choose how long to work within a shift. Fehr and Goette (2007) randomly assign 44 messengers into two groups. Each group receives a 25 percent higher commission for the deliveries for just one month in two different months. This design solves both problems discussed above, since the increase in wage is exogenous, and the wage and the actual deliveries are exactly measured.

Fehr and Goette show that bike messengers in the treatment group respond in two ways to the exogenous (and anticipated) temporary increase in wage: (i) they work 30 percent more shifts; (ii) within each shift, they do 6 percent fewer deliveries. The first finding is consistent with both the standard model and the reference-dependent model. (When deciding on which day to work, reference-dependent workers will sign up for shifts on days in which it is easier to reach the daily target.) The second finding is consistent with target earning, and not with the standard model, which predicts an increase in the number of hours worked within each shift. However, this second finding, while statistically significant, is quantitatively small, suggesting the need for further evidence. In addition, this finding is consistent with an extension of the standard model in which workers in the treatment group get more tired, and hence do fewer deliveries, because they work more shifts.

With a clever design twist, Fehr and Goette (2007) provide additional evidence in support of reference-dependence using laboratory tests of risk-taking. The bike messengers that display loss aversion in the lab—i.e., they reject a  $(-5,.5;8,.5)$  lottery—exhibit a more negative response (though not significantly so) in their deliveries to the wage increase. The correlation between the laboratory and the field evidence of loss-aversion lends more credence to the reference-dependence interpretation. Still, the debate on reference dependence and labor supply is open.

**Finance.** Two of the most important applications of reference-dependent preferences are to the field of finance.<sup>16</sup> The first application is to the equity premium puzzle: equity returns outperformed bond returns by on average 3.9 percentage points during the period 1871-1993 (Campbell and Cochrane, 1999), a premium too large to be reconciled with the standard model, except for extremely high risk aversion (Mehra and Prescott, 1985). Benartzi and Thaler (1995) use a calibration<sup>17</sup> to show that this is the premium that loss-averse investors would require to invest in stocks, provided that they evaluate their portfolio performance annually. At horizons as short as a year, the likelihood that stocks underperform relative to bonds requires a substantial compensation in terms of returns, given loss aversion. In a paper that carefully formalizes the idea of Benartzi and Thaler (1995), Barberis, Huang, and Santos (2001) show that reference-dependent preferences can match the observed equity premium. This paper uses the simplified prospect-theory model with piece-wise linear function as in (5), relying on reference dependence and loss aversion for the predictions.

---

<sup>16</sup>Barberis and Thaler (2003) present a more comprehensive survey of these applications.

<sup>17</sup>The calibration uses the loss-aversion parameter estimated from the experiments.

The second application is to the so-called disposition effect, which denotes the tendency to sell ‘winners’ and hold on to ‘losers’<sup>18</sup>. Odean (1998) documents this phenomenon using individual trading data from a discount brokerage house during the period 1987-1993. Defining gains and losses relative to the purchase price of a share, Odean computes the share of realized gains  $PGR = (\text{Realized Gains}) / (\text{Realized Gains} + \text{Paper Gains})$  to equal .148. The share of realized losses  $PLR = (\text{Realized Losses}) / (\text{Realized Losses} + \text{Paper Losses})$  equals .098. Odean (1998) shows that the large difference between the propensity to realize gains ( $PGR$ ) and the propensity to realize losses ( $PLR$ ) is not due to portfolio rebalancing, or to ex-post higher returns for ‘losers’ (if anything, ‘winners’ outperform ‘losers’), or to transaction costs. The disposition effect is puzzling for the standard theory, since capital gain taxation would lead to expect that investors liquidate ‘losers’ sooner. This puzzle is a robust finding, replicated more recently by Ivkovich, Poterba, and Weisbenner (2005), who show that the effect is present in both taxable and tax-deferred accounts (though larger in tax-deferred accounts).

Prospect theory is viewed as a possible explanation for this phenomenon. The concavity over gains induces less risk-taking for ‘winner’ stocks, and hence more sales of ‘winners’. The convexity over losses induces more risk-taking for ‘loser’ stocks, and hence more purchases of ‘losers’. Barberis and Xiong (2006), however, point out that this argument does not take into account the impact of the kink at the reference point. When they simulate a calibrated model of reference-dependent preferences, Barberis and Xiong (2006) find that they obtain the disposition effect only for certain ranges of the parameters, and they obtain the opposite pattern for other ranges. More research is necessary to say whether reference-dependent preferences are a plausible explanation for the disposition effect.

**Insurance.** A puzzling feature of insurance behavior is the pervasiveness of small-scale insurance. Insurance policies on, for example, the telephone wiring are commonplace despite the fact that, in case of an accident, the losses amount to at most \$50 (Cicchetti and Dubin, 1994). This is a puzzle for expected utility, which implies local risk-neutrality and hence no demand for small-scale insurance (except in the unrealistic case of fair pricing). Sydnor (2006) provides evidence of excess small-scale insurance for the \$36 billion home insurance industry. Since mortgage companies require home insurance, the consumer choice is limited to the level of deductible in a standard menu: \$250 vs. \$500 vs. \$1000. Using a random sample of 50,000 members of a major insurance company in one year, Sydnor documents that 83% of customers and 61% of new customers choose deductibles lower than \$1000. The modal homeowner chooses a \$500 deductible, thereby paying on average \$100 of additional premium relative to a \$1000 deductible. However, the claim rate is under 5%, which implies that the value of a low deductible is about \$25 in expectation. The standard homeowner, therefore, is sacrificing \$100-\$25=\$75 in expectations to insure against, at worst, a \$500-\$100=\$400 risk.

---

<sup>18</sup>In the housing market, Genesove and Mayer (2001) document that house-owners are less willing to sell houses when housing prices are below the initial buying price, a phenomenon related to the disposition effect.

This indicates a strong preference for insuring against small risks that is a puzzle for the standard theory, unless one assumes three-digit coefficients of relative risk aversion. This deviation from the standard model involves substantial stakes. If, instead of choosing a low deductible, homeowners selected the \$1000 deductible from age 30 to age 65 and invested the money in a money market fund, their wealth at retirement would be \$6,000 higher. Sydnor (2006) shows that a calibrated version of prospect-theory can match the findings by the over-weighting of the small probability of an accident and the loss aversion with respect to future losses<sup>19</sup>. The two components of prospect theory each account for about half of the observed discrepancy between the predicted and the observed willingness to pay for low deductibles. Social pressure by the salesmen (who are paid a percentage of the premium as commission) may also contribute to the prevalence of low-deductible contracts.

**Employment.** Mas (2006) estimates the impact of reference points for the New Jersey police. In the 9 percent of cases in which the police and the municipality do not reach an agreement, the contract is determined by final offer arbitration. The police and the municipality submit their offers to the arbitrator, who has to choose one of the two offers. In theory (Mas, 2006), if the disputing parties are equally risk-averse, the winner in arbitration is determined by a coin toss.<sup>20</sup> Mas (2006) exploits this prediction of quasi-random assignment to present evidence on how police pay affects performance for 383 arbitration cases from 1978 to 1995. Mas documents that, in the cases in which the offer of the employer is chosen, the share of crimes solved by the police (the clearance rate) decreases by 12 percent compared to the cases in which the police offer is chosen. The author also documents a smaller increase in crime. Lower than expected pay therefore induces the police to devote less effort to fighting crime.

Mas (2006) provides additional evidence that reference points mediate this effect of pay on performance. Mas uses the predicted award based on a set of observables as a proxy for the reference point, and computes how the clearance rate responds to differences between the award and the predicted award. The response is significantly higher for cases in which the police loses—and hence is on the loss side—than for cases in which the police wins—and hence is on the gain side. This finding is consistent with reference-dependent preferences with loss aversion. Assume for example that the utility function of the police is  $[V + v(w|r)]e - \theta e^2/2$ , where  $v(w|r)$  is as in (5). This assumes a complementarity between police pay  $w$  and effort  $e$  in the utility function, capturing a form of reference-dependent reciprocity. The first-order condition, then, implies  $e^*(w) = [V + v(w|r)]/\theta$ . Given loss aversion in  $v(w|r)$ , this predicts indeed a stronger response for  $w$  below  $r$  than for  $w$  above  $r$ .

---

<sup>19</sup>Loss aversion could in principle go the other way, since individuals that are loss-averse to paying a high premium may as well prefer the high deductible. Experimental evidence, however, suggests that consumers will adjust their reference point on the premium side, since they are expecting to pay the premium for sure, but cannot adjust the reference point on the future uncertain loss.

<sup>20</sup>In reality, the arbitrator rules for the municipality in 34.4 percent of cases, suggesting that the unions are more risk-averse than the employers.

**Summary.** Reference-dependent preferences help explain: (i) excessive aversion to small risks in the laboratory; (ii) endowment effect for inexperienced traders; (iii) (some evidence of) target earnings in labor supply decisions; (iv) equity premium puzzle in asset returns; (v) (possibly) the tendency to sell ‘winners’ rather than ‘losers’ in financial markets; (vi) the tendency to insure against small risks; (vii) effort in the employment relationship. I have discussed cases in which the evidence is more controversial (labor supply and endowment effect) and cases in which it is unclear whether reference-dependence is an explanation for the phenomenon (disposition effect). I have also discussed how the original model in Kahneman and Tversky (1979) (and the calibrated version in Tversky and Kahneman, 1992) is rarely applied in its entirety, often appealing just to reference dependence and loss-aversion.

A key issue in this literature is the determination of the reference point  $r$ . Often, different assumptions about the reference point are plausible, which makes the application of the theory difficult. Kőszegi and Rabin (2006) have proposed a solution. They suggest that the reference point be modeled as the (stochastic) rational-expectations equilibrium of the transaction. In any given situation, this model makes a prediction for the reference point, without the need for additional parameters (though there can often be multiple equilibria, and hence multiple possible reference points). This theory also provides a plausible explanation for some of the puzzles in this literature. For example, as we discussed above, it predicts the absence of endowment effect among experienced traders (List, 2003 and Plott and Zeiler, 2004), even if these traders are loss-averse. Experienced traders expect to trade any item they receive, and hence their reference point is unaffected by the initial allocation of objects.

### 2.3 Social Preferences

The standard model, in its starker form as in (1), assumes purely self-interested consumers, that is, utility  $U(x_i|s)$  depends only on own payoff  $x_i$ .

**Laboratory Experiments.** An extensive number of laboratory experiments calls into question the assumption of pure self-interest. I present here the results of two classical experiments, which we relate to the field evidence below. (i) *Dictator game*. In this experiment (Forsythe et al., 1994) a subject (the dictator) has an endowment of \$10 and chooses how much to transfer of the \$10 to an anonymous partner. While the standard theory of self-interested consumers predicts that the dictator would keep the whole endowment, Forsythe et al. (1994) find that sixty percent of subjects transfers a positive amount. (ii) *Gift Exchange game*. This experiment (Fehr, Kirchsteiger, and Riedl, 1993) is designed to mirror a labor market. It tests efficiency wages models according to which the workers reciprocate a generous wage by working harder (Akerlof, 1982). The first subject (the firm) decides a wage  $w \in \{0, 5, 10, \dots\}$ . After observing  $w$ , the second subject (the worker) responds by choosing an effort level  $e \in [.1, 1]$ . The firm payoff is  $(126 - w)e$  and the worker payoff is  $w - 26 - c(e)$ , with  $c(e)$  increasing

and slightly convex. The standard theory predicts that the worker, no matter what the firm chooses, exerts the minimal effort and that, in response, the firm offers the lowest wage that satisfies the participation constraint for the workers ( $w = 30$ ). Fehr et al. (1993) instead find that the workers respond to a higher wage  $w$  by providing a higher effort  $e$ . The firms, anticipating this, offer a wage above the market-clearing one (the average  $w$  is 72). These results have been widely replicated and have given rise to a rich literature on social preferences in the laboratory, summarized in Charness and Rabin (2002) and Fehr and Gächter (2000).

**Model.** Several models have been proposed to rationalize the behavior in these experiments; we introduce a simplified version of the social preference model in Charness and Rabin (2002), which builds on the formulation of Fehr and Schmidt (1999).<sup>21</sup> In a two-player experiment, the utility of subject 1 is defined as a function of the own payoff ( $x_1$ ) and other-player's payoff ( $x_2$ ):

$$U_1(x_1, x_2) \equiv \begin{cases} \rho x_2 + (1 - \rho)x_1 & \text{when } x_1 \geq x_2; \\ \sigma x_2 + (1 - \sigma)x_1 & \text{when } x_1 < x_2. \end{cases} \quad (7)$$

The standard model is a special case for  $\rho = \sigma = 0$ . The case of baseline altruism is  $\rho > 0$  and  $\sigma > 0$ , that is, player 1 cares positively about player 2, whether 1 is ahead or not. In addition, Charness-Rabin (2002) assume  $\rho > \sigma$ , that is, player 1 cares more about player 2 when 1 is ahead. Fehr and Schmidt (1999) propose an equivalent representation of preferences<sup>22</sup> and assume  $0 < \rho < 1$ , like Charness-Rabin (2002), but also  $\sigma < -\rho < 0$ . When player 1 is behind, therefore, she prefers to lower the payoff of player 2 (since she is inequality-averse). These two models can explain giving in a Dictator Game with a \$10 endowment. The utility of giving \$5 is higher than the utility of giving \$0 if  $5 \geq \max((1 - \rho)10, \sigma10)$ , that is, if  $\rho \geq .5 \geq \sigma$  (altruism is high enough, but not so high that a player would transfer all the surplus to the opponent.) Fehr and Schmidt (1999) show that model (7) can also rationalize the average behavior in the Gift Exchange game for high enough  $\rho$ : altruistic workers provide effort to lower the inequality with the firm; the firm, anticipating this, raises  $w$ .

**Charitable Giving.** The size of charitable giving is suggestive of social preferences in the field. In the US, in 2002, 240.9 billion dollars were donated to charities, representing an approximate 2 percent share of GDP (Andreoni, 2006). Donations of time in the form of volunteer work were also substantial: 44 percent of respondents to a survey reported giving time to a charitable organization in the prior year, with volunteers averaging about 15 hours

---

<sup>21</sup>In these models, players care about the inequality of outcomes, but not about the intentions of the players (though the general model in Charness and Rabin (2002) allows for it). Another class of models (including Rabin, 1993 and Dufwenberg and Kirchsteiger, 2004), based on psychological games, instead assumes that subjects care about the intentions that lead to specific outcomes. A common concept is reciprocity—subjects are nice to subjects that are helpful to them, but not to subjects that take advantage of them. These models also explain the laboratory findings.

<sup>22</sup>Fehr-Schmidt preferences take the form:  $U_1(\pi_1, \pi_2) = \pi_1 - \alpha \min(\pi_2 - \pi_1, 0) - \beta \min(\pi_1 - \pi_2, 0)$ ; they are equivalent to the preferences in (7) for  $\beta = \rho$  and  $\alpha = -\sigma$ .

per month (Andreoni, 2006). Altogether, a substantial share of GDP reflects a concern for others, a finding qualitatively consistent with the experimental findings. However, while social preferences are a leading interpretation for giving, charitable donations may also be motivated by other factors, such as desire for status and social pressure by the fund-raisers.

Even if we take it for granted that giving is an expression of social preferences, it is difficult to use models such as (7) to explain quantitatively the patterns of giving in the field for three reasons. (i) These models are designed to capture the interaction of two players, or at most a small number of players. Charitable giving instead involves a large number of potential recipients, from local schools to NGOs in Africa. (ii) The utility representation (7) implicitly assumes that  $x_1$  and  $x_2$  include only the experimental payoffs from, say, the dictator game. In the field, it is difficult to determine to what extent  $x_1$  and  $x_2$  should include, for example, the disposable income. (iii) In one-to-one fund-raising situations, (hence side-stepping issue (i)), models such as (7) over-predict giving. Suppose, for example, that  $x_1 = \$1,000$  is the disposable income of person 1 and  $x_2 = \$0$  is the disposable income of person 2, for example, a homeless person. For  $\rho \geq .5 \geq \sigma$ , the model predicts that person 1 should transfer  $(\$1000 - \$0)/2 = \$500$ , a level of giving much higher than 2 percent of GDP. One has to make ad-hoc assumptions on  $x_1$  to reproduce the observed level of giving. For these reasons, while models of social preferences are very useful to understand behavior in the laboratory, they are less directly applicable to the field, compared to models of self-control and of reference-dependence. Andreoni (2006) overviews models that better predict patterns of giving, such as models of warm glow.

There are, however, field settings which resemble more closely the laboratory set-up. When a fund-raiser contacts a person directly, the situation resembles a dictator game, except for the lack of anonymity. Field experiments in fund-raising, starting from List and Lucking-Reilly (2002), estimate the effect on giving of variables such as the seed money (the funds raised early on), the match rate, and the identity of the solicitor. These experiments find, for example, that charitable giving is increasing in the seed money (List and Lucking-Reilly, 2002) presumably because of signaling of quality of the charity. These results, however, do not address some of the key questions on giving, such as why people give, and to whom they choose to give. These questions are likely to be the focus of future research.

**Workplace Relations.** Workplace relations between employees and employer can be upset at the time of contract renewal, and workers may respond by sabotaging production. Krueger and Mas (2004) examine the impact of a three-year period of labor unrest at a unionized Bridgestone-Firestone plant on the quality of the tires produced at the plant. The workers went on strike in July 1994 and were replaced by replacement workers. The union workers were gradually reintegrated in the plant in May 1995 after the union, running out of funds, accepted the demands of the company. An agreement was not reached until December 1996. Krueger and Mas (2004) finds that the tires produced in this plant in the 1994-1996 years were ten

times more likely to be defective. The increase in defects does not appear due to lower quality of the replacement workers. The number of defects is higher in the months preceding the strike (early 1994) and in the period in which the union workers and the replacement workers work side-by-side (and of 1995 and 1996). This indicates that negative reciprocity is response to what workers perceive as unfair treatment can have a large impact on worker productivity.

Bandiera, Barankay, and Rasul (2005) test for the impact of social preferences in the workplace among employees. They use personnel data from a fruit farm in the UK and measure changes in the productivity as a function of changes in the compensation scheme. In the first 8 weeks of the 2002 picking season, the fruit-pickers were compensated on a relative performance scheme in which the per-fruit piece rate is decreasing in the average productivity. In this system, workers that care about others have an incentive to keep the productivity low, given that effort is costly. In the next 8 weeks, the compensation scheme switched to a flat piece rate per fruit. The change was announced on the day of the switching. Bandiera et al. (2005) find that the, after the change to piece rate, the productivity of each worker increases by 51.5 percent; the estimate holds after controlling for worker fixed effects and is higher for workers with a larger network of friends. These results can be evidence for social preferences; they can, however, also be evidence of collusion in a repeated game, especially since in the field each worker can monitor the productivity of the other workers. To test for these explanations, the authors examine the effect of the change in compensation for growers of a different fruit where the height of the plant makes monitoring among workers difficult. For this other fruit, the authors find no impact on productivity of the switch to piece rate. This implies that the findings are due to collusion, rather than to social preferences.

**Gift Exchange in the Field.** The Bandiera et al. (2005) paper underscores the importance of controlling for repeated game effects in tests of social preferences. We now consider a set of field experiments that tests for Gift Exchange and carefully controls for these effects. Falk (forthcoming) examines the importance of gifts in fund-raising. The context is the mailing of 9,846 solicitation letters in Switzerland to raise money for schools in Bangladesh. One third of the recipients receives a postcard designed by the students of the school, another third receives four such postcards, and the remaining third receives no postcards. The three mailings are otherwise identical, except for the mention of the postcard as a gift in the two treatment conditions. The donations are increasing in the size of the gifts. Compared to the 12.2 percent frequency of donation in the control group, the frequency is 14.4 percent in the small gift and 20.6 percent in the large gift treatment. Conditional on a donation, the average amount donated is slightly smaller in the large-gift treatment, but this effect is small relative to the effect on the frequency of donors. The large treatment effects do not appear to affect the donations at next year's solicitation letter, when no gift is sent. A gift, therefore, appears to trigger substantial positive reciprocity, as in the laboratory version of the Gift Exchange.

Gneezy and List (2006) test the gift exchange with two field experiments in workplace

settings. In the first experiment, they hire 19 workers for a six-hour data entry task at a wage of \$12 per hour; in the second experiment, they hire 23 workers to do door-to-door fund-raising for one weekend at a wage of \$10 per hour. In both cases, they divide the workers into a control and a treatment group. The control group is paid as promised, while the treatment group is told after recruitment that the pay for the task was increased to \$20 per hour. The authors test whether the treatment group exerts more effort than the control group, as predicted by the gift exchange hypothesis, or the same effort, as predicted by the standard model. The findings are two-fold. At first, the treatment group exerts substantially more effort, consistent with gift exchange: treated workers log 20 percent more books in the first hour and raise 80 percent more money in the morning hours. The difference however is short-lived: the performances of control and treatment group are indistinguishable after two hours of data entry and after three hours of fund-raising. In these two applications, the increase in wage does not pay for itself (though it may for different experimental designs). These experiments suggest that the gift exchange may have an emotional component which dissipates over time.

Kube, Maréchal, and Puppe (2006) use a similar design for a six-hour library work in Germany, but they add a negative gift exchange treatment. This group of subjects, upon showing up, is notified that the pay is 10 Euro per hour, compared to the promised pay of ‘presumably’ 15 Euro per hour. (No one quits) This group logs 25 percent fewer books compared to the control group, a difference that, unlike in the Gneezy and List (2006) paper, does not decline over time. The group in the positive gift exchange treatment (paid 20 Euro) logs only 5 percent more books, an increase which also does not dissipate over time. The finding that negative reciprocity is stronger than positive reciprocity is consistent with laboratory findings.

Finally, List (2006) presents evidence that not everyone reciprocates a generous transfer. Attendees of a sports card fair participate in a field experiment involving buying a card from a dealer. One group is instructed to offer \$20 for a good-quality card, while another group is instructed to offer \$65 for a top-quality card. The quality of the card can be verified by an expert but is not apparent on inspection. Dealers that are ‘non-local’ (and hence are not concerned with reputation) offer cards of the same average quality to the two groups, displaying no gift-exchange behavior.<sup>23</sup> These dealers, however, display gift-exchange-type behavior in laboratory experiments designed to mirror the Fehr, Kirchsteiger, and Riedl (1993) experiment. These findings raise interesting questions on when gift-exchange behavior does and does not arise. One explanation of the findings is that bargaining in a market setting is not construed as a situation where norms of gift exchange apply. Hence, the dealers do not display such norms, but they do instead in an experiment in which they play the role of subjects. More broadly, this suggests that we need to understand the economic settings in which gift-exchange norms apply (such as charitable giving and, to some extent, employment relationships) and the ones

---

<sup>23</sup>Dealers that are ‘local’, that is, that attend the fair frequently, offer higher-quality card to the \$65 group, presumably because of reputation-building.

where they do not apply (such as market bargaining).

**Summary.** Social preferences help explain: (i) giving to charities; (ii) the response of striking workers to wage cuts; (iii) the response of giving to gifts in fund-raisers; (iv) the response of effort to unanticipated changes in pay, at least in the short-run. However, the research on social preferences displays more imbalance between laboratory and field, compared to the research on self-control and on reference dependence. The models of social preferences which match the laboratory findings are not easily applicable to the field, overpredicting, for example, the amount of giving. It will be important to see more papers linking the findings in the laboratory, which allows the most control on the design, to the evidence in the field; the recent literature on Gift Exchange is a good example. A separate issue is the difficulty of distinguishing in the field social preferences from repeated game strategies (as in Bandiera et al., 2005) and other alternative explanations. For example, social pressure (Section 4.3) can explain regularities in giving, such as the higher effectiveness of high-pressure fund-raising methods (such as phone calls) relative to low-pressure ones (such as mailings). Creative field experiments such as those in this Section can be designed to distinguish different explanations.

### 3 Non-standard Beliefs

The standard model in (1) assumes that consumers are on average correct about the distribution of the states  $p(s)$ . Experiments suggest instead that consumers have systematically incorrect beliefs in at least three ways: (i) *Overconfidence*. Consumers over-estimate their performance in tasks requiring ability, including the precision of their information; (ii) *Law of Small Numbers*. Consumers expect small samples to exhibit large-sample statistical properties; (iii) *Projection Bias*. Consumers project their current preferences onto future periods.

#### 3.1 Overconfidence

Surveys and laboratory experiments present evidence of overconfidence about ability. In Svensson (1981), 93 percent of subjects rated their driving skill as above the median, compared to the other subjects.<sup>24</sup> Most individuals underestimate the probability of negative events such as hospitalization (Weinstein, 1980) and the time needed to finish a project (Buehler, Griffin, and Ross 1994). In Camerer and Lovallo (1999), subjects play multiple rounds of an entry game in which only the top  $c$  out of  $n$  entrants make positive profits. In the luck treatment the top  $c$  subjects are determined by luck, while in the skill treatment the top  $c$  subjects are determined by ability in solving a puzzle. More subjects enter in the skill treatment than in the luck treatment, indicating that subjects overestimate their (relative) ability to solve puzzles.

---

<sup>24</sup>This finding admits alternative interpretations, such as that each individual may define driving ability in a self-serving way. These interpretations, however, are addressed in the follow-up literature.

The first example of overconfidence in the field is the naïveté about future self-control by consumers, as documented in Section 2.1. (Self-control is an ability.) In a second example, Malmendier and Tate (2005, forthcoming) provide evidence on overconfidence by CEOs about their ability to manage a company. They assume that CEOs are likely to overestimate their ability to pick successful projects and to run companies. As such, these top managers are likely to invest in too many projects, and to over-pay for mergers. To test these hypotheses, Malmendier and Tate identify a proxy for overconfidence, and examine the correlation of this proxy with corporate behavior. In particular, they identify as overconfident CEOs who hold on to their stock options until expiration, despite the fact that most CEOs are heavily under-diversified. They interpret the lack of exercise as overestimation of future performance of their company. In Malmendier and Tate (forthcoming) they find that these CEOs are 55 percent more likely to undertake a merger, and particularly so if they can finance the deal with internal funds. (Overconfident CEOs are averse to seeking external financing, since they deem it overpriced.) The correlation between option exercise and corporate behavior does not appear to be due to insider information, since the CEOs that delay exercising stock options do not gain money by doing so. Managerial overconfidence provides one explanation for the underperformance of companies undertaking mergers. Malmendier and Tate (2005) use the same proxies to show that overconfidence explains in part the excess sensitivity of corporate investment to the availability of cash flows, a long-standing puzzle in corporate finance.

A third example of overconfidence is the tendency to overestimate the precision of own information, which is also a skill. For example, Alpert and Raiffa (1982) ask people to provide answers with 98 percent confidence intervals for a number of questions. These intervals contain the correct answer only 60 percent of the time. Odean (1999) provides field evidence using data from a discount broker on all the trades of 10,000 individual investors for the years 1987-1993. If the investors overestimate the precision of their information about individual companies, they will trade too much. Indeed, the investors trade on average 1.3 times per year, with a commission cost for buying or for selling a security of over 2 percent per transaction. In addition to these substantial transaction costs, the individual investors pay a return cost to trading, since the stocks sold over-perform the purchases by about 3 percent over the next year. For individual investors, therefore, overconfidence has a substantial impact on returns.

Overconfidence about the precision of private information, coupled with self-attribution bias, can also explain other anomalies in financial markets, such as short-term positive correlation of returns (momentum) and long-term negative correlation (long-term reversal) (Daniel, Hirshleifer, and Subrahmanyam, 1998). Overconfidence induces individuals to trade excessively in response to private information; in the long-run, the public information prevails and the valuation returns to fundamentals, inducing a long-term reversal. The self-attribution bias is responsible for momentum: in the short-term, as investors receive additional private information, they interpret as more informative the information that confirm to their beliefs, and

hence become even more overconfident. In Sections 3.2 and 4.1 we discuss how the law of small numbers and limited attention provide alternative explanations for these phenomena.

**Summary.** Overconfidence helps explain: (i) patterns in credit card take-up and default effects, presented in Section 2.1 (overconfidence about self-control); (ii) value-destroying mergers and investment-cash-flow sensitivity (overconfidence about managerial ability); (iii) excess trading, momentum, and long-term reversal (overconfidence about precision of information). These applications are settings in which overconfidence is particularly likely according to the laboratory evidence: overconfidence is more common when feedback is noisy (i.e., for stock returns) and the decision-maker has an illusion of control (i.e., for managers).

### 3.2 Law of Small Numbers

Overconfidence is only one form of non-Bayesian beliefs detected in experiments. Tversky and Kahneman (1974) describe a number of deviations from rational updating, including neglect of base rate and overweighting of information that is available and representative. I focus on two phenomena—‘gambler’s fallacy’ and overinference—captured by Rabin (2002b)’s model of the law of small numbers. Rabin (2002b) assumes that subjects, observing a sequence of signals drawn from an i.i.d. process, believe (incorrectly) that the signals are drawn from an urn of size  $N < \infty$  without replacement. If the distribution of the signals is known, this induces a ‘gambler’s fallacy’ belief: after a draw of a signal, subjects expect the next draw to be a different signal (since the draw is without replacement). For example, suppose that the return to a mutual fund is drawn from an urn with 10 balls, 5 Up and 5 Down, with replacement. After two draws of Up, a rational investor expects the probability of another Up to be .5. However, a believer in the law of small number computes such probability as  $3/8 < .5$ , since two balls ‘Up’ have already been drawn. This is an example of the representativeness heuristics, in that the sequence ‘Up, Up, Down’ is judged as more representative than the sequence ‘Up, Up, Up’.

Terrell (1994) provides field evidence in New Jersey’s pick-three-numbers game. The lottery is a pari-mutuel betting system: the fewer individuals bet on a number, the higher is the expect payout. Terrell (1994) finds that the payout for a number that won one or two weeks before is 33 percent higher than for an average number. Belief in ‘gambler’s fallacy’ leads lottery players to bet less on numbers that won recently, at the cost of a lower expected payoff.

The model in Rabin (2002b) delivers a second prediction. In the case of uncertain distribution of signals, the subjects overinfer from a sequence of signals of one type that the next signal will be of the *same* type. While this appears to be the opposite of the ‘gambler’s fallacy’, it is the complementary phenomenon. Consider a mutual fund with a manager of uncertain ability. The return is drawn with replacement from an urn with 10 balls. With probability .5 the fund is well managed (7 balls Up and 3 Down) and with probability .5 the fund is poorly managed (3 Up and 7 Down). After observing the sequence ‘Up, Up,

Up', a rational investor computes the probability that the mutual fund is well-managed as  $P(Well|UUU) = .5P(UUU|Well)/[.5P(UUU|Well) + .5P(UUU|Poor)] = .7^3 / (.7^3 + .3^3) \approx .927$ . A Law-of-Small-Number investor also applies Bayes Rule but has the wrong model for  $P(UUU|Well)$  and  $P(UUU|Poor)$ . Hence, her forecasted probability for  $P(Well|UUU)$  equals  $(7/10 * 6/9 * 5/8) / [(7/10 * 6/9 * 5/8) + (3/10 * 2/9 * 1/8)] \approx .972$ . Hence, this investor overinfers about the ability of the mutual-fund manager after three good performances. Assume now that the Law-of-Small-Number investor believes that the urn is replenished after 3 periods. When forecasting the performance in the next period, the rational investor expects an Up performance with probability  $.927 * .7 + (1 - .927) * .3 \approx .671$ , while the Law-of-Small-Number investor expects Up with probability  $.972 * .7 + (1 - .972) * .3 \approx .689$ , which is higher.

Benartzi (2001) provides field evidence of overinference (also called extrapolation): the degree to which employees invest in employer stock depends strongly on the past performance of the stock. In companies in the bottom quintile of performance in the past ten years, 10.4 percent of employee savings are allocated to employer stock, compared to 39.7 percent for companies in the top quintile. This difference does not reflect information about future returns. Companies with a higher fraction of employees investing in employer stock underperform relative to companies with a lower fraction.

Barberis, Shleifer, and Vishny (1998) present an alternative model of the law of small number and apply it to financial markets. While the draws are i.i.d., investors believe that the draws come from either a ‘mean-reverting’ regime or a ‘trending’ regime; in addition, the investors believe that the first regime is more likely ex ante. If investors observe a sequence of identical signals, in the short-run they expect a mean-reverting regime (the gambler’s fallacy); hence, the returns under-react to information, inducing short-term positive correlation (momentum). However, after a longer sequence, the individuals overinfer, as in Rabin (2002b), and expect a ‘trending’ regime; this induces a long-term negative correlation of returns.

### 3.3 Projection Bias

A third way in which individuals have systematically incorrect beliefs is that they expect their future preferences to be too close to the present ones; for example, they project current hunger levels on the future. Read and van Leeuwen (1998) asked office workers to choose a healthy snack or an unhealthy snack to be delivered a week later (in the late afternoon). Workers were asked either when they were plausibly hungry (in the late afternoon) or when satiated (after lunch). In the first group, 78 percent chose an unhealthy snack, compared to 42 percent in the second group. Similarly, individuals under-appreciate the extent to which they adapt to future circumstances. (Gilbert et al., 1998).

Loewenstein, O’Donoghue and Rabin (2003) propose a simple model of projection bias. Assume that utility  $u$  is a function of consumption  $c$  and of a state variable  $s$ , that is,  $u =$

$u(c, s)$ . The current state is  $s'$  and the future state is  $s$ . Then, when predicting the future utility  $\hat{u}(c, s)$ , an individual with projection bias expects utility

$$\hat{u}(c, s) = (1 - \alpha) u(c, s) + \alpha u(c, s') \quad (8)$$

rather than  $u(c, s)$ . The parameter  $\alpha \in [0, 1]$  captures the extent of projection bias, with  $\alpha = 0$  denoting the standard case and  $\alpha = 1$  the case of full projection bias. This model can capture the mis-prediction of future hunger, as well as the under-appreciation of adaptation.

Conlin, O'Donoghue, and Vogelsang (forthcoming) present evidence of projection bias using a data set of 2 million orders of cold-weather apparel items. They consider the effect of weather at the time of purchase on the probability that an item is returned, conditional on purchase. According to the standard model, colder weather at the time of purchase should not affect the probability of a return, or may affect it negatively (since colder weather at the time of purchase is correlated with colder weather over the subsequent days). Projection bias, instead, makes the opposite prediction. On colder days, individuals overestimate the use that they will make of a cold-weather item, and hence are ex post more likely to return the item. This prediction holds whether the projection bias regards future utility, as in (8) ('I expect to like cold-weather items very much'), or future weather ('I expect the coming winter to be very cold').

Conlin et al. (forthcoming) find that a reduction in the order-date temperature of 30°F—corresponding to a decrease, for example, from 40°F to 10°F—increases the average return rate of a cold-weather item by 3.96 percent, consistent with projection bias. A simple structural model of projection bias as in (8) implies estimates for  $\hat{\alpha} \approx 0.5$ , implying that consumers predict future tastes roughly half-way between present tastes and actual future tastes.

## 4 Non-standard Decision-Making

Even given utility  $U(x|s)$  and belief  $p(s)$ , individuals make non-standard decisions. We analyze: (i) the neglect (or overweighting) of information because of limited attention; (ii) sub-optimal heuristics used for choices out of menu sets; (iii) social pressure—explicit pressure by others—and persuasion—excess impact of the beliefs of others; (iv) emotions.

### 4.1 Limited Attention

In the starker form of the standard model, individuals make decisions using all the available information. Since Simon (1955), economists have attempted to relax this strong assumption and have proposed models in which individuals simplify complex decisions, for example by processing only a subset of information<sup>25</sup>. In economic experiments, the simplifying heuristics include thinking only one step ahead in dynamic problems (Gabaix et al., 2006).

---

<sup>25</sup>Conlisk (1996) provides an early survey of this literature. We discuss the model of inattention by Gabaix and Laibson (2006) in Section 5.

The laboratory studies in psychology indicate that attention is a limited resource. In studies of dichotic listening (Broadbent, 1958), for example, subjects hear different messages in the right ear and in the left ear, and are instructed to attend to one of the messages. When asked about the other message, they remember very little of it. Moreover, in treatments in which they have to rehearse a sentence or a sequence of numbers while listening, their capacity to attend to a message is substantially lower.

We model attention as a scarce resource. Consider a good whose value  $V$  (inclusive of price) is determined by the sum of two components, a visible component  $v$  and an opaque component  $o$ :  $V = v + o$ . Due to inattention, the consumer perceives the value to be  $\hat{V} = v + (1 - \theta)o$ , where  $\theta$  denotes the degree of inattention, with  $\theta = 0$  as the standard case of full attention. The interpretation of  $\theta$  is that each individual sees the opaque information  $o$ , but then processes it only partially, to the degree  $\theta$ .<sup>26</sup> The inattention parameter  $\theta$  is itself a function of the salience  $s \in [0, 1]$  of  $o$  and of the number of competing stimuli  $N$ :  $\theta = \theta(s, N)$ . Based on the psychology evidence, I assume that the inattention  $\theta$  is decreasing in the salience  $s$  and increasing in the competing stimuli  $N$ :  $\theta'_s < 0$  and  $\theta'_N > 0$ . Inattention is zero for a fully salient signal:  $\theta(1, N) = 0$ . The consumer's demand is  $D[\hat{V}]$ , with  $D'[x] > 0$  for all  $x$ .

This framework suggests, broadly speaking, three strategies to identify the inattention parameter  $\theta$ , which the papers describe below undertake. The first is to compute how the valuation  $\hat{V}$  responds to a change in  $o$ ; the derivative  $\partial\hat{V}/\partial o = (1 - \theta)$  can be compared to  $\partial\hat{V}/\partial v = 1$  to test for limited attention. Hossain and Morgan (2006) and Chetty, Looney, and Kroft (2007) in the section on alcohol taxes follow this avenue. The second is to examine the response of consumer valuation to an increase in the salience  $s$ ,  $\partial\hat{V}/\partial s = -\theta'_s o$ , and test whether it differs from zero. This is the strategy of Chetty et al. (2007) in their field experiment. The third strategy is to vary the number of competing stimuli  $N$ ,  $\partial\hat{V}/\partial N = -\theta'_N o$ , and test whether this has an effect. This is the strategy of DellaVigna and Pollet (2006) and Hirshleifer, Lim, and Teoh (2007). All three of these strategies identify a piece of opaque information  $o$  with regards to which the decision-makers are not fully attentive.

This research is subject to two caveats. The first caveat is that measuring the salience of information involves a subjective judgment, similar to the judgment involved in setting the reference point in prospect theory. While in most settings (such as the ones in this Section) it is rather clear which features are visible and which are opaque, the psychology experiments do not provide a general criterion. The second caveat is that we do not address whether the inattention is rational or not. In general, models of limited attention can be rephrased as rational model with information costs in which less salient information has higher costs of acquisition. In most of the examples below, however, the opaque information is publicly

---

<sup>26</sup>An alternative model (Chetty, Looney, and Kroft, 2007) posits that  $\theta$  is the probability that an individual perceives the opaque signal, rather than the degree to which each individual incorporates the signal. This alternative model leads to similar results but a more cumbersome solution for settings like an auction.

available at a zero or small cost (for example, the information on earnings announcements), making a rational interpretation of the findings less plausible.

**Inattention to Shipping Costs.** In eBay auctions, the price of an item is more vivid than the shipping cost, because the shipping cost is not listed in the item title and also because historically most purchases have not involved shipping. Define  $v$  as the value of the object and  $o$  as the negative of the shipping cost:  $o = -c$ . Since eBay is (essentially) a second price auction, the inattentive bidders bid their value net of the (perceived) shipping cost:  $b^* = v - (1 - \theta)c$ . The revenue raised by the seller is  $b^* + c = v + \theta c$ . A \$1 increase in the shipping cost  $c$ , therefore, increases revenue by  $\theta$  dollars. In the case of full attention ( $\theta = 0$ ), increases in the shipping cost have no effect on revenue. Hossain and Morgan (2006) examine these predictions with a field experiment. In the treatment  $c_{LO}$ , they auction CDs with a \$4 reserve price and no shipping cost, while in treatment  $c_{HI}$  they auction CDs with a \$.01 reserve price and a \$3.99 shipping cost. The change in reserve price guarantees that the two auction are equivalent for a fully attentive bidder. The average revenue raised in treatment  $c_{HI}$  is \$1.79 higher (\$10.16 vs. \$8.37) than in treatment  $c_{LO}$ , and is higher for 9 out of 10 CDs<sup>27</sup>. These estimates imply substantial inattention:  $\hat{\theta} = 1.79/3.99 = .45$ . A second set of auctions with higher shipping costs ( $c_{LO} = \$2$  and  $c_{HI} = \$6$ ), leads to smaller increase of revenue in the high-shipping cost condition (\$12.87 vs. \$12.15), corresponding to an inattention parameter  $\hat{\theta} = 0.72/4 = .18$ .

**Inattention to Non-Transparent Taxes.** Chetty et al. (2007) study whether consumers are inattentive to taxes that are not transparently factored in the price of a good, like indirect state taxes. They use data on the demand for items in a grocery store. Assume that demand  $D$  is a function of the visible part of the value  $v$ , including the price  $p$ , and of the less visible part  $o$ , capturing the state tax  $-tp$ :  $D = D[v - (1 - \theta)tp]$ . The change in log-demand  $\Delta \log D$  from making the tax fully salient ( $s = 1$  and hence  $\theta = 0$ ) is (linearizing the demand)  $\log D[v - tp] - \log D[v - (1 - \theta)tp] = -\theta tp * D'[v - (1 - \theta)tp]/D[v - (1 - \theta)tp] = -\theta t * \eta_{D,p}$ , where  $\eta_{D,p}$  is the price elasticity of demand. Notice that the response is zero for fully attentive investors ( $\theta = 0$ ). This implies  $\theta = -\Delta \log D/(t * \eta_{D,p})$ . Chetty et al. (2007) manipulate the salience of taxes with a field experiment. In a three-week period, the price tags of certain items indicate the after-tax price, in addition to the pre-tax price. Compared to previous-week sales for the same item, and compared to items for which tax was not made salient, the average quantity sold decreases (significantly) by 2.20 units relative to a baseline level of 25, an 8.8 percent decline. Since the price elasticity  $\eta_{D,p}$  in this sample is estimated to be  $-1.59$  and the tax is 7.375 percent, we can compute  $\hat{\theta} = -(-.088)/(-1.59 * .07375) \approx .75$ . In a separate estimation strategy, Chetty et al. (2007) identify the impact on beer consumption of changes across States and over time in the excise and sales taxes. Since the excise tax is included in the price, while the sales tax is added at the register, inattentive consumers should be more responsive to changes in the excise tax than to changes in the sales tax. Indeed, the first elasticity is

---

<sup>27</sup>We exclude CDs that do not sell from this computation; the difference would be \$2.60 if they were included.

substantially larger, leading to an estimate of the inattention parameter of  $\hat{\theta} = .94$ . Consumer inattention to non-transparent taxes is substantial.

**Inattention to Complex Information in Rankings.** In other settings, the familiarity of information depends on the simplicity of the data format. Pope (2007) studies the response of consumers to rankings of hospitals and colleges by the *US News and World Report*. Each year, the company constructs a continuous quality score from 0 to 100 largely based on reputation scores, and then creates rankings based on this score. Both the scores and the rankings are published in the yearly report. While the continuous score contains all the information, the rankings are presumably easier to process (“No. 5 Hospital” vs. “Hospital with 89/100 score”). Pope shows that, holding constant the quality score, hospital discharges respond significantly to differences in ranks among hospitals; similarly, college applications respond to differences in ranks among colleges. Pope (2007) also provides a calibration of the inattention or thinking costs necessary to justify this result.

**Inattention to Financial News.** Limited attention among investors induces under-reaction to newly-released information and hence can explain anomalies such as momentum (Hong and Stein, 1999). Huberman and Regev (2001) examines the case of the company EntreMed, an interesting example of under-reaction to information. On November 28, 1997, Nature prominently features an article reporting positive results on a cure for a type of cancer for a drug patented by EntreMed. On the same day, the New York Times reports an article on the same topic on page A28. Unsurprisingly, the stock price of EntreMed increases by 28 percent. What is surprising is what happens next. On May 4, 1998, the New York Times publishes on the front page an article on EntreMed that is very similar to the article that it had already published in November. Despite the fact that the article contains no new hard information, it leads to a 330 percent one-day return for EntreMed, and to a 7.5 percent one-day return for all bio-tech companies, moving billions in market capitalization. The stock price of EntreMed does not revert to the previous level over the whole next year.

While this is just a case study, it stresses the importance of studying systematically the response to new information. One important setting is the release of quarterly earnings news, and the consequent response of asset prices. To simplify, assume that  $v$  is the known information about cash-flows of the company, and that  $o$  is the new information contained in the earnings announcement. On the day before the announcement, the company price is  $P = v$ . On the day of the announcement, the updated company value is  $v + o$ . However, since the investors are inattentive, the asset price  $P$  responds only partially to the new information:  $P = v + (1 - \theta)o$ . Over time, as the information makes its way to the inattentive investors (for example through additional articles as in the EntreMed case), the price incorporates the full value  $v + o$ . This implies that the short-run stock return  $r_{SR}$  equals  $r_{SR} = (1 - \theta)o/v$ ; the long-run stock return  $r_{LR}$ , instead, equals  $r_{LR} = o/v$ . In this example, a measure of investor attention is  $(\partial r_{SR}/\partial o)/(\partial r_{LR}/\partial o) = (1 - \theta)$ . (The division by  $(\partial r_{LR}/\partial o)$  is a re-normalization

which makes the measure scale-invariant) The higher is the inattention, the smaller is the immediate response and the larger is the predictability of stock returns in the days following the announcement, a phenomenon known as post-earnings announcement drift (Bernard and Thomas, 1989). Inattention leads to delayed absorption of information.

While this setting is highly stylized, similar results obtain after allowing for uncertainty and arbitrage, as long as arbitrage is limited by risk aversion and short investor horizons (for example, DellaVigna and Pollet, 2006). DellaVigna and Pollet (2006) estimate the empirical counterpart of  $(\partial r_{SR}/\partial o)/(\partial r_{LR}/\partial o)$  using the response of returns  $r$  to the earnings surprise  $o$ . They measure returns in the 2 days surrounding an announcement ( $r_{SR}$ ) and over the 75 trading days from an announcement ( $r_{LR}$ ). The immediate response captures 58 percent of the overall response, implying substantial inattention:  $\hat{\theta} \approx .42$ . If the delayed response is due to attention deficits, the delay should be even stronger when a higher share of investors are distracted (higher  $\theta$ ). DellaVigna and Pollet (2006) use the weekend as a proxy of investor distraction. For announcements made on Friday, indeed, the share of immediate response  $(\partial r_{SR}/\partial o)/(\partial r_{LR}/\partial o)$  is 41 percent, implying  $\hat{\theta} \approx .59$ , consistent with higher inattention before the weekend. This provides an explanation for the observed release of worse earnings on Friday: companies maximizing short-term value release worse news on low-attention days.

In a similar context, Hirshleifer, Lim, and Teoh (2007) analyze the impact of informational overload (high  $N$  in the framework above). They find that the incorporation of earnings news into stock prices is 20 percent slower on days in which more announcements take place. Increasing the amount of competing information accentuates the effect of limited attention.

Another related study is Cohen and Frazzini (forthcoming), which analyzes how investors respond to indirect, and hence less salient news (low  $s$  in the framework above). They consider companies linked in the supplier-customer chain. When a customer company announces substantial earnings news, the news affects also the supplier, but this indirect effect is less likely to attract attention. Indeed, Cohen and Frazzini (forthcoming) show that suppliers of companies which experience declining stock returns have lower stock returns 1 to 3 months later. They measure the speed of the response of returns to news about the customer company using  $(\partial r_{SR}/\partial o)/(\partial r_{LR}/\partial o)$ , where  $r_{SR}$  is the one-month return and  $r_{LR}$  is a seven-month return. They find that for the customer company, 93 percent of the overall response occurs in the initial month; for the supplier company, instead, only 60 percent of the overall response occurs in the first month, suggesting substantial inattention to indirect links.

A final dimension of salience  $s$  is the temporal distance. Holding constant the informativeness, information that is further into the future (or past) is less likely to be salient. In general, it is difficult to control for informativeness, since information that is further away is usually less relevant or less precisely estimated. DellaVigna and Pollet (forthcoming) address this issue by considering future demand shifts due to demographics. Unlike other forecasters, cohort size shifts are highly predictable even ten years into the future. For example, if a large

cohort is born in 2006, school bus companies in 2012 are going to experience a forecastable increase in demand and, if the market is not perfectly competitive, in profits. If investors are perfectly attentive, this increase will be incorporated into returns already in 2006 and stock returns from 2006 to 2012 will not be predictable using demographic information. However, if investors neglect information beyond 5 years into the future, the stock prices will increase only in 2007, and stock returns from 2006 to 2012 will be predictable using public information on demographics. Using data for 48 industries from 1939 to 2003, DellaVigna and Pollet show that the growth rate in demand due to demographics 5 to 10 years ahead forecasts stock returns in an industry positively. These results are consistent with inattention to information further than approximately 5 years into the future.

**Summary.** Limited attention helps explain the (partial) neglect of: (i) shipping costs in eBay auctions; (ii) non-transparent taxes; (iii) complex information in rankings; (iv) earnings news, specially before weekends and on days with more competing news; (v) news about linked companies; (vi) demand shifts in the distant future. As an example of application to another field, a literature on inattention in macroeconomics developed from the models of sticky information of Mankiw and Reis (2002) and of rational inattention of Sims (2003).

## 4.2 Menu Effects

In this Section, I consider choices out of a (and typically large) menu set, such as for investment options or politicians on a ballot. The evidence in psychology suggests that individuals use (at least) five sub-optimal heuristics to simplify these decisions: (i) *Excess Diversification* (or *1/n Heuristic*); (ii) *Choice Avoidance*; (iii) *Preference for the Familiar*; (iv) *Preference for the Salient*; (v) *Confusion in Implementing the Choices*.

**Excess Diversification.** Individuals facing a complex choice may simplify it by diversifying excessively across the options. An example in psychology is Simonson (1990). In a first treatment (simultaneous condition), students in a class chose snacks to be consumed over the next three class meetings, one per meeting. In a second treatment (sequential condition), the subjects chose the snack sequentially on each of the three class meetings. In the simultaneous condition, the subjects display excess diversification: 64 percent of subjects chose three different snacks, while in the sequential condition only 9 percent of subjects made this choice.

Benartzi and Thaler (2001) study whether excess diversification applies to 401(k) investments. As a special case, they study the case of equal diversification across the  $n$  available options, the *1/n heuristics*. They use aggregate data on the 1996 plan assets for 162 companies which offer an average of 6.8 plan options. Lacking individual-level data, they study an aggregate implication of the 1/n heuristic. If individuals divide their investments approximately equally across options, their exposure to equity will be increasing in the availability of equity

options in the 401(k) plan. Across plans, Benartzi and Thaler estimate the relationship

$$\%Invested\ In\ Equity = \hat{\alpha} + .36(.04) * \%Equity\ Options + \hat{B}X \quad (9)$$

(s.e. in parentheses), where the control variable  $X$  is the availability of employer stock in the portfolio. In companies with an equity share that is 10 percentage points higher, the employees invest 3.6 percent more in equity plans. This finding is consistent with a weak form of the 1/n heuristic (If the employees followed the 1/n heuristics strictly, the coefficient should be 1 rather than .36). A confound is that the equity content of a plan may be designed to cater to the preferences of the employees, resulting in reverse causation.

Huberman and Jiang (2006) investigate the investor diversification using a data set on the individual choice of employees in 647 401(k) plans managed by Vanguard. They estimate specification (9) at the individual level with a large set of individual-level and plan-level controls  $X$ . They obtain the relationship  $\%Invested\ In\ Equity = \hat{\alpha} + .29(.06) * \%Equity\ Options + \hat{B}X$  for funds with less than 10 options and  $\%Invested\ In\ Equity = \alpha + .06(.07) * \%Equity\ Options + \hat{B}X$  for funds with more than 10 options. The relationship predicted by the 1/n heuristic, therefore, is present when the number of funds is small (as in the Benartzi and Thaler sample), but not when the number of funds is large. Huberman and Jiang provide additional evidence suggesting that the predictive power of the 1/n heuristic is low. In particular, the number of funds chosen by employees hardly responds at all to the number of investment options offered in the plan. (This test differs from the one above as it is not conditional on equity vs. non-equity choices.) There is some evidence of a *conditional 1/n heuristic*: conditional on the allocations chosen, individuals allocate their savings approximately equally. 37 percent of employees follow this behavior among employees investing in 4 funds, 26 percent among employees investing in 5 funds, and 53 percent among employees investing in 10 funds; the behavior is instead not common for non-round numbers. Overall, some employees use a version of the 1/n heuristic when the number of investment options is small; when the number is large, other heuristics, which I discuss next, are at play.

**Choice Avoidance.** Iyengar, Huberman, and Jiang (2004) analyze the effect of the number of investment options on the participation in 401(k) plans. They find that the higher is the number of options, the lower is the participation rate. On average, 75 percent of employees participate in plans with only 2 funds available, but the participation rate falls to 65 percent when employees choose between 40 or more funds. Choi, Laibson, and Madrian (2006) provide additional evidence that a smaller number of options increases 401(k) plan participation. Participation in a 401(k) plan increases by 10 percentage points when non-participating employees receive a card that allows them, if mailed back, to enroll in a default plan (3 percent contribution in a balanced fund).<sup>28</sup> These findings are surprising in light of the standard theory—more options should increase the likelihood that the marginal individual invests—and in light of an

---

<sup>28</sup>The increase may be due to a reminder effect of the card. However, in other settings, reminders, and more

endogeneity explanation—more choices should be offered where consumers value them more.

The research in psychology provides a simple explanation: When the choice set is large, individuals are less likely to make a choice. In a field experiment, Iyengar and Lepper (2000) compare the behavior of consumers in an upscale grocery store, where at some times consumers were offered the opportunity to taste 6 jams, while at other times the tasting included 24 jams. They find that in the 24-jam condition, more consumers stop to sample jams (145 versus 104 customers), but substantially *fewer* buy jams (4 versus 31 customers).

Bertrand et al. (2006) finds a similar result in a field experiment on the mailing of 50,000 loan offers in South Africa. The authors randomize, among other things, the format of the table illustrating the use of the loan. The small-table format lists only one loan size as an example, while the big-table format presents four different loan sizes. The take-up in the small-table format is .6 percentage points larger compared to a baseline of 8 percentage points, an effect size equivalent to a reduction of the (monthly) interest rate by 2.3 percentage points.

**Preference for the Familiar.** A different heuristic arising in the case of large menu sets is the choice of a familiar option. This tendency is wide-spread among individual investors. Investors in the USA, Japan, and the UK allocate 94%, 98%, and 82% of their equity investment, respectively, to domestic equities (French and Poterba, 1991). While the preference for own-country equity may be due to costs of investments in foreign assets, the same pattern appears for within-country investment. Huberman (2001) documents the geographical distribution of the shareholders of the Regional Bell companies. The fraction invested in the own-state Regional Bell is 82 percent higher than the fraction invested in the next Regional Bell company. The preference for the familiar occurs despite substantial costs of under-diversification.

A particularly egregious case is the preference for own-company stock. On average, employees invest 20-30 percent of their discretionary funds in employer stocks (Benartzi, 2001), despite the fact that the employees' human capital is already invested in their company. This choice does not reflect private information about future performance. Companies where a higher proportion of employees invest in employer stock have lower subsequent one-year returns, compared to companies with a lower proportion of employee investment.

The preference for familiar options is consistent with ambiguity aversion. As in the classical Ellsberg (1961) paradox, investors that are ambiguity-averse may prefer an investment with known distribution of returns to an investment with unknown distribution, even if the average returns are the same for the two investments, and despite the benefits of diversification.

**Preference for the Salient.** Barber and Odean (forthcoming) show that individual investors simplify complex portfolio decisions also by choosing a salient option. Using individual

---

generally financial education, do not have such large effects. For example, Choi, Laibson, and Madrian (2005) sent a survey including 5 questions on the benefits of employer match to 345 employees that were not taking advantage of the match. A control group of 344 employees received the same survey except for the 5 specific questions. The treatment had no significant effect on the savings rate.

trading data, they show that individual investors are net buyers of companies with unusually high, or low, performance in the previous day, of companies with high trading volume, and of companies in the news. The effects are large: for companies in the highest or lowest decile of the previous day's returns, the Buy-Sell imbalance ( $\text{Buy-Sell}/(\text{Buy}+\text{Sell})$ ) for individual investors is 20 percentage points higher than for companies in the fifth decile. These results suggest that individuals solve the informational overload problem of which stocks to buy by picking companies that stand out. The same problem does not present itself for stock sales, since most investors own only a small number of stocks at any given time.

The preference for the salient takes different forms in different contexts. In the choice of candidates on a ballot, the first politician on the list stands out. Ho and Imai (2004) provides evidence that order of the candidates matters even when the order is random. They exploit the natural experiment induced by the California voting system which, since 1975, explicitly randomizes the ballot order of candidates across Assembly Districts. They show that in the 1998 and 2000 general elections a minor party candidate experiences on average a 10 percent increase in votes when first on the list. The effect is instead very small for candidates of the major parties, suggesting that irrelevant information is used as a tie-breaker for cases in which the decision-maker has less information. In primary elections, in which candidates are on average less known, the effect is stronger: the impact of being first in the list is on average a 20 percent increase, roughly 1.6 percent of the party vote.

**Confusion.** A final category, confusion, differs from the previous heuristics in that it does not reflect a preference, whether to avoid difficult choices or for salient options, but simply an error in the implementation of the preferences. As such, it differs from most behavioral phenomena which reflect a directional bias. A first setting is the choice of a political candidate among those in a ballot. Shue and Luttmer (2007) consider California voters in the 2003 recall elections and exploit the random variation in the placement of candidates on the ballot, similarly to Ho and Imai (2004). They find that the vote share of minor candidates is significantly higher for candidates whose name on the ballot is adjacent to the name of a major candidate. While this phenomenon could be due to a spill-over in attention, confusion is a more likely explanation: the effect of horizontal adjacency (a name to the right or to the left of the major candidate) is almost entirely due to adjacency on the confusing side. For example, in the sequence *Bubble, Candidate A, Bubble, Schwarzenegger, Bubble, Candidate B*, it is Candidate B that benefits from the presence of a major candidate, since some voters mistake its bubble for the bubble of Schwarzenegger. Candidate A does not benefit, nor do candidates located at a diagonal adjacency. Further, the spill-over of votes is larger for more confusing voting methods (such as punch-cards) and for precincts with a larger share of lower-education demographics, that are more likely to make errors when faced with a large number of options. This method allows for a measure of confusion. Across different voting methods, about 3 in 1,000 voters meaning to vote for a major candidate instead vote for a minor candidate. The phenomenon

hence is small but not irrelevant. Importantly, it can have an aggregate effect, since confusion is likely to have a different prevalence among the voters of different major candidates.

Interestingly, Rashes (2001) identifies a similar phenomenon in the choice of stocks. The article focuses on an egregious case of potential confusion between two companies, MCI and MCIC. The ticker for the MCI communication company is MCIC, while MCI is the ticker for a little-known closed-end mutual fund, Massmutual Corporate Investors. Using days with exceptional news regarding the MCIC company, the authors estimate a degree of confusion comparable to the one in the voting data. Between 1 and 10 out of 1,000 investors attempting to trade the MCIC company purchase instead shares of the closed-end mutual fund. Since this latter company has a much smaller trading volume, the confusion has significant effects on its trading and causes a significant correlation in daily stock returns between the two companies.

**Summary.** When choosing from a large menu of options, decision-makers: (i) (to same extent) diversify excessively across the options; (ii) avoid the choice and do not invest (or do not purchase); (iii) choose familiar options, such as own-country or own-company stock; (iv) choose salient options in investment choice or at the ballot; (v) display some confusion in implementing their choices.

### 4.3 Persuasion and Social Pressure

**Persuasion.** In the standard model, individuals take into account the incentives of the information provider. The neglect of incentives can lead to excess impact of the beliefs of the information provider, which I label *persuasion*. An example from psychology is Cain, Loewenstein, and Moore (2005). The subjects are paid for the precision of the estimates of the number of coins in a jar. Since they see the jar only from a distance, they have to rely on the advice of a second group of subjects, the advisors, that inspect the jar from close. The two experimental treatments vary the incentives for the advisors. In a first treatment, the advisors are paid for how closely the subjects guess the number of coins; in a second treatment, the advisors are paid for how high the subjects' guess is. Despite the fact that the incentives are common-knowledge, the estimate of the subjects is 28 percent higher in the second treatment. The subjects do not discount enough for the conflict of incentives of the advisors.

In a financial setting, Malmendier and Shanthikumar (2007) analyze how small and large investors respond to recommendation by analysts. Analyst forecasts are notoriously biased upward—94.5 percent of recommendations are Hold, Buy, or Strong Buy—, and affiliated analysts are even more biased. Malmendier and Shanthikumar (2007) show that large investors take into account this bias, and discount the information: for example, they respond to a Hold recommendation by selling the shares of a company, and they discount heavily positive recommendations by affiliated analysts. Small investors, instead, are subject to persuasion. They follow the recommendations literally—for example holding a stock in response to a Hold

recommendation—and do not discount for the additional distortions due to analyst affiliation.

In a political setting, DellaVigna and Kaplan (2007) tests whether the information provided by a news source convinces on average its audience. They exploit the geographical variation in the introduction in the cable programming of the Fox News Channel, a more conservative channel relative to the pre-existing news sources (CNN and the networks). Fox News availability in the town cable programming in 2000 appears to be largely idiosyncratic, conditional on a set of controls. Using the voting data for 9,256 towns, they find that the vote share for Republicans in 2000 is half-a-percentage point higher in the towns offering Fox News. They estimate that Fox News convinced 5 to 30 percent of the audience that was not already Republican, depending on the audience measure. This impact can be a temporary effect for Bayesian voters that are learning about the bias of Fox News, or a persuasion effect for non-rational voters that do not take into account the political orientation of Fox News.

**Social Pressure.** A separate reason for excess impact of the beliefs of others is the pressure to conform, or social pressure (Akerlof, 1991). Two classical laboratory experiments illustrating the power of social pressure are Asch (1951) and Milgram (1963). In the Asch (1951) experiment, the subjects are given two cards: the first card has a line on it, while the second card has three lines of substantially differing length (one of the same length as in the first card). The subjects are asked which of the lines in the second card is closest in length to the line in the first card, after 4 to 8 subjects (who, unbeknownst to them, are confederates) unanimously choose the wrong answer. On average, over a third of subjects give the wrong answer to avoid disagreeing with the unanimous judgment of the other participants. In a control group with no confederates, less than 2 percent of subjects give the wrong answer. While this result could be interpreted as social learning, the learning is unlikely to be about the length of the line, but possibly about the rules of the experiment. It should also be pointed out that the subjects were not paid for accuracy.

In the Milgram (1963) experiment, a group of subjects is told that their task is to monitor the learning of another subject (a confederate), and to inflict electric shocks on this subject when he makes an error. Encouraged by the experimenter, 62 percent of the subjects escalate the electric shocks up to a level of 450 Volts, despite hearing the subject scream in pain. This proneness to obedience comes as a surprise to the subjects themselves. When a different group of 40 subjects is provided with a description of the experiment and asked to predict how far subjects would go in inflicting shocks, no one predicts that 450 Volts would be reached.

In the field, social pressure is hard to distinguish from rational diffusion of information. In some studies, however, the social pressure motive is evident. Garicano, Palacios-Huerta, and Prendergast (2005) measure the length of extra-time that referees assign at the end of a game of soccer; in the extra-time the teams can score goals. They find that referees on average give twice as much extra time (4 minutes versus 2 minutes) when the extra time is bound to advantage the local team (1 goal behind) than when it is bound to hurt it (1 goal ahead).

The effect is larger when stakes are higher (toward the end of the season) and when the social pressure is larger (larger attendance at the game). Referees respond significantly to pressure by the local public, despite official rules on what determines the length of extra-time.

Some of the *peer effect* literature also points to the importance of social pressure. Falk and Ichino (2006) measure the effect of peer pressure on task performance. High-school students in Switzerland were recruited to perform a one-time job for a flat payment; they were instructed to stuff letters into envelopes for 4 hours. The control group of 8 students did the task individually, while the treatment group of 16 students worked in pairs (but each student was instructed to stuff the envelopes individually). Students in the treatment group stuffed significantly more envelopes (221 vs. 190), and coordinated the effort within group: the within-pair standard-deviation of output is significantly less than the (simulated) between-pairs standard deviation.

While the results of Falk and Ichino (2006) could also be due to social learning, Mas and Moretti (2006) presenting direct evidence of social pressure. They find that high-productivity cashiers in a supermarket chain increase the productivity of co-workers that are present in the same shift. The effect is not due to exchange of information, such as on a price tag. The positive peer effect occurs only when the more productive co-worker is behind and therefore can observe the other worker's productivity. The effect is quite large: a one percent increase in the average permanent productivity of the workers behind increases the productivity of the peer by .23 percent; the effect is even larger for co-workers that are working at a closer distance. There is no effect of a highly-productive co-worker in front.

#### 4.4 Emotions

Some of the previous phenomena, such as self-control problems, social preferences in giving, and projection bias in food purchase are likely mediated (at least partially) by emotional states, respectively temptation, empathy, and hunger. A large literature in psychology suggests that emotions play an important role in decision-making, and that different emotions operate very differently (Loewenstein and Lerner, 2003). In this Section, I consider two examples of emotions, mood and arousal, for which field evidence is available.

In experiments, even minor mood manipulations have a substantial impact on behavior and emotions. For example, on sunnier days, subjects tip more at restaurants (Rind, 1996) and express higher levels of overall happiness (Schwarz and Clore, 1983). In the field, mood fluctuations induced by the weather affect stock returns, despite the fact that daily weather fluctuations are unlikely to affect fundamentals. Days with higher cloud cover in New York are associated with lower aggregate US stock returns (Saunders, 1993). Hirshleifer and Shumway (2003) extend this analysis to 26 countries between 1982 and 1997 using the weather of the city where the stock market is located. They find a negative relationship between cloud cover (detrended from seasonal averages) and aggregate stock returns in 18 of the 26 cities. Days with

completely covered skies have daily stock returns .09 percent lower than days with sunny skies, five percent of a standard deviation. After controlling for cloud cover, other weather variables such as rain and snow are unrelated to returns. If mood is the channel for these effects, other mood-altering events should have similar effects. Indeed, international soccer matches impact the daily stock returns for the losing country (Edmans, Garcia, and Norli, 2007). Compared to a day with no match, a loss lowers daily returns (significantly) by .21 percent. (Surprisingly, a win has essentially no effect). More important matches, such as World Cup elimination games, have larger effects. The effect does not appear to depend on whether the loss was expected or not. International matches in other sports have a consistent, though smaller, effect.

The effect of these mood-altering events on returns is likely due to (i) an impact on risk aversion or perception of volatility, or (ii) a projection of the mood to economic fundamentals. The evidence above does not allow to distinguish these two effects. Mood induced by atmospheric factors can also induce subtler changes in behavior. Simonsohn (2007) examines the role of weather on the day of campus visit to a prestigious university. Students visiting on days with more cloud cover are significantly *more* likely to enroll. Simonsohn suggests that higher cloud cover induces the students to focus more on academic attributes versus social attributes of the school, a hypothesis supported by laboratory experiments.

A second set of laboratory experiments suggests that emotional arousal has an important effect on decisions. In one experiment, subjects that are sexually aroused as part of the treatment report a substantially higher willingness to engage in behavior that may lead to date rape (Ariely and Loewenstein, 2005). In other experiments, subjects exposed to violent video clips are more likely to display more aggressive behavior, such as aggressive play during a hockey game, compared to a control group watching non-violent clips (Josephson, 1987). The impact is partly due to imitation and partly to arousal.

Dahl and DellaVigna (2007) provide field evidence on the short-run impact of exposure to media violence on violent crime. They exploit the time-series variation in movie violence at the box office and compare days where the blockbuster movies are violent to days in which the blockbuster movies are non-violent. They find that on days in which exposure to media violence is higher, violent crime is lower. This effect is not only due to incapacitation because the potential criminals are in the movie theater. In the night following the exposure (midnight to 6AM), for every million people exposed to violent movies, violent crime is 1.5 percent lower. The difference between the laboratory and the field evidence is likely due to differences in design. The laboratory experiments capture the impact of surprise exposure to violence, while the field evidence captures the impact when individuals self-select. Arousal does not induce as much aggression in the short-run for the individuals who choose to watch violent media.

## 5 Market Response

In the previous Sections, I have documented how consumers deviate from the standard model in their choices of credit cards, clothing items, eBay bidding strategies, giving, health clubs, insurance contracts, and loans. I have discussed how workers make non-standard effort, labor supply, and retirement savings decisions. I have provided evidence of disposition effect, inattention, and overtrading among investors. Finally, I documented how voters are affected by irrelevant factors such as the order of politicians.

This evidence is just the first step towards a better understanding of markets where agents display non-standard preferences and beliefs. This evidence raises a natural question: how do markets and institutions respond to these non-standard features? An important test for Psychology and Economics is whether it helps to understand markets and institutions.

This Section discusses how rational actors respond to the non-standard features of other agents. Profit-maximizing firms respond to the non-standard features of consumer behavior in their contract design and pricing (Behavioral Industrial Organization). Employers tailor their employment contracts to the non-standard behavior of the employees (Behavioral Labor Economics). In response to the non-standard behavior of investors, rational investors alter their trading strategies, and firm managers alter the capital structure (Behavioral Finance and Behavioral Corporate Finance). Politicians change their behavior to respond to voter biases (Behavioral Political Economy). Finally, policy-makers can use the findings in Psychology and Economics to inform the design of institutions and of policy (Behavioral Institutional Design).

Before I proceed, I discuss an important caveat. If consumers have non-standard features, why should one expect firms, employers, financial operators, and politicians to not have them? Experience is a key difference. Unlike individual consumers, firms can specialize, hire consultants, and obtain feedback from large data sets and capital markets. Firms are also subject to competition. Compared to consumers, therefore, firms are less likely to be affected by biases (except for principal-agent problems), and we expect them to be close to profit maximization. In addition, even if firms have non-standard features, they still have incentives to respond to the non-standard features of consumers. Similar arguments apply for employers, institutional investors, top managers, and politicians.

**Behavioral Industrial Organization.** The interaction between consumers with biases and rational, profit-maximizing firms is the central theme of the growing literature in behavioral industrial organization, surveyed in Ellison (2006). While this literature is mostly theoretical, the papers surveyed here also make predictions about observed pricing.

DellaVigna and Malmendier (2004) consider the profit-maximizing pricing with  $(\beta, \hat{\beta}, \delta)$  consumers with self-control problems. A (monopolistic) firm sells a product which, as in Section 2.1, has immediate payoff  $b_1$  and delayed payoff  $b_2$ . The set-up covers investment goods such as exercise ( $b_1 < 0$  and  $b_2 > 0$ ) and leisure goods such as gambling ( $b_1 > 0$  and

$b_2 < 0$ ). The immediate payoff  $b_1$  is stochastic with c.d.f.  $F$ . The firm produces the good at marginal cost  $c$  and sells it using a two-part tariff, with a lump-sum fee  $L$  and a unitary price  $p$ . DellaVigna and Malmendier (2004) show that the profit-maximizing price  $p^*$  satisfies

$$p^* - c = - \left(1 - \hat{\beta}\right) \delta b_2 \frac{f(\hat{\beta} \delta b_2 - p^*)}{f(\beta \delta b_2 - p^*)} - \frac{F(\hat{\beta} \delta b_2 - p^*) - F(\beta \delta b_2 - p^*)}{f(\beta \delta b_2 - p^*)}. \quad (10)$$

For standard agents ( $\beta = \hat{\beta} = 1$ ), the two terms on the right-hand side of (10) are zero: the firm prices at marginal cost,  $p^* = c$ , to align the incentives of the consumers. For sophisticated agents with self-control problems ( $\beta = \hat{\beta} < 1$ ), only the first term in (10) is non-zero: the firm prices investment goods below marginal cost ( $p^* < c$ ) and leisure goods above marginal cost ( $p^* > c$ ) to provide a commitment device—the pricing increases the consumption of investment goods and lowers the consumption of leisure goods. The deviation from marginal cost pricing,  $-(1 - \beta) \delta b_2$ , is exactly the difference in how much the current self and the future selves value the delayed payoff  $b_2$ ; hence, the firm offers a perfect commitment device. For fully naive agents with self-control problems ( $\beta < \hat{\beta} = 1$ ), only the second term in (10) is non-zero: the firm again prices investment goods below marginal cost and leisure goods above marginal cost again, but for a different reason—it takes advantage of consumer overestimation (underestimation) of the consumption of investment (leisure) goods. The deviation from marginal cost pricing is indeed a function of the mis-estimation of consumption  $F(\hat{\beta} \delta b_2 - p^*) - F(\beta \delta b_2 - p^*)$ . These results generalize to the case of perfect competition, since competition only alters the equilibrium fee  $L^*$ . This theory rationalizes the presence of contracts with no payment per visit in health clubs ( $b_2 > 0$ ), the presence of high interest rates but no annual fees for credit cards ( $b_2 < 0$ ), and cheap room rates and buffets for gamblers in Las Vegas ( $b_2 < 0$ ).

Eliaz and Spiegler (2006) generalize this analysis to allow for heterogeneity in naïveté and a more general form of time-inconsistency of preferences. They show that firms offer two types of contracts: perfect commitment devices that cater to time-inconsistent agents that are sufficiently sophisticated, and contracts that take advantage of the consumers that are sufficiently naive. Interestingly, the fully sophisticated agents do not exert any informational externality on the naïve types. Thus, the provision of the perfect commitment device does not reduce the gains that the monopolist can extract from naïve types.

Gabaix and Laibson (2006) analyze the pricing with boundedly-rational consumers that do not pay attention to hidden features of products, that they call add-ons. In equilibrium, firms charge above-marginal cost prices for the add-ons. As in DellaVigna and Malmendier (2004), the firms respond to the misprediction of future purchases. This model provides an explanation for high (hidden) fees on bank accounts and credit cards. Gabaix and Laibson (2006) also discuss how markets do not generally provide incentives for de-biasing naive consumers.

Heidhues and Köszegi (2005) study the pricing of a monopolist when consumers have reference-dependent preferences and the reference point is the rational expectations equilib-

rium (Köszegi and Rabin, 2006). Consumers are loss-averse with respect to both lower quality and higher price, relative to the reference point. The main predictions are sticky prices (despite no menu costs) and sales, two common features of pricing. In equilibrium, even if costs are stochastic, firms adjust prices seldom in response to cost shifts because consumers suffer more from price increases than they benefit from price cuts. In addition, firms offer random sales because the expectation of sales increases the likelihood of purchases at high prices.

These papers point to a dichotomy in the welfare effects of the market response. If the agents have non-standard preferences, such as self-control problems or loss aversion, but have rational expectations, the firms provide welfare-maximizing contracts. The contracts offer first-best commitment devices against the self-control problem (DellaVigna and Malmendier, 2004; Eliaz and Spiegler, 2006) or lower the probability of losses (Heidhues and Köszegi, 2005). If, instead, the agents have non-rational expectations, such as about the self-control or about the inattention, the profit-maximizing contract is likely to magnify the bias. Firms take advantage of the wrong expectations in the consumption of the tempting good (DellaVigna and Malmendier, 2004; Eliaz and Spiegler, 2006) or of the add-on (Gabaix and Laibson, 2006).

**Behavioral Labor Economics.** Contracting within a firm is also consistent with this framework. Kahneman, Knetsch, and Thaler (1986) present suggestive evidence using a survey that workers display loss aversion with respect to nominal wage losses, but not with respect to real wage losses. For example, 62 percent of respondents find unfair a wage cut of 7% in the presence of no inflation, but only 22 percent of respondents find unfair a 5% increase in salaries in presence of 12% inflation. Bewley (1999) documents similar patterns in a series of interviews. In response to a dislike for nominal wage cuts, a profit-maximizing employer should set wages such that nominal wage cuts would be rare. Card and Hyslop (1997) provide evidence on this prediction using CPS data. They consider the distribution of year-to-year changes in the nominal log wage,  $\log w_t - \log w_{t-1}$ . In the presence of aversion to nominal wage losses, we expect a discontinuity in the distribution at  $\log w_t - \log w_{t-1} = 0$ . Rather than introducing small cuts in the nominal wages that may lower morale and productivity, the employer keeps wages constant ( $\log w_t - \log w_{t-1} = 0$ ), compensating possibly by firing more workers. Card and Hyslop indeed show that a substantial fraction of the distribution of  $\log w_t - \log w_{t-1}$  is missing for negative values, despite the presence of measurement error in the wage that tends to attenuate this finding. This is an example of a market response to a bias which is likely to maximize utility for the biased agents. The observed distribution of wages is such that the employees suffer only rarely the cost of nominal wage cuts.

**Behavioral Finance.** In asset markets, arbitrage in principle is likely to limit the importance of behavioral biases such as inattention and overconfidence for price formation. If an irrational agent believes that a (fair) coin will land on tails sixty percent of the time, arbitrage by well-informed agents will keep the odds of tails around fifty percent. In actual financial markets, however, several factors limit the impact of arbitrage. DeLong et al. (1991) considers the

case of a mis-pricing that is stochastic, persistent, and correlated by so-called noise traders. If arbitrageurs are risk-averse and have a limited investment horizon, the noise traders affect the equilibrium price, despite arbitrage. If noise-traders are, for example, bullish about dot-coms, they will bid the price of dot-com shares higher. The arbitrageurs do not know whether the mis-pricing will get even worse in the next period, and given their short horizons (they have to liquidate the shares next period) they cannot short the shares aggressively enough. DeLong et al. (1991) also shows that the noise traders are not driven out of the market; under some conditions, in fact, they outperform the rational traders (since they take more risk).

The recent research in behavioral finance builds on the noise-trade models to capture the limits of arbitrage, and hence the relevance of non-standard behavior for asset prices. At the same time, this literature moved beyond these models in making explicit the source of ‘noise trading’. In Sections 3.1 and 4.1, for example, we discussed models of overconfidence and limited attention, which make specific predictions about the non-standard behavior and hence the effect on returns. The evidence on this class of models is summarized in Shleifer (2000) and Barberis and Thaler (2003).

**Behavioral Corporate Finance.** In corporate finance, the standard theory assumes that managers maximize company value subject to agency problems, given the demands of rational investors and creditors. A recent theory, known as *market timing*, expands this framework and assumes that investors may have an irrationally high or low valuation of the company. The CEO rationally responds to the mis-valuation through the equity issuance and merger decisions. CEOs provide additional shares to investors and undertake mergers when the shares are most likely to be over-priced, lowering the welfare of the biased investors. Market timing can explain the systematic underperformance of initial public offerings (IPOs) in the 3-5 years following the IPO (Loughran and Ritter, 1995). According to this interpretation, managers of private companies go public when the shares of their companies are over-priced, hence the underperformance of IPOs. Baker, Ruback, and Wurgler (2006) reviews the evidence supporting this theory. This theory complements the standard theory that issuance decisions respond to investment opportunities.

**Behavioral Political Economy.** Another setting in which we expect an asymmetry between rational and biased agents is politics. While politicians are experienced agents facing high-stake incentives and significant competition, voters make infrequent low-stake decisions—whether to vote and for whom. Therefore, we expect political settings to be well-described by the interaction of rational politicians and voters with non-standard preferences, such as imperfect memory and limited attention.

Eisensee and Stromberg (2007) provides an example of politicians responding to a bias of voters, inattention. They consider the decision by US ambassadors to release US aid in the days following a natural disaster in the country. Ambassadors presumably are more likely to release aid if they, or the government, get credit for their generosity. To capture this phenomenon,

Eisensee and Stromberg exploit variation in voter inattention due to the presence of major news items in the US television evening news, or due to major sporting event like the Olympics. They find that the probability of USAID relief is 15 percent lower for disasters occurring on days with a 2 standard deviation higher intensity of news in the US media. Similarly, the probability of relief is 30 percent lower in the period of the Olympics. On days in which the American public is less likely to notice the US generosity, generous acts are less likely to take place. This is consistent with politician response to limited attention of voters.

**Behavioral Institutional Design.** While firms, investors, managers, and politicians may respond to biases by exploiting them, the response to biases need not be predatory. Societal rules and institutions can be designed to counter-act the effect of consumer biases and improve the welfare of consumers. Benartzi and Thaler (2004)'s Save More Tomorrow (SMarT) plan is an example of one such institutional design for 401(k) savings. In a SMarT plan, the contribution rate is set to increase at each future wage increase up to a capped level. While savings increases are the default, employees can opt out of the plan at any time. This plan is an attractive commitment device to individuals with self-control problems, since the default applies to future savings rates, rather than current ones. In addition, the plan is designed with an eye to individuals that are averse to nominal wage cuts (see above), since the increases in contribution rates occur at the time of pay increases.

Benartzi and Thaler (2004) provide evidence on three implementations of this plan. In the earliest implementation, the plan is offered to 207 employees that accept to meet with a financial consultant, but do *not* accept to increase the savings rate immediately, as recommended by the consultant. Of these 207 individuals, 162 individuals accept the SMarT plan, indicating a wide-spread demand for commitment. In this subset of 162 individuals, the contribution rate increases from 3.5 percent to 13.6 percent in just four years. This increase includes the 32 individuals who opted out of the plan by the fourth year. The early results from the other two implementations of the SMarT plan indicate that the take-up of the plan is lower if it is offered as an option via mail, as opposed to with an in-person meeting. The effects conditional on take-up are, however, similarly large. These results suggest that a simple change in defaults can go a long way toward addressing under-saving. Importantly, while this plan is designed for individuals with self-control problems, it does not hurt individuals with time-consistent preferences, since these individuals can switch at any time.

While the evidence in Psychology and Economics can have policy implications, such as in this case, other considerations suggest caution regarding the policy reach of this evidence. First, unlike in the Benartzi and Thaler (2004) case, welfare-enhancing policies can be impractical—for example, no default can help people exercise more. Second, political economy considerations suggest caution in the implementation of policies (Glaeser, 2006). Nevertheless, behavioral phenomena should be taken into account alongside standard phenomena in the policy design.

## 6 Empirical Methods

The empirical research in Psychology and Economics discussed in this paper falls into five groups: Menu Choice, Natural Experiments, Field Experiments, Correlational Studies, and Structural Identification. Since these methods are broadly used in economics, I discuss the specifics of their application to Psychology and Economics.

**1. Menu Choice.** Assume that we observe both the consumer choice from a menu  $\{X_1, X_2, \dots, X_I\}$ , and an outcome  $x_i$  subsequent to the choice of  $X_i$ . The  $X_i$ s could be the contracts offered by a health club, and  $x_i$  the attendance to the club, conditional on the choice of contract  $X_i$  (DellaVigna and Malmendier, 2006). In another example, the  $X_i$ s could be different levels of deductible in home insurance, and  $x_i$  the ensuing number of accidents to the home, conditional on the deductible chosen (Sydnor, 2006).

In these examples, the standard theory makes a prediction about the outcome  $x_i$ , conditional on the choice of  $X_i$ . If an individual chooses  $X_i$ , according to the standard theory it must be the case that

$$Eg(x_i) |_{X_i} \geq \bar{g} \quad (11)$$

for some (usually linear) function  $g$  and some known threshold  $\bar{g}$ . (The direction of the inequality is immaterial) If (11) does not hold, the individual should not have chosen  $X_i$ .

If we observe the outcome  $x_i^n$  for a large number  $N$  of individuals  $n = 1, \dots, N$  choosing  $X_i$ , we can compute  $\sum_{n=1}^N g(x_i^n)/n$ , the sample mean of  $g(x_i)$  for individuals choosing  $X_i$ . Assuming that  $\sum_{n=1}^N g(x_i^n)/n$  converges to  $Eg(x_i) |_{X_i}$ ,  $\sum_{n=1}^N g(x_i^n)/n \geq \bar{g}$  must hold. If, instead,  $\sum_{n=1}^N g(x_i^n)/n < \bar{g}$  holds, the data rejects the standard model, and alternative models can be explored, in our examples models from Psychology and Economics. In the example of health club memberships, as I discussed in Section 2.1, under the standard model individuals choosing the monthly contract  $X_M$  with monthly fee  $L$  must believe that  $pE(x_M) |_{X_M} \geq L$ , or  $L/E(x_M) |_{X_M} \leq p$ : the price per expected attendance under the monthly contract should be lower than the fee  $p$  under payment-per-usage.

A strength of the menu-based approach is its simplicity: it relies on a basic principle of economics, revealed preferences, and the comparison of revealed preferences with observed behavior. If the two are inconsistent according to the standard theory, the data supports an alternative theory. A weakness of the approach is the low power of the test. If the test does not lead to rejection of the standard theory, it does not imply rejection of the alternative (behavioral) theory since condition (11) is necessary but not sufficient for the choice of  $X_i$ . An example is Agarwal et al. (2005), which cannot reject rational expectations about credit card borrowing, but is also consistent with sizeable over-/under-estimation of borrowing.

**2. Natural Experiments.** Natural Experiments evaluate the comparative statics of a model using naturally-occurring variation. Consider two situations, treatment situation  $T$  and control situation  $C$ , with respective outcomes  $x_T$  and  $x_C$ . Assume that the standard model

makes the prediction

$$Eg(x_T) \geq Eg(x_C)$$

for some function  $g$ ; the direction of the inequality is by convenience. The alternative, behavioral model predicts

$$Eg(x_T) < Eg(x_C).$$

By computing the sample equivalents of  $Eg(x_T)$  and  $Eg(x_C)$ , we can test the standard theory against the behavioral theory.

The simplest type of Natural Experiment is the **Time Series or Event Study**, in which the variable  $x$  is observed before the change of a situation (the control) and after (the treatment). An example is Madrian and Shea (2001), in which the contribution to 401(k)s  $x$  is observed before and after a default change. A **Difference-in-Difference Study** compares the change before and after the treatment to the change over the same time period for a control (placebo) group. An example is DellaVigna and Kaplan (2007) who compare the change in voting for towns where Fox News is introduced to the change in voting for control towns. The placebo group allows a control for common time trends.

An advantage of Natural Experiments is that they identify a treatment effect occurring in the field, with a guarantee of high external validity. A disadvantage is the possibility that the treatment may be endogenous and correlated with omitted variables that contaminate the causal inference. In principle, the change in default in Madrian and Shea (2001) could be due to the demand from new employees of more generous retirement plans; similarly, the introduction of Fox News could be correlated to political trends. These studies need to present evidence on the selection into the treatment group.

**3. Field Experiments.** Field Experiments (Harrison and List, 2004), like Natural Experiments, evaluate the comparative statics of the standard model and of a behavioral model. The difference is that in Field Experiments the treatment and control group are determined via an explicit randomization. An example is conducted by List (2003), who studies the endowment effect assigning sport cards to randomly-determined groups of card traders. The explicit randomization is an advantage of this approach since it guarantees internal validity, that is, the conditions for causal inference. This advantage comes sometimes at the cost of lower external validity, since in some important markets it is difficult to run Field Experiments. In addition, the sample size is often limited by the cost of the randomization.

Field Experiments, as well as Natural Experiments, in Psychology and Economics use two types of identification strategies. The first is to study environments in which the standard theory predicts there should be no effect of the treatment, while an alternative behavioral theory does. Examples are Ausubel (1999) and Bertrand et al. (2006) on credit card and loan offers among the Field Experiments, and Huberman and Regev (2001) on inattention in financial markets among the Natural Experiments. The second strategy is to consider treat-

ments that should have a sizeable effect according to the standard theory, but not according to the alternative behavioral theory. An example is Lee and McCrary (2005), who show that crime rates are unaffected by the sharp change in punishment for offenders occurring at the 18th birthday—a finding indicative of myopia. (Incidentally, this study relies on a **Regression Discontinuity** design, a variant of Natural Experiments)

**4. Correlational Studies.** Correlational Studies identify a correlation between two variables, say,  $x$  and  $y$ . Assume that the standard theory makes the prediction

$$\text{Cov}(x, y) \geq 0$$

(again, the direction of the inequality is by convenience), while a behavioral theory makes the prediction

$$\text{Cov}(x, y) < 0.$$

An example is Camerer et al. (2001) where  $x$  and  $y$  are hours worked and daily wage of cab drivers. A common application in Psychology and Economics is to the study of inter-personal psychological types. These studies use a proxy for a psychological trait (the  $x$  variable), such as overconfidence or impatience, and analyze the correlation of this proxy with a behavior (the  $y$  variable).<sup>29</sup> For example, Malmendier and Tate (2005 and forthcoming) correlate late exercise of stock option, a proxy of CEO overconfidence, with investment and merger activity.

An advantage of Correlational Studies is that they do not require a special design, as experiments do. The lack of an experimental design, however, complicates the inference and lowers the internal validity of these studies, since the observed covariance may be due to alternative explanations. For example, as I discussed above, fluctuations in cab driver wage can be driven by shifts in disutility of effort rather than shifts in demand, as assumed by Camerer et al. (2001). Similarly, the personality proxy may be correlated with unobserved variables, such as private information of the CEO. As a result, these papers need to undertake additional empirical tests to address the alternative explanations.

**5. Structural Identification.** While the above methods are used to qualitatively test the standard theory, they are typically not designed to provide point estimates for the parameters. Studies with Structural Identification are designed to draw this type of inference and quantify the extent of the non-standard preferences or biases. These papers estimate a fully-specified model. Examples are studies estimating the  $(\beta, \delta)$  model of self-control, such as Laibson, Repetto, and Tobacman (2006) on life-cycle accumulation and Paserman (forthcoming) on job search.

An advantage of this method is that the parameter estimates can be used for welfare and policy evaluations. Paserman (forthcoming), for example, simulates the effect of labor market

---

<sup>29</sup>Interestingly, the studies that emphasize personality types run counter to a tenet of social psychology, that the situation appears to explain behavior more than the personality (Ross and Nisbett, 1991).

policies for individuals with the extent of self-control problems estimated in the data. This evaluation is hard to perform without parameter estimates.

While the reliance on a model permits quantitative inferences and welfare evaluations, it often comes at the cost of reduced transparency of the results. The parameter estimates depend on the full set of assumptions. For this reason, these papers are often accompanied by reduced-form results that provide intuition. Laibson, Repetto, and Tobacman (2006), for example, builds on Angeletos et al. (2001). Conlin, O'Donoghue, and Vogelsang (forthcoming) on projection bias provides both reduced-form and structural estimates.

A middle ground between reduced-form results and Structural Identification are calibrations, in which quantitative predictions from a model are compared to the data for a feasible range of parameter values. In examples such as Benartzi and Thaler (1995) and Sydnor (2006), calibrations play two key roles: (i) showing that the standard models can explain a finding only for implausible parameter values, in these two cases implausibly high risk aversion; (ii) showing that behavioral models match a finding for parameter values that are consistent with the experimental evidence, in these cases on reference-dependent preferences.

## 7 Conclusion

In this survey, I summarized the field evidence on three classes of deviations from the standard model: non-standard preferences, non-standard beliefs, and non-standard decision-making. I discussed how rational agents in the market respond to these non-standard features. I concluded with a summary of the empirical methodologies used in this research. As this survey documented, deviations from the standard model are not confined to laboratory decisions. Most phenomena that are important in laboratory experiments also affect decisions in a variety of economic settings. Hence, I expect that economists will increasingly take behavioral phenomena into account in their analysis.

Why don't market forces eliminate non-standard behavior? While a full discussion of this objection is beyond the scope of this article, I address two related arguments, one on experience and another on aggregation. A first argument is that experience reduces non-standard behavior. Indeed, experience appears to mitigate the endowment effect (List, 2003 and 2004). Palacios-Huerta and Volji (2007) provide concordant evidence on the effect of experience on the ability to perform backward induction. They consider the centipede game. Chess players, who have to routinely perform backward induction-type reasoning, come close in their play to the predictions of backward induction, in sharp contrast to college students.

However, it would be wrong to conclude, based on this evidence, that behavioral phenomena should not matter in the field. I list four reasons. (i) In a number of economic decisions, feedback is infrequent (such as in house purchases) or noisy (such as in financial investments), and hence most individuals are inexperienced. (ii) Experience can exacerbate a bias if indi-

viduals are not Bayesian learners. Haigh and List (2004) use a simple investment game and show that professional investors display significantly more myopic loss aversion (see Section 2.2) than students. Presumably, the short-term incentives in the workplace teach these investors to frame problems narrowly, contrary to the prediction of the standard theory. (iii) In principle, debiasing by experienced agents can be a substitute for direct experience. However, as Gabaix and Laibson (2006) show, experienced agents such as firms typically have little or no incentive to debias individuals. (iv) Finally, not all non-standard features should be mitigated by experience. Experience should not affect social preferences any more than it should affect preferences for the characteristics of cars.

A second argument is that, even if experience or debiasing do not eliminate the biases, the biases will not affect aggregate market outcomes. The argument is made forcefully in financial markets: given arbitrage, the rational investors set prices. However, as we discussed, the limits to arbitrage (DeLong et al., 1991) imply that individuals with non-standard features will in general affect stock prices. In addition, in most settings, there is no plausible incentive to eliminate a bias and hence the effect of non-standard behavior aggregates linearly. If a share of the population procrastinates saving for retirement, the aggregate savings rate will reflect proportionally the under-saving by this group. This is true unless a different institutional design is put in place, such as the SMarT plan (Benartzi and Thaler, 2004). (Notice that this plan was put in place not by market forces, but by academics).

Finally, the papers on behavioral IO indicate that the non-standard features, instead of having no impact, can in fact have a disproportionate impact on market outcomes. Lee and Malmendier (2007) provide a telling example regarding overbidding in eBay auctions. Lee and Malmendier define a case of overbidding when the final auction price is higher than a posted price for the same good available on eBay itself. They focus on an item for which the posted price is essentially always available and is stable, and hence should be an upper bound for the bids in a rational model. The authors show that 42 percent of auctions end at a price above the posted price, a conclusion robust to the inclusion of shipping costs, to differences in item quality and in seller reputation. The key aggregation point is that this behavior is generated by many fewer than 42 percent of overbidders. In fact, only 17 percent of bidders ever overbid. The auction design, however, is such that the overbidders determine the final price.

To conclude, a natural question is what empirical research in Psychology and Economics will look like in the future. Methodologically, I expect future research to continue using mostly the methods encountered in this overview, field experiments (such as List, 2003 and Falk, forthcoming), natural experiments (such as Madrian and Shea, 2001 and DellaVigna and Kaplan, 2007), and inference from menu choice (such as DellaVigna and Malmendier, 2006 and Sydnor, 2006). Studies that provide structural estimates of the parameters (such as Laibson et al., 2006 and Conlin et al., forthcoming) will need to address a number of open questions. For example, can models of  $(\beta, \delta)$  preferences predict choice in different decisions for fixed

parameters  $\beta$ ,  $\hat{\beta}$ , and  $\delta$ ? The evidence from the few existing papers suggests that this may be the case, but more evidence is in order. This estimation would benefit from the availability of data sets with multiple decisions by the same individual. While individuals are likely to differ in their preferences and beliefs, we expect the same individual to behave consistently if the existing models capture the behavior accurately.<sup>30</sup> It is also possible that new, more parsimonious models of the phenomena presented in this survey will emerge, as Fudenberg (2006) predicts.

As for the topics, future research is likely to reduce the imbalance across fields in economics and across topics in psychology. While the research in behavioral finance and consumption-savings is very active, relatively few studies, instead, have tackled mortgage markets, development, and political decisions, fields ripe for exploration. Future research is also likely to explore psychological phenomena that have been largely neglected. For example, emotions, automatic processing, and implicit discrimination are likely to matter for economic decisions such as divorce, judicial sentencing, and policing. Ten years from now, we will hopefully be able to assess quantitatively which psychological factors matter in which decisions.

I identify two specific areas for future research: the market interaction between standard and non-standard agents, as in Section 5, and public policy applications. The market interaction is likely to find several additional applications, for example to the interaction between politicians and voters. In addition, this area is likely to investigate the judgmental biases, such as overconfidence, of experienced agents such as managers and politicians. The area of public policy is a recent application of the research in Psychology and Economics, mainly in the context of retirement decisions. The 2006 Congress enacted a bill on *Automatic Savings and Pension Protection Act* that was motivated by the research on defaults and on the SMarT plan. This law gives incentives to companies to adopt 401(k) plans with automatic enrollment and automatic increases in savings. Future research will tell whether this is an isolated application of Psychology and Economics or the first of several.

---

<sup>30</sup>In a laboratory experiment, Fisman, Kariv, and Markovits (forthcoming) use repeated decisions on giving to another subject to identify types of subjects with different social preferences. Their results suggest substantial heterogeneity.

## References

- [1] Abadie, Alberto, and Sebastien Gay. 2006. “The Impact of Presumed Consent Legislation on Cadaveric Organ Donation: A Cross-Country Study” *Journal of Health Economics*, 25(4): 599–620.
- [2] Agarwal, Sumit, Souphala Chomisengphet, Chunlin Liu, and Nicholas S. Souleles. 2005 “Do Consumers Choose the Right Credit Contracts?” Working Paper.
- [3] Akerlof, George A. 1982. “Labor Contracts as Partial Gift Exchange.” *Quarterly Journal of Economics*, 97(4): 543–569.
- [4] Akerlof, George A. 1991. ”Procrastination and Obedience”, *American Economic Review*, 81(2): 1-19.
- [5] Alpert, M., and H. Raiffa. 1982. “A Progress Report on the Training of Probability Assessors.” In *Judgment Under Uncertainty: Heuristics and Biases*, ed. D. Kahneman, P. Slovic and A. Tversky, 294–305. Cambridge: Cambridge University Press.
- [6] Andreoni, James. 2006. “Philantropy.”In *Handbook of Giving, Altruism, and Reciprocity*, ed. Serge-Christophe Kolm and Jean Mercier Ythier, 1201–1265. Amsterdam: Elsevier/North-Holland.
- [7] Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg. 2001. “The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation.” *Journal of Economic Perspectives*, 15(3): 47–68.
- [8] Ariely, Dan and George Loewenstein. 2005. “The Heat of the Moment: The Effect of Sexual Arousal on Sexual Decision Making”, *Journal of Behavioral Decision Making*, 18(1), 1–12.
- [9] Ariely, Dan and Klaus Wertenbroch. 2002. “Procrastination, Deadlines, and Performance: Self-control by Precommitment.” *Psychological Science*, 13(3): 219-224.
- [10] Asch, Solomon E. 1951. “Effects of Group Pressure Upon the Modification and Distortion of Judgment.” In H. Guetzkow (Ed.), *Groups, leadership, and men*. Pittsburgh: Carnegie Press.
- [11] Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. “Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines.” *Quarterly Journal of Economics*, 121(2): .
- [12] Ausubel, Lawrence. 1999. “Adverse Selection in the Credit Card Market.” Working Paper, University of Maryland.

- [13] Baker, Malcolm, Richard Ruback, and Jeffrey Wurgler. 2006. “Behavioral Corporate Finance: A Survey.” In *The Handbook of Corporate Finance: Empirical Corporate Finance*, ed. Espen Eckbo. New York: Elsevier/North Holland.
- [14] Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2005. “Social Preference and The Response To Incentives: Evidence From Personnel Data”, *Quarterly Journal of Economics*, 120(3): 917-962.
- [15] Barber, Brad and Terrance Odean. 2001. “Boys will be Boys: Gender, Overconfidence, and Common Stock Investment,” *Quarterly Journal of Economics*, 116(1), 261-292.
- [16] Barber, Brad and Terrance Odean. Forthcoming. “All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors.” *Review of Financial Studies*.
- [17] Barberis, Nicholas, Ming Huang and Tano Santos. 2001. “Prospect Theory and Asset Prices”, *Quarterly Journal of Economics*, 116(1): 1-53.
- [18] Barberis, Nicholas, Andrei Shleifer and Robert Vishny. 1998. “A model of investor sentiment”, *Journal of Financial Economics*, 49: 307-345.
- [19] Barberis, Nicholas and Richard Thaler. 2003. “A Survey of Behavioral Finance.” in *Handbook of the Economics of Finance*.
- [20] Barberis, Nicholas and Wei Xiong. 2006. “What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation.” Working Paper.
- [21] Benartzi, Shlomo. 2001. “Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock” *Journal of Finance*, 56(5):1747-1764.
- [22] Benartzi, Shlomo and Richard H. Thaler. 1995. “Myopic Loss Aversion and the Equity Premium Puzzle.” *Quarterly Journal of Economics* 100, 73-92.
- [23] Benartzi, Shlomo and Richard H. Thaler. 2001. “Naive Diversification Strategies in Defined Contribution Saving Plans.” *American Economic Review*, 91(1): 79-98.
- [24] Benartzi, Shlomo and Richard Thaler. 2004. “Save More Tomorrow (TM): Using Behavioral Economics to Increase Employee Saving.” *Journal of Political Economy*, 112(S1): S164-S187.
- [25] Bernard, Victor L. and Jacob K. Thomas. 1989. “Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?”, *Journal of Accounting Research*, 27: 1-36.

- [26] Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman. 2006. “What’s Psychology Worth? A Field Experiment in the Consumer Credit Market.” Working paper.
- [27] Bewley, Truman F. 1999. *Why Wages Don’t Fall During a Recession*, Harvard University Press.
- [28] Broadbent, Donald E. 1958. *Perception and communication*, Pergamon Press, New York, NY.
- [29] Buehler, R., Dale Griffin, and M. Ross. 1994. “Exploring the ‘Planning Fallacy’: Why People Underestimate their Task Completion Times.” *Journal of Personality and Social Psychology*, 67: 366-381.
- [30] Cain, Daylain, George Loewenstein, and Don Moore. 2005. “The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest,” *Journal of Legal Studies*, 34, 1-25.
- [31] Camerer, Colin. 2005. “Behavioral Economics” *World Congress of the Econometric Society*, London, 18-24.
- [32] Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. “Labor Supply of New York City Cab Drivers: One Day At a Time”, *Quarterly Journal of Economics*, 112(2), 407-441.
- [33] Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O’Donoghue, and Matthew Rabin. 2003. “Regulation For Conservatives: Behavioral Economics And The Case For ‘Asymmetric Paternalism’” *University of Pennsylvania Law Review*, 151, 1211–1254.
- [34] Camerer, Colin, and Lovallo. 1999. “Overconfidence and Excess Entry: An Experimental Approach,” *American Economic Review*, 89(1): 306-318.
- [35] Campbell, John Y., and John Cochrane. 1999, “By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior”, *Journal of Political Economy*. 107, 205-251.
- [36] Card, David. 1994 “Intertemporal Labor Supply: An Assessment,” in C. Sims, ed., *Advances in Econometrics: Proceedings of the Econometric Society Sixth World Congress* (New York: Cambridge University Press).
- [37] Card, David, and Dean Hyslop. 1997. “Does inflation grease the wheels of the labor market?” in Christina D. Romer and David H. Romer (eds.), *Reducing Inflation: Motivation and Strategy. NBER, Studies in Business Cycles*, 30. Chicago: University of Chicago Press.

- [38] Charness, Gary, and Rabin, Matthew. 2002. “Understanding Social Preferences with Simple Tests.” *Quarterly Journal of Economics* 117(3), 817-869.
- [39] Chetty, Raj, Adam Looney, and Kory Kroft. 2007. “Salience and Taxation: Theory and Evidence”, Working Paper.
- [40] Choi, James J., David Laibson, and Brigitte C. Madrian. 2005. “\$100 Bills on the Sidewalk: Suboptimal Saving in 401(k) Plans” Working Paper.
- [41] Choi, James J., David Laibson, and Brigitte C. Madrian. 2006. “Reducing The Complexity Costs of 401(k) Participation Through Quick Enrollment” NBER Working Paper No. 11979.
- [42] Choi, James J., David Laibson, Brigitte C. Madrian and Andrew Metrick. 2004. “For Better or For Worse: Default Effects and 401(k) Savings Behavior,” in David Wise editor *Perspectives in the Economics of Aging*. Chicago, IL: University of Chicago Press, 81-121.
- [43] Choi, James J., David Laibson, Brigitte Madrian, and Andrew Metrick. 2005. “Optimal Defaults and Active Decisions,” NBER Working Paper.
- [44] Cicchetti, Charles J and Jeffrey A. Dubin. 1994. “A Microeconometric Analysis of Risk Aversion and the Decision to Self-Insure.” *The Journal of Political Economy*, 102(1): 169-186.
- [45] Cohen, Lauren, and Andrea Frazzini. Forthcoming. “Economic Links and Predictable Returns”, *Journal of Finance*.
- [46] Conlin, Michael, Ted O’Donoghue, and Timothy J. Vogelsang. Forthcoming. “Projection Bias in Catalog Orders.” *American Economic Review*.
- [47] Conlisk, John. 1996. “Why Bounded Rationality?”, *Journal of Economic Literature*, 34(2): 669–700.
- [48] Cronqvist, Henrik, and Richard H. Thaler. 2004 “Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience.” *American Economic Review Papers and Proceedings*, 94(2): 424–428.
- [49] Dahl, Gordon and Stefano DellaVigna. 2007. “Does Movie Violence Increase Violent Crime?”, Working Paper.
- [50] Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam. 1998. “Investor psychology and security market under- and over-reactions”, *Journal of Finance*, 53: 1839–1886.

- [51] DellaVigna, Stefano and M. Daniele Paserman. 2005. "Job Search and Impatience." *Journal of Labor Economics*, 23(3): 437-466.
- [52] DellaVigna, Stefano and Ulrike Malmendier. 2004. "Contract Design and Self-Control: Theory and Evidence." *Quarterly Journal of Economics*, 119: 353-402.
- [53] DellaVigna, Stefano and Ulrike Malmendier. 2006. "Paying Not to Go to the Gym." *American Economic Review*, 96(3): 694-719.
- [54] DellaVigna, Stefano and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *Quarterly Journal of Economics*, 122(3): 1187-1234.
- [55] DellaVigna, Stefano and Joshua Pollet. Forthcoming. "Demographics and Industry Returns" *American Economic Review*.
- [56] DellaVigna, Stefano and Joshua Pollet. 2006. "Investor Inattention, Firm Reaction, and Friday Earnings Announcements" NBER Working Paper No. 11683.
- [57] DeLong, Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. "Noise Trader Risk in Financial Markets." *Journal of Political Economy* 98(4): 703-738.
- [58] Diamond, Peter and Hannu Vartiainen. 2007. *Behavioral Economics and Its Applications*, Princeton University Press.
- [59] Dufwenberg, Martin and Kirchsteiger, Georg. 2004. "A Theory of Sequential Reciprocity", *Games and Economic Behavior*. 47, 268-98.
- [60] Edmans, Alex, Diego Garcia, and Oyvind Norli. 2007. "Sports Sentiment and Stock Returns" *Journal of Finance*, 62(4).
- [61] Eisensee, Thomas and David Strömberg. 2007. "News Floods, News Droughts, and U.S. Disaster Relief" *Quarterly Journal of Economics*, 122(2).
- [62] Eliaz, Kfir, and Ran Spiegler. 2006. "Contracting with Diversely Naive Agents" *Review of Economic Studies*, 73(3): 689–714.
- [63] Ellison, Glenn. 2006. "Bounded Rationality in Industrial Organization" in Blundell, Newey and Persson (eds.), *Advances in Economics and Econometrics: Theory and Applications*, Ninth World Congress, Cambridge University Press.
- [64] Ellsberg, Daniel. 1961. "Risk, Ambiguity, and the Savage Axioms", *Quarterly Journal of Economics*, 75: 643-69.
- [65] Falk, Armin. Forthcoming. "Gift Exchange in the Field." *Econometrica*.

- [66] Falk, Armin, and Andrea Ichino. 2006. “Clean Evidence of Peer Effects,” *Journal of Labor Economics*, 24(1): 39-57.
- [67] Farber, Henry S. 2005. “Is Tomorrow Another Day? The Labor Supply of New York City Cab Drivers,” *Journal of Political Economy*, 113(1): 46-82.
- [68] Farber, Henry S. 2006. “Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers.” Working Papers.
- [69] Fehr, Ernst and Simon Gächter, 2000. “Fairness and Retaliation - The Economics of Reciprocity”, *Journal of Economic Perspectives*, 14, 159-181.
- [70] Fehr, Ernst and Lorenz Goette. 2007. “Do Workers Work More if Wages are High? Evidence from a Randomized Field Experiment.” *American Economic Review*, 97(1).
- [71] Fehr, Ernst, Georg Kirchsteiger and A. Riedl. 1993. “Does Fairness prevent Market Clearing? An Experimental Investigation”, *Quarterly Journal of Economics*, 108(2): 437-460.
- [72] Fehr, Ernst, and Schmidt, Klaus. 1999. “A Theory of Fairness, Competition and Cooperation,” *Quarterly Journal of Economics* 114(3): 817-868.
- [73] Fisman, Raymond, Shachar Kariv, and Daniel Markovits. Forthcoming. “Individual Preferences for Giving”, *American Economic Review*.
- [74] Forsythe, Robert, Joel L. Horowitz, N.E. Savin, and Martin Sefton. 1994. “Fairness in Simple Bargaining Experiments” *Games and Economic Behavior*, 6(3): 347–369.
- [75] Frederick, Shane, George Loewenstein, and Ted O’Donoghue. 2002. “Time Discounting and Time Preference: A Critical Review.” *Journal of Economic Literature*, 40(2): 351-401.
- [76] French, Kenneth R. and Poterba, James M. 1991. “Investor Diversification and International Equity Markets” *American Economic Review*, 81: 222-226.
- [77] Fudenberg, Drew. 2006. “Advancing Beyond *Advances in Behavioral Economics*,” *Journal of Economic Literature*, 44: 694-711.
- [78] Fudenberg, Drew and David Levine. 2006. “A Dual Self Model of Impulse Control.” *American Economic Review*, 96: 1449-1476.
- [79] Gabaix, Xavier, David Laibson, Guillermo Moloché and Stephen Weinberg. 2006. “Information Acquisition: Experimental Analysis of a Boundedly Rational Model”, *American Economic Review*, 96(4): 1043–1068.

- [80] Gabaix, Xavier, and David Laibson. 2006. “Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets.” *Quarterly Journal of Economics*, 121(2): 505–540.
- [81] Garicano, Luis, Ignacio Palacios-Huerta, and Canice Prendergast. 2005. “Favoritism Under Social Pressure,” *The Review of Economics and Statistics*, 87(2): 208-216.
- [82] Genesove, David and Christopher Mayer. 2001. “Loss Aversion and Seller Behavior: Evidence from the Housing Market,” *Quarterly Journal of Economics*, 116(4), 1233-1260.
- [83] Gilbert, Daniel T., Elizabeth C. Pinel, Timothy D. Wilson, Stephen J. Blumberg, and Thalia P. Wheatley. 1998. “Immune neglect: A source of durability bias in affective forecasting.” *Journal of Personality and Social Psychology*, 75(3): 617-638.
- [84] Glaeser, Edward. 2006. “Paternalism and Psychology” *The University of Chicago Law Review*, 73(1): 133-156.
- [85] Gneezy, Uri, and John List. 2006. “Putting Behavioral Economics To Work: Testing For Gift Exchange In Labor Markets Using Field Experiments”, *Econometrica*, 74(5), 1365–1384.
- [86] Gul, Faruk and Wolfgang Pesendorfer. 2001. “Temptation and Self-Control,” *Econometrica*, 69(6): 1403-1435.
- [87] Haigh, Michael, and John A. List. 2004. “Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis,” *Journal of Finance*, 60(1), 523-534.
- [88] Harrison, Glenn and John A. List. 2004. “Field Experiments,” *Journal of Economic Literature*, 42: 1009-1055.
- [89] Heidhues, Paul, and Botond Kőszegi. 2005. “The Impact of Consumer Loss Aversion on Pricing.” Working Paper.
- [90] Hirshleifer, David A., Sonya S. Lim, and Siew Hong Teoh. 2007. “Driven to Distraction: Extraneous Events and Underreaction to Earnings News” Working Paper.
- [91] Hirshleifer, David A. and Tyler Shumway. 2003. “Good Day Sunshine: Stock Returns and the Weather”, *Journal of Finance*, 58(3), 1009-1032.
- [92] Ho, Daniel E., and Kosuke Imai. 2004. “Estimating Causal Effects of Ballot Order from a Randomized Natural Experiment: California Alphabet Lottery, 1978-2002.” Working Paper.

- [93] Hong, Harrison and Jeremy C. Stein. 1999. “A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets”, *Journal of Finance*, 54(6): 2143-2184.
- [94] Hossain, Tanjim and John Morgan. 2006. “...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay,” *Advances in Economic Analysis & Policy*, 6(2): 1429-1429.
- [95] Huberman, Gur. 2001. “Familiarity Breeds Investment,” *Review of Financial Studies*, 14(3): 659-680.
- [96] Huberman, Gur, and Tomer Regev. 2001. “Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar.” *Journal of Finance*, 56(1): 387-396.
- [97] Huberman, Gur, and Wei Jiang. 2006. “Offering vs. Choice in 401(k) Plans: Equity Exposure and Number of Funds.” *Journal of Finance*, 61(2): 763-801.
- [98] Ivkovich, Zoran, James Poterba and Scott Weisbenner. 2005 “Tax-Motivated Trading by Individual Investors.” *American Economic Review*, 95(5): 1605–1630.
- [99] Iyengar, Sheena and Mark Lepper. 2000. “When Choice is Demotivating: Can One Desire Too Much of a Good Thing?” *Journal of Personality and Social Psychology*, 79(6): 995-1006.
- [100] Iyengar, Sheena S., Gur Huberman, and Wei Jiang. 2004. “How Much Choice Is Too Much? Contributions to 401(k) Retirement Plans,” in *Pension Design and Structure: New Lessons from Behavioral Finance*, ed. by Olivia S. Mitchell, and Stephen P. Utkus, Oxford University Press.
- [101] Johnson, Eric J., John Hershey, Jacqueline Msezaros, and Howard Kuhnreuther. 1993. “Framing, Probability Distortions, and Insurance Decisions.” *Journal of Risk and Uncertainty*, 7(1): 35–51.
- [102] Josephson, Wendy L. 1987. “Television Violence and Children’s Aggression: Testing the Priming, Social Script, and Disinhibition Predictions.” *Journal of Personality and Social Psychology*, 53, 882–890.
- [103] Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1986. “Fairness and the assumptions of economics.” *Journal of Business*, 59(4): S285-S300.
- [104] Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1990. “Experimental Tests of the Endowment Effect and the Coase Theorem.” *Journal of Political Economy*, 98(6): 1325-1348.

- [105] Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk" *Econometrica*, 47(2): 263-292.
- [106] Kim, E. Han, Adair Morse and Luigi Zingales. 2006. "What Has Mattered To Economics Since 1970", *Journal of Economic Perspectives*, 20(4): 189-202.
- [107] Kőszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences" *Quarterly Journal of Economics*, 121(4): 1133-1166.
- [108] Krueger, Alan and Alexandre Mas. 2004. "Strikes, Scabs and Tire Separations: Labor Strife and the Production of Defective Bridgestone/Firestone Tires," *Journal of Political Economy*, 112(2): 253-89.
- [109] Kube, Sebastian, Michel André Maréchal, and Clemens Puppe. 2006. "Putting Reciprocity to Work – Positive versus Negative Responses in the Field." Working Paper.
- [110] Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, 112(2): 443-477.
- [111] Laibson, David, Andrea Repetto, and Jeremy Tobacman. 2006. "Estimating Discount Functions with Consumption Choices over the Lifecycle," Working paper.
- [112] Lee, David S. and Justin McCrary. 2005. "Crime, Punishment, and Myopia" National Bureau of Economic Research Working Paper 11491.
- [113] Lee, Han and Ulrike Malmendier. 2007. "The Bidder's Curse", Working Paper.
- [114] Levitt, Steven and John A. List. 2007. "What Do Laboratory Experiments Measuring Social Preferences Tell Us About The Real World?" *Journal of Economic Perspectives*, 21(2).
- [115] List, John A. and David Lucking-Reiley. 2002. "The Effects of Seed Money and Refunds on Charitable Giving: Experimental Evidence from a University Capital Campaign." *Journal of Political Economy*, 110(1): 215-233.
- [116] List, John A. 2003. "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics*, 118(1): 41-71.
- [117] List, John A. 2004. "Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace," *Econometrica*, 72(2): 615-625.
- [118] List, John A. 2006. "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions", *Journal of Political Economy*, 114(1): 1-37.

- [119] Loewenstein, George and Jennifer Lerner. 2003. "The Role of Affect in Decision Making," in *Handbook of Affective Science*, eds. Richard J. Dawson, Klaus R. Scherer and H. Hill Goldsmith, Oxford: Oxford University Press, 619-42.
- [120] Loewenstein, George, Ted O'Donoghue, and Matthew Rabin. 2003. "Projection Bias In Predicting Future Utility," *Quarterly Journal of Economics*, 118(4): 1209-1248.
- [121] Loewenstein, George and Drazen Prelec. 1992. "Anomalies in Intertemporal Choice: Evidence and an Interpretation." *Quarterly Journal of Economics*, 107(2): 573-597.
- [122] Loughran, Tim and Jay R. Ritter. 1995. "The New Issues Puzzle." *Journal of Finance*, 50(1): 23-51.
- [123] Madrian, Brigitte C. and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics*, 116(4): 1149-1187.
- [124] Malmendier, Ulrike, and Devin Shanthikumar. 2007. "Are Small Investors Naive about Incentives?", *Journal of Financial Economics*, 85(2), 457-489.
- [125] Malmendier, Ulrike, and Geoffrey Tate. 2005. "CEO Overconfidence and Corporate Investment." *Journal of Finance*, 60 (6): 2661-2700.
- [126] Malmendier, Ulrike, and Geoffrey Tate. Forthcoming. "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction," *Journal of Financial Economics*.
- [127] Mankiw, N. Gregory and Ricardo Reis. 2002. "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve", *Quarterly Journal of Economics*, 117(4): 1295-1328.
- [128] Mas, Alexander. 2006. "Pay, Reference Points, And Police Performance", *Quarterly Journal of Economics*, 121(3): 783-821.
- [129] Mas, Alexander and Moretti, Enrico. 2006. "Peers At Work" Working Paper.
- [130] McClure, Samuel M., David Laibson, George Loewenstein and Jonathan D. Cohen. 2004. "Separate Neural Systems Value Immediate and Delayed Monetary Rewards." *Science*, 306.
- [131] Mehra, R., and E. Prescott. 1985, "The equity premium: a puzzle", *Journal of Monetary Economics*, 15: 145-161.
- [132] Milgram, Stanley. 1963. "Behavioral study of obedience." *Journal of Abnormal and Social Psychology*, 67: 371-378.

- [133] Mullainathan, Sendhil and Richard H. Thaler. 2001. “Behavioral Economics” in *International Encyclopedia of Social Sciences*, Pergamon Press, 1094-1100.
- [134] Odean, Terrance. 1998. “Are Investors Reluctant to Realize Their Losses?” *Journal of Finance*, 53(5): 1775-1798.
- [135] Odean, Terrance. 1999. “Do investors trade too much?” *American Economic Review*, 89: 1279-1298.
- [136] O’Donoghue, Ted and Matthew Rabin. 1999a. “Doing It Now or Later,” *American Economic Review*, 89(1): 103-124.
- [137] O’Donoghue, Ted and Matthew Rabin. 1999b. “Procrastination in Preparing for Retirement”, in *Behavioral Dimensions of Retirement Economics*, Henry Aaron, editor, The Brookings Institution.
- [138] O’Donoghue, Ted and Matthew Rabin. 2001. “Choice and Procrastination,” *Quarterly Journal of Economics*, 116(1), 121-160.
- [139] Palacios-Huerta, Ignacio and Oscar Volji. 2007. “Field Centipedes”, Working paper.
- [140] Paserman, M. Daniele. Forthcoming. “Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation”, *Economic Journal*.
- [141] Phelps, E. and R. A. Pollak. 1968 “On Second-best National Saving and Game-equilibrium Growth”, *Review of Economic Studies*, 35: 185-199.
- [142] Plott, Charlie R., and Kathryn Zeiler. 2004 “The Willingness to Pay/Willingness to Accept Gap, the ‘Endowment Effect’ Subject Misconceptions and Experimental Procedures for Eliciting Valuations”, *American Economic Review*.
- [143] Pope, Devin. 2007. “Reacting to Rankings: Evidence from ’America’s Best Hospitals and Colleges””, Working Paper.
- [144] Rabin, Matthew. 1993. “Incorporating Fairness into Game Theory and Economics”, *American Economic Review*, 83, 1281-1302.
- [145] Rabin, Matthew. 1998. “Psychology and Economics.” *Journal of Economic Literature*, 36(1): 11-46.
- [146] Rabin, Matthew. 2000. “Risk Aversion and Expected-Utility Theory: A Calibration Theorem,” *Econometrica* 68(5), 1281-1292.
- [147] Rabin, Matthew. 2002a. “A Perspective on Psychology and Economics” *European Economic Review*, 46(4-5): 657-685.

- [148] Rabin, Matthew. 2002b. "Inference by Believers in the Law of Small Numbers." *Quarterly Journal of Economics*, 117(3): 775-816.
- [149] Rind, B. 1996. "Effect of beliefs about weather conditions on tipping". *Journal of Applied Social Psychology*, 26(2), 137–147.
- [150] Michael S. Rashes. 2001. "Massively Confused Investors Making Conspicuously Ignorant Choices (MCI-MCIC)." *Journal of Finance*, 56(5): 1911-1927.
- [151] Read, Daniel and Barbara van Leeuwen. 1998. "Predicting hunger. The Effects of Appetite and Delay on Choice." *Organizational Behavior and Human Decision Processes*, 76, 189-205.
- [152] Ross, L. and R. E. Nisbett. 1991. *The Person and the Situation: Perspectives on Social Psychology*. New York, McGraw-Hill.
- [153] Samuelson, William and Richard Zeckhauser. 1988. "Status-Quo Bias". *Journal of Risk and Uncertainty*.
- [154] Saunders, E. M. J., 1993. "Stock Prices and Wall Street Weather", *American Economic Review*, 83, 1337-1345.
- [155] Schwarz, Norbert, & Clore, G. L. 1983. "Mood, Misattribution, and Judgments of Well-being: Informative And Directive Functions of Affective States". *Journal of Personality and Social Psychology*, 45(3), 513–523.
- [156] Shefrin, Hersh M. and Richard H. Thaler. 1981. "An Economic Theory of Self-Control." *Journal of Political Economy*, 89(2), 392-406.
- [157] Shleifer, Andrei. 2000. *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford University Press.
- [158] Shue, Kelly and Erzo Luttmer. 2007. "Who Misvotes? The Effect of Differential Cognition Costs on Election Outcomes," NBER Working Paper No. 12709.
- [159] Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*, 69(1), 99–118.
- [160] Simonsohn, Uri. 2007. "Weather to Go To College", Working Paper.
- [161] Simonson, Itamar. 1990. "The Effect of Purchase Quantity and Timing on Variety-Seeking Behavior," *Journal of Marketing Research*, 27(2), 150–162.
- [162] Sims, Christopher. 2003. "Implications of Rational Inattention", *Journal of Monetary Economics*, 50(3), 665-690.

- [163] Strotz, R. H. 1956. "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies*, 23: 165–180.
- [164] Svenson, Ola. 1981. "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica* 47: 143-148.
- [165] Sydnor, Justin. 2006. "Sweating the Small Stuff: The Demand for Low Deductibles in Homeowner's Insurance." Working paper.
- [166] Terrell, Dek. 1994. "A Test of the Gambler's Fallacy – Evidence from Pari-Mutuel Games," *Journal of Risk and Uncertainty*, 8(3): 309-317.
- [167] Thaler, Richard H. 1981. "Some Empirical Evidence on Dynamic Inconsistency," *Economics Letters*, 8: 201-7.
- [168] Tversky, Amos and Daniel Kahneman. 1974. "Judgement under uncertainty: Heuristics and biases" *Science*, 125: 1124–1131.
- [169] Tversky, Amos, and Daniel Kahneman. 1992. "Advances in prospect theory: cumulative representation of uncertainty", *Journal of Risk and Uncertainty*, 5: 297-323.
- [170] Weinstein, Neil D. 1980. "Unrealistic Optimism about Future Life Events." *Journal of Personality and Social Psychology*, 39: 806-820.

*Chapter 6*

## INSTITUTIONS AS A FUNDAMENTAL CAUSE OF LONG-RUN GROWTH

DARON ACEMOGLU

*Department of Economics, MIT, 50 Memorial Drive E52-380b, Cambridge, MA 02142*  
*e-mail: [daron@mit.edu](mailto:daron@mit.edu)*

SIMON JOHNSON

*Sloan School of Management, MIT, 50 Memorial Drive, Cambridge, MA 02142*  
*e-mail: [sjohnson@mit.edu](mailto:sjohnson@mit.edu)*

JAMES A. ROBINSON

*Department of Government, WCFIA, Harvard University, 1033 Massachusetts Avenue, Cambridge, MA 02138*  
*e-mail: [jrobinson@gov.harvard.edu](mailto:jrobinson@gov.harvard.edu)*

### Contents

Abstract	386
Keywords	387
1. Introduction	388
1.1. The question	388
1.2. The argument	389
1.3. Outline	396
2. Fundamental causes of income differences	396
2.1. Three fundamental causes	397
2.1.1. Economic institutions	397
2.1.2. Geography	399
2.1.3. Culture	400
3. Institutions matter	402
3.1. The Korean experiment	404
3.2. The colonial experiment	407
4. The Reversal of Fortune	407
4.1. The reversal among the former colonies	408
4.2. Timing of the reversal	412
4.3. Interpreting the reversal	412
4.4. Economic institutions and the reversal	414

4.5. Understanding the colonial experience	416
4.6. Settlements, mortality and development	417
<b>5. Why do institutions differ?</b>	<b>421</b>
5.1. The efficient institutions view – the Political Coase Theorem	422
5.2. The ideology view	424
5.3. The incidental institutions view	425
5.4. The social conflict view	427
<b>6. Sources of inefficiencies</b>	<b>428</b>
6.1. Hold-up	430
6.2. Political losers	432
6.3. Economic losers	434
6.4. The inseparability of efficiency and distribution	436
6.5. Comparative statics	437
6.6. The colonial experience in light of the comparative statics	438
6.7. Reassessment of the social conflict view	439
<b>7. The social conflict view in action</b>	<b>439</b>
7.1. Labor markets	440
7.2. Financial markets	441
7.3. Regulation of prices	443
7.4. Political power and economic institutions	445
<b>8. A theory of institutions</b>	<b>448</b>
8.1. Sources of political power	448
8.2. Political power and political institutions	449
8.3. A theory of political institutions	451
<b>9. The theory in action</b>	<b>452</b>
9.1. Rise of constitutional monarchy and economic growth in early modern Europe	452
9.2. Summary	457
9.3. Rise of electoral democracy in Britain	458
9.4. Summary	462
<b>10. Future avenues</b>	<b>463</b>
Acknowledgements	464
References	464

## Abstract

This paper develops the empirical and theoretical case that differences in economic institutions are the fundamental cause of differences in economic development. We first document the empirical importance of institutions by focusing on two “quasi-natural experiments” in history, the division of Korea into two parts with very different economic institutions and the colonization of much of the world by European powers starting in the fifteenth century. We then develop the basic outline of a framework for thinking about why economic institutions differ across countries. Economic institutions determine the incentives of and the constraints on economic actors, and shape economic

outcomes. As such, they are social decisions, chosen for their consequences. Because different groups and individuals typically benefit from different economic institutions, there is generally a conflict over these social choices, ultimately resolved in favor of groups with greater political power. The distribution of political power in society is in turn determined by political institutions and the distribution of resources. Political institutions allocate *de jure* political power, while groups with greater economic might typically possess greater *de facto* political power. We therefore view the appropriate theoretical framework as a dynamic one with political institutions and the distribution of resources as the state variables. These variables themselves change over time because prevailing economic institutions affect the distribution of resources, and because groups with *de facto* political power today strive to change political institutions in order to increase their *de jure* political power in the future. Economic institutions encouraging economic growth emerge when political institutions allocate power to groups with interests in broad-based property rights enforcement, when they create effective constraints on power-holders, and when there are relatively few rents to be captured by power-holders. We illustrate the assumptions, the workings and the implications of this framework using a number of historical examples.

## Keywords

institutions, growth, development, political power, rents, conflict, property rights, efficiency, distributions

*JEL classification:* D7, H1, O10, O40

## 1. Introduction

### 1.1. The question

The most trite yet crucial question in the field of economic growth and development is: Why are some countries much poorer than others? Traditional neoclassical growth models, following Solow (1956), Cass (1965) and Koopmans (1965), explain differences in income per capita in terms of different paths of factor accumulation. In these models, cross-country differences in factor accumulation are due either to differences in saving rates (Solow), preferences (Cass–Koopmans), or other exogenous parameters, such as total factor productivity growth. In these models there are institutions, for example agents have well defined property rights and exchange goods and services in markets, but differences in income and growth are not explained by variation in institutions.

The first wave of the more recent incarnations of growth theory, following Romer (1986) and Lucas (1988) differed in the sense that they emphasized that externalities from physical and human capital accumulation could induce sustained steady-state growth. However, they also stayed squarely within the neoclassical tradition of explaining differences in growth rates in terms of preferences and endowments. The second wave of models, particularly Romer (1990), Grossman and Helpman (1991) and Aghion and Howitt (1992), endogenized steady-state growth and technical progress, but their explanation for income differences is similar to that of the older theories. For instance, in the model of Romer (1990), a country may be more prosperous than another if it allocates more resources to innovation, but what determines this is essentially preferences and properties of the technology for creating ‘ideas’.<sup>1</sup>

Though this theoretical tradition is still vibrant in economics and has provided many insights about the mechanics of economic growth, it has for a long time seemed unable to provide a *fundamental* explanation for economic growth. As North and Thomas (1973, p. 2) put it: “the factors we have listed (innovation, economies of scale, education, capital accumulation, etc.) are not causes of growth; they *are* growth” (italics in original). Factor accumulation and innovation are only *proximate* causes of growth. In North and Thomas’s view, the fundamental explanation of comparative growth is differences in *institutions*.

What are institutions exactly? North (1990, p. 3) offers the following definition: “Institutions are the rules of the game in a society or, more formally, are the humanly devised constraints that shape human interaction”. He goes on to emphasize the key implications of institutions since, “In consequence they structure incentives in human exchange, whether political, social, or economic”.

<sup>1</sup> Although, as we discuss later, some recent contributions to growth theory emphasize the importance of economic policies, such as taxes, subsidies to research, barriers to technology adoption and human capital policy, they typically do not present an explanation for why there are differences in these policies across countries.

Of primary importance to economic outcomes are the *economic institutions* in society such as the structure of property rights and the presence and perfection of markets. Economic institutions are important because they influence the structure of economic incentives in society. Without property rights, individuals will not have the incentive to invest in physical or human capital or adopt more efficient technologies. Economic institutions are also important because they help to allocate resources to their most efficient uses, they determine who gets profits, revenues and residual rights of control. When markets are missing or ignored (as they were in the Soviet Union, for example), gains from trade go unexploited and resources are misallocated. Societies with economic institutions that facilitate and encourage factor accumulation, innovation and the efficient allocation of resources will prosper.

Central to this chapter and to much of political economy research on institutions is that economic institutions, and institutions more broadly, are *endogenous*; they are, at least in part, determined by society, or a segment of it. Consequently, the question of why some societies are much poorer than others is closely related to the question of why some societies have much “worse economic institutions” than others.

Even though many scholars including John Locke, Adam Smith, John Stuart Mill, Arthur Lewis, Douglass North and Robert Thomas, and recently many papers in the literature on economic growth and development, have emphasized the importance of economic institutions, we are far from a useful framework for thinking about how economic institutions are determined and why they vary across countries. In other words, while we have good reason to believe that economic institutions matter for economic growth, we lack the crucial *comparative static* results which will allow us to explain why equilibrium economic institutions differ (and perhaps this is part of the reason why much of the economics literature has focused on the proximate causes of economic growth, largely neglecting fundamental institutional causes).

This chapter has three aims. First, we selectively review the evidence that differences in economic institutions are a fundamental cause of cross-country differences in prosperity. Second, we outline a framework for thinking about why economic institutions vary across countries. We emphasize the potential comparative static results of this framework and also illustrate the key mechanisms through a series of historical examples and case studies. Finally, we highlight a large number of areas where we believe future theoretical and empirical work would be very fruitful.

## 1.2. The argument

The basic argument of this chapter can be summarized as follows:

1. Economic institutions matter for economic growth because they shape the incentives of key economic actors in society, in particular, they influence investments in physical and human capital and technology, and the organization of production. Although cultural and geographical factors may also matter for economic performance, differences in economic institutions are the major source of cross-country differences in economic growth and prosperity. Economic institutions not only determine the ag-

gregate economic growth potential of the economy, but also an array of economic outcomes, including the distribution of resources in the future (i.e., the distribution of wealth, of physical capital or human capital). In other words, they influence not only the size of the aggregate pie, but how this pie is divided among different groups and individuals in society. We summarize these ideas schematically as (where the subscript  $t$  refers to current period and  $t + 1$  to the future):

$$\text{economic institutions}_t \implies \begin{cases} \text{economic performance}_t \\ \text{distribution of resources}_{t+1} \end{cases}.$$

2. Economic institutions are endogenous. They are determined as collective choices of the society, in large part for their economic consequences. However, there is no guarantee that all individuals and groups will prefer the same set of economic institutions because, as noted above, different economic institutions lead to different distributions of resources. Consequently, there will typically be a *conflict of interest* among various groups and individuals over the choice of economic institutions. So how are equilibrium economic institutions determined? If there are, for example, two groups with opposing preferences over the set of economic institutions, which group's preferences will prevail? The answer depends on the *political power* of the two groups. Although the efficiency of one set of economic institutions compared with another may play a role in this choice, political power will be the ultimate arbiter. Whichever group has more political power is likely to secure the set of economic institutions that it prefers. This leads to the second building block of our framework:

$$\text{political power}_t \implies \text{economic institutions}_t.$$

3. Implicit in the notion that political power determines economic institutions is the idea that there are conflicting interests over the distribution of resources and therefore indirectly over the set of economic institutions. But why do the groups with conflicting interests not agree on the set of economic institutions that maximize aggregate growth (the size of the aggregate pie) and then use their political power simply to determine the distribution of the gains? Why does the exercise of political power lead to economic inefficiencies and even poverty? We will explain that this is because there are commitment problems inherent in the use of political power. Individuals who have political power cannot commit not to use it in their best interests, and this commitment problem creates an inseparability between efficiency and distribution because credible compensating transfers and side-payments cannot be made to offset the distributional consequences of any particular set of economic institutions.

4. The distribution of political power in society is also endogenous, however. In our framework, it is useful to distinguish between two components of political power, which we refer to as *de jure (institutional)* and *de facto political power*. Here *de jure* political power refers to power that originates from the *political institutions* in society. Political institutions, similarly to economic institutions, determine the constraints on and the incentives of the key actors, but this time in the political sphere. Examples of political institutions include the form of government, for example, democracy vs. dictatorship or

autocracy, and the extent of constraints on politicians and political elites. For example, in a monarchy, political institutions allocate all de jure political power to the monarch, and place few constraints on its exercise. A constitutional monarchy, in contrast, corresponds to a set of political institutions that reallocates some of the political power of the monarch to a parliament, thus effectively constraining the political power of the monarch. This discussion therefore implies that:

$$\text{political institution}_t \implies \text{de jure political power}_t.$$

5. There is more to political power than political institutions, however. A group of individuals, even if they are not allocated power by political institutions, for example as specified in the constitution, may nonetheless possess political power. Namely, they can revolt, use arms, hire mercenaries, co-opt the military, or use economically costly but largely peaceful protests in order to impose their wishes on society. We refer to this type of political power as de facto political power, which itself has two sources. First, it depends on the ability of the group in question to solve its collective action problem, i.e., to ensure that people act together, even when any individual may have an incentive to free ride. For example, peasants in the Middle Ages, who were given no political power by the constitution, could sometimes solve the collective action problem and undertake a revolt against the authorities. Second, the de facto power of a group depends on its economic resources, which determine both their ability to use (or misuse) existing political institutions and also their option to hire and use force against different groups. Since we do not yet have a satisfactory theory of when groups are able to solve their collective action problems, our focus will be on the second source of de facto political power, hence:

$$\text{distribution of resources}_t \implies \text{de facto political power}_t.$$

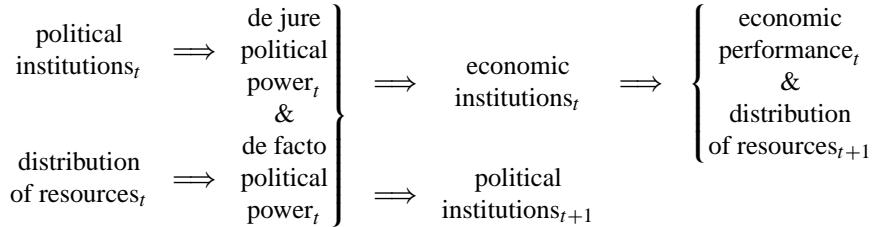
6. This brings us to the evolution of one of the two main *state variables* in our framework, political institutions (the other state variable is the distribution of resources, including distribution of physical and human capital stocks, etc.). Political institutions and the distribution of resources are the state variables in this dynamic system because they typically change relatively slowly, and more importantly, they determine economic institutions and economic performance both directly and indirectly. Their direct effect is straightforward to understand. If political institutions place all political power in the hands of a single individual or a small group, economic institutions that provide protection of property rights and equal opportunity for the rest of the population are difficult to sustain. The indirect effect works through the channels discussed above: political institutions determine the distribution of de jure political power, which in turn affects the choice of economic institutions. This framework therefore introduces a natural concept of *a hierarchy of institutions*, with political institutions influencing equilibrium economic institutions, which then determine economic outcomes.

Political institutions, though slow changing, are also endogenous. Societies transition from dictatorship to democracy, and change their constitutions to modify the constraints on power holders. Since, like economic institutions, political institutions are collective

choices, the distribution of political power in society is the key determinant of their evolution. This creates a tendency for persistence: political institutions allocate de jure political power, and those who hold political power influence the evolution of political institutions, and they will generally opt to maintain the political institutions that give them political power. However, de facto political power occasionally creates changes in political institutions. While these changes are sometimes discontinuous, for example when an imbalance of power leads to a revolution or the threat of revolution leads to major reforms in political institutions, often they simply influence the way existing political institutions function, for example, whether the rules laid down in a particular constitution are respected as in most functioning democracies, or ignored as in current-day Zimbabwe. Summarizing this discussion, we have:

$$\text{political power}_t \implies \text{political institutions}_{t+1}.$$

Putting all these pieces together, a schematic (and simplistic) representation of our framework is as follows:



The two state variables are political institutions and the distribution of resources, and the knowledge of these two variables at time  $t$  is sufficient to determine all the other variables in the system. While political institutions determine the distribution of de jure political power in society, the distribution of resources influences the distribution of de facto political power at time  $t$ . These two sources of political power, in turn, affect the choice of economic institutions and influence the future evolution of political institutions. Economic institutions determine economic outcomes, including the aggregate growth rate of the economy and the distribution of resources at time  $t+1$ . Although economic institutions are the essential factor shaping economic outcomes, they are themselves endogenous and determined by political institutions and distribution of resources in society.

There are two sources of persistence in the behavior of the system: first, political institutions are durable, and typically, a sufficiently large change in the distribution of political power is necessary to cause a change in political institutions, such as a transition from dictatorship to democracy. Second, when a particular group is rich relative to others, this will increase its de facto political power and enable it to push for economic and political institutions favorable to its interests. This will tend to reproduce the initial relative wealth disparity in the future. Despite these tendencies for persistence, the framework also emphasizes the potential for change. In particular, “shocks”, including changes in technologies and the international environment, that modify the balance of

(de facto) political power in society and can lead to major changes in political institutions and therefore in economic institutions and economic growth.

A brief example might be useful to clarify these notions before commenting on some of the underlying assumptions and discussing comparative statics. Consider the development of property rights in Europe during the Middle Ages. There is no doubt that lack of property rights for landowners, merchants and proto-industrialists was detrimental to economic growth during this epoch. Since political institutions at the time placed political power in the hands of kings and various types of hereditary monarchies, such rights were largely decided by these monarchs. Unfortunately for economic growth, while monarchs had every incentive to protect their own property rights, they did not generally enforce the property rights of others. On the contrary, monarchs often used their powers to expropriate producers, impose arbitrary taxation, renege on their debts, and allocate the productive resources of society to their allies in return for economic benefits or political support. Consequently, economic institutions during the Middle Ages provided little incentive to invest in land, physical or human capital, or technology, and failed to foster economic growth. These economic institutions also ensured that the monarchs controlled a large fraction of the economic resources in society, solidifying their political power and ensuring the continuation of the political regime.

The seventeenth century, however, witnessed major changes in the economic and political institutions that paved the way for the development of property rights and limits on monarchs' power, especially in England after the Civil War of 1642 and the Glorious Revolution of 1688, and in the Netherlands after the Dutch Revolt against the Hapsburgs. How did these major institutional changes take place? In England, for example, until the sixteenth century the king also possessed a substantial amount of de facto political power, and leaving aside civil wars related to royal succession, no other social group could amass sufficient de facto political power to challenge the king. But changes in the English land market [Tawney (1941)] and the expansion of Atlantic trade in the sixteenth and seventeenth centuries [Acemoglu, Johnson and Robinson (2005)] gradually increased the economic fortunes, and consequently the de facto power of landowners and merchants. These groups were diverse, but contained important elements that perceived themselves as having interests in conflict with those of the king: while the English kings were interested in predating against society to increase their tax incomes, the gentry and merchants were interested in strengthening their property rights.

By the seventeenth century, the growing prosperity of the merchants and the gentry, based both on internal and overseas, especially Atlantic, trade, enabled them to field military forces capable of defeating the king. This de facto power overcame the Stuart monarchs in the Civil War and Glorious Revolution, and led to a change in political institutions that stripped the king of much of his previous power over policy. These changes in the distribution of political power led to major changes in economic institutions, strengthening the property rights of both land and capital owners and spurred a process of financial and commercial expansion. The consequence was rapid economic growth, culminating in the Industrial Revolution, and a very different distribution of economic resources from that in the Middle Ages.

It is worth returning at this point to two critical assumptions in our framework. First, why do the groups with conflicting interests not agree on the set of economic institutions that maximize aggregate growth? So in the case of the conflict between the monarchy and the merchants, why does the monarchy not set up secure property rights to encourage economic growth and tax some of the benefits? Second, why do groups with political power want to change political institutions in their favor? For instance, in the context of the example above, why did the gentry and merchants use their de facto political power to change political institutions rather than simply implement the policies they wanted? The answers to both questions revolve around issues of *commitment* and go to the heart of our framework.

The distribution of resources in society is an inherently conflictual, and therefore political, decision. As mentioned above, this leads to major commitment problems, since groups with political power cannot commit to not using their power to change the distribution of resources in their favor. For example, economic institutions that increased the security of property rights for land and capital owners during the Middle Ages would not have been credible as long as the monarch monopolized political power. He could promise to respect property rights, but then at some point, renege on his promise, as exemplified by the numerous financial defaults by medieval kings [e.g., [Veitch \(1986\)](#)]. Credible secure property rights necessitated a reduction in the political power of the monarch. Although these more secure property rights would foster economic growth, they were not appealing to the monarchs who would lose their rents from predation and expropriation as well as various other privileges associated with their monopoly of political power. This is why the institutional changes in England as a result of the Glorious Revolution were not simply conceded by the Stuart kings. James II had to be deposed for the changes to take place.

The reason why political power is often used to change political institutions is related. In a dynamic world, individuals care not only about economic outcomes today but also in the future. In the example above, the gentry and merchants were interested in their profits and therefore in the security of their property rights, not only in the present but also in the future. Therefore, they would have liked to use their (de facto) political power to secure benefits in the future as well as the present. However, commitment to future allocations (or economic institutions) was not possible because decisions in the future would be decided by those who had political power in the future with little reference to past promises. If the gentry and merchants would have been sure to maintain their de facto political power, this would not have been a problem. However, de facto political power is often transient, for example because the collective action problems that are solved to amass this power are likely to resurface in the future, or other groups, especially those controlling de jure power, can become stronger in the future. Therefore, any change in policies and economic institutions that relies purely on de facto political power is likely to be reversed in the future. In addition, many revolutions are followed by conflict within the revolutionaries. Recognizing this, the English gentry and merchants strove not just to change economic institutions in their favor following their victories against the Stuart monarchy, but also to alter political institutions and the

future allocation of de jure power. Using political power to change political institutions then emerges as a useful strategy to make gains more durable. The framework that we propose, therefore, emphasizes the importance of political institutions, and changes in political institutions, as a way of manipulating future political power, and thus indirectly shaping future, as well as present, economic institutions and outcomes.

This framework, though abstract and highly simple, enables us to provide some preliminary answers to our main question: why do some societies choose “good economic institutions”? At this point, we need to be more specific about what good economic institutions are. A danger we would like to avoid is that we define good economic institutions as those that generate economic growth, potentially leading to a tautology. This danger arises because a given set of economic institutions may be relatively good during some periods and bad during others. For example, a set of economic institutions that protects the property rights of a small elite might not be inimical to economic growth when all major investment opportunities are in the hands of this elite, but could be very harmful when investments and participation by other groups are important for economic growth [see [Acemoglu \(2003b\)](#)]. To avoid such a tautology and to simplify and focus the discussion, throughout we think of good economic institutions as those that provide security of property rights and relatively equal access to economic resources to a broad cross-section of society. Although this definition is far from requiring equality of opportunity in society, it implies that societies where only a very small fraction of the population have well-enforced property rights do not have good economic institutions. Consequently, as we will see in some of the historical cases discussed below, a given set of economic institutions may have very different implications for economic growth depending on the technological possibilities and opportunities.

Given this definition of good economic institutions as providing secure property rights for a broad cross-section of society, our framework leads to a number of important comparative statics, and thus to an answer to our basic question. First, political institutions that place checks on those who hold political power, for example, by creating a balance of power in society, are useful for the emergence of good economic institutions. This result is intuitive; without checks on political power, power holders are more likely to opt for a set of economic institutions that are beneficial for themselves and detrimental for the rest of society, which will typically fail to protect property rights of a broad cross-section of people. Second, good economic institutions are more likely to arise when political power is in the hands of a relatively broad group with significant investment opportunities. The reason for this result is that, everything else equal, in this case power holders will themselves benefit from secure property rights.<sup>2</sup> Third, good economic institutions are more likely to arise and persist when there are only limited rents that power holders can extract from the rest of society, since such rents would

<sup>2</sup> The reason why we inserted the caveat of “a relatively broad group” is that when a small group with significant investment opportunities holds power, they may sometimes opt for an oligarchic system where their own property rights are protected, but those of others are not [see [Acemoglu \(2003b\)](#)].

encourage them to opt for a set of economic institutions that make the expropriation of others possible. These comparative statics therefore place political institutions at the center of the story, as emphasized by our term “hierarchy of institutions” above. Political institutions are essential both because they determine the constraints on the use of (de facto and de jure) political power and also which groups hold de jure political power in society. We will see below how these comparative statics help us understand institutional differences across countries and over time in a number of important historical examples.

### 1.3. Outline

In the next section we discuss how economic institutions constitute the basis for a fundamental theory of growth, and we contrast this with other potential fundamental theories. In Section 3 we consider some empirical evidence that suggests a key role for economic institutions in determining long-run growth. We also emphasize some of the key problems involved in establishing a causal relationship between economic institutions and growth. We then show in Section 4 how the experience of European colonialism can be used as a ‘natural experiment’ which can address these problems. Having established the central causal role of economic institutions and their importance relative to other factors in cross-country differences in economic performance, the rest of the paper focuses on developing a theory of economic institutions. Section 5 discusses four types of explanation for why countries have different institutions, and argues that the most plausible is the *social conflict view*. According to this theory, bad institutions arise because the groups with political power benefit from bad institutions. The emphasis on social conflict arises naturally from our observation above that economic institutions influence the distribution of resources as well as efficiency. Different groups or individuals will therefore prefer different institutions and conflict will arise as each tries to get their own way. Section 6 delves deeper into questions of efficiency and asks why a political version of the Coase Theorem does not hold. We emphasize the idea that commitment problems are intrinsic to the exercise of political power. In Section 7 we argue that a series of historical examples of diverging economic institutions are best explained by the social conflict view. These examples illustrate how economic institutions are determined by the distribution of political power, and how this distribution is influenced by political institutions. Section 8 puts these ideas together to build our theory of institutions. In Section 9 we then consider two more extended examples of the theory in action, the rise of constitutional rule in early modern Europe, and the creation of mass democracy, particularly in Britain, in the nineteenth and twentieth centuries. Section 10 concludes with a discussion of where this research program can go next.

## 2. Fundamental causes of income differences

We begin by taking a step back. The presumption in the introduction was that economic institutions matter, and should in fact be thought of as one of the key fundamental causes

of economic growth and cross-country differences in economic performance. How do we know this?

## 2.1. Three fundamental causes

If standard economic models of factor accumulation and endogenous technical change only provide proximate explanations of comparative growth, what types of explanations would constitute fundamental ones? Though there is no conventional wisdom on this, we can distinguish three such theories: the first set of theories, our main focus in this chapter, emphasize the importance of economic institutions, which influence economic outcomes by shaping economic incentives; the second emphasize geography, and the third emphasize the importance of culture (a fourth possibility is that differences are due to “luck”, some societies were just lucky; however we do not believe that differences in luck by themselves constitute a sufficient fundamental causes of cross-country income differences).

### 2.1.1. Economic institutions

At its core, the hypothesis that differences in economic institutions are the fundamental cause of different patterns of economic growth is based on the notion that it is the way that humans themselves decide to organize their societies that determines whether or not they prosper. Some ways of organizing societies encourage people to innovate, to take risks, to save for the future, to find better ways of doing things, to learn and educate themselves, solve problems of collective action and provide public goods. Others do not.

The idea that the prosperity of a society depends on its economic institutions goes back at least to Adam Smith, for example in his discussions of mercantilism and the role of markets, and was prominent in the work of many nineteenth century scholars such as John Stuart Mill [see the discussion in [Jones \(1981\)](#)]: societies are economically successful when they have ‘good’ economic institutions and it is these institutions that are the *cause* of prosperity. We can think of these good economic institutions as consisting of an inter-related cluster of things. There must be enforcement of property rights for a broad cross-section of society so that all individuals have an incentive to invest, innovate and take part in economic activity. There must also be some degree of equality of opportunity in society, including such things as equality before the law, so that those with good investment opportunities can take advantage of them.<sup>3</sup>

One could think of other types of economic institutions and many explanations for growth and development have moved beyond models based on preferences, technology and factor endowments to focus on what might loosely be called institutions. One

<sup>3</sup> In [Acemoglu, Johnson and Robinson \(2001\)](#), we coined the term *institutions of private property* for a cluster of good economic institutions, including the rule of law and the enforcement of property rights, and the term *extractive institutions* to designate institutions under which the rule of law and property rights are absent for large majorities of the population.

set of ideas, important for our work, has emphasized that conflict over resources and predation, as well as production, are fundamental forces in society. Scholars such as [Skaperdas \(1992\)](#), [Grossman and Kim \(1995, 1996\)](#), [Hirshleifer \(2001\)](#) and [Dixit \(2004\)](#) have examined how stable property rights can emerge in such circumstances. These scholars have studied almost institution free models and asked how the type of social order that underlies standard economic models might emerge endogenously. Closely related to this work is the research that shows how rent-seeking and redistributive conflict more generally has important implications for growth [e.g., [Tornell and Velasco \(1992\)](#), [Murphy, Shleifer and Vishny \(1991\)](#) [Acemoglu \(1995\)](#), [Alesina and Perotti \(1996\)](#), [Benhabib and Rustichini \(1996\)](#)].

Another literature, following in the footsteps of traditional accounts of economic growth by historians, following the lead of Adam Smith, has emphasized the perfection and spread of markets, clearly a key economic institution [[Pirenne \(1937\)](#), [Hicks \(1969\)](#)]. Problems of the imperfection or absence of markets can clearly have important ramifications for resource allocation, incentives and growth. A central role here has been played by capital markets. For example, [Banerjee and Newman \(1993\)](#) and [Galor and Zeira \(1993\)](#) propose canonical models of how imperfect financial markets can impede growth and development. Models of poverty traps in the tradition of [Rosenstein-Rodan \(1943\)](#), [Murphy, Vishny and Shleifer \(1989a, 1989b\)](#) and [Acemoglu \(1995, 1997\)](#), are based on the idea that market imperfections can lead to the existence of multiple Pareto-ranked equilibria. As a consequence a country can get stuck in a Pareto inferior equilibrium, associated with poverty, but getting out of such a trap necessitates coordinated activities that the market cannot deliver. Other mechanisms, such as increasing returns to scale, can lead to similar situations [e.g., [Durlauf \(1993\)](#), [Krugman and Venables \(1995\)](#), see [Azariadis and Stachurski \(2005\)](#), for other mechanisms and examples]. The implications of many other types of market imperfections have been considered, for example in the labor market [[Aghion and Howitt \(1994\)](#), [Pissarides \(2000\)](#)] and other scholars have examined the implications of industrial organization, market structure and the nature of competition [e.g., [Acemoglu and Zilibotti \(1997\)](#), [Aghion et al. \(2001\)](#), [Aghion and Howitt \(2005\)](#)].

The idea that market imperfections and economic institutions play a central role in development has also been important in the academic literature on development economics since its initiation. Both Adam Smith and Alfred Marshall argued that share-cropping was an inefficient way of organizing agriculture because it gave incorrect incentives to tenants. This argument has been formalized, and at the heart of a large literature on development are imperfections in tenancy, labor, land and credit markets [see [Ray \(1998\)](#), [Bardhan and Udry \(1999\)](#), [Banerjee and Duflo \(2005\)](#)].

Finally, the literature that one might broadly class as institutional has extensively discussed political economy models. Most influential is the early work of [Perotti \(1993\)](#), [Saint-Paul and Verdier \(1993\)](#), [Alesina and Rodrik \(1994\)](#) and [Persson and Tabellini \(1994\)](#) who developed dynamic models to examine the effect of redistributive taxation on growth. There are now many models where political mechanisms and outcomes can have important influences on the growth rate [see [Ades and Verdier \(1996\)](#), [Krusell and](#)

Ríos-Rull (1999), Bourguignon and Verdier (2000) and other contributions which we discuss in the body of the paper].

At some level then there is a bewildering array of ideas connecting institutions, both economic and political, to growth and development. In this chapter however, as will already be apparent, we do not attempt to survey all of these theories. Rather, we attempt to develop a perspective on this topic which revolves around what we see as the key issues. From the empirical side this entails really establishing the causal role of institutions in development. From the theoretical side this involves emphasizing the importance of understanding why institutions differ across countries. From the perspective of this chapter the main problem with most of the existing research is the lack of comparative statics and the absence of a truly comparative focus. For instance, in the model of Grossman and Kim (1995) stable property rights may emerge as an equilibrium, but whether they do so or not depends on parameters in the fighting technology which are hard to interpret in reality. Most models of imperfect markets and multiple equilibria fail to provide explanations either for why markets are incomplete or imperfect, or for how some societies manage to get into good equilibria while others do not. To the extent that imperfect market are grounded in imperfections in information or possibilities for opportunism, one would like to know how and why these vary across countries in ways which are consistent with the basic facts about relative economic outcomes. We believe that the structure of markets is endogenous, and partly determined by property rights. Once individuals have secure property rights and there is equality of opportunity, the incentives will exist to create and improve markets (even though achieving perfect markets would be typically impossible). Thus we expect differences in markets to be an outcome of differing systems of property rights and political institutions, not unalterable characteristics responsible for cross-country differences in economic performance. This motivates our focus on economic institutions related to the enforcement of the property rights of a broad cross-section of society.

There are some genuinely comparative studies in the literature. For example, Banerjee and Newman (1993), Alesina and Rodrik (1994) and Persson and Tabellini (1994) all point to differences in wealth distribution as the key to success or failure. We will discuss other such theories, for example those connected to legal origins [e.g., La Porta et al. (1998)] later. Nevertheless, these studies are very different from the approach we propose in this chapter.

### 2.1.2. Geography

While institutional theories emphasize the importance of man-made factors shaping incentives, an alternative is to focus on the role of “nature”, that is, on the physical and geographical environment. In the context of understanding cross-country differences in economic performance, this approach emphasizes differences in geography, climate and ecology that determine both the preferences and the opportunity set of individual economic agents in different societies. We refer to this broad approach as the “geography

hypothesis". There are at least three main versions of the geography hypothesis, each emphasizing a different mechanism for how geography affects prosperity.

First, climate may be an important determinant of work effort, incentives, or even productivity. This idea dates back at least to the famous French philosopher, [Montesquieu \(1748\)](#), who wrote in his classic book *The Spirit of the Laws*: "The heat of the climate can be so excessive that the body there will be absolutely without strength. So, prostration will pass even to the spirit; no curiosity, no noble enterprise, no generous sentiment; inclinations will all be passive there; laziness there will be happiness", and "People are ... more vigorous in cold climates. The inhabitants of warm countries are, like old men, timorous; the people in cold countries are, like young men, brave." One of the founders of modern economics Marshall is another prominent figure who emphasized the importance of climate, arguing: "vigor depends partly on race qualities: but these, so far as they can be explained at all, seem to be chiefly due to climate" [[Marshall \(1890, p. 195\)](#)].

Second, geography may determine the technology available to a society, especially in agriculture. This view is developed by an early Nobel Prize winner in economics, Myrdal, who wrote "serious study of the problems of underdevelopment ... should take into account the climate and its impacts on soil, vegetation, animals, humans and physical assets – in short, on living conditions in economic development" [[Myrdal \(1968, vol. 3, p. 212\)](#)]. More recently, Diamond espouses this view, "... proximate factors behind Europe's conquest of the Americas were the differences in all aspects of technology. These differences stemmed ultimately from Eurasia's much longer history of densely populated ... [societies dependent on food production]", which was in turn determined by geographical differences between Europe and the Americas [[Diamond \(1997, p. 358\)](#)]. The economist Sachs has been a recent and forceful proponent of the importance of geography in agricultural productivity, stating that "By the start of the era of modern economic growth, if not much earlier, temperate-zone technologies were more productive than tropical-zone technologies ..." [[Sachs \(2001, p. 2\)](#)].

The third variant of the geography hypothesis, especially popular over the past decade, links poverty in many areas of the world to their "disease burden", emphasizing that: "The burden of infectious disease is similarly higher in the tropics than in the temperate zones" [[Sachs \(2000, p. 32\)](#)]. [Bloom and Sachs \(1998\)](#) claim that the prevalence of malaria, a disease which kills millions of children every year in sub-Saharan Africa, reduces the annual growth rate of sub-Saharan African economies by more than 1.3 percent a year (this is a large effect, implying that had malaria been eradicated in 1950, income per capita in sub-Saharan Africa would be double what it is today).

### 2.1.3. *Culture*

The final fundamental explanation for economic growth emphasizes the idea that different societies (or perhaps different races or ethnic groups) have different cultures, because of different shared experiences or different religions. Culture is viewed as a

key determinant of the values, preferences and beliefs of individuals and societies and, the argument goes, these differences play a key role in shaping economic performance.

At some level, culture can be thought to influence equilibrium outcomes for a given set of institutions. Possibly there are multiple equilibria connected with any set of institutions and differences in culture mean that different societies will coordinate on different equilibria. Alternatively, as argued by Greif (1994), different cultures generate different sets of beliefs about how people behave and this can alter the set of equilibria for a given specification of institutions (for example, some beliefs will allow punishment strategies to be used whereas others will not).

The most famous link between culture and economic development is that proposed by Weber (1930) who argued that the origins of industrialization in western Europe could be traced to the Protestant reformation and particularly the rise of Calvinism. In his view, the set of beliefs about the world that was intrinsic to Protestantism were crucial to the development of capitalism. Protestantism emphasized the idea of predestination in the sense that some individuals were ‘chosen’ while others were not. “We know that a part of humanity is saved, the rest damned. To assume that human merit or guilt play a part in determining this destiny would be to think of God’s absolutely free decrees, which have been settled from eternity, as subject to change by human influence, an impossible contradiction” [Weber (1930, p. 60)].

But who had been chosen and who not? Calvin did not explain this. Weber (1930, p. 66) notes “Quite naturally this attitude was impossible for his followers . . . for the broad mass of ordinary men . . . So wherever the doctrine of predestination was held, the question could not be suppressed whether there was any infallible criteria by which membership of the *electi* could be known”. Practical solutions to this problem were quickly developed, “. . . in order to attain that self-confidence intense worldly activity is recommended as the most suitable means. It and it alone disperses religious doubts and gives the certainty of grace” [Weber (1930, pp. 66–67)].

Thus “however useless good works might be as a means of attaining salvation . . . nevertheless, they are indispensable as a sign of election. They are the technical means, not of purchasing salvation, but of getting rid of the fear of damnation” (p. 69). Though economic activity was encouraged, enjoying the fruits of such activity was not. “Waste of time is . . . the first and in principle the deadliest of sins. The span of human life is infinitely short and precious to make sure of one’s own election. Loss of time through sociability, idle talk, luxury, even more sleep than is necessary for health . . . is worthy of absolute moral condemnation . . . Unwillingness to work is symptomatic of the lack of grace” (pp. 104–105).

Thus Protestantism led to a set of beliefs which emphasized hard work, thrift, saving, and where economic success was interpreted as consistent with (if not actually signaling) being chosen by God. Weber contrasted these characteristics of Protestantism with those of other religions, such as Catholicism, which he argued did not promote capitalism. For instance on his book on Indian religion he argued that the caste system blocked capitalist development [Weber (1958, p. 112)].

More recently, scholars, such as [Landes \(1998\)](#), have also argued that the origins of Western economic dominance are due to a particular set of beliefs about the world and how it could be transformed by human endeavor, which is again linked to religious differences. Although [Barro and McCleary \(2003\)](#) provide evidence of a positive correlation between the prevalence of religious beliefs, notably about hell and heaven, and economic growth, this evidence does not show a causal effect of religion on economic growth, since religious beliefs are endogenous both to economic outcomes and to other fundamental causes of income differences [points made by [Tawney \(1926\)](#), and [Hill \(1961b\)](#), in the context of Weber's thesis].

Ideas about how culture may influence growth are not restricted to the role of religion. Within the literature trying to explain comparative development there have been arguments that there is something special about particular cultural endowments, usually linked to particular nation states. For instance, Latin America may be poor because of its Iberian heritage, while North America is prosperous because of its Anglo-Saxon heritage [[Vélez \(1994\)](#)]. In addition, a large literature in anthropology argues that societies may become ‘dysfunctional’ or ‘maladapted’ in the sense that they adopt a system of beliefs or ways of operating which do not promote the success or prosperity of the society [see [Edgerton \(1992\)](#), for a survey of this literature]. The most famous version of such an argument is due to [Banfield \(1958\)](#) who argued that the poverty of Southern Italy was due to the fact that people had adopted a culture of “amoral familism” where they only trusted individuals of their own families and refused to cooperate or trust anyone else. This argument was revived in the extensive empirical study of [Putnam, Leonardi and Nanetti \(1993\)](#) who characterized such societies as lacking “social capital”. Although Putnam and others, for example, [Knack and Keefer \(1997\)](#) and [Durlauf and Fafchamps \(2004\)](#), document positive correlations between measures of social capital and various economic outcomes, there is no evidence of a causal effect, since, as with religious beliefs discussed above, measures of social capital are potentially endogenous.

### **3. Institutions matter**

We now argue that there is convincing empirical support for the hypothesis that differences in economic institutions, rather than geography or culture, *cause* differences in incomes per-capita. Consider first [Figure 1](#).

This shows the cross-country bivariate relationship between the log of GDP per capita in 1995 and a broad measure of property rights, “protection against expropriation risk”, averaged over the period 1985 to 1995. The data on economic institutions come from Political Risk Services, a private company which assesses the risk that investments will be expropriated in different countries. These data, first used by [Knack and Keefer \(1995\)](#) and subsequently by [Hall and Jones \(1999\)](#) and [Acemoglu, Johnson and Robinson \(2001, 2002\)](#) are imperfect as a measure of economic institutions, but the findings are robust to using other available measures of economic institutions. The scatter plot

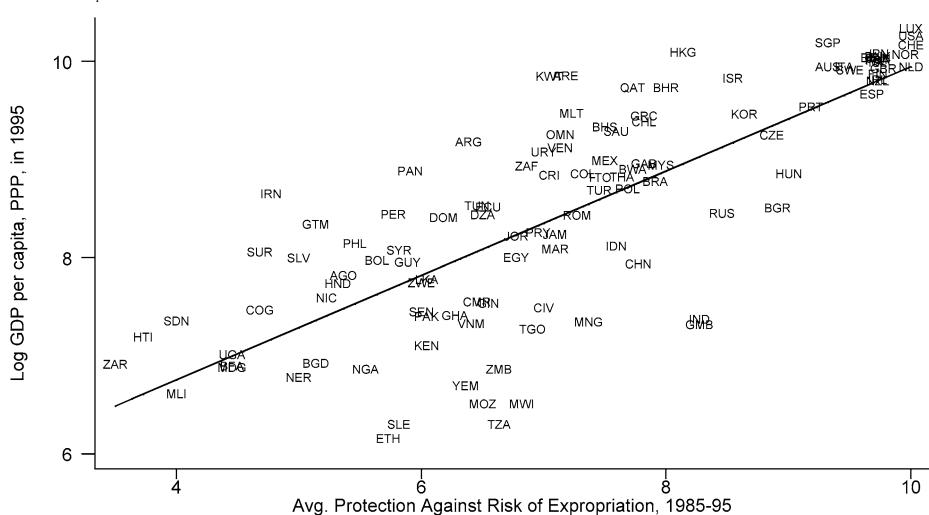


Figure 1. Average protection against risk of expropriation 1985–95 and log GDP per capita 1995.

shows that countries with more secure property rights, i.e., better economic institutions, have higher average incomes.

It is tempting to interpret Figure 1 as depicting a causal relationship (i.e., as establishing that secure property rights cause prosperity). Nevertheless, there are well-known problems with making such an inference. First, there could be reverse causation – perhaps only countries that are sufficiently wealthy can afford to enforce property rights. More importantly, there might be a problem of omitted variable bias. It could be something else, e.g., geography, that explains both why countries are poor and why they have insecure property rights. Thus if omitted factors determine institutions and incomes, we would spuriously infer the existence of a causal relationship between economic institutions and incomes when in fact no such relationship exists. Trying to estimate the relationship between institutions and prosperity using Ordinary Least Squares, as was done by Knack and Keefer (1995) and Barro (1997) could therefore result in biased regression coefficients.

To further illustrate these potential *identification* problems, suppose that climate, or geography more generally, matters for economic performance. In fact, a simple scatterplot shows a positive association between latitude (the absolute value of distance from the equator) and income per capita. Montesquieu, however, not only claimed that warm climate makes people lazy and thus unproductive, but also unfit to be governed by democracy. He argued that despotism would be the political system in warm climates. Therefore, a potential explanation for the patterns we see in Figure 1 is that there is an omitted factor, geography, which explains both economic institutions and economic performance. Ignoring this potential third factor would lead to mistaken conclusions.

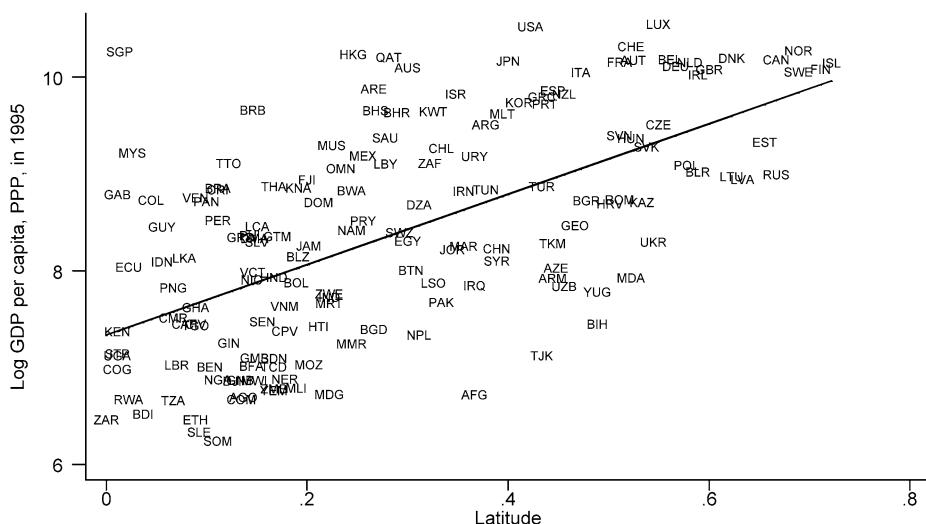


Figure 2. Latitude and log GDP per capita 1995.

Even if Montesquieu's story appears both unrealistic and condescending to our modern sensibilities, the general point should be taken seriously: the relationship shown in Figure 1, and for that matter that shown in Figure 2, is not causal. As we pointed out in the context of the effect of religion or social capital on economic performance, these types of scatterplots, correlations, or their multidimensional version in OLS regressions, *cannot* establish causality.

What can we do? The solution to these problems of inference is familiar in microeconomics: find a source of variation in economic institutions that should have no effect on economic outcomes, or depending on the context, look for a natural experiment. As an example, consider first one of the clearest natural experiments for institutions.

### 3.1. The Korean experiment

Until the end of World War II, Korea was under Japanese occupation. Korean independence came shortly after the Japanese Emperor Hirohito announced the Japanese surrender on August 15, 1945. After this date, Soviet forces entered Manchuria and North Korea and took over the control of these provinces from the Japanese. The major fear of the United States during this time period was the takeover of the entire Korean peninsula either by the Soviet Union or by communist forces under the control of the former guerrilla fighter, Kim Il Sung. U.S. authorities therefore supported the influential nationalist leader Syngman Rhee, who was in favor of separation rather than a united communist Korea. Elections in the South were held in May 1948, amidst a widespread

boycott by Koreans opposed to separation. The newly elected representatives proceeded to draft a new constitution and established the Republic of Korea to the south of the 38th parallel. The North became the Democratic People's Republic of Korea, under the control of Kim II Sung. These two independent countries organized themselves in very different ways and adopted completely different sets of institutions. The North followed the model of Soviet socialism and the Chinese Revolution in abolishing private property of land and capital. Economic decisions were not mediated by the market, but by the communist state. The South instead maintained a system of private property and the government, especially after the rise to power of Park Chung Hee in 1961, attempted to use markets and private incentives in order to develop the economy.

Before this “natural experiment” in institutional change, North and South Korea shared the same history and cultural roots. In fact, Korea exhibited an unparalleled degree of ethnic, linguistic, cultural, geographic and economic homogeneity. There are few geographic distinctions between the North and South, and both share the same disease environment. For example, the CIA Factbook describes the climate of North Korea as “temperate with rainfall concentrated in summer” and that of South Korea as “temperate, with rainfall heavier in summer than winter”. In terms of terrain North Korea is characterized as consisting of “mostly hills and mountains separated by deep, narrow valleys; coastal plains wide in west, discontinuous in east”, while South Korea is “mostly hills and mountains; wide coastal plains in west and south”. In terms of natural resources North Korea is better endowed with significant reserves of coal, lead, tungsten, zinc, graphite, magnesite, iron ore, copper, gold, pyrites, salt, fluorspar, hydropower. South Korea’s natural resources are “coal, tungsten, graphite, molybdenum, lead, hydropower potential”. Both countries share the same geographic possibilities in terms of access to markets and the cost of transportation.

Other man-made initial economic conditions were also similar, and if anything, advantaged the North. For example, there was significant industrialization during the colonial period with the expansion of both Japanese and indigenous firms. Yet this development was concentrated more in the North than the South. For instance, the large Japanese zaibatsu of Noguchi, which accounted for one third of Japanese investment in Korea, was centered in the North. It built large hydroelectric plants, including the Suiho dam on the Yalu river, second in the world only to the Boulder dam on the Colorado river. It also created Nippon Chisso, the second largest chemical complex in the world that was taken over by the North Korean state. Finally, in Ch’ongjin North Korea also had the largest port on the Sea of Japan. All in all, despite some potential advantages for the North,<sup>4</sup> Maddison (2001) estimates that at the time of separation, North and South Korea had approximately the same income per capita.

We can therefore think of the splitting on the Koreas 50 years ago as a natural experiment that we can use to identify the causal influence of a particular dimension of

<sup>4</sup> Such initial differences were probably eradicated by the intensive bombing campaign that the United States unleashed in the early 1950’s on North Korea [see Cumings (2004, Chapter 1)].

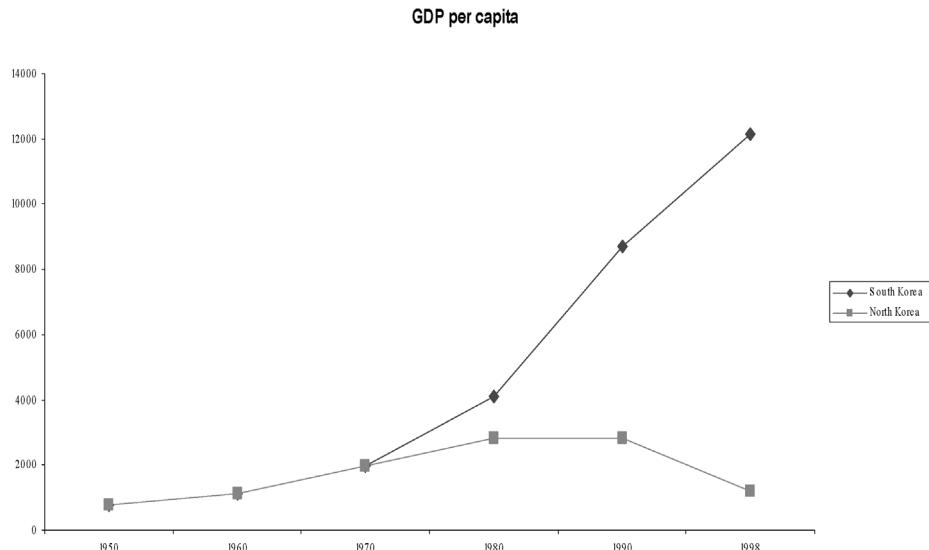


Figure 3. GDP per capita in North and South Korea, 1950–98.

institutions on prosperity. Korea was split into two, with the two halves organized in radically different ways, and with geography, culture and many other potential determinants of economic prosperity held fixed. Thus any differences in economic performance can plausibly be attributed to differences in institutions.

Consistent with the hypothesis that it is institutional differences that drive comparative development, since separation, the two Koreas have experienced dramatically diverging paths of economic development (Figure 3).

By the late 1960's South Korea was transformed into one of the Asian "miracle" economies, experiencing one of the most rapid surges of economic prosperity in history while North Korea stagnated. By 2000 the level of income in South Korea was \$16,100 while in North Korea it was only \$1,000. By 2000 the South had become a member of the Organization of Economic Cooperation and Development, the rich nations club, while the North had a level of per-capita income about the same as a typical sub-Saharan African country. There is only one plausible explanation for the radically different economic experiences on the two Koreas after 1950: their very different institutions led to divergent economic outcomes. In this context, it is noteworthy that the two Koreas not only shared the same geography, but also the same culture.

It is possible that Kim Il Sung and Communist Party members in the North believed that communist policies would be better for the country and the economy in the late 1940s. However, by the 1980s it was clear that the communist economic policies in the North were not working. The continued efforts of the leadership to cling to these policies and to power can only be explained by those leaders wishing to look after their own interests at the expense of the population at large. Bad institutions are therefore

kept in place, clearly not for the benefit of society as a whole, but for the benefit of the ruling elite, and this is a pattern we encounter in most cases of institutional failure that we discuss in detail below.

However convincing on its own terms, the evidence from this natural experiment is not sufficient for the purposes of establishing the importance of economic institutions as the primary factor shaping cross-country differences in economic prosperity. First, this is only one case, and in the better-controlled experiments in the natural sciences, a relatively large sample is essential. Second, here we have an example of an extreme case, the difference between a market-oriented economy and a communist one. Few social scientists today would deny that a lengthy period of totalitarian centrally planned rule has significant economic costs. And yet, many might argue that differences in economic institutions among capitalist economies or among democracies are not the major factor leading to differences in their economic trajectories. To establish the major role of economic institutions in the prosperity and poverty of nations we need to look at a larger scale “natural experiment” in institutional divergence.

### 3.2. *The colonial experiment*

The colonization of much of the world by Europeans provides such a large scale natural experiment. Beginning in the early fifteenth century and massively intensifying after 1492, Europeans conquered many other nations. The colonization experience transformed the institutions in many diverse lands conquered or controlled by Europeans. Most importantly, Europeans imposed very different sets of institutions in different parts of their global empire, as exemplified most sharply by the contrast to the economic institutions in the northeast of America to those in the plantation societies of the Caribbean. As a result, while geography was held constant, Europeans initiated large changes in economic institutions, in the social organization of different societies. We will now show that this experience provides evidence which conclusively establishes the central role of economic institutions in development. Given the importance of this material and the details we need to provide, we discuss the colonial experience in the next section.

## 4. The Reversal of Fortune

The impact of European colonialism on economic institutions is perhaps most dramatically conveyed by a single fact – historical evidence shows that there has been a remarkable Reversal of Fortune in economic prosperity within former European colonies. Societies like the Mughals in India, and the Aztecs and the Incas in the Americas were among the richest civilizations in 1500, yet the nation states that now coincide with the boundaries of these empires are among the poorer societies of today. In contrast, countries occupying the territories of the less-developed civilizations in North America, New Zealand and Australia are now much richer than those in the lands of the Mughals, Aztecs and Incas.

#### 4.1. The reversal among the former colonies

The Reversal of Fortune is not confined to such comparisons. Using reasonable proxies for prosperity before modern times, we can show that it is a much more systematic phenomenon. Our proxies for income per capita in pre-industrial societies are urbanization rates and population density. Only societies with a certain level of productivity in agriculture and a relatively developed system of transport and commerce can sustain large urban centers and a dense population. Figure 4 shows the relationship between income per capita and urbanization (fraction of the population living in urban centers with greater than 5000 inhabitants) today, and demonstrates that in the current era there is a significant relationship between urbanization and prosperity.

Naturally, high rates of urbanization do not mean that the majority of the population lived in prosperity. In fact, before the twentieth century urban areas were centers of poverty and ill health. Nevertheless, urbanization is a good proxy for average income per capita in society, which closely corresponds to the measure we are using to look at prosperity.

Figures 5 and 6 show the relationship between income per capita today and urbanization rates and (log) population density in 1500 for the sample of European colonies.<sup>5</sup>

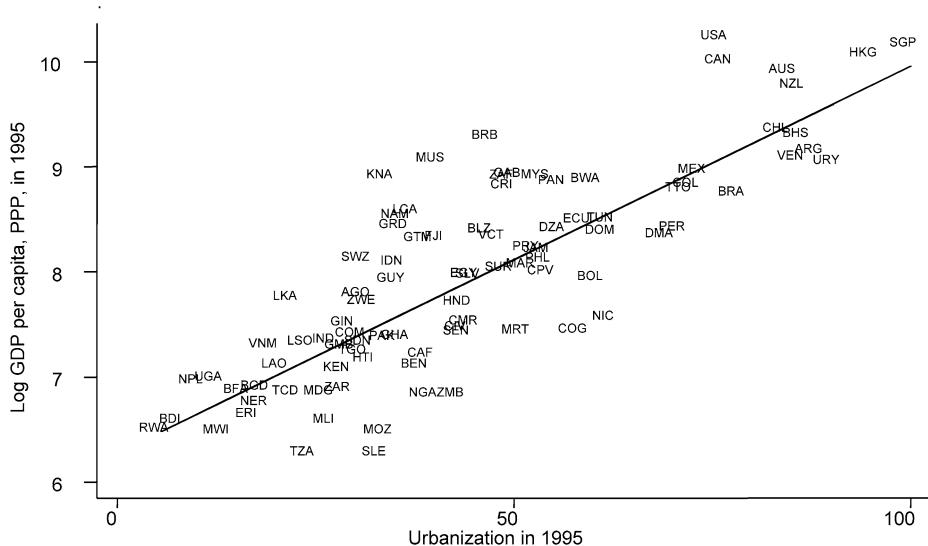


Figure 4. Urbanization in 1995 and log GDP per capita in 1995.

<sup>5</sup> The sample includes the countries colonized by the Europeans between the 15th and the 19th centuries as part of their overseas expansion after the discovery of the New World and the rounding of the Cape of Good Hope. It therefore excludes Ireland, parts of the Russian Empire and also the Middle East and countries briefly controlled by European powers as U.N. Mandates during the 20th century.

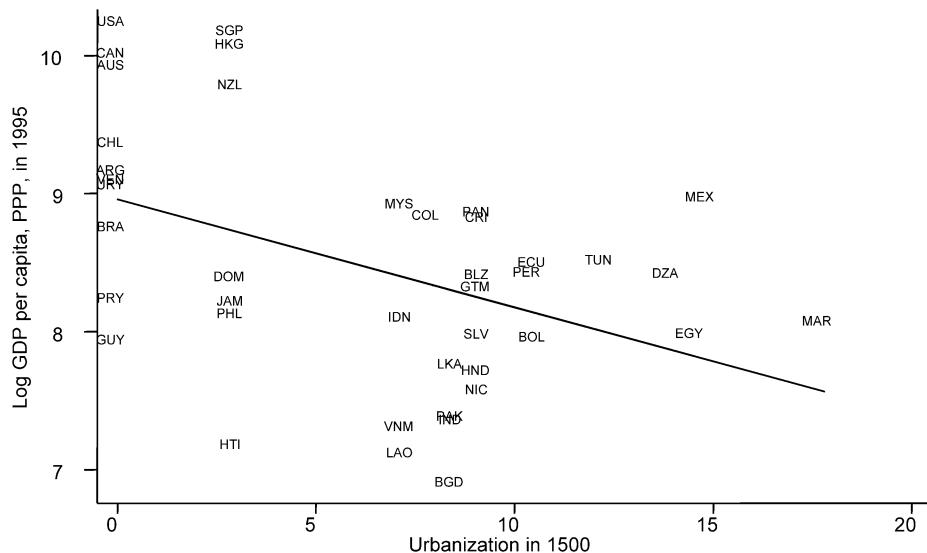


Figure 5. Urbanization in 1500 and log GDP per capita in 1995, among former European colonies.

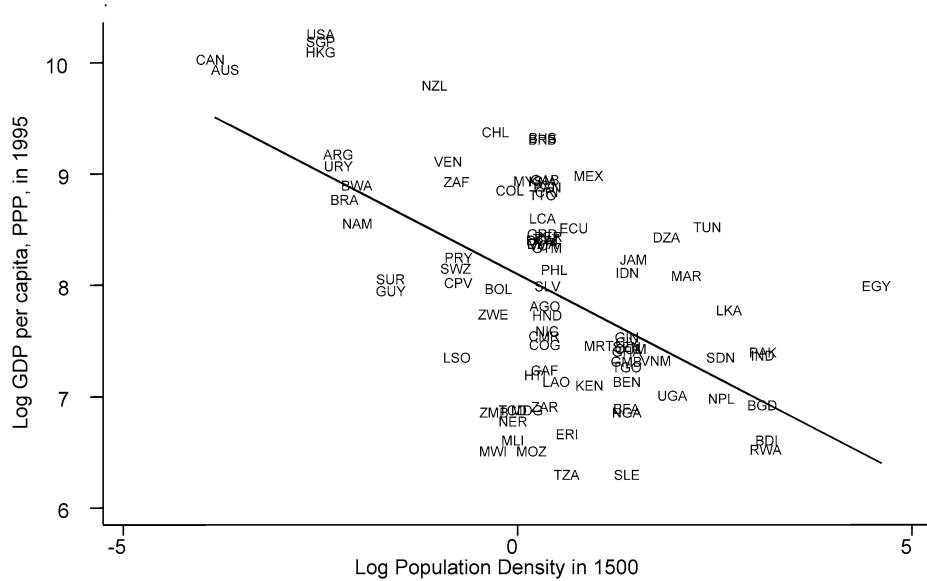


Figure 6. Log population density in 1500 and log GDP per capita in 1995, among former European colonies.

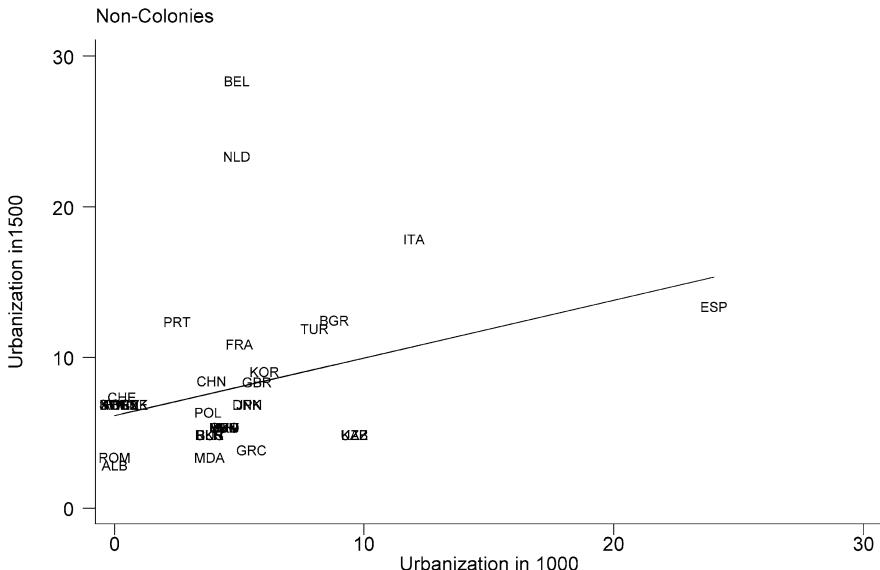


Figure 7. Urbanization in 1000 and 1500, among non-colonies.

We pick 1500 since it is before European colonization had an effect on any of these societies. A strong negative relationship, indicating a reversal in the rankings in terms of economic prosperity between 1500 and today, is clear in both figures. In fact, the figures show that in 1500 the temperate areas were generally less prosperous than the tropical areas, but this pattern too was reversed by the twentieth century.

The urbanization data for these figures come from [Bairoch \(1988\)](#), [Bairoch, Batou and Chèvre \(1988\)](#), [Chandler \(1987\)](#), and [Eggimann \(1999\)](#). The data on population density are from [McEvedy and Jones \(1978\)](#). Details and further results are in [Acemoglu, Johnson and Robinson \(2002\)](#).

There is something extraordinary about this reversal. For example, after the initial spread of agriculture there was remarkable persistence in urbanization and population density for all countries, including those which were to be subsequently colonized by Europeans. In [Figures 7 and 8](#) we show the relationships for urbanization plotting separately the relationship between urbanization in 1000 and in 1500 for the samples of colonies and all other countries. Both figures show persistence, not reversal. Although Ancient Egypt, Athens, Rome, Carthage and other empires rose and fell, what these pictures show is that there was remarkable persistence in the prosperity of regions.

Moreover, reversal was not the general pattern in the world after 1500. [Figure 9](#) shows that within countries not colonized by Europeans in the early modern and modern period, there was no reversal between 1500 and 1995.

There is therefore no reason to think that what is going on in [Figures 5 and 6](#) is some sort of natural reversion to the mean.

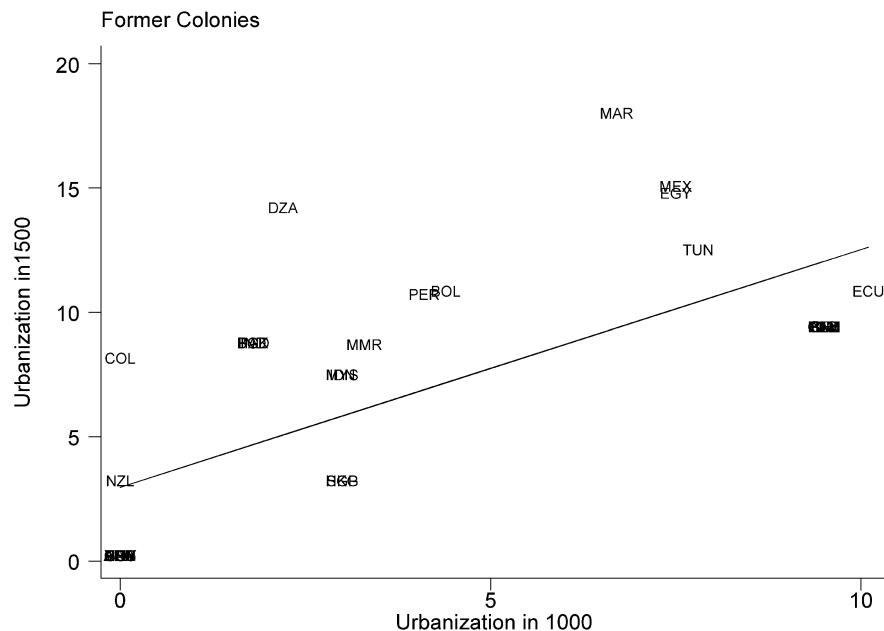


Figure 8. Urbanization in 1000 and 1500, among former European colonies.

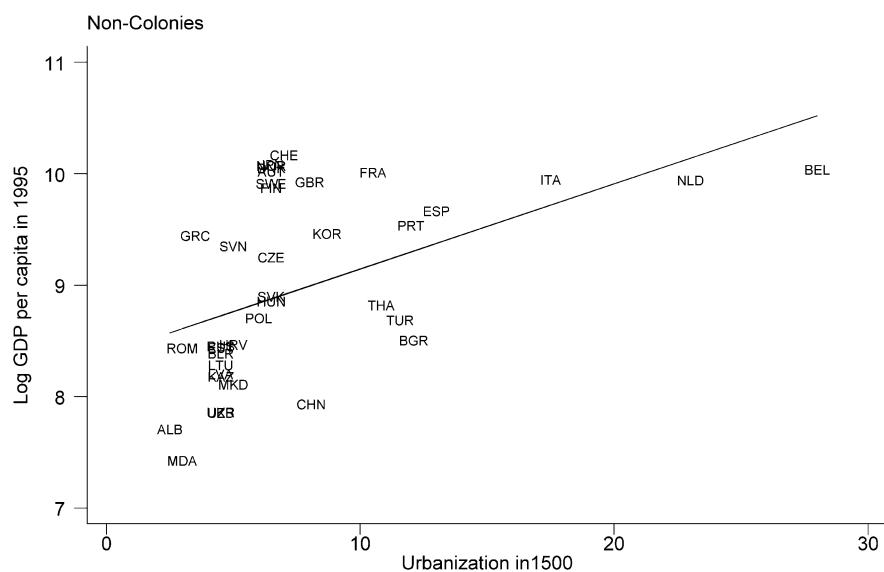


Figure 9. Urbanization in 1500 and log GDP per capita in 1995, among non-colonies.

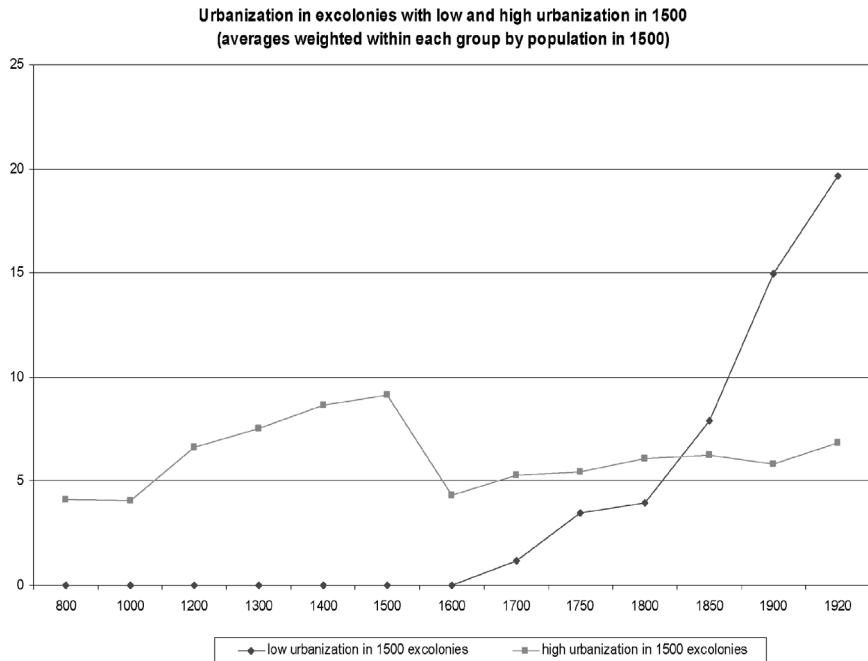


Figure 10. Evolution of urbanization among former European colonies.

#### 4.2. Timing of the reversal

When did the reversal occur? One possibility is that it arose shortly after the conquest of societies by Europeans but [Figures 10 and 11](#) show that the previously-poor colonies surpassed the former highly-urbanized colonies starting in the late eighteenth and early nineteenth centuries, and this went hand in hand with industrialization. [Figure 10](#) shows average urbanization in colonies with relatively low and high urbanization in 1500. The initially high-urbanization countries have higher levels of urbanization and prosperity until around 1800. At that time the initially low-urbanization countries start to grow much more rapidly and a prolonged period of divergence begins. [Figure 11](#) shows industrial production per capita in a number of countries. Although not easy to see in the figure, there was more industry (per capita and total) in India in 1750 than in the United States. By 1860, the United States and British colonies with relatively good economic institutions, such as Australia and New Zealand, began to move ahead rapidly, and by 1953, a huge gap had opened up.

#### 4.3. Interpreting the reversal

Which of the three broad hypotheses about the sources of cross-country income differences are consistent with the reversal and its timing? These patterns are clearly

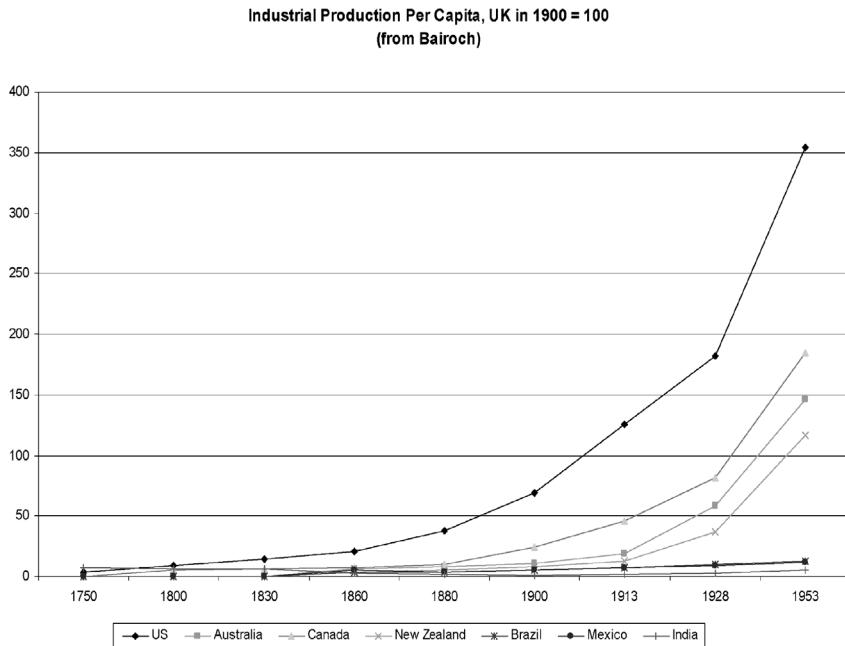


Figure 11. Evolution of industrial production per capita among former European colonies.

inconsistent with simple geography based views of relative prosperity. In 1500 it was the countries in the tropics which were relatively prosperous, in 2003 it is the reverse. This makes it implausible to base a theory of relative prosperity today, as [Sachs \(2000, 2001\)](#) does, on the intrinsic poverty of the tropics. This argument is inconsistent with the historical evidence.

Nevertheless, following [Diamond \(1997\)](#), one could propose what [Acemoglu, Johnson and Robinson \(2002\)](#) call a “sophisticated geography hypothesis” which claims that geography matters but in a time varying way. For example, Europeans created “latitude specific” technology, such as heavy metal ploughs, that only worked in temperate latitudes and not with tropical soils. Thus when Europe conquered most of the world after 1492, they introduced specific technologies that functioned in some places (the United States, Argentina, Australia) but not others (Peru, Mexico, West Africa). However, the timing of the reversal, coming as it does in the nineteenth century, is inconsistent with the most natural types of sophisticated geography hypotheses. Europeans may have had latitude specific technologies, but the timing implies that these technologies must have been industrial, not agricultural, and it is difficult to see why industrial technologies do not function in the tropics (and in fact, they have functioned quite successfully in tropical Singapore and Hong Kong).<sup>6</sup>

<sup>6</sup> A possible link is that proposed by [Lewis \(1978\)](#) who argued that tropical agriculture is less productive than temperate agriculture, and that an ‘agricultural revolution’ is a prerequisite to an industrial revolution because

Similar considerations weigh against the culture hypothesis. Although culture is slow-changing the colonial experiment was sufficiently radical to have caused major changes in the cultures of many countries that fell under European rule. In addition, the destruction of many indigenous populations and immigration from Europe are likely to have created new cultures or at least modified existing cultures in major ways [see [Vargas Llosa \(1989\)](#), for a fictionalized account of just such a cultural change]. Nevertheless, the culture hypothesis does not provide a natural explanation for the reversal, and has nothing to say on the timing of the reversal. Moreover, we discuss below how econometric models that control for the effect of institutions on income do not find any evidence of an effect of religion or culture on prosperity.

The most natural explanation for the reversal comes from the institutions hypothesis, which we discuss next.

#### *4.4. Economic institutions and the reversal*

Is the Reversal of Fortune consistent with a dominant role for economic institutions in comparative development? The answer is yes. In fact, once we recognize the variation in economic institutions created by colonization, we see that the Reversal of Fortune is exactly what the institutions hypothesis predicts.

In [Acemoglu, Johnson and Robinson \(2002\)](#) we tested the connection between initial population density, urbanization, and the creation of good economic institutions. We showed that, others things equal, the higher the initial population density or the greater initial urbanization, the worse were subsequent institutions, including both institutions right after independence and today. [Figures 12 and 13](#) show these relationships using the same measure of current economic institutions used in [Figure 1](#), protection against expropriation risk today. They document that the relatively densely settled and highly urbanized colonies ended up with worse (or ‘extractive’) institutions, while sparsely-settled and non-urbanized areas received an influx of European migrants and developed institutions protecting the property rights of a broad cross-section of society. European colonialism therefore led to an institutional reversal, in the sense that the previously-richer and more-densely settled places ended up with worse institutions.<sup>7</sup>

To be fair, it is possible that the Europeans did not actively introduce institutions discouraging economic progress in many of these places, but inherited them from previous civilizations there. The structure of the Mughal, Aztec and Inca empires were already very hierarchical with power concentrated in the hands of narrowly based ruling elites and structured to extract resources from the majority for the benefit of a minority. Often

high agricultural productivity is needed to stimulate the demand for industrial goods. Though obviously such an explanation is not relevant for explaining industrialization in Singapore or Hong Kong, it may be relevant in other places.

<sup>7</sup> The institutional reversal does not mean that institutions were necessarily better in the previously more densely-settled areas (see next paragraph). It only implies a tendency for the relatively poorer and less densely-settled areas to end up with better institutions than previously-rich and more densely-settled areas.

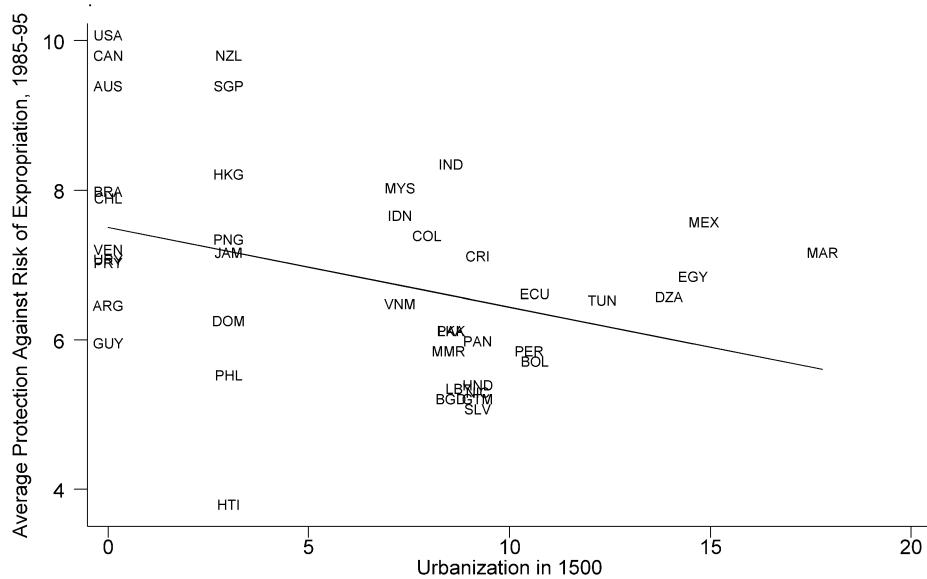


Figure 12. Urbanization in 1500 and average protection against risk of expropriation 1985–95.

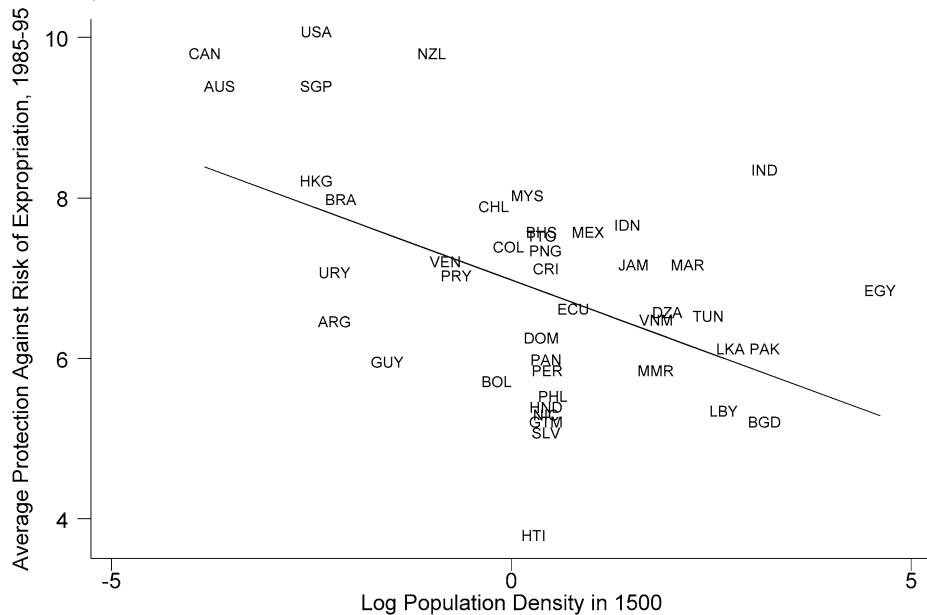


Figure 13. Log population density in 1500 and average protection against risk of expropriation 1985–95.

Europeans simply took over these existing institutions. Whether this is so is secondary for our focus, however. What matters is that in densely-settled and relatively-developed places it was in the interests of Europeans to have institutions facilitating the extraction of resources thus not respecting the property rights of the majority, while in the sparsely-settled areas it was in their interests to develop institutions protecting property rights. These incentives led to an institutional reversal.

The institutional reversal, combined with the institutions hypothesis, predicts the Reversal of Fortune: relatively rich places got relatively worse economic institutions, and if these institutions are important, we should see them become relatively poor over time. This is exactly what we find with the Reversal of Fortune.

Moreover, the institutions hypothesis is consistent with the timing of the reversal. Recall that the institutions hypothesis links incentives to invest in physical and human capital and in technology to economic institutions, and argues that economic prosperity results from these investments. Therefore, economic institutions should become more important when there are major new investment opportunities. The opportunity to industrialize was the major investment opportunity of the nineteenth century. Countries that are rich today, both among the former European colonies and other countries, are those that industrialized successfully during this critical period.

#### *4.5. Understanding the colonial experience*

The explanation for the reversal that emerges from our discussion so far is one in which the economic institutions in various colonies were shaped by Europeans to benefit themselves. Moreover, because conditions and endowments differed between colonies, Europeans consciously created different economic institutions, which persisted and continue to shape economic performance. Why did Europeans introduce better institutions in previously-poor and unsettled areas than in previously-rich and densely-settled areas? The answer to this question relates to the comparative statics of our theoretical framework. Leaving a full discussion to later, we can note a couple of obvious ideas.

Europeans were more likely to introduce or maintain economic institutions facilitating the extraction of resources in areas where they would benefit from the extraction of resources. This typically meant areas controlled by a small group of Europeans, and areas offering resources to be extracted. These resources included gold and silver, valuable agricultural commodities such as sugar, but most importantly people. In places with a large indigenous population, Europeans could exploit the population, be it in the form of taxes, tributes or employment as forced labor in mines or plantations. This type of colonization was incompatible with institutions providing economic or civil rights to the majority of the population. Consequently, a more developed civilization and a denser population structure made it more profitable for the Europeans to introduce worse economic institutions.

In contrast, in places with little to extract, and in sparsely-settled places where the Europeans themselves became the majority of the population, it was in their interests to introduce economic institutions protecting their own property rights.

#### 4.6. Settlements, mortality and development

The initial conditions we have emphasized so far refer to indigenous population density and urbanization. In addition, the disease environments differed markedly among the colonies, with obvious consequences on the attractiveness of European settlement. As we noted above, when Europeans settled, they established institutions that they themselves had to live under. Therefore, whether Europeans could settle or not had an exogenous effect on the subsequent path of institutional development. In other words, if the disease environment 200 or more years ago affects outcomes today only through its effect on institutions today, then we can use this historical disease environment as an exogenous source of variation in current institutions. From an econometric point of view we have a valid instrument which will enable us to pin down the causal effect of economic institutions on prosperity.<sup>8</sup>

We developed this argument in [Acemoglu, Johnson and Robinson \(2001\)](#) and investigated it empirically. We used initial conditions in the European colonies, particularly data from [Curtin \(1989, 1998\)](#) and [Gutierrez \(1986\)](#) on the mortality rates faced by Europeans (primarily soldiers, sailors, and bishops), as instruments for current economic institutions. The justification for this is that, outside of its effect on economic institutions during the colonial period, historical European mortality has no impact on current income levels. [Figures 14 and 15](#) give scatter plots of this data against contemporaneous economic institutions and GDP per-capita. The sample is countries which were colonized by Europeans in the early modern and modern periods and thus excludes, among others, China, Japan, Korea, Thailand.

[Figure 14](#) shows the very strong relationship between the historical mortality risk faced by Europeans and the current extent to which property rights are enforced. A bivariate regression has an  $R^2$  of 0.26. It also shows that there were very large differences in European mortality. Countries such as Australia, New Zealand and the United States were very healthy with life expectancy typically greater than in Britain. On the other hand mortality was extremely high in Africa, India and South-East Asia. Differential mortality was largely due to tropical diseases such as malaria and yellow fever and at the time it was not understood how these diseases arose nor how they could be prevented or cured.

In [Acemoglu, Johnson and Robinson \(2001\)](#) we showed, using European mortality as an instrument for the current enforcement of property rights, that most of the gap between rich and poor countries today is due to differences in economic institutions. More precisely, we showed (p. 1387) that if one took two typical – in the sense that they both lie on the regression line – countries with high and low expropriation risk, like Nigeria and Chile, then almost the entire difference in incomes per-capita between

<sup>8</sup> Although European mortality is potentially correlated with indigenous mortality, which may determine income today, in practice local populations have developed much greater immunity to malaria and yellow fever. Thus the historical experience of European mortality is a valid instrument for institutional development. See [Acemoglu, Johnson and Robinson \(2001\)](#).

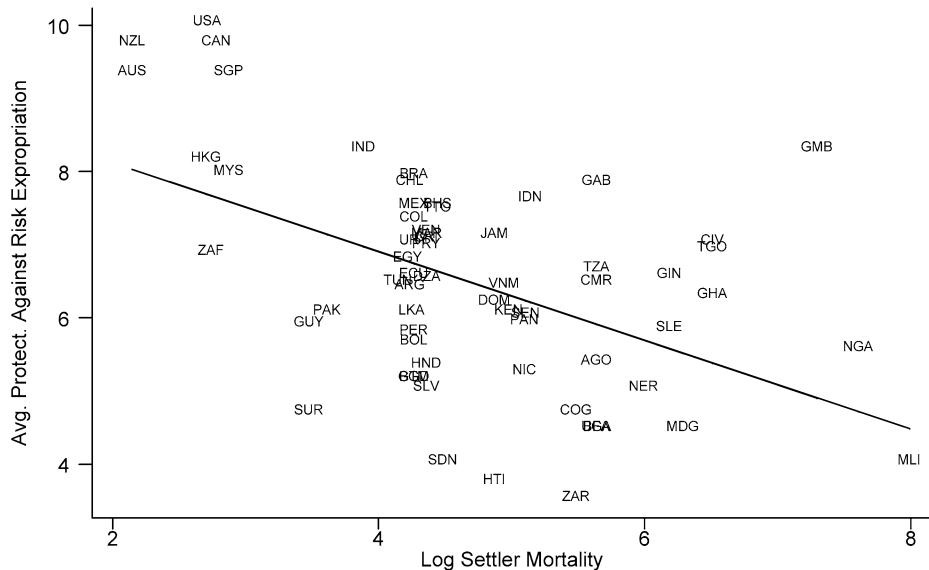


Figure 14. Log mortality of potential European settlers and average protection against risk of expropriation 1985–95.

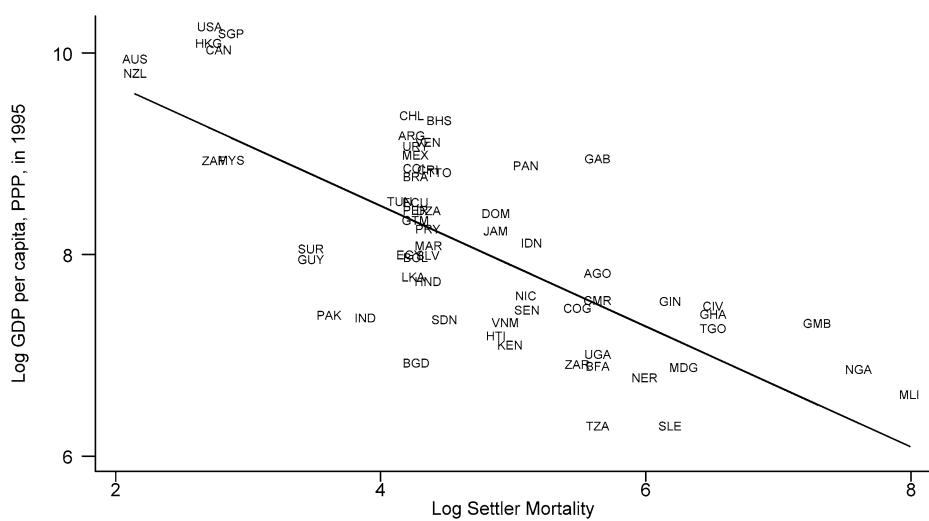


Figure 15. Log mortality of potential European settlers and log GDP per capita in 1995.

them could be explained by the differences in the security of property rights. We also presented regression evidence that showed that once the effect of economic institutions on GDP per-capita was properly controlled for, geographical variables, such as latitude, whether or not a country is land-locked and the current disease environment, have no explanatory power for current prosperity.

These ideas and results provide an interpretation of why there are strong correlations between geographical variables such as latitude and income per-capita. Basically this is because Europeans did not have immunity to tropical diseases during the colonial period and thus settler colonies tended, other things equal, to be created in temperate latitudes. Thus the historical creation of economic institutions was correlated with latitude. Without considering the role of economic institutions it is easy to find a spurious relationship between latitude and income per-capita. However, once economic institutions are properly controlled for, these relationships go away. There is no causal effect of geography on prosperity today, though geography may have been important historically in shaping economic institutions.

What about the role of culture? On the face of it, the Reversal of Fortune is consistent with cultural explanations of comparative growth. The Europeans not only brought new institutions, they also brought their own cultures. There seem to be three main ways to test this idea. First, cultures may be systematically related to the national identity of the colonizing power. For example, the British may have implanted a ‘good’ Anglo-Saxon culture into colonies such as Australia and the United States, while the Spanish may have condemned Latin America by endowing it with a Hispanic or Iberian culture [the academic literature is full of ideas like this, for recent versions see [Vélez \(1994\)](#), [North, Summerhill and Weingast \(2000\)](#) and [Wiarda \(2001\)](#)]. Second, following [Landes \(1998\)](#), Europeans may have had a culture, for example a work ethic or set of beliefs, which was uniquely propitious to prosperity. Finally, following [Weber \(1930\)](#), Europeans also brought different religions with different implications for prosperity. Such a hypothesis could explain why Latin America is relatively poor since its citizens are primarily Roman Catholic, while North America is relatively rich because its citizens are mostly Protestant.

However, the econometric evidence in [Acemoglu, Johnson and Robinson \(2001\)](#) is not consistent with any these views. Once we control properly for the effects of economic institutions, neither the identity of the colonial power, nor the contemporary fraction of Europeans in the population, nor the proportions of the populations of various religions, are significant determinants of income per capita.

These econometric results are supported by historical examples. For instance, with respect to the identity of the colonizing power, in the 17th century the Dutch had perhaps the best domestic economic institutions in the world but the colonies they created in South-East Asia ended up with institutions designed for the extraction of resources, providing little economic or civil rights to the indigenous population.

It is also be clear that the British in no way simply re-created British institutions in their colonies. For example, by 1619 the North American colony of Virginia had a representative assembly with universal male suffrage, something that did not arrive in Britain

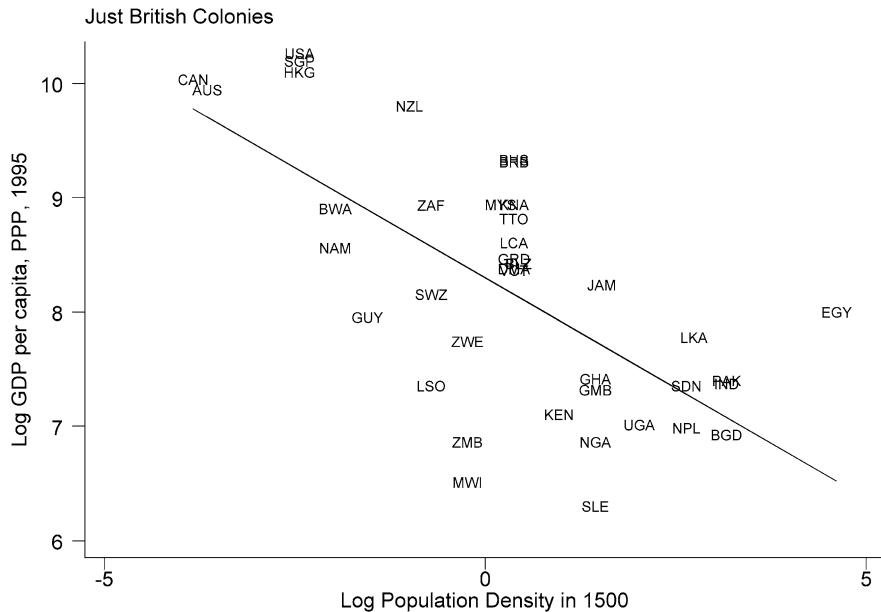


Figure 16. Log population density in 1500 and log GDP per capita in 1995, among former British colonies.

itself until 1919. Another telling example is due to [Newton \(1914\)](#) and [Kupperman \(1993\)](#), who showed that the Puritan colony in Providence Island in the Caribbean quickly became just like any other Caribbean slave colony despite the Puritanical inheritance. Although no Spanish colony has been as successful economically as British colonies such as the United States, it is also important to note that Britain had many unsuccessful colonies (in terms of per capita income), such as in Africa, India and Bangladesh.

To emphasize that the culture or the religion of the colonizer was not at the root of the divergent economic performances of the colonies, [Figure 16](#) shows the reversal among the British colonies (with population density in 1500 on the horizontal axis). Just as in [Figure 6](#), there is a strong negative relationship between population density in 1500 and income per capita today.

With respect to the role of Europeans, Singapore and Hong Kong are now two of the richest countries in the world, despite having negligible numbers of Europeans. Moreover, Argentina and Uruguay have higher proportions of people of European descent than the United States and Canada, but are much less rich. To further document this, [Figure 17](#) shows a similar reversal of fortune for countries where the fraction of those with European descent in 1975 is less than 5 percent of the population.

Overall, the evidence is not consistent with a major role of geography, religion or culture transmitted by the identity of the colonizer or the presence of Europeans. Instead,

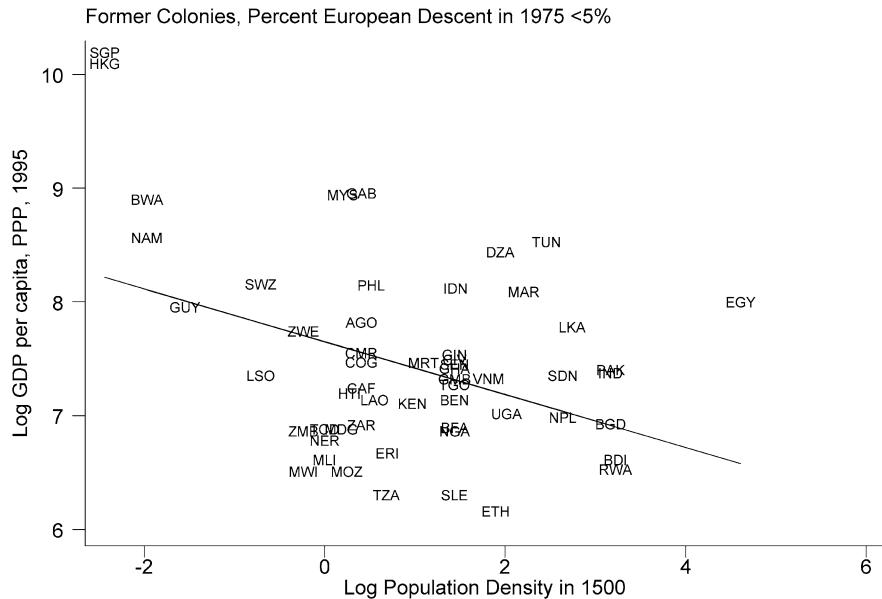


Figure 17. Log population density in 1500 and log GDP per capita in 1995, among former European colonies with current population less than 5% of European descent.

differences in economic institutions appear to be the robust causal factor underlying the differences in income per capita across countries. Institutions are therefore the fundamental cause of income differences and long-run growth.

## 5. Why do institutions differ?

We saw that economic institutions matter, indeed are central in determining relative prosperity. In terms of the different fundamental theories that we discussed, there is overwhelming support for the emphasis of North and Thomas on institutions, as opposed to alternative candidate explanations which emphasize geography or culture. Yet, as we discussed in the introduction, finding that differences in economic institutions can account for the preponderance of differences in per-capita income between countries creates as many questions as it answers. For example, why do countries have different economic institutions? If poor countries are poor because they have bad economic institutions why do they not change them to better institutions? In short, to explain the evidence presented in the last two sections we need a theory of economic institutions. The theory will help to explain the equilibrium set of economic institutions in a particular country and the comparative statics of this theory will help to explain why economic institutions differ across countries.

In the Introduction (Section 1.2), we began to develop such a theory based on social conflict over economic institutions. We have now substantiated the first point we made there, that economic institutions determine prosperity. We must now move to substantiate our second point, that economic institutions must be treated as endogenous and what which economic institutions emerge depends on the distribution of political power in society. This is a key step towards our theory of economic institutions. In the process of substantiating this point however it is useful to step back and discuss other alternative approaches to developing a theory of economic institutions. Broadly speaking, there are four main approaches to the question of why institutions differ across countries, one of which coincides with the approach we are proposing, the social conflict view. We next discuss each of these separately and our assessment as to whether they provide a satisfactory framework for thinking about differences in economic institutions [see [Acemoglu \(2003a\)](#) and [Robinson \(1998\)](#), for related surveys of some of these approaches]. We shall conclude that the approach we sketched in Section 1.2 is by far the most promising one.

### *5.1. The efficient institutions view – the Political Coase Theorem*

According to this view, societies will choose the economic institutions that are socially efficient. How this surplus will be distributed among different groups or agents does not affect the choice of economic institutions. We stress here that the concept of efficiency is stronger than simply Pareto Optimality; it is associated with surplus, wealth or output maximization.

The underlying reasoning of this view comes from the Coase Theorem. [Coase \(1960\)](#) argued that when different economic parties could negotiate costlessly, they will be able to bargain to internalize potential externalities. A farmer, who suffers from the pollution created by a nearby factory, can pay the factory owner to reduce pollution. Similarly, if the current economic institutions benefit a certain group while creating a disproportionate cost for another, these two groups can negotiate to change the institutions. By doing so they will increase the size of the total surplus that they can divide between themselves, and they can then bargain over the distribution of this additional surplus.

Many different versions of the efficient economic institutions view have been proposed. Indeed, assuming that existing economic institutions are efficient is a standard methodological approach of economists, i.e., observing an institution, one tries to understand what are the circumstances that lead it to be efficient. For instance, [Demsetz \(1967\)](#) argued that private property emerged from common property when land became sufficiently scarce and valuable that it was efficient to privatize it. More recently, [Williamson's \(1985\)](#) research, as well as [Coase's \(1937\)](#) earlier work and the more formal analysis by [Grossman and Hart \(1986\)](#), argues that the governance of firms or markets is such as to guarantee efficiency (given the underlying informational and contractual constraints). Williamson argued that firms emerged as an efficient response to contractual problems that plague markets, particularly the fact that there may be ex-post opportunism when individuals make relationship specific investments. Another famous

application of the efficient institutions view is due to [North and Thomas \(1973\)](#) who argued that feudal economic institutions, such as serfdom, were an efficient contract between serfs and lords. The lords provided a public good, protection, in exchange for the labor of the serfs on their lands. In this view, without a modern fiscal system this was an efficient way to organize this exchange. [See [Townsend \(1993\)](#), for a recent version of the idea that other economic institutions of Medieval Europe, such as the open field system, were efficient.]

Williamson and North and Thomas do not specify how different parties will reach agreement to achieve efficient economic institutions, and this may be problematical in the sense that many economic institutions relevant for development are collective choices not individual bargains. There may therefore be free riding problems inherent in the creation of efficient economic institutions. Nevertheless, the underlying idea, articulated by [Becker \(1958\)](#) and [Wittman \(1989\)](#), is that, at least in democracies, competition among pressure groups and political parties will lead to efficient policies and collective choices. In their view, an inefficient economic institution cannot be stable because a political entrepreneur has an incentive to propose a better economic institution and with the extra surplus generated will be able to make himself more attractive to voters. The efficient institutions view regards the structure of political institutions or power as irrelevant. This may matter for the distribution of total surplus, but it will not matter for efficiency itself. The ‘efficient’ set of political institutions is therefore indeterminate.

The notion that a Coasian logic applies in political life as well as in economics is referred to by [Acemoglu \(2003a\)](#) as the Political Coase Theorem. Although the intuition that individuals and groups will strive towards efficient economic outcomes is appealing, there are both theoretical and empirical limits to the Political Coase Theorem. First, as argued by [Acemoglu \(2003a\)](#) and further discussed below, in politics there is an inherent commitment problem, often making the Political Coase Theorem inapplicable.

Second, the Political Coase Theorem does not take us very far in understanding the effect of economic (or indeed political) institutions on economic outcomes – in this view, economic institutions are chosen efficiently, and all societies have the best possible economic institutions given their needs and underlying structures; hence, with the Political Coase Theorem, economic institutions cannot be the fundamental cause of income differences. However, the empirical results we discussed above suggest a major role for such institutional differences.

The only way to understand these patterns is to think of economic institutions varying for reasons other than the underlying needs of societies. In fact, the instrumental variables and natural experiment strategies we exploited above make use precisely of a source of variation unrelated to the underlying needs of societies. For example, South and North Korea did not adopt very different economic systems because they had different needs, but because different systems were imposed on them for other exogenous reasons. In sum, we need a framework for understanding why certain societies consistently end up with economic institutions that are not, from a social point of view, in their best interests. We need a framework other than the Political Coase Theorem.

## 5.2. The ideology view

A second view is that economic institutions differ across countries because of ideological differences – because of the similarity between this and the previous view, [Acemoglu \(2003a\)](#) calls this the Modified Political Coase Theorem. According to this view, societies may choose different economic institutions, with very different implications, because they – or their leaders – disagree about what would be good for the society. According to this approach, there is sufficient uncertainty about the right economic institutions that well-meaning political actors differ about what's good for their own people. Societies where the leaders or the electorate turn out to be right *ex post* are those that prosper. The important point is that, just as with the efficient institutions view, there are strong forces preventing the implementation of policies that are *known* to be bad for the society at large.

Several theoretical models have developed related ideas. For example, [Piketty \(1995\)](#) examined a model where different people have different beliefs about how much effort is rewarded in society. If effort is not rewarded then taxation generates few distortions and agents with such beliefs prefer a high tax rate. On the other hand if one believes that effort is rewarded then low taxes are preferable. Piketty showed that dispersion of beliefs could create dispersion of preferences over tax rates, even if all agents had the same objective. Moreover, incorrect beliefs could be self-fulfilling and persist over time because different beliefs tend to generate information consistent with those beliefs. [Romer \(2003\)](#) also presents a model where voters have different beliefs and showed that if mistakes are correlated, then society can choose a socially inefficient outcome. These models show that if different societies have different beliefs about what is socially efficient they can rationally choose different economic institutions.

Belief differences clearly do play a role in shaping policies and institutions. Several interesting examples of this come from the early experience of independence in former British colonies. For example, it is difficult to explain Julius Nyerere's policies in Tanzania without some reference to his and other leading politicians' beliefs about the desirability of a socialist society. It also appears true that in India the Fabian socialist beliefs of Jawaharlal Nehru were important in governing the initial direction that Indian economic policies took.

Nevertheless, the scope of a theory of institutional divergence and comparative development based on ideology seems highly limited. Can we interpret the differences in institutional development across the European colonies or the divergence in the economic institutions and policies between the North and South of Korea as resulting from differences in beliefs? For example, could it be the case that while Rhee, Park, and other South Korean leaders believed in the superiority of capitalist institutions and private property rights enforcement, Kim Il Sung and Communist Party members in the North believed that communist policies would be better for the country?

In the case of South versus North Korea, this is certainly a possibility. However, even if differences in beliefs could explain the divergence in economic institutions in the immediate aftermath of separation, by the 1980s it was clear that the communist economic

policies in the North were not working. The continued effort of the leadership to cling to these policies, and to power, can only be explained by leaders looking after their own interests at the expense of the population at large. Most likely, North Korean leaders, the Communist Party, and bureaucratic elites are prolonging the current system, which gives them greater economic and political returns than the alternative, even though they fully understand the costs that the system imposes on the North Korean people.

Differences in colonial policies are even harder to explain on the basis of differences in ideology. British colonists established different economic institutions in very different parts of the world: in the Caribbean they set up plantation societies based on slavery, supported by highly oppressive economic institutions. In contrast, the economic institutions that developed in areas where the British settled, and where there was no large population of indigenous people to be captured and put to work, and where slavery could not be profitably used, such as northeastern United States, Canada, Australia and New Zealand, were very different. Moreover, differences in the incentives of the colonists in various colonies are easy to understand: when they did not settle, they were choosing economic institutions simply to extract resources from the native population. When they settled in large numbers, economic institutions and policies emerged in order to protect them in the future and encourage investment and prosperity.

These considerations make us tend towards a view which emphasizes the actions of key economic and political agents that are taken rationally and in recognition of their consequences, not simply differences in beliefs. We do not deny that belief differences and ideology often play important roles but we do not believe that a satisfactory theory of institutional differences can be founded on differences in ideology.

### 5.3. The incidental institutions view

The efficient institutions view is explicitly based on economic reasoning: the social costs and benefits of different economic institutions are weighed against each other to determine which economic institutions should prevail. Efficiency arises because individuals ultimately calculate according to social costs and benefits. Institutions are therefore choices. A different approach, popular among many political scientists and sociologists, but also some economists, is to downplay choices and to think of institutions, both economic and political, as the by-product or unintended consequence of other social interactions or historical accidents. In other words, historical accidents at critical junctures determine institutions, and these institutions persist for a long time, with significant consequences.

Here, we discuss two such theories. The first is the theory of political institutions developed by [Moore \(1966\)](#) in his *Social Origins of Dictatorship and Democracy*, the second is the recent emphasis in the economics literature on legal origins, for example as in the work of Shleifer and his co-authors [[La Porta et al. \(1998, 1999\)](#), [Djankov et al. \(2002, 2003\)](#), [Glaeser and Shleifer \(2002\)](#)].

Moore attempted to explain the different paths of political development in Britain, Germany and Russia. In particular, he investigated why Britain evolved into a democ-

racy, while Germany succumbed to fascism and Russia had a communist revolution. Moore stressed the extent of commercialization of agriculture and resulting labor relations in the countryside, the strength of the ‘bourgeoisie’, and the nature of class coalitions. In his theory, democracy emerged when there was a strong, politically assertive, commercial middle class, and when agriculture had commercialized so that there were no feudal labor relations in the countryside. Fascism arose when the middle classes were weak and entered into a political coalition with landowners. Finally, a communist revolution resulted when the middle classes were non-existent, agriculture was not commercialized, and rural labor was repressed through feudal regulations. In Moore’s theory, therefore, class coalitions and the way agriculture is organized determine which political institutions will emerge. However, the organization of agriculture is not chosen with an eye to its effects on political institutions, so these institutions are an unintended consequence. Although Moore was not explicitly concerned with economic development, it is a direct implication of his analysis that societies may end up with institutions that do not maximize income or growth, for example, when they take the path to communist revolution.

Beginning with the work on shareholder rights [La Porta et al. (1998)], continuing to the efficiency of government [La Porta et al. (1999)] and more recently the efficiency of the legal system [Djankov et al. (2003)], Shleifer and his co-authors have argued that a central source of variation in many critical economic institutions is the origin of the legal system. For example, “Civil laws give investors weaker legal rights than common laws do, independent of the level of per-capita income. Common-law countries give both shareholders and creditors – relatively speaking – the strongest, and French-civil-law countries the weakest, protection” [La Porta et al. (1998, p. 1116)].

These differences have important implications for resource allocation. For example, when shareholders have poor protection of their rights, ownership of shares tends to be more highly concentrated. Djankov et al. (2003) collected a cross-national dataset on how different countries legal systems dealt with the issue of evicting a tenant for non-payment of rent and collecting on a bounced check. They used these data to construct an index of procedural formalism of dispute resolution for each country and showed that such formalism was systematically greater in civil than in common law countries, and is associated with higher expected duration of judicial proceedings, less consistency, less honesty, less fairness in judicial decisions, and more corruption. Legal origins therefore seems to matter for important institutional outcomes.

Where do legal origins come from? The main argument is that they are historical accidents, mostly related to the incidence of European colonialism. For example, Latin American countries adopted the Napoleonic codes in the nineteenth century because these were more compatible with their Spanish legal heritage. Importantly, the fact that Latin American countries therefore have ‘French legal origin’ is due to a historical accident and can be treated as exogenous with respect to current institutional outcomes. What about the difference between common law and civil law? Glaeser and Shleifer (2002) argue that the divergence between these systems stems from the medieval period

and reflects the balance of power between the lords and the king in England and France. Once these systems established, they persisted long after the initial rationale vanished.

Although we believe that historical accidents and persistence are important, in reality the aspect of choice over institutions seems too important to be denied. Even if institutions have a tendency to persist, their persistence is still a choice, in the sense that if the agents decided to change institutions, change would be possible. There are important examples from history of countries radically changing their legal systems such as in Japan after the Meiji restoration, Russia after the Crimean War, and Turkey under Mustafa Kemal in the 1920's. Another example might be central planning of the economy. Though many countries adopted this way of organizing the economy some abandoned it while others, such as North Korea and Cuba, still maintain it. The point here is that though institutions may in some circumstances be the incidental outcome of history, at some point people will start to ask why society has the institutions that it does and to consider other alternatives. At this point we are back in the realm of choice.

#### *5.4. The social conflict view*

According to this view, economic (and political) institutions are not always chosen by the whole society (and not for the benefit of the whole society), but by the groups that control political power at the time (perhaps as a result of conflict with other groups). These groups will choose the economic institutions that maximize their own rents, and the economic institutions that result may not coincide with those that maximize total surplus, wealth or income. For example, economic institutions that enforce property rights by restricting state predation may not be in the interest of a ruler who wants to appropriate assets in the future. By establishing property rights, this ruler would be reducing his own future rents, so may well prefer economic institutions other than enforced private property. Therefore, equilibrium economic institutions will not be those that maximize the size of the overall pie, but the slice of the pie taken by the powerful groups.

The first systematic development of this point of view in the economics literature is [North \(1981\)](#), who argued in the chapter on “A Neoclassical Theory of the State” that agents who controlled the state should be modeled as self-interested. He then argued that the set of property rights that they would choose for society would be those that maximized their payoff and because of ‘transactions costs’, these would not necessarily be the set that maximized social welfare. One problem with North’s analysis is that he does not clarify what the transactions costs creating a divergence between the interests of the state and the citizens are. Here, we will argue that commitment problems are at the root of this divergence.

The notion that elites, i.e., the politically powerful, may opt for economic institutions which increase their incomes, often at the expense of society, is of course also present in much of the Marxist and dependency theory literature. For example, [Dobb \(1948\)](#), [Brenner \(1976, 1982\)](#) and [Hilton \(1981\)](#) saw feudalism, contrary to [North and Thomas's \(1973\)](#) model, as a set of institutions designed to extract rents from the peasants at the

expense of social welfare.<sup>9</sup> Dependency theorists such as Williams (1944), Wallerstein (1974–1980), Rodney (1972), Frank (1978) and Cardoso and Faletto (1979) argued that the international trading system was designed to extract rents from developing countries to the benefit of developed countries.

The social conflict view includes situations where economic institutions may initially be efficient for a set of circumstances but are no longer efficient once the environment changes. For example, Acemoglu, Aghion and Zilibotti (2002) show that though certain sorts of organizations may be useful for countries a long way from the technological frontier, it may be socially efficient to change them subsequently. This may not happen however because it is not privately rational. An interesting example may be the large business enterprises (the *chaebol*) of South Korea. In the context of political institutions, one might then develop a similar thesis. Certain sets of institutions are efficient for very poor countries but they continue to exist even after they cease to be the efficient institutional arrangement.

In stark contrast to the efficient institutions view, political institutions play a crucial role in the social conflict view. Which economic institutions arise depends on who has political power to create or block different economic institutions. Since political institutions play a central role in the allocation of such power they will be an intimate part of a social conflict theory of economic institutions.

What distinguishes the social conflict view from the ideological view is that social conflict can lead to choices of economic institutions which cause underdevelopment even when all agents have common knowledge that this is so. What distinguishes it from the incidental view is that it emphasizes that institutional choices which cause underdevelopment are conscious choices, rather than the result of some historical accident. The aspect that distinguishes the social conflict view from the efficient institutions view is that it does not assume that institutions are always efficient. This is one possible outcome but it is not the only one or indeed the most likely. Why is this? Why cannot efficiency be separated from distribution? We discuss this issue in the next section.

## 6. Sources of inefficiencies

Having motivated our first two assertions in Section 1.2, we are now in a position to discuss the third, related to the importance of commitment problems. The inability to commit to how political power will be used in the future means that the impact of economic institutions on efficiency cannot be separated from their effects on distribution.<sup>10</sup>

In any market situation where economic exchange takes place, and the quid is separated from the pro quo, issues of commitment will arise. That these issues are of crucial

<sup>9</sup> Postan (1966, pp. 603–604) famously estimated that lords extracted about 50% of the entire production of peasants.

<sup>10</sup> An alternative approach would be to stress informational asymmetries [Farrell (1987)].

importance has been recognized in the literatures on incomplete contracts and renegotiation [e.g., [Hart \(1995\)](#)]. Nevertheless, if the legal system functions properly, there is an array of enforceable contracts that owners can sign with managers, workers with employers, borrowers with lenders, etc. These contracts can be enforced because there is an authority, a third party, with the power to enforce contracts. Although the authority that is delegated to enforce contracts and to resolve disputes varies depending on the exact situation, all such power ultimately emanates from the state, which, in modern society, has a near-monopoly on the use of legitimate coercion. An owner and manager can write a contract because they believe that the state, and its agents the courts, would be impartial enforcers of the contract.

In contrast, if, for example, a manager believed that the state would be aligned with the interests of the owner and refuse to punish the owner if and when he failed to make a payment stipulated by the contract, then the contract would have little value. Therefore, the presence of an impartial enforcer is important for contracting. The problem when it comes to institutional choices is that there is no such impartial third party that can be trusted to enforce contracts. This is the origin of the commitment problem in politics.<sup>11</sup>

To elaborate on this point, let us consider a situation where society can be governed as a dictatorship or as a democracy. Imagine that the dictator does not relinquish his power, but instead he promises that he will obey the rules of democracy, so that individuals can undertake the same investments as they would in democracy. This promise would not necessarily be credible. As long as the political system remains a dictatorship, there is no higher authority to make the dictator stick to his promise. There is no equivalent of a contract that can be enforced by an impartial third-party. After all, the dictator has the monopoly of military and political power, so he is the final arbiter of conflicting interests. There is no other authority to force the dictator to abide by his promises.

A similar problem plagues the reverse solution, whereby the dictator agrees to a voluntary transition to democracy in return for some transfers in the future to compensate him for the lost income and privileges. Those who will benefit from a transition to democracy would be willing to make such promises, but once the dictator relinquishes his political power, there is no guarantee that citizens would agree to tax themselves in order to make payments to this former dictator. Promises of compensation to a former dictator are typically not credible.

The essence of the problem is commitment. Neither party can commit to compensate the other nor can they commit to take actions that would not be in their interests ex post. The reason why commitment problems are severe in these examples is because we are

<sup>11</sup> Many scholars have emphasized the fact that a key feature of political economy is that there is no third party which can enforce the promises made by the state and that this leads to problems of commitment and endemic inefficiencies. This idea is discussed by [North \(1990\)](#) and [Olson \(1993\)](#), is central to the work of [Weingast \(1997, 1998\)](#) and is implicit in many other studies. See also [Grossman and Noh \(1994\)](#), [Dixit \(1996\)](#), [Dixit and Londregan \(1995\)](#), [Besley and Coate \(1998\)](#) and [Powell \(2004\)](#) for discussions of how inability to commit generates inefficiencies in political outcomes.

dealing with political power. Different institutions are associated with different distributions of political power, and there is no outside impartial party with the will and the power to enforce agreements. In some cases, there may be self-enforcing promises that maintain an agreement. [Acemoglu \(2003a\)](#) discusses such possibilities, but in general, there are limits to such self-enforcing agreements, because they require the participants to be sufficiently patient, and when it comes to matters of political power, the future is uncertain enough that no party would behave in a highly patient manner.

Based on this reasoning, we can now discuss three different channels via which the presence of commitment problems will lead to the choice and persistence of inefficient institutions.

### *6.1. Hold-up*

Imagine a situation in which an individual or a group holds unconstrained political power. Also suppose that productive investments can be undertaken by a group of citizens or producers that are distinct from the “political elites”, i.e., the current power holders. The producers will only undertake the productive investments if they expect to receive the benefits from their investments. Therefore, a set of economic institutions protecting their property rights are necessary for investment. Can the society opt for a set of economic institutions ensuring such secure property rights? The answer is often no (even assuming that “society” wants to do so).

The problem is that the political elites – those in control of political power – cannot commit to respect the property rights of the producers once the investment are undertaken. Naturally, *ex ante*, before investments are undertaken, they would like to promise secure property rights. But the fact that the monopoly of political power in their hands implies that they cannot commit to not hold-up producers once the investments are sunk.

This is an obvious parallel to the hold-up problem in the theory of the firm, where once one of the parties in a relationship has undertaken investments specific to the relationship, other parties can hold her up, and capture some of the returns from her investments. As in the theory of the firm, the prospect of hold-up discourages investment. But now the problem is much more severe, since it is not only investments that are specific to a relationship that are subject to hold-up, but all investments.

This is therefore an example of how inefficient economic institutions arise because of a monopoly of political power. Those with political power cannot commit not to use their political power *ex post*, and this translates directly into a set of economic institutions that do not provide secure property rights to groups without political power. The consequence is clear: without such protection, productive investments are not undertaken, and opportunities for economic growth go unexploited.

The reason why these inefficient economic institutions persist (or may be the equilibrium institutions of the society) is related to commitment problems. Parallel to our above example of inducing the dictator to relinquish power, there are two ways to introduce secure property rights. First, in principle, political elites could promise to respect property rights. However, mere promises would not be credible, unless backed up by

the political elites relinquishing power, and this would mean relinquishing their rents and privileges. Second, political elites can be bought off by the beneficiaries of a system of more secure property rights. This would typically be achieved by a promise of future payments. For example, after investments are undertaken and output is produced, a share can be given to the political elites. But, as pointed out above, there is another, reverse commitment problem here; the beneficiaries of the new regime cannot commit to make the promised payments to the previous political elites.

Many real world examples illustrate the commitment problems involved in limiting the use of political power. In practice, although buying off dictators and persuading them to leave power is difficult, there have been many attempts to do so, usually by trying to guarantee that they will not be persecuted subsequently. One way of doing this is to give them asylum in another country. Nevertheless, such attempts rarely succeed, most likely again because of commitment problems (the new regime cannot commit to abide by its promises). An illustrative example of this is the attempts by the Reagan administration to persuade Jean-Claude ('Baby Doc') Duvalier to relinquish power in Haiti in 1986. In the face of a popular uprising and rising social and economic chaos, the Reagan administration, via the intermediation of the Jamaican Prime Minister Edward Seaga, tried to persuade Duvalier to go into exile. He at first agreed and the White House announced his departure on January 30th, but the next day he changed his mind, unsure that he would really be protected, and stayed in Haiti. One month later he was forced into exile in France by the military.

A more common, and in many ways more interesting strategy to induce dictators to relinquish power is to try to structure political institutions so as to guarantee that they will not be punished. Such institutional changes are sometimes important in transitions to democracy. For example, President Pinochet was willing to abide by the results of the 1989 plebiscite he lost in Chile because as a senator the Constitution protected him from prosecution. It was only when he left the country that he was vulnerable.

Although Pinochet's experience illustrates an example of structuring political institutions to achieve commitment, to create durable institutions constraining future use of political power is difficult in practice. These difficulties are well illustrated by the transition from white rule in Rhodesia to majority rule in Zimbabwe. Facing an un-winnable guerrilla war, the white elite in Rhodesia sought to negotiate a transition of majority rule, but with enough institutional safeguards that their rents would be protected. These safeguards included the electoral system they wanted, which was used for the first post-independence elections, and massive over-representation in parliament [Reynolds (1999, p. 163)]. Whites were guaranteed 20% of the seats in the legislature for seven years despite making up only 2–3% of the population and were guaranteed 10 seats of the 40 seat senate. Clauses of the 1980 Constitution were also aimed at directly guaranteeing the property rights of the whites. In particular land reform was outlawed for 10 years after which it could only take place if compensated.

The white negotiators at the Lancaster House talks in 1979 that produced these agreements understood that any promises made by the black majority negotiators about what would happen after independence could not be believed. They sought therefore to find

a set of rules that would get around this problem [Herbst (1990, pp. 13–36)]. Nevertheless, these guarantees were not enough to protect the property rights (and rents) of the whites in anything other than the short run. The Mugabe regime quickly absorbed the other factions from among the African guerrilla opposition, and more moderate relatively pro-white groups, such as Abel Muzorewa's United African National Council, crumbled. In 1985 the Mugabe regime switched back to the electoral system it preferred [Reynolds (1999, p. 164)] and in 1987, at the first possible opportunity, it removed the guaranteed representation for whites. Though in 1987 Mugabe nominated white candidates for these seats [Horowitz (1991, pp. 135–136)], this did not last for long. In 1990 the senate was abolished. Finally, in 1990 the Constitution was amended to allow for the redistribution of land. Since this time the Mugabe government has begun a sustained policy of land redistribution away from whites through legal and extra-legal means.

## *6.2. Political losers*

Another related source of inefficient economic institutions arises from the desire of political elites to protect their political power. Political power is the source of the incomes, rents, and privileges of the elite. If their political power were eroded, their rents would decline. Consequently, the political elite should evaluate every potential economic change not only according to its economic consequences, such as its effects on economic growth and income distribution, but also according to its political consequences. Any economic change that will erode the elite's political power is likely to reduce their economic rents in the long run.

As an example, imagine a change in economic institutions that will increase economic growth, but in doing so, will also enrich groups that could potentially contest political power in the future. Everything else equal, greater economic growth is good for those holding political power. It will create greater returns on the assets that they possess, and also greater incomes that they can tax or expropriate. However, if their potential enemies are enriched, this also means greater threats against their power in the future. Fearing these potential threats to their political power, the elites may oppose changes in economic institutions that would stimulate economic growth.

That the threat of becoming a political loser impedes the adoption of better institutions is again due to a commitment problem. If those who gained political power from institutional change could promise to compensate those who lost power then there would be no incentive to block better institutions.

There are many historical examples illustrating how the fear of losing political power has led various groups of political and economic elites to oppose institutional change and also the introduction of new technologies. Perhaps the best documented examples come from the attitude of the elites to industrialization during the nineteenth century [see Acemoglu and Robinson (2000b, 2002)]. There were large differences between the rates at which countries caught up with British industrialization with many countries completely failing to take advantage of the new technologies and opportunities. In most of these cases, the attitudes of political elites towards industrialization, new technology

and institutional change appear to have been the decisive factor, and these attitudes were driven by their fears of becoming political losers. These issues are best illustrated by the experiences of Russia and Austria-Hungary.

In both Russia and Austria-Hungary, absolutist monarchies feared that promoting industrialization would undermine their political power. In Russia, during the reign of Nikolai I between 1825 and 1855 only one railway line was built in Russia, and this was simply to allow the court to travel between Moscow and St. Petersburg. Economic growth and the set of institutions that would have facilitated it were opposed since, as [Mosse \(1992, p. 19\)](#) puts it “it was understood that industrial development might lead to social and political change”. In a similar vein, [Gregory \(1991, p. 74\)](#) argues: “Prior to the about face in the 1850s, the Russian state feared that industrialization and modernization would concentrate revolution minded workers in cities, railways would give them mobility, and education would create opposition to the monarchy”.

It was only after the defeat in the Crimean War that Nikolai’s successor, Alexander II, initiated a large scale project of railway building and an attempt to modernize the economy by introducing a western legal system, decentralizing government, and ending feudalism by freeing the serfs. This period of industrialization witnessed heightened political tensions, consistent with the fears of the elites that times of rapid change would destabilize the political status quo and strengthen their opposition [[McDaniel \(1991\)](#) gives a detailed account of these events, see also [Mosse \(1958\)](#)].

The consensus view amongst historians also appears to be that the main explanation for the slow growth of Austria-Hungary in the nineteenth century was lack of technology adoption and institutional change, again driven by the opposition of the state to economic change. This view was proposed by Gerschenkron who argued that the state not only failed to promote industrialization, but rather, “economic progress began to be viewed with great suspicion and the railroads came to be regarded, not as welcome carriers of goods and persons, but as carriers of the dreaded revolution. Then the state clearly became an obstacle to the economic development of the country” ([1970, p. 89](#)). See also [Gross \(1973\)](#).

The analysis of [Freudenberger \(1967, pp. 498–499\)](#) is similar. As with the Tsar, the Hapsburg emperors opposed the building of railways and infrastructure and there was no attempt to develop an effective educational system. [Blum \(1943\)](#) pointed to the pre-modern institutional inheritance as the major blockage to industrialization arguing (p. 26) that

“these living forces of the traditional economic system were the greatest barrier to development. Their chief supporter was . . . Emperor Francis. He knew that the advances in the techniques of production threatened the life of the old order of which he was so determined a protector. Because of his unique position as final arbiter of all proposals for change he could stem the flood for a time. Thus when plans for the construction of a steam railroad were put before him, he refused to give consent to their execution ‘lest revolution might come into the country’.”

### 6.3. Economic losers

A distinct but related source of inefficiency stems from the basic supposition of the social conflict view that different economic institutions imply different distributions of incomes. This implies that a move from a bad to a better set of economic institutions will make some people or groups worse off (and will not be Pareto improving). This in turn implies that such groups will have an incentive to block or impede such institutional changes even if they benefit the whole of society in some aggregate sense.

The idea that economic losers impede the choice of efficient economic institutions and economic policies is widespread in economics and was seen earliest in the literature on international trade. Even though free trade may be socially desirable, individuals invested in sectors in which an economy does not enjoy comparative advantage will lose economically from free trade. Since at least the work of [Schattschneider \(1935\)](#) the role of economic losers has been central in understanding why free trade is not adopted. In the context of development economics, this idea was first discussed by [Kuznets \(1968\)](#), developed at length by [Olson \(1982, 2000\)](#) and [Mokyr \(1990\)](#), and formalized by [Krusell and Rfos-Rull \(1996\)](#) and [Parente and Prescott \(1999, 2005\)](#). Most of the examples discussed in the development literature on economic losers are about technological change – people with specific investments in obsolete technology try to block the introduction of better technology. The most celebrated example is the case of the Luddites, skilled weavers in early nineteenth century England who smashed new mechanized looms which threatened to lead to massive cuts in their wages [see [Thomis \(1970\), Randall \(1991\)](#)]. [Scott \(2000, p. 200\)](#) relates a similar example from modern Malaysia, “When, in 1976, combine harvesters began to make serious inroads into the wages of poor villagers, the entire region experienced a rash of machine-breaking and sabotage reminiscent of the 1830s in England”.

That better economic institutions are blocked by individuals whose incomes are threatened by such change is again due to a problem of commitment. If those whose incomes rose when economic institutions changed could promise to compensate those whose incomes fell then there would be no incentive to block better economic institutions. Nevertheless, it is difficult to commit to such transfers. To consider again the example of the Luddites, the factory owners could have promised to pay the weavers high wages in the future even though their skills were redundant. Once the new technology was in place however, owners would have a clear incentive to fire the weavers and hire much cheaper unskilled workers.<sup>12</sup>

Although the problem of economic losers is appealing at first sight, has received some attention in the economics literature, and fits into our framework by emphasizing the importance of commitment problems, we view it both theoretically and empirically

<sup>12</sup> One possible way round this problem would be for the owners, if they could afford it, to compensate the weavers in advance for their lower future wages. But this would raise the reverse commitment problem: the weavers would have an incentive to take the money and still break the machines – i.e., they could not commit to not blocking the innovations that would reduce their wages even after they had taken the money.

less important than the holdup and the political loser problems. First, as pointed out in [Acemoglu and Robinson \(2000b\)](#), in theories emphasizing issues of economic losers, there are implicit assumptions about politics, which, when spelled out, imply that political concerns must be important whenever issues of economic losers are present. The idea of economic losers is that certain groups, fearing that they will lose their economic rents, prevent adoption of beneficial economic institutions or technologies. The assumption in this scenario is that these groups have the political power to block socially beneficial changes. But then, if they have the political power to block change, why would not they allow the change to take place and then use their political power to redistribute some of the gains to themselves? The implicit assumption must therefore be that groups losing economically also experience a reduction in their political power, making it impossible for them to redistribute the gains to themselves after the change takes place. This reasoning therefore suggests that whether certain groups will lose economically or not is not as essential to their attitudes towards change as *whether their political power will be eroded*. Problems of political losers therefore seem much more important than problems of economic losers.

Possibly for this reason, advocates of the economic losers view have been unable to come up with any well documented examples where the economic losers hypothesis can actually explain first-order patterns of development. For instance, while it is true that the Luddites tried to break machines, they singularly failed to halt the progress of agricultural technology in nineteenth century Britain. The same is true for Malaysia in the 1970s, one of the fastest growing economies in the world at that time. Neither set of workers had sufficient political power to stop change. Indeed, when political powerful groups became economic losers, such as landowners in nineteenth century England who saw land prices and agricultural rents fall rapidly after 1870, they did nothing to block change because their political power allowed them to benefit from efficient economic institutions [\[Acemoglu and Robinson \(2002\)\]](#).

Perhaps the most interesting failure of economic losers to halt progress in English economic history comes from the impact of the enclosure of common lands. Land has not always been privately owned as property. In much of Africa land is still owned communally, rather than individually, and this was true in Medieval Britain. Starting around 1550 however an ‘enclosure movement’ gathered pace where ‘common land’ was divided between cultivators and privatized. By 1850 this process of enclosures had made practically all of Britain private property.

Enclosure was a heterogenous process [\[Overton \(1996, p. 147\)\]](#) and it also took place at different times in different places. Nevertheless, most of it was in two waves, the so called ‘Tudor enclosures’ between 1550 and 1700 and the ‘parliamentary enclosures’ in the century after 1750.

“From the mid-eighteenth century the most usual way in which common rights were removed was through a specific act of parliament for the enclosure of a particular locality. Such acts . . . made the process easier because enclosure could be secured provided the owners of a majority (four fifths) of the land, the lord of the manor, and the owner of the tithe agreed it should take place. Thus the law of

parliament (statue law) only took account of the wishes of those *owning* land as opposed to the common law which took account of all those who had both ownership rights and *use* rights to land. Moreover ... in some parishes the ... majority could be held by a single landowner ... parliamentary enclosure often resulted in a minority of owners imposing their will on the majority of farmers” [Overton (1996, p. 158), *italics in original*].

The historical evidence is unanimous that the incentive to enclose was because “enclosed land was worth more than open common field land ... the general consensus has been that rents doubled” [Overton (1996, p. 162)]. More controversial is the source of this increase in rent. Overton continues (pp. 162–163) “The proportion of profits taken as rents from tenants by landlords is the outcome of a power struggle between the two groups, and the increase in rent with enclosure may simply reflect an increase in landlord power”. Allen (1982, 1992) showed, in his seminal study of the enclosure movement in the South Midlands, that the main impact was a large increase in agricultural rents and a redistribution of income away from those cultivators who had previously used the commons.

The enclosure of common land thus led to a huge increase in inequality in early modern England. Many peasants and rural dwellers had their traditional property rights expropriated. In protest, groups of citizens dispossessed by enclosure attempted to oppose it through collective action and riots – attempting to influence the exercise of political power. These groups were no match for the British state, however. Kett’s rebellion of 1549, the Oxfordshire rebellion of 1596, the Midland Revolt of 1607, and others up to the Swing Riots of 1830–1831 were all defeated [see Charlesworth (1983)]. The presence of economic losers did not prevent this huge change in economic institutions and income distribution.

#### *6.4. The inseparability of efficiency and distribution*

Commitment problems in the use and the allocation of political power therefore introduce a basic trade-off between efficiency and distribution. For example, when lack of commitment causes hold-ups, those who hold political power know that people will not have the right incentives to invest so growth will be low. In response to this, they might voluntarily give away their power or try to create political institutions that restricted their power. Such a change in political institutions would create better investment incentives. Though this situation is hypothetically possible and has formed the basis for some theories of institutional change [e.g., Barzel (2001)] it appears to be insignificant in reality. Even faced with severe underinvestment, political elites are reluctant to give away their power because of its *distributional* implications, i.e., because this would reduce their ability to extract rents from the rest of society. Thus poor economic institutions, here lack of property rights and hold-up, persist in equilibrium because to solve the problem, holders of political power have to voluntarily constrain their power or give it away. This may increase the security of property in society and increase incentives to

invest, but it also undermines the ability of rulers to extract rents. They may be better off with a large slice of a small pie.

Similar phenomena are at work when there are either political or economic losers. In the first case, namely a situation where political power holders anticipate being political losers, promoting good institutions directly reduces the political power and rents of incumbents and a similar trade-off emerges. Adopting efficient economic institutions will stimulate growth, but when the political status quo is simultaneously eroded the amount of rent accruing to the initially powerful may fall. In the second case, the incomes of those with political power to determine economic institutions falls directly when better economic institutions are introduced. In the absence of credible commitments to side-payments, those whose incomes fall when better economic institutions are introduced have an incentive to block such institutions.

Because commitment problems seem so endemic in collective choice and politics, it seems natural to believe that institutional change has significant distributional consequences and as a result there will be conflict over the set of institutions in society.

### 6.5. Comparative statics

Our analysis so far has made some progress towards our theory of differences in economic institutions. Although our full theory is yet to be developed in the later sections, the different mechanisms discussed in this section already point out the major comparative static implications of our approach regarding when economic institutions protecting the property rights of a broad cross-section of society are likely to be adopted, and when they are likely to be opposed and blocked. We now briefly discuss these comparative statics.

Hold-up, political loser and economic loser considerations lead to some interesting comparative static results which can be derived by considering the political institutions that lie behind these phenomena.

1. First, the perspective of hold-ups immediately suggests that situations in which there are constraints on the use of political power, for example, because there is a balance of political power in society or a form of separation of powers between different power-holders, are more likely to engender an environment protecting the property rights of a broad cross-section of society. When political elites cannot use their political power to expropriate the incomes and assets of others, even groups outside the elite may have relatively secure property rights. Therefore, constraints and checks on the use of political power by the elite are typically conducive to the emergence of better economic institutions
2. Second, a similar reasoning implies that economic institutions protecting the rights of a broad cross-section are more likely to arise when political power is in the hands of a relatively broad group containing those with access to the most important investment opportunities. When groups holding political power are narrower, they may protect their own property rights, and this might encourage their own

investments, but the groups outside the political elites are less likely to receive adequate protection for their investments [see [Acemoglu \(2003b\)](#)].

3. Third, good economic institutions are more likely to arise and persist when there are only limited rents that power holders can extract from the rest of society, since such rents would encourage them to opt for a set of economic institutions that make the expropriation of others possible.
4. Finally, considerations related to issues of political losers suggest that institutional reforms that do not threaten the power of incumbents are more likely to succeed. Therefore, institutional changes that do not strengthen strong opposition groups or destabilize the political situation are more likely to be adopted.

### *6.6. The colonial experience in light of the comparative statics*

We now briefly return to the colonial experience, and discuss how the comparative statics discussed here shed light on the differences in economic institutions across the former colonies and the institutional reversal.

The second comparative static result above suggests a reason why better economic institutions developed in places where Europeans settled. In these societies, a relatively broad-based group of Europeans came to dominate political power, and they opted for a set of economic institutions protecting their own property rights. In contrast, in places where Europeans did not settle, especially where they were a small minority relative to a large indigenous population, they did not have the incentives to develop good economic institutions because such institutions would have made it considerably more difficult for them to extract resources from the rest of society.

The third comparative static suggests an important reason why in places with more wealth, resources and also a high density of indigenous population to be exploited, Europeans were more likely to opt for worse institutions, without any protection for the majority of the population, again because such institutions facilitated the extraction of resources by the Europeans.

The first comparative static result, in turn, is related to the persistence of the different types of economic institutions that Europeans established, or maintained, in different colonies. In colonies where Europeans settled in large numbers, they also developed political institutions placing effective checks on economic and political elites. In contrast, the political institutions in colonies with high population density, extractive systems of production, and few Europeans, concentrated power in the hands of the elite, and built a state apparatus designed to use coercion against the majority of the population. These different political institutions naturally implied different constraints on political and economic elites. In the former set of colonies, there were constraints on the development of economic institutions that would favor a few at the expense of the majority. Such constraints were entirely absent in the latter set of colonies.

Finally, the fourth comparative static is useful in thinking about why many colonies did not attempt to change their economic institutions during the nineteenth century when new economic opportunities made their previous system based on forced labor, slavery,

or tribute-taking much less beneficial relative to one encouraging investment in industry and commerce. Part of the answer appears to lie in the fact that the political power of the elites, for example of the plantation owners in the Caribbean, was intimately linked to the existing economic system. A change in the economic system would turn them into political losers, an outcome they very much wanted to avoid.

### 6.7. Reassessment of the social conflict view

So far we have shown that the econometric evidence is convincing that differences in economic institutions are the root cause of differences in prosperity. We then argued that although there are different approaches which can account for variation in economic institutions, the most plausible approach is the social conflict view. Though we believe that there certainly are instances where history and ideology matter for the institutional structure of society, and clearly institutions are highly persistent, the most promising approach to understanding why different countries have different institutions is to focus on choices and their subsequent consequences. The social conflict view emphasizes the distributional implication of economic institutions and how commitment problems imply that efficiency and distribution cannot be separated. Hence the fundamental conflict within society over the nature of economic institutions has important implications for economic performance. Some economic institutions will promote growth, but they will not necessarily benefit all groups in society. Alternative economic institutions may induce economic stagnation, but may nevertheless enrich some groups. Which set of institutions results and whether or not a society prospers will be determined by which of these groups has the political power to get the institutions that differentially benefit them. At this point we have therefore substantiated the first three points we made in the introduction. To develop our theory of economic institutions further we need to be more specific about political power – where it comes from and why some people have it and not others. We undertake this task in Section 8. Before doing this however the next section discusses three important historical examples of the evolution of economic institutions. We use these examples to show the explanatory power of the social conflict view and to begin to illustrate in concrete settings how political power works.

## 7. The social conflict view in action

We now discuss three important examples to bring out the fact that conflict over economic institutions is critical to the functioning of the economy and that this conflict stems, not from differences in beliefs, ideology or historical accidents, but from the impact of economic institutions on distribution. The examples also show that those with political power have a disproportionate effect on economic institutions and they illustrate how the distribution of political power is influenced by different factors. These factors include the allocation of de jure political power through the structure of political institutions and the ability of groups to solve the collective action problem, or exercise

what we called de facto political power. With these examples in mind in Section 8 we move to discuss in more detail the nature and sources of political power.

### 7.1. Labor markets

A market – an opportunity for individuals to exchange a commodity or service – is obviously a fundamental economic institution relevant for development. As Adam Smith (1776) argued, markets allow individuals to take advantage of the benefits of specialization and the division of labor, and scholars such as Pirenne (1937) and Hicks (1969) argued that the expansion of markets was perhaps *the* driving forces in long-run development.

In the history of Europe a key transformation was from feudal labor market institutions towards modern notions of a free labor market where individuals were able to decide who to work for and where to live. This process of institutional change was intimately connected to the transition from a whole set of feudal economic institutions to the economic institutions we think of as ‘capitalist’. Most historians see this as key to the economic take-off that began in the nineteenth century. It was the countries which had made the transition away from feudalism most completely, such as England, the Netherlands and France, thanks to the revolution of 1789, which developed most rapidly. It was those where feudalism was still in operation, such as Russia and Austria-Hungary, which lagged far behind.

What can account for this differential evolution of feudalism? Scholars beginning with Postan (1937) saw the demographic collapse caused by the black death in the 1340’s as demolishing feudalism in Western Europe. By dramatically altering the land/labor ratio as approximately 40% of the population of Europe died [e.g., Cantor (2001)], the Black Death greatly increased the bargaining power of peasants and allowed them to negotiate a free status ending feudal obligations, particularly with respect to labor. Therefore, Postan’s demographic theory implicitly emphasizes the role of political power in the decline of feudalism: this set of economic institutions started to disappear when the political power of the peasants increased and that of lords declined.

In fact, the distribution of power may be even more important in the whole story than Postan’s theory suggests. As first pointed out by Brenner (1976), the demographic theory of the decline of feudalism is not consistent with the comparative evidence. Although demographic trends were similar all over Europe and

“it is true that . . . in most of Western Europe serfdom was dead by the early sixteenth century. On the other hand, in Eastern Europe, in particular Pomerania, Brandenburg, East Prussia and Poland, decline in population from the late fourteenth century was accompanied by an ultimately successful movement towards imposing extra-economic controls, that is serfdom, over what had been, until then, one of Europe’s freest peasantries. By 1500 the same Europe-wide trends had gone a long way towards establishing one of the great divides in European history, the emergence of an almost totally free peasant population in Western Europe, the

debasement of the peasantry to unfreedom in Eastern Europe.” [Brenner (1976, p. 41)].

What can explain these divergent outcomes? Brenner notes (p. 51): “It was the logic of the peasant to try to use his apparently improved bargaining position to get his freedom. It was the logic of the landlord to protect his position by reducing the peasants’ freedom”. The outcome “obviously came down to a question of power” (p. 51); whether the peasants or the lords had more political power determined whether serfdom declined or became stronger.

Although we are far from an understanding of the determinants of the relative structure of political power in different parts of Europe, Brenner suggests that an important element was the “patterns of the development of the contending agrarian classes and their relative strength in the different European societies: their relative levels of internal solidarity, their self-consciousness and organization, and their general political resources – especially their relationships to the non-agricultural classes (in particular, potential urban class allies) and to the state” (p. 52). To substantiate this view, Brenner studies how villages tended to be organized differently in Eastern Europe, there was “more of a tendency to individualistic farming; less developed organization of collaborative agricultural practices at the level of the village or between villages; and little of the tradition of the ‘struggle for commons rights’ against the lords which was so characteristic of western development” (p. 57). This differential organization was due to the process of initial occupation of these Eastern lands.

Although many parts of Brenner’s analysis remain controversial, there is general agreement that the decline of feudalism and the transformation of European labor markets were intimately related to the political power of the key groups with opposing interests, the peasants and the lords [see, for example, Aston and Philpin (1985) on reactions to Brenner’s interpretation]. Feudal institutions, by restricting labor mobility and by removing the role of the labor market in allocating labor to jobs, undermined incentives and resulted in underdevelopment. But these same economic institutions created large rents for the aristocracy. As a consequence, aristocracies all over Europe attempted to maintain them. It was when their political power weakened that the process of transformation got underway.

## 7.2. Financial markets

Much recent work on growth and development has focused on capital markets. Growth requires investment, so poor agents without access to financial markets will not have the resources to invest. Empirically many scholars have found correlations between the depth of financial markets and growth [see Levine (2005)] and absence of financial markets is at the heart of ambitious theories of comparative development by Banerjee and Newman (1993) and Galor and Zeira (1993).

If the stress on financial markets and financial intermediation is correct, a central issue is to understand why financial systems differ. For example, studies of the development

of banking in the United States in the nineteenth century demonstrate a rapid expansion of financial intermediation which most scholars see as a crucial facilitator of the rapid growth and industrialization that the economy experienced. In his recent study [Haber \(2001, p. 9\)](#) found that in the United States, “In 1818 there were 338 banks in operation, with a total capital of \$160 million—roughly three times as many banks and bank capital as in 1810. Circa 1860, the United States had 1,579 banks, with a total capital of \$422.5 million. Circa 1914 there were 27,864 banks in the United States. Total bank assets totaled \$27.3 billion”.

One might see this rapid expansion of banking and financial services as a natural feature. Yet [Haber \(2001\)](#) shows that the situation was very different in Mexico (p. 24). “Mexico had a series of segmented monopolies that were awarded to a group of insiders. The outcome, circa 1910 could not have been more different: the United States had roughly 25,000 banks and a highly competitive market structure; Mexico had 42 banks, two of which controlled 60 percent of total banking assets, and virtually none of which actually competed with another bank.”

The explanation for this huge difference is not obvious. The relevant technology was certainly readily available everywhere and it is difficult to see why the various types of moral hazards or adverse selection issues connected with financial intermediation should have limited the expansion of banks in Mexico but not the United States. Haber then shows that (p. 9), “at the time that the U.S. Constitution was put into effect in 1789, ... [U.S. banking] was characterized by a series of segmented monopolies that shared rents with state governments via taxes or state ownership of bank stock. In some cases, banks also shared rents directly with the legislators who regulated them.”

This structure, which looked remarkably like that which arose subsequently in Mexico, emerged because state governments had been stripped of revenues by the Constitution. In response, states started banks as a way to generate tax revenues. State governments restricted entry “in order to maximize the amount of rent earned by banks, rent which would then be shared with the state government in the form of dividends, stock distributions, or taxes of various types”.

Thus in the early nineteenth century, U.S. banks evolved as monopolies with regulations aimed at maximizing revenues for the state governments. Yet this system did not last because states began competing among themselves for investment and migrants.

“The pressure to hold population and business in the state was reinforced by a second, related, factor: the broadening of the suffrage. By the 1840s, most states had dropped all property and literacy requirements, and by 1850 virtually all states (with some minor exceptions) had done so. The broadening of the suffrage, however, served to undermine the political coalitions that supported restrictions on the number of bank charters. That is, it created a second source of political competition—competition within states over who would hold office and the policies they would enact.”

The situation was very different in Mexico. After 50 years of endemic political instability the country unified under the highly centralized 40 year dictatorship of Porfirio Diaz until the Revolution in 1910.

In Haber's argument political institutions in the United States allocated political power to people who wanted access to credit and loans. As a result they forced state governments to allow free competitive entry into banking. In Mexico political institutions were very different. There were no competing federal states and the suffrage was highly restrictive. As a result the central government granted monopoly rights to banks who restricted credit to maximize profits. The granting of monopolies turned out to be a rational way for the government to raise revenue and redistribute rents to political supporters [see [North \(1981, Chapter 3\)](#)].

A priori, it is possible that the sort of market regulation Haber found in Mexico might have been socially desirable. Markets never function in a vacuum, but rather within sets of rules and regulations which help them to function. Yet it is hard to believe that this argument applies to Mexico [see also [Maurer \(2002\)](#)]. [Haber \(2001\)](#) documents that market regulation was aimed not at solving market failures and it is precisely during this period that the huge economic gap between the United States and Mexico opened up [on which see [Coatsworth \(1993\)](#), [Engerman and Sokoloff \(1997\)](#)]. Indeed, [Haber and Maurer \(2004\)](#) examined in detail how the structure of banking influenced the Mexican textile industry between 1880 and 1913. They showed that only firms with personal contacts with banks were able to get loans. They conclude (p. 5):

“Our analysis demonstrates that textile mills that were related to banks were less profitable and less technically efficient than their competitors. Nevertheless, access to bank credit allowed them to grow faster, become larger, and survive longer than their more productive competitors. The implication for growth is clear: relatively productive firms lost market share to relatively unproductive (but bank-related) competitors.”

Despite the fact that economic efficiency was hurt by regulations, those with the political power were able to sustain these regulations.

### 7.3. Regulation of prices

As our final example we turn to the regulation of prices in agricultural markets (which is intimately related to the set of agricultural policies adopted by governments). The seminal study of agricultural price regulation in Africa and Latin America is by [Bates \(1981, 1989, 1997\)](#). [Bates \(1981\)](#) demonstrated that poor agricultural performance in Ghana, Nigeria and Zambia was due to government controlled marketing boards systematically paying farmers prices for their crops much below world levels.

“Most African states possess publicly sanctioned monopsonies for the purchase and export of agricultural goods ... These agencies, bequeathed to the governments of the independent states by their colonial predecessors, purchase cash crops

for export at administratively determined domestic prices, and then sell them at the prevailing world market prices. By using their market power to keep the price paid to the farmer below the price set by the world market, they accumulate funds from the agricultural sector” [Bates (1981, p. 12)].

The marketing boards made surpluses which were given to the government as a form of taxation. Bates (1981, p. 15) notes

“A major test of the intentions of the newly independent governments occurred ... [when] between 1959–60 and 1961–62, the world price of cocoa fell approximately £50 a ton. If the resources generated by the marketing agencies were to be used to stabilize prices, then surely this was the time to use the funds for that purpose. Instead ... the governments of both Ghana and Nigeria passed on the full burden of the drop in price to the producers.”

Bates continues “Using the price setting power of the monopsonistic marketing agencies, the states have therefore made the producers of cash crops a significant part of their tax base, and have taken resources from them without compensation in the form of interest payments or of goods and services returned” (pp. 181–189). As a result of this pernicious taxation, reaching up to 70% of the value of the crop in Ghana in the 1970s, investment in agriculture collapsed as did output of cocoa and other crops. In poor countries with comparative advantage in agriculture such a situation mapped into negative rates of economic growth.

Why were resources extracted in this way? Though part of the motivation was to promote industrialization, the main one is to generate resources that could be either expropriated or redistributed to maintain power

“governments face a dilemma: urban unrest, which they cannot successfully eradicate through co-optation or repression, poses a serious challenge to their interests ... Their response has been to try to appease urban interests not by offering higher money wages but by advocating policies aimed at reducing the cost of living, and in particular the cost of food. Agricultural policy thus becomes a by-product of political relations between governments and urban constituents.” [Bates (1981, p. 33)].

In contrast to the situation in Ghana, Zambia and Nigeria, Bates (1981, 1989, 1997) showed that agricultural policy in Kenya and Colombia over this period was much more pro-farmer. The difference was due to who controlled the marketing board. In Kenya, farmers were not smallholders, as they were in Ghana, Nigeria and Zambia, and concentrated landownership made it much easier to solve the collective action problem. Moreover, farming was important in the Kikuyu areas, an ethnic group closely related to the ruling political party, KANU, under Jomo Kenyatta [Bates (1981, p. 122)]. Farmers in Kenya therefore formed a powerful lobby and were able to guarantee themselves high prices. Even though the government of Kenya engaged in land reform after independence

“80% of the former white highlands were left intact and . . . the government took elaborate measures to preserve the integrity of the large-scale farms . . . [which] readily combine in defense of their interests. One of the most important collective efforts is the Kenya National Farmer’s Union (KNFU) . . . The organization . . . is dominated by the large-scale farmers . . . [but] it can be argued that the KNU helps to create a framework of public policies that provides an economic environment favorable to all farmers.” [Bates (1981, pp. 93–94)].

Bates concludes (p. 95) that in Kenya

“large farmers . . . have secured public policies that are highly favorable by comparison to those in other nations. Elsewhere the agrarian sector is better blessed by the relative absence of inequality. But is also deprived of the collective benefits which inequality, ironically, can bring.”

In Colombia, farmers were favored because of competition for their votes from the two main political parties. Bates (1997, p. 54) notes

“Being numerous and small, Colombia’s coffee producers, like peasants elsewhere, encountered formidable costs of collective action. In most similar instances such difficulties have rendered smallholders politically powerless. And yet . . . Colombia’s peasants elicited favorable policies from politicians, who at key moments themselves bore the costs of collective action, provisioning the coffee sector with economic institutions and delegating public power to coffee interests.”

How could the coffee growers gain such leverage over national policy?

“A major reason they could do so . . . is because the structure of political institutions, and in particular the structure of party competition, rendered them pivotal, giving them the power over the political fortunes of those with ambition for office and enabling them to make or break governments. They thereby gained the power to defeat government officials who sought to orchestrate or constrain their behavior.” [Bates (1997, pp. 51, 54)].

A telling piece of evidence in favor of this thesis is that during the 1950s when a civil war broke out between the two parties, there was five years of military rule and policy turned decisively again the coffee growers, only to switch back again with the peaceful resumption of democracy in 1958.

#### 7.4. Political power and economic institutions

These three examples of the creation of economic institutions have certain features in common. All these institutions, labor market regulation/feudalism, the rules governing financial market development, and agricultural price regulation, clearly reflect the outcome of conscious choices. Feudalism did not end in England for incidental or ideological reasons, but because those who were controlled and impoverished by feudal

regulations struggled to abolish them. In Eastern Europe the same struggle took place but with a different outcome. Similarly, Mexico did not end up with different financial institutions than the United States by accident, because of different beliefs about what an efficient banking system looked like, or because of some historical factor independent of the outcome. The same is true for differences in economic policies in Kenya and Ghana. Moreover, different sets of economic institutions arising in different places cannot be argued to be efficient adaptations to different environments. Most historians believe that the persistence of feudal institutions in Eastern Europe well into the nineteenth century explains why it lagged far behind Western Europe in economic development. The difference between the financial institutions of Mexico and the United States also plausibly played a role in explaining why they diverged economically in the nineteenth century. The same holds with respect to agricultural price regulation.

The driving force behind all three examples is that economic institutions are chosen for their distributional consequences. Which specific economic institutions emerge depends on who is able to get their way – who has political power. In England, peasant communities had developed relatively strong local political institutions and were able to consolidate on the shock of the Black Death to put an end to feudal regulations. In Eastern Europe it was the lords who had relatively more power and they were able to intensify feudalism in the face of the same demographic shock [as Domar (1970) pointed out, the Black Death actually made serfdom more attractive to the lords even if at the same time it increased the bargaining power of the peasants]. In the case of banking in the nineteenth century, Haber's research shows while the authoritarian regime in Mexico had the political power to freely create monopolies and create rents in the banking industry, the United States was different because it was federal and much more democratic. The political institutions of the United States prevented politicians from appropriating the rents that could flow from the creation of monopolies. Finally, in Bates's analysis, distortionary price regulations arose in Ghana and Zambia, but not in Kenya and Colombia, because in the latter countries agricultural producers had more political power and so could prevent the distortionary policies that would harm their interests.

It is also useful to consider in the context of these examples the mechanisms we discussed in Section 6 which underlie the adoption of inefficient economic institutions. Why could not the peasants and lords of feudal Europe negotiate and allow the introduction of a set of economic institutions that would have given peasants incentives to innovate and would have allowed for the efficient allocation of labor? Why could not either the lords have promised not to expropriate any benefits that accrued from innovation, or alternatively the peasants agreed to compensate the lords if feudal labor institutions were abolished? Though it is difficult to find direct evidence on such counterfactuals from the Medieval period, the most plausible explanation is that such 'deals' were impossible to make credible. The political power of the lords was intimately connected to feudal institutions and thus dismantling these would not only have increased peasant incentives to innovate, but would also have dramatically altered the balance of political power and the distribution of rents in society. Moreover, under feudal regulations peasants were tied to the land. The introduction of free labor mobility would have

given workers an exit option, thus increasing their bargaining power with the lords over the division of output. Thus lords might anticipate being both political and economic losers from the ending of feudalism, even if total output would have increased.

In the case of agricultural price regulation, similar arguments are plausible. Cocoa farmers in Ghana would not have believed promises by governments that they would not expropriate the fruits of higher investment, and the governments themselves would not have believed promises by the farmers to compensate them if they left office. Moreover, efficient sets of economic institutions in Ghana or Nigeria would have strengthened the economic base of the rural sector at the expense of the political power of the then dominant urban sector. Indeed, for Ghana in the 1960s, we have direct evidence from the urban economy that the threat of being a political loser led to inefficient economic institutions. This emerges in the analysis of Killick (1978, p. 37) of the attempt by the government of Kwame Nkrumah to promote industrialization. Killick notes:

“Even had there been the possibility [of creating an indigenous entrepreneurial class] it is doubtful that Nkrumah would have wanted to create such a class, for reasons of ideology and political power. He was very explicit about this saying ‘we would be hampering our advance to socialism if we were to encourage the growth of Ghanaian private capitalism in our midst’. There is evidence that he also feared the threat that a wealthy class of Ghanaian businessmen might pose to his own political power.”

Further evidence on the importance of political loser considerations comes from E. Ayeh-Kumi one of Nkrumah’s main economic advisers who noted after the coup that ousted Nkrumah in 1966 that Nkrumah: “informed me that if he permitted African business to grow, it will grow to becoming a rival power to his and the party’s prestige, and he would do everything to stop it, which he actually did” [Killick (1978, p. 60)].

In this context, it is interesting that Nkrumah’s solution to consolidate his power was to limit the size of businesses that Ghanaians could own. This caused problems for his industrialization policy which he got round by allowing foreign businessmen to enter Ghana. Though this was inconsistent with his aggressively nationalistic and anti-imperialistic rhetoric, these businessmen did not pose a domestic political threat. Killick notes “Given Nkrumah’s desire to keep Ghanaian private businesses small, his argument that ‘Capital investment must be sought from abroad since there is no bourgeois class amongst us to carry on the necessary investment’ was disingenuous” (p. 37). He goes on to add that Nkrumah “had no love of foreign capitalists but he preferred to encourage them rather than local entrepreneurs, whom he wished to restrict” (p. 40).

All these examples show that the distribution of political power in society is crucial for explaining when economic institutions are good and when they are bad. But where does political power come from and who has political power? In addressing these questions we will develop our theory of economic institutions. In a theory based on social conflict where economic institutions are endogenous, it will be to differences in political institutions and the distribution of political power that we must look to explain variation in economic institutions.

## 8. A theory of institutions

### 8.1. Sources of political power

Who has political power and where does it come from? As we noted in the Introduction (Section 1.2, point 4), political power comes from two sources. First, an individual or group can be allocated *de jure* power by *political institutions*. But institutions are not the only source of power. A second type of political power accrues to individuals or groups if they can solve the collective action problem, create riots, revolts, or demonstrations, own guns, etc. We call this type of power *de facto* political power [see [Acemoglu and Robinson\(2003, Chapter 5\)](#)].

Actual political power is the composition, the joint outcome, of de jure and de facto power. To see how this works out in practice, consider the situation in Chile in the early 1970's. Salvador Allende was elected President with a plurality of the popular vote. The formal political institutions of democracy in Chile allocated power to him to propose legislation, issue decrees, etc. Consequently, even though he did not have an absolute majority in congress, Allende had a great deal of de jure political power. Political power is not just de jure however; it does not simply stem from political institutions. Allende, despite being empowered under the Chilean Constitution, was overthrown by a military coup in 1973. Here, the military under the leadership of General Pinochet, were able to use brute force and guns to over-ride the formal political institutions. The ability to use force is one example of de facto political power.

As we suggested in the introduction, the relationship between political power and economic and political institutions is complex and dynamic. Consider the example we discussed in Section 7.2, the research by Haber on the comparative financial evolution of Mexico and the United States in the nineteenth century. Haber traced the different evolution of economic institutions to differences in initial political institutions. These political institutions led to different distributions of power and this was critical for the emergence of good financial institutions in the United States, whereby those who benefited from a competitive banking industry were able to force politicians to provide the rules which would guarantee it. But where did these differences in political institutions come from? These differences were partly a result of political events in the nineteenth century, and partially a result of different colonial political institutions. In the United States, during the initial phase of colonization in the early seventeenth century very low population density and lack of easily exploitable resources forced colonizing companies and the British state to make both economic and political concessions; they granted the settlers access to land and accepted the formation of representative democratic institutions [[Morgan \(1975\)](#)]. Consequently, even at independence the United States had relatively democratic political institutions [[Keyssar \(2000\)](#)]. Moreover, the initial egalitarian distribution of assets and the high degree of social mobility made for a situation where, at least in the northern states, the distribution of economic resources, and thus de facto power, was relatively equal. The relatively representative political institutions therefore persisted and were supported by the balance of de facto power in society.

In Mexico there were very different initial conditions during the colonial period with a large indigenous population and rich silver mines to exploit. This led to a much more hierarchical and authoritarian balance of political power and very different colonial economic institutions [see [Engerman and Sokoloff \(1997\)](#)]. These conditions fed into the different institutional structures at independence, the United States with its constitution, checks and balances and federalism, Mexico with its much more centralized, unchecked, unbalanced and absolutist state. These different political institutions then led to very different economic institutions and economic outcomes after independence. Thus, in some ultimate sense, the source of different political institutions were different initial conditions during the colonial period.

Consider now the evidence presented by Bates. Agricultural policies were better in Kenya because large farmers could solve the collective action problem and exercise de facto political power. But the main reason for the existence of large farms was that British settlers expropriated the land from Africans during the expansion of colonialism [see [Berman and Lonsdale \(1992\)](#)]. Thus previous combinations of formal political institutions (colonial institutions) and de facto power (the military might of the British Empire) determined economic institutions, feeding into the future distribution of de facto power even after the nature of de jure power changed dramatically with independence.

We can now see that these examples substantiate the dynamic model that we sketched in Section 1.2. There we showed that at any date, political power is shaped by political institutions, which determine de jure power, and the inherited distribution of resources, which affect the balance of de facto power. Political power then determines economic institutions and economic performance. It also influences the future evolution of political power and prosperity. Economic institutions determine the distribution of resources at that point, which, in turn, influences the distribution of de facto power in the future. Similarly, the distribution of power at any point determines not just the economic institutions then, but also the future political institutions. Thus the allocation of political power at one date, because of the way it influences the distribution of resources and future political institutions, has a crucial effect on the future allocation of both de facto and de jure political power.

Both the comparison Haber made between Mexico and the United States, and that which Bates made between Ghana, Zambia, Kenya and Colombia illustrate this diagram in action. They show how political institutions and de facto power combine to generate different set of economic institutions, how these institutions determine both the distribution of resources and the growth rate of the economy, and how power and institutions evolve over time, often in ways that tend to reinforce particular initial conditions.

## 8.2. Political power and political institutions

The examples we discussed above showed how political power depends on political institutions and de facto power, and how this determines economic institutions. Moreover, we saw that at any time the pre-existing economic institutions will be an important de-

terminant of the distribution of de facto power. The final element to emphasize is how political institutions evolve over time and how they influence the distribution of political power.

To see why political institutions are so important as a source of political power think of a situation where a group, say the Chilean army in the early 1970s, has a great deal of de facto power. Indeed, it has so much de facto power that it can overrule the Chilean Constitution, making the political institutions largely irrelevant. In fact in Chile the de facto power of the military was able to overthrow the legitimate government and completely reverse the economic policies and economic institutions chosen by the Allende government (including land reform and mass nationalization of industry). Not only did the military reverse the economic institutions preferred by Allende and the groups who elected him, they then implemented their own preferred set of economic institutions, in particular deregulating the trade regime and the economy. Yet the Pinochet regime was heavily concerned with political institutions, and in 1980 Pinochet re-wrote the constitution.

If de facto power was decisive in Chile what is the role for political institutions? If the constitution can be overthrown, why bother to re-write it? The secret to this lies in the intrinsically transitory nature of de facto power.<sup>13</sup> Yes, the military were able to organize a coup in 1973 but this was only because times were uniquely propitious. There was a world-wide economic crisis, and factions of the military that opposed the coup could be marginalized. Moreover, the United States government at the time was happy to encourage and endorse the overthrow of a socialist government, even if it had been democratically elected. The coming together of such circumstances could not be expected to happen continually, hence once Chilean society re-democratized, as it did after 1990, the military would not be able to continually threaten a coup. In response to this Pinochet changed the political institutions in order to attempt to lock in the power of the military, and thus the economic institutions that he/they preferred. Therefore, the important role for political institutions is that they influence the future allocation of political power. This dynamic role is crucial because it explains the major desire of agents to change political institutions when they get the chance – this is how they can attempt to enduringly alter the balance of political power in their favor [see [Acemoglu and Robinson \(2003\)](#)].

<sup>13</sup> The empirical literature on the collective action problem has recognized that the difficulty of solving the collective action problems lead collective action to typically be transitory. [Lichbach \(1995, p. 17\)](#) notes “collective action, if undertaken on a short-term basis, may indeed occur; collective action that requires long periods of time does not . . . Given that most people’s commitments to particular causes face inevitable decline, most dissident groups are ephemeral, most dissident campaigns brief”. This transitory nature of collective action is echoed by [Tarrow \(1991, p. 15\)](#) who notes “the exhaustion of mass political involvement”, while [Ross and Gurr \(1989, p. 414\)](#) discuss political “burnout”. Similarly, [Hardin \(1995, p. 18\)](#) argues that “the extensive political participation of civil society receives enthusiastic expression only in moments of state collapse or great crisis. It cannot be maintained at a perpetually high level.”

### 8.3. A theory of political institutions

We now have in place the outlines of our theory of institutions. There are seven points to emphasize, paralleling the discussion in Section 1.2 and our diagrammatic exposition there. First, individuals have preferences over economic institutions because of the allocation of resources that these institutions induce.

Second, people's preferences typically do not agree because efficiency and distribution cannot be separated. Different economic institutions will benefit different groups, and this will determine the preferences of these individuals and groups with respect to economic institutions.

Third, the problem of commitment explains why efficiency and distribution are inseparable. Economic institutions are collective choices, and they are chosen and sustained by the state. Since there is no third party to enforce the decisions of the state, problems of commitment are particularly severe in the political realm.

Fourth, the equilibrium structure of economic institutions will therefore be determined by who has the power to get their way, i.e., who can create and sustain economic institutions that benefit themselves. The distribution of political power thus determines economic institutions, the allocation of resources and the rate of economic growth.

Fifth, political power has two forms: de jure power determined by the political institutions, such as the constitution and the electoral rules, and de facto power, which stems from the ability to solve the collective action problem, mobilize weapons, etc. De facto power can influence political outcomes independently of the political institutions, and its distribution often critically determines how a given set of institutions works in practice and whether or not they are actually obeyed.

Sixth, the distribution of de facto political power at any date is influenced to a large degree by the distribution of resources in society, since those with greater resources can command more power both through legitimate and intimate means, and perhaps can also solve the collective action problem more efficiently. Naturally, the distribution of resources at this point is influenced by economic institutions and economic outcomes in the past.

Finally, political institutions are also endogenous; the current balance of political power, incorporating both de jure and de facto elements, also determines future political institutions. Political institutions are important because they allocate, at least within the limits defined by the exercise of future de facto power, the allocation of future de jure political power. Since de facto power, because of the nature of the collective action problem, is intrinsically transitory and difficult to wield, political institutions are often crucial in creating a source of durable political power. This makes it very attractive for groups to use their de facto political power to change political institutions so as to modify the distribution of future political power in their favor.

## 9. The theory in action

We now consider two more examples that further demonstrate our theory of institutions in action. Like the examples discussed in Section 7, these examples contain all the elements of our theory laid out in a skeletal way in Section 1.2. They show the role of political power in determining economic institutions, they demonstrate the different factors, both de facto and de jure, that determine political power, and they illustrate how de facto political power is often used to change political institutions in order to influence the future distribution of de jure political power.

### 9.1. *Rise of constitutional monarchy and economic growth in early modern Europe*

Our first example is the rise of constitutional monarchy in Europe. In the medieval period most European nations were governed by hereditary monarchies. However, as the feudal world changed, various groups struggled to gain political rights and reduce the autocratic powers of monarchies. In England, this process began as early as 1215 when King John was forced by his barons to sign the Magna Carta, a document which increased the powers of the barons, introduced the concept of equality before the law, and forced subsequent kings to consult with them. Many other European nations also developed ‘parliaments’ which kings could summon to discuss taxation or warfare [see [Graves \(2001\)](#), [Ertman \(1997\)](#)]. Nevertheless, the movement towards limited, constitutional monarchy was not linear or simple. Indeed, in France, certainly from the beginning of Louis XIV’s reign in 1638, a more powerful absolutist monarchy appeared with very few controls. Indeed the feudal French parliaments, the Estates General, were not summoned between 1614 and 1788, just before the Revolution.

In England, the Tudor monarchs, particularly Henry VIII and then Elizabeth I, followed by the first Stuart kings, James I and Charles I, also attempted to build an absolutist monarchy. They failed, however, mostly because of Parliament, which blocked attempts to concentrate power. The constitutional outcome in England was settled by the Civil War from 1642–1651 and the Glorious Revolution in 1688. In the first of these conflicts the forces of Parliament defeated those loyal to Charles I and the king was beheaded. In 1660 the monarchy was restored when Charles II became king, but his brother James II was deposed in 1688 and Parliament invited William of Orange to become king.

Other places in Europe, particularly the Netherlands, saw similar developments to those in England. Under the Dukes of Burgundy, the Netherlands had won a considerable amount of political and economic freedom, particularly under the Grand Privilege of 1477 which gave the States General of the Burgundian Netherlands the right to gather on their own initiative and curbed the right of the ruler to raise taxes. However, the Netherlands were inherited by the Hapsburgs through marriage, and by 1493 Maximilian of Hapsburg had reversed the Grand Privilege. After 1552, war with France increased the Hapsburgs’ fiscal needs and led them to impose a large tax burden on the Netherlands, already a prosperous agricultural and mercantile area. Growing fiscal and

religious resentment in 1572 led to a series of uprisings against the Hapsburgs, mostly orchestrated by commercial interests. These culminated in the War of Independence which was finally won in 1648.

While England and the Netherlands were developing limited constitutional governments, Spain and Portugal were moving in the same direction as France, towards greater absolutism. [Davis \(1973a, p. 66\)](#) notes [in Castille] “the king ruled subject only to weak constitutional restraints. In the first decades of the sixteenth century the crown had reduced the pretensions of the Castilian nobility and towns, so that the representative body, the Cortes, could obstruct but not in the last resort prevent royal tax raising.”

These differential institutional trajectories were of enormous consequence. The economies of the Netherlands and England moved ahead of the rest of Europe precisely because these countries developed limited, constitutional government. This form of government led to secure property rights, a favorable investment climate and had rapid multiplier effects on other economic institutions, particularly financial markets [see, e.g., [North and Weingast \(1989\)](#), [de Vries and van der Woude \(1997\)](#)]. While the Netherlands and Britain prospered, France was convulsed by the French Revolution, and by the nineteenth century Spain and Portugal were impoverished backward nations. How can we account for these diverging paths in the early modern period? Why did England and the Netherlands develop limited constitutional rule, while France, Spain and Portugal did not?

We proposed an explanation in [Acemoglu, Johnson and Robinson \(2005\)](#) related to the differential responses of these countries to the opportunities of ‘Atlantic trade’, that is, overseas trade and colonial activity unleashed by the discovery of the New World and the rounding of the Cape of Good Hope at the end of the fifteenth century. All five nations engaged in Atlantic trade, but they did so in different ways, with very different implications for the organization of society, political institutions and subsequent economic growth.

In England “most trade was carried on by individuals and small partnerships, and not by the Company of Merchant Adventurers, the Levant Company … or others of their kind” [[Davis \(1973b, p. 41\)](#)]. At least by 1600 there was quite free entry into the English merchant class. The same was true in the Netherlands. In contrast, [Cameron \(1993, p. 127\)](#) describes the Portuguese situation as follows: “The spice trade in the East Indies of the Portuguese Empire was a crown monopoly; the Portuguese navy doubled as a merchant fleet, and all spices had to be sold through the *Casa da India* (India House) in Lisbon … no commerce existed between Portugal and the East except that organized and controlled by the state”. In Spain, similarly, colonial trade was a monopoly of the Crown of Castille, which they delegated to the *Casa de Contratación* (House of Trade) in Seville. This merchants guild was closely monitored by the government [[Parry \(1966, Chapter 2\)](#)]. The main aim of these regulations was to make sure that all of the gold and silver from the Americas flowed back to Spain, creating a source of direct tax revenues for the crown. As a result, Latin American colonies were forbidden to buy manufactured goods from anywhere other than Spain, and all exports and imports had to pass through controlled channels. For example, until the Bourbon reforms of the mid eighteenth cen-

tury, nothing could be exported directly from Buenos Aires, and if somebody produced anything for export on the Pampas, it had to be carried over the Andes and exported from Lima in Peru!

The source of the differences in the organization of trade, in turn, reflected the different political institutions of these countries. At the time, the granting of trade monopolies was a key fiscal instrument to raise revenues; the more powerful monarchs could increase their revenues by granting trade monopolies or by directly controlling overseas trade, while weaker monarchs could not. At the turn of the fifteenth century, the crown was much stronger in France, Spain and Portugal than in Britain and the Netherlands, and this was the most important factor in the differences in the organization of overseas trade. In fact, when both Tudor and Stuart monarchs attempted to create monopolies similar to those in Spain and Portugal, this was successfully blocked by the English Parliament [see, for example, [Hill \(1969\)](#)]. Consequently, as world trade expanded in the sixteenth and early seventeenth centuries, in England and the Netherlands it enriched merchants engaged in overseas trade, but in France, Spain and Portugal it enriched the crown and groups allied with it. In England and the Netherlands, but not in France, Spain and Portugal, a new class of merchants (and gentry in England) arose with interests directly opposed to those of the Stuarts and the Hapsburgs, and this group was to play a central part in subsequent political changes.

In the case of the Netherlands, [de Vries and van der Woude \(1997\)](#) argue that “urban economic interests ultimately believed it advantageous to escape the Hapsburg imperial framework” (p. 369), and that it was “the traditional pillars of the maritime economy ... that supported and strengthened the young Republic in its hour of need” (p. 366). Moreover, in the case of Amsterdam, “[Hapsburgs’] opponents included most of the city’s international merchants ... In 1578 a new Amsterdam city council threw the city’s lot in with the Prince of Orange ... among the merchants returning from ... exile were [those whose families] and several generations of their descendants would long dominate the city” ([1997](#), p. 365). The expansion of world trade enriched and expanded precisely those groups within Dutch society most opposed to Hapsburg rule. [Israel \(1995, pp. 241–242\)](#) writes: ‘From 1590, there was a dramatic improvement in the Republic’s economic circumstances. Commerce and shipping expanded enormously, as did the towns. As a result, the financial power of the states rapidly grew, and it was possible to improve the army vastly, both qualitatively, and quantitatively, within a short space of time. The army increased from 20,000 men in 1588 to 32,000 by 1595, and its artillery, methods of transportation, and training were transformed’ [see also [Israel \(1989, Chapter 3\)](#)]. By 1629, the Dutch were able to field an army of 77,000 men, 50% larger than the Spanish army of Flanders [[Israel \(1995, p. 507\)](#)]. As a consequence of the Dutch revolt, the Netherlands developed a republican form of government closely attuned to mercantile interests. [De Vries and van der Woude \(1997, p. 587\)](#) describe the new political elite following the Dutch Revolt as: “6 to 8% of urban households with incomes in excess of 1,000 guilders per year. This was the *grote burgerij* from whom was drawn the political and commercial leadership of the country. Here we find, first and

foremost, the merchants”, and point out how merchants dominated the governments of Leiden, Rotterdam and the cities in two largest states, Zeeland and Holland.

In England, the Civil War and Glorious Revolution coincided with the great expansion of English mercantile groups into the Atlantic. The East India Company was founded in 1600 as the culmination of a series of efforts to develop trade routes with Asia. The 1620s saw the great expansion of tobacco cultivation in Virginia and this was shortly followed by the development of the highly profitable English sugar colonies in the Caribbean. Finally, in the 1650s the English began to take over the Atlantic slave trade. Both the Civil War and the Glorious Revolution were at root battles over the rights and prerogatives of the monarchy. In both cases new merchant interests predominantly sided with those in the gentry demanding restrictions on the powers of the monarchy in order to protect their property and commerce.

The majority of merchants trading with the Americas and in Asia supported Parliament during the Civil War. [Brunton and Pennington \(1954, p. 62\)](#) also note “in the country as a whole there was probably a preponderance of Parliamentarian feeling among merchants”. Detailed analyses of the initial members of the Long Parliament in 1640 show that a significant majority of merchants supported the Parliamentarian cause [see [Brenner \(1973, 1993\)](#), [Keeler \(1954\)](#) and [Brunton and Pennington \(1954\)](#)]. Members of the Commons from the City of London (the main center of mercantile activity), as well as many non-London commercial constituencies, such as Southampton, Newcastle and Liverpool, supported Parliament against the King. These men included both professional merchants and aristocrats who invested in colonizing the Americas. These new merchants also provided the financial support needed by Parliament in the difficult early days of the war. They became the customs farmers for the new regime and therefore advanced tens of thousands of pounds that were essential in building up the army [[Brenner \(1973, p. 82\)](#)].

[Pincus \(1998, 2001, 2002\)](#) further documents the critical role of mercantile interests in the Glorious Revolution. He concludes ([2002, p. 34](#)) “England’s merchant community actively supported William’s plan for invasion, and provided a key financial prop to the regime in the critical early months”. He notes that James II favored the East India Company and granted various monopoly privileges, alienating the merchant class. Thus, “no wonder the merchant community poured money into William of Orange’s coffers in 1688” [[Pincus \(2002, pp. 32–33\)](#)].

The changes in the distribution of political power, political institutions and thus economic institutions that took place in England and the Netherlands had no counterparts in countries with relatively absolutist institutions, like Spain and Portugal, where the crown was able to closely control the expansion of trade. In these countries it was the monarchy and groups allied with it that were the main beneficiaries of the early profits from Atlantic trade, and groups favoring political and economic change did not become strong enough to induce such change. As a result, only in the Netherlands and England did constitutional rule emerge, and only in these two countries were property rights secure. As a result it was these same two countries that prospered.

Why could the monarchies of Spain and Portugal not negotiate a more efficient set of institutions? Alternatively why did the Stuart monarchs in England have to be beheaded or forced from power before better economic institutions could emerge?

It seems quite clear that a change to a more efficient set of institutions in Spain and Portugal would not have been possible under the auspices of the absolutist state, and a reduction in the power of the state was certainly inimical to the interest of the crown. In the case of England, [Hill \(1961a, p. 22\)](#) argues directly that the reason that the Tudor and Stuart monarchs were not in favor of efficient economic institutions is because they feared that this would undermine their political power. He notes:

“in general the official attitude to industrial advance was hostile, or at best indifferent. It was suspicious of social change and social mobility, the rapid enrichment of capitalists, afraid of the fluctuations of the market and of unemployment, of vagabondage, and social unrest . . . the Elizabethan codes aimed at stabilizing the existing class structure, the location of industry and the flow of labor supply by granting privileges and by putting hindrances in the way of the mobility and the freedom of contract.”

The account so far explains why a change in the balance of (de facto) political power in England and the Netherlands led to a set of economic institutions favoring the interests of merchants. But in fact much more happened during the seventeenth century; an entirely new set of political institutions, constitutional regimes, restricting the power of the monarchy, were introduced. The reason why the merchants and the gentry in England (and the merchants in the Netherlands) used their newfound powers for political reform illustrates the dynamics of political power emphasized by our theoretical framework.

For example in the case of England, although in 1688 the Parliament might have been strong, it could not be sure that this power would endure. Indeed, the ability to solve the collective action problem and wield de facto power is intrinsically transitory. For instance, the Parliament vanquished James II with the help of a Dutch army, after which they invited William of Orange to take the throne. But how could they anticipate whether or not William would try to assert the absolutist prerogatives that James II had demanded?

The way to make transitory power permanent is to embody it into the rules of the game which is exactly what the English Parliament did after 1688. The changes in institutions after 1688 had large and important effects. For instance, in the eighteenth century the English monarchy was able to borrow huge amounts of money because the fiscal control of Parliament guaranteed that it would not default [see [Brewer \(1988\)](#), [Stasavage \(2003\)](#)]. This borrowing has been seen as crucial to the success of the English war machine. Moreover, with Parliament in control of fiscal policy, the crown was no longer able either to raise money through arbitrary taxation, or to grant monopoly rights in exchange for money – issues which had previously been constant sources of friction between the crown and Parliament. Similarly, after 1688, the greater security of property rights in England led to a huge expansion of financial institutions and markets [[Neal](#)

(1990)], which, [North and Weingast \(1989\)](#) argue, laid the institutional foundations for the Industrial Revolution.

Of course the English crown was not without some residual power and might have attempted to mount a coup against Parliament to change political institutions back in its favor. This certainly happened in some places, such as in France after 1849 when Louis Napoleon mounted a successful coup to restore absolutist privileges lost in 1848. Nevertheless, changes in political institutions altered the nature of the status quo in significant ways, and therefore, influenced the future distribution of de jure political power. Political institutions are not cast stone, and they can change, but they still create a source of political power more durable than mere de facto power.

## 9.2. Summary

The emergence of constitutional rule in some societies of early modern Europe therefore provides a nice example of how economic institutions, which shape economic outcomes, are determined by political power, which is in turn determined by political institutions and the distribution of resources in society. The Netherlands and England prospered in this period because they had good economic institutions, particularly secure property rights and well developed financial markets. They had these economic institutions because their governments were controlled by groups with a strong vested interest in such economic institutions. These groups wielded political power because of the structure of political institutions, i.e., they received de jure power in the Netherlands after the Dutch Revolt and in England after the Civil War and Glorious Revolution.

Moving one step back, we see that political institutions allocated more de jure political power to commercial interests in England and the Netherlands than in France, Spain and Portugal because of major changes in political institutions during the 1600s. These changes took place because commercial interests in England and the Netherlands acquired significant de facto political power as a result of their improving economic fortunes, itself a consequence of the interaction of Atlantic trade and the organization of overseas trade in these countries. Crucially for our framework, these commercial interests used their de facto power to reform (or revolutionize) political institutions so as to acquire de jure political power and solidify their gains.

These events, therefore, illustrate the various elements of our theoretical framework. In particular, they show how it is useful to think of political institutions and the distribution of economic resources as the state variables of the dynamic system, which determine the distribution of political power, and via this channel, economic institutions and economic outcomes. Political institutions and the distribution of economic resources are, themselves, endogenous, determined by political power and economic institutions, as exemplified by the fact that the distribution of economic resources changed significantly during the sixteenth century as a result of the new economic opportunities presented by the rise of Atlantic trade, and these changes were crucially influenced by the existing economic institutions (the organization of overseas trade). Furthermore,

the change in the balance of political power led to the changes in political institutions through the English Civil War, the Glorious Revolution and the Dutch Revolt.

### *9.3. Rise of electoral democracy in Britain*

Our second example, based on [Acemoglu and Robinson \(2000a, 2001, 2003\)](#), is the rise of mass democracy. In the early nineteenth century, European countries were run by small elites. Most had elected legislatures, often descendants of medieval parliaments, but the franchise was highly restricted to males with relatively large amounts of assets, incomes or wealth. However, as the century and the Industrial Revolution progressed, this political monopoly was challenged by the disenfranchised who engaged in collective action to force political change.

In response to these developments, the elites responded in three ways. First by using the military to repress the opposition, as in the responses to the revolutions of 1848. Second, by making economic concessions to buy off opposition – this is the standard explanation for the beginnings of the welfare state in Germany under Bismarck. Finally, if neither repression nor concessions were attractive or effective, elites expanded the franchise and gave political power to the previously disenfranchised – they created the precedents of modern democracy.

The history of the rise of democracy in Britain is in many ways representative of the experiences of many other European countries. The first important move towards democracy in Britain came with the First Reform Act of 1832. This act removed many of the worst inequities under the old electoral system, in particular the ‘rotten boroughs’ where several members of parliament were elected by very few voters. The 1832 reform also established the right to vote based uniformly on the basis of property and income. The reform was passed in the context of rising popular discontent at the existing political status quo in Britain.

By the 1820s the Industrial Revolution was well under way and the decade prior to 1832 saw continual rioting and popular unrest. Notable were the Luddite Riots from 1811–1816, the Spa Fields Riots of 1816, the Peterloo Massacre in 1819, and the Swing Riots of 1830 [see [Stevenson \(1979\)](#) for an overview]. Another catalyst for the reforms was the July revolution of 1830 in Paris. Much of this was led and orchestrated by the new middle-class groups who were being created by the spread of industry and the rapid expansion of the British economy. For example, under the pre-1832 system neither Manchester nor Sheffield had any members of the House of Commons.

There is little dissent amongst historians that the motive for the 1832 Reform was to avoid social disturbances [e.g., [Lang \(1999, p. 36\)](#)]. The 1832 Reform Act increased the total electorate from 492,700 to 806,000, which represented about 14.5% of the adult male population. Yet, the majority of British people could not vote, and the elite still had considerable scope for patronage, since 123 constituencies still contained less than 1000 voters. There is also evidence of continued corruption and intimidation of voters until the Ballot Act of 1872 and the Corrupt and Illegal Practices Act of 1883. The Reform Act therefore did not create mass democracy, but rather was designed as

a strategic concession. In presenting his electoral reform to the British Parliament in 1831, the Prime Minister Earl Grey was well aware that this was a measure necessary to prevent a likely revolution. He argued:

“The Principal of my reform is to prevent the necessity for revolution . . . re-forming to preserve and not to overthrow.” [Quoted in [Evans \(1983, p. 212\)](#)].

Unsurprisingly therefore, the issue of parliamentary reform was still very much alive after 1832, and it was taken up centrally by the Chartist movement. But as [Lee \(1994, p. 137\)](#) notes “The House of Commons was largely hostile to reform because, at this stage, it saw no need for it”. This had changed by 1867, largely due to a juxtaposition of factors, including the sharp business cycle downturn that caused significant economic hardship and the increased threat of violence. Also significant was the founding of the National Reform Union in 1864 and the Reform League in 1865, and the Hyde Park riots of July 1866 provided the most immediate catalyst.

[Lang \(1999, p. 75\)](#) sums up his discussion by saying “The Hyde Park affair, coupled with other violent outbursts, helped to underscore the idea that it would be better to keep the goodwill of the respectable workers than to alienate them”. Reform was initially proposed by the Liberal Prime Minister Russell in 1866 but was defeated by the Conservatives and dissident MP’s. As a result Russell’s government fell, and the Conservatives formed a minority administration with Lord Derby as their leader in the House of Lords, and Disraeli in charge of the House of Commons. It was Disraeli who then constructed a coalition to pass the Second Reform Act in 1867. As a result of these reforms, the total electorate was expanded from 1.36 million to 2.48 million, and working class voters became the majority in all urban constituencies. The electorate was doubled again by the Third Reform Act of 1884, which extended the same voting regulations that already existed in the boroughs (urban constituencies) to the counties (rural constituencies). The Redistribution Act of 1885 removed many remaining inequalities in the distribution of seats and from this point on Britain only had single member electoral constituencies (previously many constituencies had elected two members – the two candidates who gained the most votes). After 1884 about 60% of adult males were enfranchised. Once again social disorder appears to have been an important factor behind the 1884 act.

In Britain, the Reform Acts of 1867–1884 were a turning point in the history of the British state. Economic institutions also began to change. In 1871 Gladstone reformed the civil service, opening it to public examination, making it meritocratic. Liberal and Conservative governments introduced a considerable amount of labor market legislation, fundamentally changing the nature of industrial relations in favor of workers. During 1906–1914, the Liberal Party, under the leadership of Asquith and Lloyd George, introduced the modern redistributive state into Britain, including health and unemployment insurance, government financed pensions, minimum wages, and a commitment to redistributive taxation. As a result of the fiscal changes, taxes as a proportion of National Product more than doubled in the 30 years following 1870, and then doubled again. In the meantime, the progressivity of the tax system also increased [[Lindert \(2004\)](#)]. Fi-

nally, there is also a consensus amongst economic historians that inequality in Britain fell after the 1870's [see [Lindert \(2000, 2004\)](#)].

Meanwhile, the education system, which was either primarily for the elite or run by religious denominations during most of the nineteenth century, was opened up to the masses; the Education Act of 1870 committed the government to the systematic provision of universal education for the first time, and this was made free in 1891. The school leaving age was set at 11 in 1893, then in 1899, it increased to 12 and special provisions for the children of needy families were introduced [[Mitch \(1993\)](#)]. As a result of these changes, the proportion of 10-year olds enrolled in school that stood at 40 percent in 1870 increased to 100 percent in 1900 [[Ringer \(1979, p. 207\)](#)]. Finally, a further act in 1902 led to a large expansion in the resources for schools and introduced the grammar schools which subsequently became the foundation of secondary education in Britain.

Following the Great War, the Representation of the People Act of 1918 gave the vote to all adult males over the age of 21, and women over the wage of 30 who were ratepayers or married to ratepayers. Ultimately, all women received the vote on the same terms as men in 1928. The measures of 1918 were negotiated during the war and may reflect to some extent a quid pro quo between the government and the working classes who were needed to fight and produce munitions. Nevertheless, [Garrard \(2002, p. 69\)](#) notes "most assumed that, if the system was to survive and 'contentment and stability prevail', universal citizenship could not be denied men, perceived to have suffered so much and to have noticed Russia's Revolution".

Overall, the picture which emerges from British political history is clear. Beginning in 1832, when Britain was governed by the relatively rich, primarily rural aristocracy, a series of strategic concessions were made over an 86 year period. These concessions were aimed at incorporating the previously disenfranchised into politics since the alternative was seen to be social unrest, chaos and possibly revolution. The concessions were gradual because in 1832, social peace could be purchased by buying off the middle classes. Moreover, the effect of the concessions was diluted by the specific details of political institutions, particularly the continuing unrepresentative nature of the House of Lords. Although challenged during the 1832 reforms, the House of Lords provided an important bulwark for the wealthy against the potential of radical reforms emanating from a democratized House of Commons. Later, as the working classes reorganized through the Chartist movement and later through trade unions, further concessions had to be made. The Great War and the fallout from it sealed the final offer of full democracy. Though the pressure of the disenfranchised played less of a role in some reforms than others, and other factors undoubtedly played a role, the threat of social disorder was the main driving force behind the creation of democracy in Britain.

The story of the rise of mass democracy that emerges from the British evidence is one where economic and social changes connected with industrialization (for example, rising inequality) and urbanization increased the de facto power of the disenfranchised. In response, they demanded political rights, in particular changes in the political institutions which would allocate future political power to them. These changes in political

institutions were, in many ways, the direct cause of the changes in economic institutions, in particular, in the labor market, in government policy, in the educational system, with major distributional implications, including the fall in inequality.

Why did elites in Britain create a democracy? Our discussion makes it clear that democracy did not emerge from the voluntary acts of an enlightened elite. Democracy was, in many ways, forced on the elite, because of the threat of revolution. Nevertheless, democratization was not the only potential outcome in the face of pressure from the disenfranchised, or even in the face of the threat of revolution. Many other countries faced the same pressures and political elites decided to repress the disenfranchised rather than make concessions to them. This happened with regularity in Europe in the nineteenth century, though by the turn of the twentieth century most had accepted that democracy was inevitable. Repression lasted much longer as the favorite response of elites in Latin America, and it is still the preferred option for current political elites in China or Burma.

The problem with repression is that it is costly. Faced with demands for democracy political elites face a trade-off. If they grant democracy, then they lose power over policy and face the prospect of, possibly radical, redistribution. On the other hand, repression risks destroying assets and wealth. In the urbanized environment of nineteenth century Europe (Britain was 70% urbanized at the time of the Second Reform Act), the disenfranchised masses were relatively well organized and therefore difficult to repress. Moreover, industrialization had led to an economy based on physical, and increasing human, capital. Such assets are easily destroyed by repression and conflict, making repression an increasingly costly option for elites. In contrast, in predominantly agrarian societies like many parts of Latin America earlier in the century or current-day Burma, physical and human capital are relatively unimportant and repression is easier and cheaper. Moreover, not only is repression cheaper in such environments, democracy is potentially much worse for the elites because of the prospect of radical land reform. Since physical capital is much harder to redistribute, elites in Western Europe found the prospect of democracy much less threatening.

Faced with the threat of revolt and social chaos, political elites may also attempt to avoid giving away their political power by making concessions, such as income redistribution or other policies that favor non-elites and the disenfranchised. The problem with concessions however is their credibility, particularly when de facto power is transitory. For example, if a crisis, such as a harvest failure or business cycle recession creates a window of opportunity to solve the collective action problem and challenge the existing regime, the elites would like to respond with the promise of policy concessions. Yet windows of opportunity disappear and it is difficult to sustain collective action which entails people protesting in the streets and being away from their families and jobs. Thus collective action quickly dissipates and once it does so, the government has an incentive to renege on its promise of concessions. The promise of concessions, which people know to be non-credible is unlikely to defuse collective action. Hence, [Acemoglu and Robinson \(2000a, 2001, 2003\)](#) argue that democratization occurred as a way of making credible commitments to the disenfranchised. Democratization was a credible commitment to future redistribution, because it reallocated de jure political power away from

the elites to the masses. In democracy, the poorer segments of the society would be more powerful and could vote, in other words, could use their de jure political power, to implement economic institutions and policies consistent with their interests. Therefore, democratization was a way of transforming the transitory de facto power of the disenfranchised poor into more durable de jure political power.

#### 9.4. Summary

The emergence of mass democracy is another example illustrating our theory of institutions. Into the nineteenth century, economic institutions, particularly in the labor market, disadvantaged the poor. For example, trade unions were illegal and as late as the 1850 in Britain workers trying to organize a union could be shipped to the penal colony in Tasmania, Australia. The poor could not alter economic institutions in their favor because, being disenfranchised, they had little de jure political power, and also limited de facto power because they were often unable to solve their collective action problems.

However, changes in the structure of society and the economy during the early nineteenth century altered the balance of political power, in particular making the exercise of de facto power by the politically disenfranchised much easier [Tilly (1995) and Tarrow (1998) document the changing qualitative nature of collective action over this period]. The rise in the de facto political power of the poor necessitated a change in political institutions in their favor to defuse the threat of revolution. This was to tilt the future allocation of de jure political power, and consequently to ensure future economic institutions and policies consistent with their interests.

Whether or not increases in de facto power translated into democracy depended on a number of factors, in particular how difficult and costly it was for elites to use repression to counter the increase in the power of the masses, and how costly the prospect of democracy was. The changes in political institutions that occurred with democracy had profound implications for economic institutions. In the case of Britain, the period after the Second Reform Act of 1867 led the British state to commit itself to providing universal education and it also led to radical changes in labor market institutions allowing trade unions to form legally for the first time and increasing the bargaining power of labor. Hence economic institutions changed radically in favor of those newly endowed with de jure political power, mostly the relatively poor. This is in fact a relatively general result of democratization. Democracy enfranchises the poor, and the poor are able to use democracy to tilt economic institutions and the distribution of income in society in their favor [Li, Squire and Zou (1998), Rodrik (1999)].

The emergence of democracy in the nineteenth-century Europe therefore also illustrates the workings of our theoretical framework. In particular, it shows how political institutions determine economic institutions and policies, and thus the distribution of resources, and it shows how political institutions change, especially in response to an imbalance of de facto political power, as a credible way of influencing the future allocation of de jure political power.

## 10. Future avenues

In this chapter we have proposed a framework for thinking about why some countries grow faster and are richer than others. We emphasized, following [North and Thomas \(1973\)](#), that most economic growth theory focuses only on proximate determinants of prosperity. Although this body of work has been useful in helping us understand the mechanics of growth, it fails to provide a satisfactory account of why some countries grow while others do not. Even more recent analyses which have emphasized market institutions and imperfections, and even political economy, have not provided convincing explanations for why countries differ in their equilibrium set of institutions. A major research goal must now be to get beyond the neoclassical growth model and its modern extensions, and search for the deeper causes, i.e., the fundamental determinants of growth.

We argued that the available evidence is consistent with the view that whether or not a society grows depends on how its economy is organized – on its economic institutions. We then proposed the outlines of a theory of institutions and illustrated it through a series of historical examples. We emphasized that a theory of why different countries have different economic institutions must be based on politics, on the structure of political power, and the nature of political institutions. Much remains to be done. First, the framework we outlined was largely verbal rather than mathematical, and thus, by its very nature, not fully specified. Constructing formal models incorporating and extending these ideas is the most important task ahead. Although some of our past work [e.g., [Acemoglu and Robinson \(2000a, 2001\)](#), [Acemoglu \(2003b\)](#)] formalizes parts of this framework, the full model has not been developed yet.

There are also many important issues left out of our framework, which appear to offer fruitful areas for future research. First, though we know that institutions, both economic and political, persist for long periods of time, often centuries (and sometimes millennia), we do not as yet have a satisfactory understanding of the mechanisms through which institutions persist.

Second, and closely related, although institutions do generally persist, sometimes they change. We have important examples of societies which have radically changed their political and economic institutions. Some do so for internal reasons, such as France after the Revolution of 1789, and some do because of external pressures such as Japan after the Meiji restoration or Russia after the Crimean War.

The important point here is that both institutional persistence and institutional change are equilibrium outcomes. Approaches positing institutional persistence as a matter of fact, and then thinking of institutional changes as unusual events will not be satisfactory. Both phenomena have to be analyzed as part of the same dynamic equilibrium framework.

One type of institutional change, consistent with the examples we discussed in this chapter, takes place when those who benefit from the existing set of institutions are forced to accept change, either because they are the losers in a process of fighting or because of the threat of internal revolution (another possibility is that they might accept

change because of the threat of external invasion). However, institutional change can also take place because the set of economic institutions that is optimal for a particular group with political power may vary over time as the state variables in the system and economic opportunities evolve. One example may be the end of slavery in the British Empire and another may be the economic and political changes introduced by Mikhail Gorbachev in the Soviet Union in the 1980s. We need more research on the dynamic mechanisms at work [see [Tornell \(1997\)](#) for a model of such a process].

Finally, it is important to understand the role of policy and interventions in changing the institutional equilibrium. Though social science research is of intrinsic interest, one would hope that a convincing fundamental theory of comparative growth based on institutions would lead to policy conclusions that would help us improve the institutions and thus the lives and welfare of people in poor countries. It should be obvious that, at the moment, we are a long way from being in a position to draw such conclusions. In a world where political choices are made rationally and are endogenous to the structure of institutions, which are themselves ultimately endogenous, giving policy advice is a conceptually complex issue [see [Acemoglu et al. \(2003\)](#) for reflections on this]. Recognizing our current ignorance on this topic in no way diminishes its importance, and its role as the Holy Grail of political economy research, however. And we believe that better and empirically more realistic theoretical frameworks in the future will take us closer to this Holy Grail.

## Acknowledgements

We thank the editors for their outstanding patience and Philippe Aghion, Robert Barro, Avinash Dixit, Leopoldo Fergusson, Herschel Grossman, Pablo Querubín and Barry Weingast for their helpful suggestions.

## References

- Acemoglu, D. (1995). "Reward structures and the allocation of talent". *European Economic Review* 39, 17–33.
- Acemoglu, D. (1997). "Training and innovation in an imperfect labor market". *Review of Economic Studies* 64, 445–464.
- Acemoglu, D. (2003a). "Why not a Political Coase Theorem?". *Journal of Comparative Economics* 31, 620–652.
- Acemoglu, D. (2003b). "The form of property rights: Oligarchic versus democratic societies". NBER Working Paper #10037.
- Acemoglu, D., Aghion, P., Zilibotti, F. (2002). "Distance to frontier, selection, and economic growth". NBER Working Paper #9066.
- Acemoglu, D., Johnson, S., Robinson, J.A. (2001). "The colonial origins of comparative development: An empirical investigation". *American Economic Review* 91 (5), 1369–1401.
- Acemoglu, D., Johnson, S., Robinson, J.A. (2002). "Reversal of fortune: Geography and institutions in the making of the modern world income distribution". *Quarterly Journal of Economics* 118, 1231–1294.

- Acemoglu, D., Johnson, S., Robinson, J.A. (2005). "The rise of Europe: Atlantic trade, institutional change and economic growth." *American Economic Review*, in press.
- Acemoglu, D., Robinson, J.A. (2000a). "Why did the West extend the franchise? Democracy, inequality and growth in historical perspective". *Quarterly Journal of Economics* 115, 1167–1199.
- Acemoglu, D., Robinson, J.A. (2000b). "Political losers as a barrier to economic development". *American Economic Review* 90, 126–130.
- Acemoglu, D., Robinson, J.A. (2001). "A theory of political transitions". *American Economic Review* 91, 938–963.
- Acemoglu, D., Robinson, J.A. (2002). "Economic backwardness in political perspective". NBER Working Paper #8831.
- Acemoglu, D., Robinson, J.A. (2003). *Economic Origins of Dictatorship and Democracy*, forthcoming book to be published by Cambridge University Press.
- Acemoglu, D., Zilibotti, F. (1997). "Was Prometheus unbound by chance? Risk, diversification and growth". *Journal of Political Economy* 105, 709–751.
- Acemoglu, D., Johnson, S., Robinson, J.A., Thaicharoen, Y. (2003). "Institutional causes, macroeconomic symptoms: Volatility, crises and growth". *Journal of Monetary Economics* 50, 49–123.
- Ades, A., Verdier, T. (1996). "The rise and fall of elites: A theory of economic development and social polarization in rent-seeking societies". CEPR Discussion Paper #1495.
- Aghion, P., Howitt, P.W. (1992). "A model of growth through creative destruction". *Econometrica* 60, 323–351.
- Aghion, P., Howitt, P.W. (1994). "Growth and unemployment". *Review of Economic Studies* 61, 477–494.
- Aghion, P., Howitt, P.W. (2005). "Growth with quality-improving innovations: An integrated framework". In: Aghion, P., Durlauf, S.N. (Eds.), *Handbook of Economic Growth*, vol. 1A. Elsevier, Amsterdam.
- Aghion, P., Harris, C., Howitt, P.W., Vickers, J. (2001). "Competition, imitation, and growth with step-by-step innovation". *Review of Economic Studies* 68, 467–492.
- Alesina, A., Perotti, R. (1996). "Income distribution, political instability and investment". *European Economic Review* 40, 1203–1225.
- Alesina, A., Rodrik, D. (1994). "Distributive politics and economic growth". *Quarterly Journal of Economics* 109, 465–490.
- Allen, R.C. (1982). "The efficiency and distributional consequences of eighteenth century enclosures". *Economic Journal* 92, 937–953.
- Allen, R.C. (1992). *Enclosure and the Yeoman*. Oxford University Press, New York.
- Aston, T.H., Philpin, C.H.E. (Eds.) (1985). *The Brenner Debate: Agrarian Class Structure and Economic Development in Pre-industrial Europe*. Cambridge University Press, New York.
- Azariadis, C., Stachurski, J. (2005). "Poverty traps". In: Aghion, P., Durlauf, S.N. (Eds.), *Handbook of Economic Growth*, vol. 1A. Elsevier, Amsterdam.
- Bairoch, P. (1988). *Cities and Economic Development: From the Dawn of History to the Present*. University of Chicago Press, Chicago.
- Bairoch, P., Batou, J., Chèvre, P. (1988). *La Population des villes Européennes de 800 à 1850: Banque de Données et Analyse Sommaire des Résultats*. Centre D'histoire Économique Internationale de l'Uni. de Genève, Librairie Droz, Geneva.
- Banerjee, A., Duflo, E. (2005). "Growth theory through the lens of development economics". In: Aghion, P., Durlauf, S.N. (Eds.), *Handbook of Economic Growth*, vol. 1A. Elsevier, Amsterdam.
- Banerjee, A., Newman, A.F. (1993). "Occupational choice and the process of development". *Journal of Political Economy* 101, 274–298.
- Banfield, E.C. (1958). *The Moral Basis of a Backward Society*. University of Chicago Press, Chicago.
- Bardhan, P., Udry, C. (1999). *Development Microeconomics*. Oxford University Press, New York.
- Barro, R.J. (1997). *The Determinants of Economic Growth: A Cross-Country Empirical Study*. MIT Press, Cambridge.
- Barro, R.J., McCleary, R. (2003). "Religion and economic growth". NBER Working Paper #9682.
- Barzel, Y. (2001). *A Theory of the State: Economic Rights, Legal Rights, and the Scope of the State*. Cambridge University Press, New York.

- Bates, R.H. (1981). Markets and States in Tropical Africa. University of California Press, Berkeley, CA.
- Bates, R.H. (1989). Beyond the Miracle of the Market. Cambridge University Press, New York.
- Bates, R.H. (1997). Open Economy Politics. Princeton University Press, Princeton.
- Becker, G.S. (1958). "Competition and democracy". *Journal of Law and Economics* 1, 105–109.
- Benhabib, J., Rustichini, A. (1996). "Social conflict and growth". *Journal of Economic Growth* 1, 125–142.
- Berman, B.J., Lonsdale, J. (1992). Unhappy Valley. James Currey, London.
- Besley, T.F., Coate, S.T. (1998). "Sources of inefficiency in a representative democracy: A dynamic analysis". *American Economic Review* 88, 139–156.
- Bloom, D.E., Sachs, J.D. (1998). "Geography, demography, and economic growth in Africa". *Brookings Papers on Economic Activity* 2, 207–295.
- Blum, J. (1943). "Transportation and industry in Austria, 1815–1848". *Journal of Modern History* 15, 24–38.
- Bourguignon, F., Verdier, T. (2000). "Oligarchy, democracy, inequality and growth". *Journal of Development Economics* 62, 285–313.
- Brenner, R. (1973). "The Civil War politics of Londons merchant community". *Past and Present* 58, 53–107.
- Brenner, R. (1976). "Agrarian class structure and economic development in preindustrial Europe". *Past and Present* 70, 30–75.
- Brenner, R. (1982). "Agrarian roots of European capitalism". *Past and Present* 97, 16–113.
- Brenner, R. (1993). Merchants and Revolution: Commercial Change, Political Conflict, and London's Overseas Traders, 1550–1653. Princeton University Press, Princeton.
- Brewer, J. (1988). The Sinews of Power: War, Money, and the English State, 1688–1783. Harvard University Press, Cambridge.
- Brunton, D., Pennington, D.H. (1954). Members of the Long Parliament. Allen and Unwin, London.
- Cameron, R. (1993). A Concise Economic History of the World. Oxford University Press, New York.
- Cantor, N.F. (2001). In the Wake of the Plague: The Black Death and the World it Made. The Free Press, New York.
- Cardoso, F.H., Faletto, E. (1979). Dependency and Development in Latin America. University of California Press, Berkeley, CA.
- Cass, D. (1965). "Optimum growth in an aggregate model of capital accumulation". *Review of Economic Studies* 32, 233–240.
- Chandler, T. (1987). Four Thousand Years of Urban Growth: An Historical Census. St. David's University Press, Lewiston, NY.
- Charlesworth, A. (1983). An Atlas of Rural Protest in Britain 1548–1900. Croon Helm, London.
- Coase, R.H. (1937). "The nature of the firm". *Economica* 3, 386–405.
- Coase, R.H. (1960). "The problem of social cost". *Journal of Law and Economics* 3, 1–44.
- Coatsworth, J.H. (1993). "Notes on the comparative economic history of Latin America and the United States". In: Bernecker, W.L., Werner Tobler, H. (Eds.), *Development and Underdevelopment in America: Contrasts in Economic Growth in North and Latin America in Historical Perspective*. Walter de Gruyter, New York.
- Cumings, B. (2004). North Korea: Another Country. The New Press, New York.
- Curtin, P.D. (1989). Death by Migration: Europe's Encounter with the Tropical World in the Nineteenth Century. Cambridge University Press, New York.
- Curtin, P.D. (1998). Disease and Empire: The Health of European Troops in the Conquest of Africa. Cambridge University Press, New York.
- Davis, R. (1973a). The Rise of the Atlantic Economies. Cornell University Press, Ithaca.
- Davis, R. (1973b). English Overseas Trade 1500–1700. Macmillan, London.
- Demsetz, H. (1967). "Toward a theory of property right". *American Economic Review* 57, 61–70.
- de Vries, J., van der Woude, A. (1997). The First Modern Economy: Success, Failure, and Perseverance of the Dutch Economy, 1500–1815. Cambridge University Press, New York.
- Diamond, J.M. (1997). Guns Germs and Steel: The Fate of Human Societies. W.W. Norton & Co., New York.
- Dixit, A.K. (1996). The Making of Economic Policy: A Transaction Cost Politics Perspective. MIT Press, Cambridge.

- Dixit, A.K. (2004). Lawlessness and Economics: Alternative Modes of Governance. Princeton University Press, Princeton.
- Dixit, A.K., Londregan, J.B. (1995). "Redistributive politics and economic efficiency". *American Political Science Review* 89, 856–866.
- Djankov, S., La Porta, R., Lopez-de-Silanes, F., Shleifer, A. (2002). "The regulation of entry". *Quarterly Journal of Economics* 117, 1–37.
- Djankov, S., La Porta, R., Lopez-de-Silanes, F., Shleifer, A. (2003). "Courts". *Quarterly Journal of Economics* 118, 453–517.
- Dobb, M.H. (1948). Studies in the Development of Capitalism. Cambridge University Press, Cambridge, UK.
- Domar, E. (1970). "The causes of slavery or serfdom: A hypothesis". *Journal of Economic History* 30, 18–32.
- Durlauf, S.N. (1993). "Nonergodic economic growth". *Review of Economic Studies* 60, 349–366.
- Durlauf, S.N., Fafchamps, M. (2004). "Empirical studies of social capital: A critical survey". In: Aghion, P., Durlauf, S.N. (Eds.), *Handbook of Economic Growth*, vol. 1A. Elsevier, Amsterdam.
- Edgerton, R.B. (1992). *Sick Societies: Challenging the Myth of Primitive Harmony*. Free Press, New York.
- Eggimann, G. (1999). *La Population des Villes des Tiers-Mondes, 1500–1950*. Centre d'histoire économique Internationale de l'Uni. de Genève, Librairie Droz, Geneva.
- Engerman, S.L., Sokoloff, K.L. (1997). "Factor endowments, institutions, and differential growth paths among new world economies". In: Haber, S. (Ed.), *How Latin America Fell Behind*. Stanford University Press, Stanford, CA.
- Ertman, T. (1997). *Birth of the Leviathan: Building States and Regimes in Medieval and Early Modern Europe*. Cambridge University Press, New York.
- Evans, E.J. (1983). *The Forging of the Modern State: Early Industrial Britain, 1783–1870*. Longman, New York.
- Farrell, J. (1987). "Information and the Coase Theorem". *Journal of Economic Perspectives* 1, 113–129.
- Frank, A.G. (1978). *Dependent Accumulation and Underdevelopment*. Macmillan, London.
- Freudenberger, H. (1967). "State intervention as an obstacle to economic growth in the Hapsburg Monarchy". *Journal of Economic History* 27, 493–509.
- Galor, O., Zeira, J. (1993). "Income distribution and macroeconomics". *Review of Economic Studies* 40, 35–52.
- Garrard, J. (2002). *Democratization in Britain: Elites, Civil Society and Reform since 1800*. Basingstoke, Palgrave.
- Gerschenkron, A. (1970). *Europe in the Russian Mirror: Four Lectures in Economic History*. Cambridge University Press, Cambridge, UK.
- Glaeser, E.L., Shleifer, A. (2002). "Legal origins". *Quarterly Journal of Economics* 117, 1193–1230.
- Graves, M.A.R. (2001). *The Parliaments of Early Modern Europe*. Longman, New York.
- Gregory, P.R. (1991). "The role of the state in promoting economic development: The Russian case and its general implications". In: Sylla, R., Toniolo, G. (Eds.), *Patterns of European Industrialization: The Nineteenth Century*. Routledge, New York.
- Greif, A. (1994). "Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies". *Journal of Political Economy* 102, 912–950.
- Gross, N. (1973). "The industrial revolution in the Hapsburg monarchy, 1750–1914". In: Cipolla, C.M. (Ed.), *The Fontana Economic History of Europe*, vol. 4. Fontana Books, London.
- Grossman, G.M., Helpman, E. (1991). *Innovation and Growth in the Global Economy*. MIT Press, Cambridge.
- Grossman, H., Kim, M. (1995). "Swords or ploughshares? A theory of the security of claims to property rights". *Journal of Political Economy* 103, 1275–1288.
- Grossman, H., Kim, M. (1996). "Predation and accumulation". *Journal of Economic Growth* 1, 333–350.
- Grossman, H., Noh, S.J. (1994). "Proprietary public finance and economic welfare". *Journal of Public Economics* 53, 187–204.
- Grossman, S.J., Hart, O.D. (1986). "The costs and benefits of ownership: A theory of vertical and lateral integration". *Journal of Political Economy* 94, 691–719.
- Gutierrez, H. (1986). "La mortalité des évêques Latino-Américains aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles". *Annales de Demographie Historique*, 29–39.

- Haber, S.H. (2001). "Political institutions and banking systems: Lessons from the economic histories of Mexico and the United States, 1790–1914". Unpublished. Department of Political Science, Stanford University.
- Haber, S.H., Maurer, N. (2004). "Related lending and economic performance: Evidence from Mexico". Unpublished. Department of Political Science, Stanford University.
- Hall, R.E., Jones, C.I. (1999). "Why do some countries produce so much more output per worker than others?". *Quarterly Journal of Economics* 114, 83–116.
- Hardin, R. (1995). All For One. Princeton University Press, Princeton.
- Hart, O.D. (1995). Firms, Contracts, and Financial Structure. Oxford University Press, New York.
- Herbst, J.I. (1990). State Politics in Zimbabwe. University of California Press, Berkeley.
- Hicks, J.R. (1969). A Theory of Economic History. Oxford University Press, New York.
- Hill, C. (1961a). The Century of Revolution, 1603–1714. W.W. Norton & Co., New York.
- Hill, C. (1961b). "Protestantism and the rise of capitalism". In: Fisher, F.J. (Ed.), Essays in the Economic and Social History of Tudor and Stuart England. Cambridge University Press, Cambridge.
- Hill, C. (1969). From Reformation to Industrial Revolution 1530–1780. Penguin Books, Baltimore.
- Hilton, R. (1981). Bond Men Made Free. Routledge, Oxford.
- Hirschleifer, J. (2001). The Dark Side of the Force: Economic Foundations of Conflict Theory. Cambridge University Press, New York.
- Horowitz, D.L. (1991). A Democratic South Africa? Constitutional Engineering in Divided Societies. University of California Press, Berkeley.
- Israel, J.I. (1989). Dutch Primacy in World Trade, 1585–1740. The Clarendon Press, Oxford.
- Israel, J.I. (1995). The Dutch Republic: Its Rise, Greatness and Fall 1477–1806. Oxford University Press, New York.
- Jones, E.L. (1981). The European Miracle: Environments, Economies, and Geopolitics in the History of Europe and Asia. Cambridge University Press, New York.
- Keeler, M. (1954). The Long Parliament, 1640–1641; A Biographical Study of its Members. American Philosophical Society, Philadelphia.
- Keyssar, A. (2000). The Right to Vote: The Contested History of Democracy in the United States. Basic Books, New York.
- Killick, T. (1978). Development Economics in Action: A Study of Economic Policies in Ghana. Heinemann, London.
- Knack, S., Keefer, P. (1995). "Institutions and economic performance: Cross-country tests using alternative measures". *Economics and Politics* 7, 207–227.
- Knack, S., Keefer, P. (1997). "Does social capital have an economic impact? A cross-country investigation". *Quarterly Journal of Economics* 112, 1252–1288.
- Koopmans, T.C. (1965). "On the concept of optimal economic growth". In: The Economic Approach to Development Planning. North-Holland, Amsterdam.
- Krugman, P.R., Venables, A.J. (1995). "Globalization and the inequality of nations". *Quarterly Journal of Economics* 110, 857–880.
- Krusell, P., Ríos-Rull, J.-V. (1996). "Vested interests in a theory of stagnation and growth". *Review of Economic Studies* 63, 301–330.
- Krusell, P., Ríos-Rull, J.-V. (1999). "On the size of government: Political economy in the neoclassical growth model". *American Economic Review* 89, 1156–1181.
- Kupperman, K.O. (1993). Providence Island, 1630–1641: The other Puritan Colony. Cambridge University Press, New York.
- Kuznets, S. (1968). Towards a Theory of Economic Growth. Yale University Press, New Haven, CT.
- Landes, D.S. (1998). The Wealth and Poverty of Nations: Why Some Are so Rich and Some so Poor. Norton & Co, New York.
- Lang, S. (1999). Parliamentary Reform, 1785–1928. Routledge, New York.
- La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R. (1998). "Law and finance". *Journal of Political Economy* 106, 1113–1155.

- La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R. (1999). "The quality of government". *Journal of Law, Economics and Organization* 15, 222–279.
- Lee, S.J. (1994). *Aspects of British Political History, 1815–1914*. Routledge, New York.
- Levine, R. (2005). "Finance and growth: Theory and evidence". In: Aghion, P., Durlauf, S.N. (Eds.), *Handbook of Economic Growth*, vol. 1A. Elsevier, Amsterdam.
- Lewis, W.A. (1978). *The Emergence of the International Economic Order*. Princeton University Press, Princeton.
- Li, H., Squire, L., Zou, H.-f. (1998). "Explaining international and intertemporal variations in income inequality". *Economic Journal* 108, 26–43.
- Lichbach, M.I. (1995). *The Rebel's Dilemma*. University of Michigan Press, Ann Arbor.
- Lindert, P.H. (2000). "Three centuries of inequality in Britain and America". In: Atkinson, A.B., Bourguignon, F. (Eds.), *Handbook of Income Distribution*. North-Holland, Amsterdam.
- Lindert, P.H. (2004). *Growing Public: Social Spending and Economics Growth since the Eighteenth Century*, vols. 1, 2. Cambridge University Press, Cambridge, MA.
- Lucas, R.E. (1988). "On the mechanics of economic development". *Journal of Monetary Economics* 22, 3–42.
- Maddison, A. (2001). *The World Economy: A Millennial Perspective*. Development Centre of the Organization for Economic Cooperation and Development, OECD, Paris.
- Marshall, A. (1890). *Principles of Economics*. Macmillan, London. 1949.
- Maurer, N. (2002). *The Power and the Profits: The Mexican Financial System, 1876–1932*. Stanford University Press, Stanford.
- McDaniel, T. (1991). *Autocracy, Modernization and Revolution in Russia and Iran*. Princeton University Press, Princeton.
- McEvedy, C., Jones, R. (1978). *Atlas of World Population History*. Facts on File, New York.
- Mitch, D. (1993). "The role of human capital in the first industrial revolution". In: Mokyr, J. (Ed.), *The British Industrial Revolution: An Economic Perspective*. Westview Press, San Francisco.
- Mokyr, J. (1990). *The Lever of Riches: Technological Creativity and Economic Progress*. Oxford University Press, New York.
- Montesquieu, C.S. (1748). *The Spirit of the Laws*. Cambridge University Press, New York. 1989.
- Moore, B. Jr. (1966). *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Beacon Press, Boston.
- Morgan, E.S. (1975). *American Slavery, American Freedom: The Ordeal of Colonial Virginia*. W.W. Norton & Co., New York.
- Mosse, W.E. (1958). *Alexsandr II and the Modernization of Russia*. University of London Press, London, UK.
- Mosse, W.E. (1992). *An Economic History of Russia, 1856–1914*. I.B. Taurus Press, London, UK.
- Murphy, K.M., Shleifer, A., Vishny, R.W. (1989a). "Industrialization and the big push". *Journal of Political Economy* 97, 1003–1026.
- Murphy, K.M., Shleifer, A., Vishny, R.W. (1989b). "Income distribution, market size and industrialization". *Quarterly Journal of Economics* 104, 537–564.
- Murphy, K.M., Shleifer, A., Vishny, R.W. (1991). "The allocation of talent: Implications for growth". *Quarterly Journal of Economics* 106, 503–530.
- Myrdal, G. (1968). *Asian Drama: An Inquiry into the Poverty of Nations*, vols. 1–3. Twentieth Century Fund, New York.
- Neal, L. (1990). *The Rise of Financial Capitalism: International Capital Markets in the Age of Reason*. Cambridge University Press, New York.
- Newton, A.P. (1914). *The Colonizing Activities of the English Puritans*. Yale University Press, New Haven.
- North, D.C. (1981). *Structure and Change in Economic History*. W.W. Norton & Co., New York.
- North, D.C. (1990). *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press, New York.
- North, D.C., Thomas, R.P. (1973). *The Rise of the Western World: A New Economic History*. Cambridge University Press, Cambridge, UK.

- North, D.C., Summerhill, W., Weingast, B.R. (2000). "Order, disorder, and economic change: Latin America versus North America". In: de Mesquita, B.B., Root, H.L. (Eds.), *Governing for Prosperity*. Yale University Press, New Haven.
- North, D.C., Weingast, B.R. (1989). "Constitutions and commitment: Evolution of institutions governing public choice in seventeenth century England". *Journal of Economic History* 49, 803–832.
- Olson, M.C. (1982). *The Rise and Decline of Nations: Economic Growth, Stagflation, and Economic Rigidities*. Yale University Press, New Haven and London.
- Olson, M.C. (1993). "Dictatorship, democracy and development". *American Political Science Review* 87, 567–575.
- Olson, M.C. (2000). *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*. Basic Books, New York.
- Overton, M. (1996). *Agricultural Revolution in England: The Transformation of the Agrarian Economy 1500–1850*. Cambridge University Press, New York.
- Parente, S., Prescott, E.C. (1999). "Monopoly rights as barriers to riches". *American Economic Review* 89, 1216–1233.
- Parente, S., Prescott, E.C. (2005). "A unified theory of the evolution of international income levels". In: Aghion, P., Durlauf, S.N. (Eds.), *Handbook of Economic Growth*, vol. 1B. Elsevier, Amsterdam.
- Parry, J.H. (1966). *The Spanish Seaborne Empire*. University of California Press, Berkeley.
- Perotti, R. (1993). "Political equilibrium, income distribution and growth". *Review of Economic Studies* 61, 755–776.
- Persson, T., Tabellini, G. (1994). "Is inequality harmful for growth?". *American Economic Review* 84, 600–621.
- Piketty, T. (1995). "Social mobility and redistributive politics". *Quarterly Journal of Economics* 100, 551–584.
- Pincus, S. (1998). "Neither Machiavellian moment nor possessive individualism: Commercial society an the defenders of the English commonwealth". *American Historical Review* 103, 705–736.
- Pincus, S. (2001). "From holy cause to economic interest: The study of population and the invention of the state". In: Houston, A., Pincus, S. (Eds.), *A Nation Transformed: England after the Restoration*. Cambridge University Press, New York.
- Pincus, S. (2002). "Civic republicanism and political economy in an age of revolution: Law, politics, and economics in the revolution of 1688–89". Unpublished. Department of History, University of Chicago.
- Pirenne, H. (1937). *Economic and Social History of Medieval Europe*. Harcourt, Brace and Company, New York.
- Pissarides, C. (2000). *Equilibrium Unemployment Theory*, second edn. MIT Press, Cambridge.
- Postan, M.M. (1937). "The chronology of labour services". *Transactions of the Royal Historical Society* 20.
- Postan, M.M. (1966). "Medieval agrarian society in its prime: England". In: Postan, M.M. (Ed.), *The Cambridge Economic History of Europe*. Cambridge University Press, London.
- Powell, R. (2004). "The inefficient use of power: Costly conflict with complete information". *American Political Science Review* 98, 231–241.
- Putnam, R.D., Leonardi, R., Nanetti, R.Y. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton.
- Randall, A. (1991). *Before the Luddites: Custom, Community, and Machinery in the English Woollen Industry, 1776–1809*. Cambridge University Press, New York.
- Ray, D. (1998). *Development Economics*. Princeton University Press, Princeton.
- Reynolds, A. (1999). *Electoral Systems and Democratization in Southern Africa*. Oxford University Press, New York.
- Ringer, F. (1979). *Education and Society in Modern Europe*. University of Indiana Press, Bloomington.
- Robinson, J.A. (1998). "Theories of bad policy". *Journal of Policy Reform* 3, 1–46.
- Rodney, W. (1972). *How Europe Underdeveloped Africa*. Howard University Press, Washington, DC.
- Rodrik, D. (1999). "Democracies pay higher wages". *Quarterly Journal of Economics* CXIV, 707–738.
- Romer, D. (2003). "Misconceptions and political outcomes". *Economic Journal* 113, 1–20.

- Romer, P.M. (1986). "Increasing returns and long-run growth". *Journal of Political Economy* 94, 1002–1037.
- Romer, P.M. (1990). "Endogenous technical change". *Journal of Political Economy* 98, 71–102.
- Rosenstein-Rodan, P. (1943). "Problems of industrialization in Eastern and South-Eastern Europe". *Economic Journal* 53, 202–211.
- Ross, J.I., Gurr, T.R. (1989). "Why terrorism subsides: A comparative study of Canada and the United States". *Comparative Politics* 21, 405–426.
- Sachs, J.D. (2000). "Notes on a new sociology of economic development". In: Harrison, L.E., Huntington, S.P. (Eds.), *Culture Matters: How Values Shape Human Progress*. Basic Books, New York.
- Sachs, J.D. (2001). "Tropical underdevelopment". NBER Working Paper #8119.
- Saint-Paul, G., Verdier, T. (1993). "Education, democracy, and growth". *Journal of Development Economics* 42, 399–407.
- Schattschneider, E.E. (1935). *Politics, Pressures and the Tariff; a Study of Free Private Enterprise in Pressure Politics*, as Shown in the 1929–1930 Revision of the Tariff. Prentice-Hall, New York.
- Scott, J.C. (2000). "The moral economy as an argument and as a fight". In: Randall, A., Charlesworth, A. (Eds.), *Moral Economy and Popular Protest: Crowds, Conflict and Authority*. MacMillan, London.
- Skaperdas, S. (1992). "Cooperation, conflict, and power in the absence of property rights". *American Economic Review* 82, 720–739.
- Smith, A. (1776). *The Wealth of Nations*, vols. 1, 2. Penguin Classics, London. 1999.
- Solow, R.M. (1956). "A contribution to the theory of economic growth". *Quarterly Journal of Economics* 70, 65–94.
- Stasavage, D. (2003). *Public Debt and the Birth of the Democratic State: France and Great Britain, 1688–1789*. Cambridge University Press, New York.
- Stevenson, J. (1979). *Popular Disturbances in England, 1700–1870*. Longman, New York.
- Tarrow, S. (1991). "Aiming at a moving target: Social science and the recent rebellions in Eastern Europe". *PS: Political Science and Politics* 24, 12–20.
- Tarrow, S. (1998). *Power in Movement: Social Movements and Contentious Politics*, second edn. Cambridge University Press, New York.
- Tawney, R.H. (1926). *Religion and the Rise of Capitalism: A Historical Study*. J. Murray, London.
- Tawney, R.H. (1941). "The rise of the gentry, 1558–1640". *Economic History Review* 11, 1–38.
- Thomis, M.I. (1970). *The Luddites; Machine-Breaking in Regency England*. David & Charles, Newton Abbot.
- Tilly, C. (1995). *Popular Contention in Britain, 1758–1834*. Harvard University Press, Cambridge.
- Tornell, A. (1997). "Economic growth and decline with endogenous property rights". *Journal of Economic Growth* 2, 219–250.
- Tornell, A., Velasco, A. (1992). "Why does capital flow from poor to rich countries? The tragedy of the commons and economic growth". *Journal of Political Economy* 100, 1208–1231.
- Townsend, R.M. (1993). *The Medieval Village Economy: A Study of the Pareto Mapping in General Equilibrium Models*. Princeton University Press, Princeton.
- Vargas Llosa, M. (1989). *The Storyteller*. Farrar, Straus, Giroux, New York.
- Veitch, J.M. (1986). "Repudiations and confiscations by the medieval state". *Journal of Economic History* 46, 31–36.
- Vélez, C. (1994). *The New World of the Gothic Fox: Culture and Economy in English and Spanish America*. University of California Press, Berkeley.
- Wallerstein, I.M. (1974–1980). *The Modern World-System*, vols. 1–3. Academic Press, New York.
- Weber, M. (1930). *The Protestant Ethic and the Spirit of Capitalism*. Allen and Unwin, London.
- Weber, M. (1958). *The Religion of India*. Free Press, Glencoe.
- Weingast, B.R. (1997). "The political foundations of democracy and the rule of law". *American Political Science Review* 91, 245–263.
- Weingast, B.R. (1998). "Political stability and Civil War: Institutions, commitment, and American democracy". In: Bates, R., Greif, A., Levi, M., Rosenthal, J.-L., Weingast, B.R. (Eds.), *Analytic Narratives*. Princeton University Press, Princeton.
- Wiarda, H.J. (2001). *The Soul of Latin America: The Cultural and Political Tradition*. Yale University Press, New Haven.

- Williams, E.E. (1944). Capitalism and Slavery. University of North Carolina Press, Chapel Hill.
- Williamson, O. (1985). The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting. Free Press, New York.
- Wittman, D. (1989). "Why democracies produce efficient results". *Journal of Political Economy* 97, 1395–1424.



Journal of Economic Growth, 9, 131–165, 2004  
© 2004 Kluwer Academic Publishers. Manufactured in The Netherlands.

# Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development\*

DANI RODRIK

*Kennedy School of Government, Harvard University, Cambridge, MA, USA*

ARVIND SUBRAMANIAN

*International Monetary Fund (IMF), Washington DC, USA*

FRANCESCO TREBBI

*Department of Economics, Harvard University, Littauer center, Cambridge, MA, USA*

We estimate the respective contributions of institutions, geography, and trade in determining income levels around the world, using recently developed instrumental variables for institutions and trade. Our results indicate that the quality of institutions “trumps” everything else. Once institutions are controlled for, conventional measures of geography have at best weak direct effects on incomes, although they have a strong indirect effect by influencing the quality of institutions. Similarly, once institutions are controlled for, trade is almost always insignificant, and often enters the income equation with the “wrong” (i.e., negative) sign. We relate our results to recent literature, and where differences exist, trace their origins to choices on samples, specification, and instrumentation.

**Keywords:** growth, institutions, openness, geography

**JEL classification:** F1, N7, O1

Commerce and manufactures can seldom flourish long in any state which does not enjoy a regular administration of justice, in which the people do not feel themselves secure in the possession of their property, in which the faith of contracts is not supported by law, and in which the authority of the state is not supposed to be regularly employed in enforcing the payment of debts from all those who are able to pay. Commerce and manufactures, in short, can seldom flourish in any state in which there is not a certain degree of confidence in the justice of government.

Adam Smith, *Wealth of Nations*

---

\* The views expressed in this paper are the authors' own and not of the institutions with which they are affiliated.

## 1. Introduction

Average income levels in the world's richest and poorest nations differ by a factor of more than 100. Sierra Leone, the poorest economy for which we have national income statistics, has a per capita GDP of \$490, compared to Luxembourg's \$50,061.<sup>1</sup> What accounts for these differences, and what (if anything) can we do to reduce them? It is hard to think of any question in economics that is of greater intellectual significance, or of greater relevance to the vast majority of the world's population.

In the voluminous literature on this subject, three strands of thoughts stand out. First, there is a long and distinguished line of theorizing that places geography at the center of the story. Geography is a key determinant of climate, endowment of natural resources, disease burden, transport costs, and diffusion of knowledge and technology from more advanced areas. It exerts therefore a strong influence on agricultural productivity and the quality of human resources. Recent writings by Jared Diamond and Jeffrey Sachs are among the more notable works in this tradition (see Diamond, 1997; Gallup et al., 1998; Sachs, 2001).

A second camp emphasizes the role of international trade as a driver of productivity change. We call this the integration view, as it gives market integration, and impediments thereof, a starring role in fostering economic convergence between rich and poor regions of the world. Notable recent research in this camp includes Frankel and Romer (1999) and the pre-geography work of Sachs (Sachs and Warner, 1995).<sup>2</sup> It may be useful to distinguish between "moderate" and "maximal" versions of this view. Much of the economics profession would accept the hypothesis that trade can be an underlying source of growth once certain institutional pre-requisites have been fulfilled. But a more extreme perspective, and one that has received wide currency in public debates, is that trade/integration is the major determinant of whether poor countries grow or not. It is the latter perspective that characterizes such widely cited papers as Sachs and Warner (1995) and Dollar and Kraay (2004).

Finally, a third group of explanations centers on institutions, and in particular the role of property rights and the rule of law. In this view, what matters are the rules of the game in a society and their conduciveness to desirable economic behavior. This view is associated most strongly with Douglass North (1990). It has received careful econometric treatment recently in Hall and Jones (1999), who focus on what they call "social infrastructure," and in Acemoglu et al. (2001), who focus on the expropriation risk that current and potential investors face.

Growth theory has traditionally focussed on physical and human capital accumulation, and, in its endogenous growth variant, on technological change. But accumulation and technological change are at best proximate causes of economic growth. No sooner have we ascertained the impact of these two on growth—and with some luck their respective roles

1 These are figures for 2000, and they are expressed in current "international" dollars, adjusted for PPP differences. The source is the World Development Indicators CD-ROM of the World Bank.

2 One can question whether it is appropriate to treat trade as one of the ultimate determinants of economic prosperity, but here we are simply following a long literature that has attached central causal importance to it.

also—that we want to ask: But why did some societies manage to accumulate and innovate more rapidly than others? The three-fold classification offered above—geography, integration, and institutions—allows us to organize our thoughts on the “deeper” determinants of economic growth. These three are the factors that determine which societies will innovate and accumulate, and therefore develop, and which will not.

Since long-term economic development is a complex phenomenon, the idea that any one (or even all) of the above deep determinants can provide an adequate accounting of centuries of economic history is, on the face of it, preposterous. Historians and many social scientists prefer nuanced, layered explanations where these factors interact with human choices and many other not-so-simple twists and turns of fate. But economists like parsimony. We want to know how well these simple stories do, not only on their own or collectively, but more importantly, vis-à-vis each other. How much of the astounding variation in cross-national incomes around the world can geography, integration, and institutions explain? Do these factors operate additively, or do they interact? Are they all equally important? Does one of the explanations “trump” the other two?

The questions may be simple, but devising a reasonable empirical strategy for answering them is far from straightforward. This is not because we do not have good empirical proxies for each of these deep determinants. There are many reasonable measures of “geography,” such as distance from the equator (our preferred measure), percentage land mass located in the tropics, or average temperature. The intensity of an economy’s integration with the rest of the world can be measured by flows of trade or the height of trade barriers. The quality of institutions can be measured with a range of perceptions-based indicators of property rights and the rule of law. The difficulty lies instead in sorting out the complex web of causality that entangles these factors.

The extent to which an economy is integrated with the rest of the world and the quality of its institutions are both endogenous, shaped potentially not just by each other and by geography, but also by income levels. Problems of endogeneity and reverse causality plague any empirical researcher trying to make sense of the relationships among these causal factors. We illustrate this with the help of Figure 1, adapted from Rodrik (2003). The plethora of arrows in the figure, going in both directions at once in many cases, exemplifies the difficulty.

The task of demonstrating causality is perhaps easiest for the geographical determinists. Geography is as exogenous a determinant as an economist can ever hope to get, and the main burden here is to identify the main channel(s) through which geography influences economic performance. Geography may have a direct effect on incomes, through its effect on agricultural productivity and morbidity. This is shown with arrow (1) in Figure 1. It can also have an indirect effect through its impact on distance from markets and the extent of integration (arrow (2)) or its impact on the quality of domestic institutions (arrow (3)). With regard to the latter, economic historians have emphasized the disadvantageous consequences for institutional development of certain patterns of factor endowments, which engender extreme inequalities and enable the entrenchment of a small group of elites (e.g., Engerman and Sokoloff, 1994). A similar explanation, linking ample endowment of natural resources with stunted institutional development, also goes under the name of “resource curse” (Sala-i-Martin and Subramanian, 2003).

Trade fundamentalists and institutionalists have a considerably more difficult job to do,

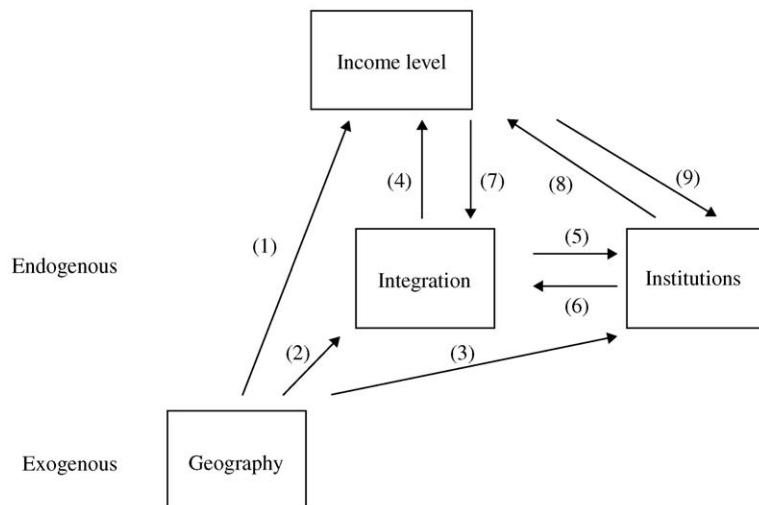


Figure 1. The “deep” determinants of income.

since they have to demonstrate causality for their preferred determinant, as well as identify the effective channel(s) through which it works. For the former, the task consists of showing that arrows (4) and (5)—capturing the direct impact of integration on income and the indirect impact through institutions, respectively—are the relevant ones, while arrows (6) and (7)—reverse feedbacks from incomes and institutions, respectively—are relatively insignificant. Reverse causality cannot be ruled out easily, since expanded trade and integration can be mainly the result of increased productivity in the economy and/or improved domestic institutions, rather than a cause thereof.

Institutionalists, meanwhile, have to worry about different kinds of reverse causality. They need to show that improvements in property rights, the rule of law and other aspects of the institutional environment are an independent determinant of incomes (arrow (8)), and are not simply the consequence of higher incomes (arrow (9)) or of greater integration (arrow (5)).

In econometric terms, what we need to sort all this out are good instruments for integration and institutions—sources of exogenous variation for the extent of integration and institutional quality, respectively, that are uncorrelated with other plausible (and excluded) determinants of income levels. Two recent papers help us make progress by providing plausible instruments. Frankel and Romer (1999) suggests that we can instrument for actual trade/GDP ratios by using trade/GDP shares constructed on the basis of a gravity equation for bilateral trade flows. The Frankel and Romer approach consists of first regressing bilateral trade flows (as a share of a country’s GDP) on measures of country mass, distance between the trade partners, and a few other geographical variables, and then constructing a predicted aggregate trade share for each country on the basis of the coefficients estimated. This constructed trade share is then used as an instrument for actual trade shares in estimating the impact of trade on levels of income.

Acemoglu et al. (2001) use mortality rates of colonial settlers as an instrument for

institutional quality. They argue that settler mortality had an important effect on the type of institutions that were built in lands that were colonized by the main European powers. Where the colonizers encountered relatively few health hazards to European settlement, they erected solid institutions that protected property rights and established the rule of law. In other areas, their interests were limited to extracting as much resources as quickly as possible, and they showed little interest in building high-quality institutions. Under the added assumption that institutions change only gradually over time, Acemoglu et al. argue that settler mortality rates are therefore a good instrument for institutional quality. Frankel and Romer (1999) and Acemoglu et al. (2001) use their respective instruments to demonstrate strong causal effects from trade (in the case of Frankel and Romer) and institutions (in the case of Acemoglu et al.) to incomes. But neither paper embeds their estimation in the broader framework laid out above. More specifically, Acemoglu et al. control for geographical determinants, but do not check for the effects of integration. Frankel and Romer do not control for institutions.

Our approach in this paper consists of using the Frankel and Romer and Acemoglu et al. instruments simultaneously to estimate the structure shown in Figure 1. The idea is that these two instruments, having passed what might be called the American Economic Review (AER)-test, are our best hope at the moment of unraveling the tangle of cause-and-effect relationships involved. So we systematically estimate a series of regressions in which incomes are related to measures of geography, integration, and institutions, with the latter two instrumented using the Frankel and Romer and Acemoglu et al. instruments, respectively. These regressions allow us to answer the question: what is the independent contribution of these three sets of deep determinants to the cross-national variation in income levels? The first stage of these regressions provides us in turn with information about the causal links among the determinants.

This exercise yields some sharp and striking results. Most importantly, we find that the quality of institutions trumps everything else. Once institutions are controlled for, integration has no direct effect on incomes, while geography has at best weak direct effects. Trade often enters the income regression with the “wrong” (i.e., negative) sign, as do many of the geographical indicators. By contrast, our measure of property rights and the rule of law always enters with the correct sign, and is statistically significant, often with *t*-statistics that are very large.

On the links among determinants, we find that institutional quality has a positive and significant effect on integration. Our results also tend to confirm the findings of Easterly and Levine (2003), namely that geography exerts a significant effect on the quality of institutions, and via this channel on incomes.<sup>3</sup>

---

3 The Easterly and Levine approach is in some ways very similar to that in this paper. Easterly and Levine estimate regressions of the levels of income on various measures of endowments, institutions, and “policies.” They find that institutions exert an important effect on development, while endowments do not, other than through their effect on institutions. Policies also do not exert any independent effect on development. The main differences between our paper and Easterly and Levine are the following. First, we use a larger sample of countries (79 and 137) to run the “horse” race between the three possible determinants. The Easterly and Levine sample is restricted to 72 countries. Second, Easterly and Levine do not test in any detail whether integration has an effect on development. For them, integration or open

Our preferred specification “accounts” for about half of the variance in incomes across the sample, with institutional quality (instrumented by settler mortality) doing most of the work. Our estimates indicate that an increase in institutional quality of one standard deviation, corresponding roughly to the difference between measured institutional quality in Bolivia and South Korea, produces a two log-points rise in per capita incomes, or a 6.4-fold difference—which, not coincidentally, is also roughly the income difference between the two countries. In our preferred specification, trade and distance from the equator both exert a negative, but insignificant effect on incomes (see Table 3, panel A, column (6)).

Much of our paper is devoted to checking the robustness of our central results. In particular, we estimate our model for three different samples: (a) the original 64-country sample used by Acemoglu et al.; (b) a 79-country sample which is the largest sample we can use while still retaining the Acemoglu et al. instrument; and (c) a 137-country sample that maximizes the number of countries at the cost of replacing the Acemoglu et al. instrument with two more widely available instruments (fractions of the population speaking English and Western European languages as the first language, from Hall and Jones, 1999.) We also use a large number of alternative indicators of geography and integration. In all cases, institutional quality emerges as the clear winner of the “horse race” among the three. Finally, we compare and contrast our results to those in some recent papers that have undertaken exercises of a similar sort. Where there are differences in results, we identify and discuss the source of the differences and explain why we believe our approach is superior on conceptual or empirical grounds.<sup>4</sup>

One final word about policy. As we shall emphasize at the end of the paper, identifying the deeper determinants of prosperity does not guarantee that we are left with clearcut policy implications. For example, finding that the “rule of law” is causally implicated in development does not mean that we actually know how to increase it under the specific conditions of individual countries. Nor would finding that “geography matters” necessarily imply geographic determinism—it may simply help reveal the roadblocks around which policy makers need to navigate. The research agenda to which this paper contributes is one that clarifies the priority of pursuing different objectives—improving the quality of domestic institutions, achieving integration into the world economy, or overcoming geographical adversity—but says very little about how each one of these is best achieved.

---

trade policy is part of a wider set of government policies that can affect development. Testing for the effect of policies in level regressions is, however, problematic as discussed in greater detail below. Policies pursued over a short time span, say 30–40 years, are like a flow variable, whereas development, the result of a much longer cumulative historical process, is more akin to a stock variable. Thus, level regressions that use policies as regressors conflate stocks and flows.

<sup>4</sup> We note that many of the papers already cited as well as others have carried out similar robustness tests. For example, Acemoglu et al. (2001) document that geographic variables such as temperature, humidity, malaria risk exert no independent direct effects on income once institutions are controlled for. A follow-up paper by the same authors (Acemoglu et al., 2003) shows that macroeconomic policies have limited effects after institutions are controlled. Easterly and Levine (2003) produce similar robustness results on the geography front. Our contribution is to put these and other tests in a broader framework, including trade, and to provide an interpretation of the results which we think is more appropriate.

The plan of the paper is as follows. Section 2 presents the benchmark results and robustness tests. Section 3 provides a more in-depth interpretation of our results and lays out a research agenda.

## 2. Core Results and Robustness

### 2.1. Data and Descriptive Statistics

Table 1 provides descriptive statistics for the key variables of interest. The first column covers the sample of 79 countries for which data on settler mortality have been compiled by Acemoglu et al.<sup>5</sup> Given the demonstrated attractiveness of this variable as an instrument that can help illuminate causality, this will constitute our preferred sample. The second column contains summary statistics for a larger sample of 137 countries for which we have data on alternative instruments for institutions (fractions of the population speaking

*Table 1.* Descriptive statistics.

	Extended Acemoglu et al. Sample (79 countries)	Large Sample (137 countries)
Log GDP per capita (PPP) in 1995 (LCGDP95)	8.03 (1.05)	8.41 (1.14)
Rule of law (RULE)	-0.25 (0.86)	0.08 (0.95)
Log openness (LCOPEN)	3.94 (0.61)	4.01 (0.57)
Distance from equator in degrees (DISTEQ)	15.37 (11.16)	23.98 (16.26)
Log European settler mortality (LOGEM4) (deaths per annum per 1,000 population)	4.65 (1.22)	— —
Log constructed openness (LOGFRANKROM)	2.76 (0.76)	2.91 (0.79)
Fraction of population speaking other European language (EURFRAC)	0.30 (0.41)	0.24 (0.39)
Fraction of population speaking English (ENGFRAC)	0.11 (0.29)	0.08 (0.24)

Notes: Standard deviations are reported below the means. Rule of law ranges between -2.5 and +2.5. Openness is measured as the ratio of trade to GDP. Constructed openness—the instrument for openness—is the predicted trade share and is from Frankel and Romer (1999). The Appendix describes in detail all the data and their sources.

---

<sup>5</sup> Acemoglu et al. actually compiled data on settler mortality for 81 countries, but data on our other variables are unavailable for Afghanistan (for per capita PPP GDP for 1995) and the Central African Republic (for rule of law).

English and other European languages). Data for the Frankel and Romer instrument on trade, on which we will rely heavily, are also available for this larger sample.

GDP per capita on a PPP basis for 1995 will be our measure of economic performance. For both samples, there is substantial variation in GDP per capita: for the 79-country sample, mean GDP in 1995 is \$3,072, the standard deviation of log GDP is 1.05, with the poorest country's (Congo, DRC) GDP being \$321 and that of the richest (Singapore) \$28,039. For the larger sample, mean income is \$4,492, the standard deviation is 1.14, with the richest country (Luxembourg) enjoying an income level of \$34,698.

The institutional quality measure that we use is due to Kaufmann et al. (2002). This is a composite indicator of a number of elements that capture the protection afforded to property rights as well as the strength of the rule of law.<sup>6</sup> This is a standardized measure that varies between -2.5 (weakest institutions) and 2.5 (strongest institutions). In our sample of 79 countries, the mean score is -0.25, with Zaire (score of -2.09) having the weakest institutions and Singapore (score of 1.85) the strongest.

Integration, measured using the ratio of trade to GDP, also varies substantially in our sample. The average ratio is 51.4 percent, with the least "open" country (India) posting a ratio of 13 percent and the most "open" (Singapore) a ratio of 324 percent. Our preferred measure of geography is a country's distance from the equator (measured in degrees). The typical country is about 15.4 degrees away from the equator.

## 2.2. OLS and IV Results in the Core Specifications

Our paper represents an attempt to estimate the following equation:

$$\log y_i = \mu + \alpha \text{INS}_i + \beta \text{INT}_i + \gamma \text{GEO}_i + \varepsilon_i, \quad (1)$$

where  $y_i$  is income per capita in country  $i$ ,  $\text{INS}_i$ ,  $\text{INT}_i$ , and  $\text{GEO}_i$  are respectively measures for institutions, integration, and geography, and  $\varepsilon_i$  is the random error term. Throughout the paper, we will be interested in the size, sign, and significance of the three coefficients  $\alpha$ ,  $\beta$ , and  $\gamma$ . We will use standardized measures of  $\text{INS}_i$ ,  $\text{INT}_i$ , and  $\text{GEO}_i$  in our core regressions, so that the estimated coefficients can be directly compared.

Before we discuss the benchmark results, it is useful to look at the simple, bivariate relationships between income and each of the "deep determinants." Figure 2 shows these scatter plots, with the three panels on the left hand side corresponding to the sample of 79 countries and the three panels on the right to the larger sample of 137 countries. All the plots show a clear and unambiguously positive relationship between income and its possible determinants. Thus, any or all of them have the potential to explain levels of income. This positive relationship is confirmed by the simple OLS regression of equation

---

<sup>6</sup> Acemoglu et al. (2001) use an index of protection against expropriation compiled by Political Risk Services. The advantage of the rule of law measure used in this paper is that it is available for a larger sample of countries, and in principle captures more elements that go toward determining institutional quality. In any case, measures of institutional quality are highly correlated: in our 79-country sample, the two measures have a simple correlation of 0.78.

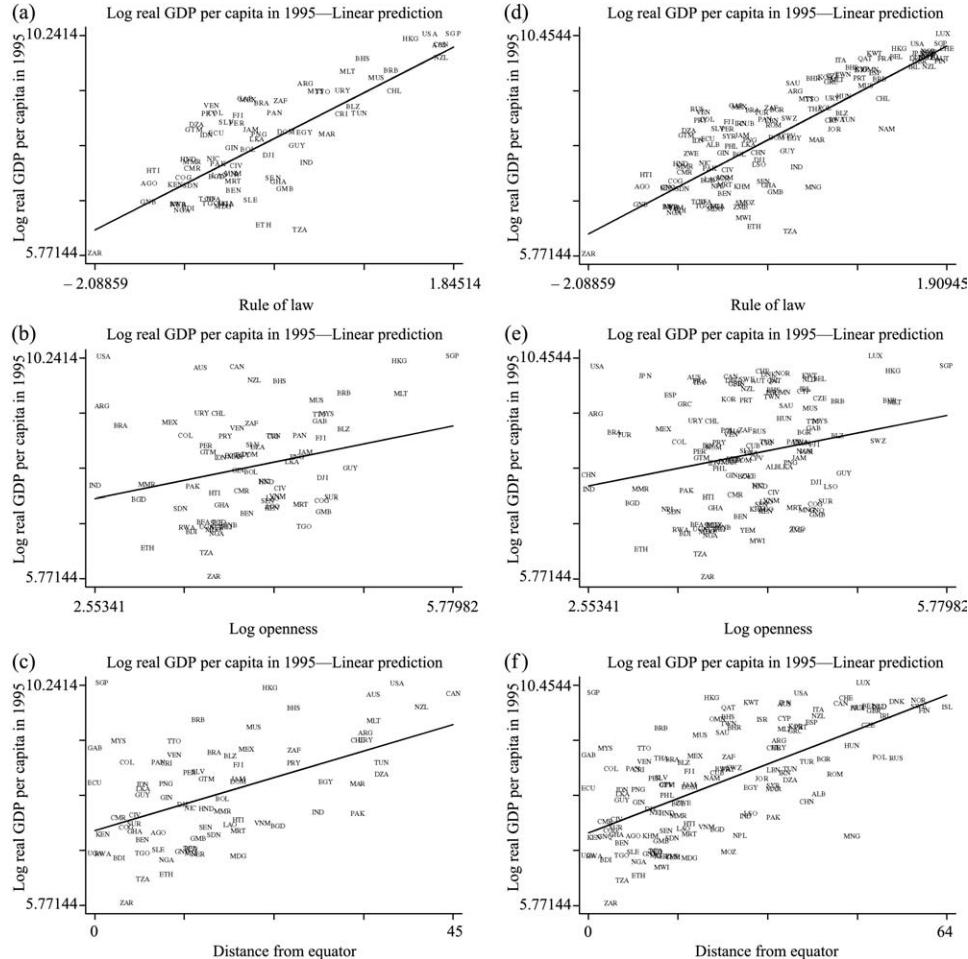


Figure 2. Simple correlations between income and its determinants (sample of 79 countries for (a)–(c); sample of 137 countries for (d)–(f)).

(1) reported in column (6) of Table 2. The signs of institution, openness, and geography are as expected and statistically significant or close to being so. Countries with stronger institutions, more open economies, and more distant from the equator are likely to have higher levels of income.

To get a sense of the magnitude of the potential impacts, we can compare two countries, say Nigeria and Mauritius, both in Africa. If the OLS relationship is indeed causal, the coefficients in column (6) of Table 2 would suggest that Mauritius's per capita GDP should be 10.3 times that of Nigeria, of which 77 percent would be due to better institutions, 9 percent due to greater openness, and 14 percent due to better location. In practice, Mauritius's income (\$11,400) is 14.8 times that of Nigeria (\$770).

Of course, for a number of reasons described extensively in the literature—reverse

Table 2. Determinants of development: Core specifications, ordinary least squares estimates.

Dependent Variable	Log GDP per capita								
	Acemoglu et al. Sample			Extended Acemoglu et al. Sample			Large Sample		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Geography (DISTEQ)	0.74 (4.48)*	0.20 (1.34)	0.32 (1.85)**	0.80 (5.22)*	0.22 (1.63)	0.33 (2.11)**	0.76 (10.62)*	0.20 (2.48)**	0.23 (2.63)*
Institutions (RULE)		0.78 (7.56)*	0.69 (6.07)*		0.81 (9.35)*	0.72 (6.98)*		0.81 (12.12)*	0.78 (10.49)*
Integration (LCOPEN)			0.16 (1.48)			0.15 (1.53)			0.08 (1.24)
Observations	64	64	64	79	79	79	137	137	137
R-square	0.25	0.57	0.59	0.26	0.61	0.62	0.42	0.71	0.71

Notes: The dependent variable is per capita GDP in 1995, PPP basis. There are three samples for which the core regressions are run: (i) the first three columns correspond to the sample of 64 countries in Acemoglu et al. (2001); (ii) columns (4)–(6) use a sample of 79 countries for which data on settler mortality (LOGEM4) have been compiled by Acemoglu et al.; and (iii) columns (7)–(9) use a larger sample of 137 countries. The regressors are: (i) DISTEQ, the variable for geography, which is measured as the absolute value of latitude of a country; (ii) Rule of law (RULE), which is the measure for institutions; and (iii) LCOPEN, the variable for integration, which is measured as the ratio of nominal trade to nominal GDP. All regressors are scaled in the sense that they represent deviations from the mean divided by the standard deviation. All regressors, except DISTEQ and RULE, in the three panels are in logs. See the Appendix for more detailed variable definitions and sources. *t*-statistics are reported under coefficient estimates. Significance at the 1, 5, and 10 percent levels are denoted respectively by \*, \*\*, and \*\*\*.

causality, omitted variables bias, and measurement error—the above relationship cannot be interpreted as causal or accurate. To address these problems, we employ a two-stage least squares estimation procedure. The identification strategy is to use the Acemoglu et al. settler mortality measure as an instrument for institutions and the Frankel and Romer measure of constructed trade shares as an instrument for integration. In the first-stage regressions,  $INS_i$  and  $INT_i$  are regressed on all the exogenous variables. Thus

$$INS_i = \lambda + \delta SM_i + \phi CONST_i + \psi GEO_i + \varepsilon_{INS_i}, \quad (2)$$

$$INT_i = \theta + \sigma CONST_i + \tau SM_i + \omega GEO_i + \varepsilon_{INT_i}, \quad (3)$$

where  $SM_i$  refers to settler mortality and  $CONST_i$  to the Frankel and Romer instrument for trade/GDP. The exclusion restrictions are that  $SM_i$  and  $CONST_i$  do not appear in equation (1).

Equations (1)–(3) are our core specification. This specification represents, we believe, the most natural framework for estimating the respective impacts of our three deep determinants. It is general, yet simple, and treats each of the three deep determinants symmetrically, giving them all an equal chance. Our proxies for institutions, integration, and geography are the ones that the advocates of each approach have used. Our instruments for institutions and integration are sensible, and have already been

demonstrated to “work” in the sense of producing strong second-stage results (albeit in estimations not embedded in our broader framework).

Panel A of Table 3 reports the two-stage least squares estimates of the three coefficients of interest. The estimation is done for three samples of countries: (i) for the sample of 64 countries analyzed by Acemoglu et al.; (ii) for an extended sample of 79 countries for which Acemoglu et al. had compiled data on settler mortality; and (iii) for a larger sample of 137 countries that includes those that were not colonized. In Acemoglu et al., the quality of institutions was measured by an index of protection against expropriation. We use a rule of law index because it is available for a larger sample. The IV estimates of the coefficient on institutions in the first three columns of panel A are very similar to those in Acemoglu et al., confirming that these two indexes are capturing broadly similar aspects of institutions, and allowing us to use the larger sample for which data on settler mortality are available.

Columns (4)–(6) report our estimates for the extended Acemoglu et al. sample (which as we shall explain below will be our preferred sample in this paper). Columns (5) and (6) confirm the importance of institutions in explaining the cross-country variation in development. Once the institutional variable is added, geography and openness do not have any additional power in explaining development. Institutions trump geography and openness. In our preferred specification (column (6)), not only are institutions significant, their impact is large, and the estimated coefficients on geography and openness have the “wrong” sign! The coefficient on institutions in the IV estimation is nearly three times as large as in the corresponding OLS estimation (2 versus 0.7), suggesting that the attenuation bias from measurement error in the institution variables swamps the reverse causality bias that would tend to make the OLS estimates greater than the IV estimates.

The results are similar for the larger sample of countries (panel A, columns (6)–(9)). In this sample, we follow Hall and Jones (1999) in using the following two variables as instruments for institutional quality (in lieu of settler mortality): ENGFRAC, fraction of the population speaking English, and EURFRAC, fraction of the population speaking other European languages. Once again, institutions trump geography and openness, although the size of the estimated coefficient is smaller than that for the smaller sample. Figure 3 plots the conditional relationship between income and each of the three determinants for the 79-country (left panels) and 137-country (right panels) samples. In contrast to Figure 2, which showed a positive partial relationship between income and all its determinants, Figure 3 shows that only institutions have a significant and positive effect on income once the endogenous determinants are instrumented.<sup>7</sup>

The first-stage regressions (reported in panel B) are also interesting. In our preferred specification, settler mortality has a significant effect on integration: the coefficient is correctly signed and significant at the 1 percent level. This result holds for the range of specifications that we estimate as part of the robustness checks reported below. The

---

<sup>7</sup> The finding that trade enters with a negative (albeit insignificant) coefficient may be considered puzzling. In further results (available from the authors), we found that this is due largely to the adverse effects of trade in primary products. When total trade is broken into manufactures and non-manufactures components, it is only the latter that enters with a negative coefficient.

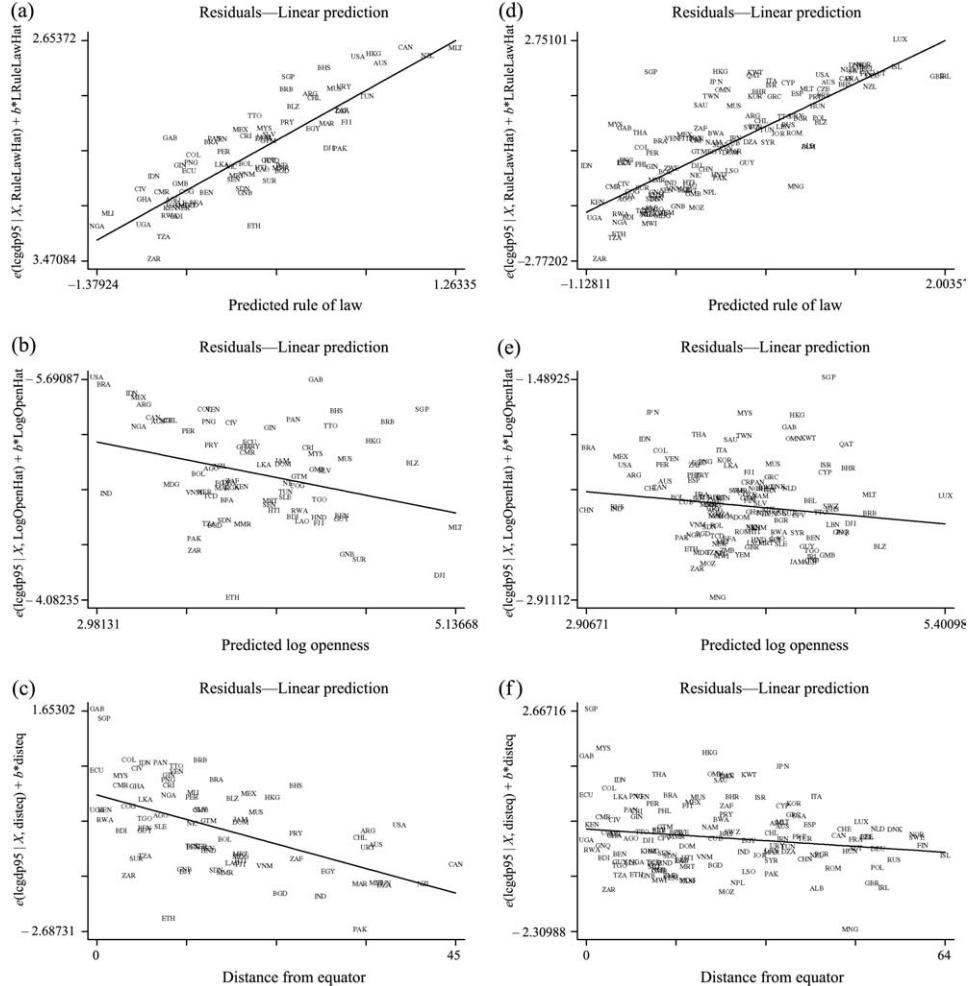


Figure 3. Conditional correlations between income and its determinants (sample of 79 countries for (a)–(c); sample of 137 countries for (d)–(f)).

geography variable has a significant impact in determining the quality of institutions as does integration, although its coefficient is significant only at the 5 percent level. The table also reports a number of diagnostic statistics on weak instruments. These provide little evidence that our results suffer from the presence of weak instruments. The  $F$ -statistic for both first-stage regressions is well above the threshold of 10 suggested by Staiger and Stock (1997);<sup>8</sup>

8 Although this threshold applies strictly to the case where there is a single endogenous regressor, it is nevertheless reassuring that our specifications yield  $F$ -statistics well above it.

the partial  $R$ -squares are reasonable; and the correlation between the fitted values of the endogenous variables appears to be sufficiently low.<sup>9</sup>

While all three samples provide qualitatively similar results, our preferred sample will be the 79-country sample: obviously this sample Pareto-dominates the 64-country sample. We also prefer this sample to the 137-country sample because settler mortality appears to be a superior instrument to those used in the 137-country sample (ENGFRAC and EURFRAC). Panel B shows that the instruments for the IV regressions in the 137-country sample fail to pass the over-identification tests despite the well-known problems of these tests having low power. Indeed, this turns out to be true not just for the core specifications in Table 3, but for many of the robustness tests that we discuss below. Thus, while it is reassuring that the main result regarding the primacy of institutions also holds in the larger sample, we will focus mainly on the 79-country sample in the rest of the paper (referring to results for the larger sample in passing).<sup>10</sup> We shall examine the robustness of our main results in the following section.

Table 4 illustrates the inter-relationships between integration and institutions in the 79-country sample. We regress trade and institutional quality separately on geography and on each other (instrumenting the endogenous variables in the manner discussed previously). While it is possible to envisage non-linear relationships among these determinants—trade may have sometimes positive and sometimes negative effects on the quality of institutions, for example,—we keep the specifications simple and linear as in all our core specifications. The IV regressions show that each of these exerts a positive impact on the other, with the larger quantitative and statistically significant impact being that of institutional quality on trade. A unit increase in institutional quality increases the trade share by 0.45 units, while a unit increase in trade increases institutional quality by 0.22 units.<sup>11</sup>

Taking these indirect effects into account, we can calculate the total impacts on incomes of these two determinants by combining the estimated parameters. Our estimates of  $\alpha$  and  $\beta$  (the direct effects) in our preferred sample and specification are 1.98 and  $-0.31$ , respectively (column (6)). We can solve the system of equations implied by the additional results in columns (1) and (2) of Table 4 to calculate the total effects on log incomes of “shocks” to the error terms in the institution and trade equations.<sup>12</sup>

The results are as follows. If we consider the point estimates in column (6) of Table 3 and in columns (1) and (2) in Table 4 as our best estimate of the various effects, a unit (positive) shock to the institutional quality equation ultimately produces an increase in log incomes of 1.85; a unit (positive) shock to the trade equation ultimately produces an

9 In Appendix A, we explain that our core specification also passes a more formal test (suggested by Stock and Yogo, 2002) for weak instruments in the presence of two endogenous regressors.

10 We emphasize that we have not found an instance in which the use of one sample or another makes a qualitative difference to our results.

11 Breaking trade into manufactures and non-manufactures components as before, we find that it is only manufactures trade that has significant positive effect on institutional quality (results are available from the authors).

12 Note that these calculations omit the feedback effect from income to trade and institutions, since we are unable to estimate these. Our numbers can hence be viewed as impact effects, taking both direct and indirect channels into account, but ignoring the feedback from income.

Table 3. Determinants of development: Core specifications, instrumental variables estimates.

	Acemoglu et al. Sample			Extended Acemoglu et al. Sample			Large Sample		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Panel A. Second-stage: Dependent variable = Log GDP per capita</i>									
Geography (DISTEQ)	0.74 (4.48)*	-0.42 (-1.19)	-0.56 (-1.23)	0.80 (5.22)*	-0.45 (-1.26)	-0.72 (-1.38)	0.76 (10.62)*	-0.06 (-0.5)	-0.14 (-0.93)
Institutions (RULE)	1.68 (4.29)*	1.78 (3.78)*		1.75 (4.42)*	1.98 (3.56)*	1.98 (8.02)*	1.19 (7.09)*	1.30	
Integration (LCOPEN)		-0.18 (-0.71)			-0.31 (-1.10)		-0.15 (-1.09)		
No. of observations	64	64	64	79	79	79	137	137	137
R-square	0.25	0.54	0.56	0.26	0.51	0.52	0.417	0.51	0.56
Test for over-identifying restrictions (p-value)							(0.0089)	(0.0354)	
<i>Panel B: First Stage for Endogenous Variables (Institutions (RULE) and Integration (LCOPEN))</i>									
Dependent variable	RULE	RULE	LCOPEN	RULE	RULE	LCOPEN	RULE	RULE	LCOPEN
Geography (DISTEQ)	0.41 (2.8)*	0.47 (3.21)*	-0.25 (-2.00)**	0.47 (3.34)*	0.54 (3.87)*	-0.18 (-1.37)	0.67 (10.81)*	0.66 (11.23)*	-0.05 (-0.84)
Settler mortality	-0.39 (-3.87)*	-0.40 (-4.1)*	-0.30 (-3.51)*	-0.34 (-3.69)*	-0.34 (-3.82)*	-0.27 (-3.22)*			
Population speaking English (ENGFRAC)							0.19 (2.69)*	0.18 (2.69)*	0.17 (2.65)*
Population speaking other European languages (EURFRAC)							0.14 (1.94)**	0.17 (2.55)**	-0.11 (-1.67)**
Constructed openness (LOGFRANKROM)	na (1.95)**	0.20 (10.32)*	0.90 (21.6)**	na (9.67)*	0.19 (2.16)**	0.80 (9.67)*	na (3.99)*	0.23 (12.33)*	0.70 (41.39)
F-statistic	22.9	17.2	41.7	24	18.5	36.9	50.09	45.79	41.39
R-square	0.41	0.44	0.66	0.37	0.40	0.58	0.52	0.57	0.54
Partial R-square	0.16	0.16	0.58	0.12	0.12	0.51	0.18	0.18	0.52
corr(RULEFIT, LCOPENFIT)			0.14			0.21			0.27

Notes: The dependent variable in panel A is per capita GDP in 1995, PPP basis. There are three samples for which the core regressions are run: (i) the first three columns correspond to the sample of 64 countries in Acemoglu et al. (2001); (ii) columns (4)–(6) use a sample of 79 countries for which data on settler mortality (LOGEM4) have been compiled by Acemoglu et al.; and (iii) columns (7)–(9) use a larger sample of 137 countries for which the instrument for institutions is similar to that in Hall and Jones (1999). The regressors in panel A are: (i) DISTEQ, the variable for geography, which is measured as the absolute value of latitude of a country; (ii) Rule of law (RULE), which is the measure for institutions; and (iii) LCOPEN, the variable for integration, which is measured as the ratio of nominal trade to nominal GDP. All regressors are scaled in the sense that they represent deviations from the mean divided by the standard deviation. The dependent variables in panel B are measures of institutions (RULE) and/or integration (LCOPEN) depending on the specification. The regressors in panel B are: (i) DISTEQ described above; (ii) settler mortality (LOGEM4) in the first six columns; (iii) the proportion of the population of a country that speaks English (ENGFRAC) and the proportion of the population that speaks any European language (EURFRAC) in the last three columns; (iv) instrument for openness (LOGFRANKROM) obtained from Frankel and Romer (1999). All regressors, except DISTEQ and RULE, in the three panels are in logs. See the Appendix for more detailed variable definitions and sources. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. *t*-statistics are reported under coefficient estimates. Significance at the 1, 5, and 10 percent levels are denoted respectively by \*, \*\*, and \*\*\*.

Table 4. Inter-relations between integration and institutions.

Dependent Variable	Extended Acemoglu et al. Sample		Large Sample	
	RULE (1)	LCOPEN (2)	RULE (3)	LCOPEN (4)
<i>Panel A. OLS Estimates</i>				
Geography (DISTEQ)	0.79 (5.79)*	-0.43 (-3.48)*	0.66 (10.81)*	-0.2 (-2.8)*
Institutions (RULE)		0.34 (4.17)*		0.27 (3.8)*
Integration (LCOPEN)	0.32 (3.35)*		0.27 (3.55)*	
<i>R</i> -square	0.42	0.23	0.52	0.12
<i>Panel B. Second Stage IV Estimates</i>				
Geography (DISTEQ)	0.75 (5.19)*	-0.51 (-3.08)*	0.66 (10.88)*	-0.02 (-0.16)
Institutions (RULE)		0.45 (2.33)**		0.02 (0.09)
Integration (LCOPEN)	0.22 (1.02)		0.27 (1.34)	
No. of observations	79	79	137	137
<i>R</i> -square	0.31	0.09	0.49	0.02
<i>Panel C. First Stage IV Estimates</i>				
Dependent variable	LCOPEN	RULE	LCOPEN	RULE
Geography (DISTEQ)	0.02 (0.13)	0.47 (3.34)*	-0.04 (-0.7)	0.67 (10.81) *
Settler mortality (LOGEM4)		-0.34 (-3.69)*		
Constructed openness (LOGFRANKROM)	0.79 (9.08)*		0.71 (12.35)*	
Population speaking English (ENGFRAC)				0.19 (2.69)*
Population speaking other European languages (EURFRAC)				0.14 (1.94)***
<i>F</i> -statistic	44.7	24.0	76.3	50.1
<i>R</i> -square	0.54	0.39	0.53	0.53

Notes: The dependent variable in panels A and B are per capita GDP in 1995, PPP basis. There are two samples for which the core regressions are run: (i) the first two columns correspond to the sample of 79 countries for which data on settler mortality (LOGEM4) have been compiled by Acemoglu et al.; and (ii) columns (3) and (4) use a larger sample of 137 countries for which the instrument for institutions is similar to that in Hall and Jones (1999). The regressors in panels A and B are: (i) DISTEQ, the variable for geography, which is measured as the absolute value of latitude of a country; (ii) Rule of law (RULE), which is the measure for institutions; and (iii) LCOPEN, the variable for integration, which is measured as the ratio of nominal trade to nominal GDP. All regressors are scaled in the sense that they represent deviations from the mean divided by the standard deviation. The dependent variables in panel C are measures of institutions (RULE) and/or integration (LCOPEN) depending on the specification. The regressors in panel C are: (i) DISTEQ described above; (ii) settler mortality (LOGEM4) in the first two columns; (iii) the proportion of the population of a country that speaks English (ENGFRAC) and the proportion of the population that speaks any European language (EURFRAC) in the last two columns; (iv) instrument for openness (LOGFRANKROM) obtained from Frankel and Romer (1999). All regressors, except DISTEQ and RULE, in the three panels are in logs. See the Appendix for more detailed variable definitions and sources. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. *t*-statistics are reported under coefficient estimates. Significance at the 1, 5, and 10 percent levels are denoted respectively by \*, \*\*, and \*\*\*.

increase in log incomes of 0.09. This is a 22-fold difference. Alternatively, we could consider only those impacts that are statistically significant. Under this assumption, a unit shock to the institutional quality equation is the estimate from column (6), namely 1.98. The corresponding unit shock to the trade equation has no impact on income at all. Institutions overwhelmingly trump integration.

The much greater impact of institutions is the consequence of four factors: (i) the estimated direct effect of institutions on incomes is positive and large; (ii) the estimated direct effect of trade on incomes is negative (but statistically insignificant); (iii) the estimated indirect effect of trade on institutions is positive, but small and statistically insignificant; and (iv) the estimated indirect effect of institutions on trade is large and statistically significant but this has either a negative or no impact on incomes because of (ii).

Repeating this exercise, and taking into account only the statistically significant coefficients, we find that the total impact of a unit shock to geography on income is about 1.49, only a quarter less than that of institutions. The large impact of geography stems from the sizable indirect impact that it has in determining institutional quality (coefficient of 0.75 in Panel B, column (1) of Table 4).<sup>13</sup>

We next analyze the channels through which the deep determinants influence incomes. The proximate determinants of economic growth are accumulation (physical and human) and productivity change. How do the deep determinants influence these channels? To answer this question, we regressed income per worker and its three proximate determinants, physical, human capital per worker, and total factor productivity (strictly speaking a labor-augmenting technological progress parameter) on the deep determinants. Data for the left hand side variables for these regressions that is, income, physical, and human capital per worker, and factor productivity are taken from Hall and Jones (1999). These results are reported in Table 5 for both the 79-country sample (columns (1)–(4)) and the 137-country sample (columns (5)–(8)).<sup>14</sup> Three features stand out.

First, the regression for income per worker is very similar to the regressions for per capita income reported in Tables 2–4, with institutions exerting a positive and significant effect on income, while integration and geography remain insignificant. Second, and interestingly, the same pattern holds broadly for the accumulation and productivity regressions; that is, institutions are an important determinant of both accumulation and productivity, while integration and geography are not influential in determining either accumulation or productivity.<sup>15</sup> Finally, it is interesting to note that institutions have a quantitatively larger impact on physical accumulation than on human capital accumulation or productivity; for example, in the 79-country sample the coefficient on physical capital accumulation is about six times greater than on human capital accumulation and about 3.2 times greater than on productivity. One possible interpretation

13 In light of our quantitative estimates, our main difference with Sachs (2003) seems to relate to whether the sizable effects of geography are direct (the Sachs position) or indirect, operating via institutions (our position).

14 Actual sample sizes are smaller than for our core specifications because of the unavailability of data for some countries in the Hall and Jones (1999) data set.

15 In the larger sample, integration has a negative and significant effect on income and accumulation but this result is not robust to the inclusion of additional variables such as land and area.

Table 5. Determinants of Development: Channels of Influence.

Dependent Variable	Extended Acemoglu et al. Sample				Larger Sample			
	Income per worker (1)	Capital per worker (2)	Human Capital per worker (3)	Total Factor productivity (4)	Income per worker (5)	Capital per worker (6)	Human Capital per worker (7)	Total Factor productivity (8)
Geography	-0.97 (-1.52)	-1.72 (-1.63)	-0.26 (-1.54)	-0.33 (-1.02)	-0.25 (-1.18)	-0.38 (-1.14)	-0.05 (-1.00)	-0.13 (-0.85)
Institutions	2.21 (3.30)*	3.39 (3.03)*	0.56 (3.14)*	1.06 (3.08)*	1.32 (5.30)*	1.90 (4.72)*	0.34 (5.64)*	0.69 (3.74)*
Integration	-0.42 (-1.36)	-0.70 (-1.30)	-0.15 (-1.86)***	-0.13 (-0.84)	-0.30 (-1.79)**	-0.46 (-2.10)**	-0.11 (-3.00)*	-0.11 (-0.84)
R-square	0.60	0.52	0.51	0.44	0.58	0.54	0.58	0.36
No. of observations	73	73	73	73	119	119	119	119

Notes: The four dependent variables—income per worker, capital per worker, human capital per worker, and the level of total factor productivity—are expressed in natural logarithms and are from Hall and Jones (1999). The IV estimates for the Acemoglu et al. sample use settler mortality (LOGEM4) as the instrument for institutions and EURFRAC and ENGFRAC as the instrument for the larger sample. All regressors, except RULE, are in logarithms and are scaled. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. *t*-statistics are reported under coefficient estimates. Significance at the 1, 5, and 10 percent levels are denoted respectively by \*, \*\*, and \*\*\*.

is that these results emphasize the particularly important role that institutions play in preventing expropriability of property which serves as a powerful incentive to invest and accumulate physical capital.

### 2.3. Robustness Checks

Tables 6–8 present our robustness checks. In Table 6 we test whether our results are driven by certain influential observations or by the four neo-European countries in our sample (Australia, Canada, New Zealand, and Australia), which are arguably different from the rest of the countries included. We also check to see whether the inclusion of regional dummies affects the results.

In columns (1)\* and (1)\*\* of Table 6 we use the Belsey-Kuh-Welsch (1980) test to check whether individual observations exert unusual leverage on the coefficient estimates, discarding those which do so. In the specification without regional dummies ((1)\*), two observations—Ethiopia and Singapore—are influential. Once these are dropped, the coefficient estimate for institutions not only remains statistically unaffected, but increases in magnitude. In the equation with regional dummies, the test requires the observation for Ethiopia to be omitted, and the revised specification (column (1)\*\*) yields results very similar to the baseline specification, with the coefficient estimate on institutions remaining strong and significant. The inclusion of regional dummies for Latin America, sub-Saharan Africa, and Asia tends to lower somewhat the estimated coefficient on institutions, but its significance level remains unaffected. Note also that none of the regional dummies enters

Table 6. Determinants of development: Robustness to “influential” observations, neoeuropes, legal systems, origin of colonizer, and religion.

	Baseline 1	(1)*	(1)**	(1)***	(1)***	Baseline 2	(2)*	(2)***	(2)***	(3)	(4)	(5)
Two-stage Least Squares: Dependent Variable is log GDP per Capita in 1995												
Geography (DISTEQ)	-0.72 (-1.38)	-1.37 (-1.71)***	-0.71 (-1.42)	-0.92 (-1.18)	-0.62 (-0.82)	-0.14 (-0.93)	0.02 (0.20)	-0.34 (-1.48)	-1.00 (-1.53)	-0.78 (-1.05)	-0.82 (-1.31)	
Institutions (RULE)	1.98	2.66	1.86	2.77	1.99	1.30	1.30	0.90	1.64	2.43	2.32	2.09
Integration (LCOPEN)	(3.56)* (-0.31)	(3.06)* -0.45	(3.26)* -0.33	(2.45)*** -0.74	(1.64) -0.44	(7.09)* -0.15	(7.14)* -0.15	(8.54)* 0.02	(5.15)* -0.31	(3.08)* -0.42	(2.46)*** -0.28	(3.04)* (-0.32)
Regional Dummies												
Latin America (LAA.M)												
Sub-Saharan Africa (SAFRICA)					0.42 (1.18)	0.15 (0.28)	0.15 (0.28)	0.25 (1.65)***				
East Asia (ASIAE)					-0.17 (-0.43)	-0.41 (-1.05)	-0.41 (-1.05)	-0.62 (-3.70)*				
Legal origin												
Identity of colonizer												
Religion												
R-square	0.52	0.56	0.65	0.44	0.63	0.56	0.59	0.68	0.55	0.54	0.55	0.59
No. of observations	79	77	78	75	75	137	136	134	133	79	79	79
Omitted observations	None	Singapore	Ethiopia	Australia	Australia	None	Singapore	Cuba	Australia	None	None	None
		Ethiopia		Canada	Canada			Czech Rep.	Canada			
				New Zealand	New Zealand			Germany	New Zealand			
				USA	USA			USA	USA			
										[0.135]	[0.080]***	[0.021]***

Notes: The dependent variable is per capita GDP in 1995, PPP basis. Baseline 1 corresponds to the specification in column (6) of Table 3. Baseline 2 corresponds to the specification in column (9) of Table 3. In columns labeled with one and two asterisks, influential observations are defined according to the Belsey et al. (1980) DFITS statistic, which requires omitting those observations for which DFITS exceeds  $2(k/n)^{(1/2)}$ , where  $k$  is the number of regressors and  $n$  is the sample size. In columns labeled with three or four asterisks, observations for Australia, Canada, New Zealand, and Canada (Neueuropes) are omitted. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated.  $t$ -statistics are reported under coefficient estimates. For legal origin, identity of colonizer, and religion,  $p$ -values for joint significance of the underlying variables (LEGFR and LEGSO for legal origin, COLUK and COLFR for colonizer's identity, and CATH, PROT, and MUSL for religion) are reported. Significance at the 1, 5, and 10 percent levels are denoted respectively by \*, \*\*, and \*\*\*. All regressors are scaled as described in the notes to Tables 2-4.

significantly, which is reassuring regarding the soundness of our parsimonious specification.

The tests for influential observations suggest that there is no statistical basis for discarding neo-European countries. Nevertheless to confirm that these countries are not driving the results, we re-estimated the baseline specification without these observations. As the column labeled (1)\*\*\* confirms, the coefficient estimates are unaffected; indeed, once again the size of the coefficient on institutions rises substantially, suggesting the greater importance of institutions for the non-neo-European colonized countries. The remaining columns (columns (2)\* and (2)\*\*) confirm that our results are robust also for the larger sample of countries.

We then check whether our results are robust to the inclusion of dummies for legal origin (column (3)), for the identity of colonizer (column (4)), and religion (column (5)). La Porta et al. (1998) argue that the type of legal system historically adopted in a country or imported through colonization has an important bearing on the development of institutions and hence on income levels. Similar claims are made on behalf of the other variables. In all cases, while these variables themselves tend to be individually and in some cases jointly significant, their inclusion does not affect the core results about the importance of institutions and the lack of any direct impact of geography and integration on incomes. Indeed, controlling for these other variables, the coefficient of the institutions variable increases: for example, in the 79-country sample, this coefficient increases from two in the baseline to 2.43 when the legal origin dummies are included.<sup>16</sup>

In Table 7 we check whether our particular choice of measure for geography (distance from the equator) influences our results. We successively substitute in our baseline specification a number of alternative measures of geography used in the literature. These include percent of a country's land area in the tropics (TROPICS), access to the sea (ACCESS), number of frost days per month in winter (FROSTDAYS), the area covered by frost (FROSTAREA), whether a country is an oil exporter (OIL), and mean temperature (MEAN TEMPERATURE). (Recall that we had already introduced regional dummies as part of our basic robustness check above.) The variables FROSTDAYS and FROSTAREA are taken from Masters and McMillan (2001), who argue that the key disadvantage faced by tropical countries is the absence of winter frost. (Frost kills pests, pathogens and parasites, thereby improving human health and agricultural productivity.) We find that none of these variables, with the exception of the oil dummy, is statistically significant in determining incomes. Equally importantly, they do not change qualitatively our estimates of the institution variable, which remains significant, nor of the integration variable, which remains insignificant and “wrongly” signed.<sup>17</sup>

---

16 We do not report the results for the larger sample but they are very similar. For the 79-country sample, interesting results are obtained for some of the individual legal origin and other variables. For example, as in Acemoglu et al. (2001), the French legal origin dummy has a positive total effect on incomes; the total impact of having been colonized by the United Kingdom is negative and statistically significant even though former UK-colonies have better quality of institutions on average. As for religion, suffice it to say that Weber is not vindicated!

17 In most of these regressions (columns (1)–(7)), the geography variable is a significant determinant of institutions in the first stage regressions.

Table 7. Determinants of development: Robustness to alternative measures of geography.

	Baseline	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Two-stage Least Squares: Dependent Variable is log GDP per Capita in 1995												
Institutions (RULE)	1.98 (3.56)*	1.47 (2.96)*	1.99 (3.57)*	1.48 (6.09)*	1.45 (7.03)*	1.97 (3.39)*	1.89 (3.03)*	1.81 (2.23)*	0.65 (2.34)*	0.94 (3.90)*	0.91 (1.67)***	1.04 (2.47)*
Integration (LCOPEN)	-0.31 (-1.10)	-0.19 (-0.91)	-0.35 (-1.00)	-0.10 (-0.52)	0.00 (0.01)	-0.42 (-1.10)	0.00 (0.01)	-0.25 (-0.65)	-0.02 (-0.01)	-0.11 (-0.71)	-0.08 (-0.46)	-0.14 (-1.00)
Geography (DISTEQ)	-0.72 (-1.38)	-0.41 (-0.97)						-0.25 (-0.91)	-0.25 (-1.35)	-0.34 (-1.07)	-0.26 (-1.11)	-0.28 (-1.11)
Regional dummies												
Latin America (LAAM)	0.43 (1.57)										0.25 (0.61)	0.29 (0.74)
Sub-Saharan Africa (SAFRICA)	-0.31 (-0.97)										-0.20 (-0.60)	-0.06 (0.10)
East Asia (ASIAE)	0.29 (0.83)										0.26 (0.97)	0.34 (1.22)
Area under tropics (TROPICS)												
Access to sea (ACCESS)												
Major oil exporter (OIL)												
Days under frost (FROSTDAYS)								-1.07 (-1.47)	-1.07 (-1.47)			
Area under frost (FROSTAREA)								-0.62 (-1.17)	-0.62 (-1.17)			
Temperature (MEANTEMP)									0.52 (1.27)	0.52 (1.27)		
Malaria proportion in population (MALFAL94)												
Malaria proportion in population * proportion of fatal species (MALFAL)												
R-square	0.52 79	0.63 79	0.54 76	0.52 67	0.54 76	0.53 67	0.51 66	0.64 69	0.52 75	0.52 75	-0.47 (-5.26)*	-0.36 (-0.67)
No. of observations												

Notes: The dependent variable is per capita GDP in 1995, PPP basis. Baseline corresponds to the specification in column (6) of Tables 2-4. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. *t*-statistics are reported under coefficient estimates. Significance at the 1, 5, and 10 percent levels are denoted respectively by \*, \*\*, and \*\*\*.

In response to the findings reported in an earlier version of this paper, Sachs (2003) has produced new empirical estimates which attribute a more significant causal role to geography. Arguing that distance from equator is a poor proxy for geographical advantage, Sachs uses two explanatory variables related to malaria incidence. The first of these is an estimate of the proportion of a country's population that lives with risk of malaria transmission (MAL94P), while the second multiplies MAL94P by an estimate of the proportion of malaria cases that involve the fatal species, *Plasmodium falciparum* (MALFAL). Since malaria incidence is an endogenous variable, Sachs instruments for both of these using an index of "malaria ecology" (ME) taken from Kiszewski et al. (2003). In columns (8) and (9) we add these variables to our core specification, instrumenting them in the same manner as in Sachs (2003). Our results are similar to Sachs', namely that malaria appears to have a strong, statistically significant, and negative effect on income. Note that the statistical significance of institutional quality is unaffected by the addition of the malaria variables, something that Sachs (2003) notes as well.

We are inclined to attach somewhat less importance to these results than Sachs does. First, it is difficult to believe that malaria, which is a debilitating rather than fatal disease, can by itself exert such a strong effect on income levels. If it is a proxy for something else, it would be good to know what that something else is and measure it more directly. Second, we are a bit concerned about the exogeneity of the instrument (ME). Sachs (2003, 7) asserts that ME is exogenous because it is "built upon climatological and vector conditions on a country-by-country basis," but he does not go into much further detail. The original source for the index (Kiszewski et al. 2003), written for a public health audience, has no discussion of exogeneity at all. Third, the malaria variables are very highly correlated with location in sub-Saharan Africa.<sup>18</sup> The practical import of this is that it is difficult to tell the effect of malaria variables apart from those of regional dummies. This is shown in columns (10) and (11) where we add regional dummies to the specifications reported earlier. Both malaria variables now drop very far below statistical significance, while institutional quality remains significant (albeit at the 90 percent level in one case).

Finally, we experimented with a series of specifications (not reported) that involved interacting the different geography variables with each other as well as introducing different functional forms (e.g., exponential) for them. These did not provide evidence in favor of additional significant direct effects of geography on income. Overall, we conclude that there seems to be some, albeit modest, support for the direct impact of geography on income, although one that is not as robust as that of institutional quality.<sup>19</sup> The first stage regressions, however, point clearly in favor of an important indirect role of geography via institutions.

In Table 8, we undertake a series of robustness checks on the side of trade. First, we check whether our results are sensitive to our omission of market size variables. Frankel and Romer (1999) argue that smaller countries tend to trade more, and that one should

<sup>18</sup> Regressing MALFAL and MAL94P, respectively, on a sub-Saharan African dummy yields *t*-statistics of above 10 and *R*<sup>2</sup>'s above 0.40.

<sup>19</sup> See also Acemoglu et al. (2002) who document that the regions of the world that were relatively rich around 1,500 underwent a "reversal of fortune" subsequently. They argue that this militates against a geographic determinist view of income difference.

Table 8. Determinants of development: Robustness to alternative measures and instruments for integration.

	Baseline	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Two-Stage Least Squares: Dependent Variable is log GDP per Capita in 1995								
Geography (DISTEQ))	-0.72 (-1.38)	-0.56 (-0.83)	-0.63 (-0.88)	0.13 (0.38)	0.12 (0.35)	-1.16 (-1.25)	-1.24 (-1.14)	-0.86 (-1.17)
Institutions (RULE)	1.98 (3.56)*	1.83 (2.64)**	1.90 (2.58)**	0.97 (2.39)**	0.99 (2.46)**	2.70 (2.14)**	2.84 (1.82)**	2.55 (2.11)**
Integration (LCOPEN)	-0.31 (-1.38)	0.12 (0.10)	-0.01 (-0.01)	-0.87 (-0.90)	-0.85 (-0.94)			
Land area (AREA)		0.27 (0.77)	0.24 (0.67)	-0.40 (-0.97)	-0.39 (-0.99)			
Population (POP)		0.11 (0.16)	0.39 (0.05)	-0.43 (-0.63)	-0.42 (-0.65)			
“Real openness” (LNOOPEN)						-0.77 (-0.83)	-0.94 (-0.70)	
“Policy openness” (SW)								-2.04 (-1.07)
R-square	0.52	0.61	0.61	0.60	0.60	0.55	0.55	0.61
No. of observations	79	79	79	136	136	71	71	69

Notes: The dependent variable is per capita GDP in 1995, PPP basis. All regressors, except DISTEQ, RULE, and SW, are expressed in logs. Baseline corresponds to the specification in column (6) of Table 3. In columns (1), (3) and (5) the instrument for openness (LOGFRANKROM) is from Frankel and Romer (1999). In columns (2), (4) and (6), the instrument for openness (LOGFRANKROMR) is derived by re-estimating the gravity equation in Frankel and Romer (1999) with the left-hand side variable defined as nominal bilateral trade to nominal GDP. In Frankel and Romer, the left hand side variable was defined as nominal trade divided by PPP GDP. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. *t*-statistics are reported under coefficient estimates. Significance at the 1, 5, and 10 percent levels are denoted respectively by \*, \*\*, and \*\*\*. All regressors are scaled as described in the notes to Tables 2–4.

therefore control for country size when looking for the effect of trade on incomes. The column labeled (1) in Table 8 includes two measures of country size—area and population. These variables do not have additional explanatory power in the income equation, which is different from the results in Frankel and Romer (1999). The size and significance of the coefficient on institutions are unaffected. The coefficient on openness becomes positive, but is highly insignificant. Column (3) replicates this exercise for the larger sample. The coefficient on institutions does not change qualitatively (but the standard error is sharply reduced as is the coefficient estimates), while the coefficient on openness is still negatively signed.

Alcalá and Cicconé (2004) have recently advocated the use of what they call “real openness”, which is measured as the ratio of trade to PPP GDP. They argue that this is a better measure of integration than the simple ratio of trade to nominal GDP (that Frankel and Romer and we favor) in the presence of trade-driven productivity change. We have a number of conceptual and empirical problems with the Alcalá and Cicconé approach, which we discuss in Appendix A at the end of the paper and to which the interested reader can turn. Here we simply point out that our results are robust to the use of this alternative measure of openness. Column (5) presents the findings when we substitute the Alcalá and Cicconé measure of openness for ours. This integration measure is also “wrongly” signed

and insignificant, while the coefficient on institutions increases in size and remains significant, albeit at the 5 percent level.<sup>20</sup>

Columns (2), (4), and (6) replicate the three robustness checks described above but with an instrument for openness that is slightly different from that in Frankel and Romer (1999). To obtain their instruments, Frankel and Romer estimated a gravity equation with the dependent variable defined as trade to PPP GDP. Strictly speaking therefore, theirs was an appropriate instrument for Alcalá and Cicconé's "real openness." We re-estimated the gravity equation on the original Frankel and Romer sample of 63 countries, with trade to GDP as the dependent variable. We then used the coefficients from this gravity equation to construct the instrument for openness for all the 137 countries in our larger sample. The results in columns (2), (4), and (6) are very similar to those using the original Frankel and Romer instruments. The choice of instruments thus does not affect our main results.

Finally, in column (7) we substitute a "policy" measure for the trade variable. For reasons explained later, we believe that it is not appropriate to use policy variables in level regressions. We nevertheless sought to test the robustness of our results to one of the most-widely used measures in the trade and growth literature due to Sachs and Warner (1995), which has been endorsed recently by Krueger and Berg (2002).<sup>21</sup> The results show that the institutional variable remains significant at the 5 percent level and the Sachs–Warner measure is itself wrongly signed like the other openness measures.

### 3. What Does It All Mean?

The present paper represents in our view the most systematic attempt to date to estimate the relationship between integration, institutions, and geography, on the one hand, and income, on the other. In this section, we evaluate and interpret our results further. This also gives us an opportunity to make some additional comments on the related literature. We group the comments under three headings. First, we argue that an instrumentation strategy should not be confused with building and testing theories. Second, we relate our discussion on institutions to the discussion on "policies." Third, we discuss the operational implications of the results.

#### 3.1. An Instrument Does Not a Theory Make

Insofar as our results emphasize the supremacy of institutions, they are very close to those in Acemoglu et al. Note that we have gone beyond Acemoglu et al. by using larger sample

20 An alternative, albeit imperfect, way of taking the Alcalá and Cicconé concern into account is to use a measure of trade openness that excludes services (the main source of non-tradables) from the measure of GDP in the denominator. When we use this measure of openness (trade as a share of "goods GDP") we get similar results to those reported above (these results are available from authors).

21 The shortcomings of the Sachs–Warner index as a measure of trade policy are discussed at length in Rodriguez and Rodrik (2001).

sizes, and by including measures of integration in our estimation. We now want to clarify a point regarding the interpretation of results. In particular, we want to stress the distinction between using an instrument to identify an exogenous source of variation in the independent variable of interest and laying out a full theory of cause and effect. In our view, this distinction is not made adequately clear in Acemoglu et al. and is arguably blurred by Easterly and Levine (2003).

One reading of the Acemoglu et al. paper, and the one strongly suggested by their title—“The Colonial Origins of Comparative Development”—is that they regard experience under the early period of colonization as a fundamental determinant of current income levels. While the Acemoglu et al. paper is certainly suggestive on this score, in our view this interpretation of the paper’s central message would not be entirely correct. One problem is that Acemoglu et al. do not carry out a direct test of the impact of colonial policies and institutions. Furthermore, if colonial experience were the key determinant of income levels, how would we account for the variation in incomes among countries that had never been colonized by the Europeans?

To illustrate the second point, Figure 5 presents histograms of per capita incomes for 163 countries for which we have data on per capita GDP in 1995. The sample is split into two groups, a group of 103 countries that were colonized by one of the major Western European powers sometime before the twentieth century, and a group of 60 countries that were not colonized. The latter group includes some very high-income countries such as Finland and Luxembourg as well very poor countries such as Ethiopia,<sup>22</sup> Yemen, and Mongolia. (Afghanistan is another low-income non-colonized country, but we do not have income data for it.) As the figure reveals, the dispersion of incomes within the colonized sample is not hugely different than that in the non-colonized sample. The standard deviations of log income per capita are 1.01 and 0.89 for the colonized and non-colonized samples, respectively. The income gaps that separate Ethiopia from Turkey, or China from Luxembourg are huge, and can obviously not be explained by any of these countries’ colonial experience.

While colonial history does not quite provide a satisfactory account of income differences around the world, it can still provide a valid instrument. And that, in our view, is where the Acemoglu et al. paper is successful. An instrument is something that simply has some desirable statistical properties. It need not be a large part of the causal story.<sup>23</sup> To illustrate the distinction between a theory and an instrument, here is an analogy that draws on a well-known paper by Angrist and Krueger (1991).

Angrist and Krueger (1991) use quarter of birth as an instrument for the level of

22 Ethiopia was included in the Acemoglu et al. sample of colonies, even though this country has never been colonized. (It was occupied for a period of several years by Italy during 1936–1941, but this neither counts as colonization, nor could have had much to do with the settler mortality rates from the 19th century). Excluding Ethiopia from the Acemoglu et al. sample makes no difference to the basic Acemoglu et al. results—and in fact it improves these results, as eyeballing Acemoglu et al.’s Figures 1 and 2 would indicate.

23 In Acemoglu et al. (2002), the authors provide a fuller account of comparative development, emphasizing the interaction between the quality of institutions, on the one hand, and the prevailing opportunities to industrialize, on the other.

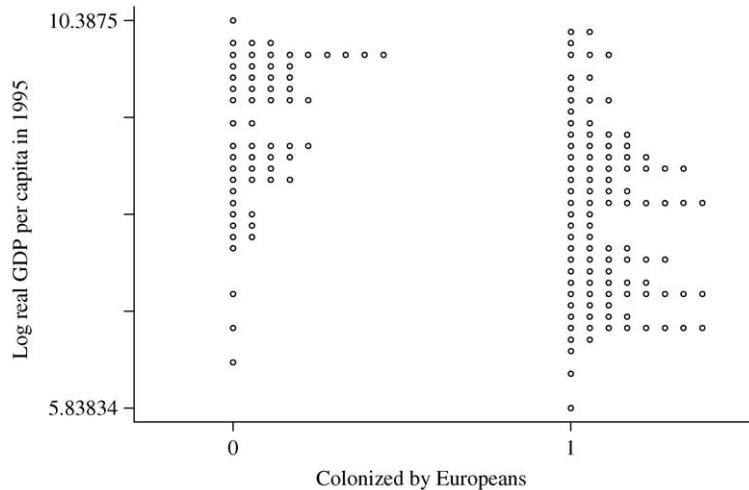


Figure 4. Distribution of incomes for colonized and non-colonized countries.

educational attainment, to disentangle the effects of schooling on personal earnings from those of unobserved attributes (such as “ability”). The story is that compulsory schooling requirements, requiring schooling until age 16 or 17, interacting with school-entry requirements, imply variation in the level of schooling that is correlated with quarter of birth but not with other personal attributes. The authors show for example that students born in the first quarter of the year have a systematically lower level of average schooling in the population. This is a plausible strategy for identification, but it obviously does not imply a quarter-of-birth theory of earnings. Similarly, the Acemoglu et al. strategy does not amount to a direct test of a colonial-origins theory of development.<sup>24</sup>

Easterly and Levine (2003) also assign a causal role to the settler mortality instrument and interpret it as a geographical determinant of institutions such as “crops and germs,” rather than viewing it as a device to capture the exogenous source of variation in institutions. Indeed, although they stress the role of institutions, they appear to come close to a geography theory of development. Our view is that we should not elevate settler mortality beyond its status as an instrument, and avoid favoring either a colonial view of development (as some readings of Acemoglu et al. would have it) or a geography-based theory of development (as some readings of Easterly and Levine would have it).

---

<sup>24</sup> Acemoglu et al. themselves are somewhat ambiguous about this. They motivate settler mortality as an instrument, but then their account gravitates towards a colonial origins theory of institutional development. And their title strongly suggests that they consider the contribution of their paper to have been a theory as opposed to an identification strategy. In personal communication, one of the authors has explained that the colonial experience allows them to exploit the exogenous source of variation in institutions and not all the variation. The fit of the first-stage regressions of about 25 percent leaves room for most of the variation to be explained by factors other than colonization.

### 3.2. *The Primacy of Institutional Quality Does Not Imply Policy Ineffectiveness*

Easterly and Levine (2003) assert that (macroeconomic) policies do not have an effect on incomes, once institutions are controlled for. Our view on the effectiveness of policy is similar to that expressed in Acemoglu et al. (2001, 1395): there are “substantial economic gains from improving institutions, for example as in the case of Japan during the Meiji Restoration or South Korea during the 1960s” or, one may add, China since the late 1970s. The distinction between institutions and policies is murky, as these examples illustrate. The reforms that Japan, South Korea, and China undertook were policy innovations that eventually resulted in a fundamental change in the institutional underpinning of their economies.

We find it helpful to think of policy as a flow variable, in contrast to institutions, which is a stock variable. We can view institutions as the cumulative outcome of past policy actions. Letting  $p_i$  denote policy on dimension  $i$  ( $i = \text{fiscal, trade, monetary, etc.}$ ),  $I$  institutional quality, and  $\delta$  the rate at which institutional quality decays absent countervailing action, the evolution of institutional quality over time can be written as  $\dot{I} = \Sigma \alpha_i p_i - \delta I$ , where  $\alpha_i$  denotes the impact of policy  $i$  on institutional quality.

This suggests that it is inappropriate to regress income levels on institutional quality and policies, as Easterly and Levine (2003) do. The problem is not just that incomes move slowly while policies can take sudden turns. In principle this could be addressed by taking long-term averages of policies. (Easterly and Levine average their policy measures over a number of decades.) It is that measures of institutional quality already contain all the relevant information about the impact of policies. If the appropriate specification for income is  $\ln y = \beta I + u$ , the effect of policies should be sought in a regression of the form  $d \ln y/dt = \beta \dot{I} + v = \alpha_0 + \beta \Sigma \alpha_i p_i + v$ . In other words, one should look for the effect of policies in a regression of growth of income on policies.

Moreover, a geography theory of institutions can underestimate the impact that policies can play in changing them over time. As an empirical matter, institutions have changed remarkably in the last three decades. For example, one indicator of institutional quality—the index measuring the constraint on the executive in the Gurr Polity IV dataset, which is available on a consistent basis for several decades—shows a marked improvement between the 1970s and 1990s. For 71 countries in our core sample, this index had a mean value of 3.21 in the 1970s, 3.52 in the 1980s, and 4.37 in the 1990s. A purely geographical theory of institutions would have difficulty in accounting for these changes. Indeed, if the first stage regressions reported in panel B of Table 3 are run over the last three decades, the coefficient on settler mortality, declines from  $-1.14$  in the 1970s to  $-1.04$  in the 1980s and  $0.86$  in the 1990s, illustrating the mutability of institutions, and the declining importance of history (on the Acemoglu et al. interpretation of settler mortality) or geography (on the Easterly and Levine interpretation of settler mortality) in explaining the cross-national variation in institutions.<sup>25</sup>

---

25 The  $R$ -squares also decline over time for these first-stage regressions.

### **3.3. *The Hard Work is Still Ahead***

How much guidance do our results provide to policymakers who want to improve the performance of their economies? Not much at all. Sure, it is helpful to know that geography is not destiny, or that focussing on increasing the economy's links with world markets is unlikely to yield convergence. But the operational guidance that our central result on the primacy of institutional quality yields is extremely meager.

Our indicators of institutional quality are investors' and other observers' ratings of the institutional environment. They quantify these observers' views as to the likelihood that investors will retain the fruits of their investments, the chances that the state will expropriate them, or that the legal system will protect their property rights. While it is important to know that these ratings matter—and matter a great deal in fact—it remains unclear how the underlying evaluations and perceptions can be altered. In terms of the formulation developed above, what we have estimated is  $\beta$ , while what policy makers need to know are the  $\alpha_i$  (policy impacts) for the policies at their disposal. In fact, since our identification strategies rely on exogenous sources of variation in these evaluations, they are doubly unhelpful from a policy perspective.

We illustrate the difficulty of extracting policy-relevant information from our findings using the example of property rights. Obviously, the presence of clear property rights for investors is a key, if not the key, element in the institutional environment that shapes economic performance. Our findings indicate that when investors believe their property rights are protected, the economy ends up richer. But nothing is implied about the actual form that property rights should take. We cannot even necessarily deduce that enacting a private property-rights regime would produce superior results compared to alternative forms of property rights.

If this seems stretching things too far, consider the experiences of China and Russia. China still retains a socialist legal system, while Russia has a regime of private property rights in place. Despite the absence of formal private property rights, Chinese entrepreneurs have felt sufficiently secure to make large investments, making that country the world's fastest growing economy over the last two decades. In Russia, by contrast, investors have felt insecure, and private investment has remained low. Our institutional quality indicators bear this out, with Russia scoring considerably lower than China despite a formal legal regime that is much more in line with European norms than China's. Credibly signaling that property rights will be protected is apparently more important than enacting them into law as a formal private property rights regime.

So our findings do not map into a determinate set of policy desiderata. Indeed, there is growing evidence that desirable institutional arrangements have a large element of context specificity, arising from differences in historical trajectories, geography, political economy, or other initial conditions. As argued in Mukand and Rodrik (2002), this could help explain why successful developing countries—China, South Korea, and Taiwan among others—have almost always combined unorthodox elements with orthodox policies. It could also account for why important institutional differences persist among the advanced countries of North America, Western Europe, and Japan—in the role of the public sector, the nature of the legal systems, corporate governance, financial markets, labor markets, and social insurance mechanisms, among others.

It is important to underscore that this does not mean economic principles work differently in different places. We need to make a distinction between economic principles and their institutional embodiment. Most first-order economic principles come institution-free. Economic ideas such as incentives, competition, hard-budget constraints, sound money, fiscal sustainability, property rights do not map directly into institutional forms. Property rights can be implemented through common law, civil law, or, for that matter, Chinese-type socialism. Competition can be maintained through a combination of free entry and laissez-faire, or through a well-functioning regulatory authority. Macroeconomic stability can be achieved under a variety of fiscal institutions. Institutional solutions that perform well in one setting may be inappropriate in other setting without the supporting norms and complementary institutions. In the words of Douglass North

“economies that adopt the formal rules of another economy will have very different performance characteristics than the first economy because of different informal norms and enforcement. The implication is that transferring the formal political and economic rules of successful Western economies to third-world and Eastern European economies is not a sufficient condition for good economic performance.” (North 1994, 366)

In addition, since policy makers always operate in second-best environments, optimal reform trajectories—even in apparently straightforward cases such as price reform—cannot be designed without regard to prevailing conditions and without weighting the consequences for multiple distorted margins.

Consequently, there is much to be learned still about what improving institutional quality means on the ground. This, we would like to suggest, is a wide open area of research. Cross-national studies of the present type are just a beginning that point us in the right direction.

#### **Appendix A: Problems with the Use of the Alcalá and Ciccone Measure of “Real Openness”**

Alcalá and Ciccone (2004) propose a different measure of integration, which they claim is superior to that used by Frankel and Romer and others. They argue that the conventional measure of openness—nominal trade divided by nominal GDP—can yield an estimate of trade on incomes that is biased downwards. The logic is as follows. Suppose that an increase in trade raises productivity, but that it does so predominantly in the tradables sector. Unless non-tradables are inferior in demand, this will raise the relative price of non-tradables. This will in turn tend to depress the ratio of trade to nominal GDP. The result is that the initial increase in the openness ratio will be attenuated. AC therefore prefer to use what they call “real openness,” nominal trade divided by PPP GDP.

#### ***A First Pass***

Alcalá and Ciccone (2004) find a relationship between “real openness” and income within their empirical framework that they claim is more robust than when the

conventional measure of openness is used. This seems to be the case even when institutional quality is entered, which shows up significantly in their regressions as well. Since we were unable to obtain their data set, we could not replicate their results exactly. However, as we have already discussed, the use of “real openness” within our empirical specification does not alter the central results of our paper, namely the importance of institutions and the insignificance of openness (see columns (5) and (6) of Table 8). Here we discuss a number of misgivings we have with the Alcalá and Ciccone approach.

As a first step, it is useful to explain the mechanics of why “real openness” (*Ropen*) does much “better” than openness (*Open*) in Alcalá and Ciccone-type regressions. *Ropen* and *Open* are linked by the identity  $\log Ropen = \log Open + \log P$ , where  $P$  is a country’s price level. We know from the Balassa–Samuelson argument that  $P$  has a close relationship to a country’s income/productivity level. This can be seen explicitly in the chart below, which shows the relationship between  $(\log Ropen - \log Open)$  and GDP per worker (The simple correlation coefficient between the two is 0.76).

Hence what Alcalá and Ciccone are doing is to augment the standard measure of openness (*Open*) with a component that is highly correlated with income/productivity. This procedure is virtually guaranteed to generate a high correlation between productivity and openness. (The correlation between  $\log Open$  and  $\log P$  is extremely weak, 0.05, while the correlation between  $\log Ropen$  and  $\log P$  is 0.66.) Even after instrumentation, what ends up doing the heavy lifting in their paper appears to be the strong correlation between price levels and labor productivity.<sup>26</sup>

### ***Theoretical Issues***

The Alcalá and Ciccone argument strikes us as being misleading on purely conceptual grounds. The use of “real openness” can yield in fact an opposite, and potentially more severe, bias. What Alcalá and Ciccone do not recognize is that the actual null hypothesis that is tested is that trade does not cause productivity. Under that null, Alcalá and Ciccone’s real openness measure generates a positive correlation between income and openness that is entirely spurious. In effect, the Alcalá and Ciccone adjustment has the consequence that any and all increases in the productivity of tradables, regardless of source, can produce a rise in their measure of openness. Any increase in tradables productivity, whether driven by trade or not, will raise non-tradables prices at home and

---

26 Indeed the Alcalá and Ciccone argument that the true relationship between trade and productivity can be ascertained only by holding the price level constant suggests estimating a more general framework of the kind:  $\log y = \alpha + \beta_1 \log Open + \beta_2 \log P + v$ . When we do so, using an instrument set close to that in Alcalá and Ciccone, we find that the coefficient on openness is negative and insignificant, and that on the price level positive and highly significant. The comparable equation estimated with real openness yields a coefficient that is positive and significant. Whatever effect *Ropen* has on productivity, it seems to be operating via  $P$ , not via *Open*. So this more general framework yields little evidence that there is a significant causal effect from openness to productivity, holding the price level constant. Indeed, if we are to interpret these results literally, they suggest that causality runs from the price level to productivity.

the price level of an economy relative to others. “Adjusting” for this by using PPP GDP as the denominator drives up measured openness. The conventional measure of openness does not suffer from that shortcoming, and hence is preferable. We explain and illustrate this point using a simple model.

Imagine a symmetric world populated with a large number of small endowment economies. Each economy  $i$  has a fixed endowment of non-traded and traded goods, denoted by  $N_i$  and  $T_i$ , respectively. Let each country produce a different traded good (the Armington assumption), but consume all the varieties of traded goods produced around the world. If there is a very large number of countries, each country’s consumption of its own endowment of the traded good will be negligible: (almost) all of its traded good will be exported in exchange for imports of the traded goods produced elsewhere. Let  $PN_i$  stand for the price of non-traded goods in country  $i$  and let the prices of all traded goods be fixed at unity. Since the sum of exports and imports are given by  $2T_i$ , conventionally measured openness in a country  $i$  can then be expressed as  $ON_i = 2T_i/(PN_i^*N_i + T_i)$ .

Assume that preferences in each country take the Cobb-Douglas form, such that non-traded goods and traded goods (in aggregate) have fixed budget shares. Under this assumption,  $2T_i/(PN_i^*N_i + T_i)$  will be constant and independent of a country’s endowments of  $T$  and  $N$ . (This is because  $dPN_i/PN_i = dT_i/T_i - dN_i/N_i$ .) Cross-country differences in conventionally measured openness,  $ON_i$ , will arise solely from differences in Cobb-Douglas budget shares. Assume also that differences in the endowment of the traded good are the only source of cross-country differences in income. That is, all countries have identical  $N_i$  but varying  $T_i$ . Countries with larger  $T_i$  are richer.

Under the above assumptions, there is no causal relationship that goes from trade to incomes. Cross-country differences in income are due entirely to differences in endowments. And if we run a regression of income on openness, we will get nothing. Trade shares either do not vary across countries, or they vary “randomly” with the Cobb-Douglas parameter. They have no systematic relationship to levels of income. So the econometrics will provide a good guide to the underlying reality.

Now suppose that we follow Alcalá and Cicconé, and construct their real openness measure,  $OR_i$ . This adjustment consists of expressing the value of  $i$ ’s non-traded production at some benchmark country’s prices,  $PB$ , instead of domestic prices,  $PN_i$ . The Alcalá and Cicconé measure of real openness is therefore  $OR_i = 2T_i/(PB^*N_i + T_i)$ . Note that  $OR_i$  is increasing in  $T_i$ . When we correlate  $OR_i$  with incomes across countries, we will get a positive relationship. This is a spurious relationship, since the only source of productivity differences in this model is differences in endowments.

Hence, in this benchmark model, the conventional measure of openness does exactly what we would like a measure of openness to do under the null hypothesis that trade does not cause productivity. The Alcalá and Cicconé variant, meanwhile, imparts a positive bias to the estimated trade-income relationship. A key feature of the model above is that the elasticity of substitution in demand between  $T$  and  $N$  is unity. This ensures that the rise in  $PN$  is just enough to keep (conventional) openness invariant to changes in the endowment (or productivity) of tradables. When the elasticity of substitution differs from one, conventional openness does not always deliver such a helpful result, but the bias is not unidirectional. So with an elasticity of substitution greater than one, a regression of income

on conventional openness will yield (misleadingly) a positive coefficient, while with an elasticity less than one, the regression will yield (misleadingly) a negative coefficient. However, the AC real openness measure is invariant to the elasticity of substitution and hence is always positively biased.

### *Empirics*

Another point relates to the choice between real openness and openness on econometric grounds. Recall that the authors' original argument on behalf of Ropen is based on the idea that there is reverse causality from productivity to Open, via the price level. If the Frankel–Romer constructed trade share is a valid instrument, in the sense of being uncorrelated with productivity through any channel other than trade, any type of reverse causality—positive or negative—is already taken care of. The reverse causality that Alcalá and Cicconé worry about should be handled by the instrument as well! For the authors' argument to be valid, instrumentation should fail when Open is used, but work when Ropen is used (even though the same instruments are used in both cases). The authors do not provide any justification for this, and it is unclear to us that any justification could be produced.

Moreover, it is possible that the Alcalá and Cicconé strategy does exactly the reverse and that it weakens the instrument. As we mentioned above, we were unable to obtain Alcalá and Cicconé's data and could not replicate their results exactly. But in our attempted replications of their baseline specification, we repeatedly found that the first-stage  $F$ -statistics were lower, sometimes substantially so, when real openness was used in lieu of openness. In fact, the  $F$ -statistic was typically below 10 when real openness was used (and always above 10 when openness was used).<sup>27</sup> A more formal test for weak instruments with two endogenous regressors (which in our case are institutions and openness) is the  $G$ -statistic proposed by Stock and Yogo (2002). Our baseline specification with Open yielded a value for this statistic of 4.85 exceeding the critical value of 4.58, while the specification with Ropen yielded a value of 3.89.<sup>28</sup>

In sum, we do not find the case for “real openness” particularly compelling. We worry that the “more robust” results that Alcalá and Cicconé claim for it derive from the interaction of strong reverse causality with imperfections of the instrument.

27 A little exploration reveals why the instruments work much better with openness than with real openness. The first stage regressions associated with estimating the equation described in the previous footnote, which is based on the decomposition of real openness into openness and price, show that the first-stage for the price level equation has an  $F$ -statistic of 1.92. Apparently, the instruments do much worse with real openness because of the very weak correlation between the instrument set and the price level. Another issue is why Alcalá and Cicconé use such an odd instrument list, entering the levels of population and land area, as well as their logs, whereas the second-stage equation has only the logs. It is hard to defend the idea that the level of land area, say, can be safely excluded from the second stage when its log belongs in it.

28 The tabulated critical values for the weak instrument test for an exactly identified problem of two endogenous variables such as ours are, respectively, 7.03, 4.58, and 3.95 for desired maximal sizes of  $r$  in a 5 percent linear restrictions Wald test of 0.10, 0.15, and 0.20.

## Appendix B: Data and Sources

AFRICA = Dummy variable taking value 1 if a country belongs to Africa, 0 otherwise.

ASIA = Dummy variable taking value 1 if a country belongs to Asia, 0 otherwise.

ACCESS = Dummy variable taking value 1 for countries without access to the sea, 0 otherwise.

AREA = Land area (thousands sq. mt.) Source: Frankel and Romer (1999).

ASIAE = Dummy variable taking value 1 if a country belongs to South-East Asia, 0 otherwise.

CATH = Dummy variable taking value 1 if the country's population is predominantly catholic.

COLFR = Dummy variable taking value 1 if the colonizer was France.

COLUK = Dummy variable taking value 1 if the colonizer was England.

DISTEQ = Distance from Equator of capital city measured as  $\text{abs}(\text{Latitude})/90$ .

ENGFRAC = Fraction of the population speaking English. Source: Hall and Jones (1999).

EURFRAC = Fraction of the population speaking one of the major languages of Western Europe: English, French, German, Portuguese, or Spanish. Source: Hall and Jones (1999).

FROSTAREA = Proportion of land with  $> 5$  frost-days per month in winter. Source: Masters and McMillan (2001).

FROSTDAYS = Average number of frost-days per month in winter. Source: CID Harvard University (2002) from Masters and McMillan (2001).

LAAM = Dummy variable taking value 1 if a country belongs to Latin America or the Caribbean, 0 otherwise.

LCGDP95 = Natural logarithm of per capita GDP in Purchasing-Power-Parity US dollars (PPP GDP). Source: Penn World Tables, Mark 6, in 1995. For the following countries the 1996 CGDP value in the PWT was utilized: Bahrain; Kuwait; Mongolia; Oman; Qatar; Saudi Arabia; Swaziland. The 1995 value for CDGP95 in PWT, Mark 6, was obtained by dividing the 1996 observation by the corresponding percentage change in real GDP computed using International Financial Statistics (IFS, IMF 2002) nominal GDP and GDP deflator series.

LCOPEN = Natural logarithm of nominal openness. Nominal openness is given by the ratio of nominal imports plus exports relative to GDP (in exchange rate US dollars). Source: Penn World Tables (PWT), Mark 6. Average overall 1950–98 available data. For the following countries the 1996 observation available in the PWT has been utilized: Bahrain; Bahamas; Bermuda; Bhutan; Djibouti; Eritrea; Georgia; Kuwait; Laos; Mongolia; Oman; Qatar; Saudi Arabia; Sudan; Swaziland; Tajikistan; Turkmenistan. For Myanmar and Suriname LCOPEN (Average over all 1950–95 available data) is constructed from IFS (IMF) using Exports and Imports in Goods and Services in national currencies and nominal GDP in national currencies for the available years.

LFR = Dummy variable taking a value of 1 if a country has a legal system deriving from that in France.

- LNOOPEN = Natural logarithm of “real” openness. Real openness is given by the ratio of nominal imports plus exports to GDP in Purchasing-Power-Parity US dollars (PPP GDP). Source: Penn World Tables, Mark 5.6 and World Bank (2002).
- LOGA = Labor-augmenting technological progress parameter in 1998. Source: Hall and Jones (1999).
- LOGEM4 = Natural logarithm of estimated European settlers’ mortality rate. Source: Acemoglu et al. (2001).
- LOGFRANKROM = Natural logarithm of predicted trade shares computed following Frankel and Romer (1999) from a bilateral trade equation with “pure geography” variables. Source: Frankel and Romer (1999).
- LOGFRANKROMR = Natural logarithm of predicted trade shares computed as for LOGFRANKROM except that the dependent variable in the bilateral trade (gravity) equation is nominal trade divided by nominal GDP (both in US dollars). Source: Authors’ estimates.
- LOGHL = Natural logarithm of human capital per worker in 1988. Source: Hall and Jones (1999).
- LOGKL = Natural logarithm of physical capital per worker in 1988. Source: Hall and Jones (1999).
- LOGYL = Natural logarithm of GDP in Purchasing-Power-Parity US dollars (PPP GDP) per worker in 1988. Source: Hall and Jones (1999).
- LSO = Dummy variable taking a value of 1 if a country has a socialist legal system.
- MALFAL94 = Malaria index, year 1994. Source: Gallup and Sachs (1998).
- MEANTEMP = Average temperature (Celsius). Source: CID Harvard University (2002).
- MUSL = Dummy variable taking value 1 if the country’s population is predominantly Muslim.
- OIL = Dummy variable taking value 1 for a country being major oil exporter, 0 otherwise.
- POP = Population.
- PROT = Dummy variable taking value 1 if the country’s population is predominantly protestant.
- RULE = Rule of Law index. Refers to 2001 and approximates for 1990’s institutions Source: Kaufmann et al. (2002).
- SAFRICA = Dummy variable taking value 1 if a country belongs to Sub-Saharan Africa, 0 otherwise.
- SW = Dummy variable taking value 0 if the country had BMP = 1, MON = 1, SOC = 1, TAR > 0.4, or NTB > 0.4; 1 otherwise. Source: Sachs and Warner (1995).
- TROPICS = Percentage of tropical land area. Source: Gallup and Sachs (1998).
- XCONST1970 = Constraint on the executive in the 1970s. Source: Polity IV dataset.

### Acknowledgments

We thank three referees, Chad Jones, James Robinson, Will Masters, and participants at the Harvard-MIT development seminar, joint IMF-World Bank Seminar, and the Harvard econometrics workshop for their comments, Daron Acemoglu for helpful conversations,

and Simon Johnson for providing us with his data. Dani Rodrik gratefully acknowledges support from the Carnegie Corporation of New York.

## References

- Acemoglu, D., S. Johnson, and J. A. Robinson. (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91, 1369–1401.
- Acemoglu, D., S. Johnson, and J. A. Robinson. (2002). "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics* 117, 1231–1294.
- Acemoglu, D., S. Johnson, J. Robinson, and Y. Thaicharoen. (2003). "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth," *Journal of Monetary Economics* 50, 49–123.
- Alcalá, F., and A. Ciccone. (2004). "Trade and Productivity," *Quarterly Journal of Economics* (forthcoming).
- Angrist, J. D., and A. B. Krueger. (1991). "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106, 979–1014.
- Belsey, D. A., E. Kuh, and R. E. Welsch. (1980). *Regression Diagnostics*. New York: John Wiley and Sons.
- Diamond, J. (1997). *Guns, Germs, and Steel*. New York: W.W. Norton & Co.
- Dollar, D., and A. Kraay. (2004). "Trade, Growth, and Poverty," *The Economic Journal* 114, 22–49.
- Easterly, W., and R. Levine. (2003). "Tropics, Germs, and Crops: How Endowments Influence Economic Development," *Journal of Monetary Economics* 50, 3–40.
- Engerman, S. L., and K. L. Sokoloff. (2004). "Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States," National Bureau of Economic Research Working Paper No. H0066.
- Frankel, J., and D. Romer. (1999). "Does Trade Cause Growth?," *American Economic Review* 89, 379–399.
- Gallup, J. L., and J. D. Sachs. (1998). *The Economic Burden of Malaria*. Center for International Development, Harvard University, mimeo.
- Gallup, J. L., J. D. Sachs, and A. D. Mellinger. (1998). "Geography and Economic Development," National Bureau of Economic Research Working Paper No. w6849.
- Hall, R., and C. I. Jones. (1999). "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114, 83–116.
- Kaufmann, D., A. Kraay, and P. Zoido-Lobatón. (2002). "Governance Matters II—Updated Indicators for 2000/01," World Bank Policy Research Department Working Paper No. 2772, Washington DC.
- Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. E. Sachs, and J. Sachs (2003). *A Global Index Representing the Stability of Malaria Transmission*. Center for International Development, Harvard University, mimeo.
- Krueger, A., and A. Berg. (2002). "Trade, Growth, and Poverty," Paper presented at the 2002 World Bank Annual Conference on Development Economics, Washington DC.
- La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. W. Vishny. (1998). "The Quality of Government," *Journal of Law, Economics, and Organization* 15, 1113–1155.
- Masters, W. A., and M. S. McMillan. (2001). "Climate and Scale in Economic Growth," *Journal of Economic Growth* 6, 167–186.
- Mukand, S., and D. Rodrik. (2002). *In Search of the Holy Grail: Policy Convergence, Experimentation, and Economic Performance*. Harvard University, mimeo.
- North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. New York: Cambridge University Press.
- North, D. C. (1994). "Economic Performance Through Time," *The American Economic Review* 84, 359–368.
- Rodríguez, F., and D. Rodrik. (2001). "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." In B. Bernanke and K. S. Rogoff (eds), *Macroeconomics Annual 2000*. Cambridge, MA: MIT Press.
- Rodrik, D. (2003). "Institutions, Integration, and Geography: In Search of the Deep Determinants of Economic Growth." In D. Rodrik (ed.), *In Search of Prosperity: Analytic Country Studies on Growth*. Princeton, NJ: Princeton University Press.

- Sachs, J. D. (2001). "Tropical Underdevelopment," National Bureau of Economic Research Working Paper No. w8119.
- Sachs, J. (2003). "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income," National Bureau of Economic Research Working Paper No. 9490.
- Sachs, J., and A. Warner. (1995). "Economic Reform and the Process of Global Integration," *Brookings Papers on Economic Activity* 1, 1–118.
- Sala-i-Martin, X., and A. Subramanian. (2003). "Addressing the Curse of Natural Resources: An Illustration from Nigeria," National Bureau of Economic Research Working Paper No. 9804.
- Staiger, D., and J. H. Stock. (1997). "Instrumental Variables Regression with Weak Instruments," *Econometrica* 65, 557–586.
- Stock, J. H., and M. Yogo. (2002). "Testing for Weak Instruments in Linear IV Regressions." In D. W. K. Andrews and J. H. Stock (eds), *Festschrift in Honor of Thomas Rothenberg*. Cambridge: Cambridge University Press, forthcoming.