4 | Studying Politics

ILLUSTRATIONS

4 | Studying Politics

This part examines some of the tools that political scientists use to make their truth claims. Two themes are prominent: the diversity of approaches that spans the subfields, and the strength of the claims that link methods to substantive areas of interest. Attention to the microfoundations of institutions as well as to choice and strategic interaction partner easily with rational-choice theory, while a focus on large processes and slow development conduces to a more sociological and historical approach to institutionalism. These essays contain many suggestions for moving across boundaries in order to use multiple methods.

David Laitin's "Comparative Politics: The State of the Subdiscipline" is a case in point. He issues a call for a multimethodological approach. As a distinct subdiscipline, comparative politics is defined both by substantive and methodological criteria. Substantively, research in comparative politics seeks to account for the variation in outcomes among political units on consequential questions that have been posed in political theory. Methodologically, an earlier consensus about the comparative method differentiated it from the statistical and case study methods and emphasized the isolation of key variables through the use of strategic controls. Today, Laitin argues, a new consensus is on the horizon, one that emphasizes a tripartite methodology, including statistics, formalization, and narrative. All three methods involve a constant interaction between theory and data (and thus are all theoretic) but with different emphases. In the first component of the tripartite method, cross-sectional or diachronic data are employed to find statistical regularities across a large number of similar units. The second component of the tripartite method is formalization. Formal modeling, in providing an internally consistent logic that accounts for the stipulated relationships among abstract variables, assures us that our causal stories are coherent and noncontradictory. The observable implications of our formal models, once they are derived, invite statistical test. The third component is narrative. Comparativists examine real (and virtual) cases to trace historically and theorize empirically the translation of values on independent variables to values on dependent variables. If statistics addresses questions of propensities, narratives address questions of process.

To be sure, the tripartite subdiscipline reported in this paper is only emerging. To illustrate the progress in the comparative politics discipline over the past decade, Laitin examines work on three consequential political outcomes, each of which is connected to perennial concerns in political theory and posseses political relevance today: democracy, civil war, and forms of capitalism. For each outcome, he reports on the progress coming from research based on each component of the tripartite methodology, while noting missed opportunities for advances because of inattention across the methodological divides. The aim of this essay is to advance the view that even with a focus on the big dependent variables, cumulative findings can be developed if done in the context of the tripartite methodology.

Barry Weingast focuses on key developments in "Rational-Choice Institutionalism." From pluralistic premises noting that no approach dominates the study of institutions, he explains the comparative advantages of the rational-choice perspective in covering three categories of questions: the effects of institutions, the necessity of institutions, and the endogenous choice of particular institutions, including their long-term durability and survival. Rooted in the economic theory of the firm, economic history, and positive political theory, this approach possesses distinctive features in providing microfoundations for institutional analysis. Methodologically, treating institutions as humanly devised constraints translates into studying how institutions constrain the sequence of interaction among the actors, the choices available to particular actors, the structure of information and hence beliefs of the actors, and the payoffs to individuals and groups.

The rational-choice approach to institutions divides into two separate levels of analysis. In the first, analysts study the effects of institutions, treating them as exogenous; in the second, treating them as endogenous, they study why institutions take particular forms, why they are needed, and why they survive. In combination, these approaches provide both a method for analyzing the effects of institutions and social and political interaction and a means for understanding the long-term evolution and survival of particular institutional forms. The study of endogenous institutions yields a distinctive theory about their stability, form, and survival. In contrast to approaches that take institutions as given, this approach allows scholars to study how actors attempt to affect the institutions themselves as conditions change. This approach potentially provides the microfoundations for macropolitical phenomena such as revolutions and critical elections. Explicit models of discontinuous political change, Weingast concludes, provide an exciting new set of applications of rational-choice theory.

Paul Pierson and Theda Skocpol assess the trajectory of "Historical

Institutionalism in Contemporary Political Science." Contemporary political scientists are familiar with leading examples of historical institutionalist research without necessarily realizing that they exemplify a coherent genre that analyzes organizational configurations where others look at particular settings in isolation; and they pay attention to critical junctures and long-term processes where others may look only at slices of time or short-term maneuvers. Researching important issues in this way, historical institutionalists, they argue, make visible and understandable the overarching contexts and interacting processes that shape and reshape states, politics, and public policymaking. Despite the great variety of their studies, historical institutionalists commonly employ distinctive and complementary strategies for framing research and developing explanations. Three important features especially characterize historical institutional scholarship. Practitionaers address big, substantive questions that are inherently of interest to broad publics as well as to fellow scholars. They take time seriously, specifying sequences and tracing transformations and processes of varying scale and temporality, and they analyze macrocontexts and hypothesize about the combined effects of institutions and processes rather than examining just one institution or process at a time. Pierson and Skocpol show both how historical institutionalism is experiencing a bold new phase of methodological development and how its focus on substance and its theoretical eclecticism open the way for fruitful cross-fertilization with sister research traditions.

Karen Orren and Stephen Skowronek evaluate recent research on the United States in the genre of historical institutionalism in their essay "The Study of American Political Development." Situating American Political Development in relation to recent intellectual trends in the study of political history, social theory, and rational choice, they argue that its growing influence reflects a more general reconsideration of the temporalities of governance and a recognition that theory building requires more careful attention to the distinctive ways in which politics changes over time. Their paper marks the field's departure in recent years from the assumptions of consensus theory and from realignment theory's scheme of periodization. Orren and Skowronek then distinguish three concurrent lines of inquiry into how institutions affect political change over time and conclude with a sketch of emergent conceptions of history and their implications for theory building.

Robert Powell's "Game Theory, International Relations Theory, and the Hobbesian Stylization" argues that strategic interaction is at the center of much of the work that has been done in international relations since at least the end of World War II and claims that one very important approach to understanding this interaction abstracts away from many of the details of international politics and foreign policy to focus on the strategic logic of a simple, stylized model of the international system. Game theory, in turn, is a central tool for analyzing strategic interaction. Not surprisingly, then, international relations theorists have tried at various times over the last half-century to use game theory to advance their understanding of these problems. Powell thus asks what game theory offers international relations theory, inquires about the extent to which what game theory has to offer is needed, and surveys some of the effects on international relations theory that come with adopting a more game theoretic approach. Above all, Powell emphasizes that game theory is not a theory of international relations but a research tool. Its use entails costs and benefits; however, he argues, after a review of some of the major debates in the field, that its benefits outweigh its costs. Formalization forces basic mechanisms underlying all political interaction to the fore and thereby facilitates a more-integrated approach to the study of politics more generally and international politics in particular.

Charles Cameron and Rebecca Morton advocate merging formal theory and empirical methods in "Formal Theory Meets Data" to fashion a new genre of political studies. First discussing what they mean by a formal model and the general relationships between models and empirical analysis, they contend that empirical analysis based on nonformal theorizing has advantages and disadvantages: it allows a researcher to be open to new and unexpected patterns in the data, but it also means that the researcher may miss important causal factors by using goodness of fit and statistical significance as his only judges as to the success of his analysis. They hold that formal theory provides empirical analysis with a guide that can help a researcher uncover causal factors by suggesting which variables should be measured and the expected functional form of the relationships between the variables in the empirical estimation. However, empirical analysis based on formal models is not easy. Working with a formal model as the basis for devising the empirical model means that the researcher must confront difficult questions about the relationship between the two. They show how researchers either work with empirical models that are inspired by the formal model or directly derive the empirical model from the formal model, using "structural" estimation techniques. And they indicate how researchers handle randomness when they undertake empirical analysis of formal models either by using experimental analysis and thus reducing the causes of randomness through experimental design control or by directly incorporating the randomness within the structure of the formal model.

In "Reclaiming the Experimental Tradition in Political Science," Donald Green and Alan Gerber argue for a different approach to the study of politics than the previous authors, one that is now gaining ground in cognate fields, including economics. They claim that the tradition of field experimentation has been largely forgotten in political science, having been supplanted by other methodologies. Nevertheless, field experimentation remains a powerful method for drawing inferences about cause and effect, often more so than the observational studies and lab experiments that currently dominate empirical research. And they contend that field experimentation based on randomized interventions is more valuable and feasible than political scientists typically suppose. While perhaps most feasible in stable domestic contexts, they see no inherent reason why other areas of politics cannot usefully appropriate such experimental methods.

DAVID D. LAITIN

Comparative Politics: The State of the Subdiscipline[1]

Comparative politics is a distinct subdiscipline of political science, defined by both substantive and methodological criteria. Substantively, research in comparative politics seeks to account for the variation in outcomes among political units on consequential questions that have been posed in political theory. Comparative politics research places these outcomes on dimensions and seeks to account for the placement of political units on these dimensions. It then seeks to account for differences in placement along these dimensions among political units and for the same political unit in different time periods.

Methodologically, an earlier consensus about the comparative method differentiated it from the statistical and case-study methods, and it emphasized through the use of strategic controls the isolation of key variables. Beholden to the discussions of J. S. Mill, comparativists worked out the implications of using the method of similarity, or the method of difference, to capture the workings of independent variables.[2] Today, a new consensus is on the horizon, one that emphasizes a tripartite methodology, including statistics, formalization, and narrative.[3] All three methods involve a constant interaction between theory and data (and thus are all theoretic), but with different emphases.

In the first component of the tripartite method, cross-sectional or diachronic data are employed to find statistical regularities across a large

1. The author would like to thank Kanchan Chandra, Peter Gourevitch, Donald Green, Peter Katzenstein, Ira Katznelson, Peter Lange, Lisa Martin, Helen Milner, and Gerald Munck, all of whom commented on earlier versions of this essay. Mary-Lee Kimber provided research assistance for it.

2. The classic statements on the comparative method, by Eckstein, by Lijphart, by Przeworski and Teune, by David Collier, and by Skocpol and Somers are all cited and neatly developed in Lichbach and Zuckerman 1997b.

3. My identification of a tripartite methodology takes a step beyond the approach advocated in Bates et al. 1998, in which formalization and narrative are highlighted but statistical work is largely ignored.

number of similar units. Whereas statistical techniques were once seen as alternatives to the comparative method, they are increasingly seen as an important element in that method but only one step toward explanation.

The second component of the tripartite method is formalization. Formal modeling, in providing an internally consistent logic that accounts for the stipulated relationships among abstract variables, assures us that our causal stories are coherent and noncontradictory. The observable implications of our formal models, once they are derived, invite statistical test. These exercises in comparative statics have set new methodological challenges in comparative politics.[4]

The third component in the tripartite methodology is narrative. Comparativists examine real (and virtual[5]) cases to trace historically and theorize empirically the translation of values on independent variables onto values on dependent variables. If statistics addresses questions of propensities, narratives address questions of process. Narratives are helpful in suggesting how statistical anomalies arising from contradictory results in cross-sectional and diachronic analyses can be resolved. Furthermore, narratives provide reliable information on the measurement of key variables. Also, in juxtaposing theory to cases, as comparative narratives demand, methods of similarity and difference are especially useful in making sense of cases that are off the regression line. These "Millian" exercises have the potential to discover formerly omitted variables that, once plugged into statistical analyses, alter previous findings.

To be sure, the tripartite subdiscipline reported in this paper is only emerging, and it has hardly been noticed by leading practitioners. Indeed, the editor of *Comparative Political Studies* wrote in the introduction to a special issue devoted to the organization of the field that comparativists have a commitment to "explanatory accuracy" that accounts for a "fragmented discipline" (Caporaso 2000, 699–700). Caporaso's judgment is correct, and the subdiscipline has paid a cost for not taking full advantage of the emergent consensus. As this paper will show, opportunities for advance have been missed because of the field's fragmentation. Thus, in identifying an *emergent* subdiscipline, this review plays the Leninist role of pushing history.

The dependent variables that engage the attention of the comparative politics subdiscipline are not timelessly and unambiguously arranged, like the unanswered conundrums that drive mathematicians. There are two crucial differences between the questions that drive political scientists and those that drive mathematicians. In comparative politics, questions are chosen because they have vital interest for the world we live in. Questions

4. On testing the observable implications of our models, see Geddes 1991. On the challenges of testing formal models statistically, see Alt and Alesina 1996.

5. See Lustick 2000 for the use of virtual data to study the construction of ethnic identities.

concerning democratization were prominent on the agenda of comparative politics for the past decade in large part because they were on the agenda of citizens, politicians, and the informed public around the world. Comparativists will drop old questions, not because they are solved but because new questions have pushed their way onto the political agenda. And so, in the wake of events of September 11, 2001, Hobbesian questions of making the world safe from terror may supplant Lockean questions of making the world open to freedom on the comparative field's agenda. Choice of the dependent variable cannot be separated from the goals, interests, and generational perspectives of researchers. Reviews of progress in comparative politics by different reviewers or written in different periods will therefore highlight different dependent variables.[6]

Also, questions comparativists ask about outcomes continually get specified anew, as the way we ask our questions about political outcomes changes over time. In the Hobbesian period, civil war meant the collapse of monarchy; in today's world, it increasingly means rebellions fought in the name of an ethnic group against state authority. While there may well be explanatory factors that cross eras and types of civil wars, small respecifications of a dependent variable can have large repercussions concerning the significance of independent variables. Concerning democracy, Moore (1966) was attentive to the protection of individual liberties. Today, some democratic theorists are attentive to the protection of property rights. These different foci imply different dimensions, although both could use "degree of democratic consolidation" to name their variable. Researchers' values can not so subtly influence the specification of a dependent variable with implications for the explanatory power of different independent variables. The downside of this incentive to respecify problems given changing concerns within the research community is that questions in comparative politics never get satisfactorily solved, as on the brink of discovery they get specified in a new way, opening up new lines of inquiry.

In light of this tendency toward fragmentation through multiple specifications of the big dependent variables, are there better ways to organize the field? My attempt to create a consensus is not the only alternative. Lichbach and Zuckerman's work (1997a) reflects a widespread belief that the field is divided by a set of paradigmatic approaches—structural, cultural and rational choice—each with its own insights. Alternatively, many remain indebted to Samuel Beer's teachings to focus on the relative power of three independent variables: interests, ideas, and institutions. There are several advantages to the establishment of coherence based on indepen-

6. My predecessor for this decadal review, Rogowski (1993), made no mention of comparative democracies as on the agenda. See Shin 1994, 138, n. 9, for the tide of publications on democracy that followed on the heels of Rogowski's review. Of the three dependent variables singled out for attention in my review, only one (forms of capitalism) received serious attention by Rogowski.

dent variables. First, it is surely the case that some independent variables are causal for more than one dependent variable, and the focus on dependent variables tends to ignore the broader implications of changes in powerful independent variables. For example, many comparativists (e.g., Collier and Collier 1991), focus on the independent affects of "world historical time" and its implications for a variety of political outcomes. The cumulative impact of this variable is missed when each study isolates a single dependent variable. Indeed, this review, with its focus on three dependent variables, gives no treatment to critical junctures in world historical time as an independent variable. Second, the focus on big dependent variables can be intractable. We may never be able to answer the question "What causes democracy?" but we can make considerable progress if we ask instead, "What is the effect of economic growth on the survival chances of democracy?" (Geddes 1991). Third, since compelling explanatory variables are few and attractive dependent variables are many, it is easier to keep a catalogue of cumulative findings if the field is organized by independent variables.

Nonetheless, the identification of competing approaches or alternative explanatory variables is, in my judgment, the wrong way to go in developing a discipline. Doing so focuses attention on the explanatory limits of a particular method or a favored variable rather than the degree to which we collectively in comparative politics have accounted for variations on important outcomes. We would not be interested in learning about the effect of x on y unless we cared about the value of y. True, the specification of big dependent variables in comparative politics induces some fragmentation. But the thrust of this paper is to show that even with a focus on the big dependent variables, fragmentation can be contained, and cumulative findings developed if done in the context of the tripartite methodology.

To illustrate the progress in the comparative politics discipline over the past decade, I shall examine work on three consequential political outcomes, each of which is connected to perennial concerns in political theory. But each has political relevance in our age with a corresponding high level of attention in comparative research. The three outcomes are democracy, civil war, and forms of capitalism. For each outcome, I will report on the progress coming from research based on each component of the tripartite methodology. I will also report on the missed opportunities for advances because of inattention across the methodological divides.

■ | Democracy

Comparative studies of democracy and its alternatives have focused on the factors that differentiate democratic countries from nondemocratic ones. Here I will begin with studies in the tradition of Lipset (1959), relying primarily on statistical analysis. I then examine new studies, in the tradition of Moore (1966), relying principally on historical narratives and the patterns uncovered by these narratives. Finally, I discuss recent formal models of democratization.

STATISTICS

What distinguishes democracies from nondemocracies? This is a question that begs for statistical analysis. Przeworski et al. (2000) have made a fundamental contribution to comparative politics in compiling a data set that enables them to provide fresh answers to this age-old question.[7] Their data are consistent with Lipset's finding (1959) that there is a strong relationship between economic development and democracy. But Lipset's data did not allow him to distinguish two possible reasons for this correlation. Are democracies the result of modernization? Or do democracies survive more successfully once a certain level of economic growth is attained? Meanwhile, poor democracies fall into dictatorship. In this scenario, democracy tends to survive if a country is modern, but democracy itself may arise randomly, exogenous to the level of economic development.

Przeworski et al. provide powerful evidence that modernization is not the cause of democracy. They collected data from 141 countries annually from 1950 through 1990 and coded them dichotomously as to whether they were democracies. Their metric, the probability of a transition to democracy, shows "dictatorships survived for years in countries that were wealthy by comparative standards . . . conversely, many dictatorships fell in countries with low income levels"[8] (2000, 94).

Meanwhile, their data show that if "the causal power of economic development in bringing down dictatorships appears paltry . . . per capita income has a strong impact on the survival of democracies" (98). In fact, over $7,000 in per capita income brings a zero probability of the collapse

7. There is no consensus on the specification of the dependent variable. Collier and Levitsky (1997) demonstrate that in political science there are nearly as many types of democracy as there are studies of it. Those working in the Moore tradition (1966) face the challenge that the master specified the dimension of democracy to dictatorship in an ad hoc way, differently for each country studied. Przeworski et al.'s minimalist specification (2000) is in the Schumpeterian tradition (1942). O'Donnell, in the tradition of Dahl (1956, Appendix), presses for a maximalist specification, as he sees no real democracy unless the informal workings of institutions squelch particularism (1997, 46). Bollen (1993) and Coppedge and Reinicke (1990) have used techniques such as factor analysis to arrive at an underlying dimension of democracy richer than Schumpeter's. Schaffer (1998) argues for a varied specification that is more sensitive to local cultural meanings. Clearly our findings on democratic outcomes are affected by these specification choices, and progress on explaining democratic outcomes will continue on multiple paths given these varied specifications.

8. The findings held up with exploratory uses of scaled democracy scores.

of a democracy, where there is a 12 percent chance if income per capita is less than $1,000. The collapse of Argentina's democracy at $6,055 is the highest in the data set. O'Donnell (1978) used the Argentina case to challenge Lipset, but he did this, according to Przeworski et al., by examining a "distant outlier." Three of the four transitions to authoritarianism at per capita incomes of greater than $4,000 occurred in Argentina, and the fourth in Uruguay (Przeworski et al. 2000, 90–98). Przeworski et al. correctly predict 77.5 percent of the 4,126 annual observations merely by knowing per capita income. Furthermore, they find, democracies survive more successfully under conditions of economic growth, whereas dictatorship fail equally under conditions of economic growth and conditions of economic decline.[9]

What about political culture? Lipset (1994), while acknowledging the importance of economic prosperity, insists that legitimacy, the key to the sustenance of democracy, requires a supportive political culture. Diamond (1999b) too insists on the importance of regime legitimacy and a political culture that favors democratic institutions. Survey research, he points out, shows that people condition their support of democracy less on economic conditions and more on the institutional workings of the political system.[10] The corruption of the regime, the behavior of parliamentarians, and the responsiveness of elected representatives all play important roles in assuring legitimization. The key criterion for legitimization is that all significant political actors believe that democracy is appropriate for their society, and all significant political competitors believe that democracy is "the only game in town." Although Diamond acknowledges that economic performance plays a role in all regressions, "many more political variables than economic ones have significant effects" on survey support for democracy (quotes from 65, 193). The strongest advocate of the political culture foundation for democracy is Inglehart who claims "that over half of the variance (in a sample of European or run-by-European states) in the persistence of democratic institutions can be attributed to the effects of political culture alone" (1990, 41).

Statistical respecifications of Inglehart's data by Muller and Seligson (1994) (who also enhance the scope of those data with material from Latin America) show that for most elements of the civic culture package, Inglehart had the causal arrows going in the wrong direction. With changes in the level of democracy across decades as their dependent variable, Muller and Seligson show that interpersonal trust is not an explanation for democ-

9. But see Remmer's findings (1991b). Through a statistical analysis of voter volatility in Latin America, she attacks those who see economic crisis as the death knell for democracies. Przeworski et al. predict correctly in Latin America on the basis of GDP per capita alone, and for them, the economic crises of the 1980s were not consequential. Nonetheless, Remmer's findings merit further testing.

10. Here Diamond (1999b, ch. 5) relies on data from Rose, Mishler, and Haerpfer 1998 and Shin and McDonough 1999.

racy but a result of having experienced a long period of democratic rule. The only variable in the civic culture package that holds up as having independent causal influence on democracy is that of the population favoring moderate reform (over revolutionary change or the suppression of reform).

While not testing political culture, Przeworski et al. (2000) examine other factors besides economic level and growth. Once economic controls are added, however, duration of democracy is not significant. Nor do cultural factors, such as the majority religion, seem to have much explanatory power. Knowing the degree of ethnic fractionalization, which many scholars have seen as an added hurdle for democratic consolidation, adds almost nothing to the predictive power of their hazard model. Educational levels, however, do add predictive power, independent of economic levels.

The one powerful noneconomic predictor of democratic longevity is parliamentary institutions, which are less subject to collapse than presidential democracies. To be sure, parliamentary regimes are more likely in rich countries and presidential regimes were especially prevalent in Latin America. But controlling for wealth and region, parliamentary regimes still survive significantly longer than do presidential regimes.

Przeworski et al. (2000) have set a new standard in research differentiating democratic from nondemocratic regimes. Yet much remains to be done. For one, political system variables have been insufficiently specified to be used in statistical analyses. The presidential-parliamentary dichotomy is especially worrisome, inasmuch as nearly 10 percent of the cases are coded as mixed and there are theoretical reasons, as I discuss in the formal models section, to believe that the finding in favor of parliamentarism may be biased by missing some underlying variable explaining regime choice.

Second, Przeworski et al. (2000) have ignored several opportunities to challenge their economic variables with a variety of institutional ones that are prevalent in the democracy literature. Political system variables tell us little about the capacity of democratic states to protect property rights, secure a rule of law, and administer laws without corruption. Linz and Stepan (1996) give conceptual foundations for newly reconstituted institutional variables. Treisman (2000a) has begun to use data on comparative corruption in a way that can be appropriated by democratic theory. Also ignored is the institutional power of the military. Not only state institutions should be considered, but societal ones as well. Przeworski et al. have no indicators for the strength or density of civil society. In light of the gaggle of books and papers that purport to show the importance of civil society for the consolidation of democracy (in addition to those I've reviewed so far, see Putnam 1993a; Schmitter 1997), it is a surprise that they did not collect systematic data (even if they would need to impute for missing years) on this factor. That Przeworski et al. (2000) do not have well-designed tests for political and societal institutions, in order to see if they alter the coefficients of the economic variables, is an invitation for new research.

Third, Przeworski et al.'s (2000) cross-sectional findings are nearly impossible to interpret causally. What are the mechanisms that undermine poor democracies or sustain rich ones? It seems impossible to narrate the progression of events from democracy to dictatorship or reverse in terms of variables such as per capita income. Here is where the other two prongs of the tripartite methodology come into play. Comparativists need to formalize the discovered relationships, and they need to get down to narrative to see if actors in the real world of democracy are conditioning their behavior on the factors that the formal models highlight.

NARRATIVES

What pushes some countries at specific historical periods into democracy? How do fledgling democracies persevere when they face crises? These are questions that require sensitivity to change over time and lend themselves better to narrative rather than statistical analysis. To be sure, Przeworski et al.'s data allow for some diachronic analysis. But these data do not tell us who precisely is doing the acting and where these people fit into the social spectrum. In the past decade, very much in the Moore tradition, research on the historical role of social classes in the making and unmaking of democracy has addressed these questions, and here I will review the studies of Luebbert (1991), Rueschemeyer, Stephens, and Stephens (1992), and Ruth Collier (1999).[11]

Luebbert examined the maintenance of democratic institutions under the challenging conditions of the interwar depression. Like Moore, he found the key to democratic strength in the interwar period to be in the middle classes. But he demonstrated that the so-called marriage of iron and rye was not, as Moore argued, the source of fascism. In fact, Luebbert showed, rural support for fascism did not require a landed elite. In Germany, Spain, and Italy, rural support for fascism was found mainly in areas in which the family peasantry rather than the landed elites predominated. Only in southern Italy was there a landed elite that could deliver votes, and they (ironically) sided with the liberals (concerned more for patronage than with class conflict). Their support for fascism came only after Mussolini attained power.

Liberal democracy survived in those countries where the middle classes were not divided by religion, language, region, or urban-rural differences. A united middle class was not afraid of workers, who were, without much struggle, incorporated into the electoral system. Workers gained in dignity what they lost in income, and were not responsive to radical unionist programs. Meanwhile, liberal hegemony failed where preindustrial

11. For a review of critical commentaries on Moore's classic, see Weiner 1976. On recent empirical tests of the hypotheses, see Valenzuela 2000 and Mahoney 2000a.

cleavages divided the middle classes. Divided among themselves, the middle classes were afraid to ally with the workers, compelling them to build trade unions for protection. Under these conditions, the workers needed an electoral alliance to defeat the middle classes. When they successfully allied with the middle peasantry or family farmers, social democracy was the result. However, whenever socialists sought to organize the agrarian proletariat in politics, the family peasantry was pushed into an alliance with the middle class, which became a fascist alliance. Thus, one of the bitter historical ironies: where interwar working-class leaders committed themselves to social justice through taking up the cause of the rural workers, they forced a coalition of middle classes and family peasants, and this was the route to fascism.

Not only Luebbert but many others in the Moore tradition give far more attention to the independent role of the working class, which is a factor that plays only a small role in Moore's alliance patterns. Rueschemeyer, Stephens, and Stephens, in their comprehensive historical treatment of western Europe, Latin America, and the Caribbean, cannot find empirical support for Moore's principal claim in regard to the bourgeoisie. *Pace* Moore, Rueschemeyer et al. find that the middle classes, after their inclusion into the power framework, are ambivalent toward democracy. Therefore, it takes the working class (which, unlike peasants, can organize themselves politically) to affect the true balance of power (where no social group can establish hegemony over the others) upon which democracy rests. But in the end, the authors amend their generalization about working class power as the key to democracy. Only under conditions where a party system can effectively protect the interests of the upper classes, they find, will these classes accommodate to the pro-democratic pressures of the working classes.[12]

Collier (1999) also seeks to delineate the role of the working classes in democratization. Examining cases from both the nineteenth and late–twentieth centuries, and from Latin America as well as western Europe, she finds several distinct paths toward democracy. By looking at seven

12. Their original supposition, going into the study, was that the working class was the "most consistently pro-democratic force," except where it was mobilized by a charismatic but authoritarian leader or a hegemonic party linked to the state apparatus. This point makes little sense theoretically. The middle classes only wanted to include themselves and no one below them. This is the same with the working class. Neither was more democratic. It is just that the working class was lower on the totem pole, and once they were included, the vast majority of the population had voting rights. The equation of a particular group's or class's outward commitment to the ideology of democracy with the attribution of causality to that group or class in explaining democratic outcomes is common, especially in the case study literature. See, e.g., Hsiao and Koo, 1997. The key question for democracy is not a group's or class's desire to undermine autocrats, but the probability that a group or class coalition in power will leave power should an out-group win an election. On this point, see Przeworski 1991, ch. 1.

distinct patterns of democratic initiation, she finds that labor plays at least some role in four of those patterns. Her narratives provide plausible evidence of labor's role in democratization across historical periods. But there is a methodological problem with this argument, foreshadowed by the Rueschemeyer et al. recognition (1992) of the need to protect the interests of the upper classes if democracy is to be successfully implemented. Collier only examines labor mobilization in the initiation of successful democracies. If she had coded labor mobilization for every year, she might have found that the higher the mobilization, the lower probability of democratic initiation. This is a real selection bias problem. It could be the case—profoundly undermining the Collier thesis—that the stronger labor shows itself, the more reluctant the right is to accept a democratic constitution. A more complete data set could determine whether Collier's thesis holds, or its opposite.

The historical expositions that accompany studies in the Moore tradition, from nondemocracy to democracy, as well as the reverse, provide a rich narrative complement to the cross-sectional studies. And as is usually the case, the implications of the cross-section and the narrative findings are in some tension with one another. As Rueschemeyer et al. (1992) point out, regional comparisons allow for a large set of sequences under different contexts that can all lead to democracy. Therefore "the similarity of the correlation between development and democracy in different contexts is fortuitous . . . the only underlying homogeneity is the overall balance of power between classes and between civil society and the state. While this is enough to produce the correlation between development and democracy observed in the statistical studies, the same balance of power between pro- and antidemocratic actors can be produced in a large number of ways" (Rueschemeyer et al. 1992, 284).

These historical comparisons are impressively detailed. There has been some cumulating of knowledge. We now have considerable historical evidence that it is the weakening of labor-dependent landed elites rather than the rise of an industrial bourgeoisie that opens the path to democratic politics (Mahoney 2000a). Yet the proliferation of paths and sequences reduces one's confidence in the generality of the findings in any study (Munck 2001a). Either the studies are historically circumscribed, with the authors' being unwilling to make projections about countries in different eras or different areas (as with the case of Luebbert), or the studies are so broad as to lead to a congeries of possibilities and little way of knowing which of many paths will be followed by a case not already in the data set (Collier 1999, Rueschemeyer et al. 1992). Rueschemeyer et al. intimate that better statistical models would include regional dummies, as the patterns seem to be regionally specific. Collier (1999), however, shows that similar patterns can cross regions but not eras. Compelling statistical work, in large part due to the degrees of freedom problem with an immense number of intervening variables, has not kept up with this narrative tradition. To be sure, Dahl (1971) set the standard in hypothesizing multiple paths yet still subjecting his cases to statistical analysis. Ragin (1987) offered an innovative approach to perform statistical analyses of complex sets of hypotheses with multiple paths such as Moore's theory, but there have only been rare and inconclusive tests relying on his techniques (e.g., Grassi 1999). Vanhanen (1997) has run large-*n* statistical tests of democracy that include proxies for Moore's variables (power balance in society). But data on the distribution of societal power hardly provide an historical mechanism that accounts for democratization. Research that is sensitive to the findings in the grand narrative tradition yet keyed into the cross-sectional findings in the statistical tradition remains on the agenda.

Linz and Stepan's work is in the narrative tradition, but they look less at class structure and examine instead the role of institutions. While they acknowledge that high GDP is favorable to democracy (consistent with Przeworski et al. as well as with Moore), they insist that this fact "does not tell us much about *when*, *how*, and *if* a transition will take place and be successfully completed . . . economic trends in themselves are less important than is the perception of alternatives, system blame, and the legitimacy beliefs of significant segments of the population or major institutional actors" (1996, 77). To support this point, and relying on a comparison of the Netherlands and Germany in the 1930s, Linz and Stepan show that the economic decline was equal in both countries but only in the latter there were strong groups able to articulate blame for the economic crisis (77). Earlier (1990a and 1990b), Linz emphasized the importance of political institutions (favoring parliamentary over presidential systems) and with Stepan (1992) the sequencing of elections (favoring a sequence from central elections to regional ones).

In their monumental comparison of transitions and consolidation in southern Europe, South America and post-Communist Europe, Linz and Stepan (1996) point to five necessary conditions for the survival of democracy, which include a vibrant civil society, an autonomous political society, the rule of law, a usable state, and a set of rules, norms, institutions and regulations that undergird an economic society. Furthermore, there are seven independent variables (each with a range of values, most often nominal) that help predict successful consolidation. They include the relationship of state to nation, the type of prior regime, the leadership base of the prior regime, the pattern of the transition to democracy, the legitimacy of major institutional actors, and the environment in which the democratic constitution was drafted.[13] The degree of freedom problem for the testing of these ideas statistically is, however, immense.

13. Before laying out their five necessary conditions and seven independent variables, Linz and Stepan warn their readers, "We will not restrict ourselves to the procrustean bed of this framework. The specificities of history are also important" (1996, xiv).

Until there are more parsimonious narrative models and better-specified variables in narrative accounts, ones that can be coded for cases outside the domain of cases in which the pattern was originally found, the narrative-based approach will not challenge sufficiently findings relying on the statistical approach. But both the statistical and narrative approaches have set new problems for the third element of the tripartite methodology, that is, formal models.

FORMAL MODELS

Formal models have been offered to explain the occurrence of democratic transitions and the creation of stable democratic outcomes. In the early 1980s, the project on transitions from authoritarian rule, edited by O'Donnell, Schmitter, and Whitehead (1986a), set as a premise of the project that democratization was less a product of long-term macro conditions and more of a project of short-term contingent bargains by elements of the political class that had a range of options. This premise set the research program for formal models of the transition period. In one of the latest versions, there are six different categories of actor (radicals and moderates among hard-line rulers, soft-line rulers, and the opposition), each with a different preference ordering among democracy, nondemocracy, and reform (Colomer 2000, ch. 2). Colomer finds three routes to democracy on the basis of an analysis of strategic games among these players. Furthermore, two chapters of narrative illustrate these routes in the Soviet Union and Poland. This formalization of democratic transitions has its illuminating moments, for example, in showing how Gorbachev had no utility-increasing move when confronted by the putschists in August 1991. But there are problems with this approach as well. Consistent with Woodruff's critique (2000) of this genre of modeling, the sociology setting up Colomer's games (e.g., in explaining why a certain set of the six actors was present or not in a particular case) carried far more weight in accounting for the outcomes than the equilibrium analysis itself.[14] But second, and consistent with my tripartite framework, there is no attempt to see if the observable implications of the formal analysis can be put to any form of statistical test. No regularities in the world were systematically accounted for.

Formal democratic theory in the past decade has focused less on transitions and more on the microdynamics of democratic stability. Przeworski (1991) has addressed the problem of why actors out of power might choose not to rebel against democratically elected rulers. Weingast (1997) has ad-

14. Colomer diverges systematically from predictions based on Nash equilibria. It is unclear whether he is using a Nash refinement or a solution concept that ignores Nash. He appeals to efficient equilibria that are Pareto superior to the Nash equilibrium without providing a general solution concept of why these outcomes are likely to be reached.

dressed the problem of why democratically elected rulers might choose not to confiscate property rights (including voting rights) from their enemies—to assure longevity of rule—and thereby undermine the democratic system. Whereas Przeworski asks the conditions under which democracy is immune from revolution from below, Weingast asks the conditions under which democracy is immune from revolution from above.

The theoretical literature remains, however, disconnected from the empirical regularities discussed earlier. The statistical relationship between per capita income and democracy is not accounted for in these models. If this is a robust empirical finding, then our theoretical models should have a parameter for per capita income such that the democratic equilibrium is more unstable to the extent that the parameter goes down in value.

Nor have formal models of democracy shown why some institutional forms are more conducive to democracy than others. What is it about parliamentary rule that makes it more stable than presidential? The dichotomous variable of parliamentary versus presidential hardly captures theoretical intuitions about institutional stability (Shugart and Carey 1992). Does presidentialism allow for the election of nonrepresentative candidates (Linz and Valenzuela 1994; Powell 2000)? Cox shows that this depends on how well voters can coordinate and how strategic they are (1997, 233). Perhaps presidentialism, associated with a two-party system, denies minorities outlets to modulate majorities, outlets that are available to them in the Proportional Representation (PR) systems associated with parliamentary rule? But minorities in two-party systems play a role in pre-election coalition building; meanwhile minorities in PR systems play a role in postelection cabinet building. Neither system is inherently more compatible with minority representation.[15]

Perhaps (in a surmise by Przeworski, personal communication) the key to the difference is that at times of deadlock between the head of state and the parliament, in a parliamentary regime there will be a vote of no confidence, and the possibility of a new government that can address the deadlock anew. Meanwhile the legislature in a presidential system has no way of compelling the head of state to step down. If the deadlock is particularly severe, the president (or the military) has an incentive to declare a constitutional crisis and disregard democratic institutions. One problem with this idea is that early elections in parliamentary regimes rarely affect the seat allocation of the parties, and thus this is hardly a method to overcome interparty deadlock. However, it may be the case that the threat of new elections creates incentives for compromise while the actual calling for elections (due to rare cases where parties have different assessments of their relative electoral strength) reflects the failure of compromise. We would thus ob-

15. These issues are addressed theoretically by Taagepera and Shugart (1989), who set up the terms for a debate that remains lively, most notably in the pages of *Electoral Studies*.

serve no changes in seat strength after such elections, but the threat of such elections could still induce compromise. A formal model focusing on the existence of this threat might show Przeworski's surmise to be correct.

There are alternatives still that could be modeled. Perhaps the answer to parliamentary longevity lies elsewhere still and is to be found by analyzing the endogenous selection of institutions, as suggested in Geddes's empirical analysis (1996) of Latin America and eastern Europe? Londregan (2000) suggests that key differences may be in more microlevel institutions, such as legislative committees. Finally, the answer may turn on whether the parliamentary-presidential dichotomy is hiding some underlying variable, such as the number of "veto points" (Tsebelis 1995), or whether the system is an integrated or bargaining one (Niou and Ordeshook 1997). Whichever is chosen, it would need to be linked to a theoretical argument telling us precisely what it is about each type of system that influences regime longevity. Accounting for the relative success of parliamentary regimes or finding the reason why this success is a robust statistical finding remains a challenge to formalists.

A final gap in the formal literature on democracy is that it ignores insights coming from the class bargaining literature. Modelers might ask, in a system with workers, a middle class, two classes of peasants, and a landed class, under what conditions will working-class mobilization yield democracy? Luebbert's approach, which finds that the bourgeoisie and workers can reach a class compromise under certain conditions, is consistent with a model developed by Przeworski and Wallerstein (1982). But this cannot be the only democratic equilibrium, and formal theorists should be modeling the patterns identified by scholars in the Moore tradition to check for equilibrium possibilities.[16]

The past decade of comparative research on democracy has been a rich one empirically, both statistically and in the exploration of historical and contemporary cases. Formal modeling of democratic transitions and consolidation is making initial inroads. But as the last few points should make evident, the great gap is in the interstices between the methods. Comparativists relying on each of the three methods have been insufficiently reflexive on the advances in their counterparts to ask new questions of data, to model statistically recurrent processes, and to adjust the focus of narrative to variables that come out as important in statistical and formal studies.

16. A more radical approach is suggested by Gourevitch (1998). He points out that Moore's core insight is to find the root of political conflict to be in the axes of cleavage. Since microregulation has replaced macroeconomic policy among the advanced industrial countries as the foundation for core cleavages, Gourevitch finds it unlikely that battles among social classes will impinge on political institutions. The fragmented specialized issues of microregulation, however, will begin to carve their way into political coalitions, conflict, and institutions. If Gourevitch is correct, the Moore tradition should find its way into the microanalytic game theoretic approach that has long been considered its rival.

Order

Since the end of World War II, Hobbesian fears of disorder informed the subfield of international relations, but students of comparative politics could forget Hobbes and ask the Lasswellian questions (1936) of who gets what, when, and how? To be sure, comparativists studied revolutions, but mostly as historical phenomena (Skocpol 1979; Goldstone 1991).

Events in the late 1980s brought Hobbes back into the center of comparative politics. The states of the "second world" collapsed. Several states in the third world, mostly in Africa and Asia, collapsed as well. And an unnoticed trend since the end of World War II became quite clear, begging for explanation. This trend is the decrease in the probability of interstate war and the increase in the probability of civil war. Furthermore, civil wars were increasing in number in large part because they were in many cases interminable, whereas interstate wars have been far more likely to end in a negotiated settlement. With the dependent variable respecified as the ability of a state to withstand collapse, or ethnic and other forms of insurgency, the number of cases facilitated statistical analysis. Data sets such as the Minorities at Risk (MAR) and State Failure allowed comparativists to sort out statistically polities that were more or less subject to rebellion (Gurr 2000; Collier and Hoeffler 2000; Fearon and Laitin 1999). In a complementary effort, formal models seek to identify the causal mechanisms that might be driving the statistical findings. Because the breakdown of order was in many cases caused by insurgents acting in the name of ethnic/national groups, many of the theoretical advances build on the seminal work of Horowitz (1985), whose focus was on ethnic conflict in general. The latest case-based narrative research relies on formal models (Kalyvas forthcoming) and statistical methods (Varshney 2002) and can potentially elucidate the workings of the formally derived mechanisms.

STATISTICS

In standard comparative politics treatments of revolution, it was argued that the number was too few to allow for standard statistical methods. Skocpol (1979) relied on critical comparisons, and Goldstone (1991) on a Boolean schema developed by Ragin (1987). The J curve (Davies 1972) and resource mobilization (Tilly 1978) hypotheses lent themselves to statistical tests, but most tests of these theories have been in the historical narrative tradition.[17]

Ted Gurr's Minorities at Risk and State Failure teams, in their books, articles, and accompanying data sets, rely primarily on the statistical method.

17. Statistical work on the sources of order (because it did not address the great revolutions), and here I refer to the work of Hibbs (1973), tended to get lost in comparative research on revolution.

Gurr (1993) reports on a data project that involved extensive coding on demographic, cultural, social, military, economic, and political variables for 233 "politicized communal groups" from 93 countries in all the world's regions. To be included, groups must either have experienced discrimination or have taken political action in support of collective interests. This data set has been criticized for several problems, probably the most severe being that the cases were chosen on the dependent variable, thereby leaving out many groups which, for lack of mobilization, were not seen as being at risk. In response, and in an attempt to improve the data, the Gurr team has worked with the research community to help correct many of the problems.

What do the data show? Gurr claims that level of group grievances and strength of the group's sense of identity are the most important independent variables (1993, 123–138). Yet, oddly, this reported finding had no statistical support. The wider research community (second-generation users of the MAR data) finds otherwise. Fearon and Laitin (1999) report that the level of GDP per capita in the country (a variable they added to the data set) is the most robust predictor of rebellion. Toft (1998) reports that the geographical concentration of groups in historic territories is a powerful predictor of rebellion. Saideman (2001) reports that foreign support is important for rebellion, and this foreign support is more likely to be forthcoming if the group borders on a country that is dominated by their ethnic kin. Meanwhile, no study controlling for GDP and geographic concentration has shown that level of economic, cultural, or political grievances can differentiate cases of high rebellion from cases with low or no rebellion. In fact Laitin (2000) reports that degrees of language grievance have no relationship at all to rebellion and that in some specifications, there is a weak negative relationship, showing lower levels of rebellion the greater the grievances over language policy.

The second-generation findings from the MAR data set are in accord with many of the central findings from the State Failure Task Force (Esty et al. 1995, 1998). Unlike MAR, the State Failure data set uses country/year as its unit of analysis, and the dependent variable is state failure, a concept that includes revolutionary and ethnic wars, mass killings, and disruptive regime changes. The robust explanatory variables include living standards in the country (measured by infant mortality), level of trade openness, and level of democracy (where "partial" and "recent" democracies are most likely to suffer failure). Consistent with the nonfindings on grievances by second-generation MAR analysts, the State Failure Task Force finds (almost) no support in their statistical models for the hypothesis that ethnic discrimination or domination generates state failure.

NARRATIVES

Case studies of the breakdown of legitimate domination are a growth industry. Here I will review two of the ways in which the dimension of disorder has been analyzed: explaining the collapse of state authority and explaining the eruption of ethnic violence and civil war.

The collapse of the Soviet state—given the wide acceptance in political science that Samuel Huntington (1966) got it right, that is, that Leninist systems might be inept in providing many public goods but could produce order—came as a shock to political scientists, even those who were specialists in Soviet studies. On the causes of the Soviet collapse, area specialists have been divided: Suny (1993) sees it as caused by the emergence of national consciousnesses, seeded by the Soviet state, that could not be contained by that state; similarly Beissinger (2001) sees it as due to the tides of nationalist mobilization that undermined the regime's ability to maintain order; Roeder (1993) sees it as inherent in the sclerotic institutional arrangement of Leninist states, where selected officials had powerful incentives not to innovate; Hough (1997) sees it as caused by the loss of will by the Soviet intelligentsia (and incredibly self-destructive policies by Gorbachev) to lead what was sure to be an extremely difficult political and economic transition; and Solnick (1998) sees it in the loss of confidence by agents of the state in the ability and will of the Soviet leadership to exert domination over government and society, and therefore these agents grabbed as much property as they were able, to ensure themselves a livelihood should the state collapse. As it was rational for any agent to steal from the state, it was rational for all to do so, and thus there was a cascade that emasculated the resources of the Soviet state.

Lohmann's discussion (1994) of informational cascades and the breakdown of the East German regime has a similar dynamic. The lesson these narratives provide for theory is the equilibrium aspects of what once was called institutionalization. Seeing political order as an equilibrium compels us to analyze it in terms of coordinated expectations; suggesting that even highly institutionalized polities, given informational cascades of possible breakdown, can unravel at breakneck speeds.

State collapse in Africa has also generated a significant narrative corpus. Despite a cogent literature elaborating on the weaknesses of the precolonial (Herbst 2000) and the postcolonial African states (Callaghy 1987; Migdal 1988; M.C. Young 1994), professional practice within political science has too long continued to give state officials and state policies priority in its analyses. But with publication by Bayart (1993), Mamdani (1996), and Reno (1995), a radical shift occurred. In Reno's image the "shadow state"—the set of informal networks of state officials, ethnic chiefs, members of secret societies, local thugs, foreign governments, international firms, and independent traders—exerts domination over countries in near-total disregard for the apparatus that claims a seat in the United Nations. The shadow state constitutes political authority; the formal state, that is, the bureaucratic apparatus that negotiates with foreign governments and makes commitments to international agencies, is a ruse. Shadow state networks can topple formal states, and the costs of sustaining a rebellion are low. Foreign patrons, such as Colonel Mohammar Qaddafi, who has been willing to supply training and weapons to support a gaggle of local insurgencies, are a resource of immense importance in organizing a rebellion. Another resource is international aid from NGOs that comes in response to the collapse of the state. This aid is confiscated and deployed by rebels with the same ruthless energy as smuggled diamonds. Ethnic ties are another resource, useful for recruiting armed bands of supporters by local tyrants, but these ties are of far less use in many cases (e.g., Bazenguissa-Ganga 1999) than is often portrayed in accounts that are based more on justifications of the rebellion by rebel leaders than on actual observance of the exploitation of resources by rebels.

Ellis (1999) offers a compelling narrative of state collapse in Liberia. In this shocking yet clear-headed exposition, readers learn of insurgents eating the hearts of their enemies, castrating civilians and keeping the excised organs as trophies. In these dramas, rebels rely on renditions of traditional magic as a resource to sustain domination. Ellis argues that the colonial state was only a thin layer covering indigenous systems of rule. This state dissipated because with the end of the cold war, there were no patrons interested in propping it up. Once dissipated, unconstrained contests for power, in which memories of traditional practices played a powerful role in insurgent strategies, reduced states into anarchy. Ellis's is hardly the last word in accounting for the collapse of the colonial state, in Liberia, in Sierra Leone, in Somalia, in the Congo (Brazzaville), in the Congo (Kinshasa), and in Cambodia, but it is inconceivable that a satisfactory theory of state breakdown will be written that is not informed by the narrative corpus in which Ellis's is a model.

The narrative mode is not limited to a fascination with state collapse. The seminal work of Adam (1971) showed the importance of tracing carefully the mechanics of bureaucratic order especially under conditions where regimes lack legitimacy. Scott (1990) keeps an ethnographer's ear to the ground in order to hear the "hidden transcripts" of resistance, ones that challenge but also reify political order. Wedeen (1999) examines the cult of former Syrian President Hafiz al-Asad, where he was portrayed as a master in all arenas of human action, from diplomacy to pharmacy. Why, Wedeen asks, would a regime spend scarce resources on a cult whose rituals of obedience are transparently phony? Indeed, most Syrians did not believe the cult's claims. The book concludes that Asad's cult operated as a disciplinary device, generating a politics of public dissimulation in which citizens made believe they revered their leader. By inundating daily life with regime discourse, Wedeen argues, the cultists enforced obedience and induced complicity. This study does not rule out other explanations for Asad's success in fostering complicity. Nor does it assess the magnitude of the discourse affect. Nonetheless Wedeen's narrative, which includes cartoons, jokes, and other forms of everyday complicity, illustrates how discourse strategies can be employed as explanatory variables.

As to narratives of ethnically based violence, the most compelling narratives have been provided by Brass (1997) from research in north India. He investigates a range of local incidents, some of which blow up into ugly riots and become classified as "communal violence" in standard accounts. Brass reports on a class of actors known as riot professionals who have an interest in turning everyday forms of local violence into a large-scale communal riot. These professionals may be politicians who need the violence to solidify their voting blocs (as confirmed by Wilkinson 1998); alternatively, they may be entrepreneurs who gain profit from the looting that riots promote. Once an incident catches the attention of riot professionals, they seek to activate the masses, who can use the violence to loot for themselves or to settle scores with local enemies. Kalyvas's microscopic study (forthcoming) of a region in the Greek civil war similarly finds a powerful alliance between urban ideologues who had macroagendas such as Communism and village actors who had local scores to settle and were willing to denounce neighbors as enemies of the occupying army in order to justify murdering them. These ethnographic studies of violence show that the solidarity between leadership and killers in civil war cannot be explained simply by preexisting solidarities. These solidarities must be accounted for in their own right. Furthermore, both the Brass and Kalyvas narratives make clear that ethnically based and ideologically based civil wars may have quite similar dynamics. The separation of ethnic war from civil war as objects of study (as suggested by Kaufmann 1996) seems not to be a useful one.

Comparative case studies complement ethnographies in the narrative tradition. Bunce (1999) compares the dismemberment of the Soviet Union, Yugoslavia, and Czechoslovakia with the goal of differentiating the violent case (Yugoslavia) from those that split apart with minimal violence. She identifies two factors of importance: the interests of the military and whether the dominant national group had its own institutions under the ancien régime. Since there were no unique Russian or Czech institutions in the Communist period, post-Communist leaders of these republics were compelled to minimize tensions between them and those republics that had their own institutional apparatuses. This minimized the level of violence. The Serbs had their own institutions under the ancien régime, and with a military that had a strong interest in maintaining the federation, violence ensued. Varshney (2002) compares the few Indian cities that have had significant communal violence with comparable cities (in terms of demographics, history, and region) where violence has been minimal. He finds that preexisting patterns of civic engagement, where Muslims and Hindus belonged to labor unions, political parties, or business associations, helped cauterize communal conflict before an ugly incident could serve as a spark for violence.

FORMAL MODELS

International relations modelers began addressing the violence that ensued after the collapse of the Soviet Union and the Yugoslav Federation. Posen (1993b)—in a quasi-formal realist model—recognized the internationality situation as quite familiar: anarchic. Relying on a security dilemma framework, he explains cases of violence on the basis of such factors as a national group's overestimation of the weakness of state authority, and the window of opportunity that chaos held for the fulfillment of long-term goals. Walter and Snyder (1999) edited a volume where security dilemma ideas were applied to cases in Africa and Asia as well. However, Fearon (1998b) discounts the mechanism of the security dilemma and hypothesizes that under conditions of newly gained independence, the ruling faction (or ethnic group) was unable—even if it wanted to—to commit to the future well-being of losing factions (or minority ethnic groups). Under such conditions, minorities would have an incentive to rebel early, rather than wait to see if the cheap talk of the ruling group was honest, because to wait so long would mean having a much lower chance of winning in a rebellion.

These international relations models assume that ethnic groups were sufficiently self-organized as to act like states, as unitary actors. The apparent rapid rise of ethnic consciousness and groupness, however, requires some explanation. Kuran (1998) suggests that levels of group solidarity have a cascade or tipping quality to them. If you have some weak ties to an ethnic identification, and an increasing number of people similarly situated begin to wear ethnic clothes, perform ethnic rituals, learn historic languages, and portray themselves as members of that ethnic group, the greater the pressure will be on you to follow suit. Depending on people's hidden preferences for ethnic attachment, it is possible to move from complete demobilization to near-total mobilization in a rather short period. Jack L. Snyder (1993b) identifies a clear signal that sets off a cascade—the weakening of the state. This signal increases demand for protection from one's national group.

Insurgencies, however, are not always the result of state failure; they arise under stable conditions as well. Thus the need for a theory to account for rebellion in light of the failure of the standard grievance models to differentiate countries susceptible to rebellion from those that are not. Collier and Hoeffler (2000) model rebellion as the apex of organized crime. Rebels don't extort from shopkeepers as do mafias, but they control the export of primary produce. Leaders of rebellions therefore need a sufficient number of followers in order to challenge the state military forces at the various choke points in the sale or export of primary products. Subject to the availability of primary products, recruitable looters, and a weak army of the state, rebellions will prosper. Fearon and Laitin (1999) develop a model of insurgency (also opposed to a grievance model) where young men choose whether to join the legitimate economy or to join a rebellion;

meanwhile the state decides how many resources to put into counterinsurgency. These two simple models, though differently constructed, help explain why country-level GNP (worse job opportunities for youths, and lower predicted levels of counterinsurgency spending), availability of primary products, and group concentration of population (especially if concentrated in mountainous zones) are better predictors of rebellion than variables that measure cultural differences or group grievances.

As with democracy, the interesting gaps in the literature on order are the missed opportunities that lie between the different methods. While formal work on the question of the breakdown of order and the rise of civil wars within states has been developing rapidly, it has not kept pace with the cross-sectional and narrative reports. Findings from state failure, for example, linking state failure to low levels of trade openness, have not been formalized. Nor has the failure in statistical models to find any relationship between grievances, discrimination, and inequality and rebellion received adequate formal treatment. Most glaringly, the narratives have portrayed consequential players (e.g., riot professionals) and have shown high levels of intragroup and interstate fragmentation, but formal modelers haven't specified the implications of wars between moderates and radicals among insurgents, or between armies and presidents within states.[18]

Formal modelers who work on signaling and "cheap talk" have not sought to wrestle with the narrative evidence of discourse regimes sustaining political order. The greater the attention to the origins of order and the details of disorder, the more powerful will be our future models. Missed opportunities exist in the other direction as well. Too many narrativists, for example, focus on the ethnic bases of civil wars without paying attention to statistical results showing ethnic difference to be inconsequential in predicting civil war. Brilliant expositions of communal violence (such as Brass 1997 and Varshney 2002) are presented as if formalization would provide no further insight. Yet more attention to the strategy space of rioters may help determine the conditions under which riot professionals will do their professional thing. The comparative literature on order is advancing on all three methodological fronts, but there is insufficient cross-fertilization, which limits solid cumulative gains.

■ | Forms of Capitalism

As Rogowski (1993) highlighted in the previous decadal review of comparative politics, the political economy of the advanced industrial states is a research program of considerable energy and growth.[19] Historical

18. The exception is DeNardo (1985) who models intrarebel dynamics.

19. For an insider's guide through this extensive literature, see P. Hall 1999.

institutionalists set the research agenda.[20] In the classic text of the period (Katzenstein 1978), political strategies among OECD states in adjustment to the collapse of the Bretton Woods system and to the oil shocks constituted the dependent variable. The independent variable was country institutions that were themselves a product of distinct historical trajectories. The institutional capacities and interests of ministries of finance, central banks, commercial banks, and labor unions set limits to and provided opportunities to respond to the economic challenges. Historical institutionalists envisaged a continuum of different types of capitalism, ranging from strong states relative to society (associated with mercantilist political strategies) to strong societies relative to state institutions (associated with liberal political strategies).

Comparative political economy did not coordinate on a single dependent variable for collective analysis. In different studies, economic growth, economic stability, wage equality, redistribution, the social groups paying most heavily for the costs of readjustment, and political strategy were featured on the left side of the field's equations. But in the 1970s there was an implicit and by the 1990s an explicit sense that these outcomes formed into coherent packages which Esping-Andersen (1990) called the different "worlds" of capitalism: liberalism, corporatism, and a third form, associated with Christian Democracy, that presents a unique package of high equality and low taxes, and by so doing, sacrifices growth (Swenson 1989; Iversen and Wren 1998). To be sure, much analysis in this field has specified relationships within each world. But the glue that holds this field together is the question of what caused and what sustains the different worlds of capitalism (Gourevitch 1978).

STATISTICS

"Forms of capitalism" is a vaguely specified dependent variable, and its values are nominal. This suited the historical institutionalists, who were more interested in coherent narratives than high r^2s.[21] But the turbulence of the field was a more formidable problem for variable specification than was any methodological resistance by the historical institutionalists. Consider the problematic of the field a decade ago, as seen by Rogowski (1993), whose dependent variable was "comparative economic growth": "Among

20. Gerschenkron 1962 is the seminal work. For a comprehensive account of historical institutionalism as used in comparative political economy, see Thelen and Steinmo 1992.

21. To be sure, there is a long tradition of statistically based research (much of it done in Europe, but Hibbs 1977 reflects research on both sides of the Atlantic) that revealed stability in the institutional patterns elaborated by the historical institutionalists, with a wide variety of policy outcomes conditioned on the type of capitalism for each Organization of Economic Cooperation and Development (OECD) country.

the economically advanced nations, the continuing Japanese 'miracle' and the quite respectable growth of the continental European economies [that] contrasted sharply with the dismal record of the U.S. and the U.K." (1993, 431). A variety of theories was offered. Hall (1986), for example, sought to account for Britain's long economic decline on the basis of a theory of the interaction of institutions and economic ideas. Others stressed the combination of institutions and interests. However, a decade later the countries were reversed in growth records, and explanations for economic decline had to account for Japan's long recession rather than the United States's. Explanatory models are difficult to nail down when the world economy has been changing so rapidly that it is hard to place political units on any important dimension and to have confidence that the relative values for those political units would stay sufficiently stable as to allow for a community of scholars to account for the variance.

Despite these modeling difficulties, available data on Organizational Economic Cooperation and Development (OECD) states encouraged statistical analysis. Some (such as inflation) were produced and standardized by governments themselves; others, such as indicators of corporatism (Schmitter and Lehmbruch 1979) or central bank independence, required careful construction by the scholarly community; others still, like union concentration, were built from virtual scratch by the scholarly community (for a preliminary analysis of a new data set, see Wallerstein, Golden, and Lange 1997). Using these data, comparativists moved beyond issues of response to economic crisis to other variables on the left side of their expressions: explaining the trade-off on inflation versus unemployment; explaining the level of trade openness; and explaining variation in wage equality (Iversen and Wren 1998).

Much work in this tradition puts forms of capitalism on the independent variable side of the expression. Social democracy is shown to be associated (in contrast to free market liberalism) with a larger government sector, greater equality, public investment in task-specific technical skills, growth advantages in some sectors, and a comparative advantage in the face of economic shocks (Cameron 1978; Hibbs 1977; Garrett and Lange 1986; Garrett 1998b).

Back to forms of capitalism as a dependent variable, perhaps the most hotly debated issue in the study of the comparative political economy of the advanced industrial states is in assessing the impact of globalization. Some see differences weakening. Rodrik (1998) foresees common demands for widespread growth in government spending, especially welfare spending, as a form of insurance against the shock of job loss and social dislocations in the face of globalization, at the terrible cost of losing all mobile capital. Scharpf (1991) and Lambert (2000) see globalization undermining the welfare state in even solidly corporatist governments. Garrett (1998a) is far more optimistic about democratic corporatism and sees it as a best response to the forces of globalization, cushioning market dislocations and providing lucrative investment sites for mobile capital. Iversen and Cusack (2000) challenge this consensus, and argue that the effects of globalization are weak, in comparison to technological changes in production. To the extent that those who see the forces of globalization to be strong, we should expect increasing convergence of structure and strategy of OECD states; to the extent the Iversen-Cusack position is correct, we should expect variations in economic growth, in growth of the welfare state, and in wage equality, depending on technological profiles of state economies.

FORMAL MODELS

Historical institutionalists did not model the patterns that they had discovered in the 1970s. As a result, there were important gaps to be filled. Questions obvious to theorists, such as why there wasn't convergence toward the institutional patterns that were most efficient, did not get addressed. Why, for example, if Britain lacked the political institutions to control the City, could it not construct them to enhance political effectiveness (Blank 1978)? Historical institutionalists did not have a well-worked-out answer to what maintained institutional patterns over time. More important, historical institutionalists emphasized the interaction between politics and economics but did not incorporate this insight into testable models. In the past decade some progress has been made along these lines.

Hall and Soskice, in specifying historical institutions as equilibria, have begun to connect findings of the historical institutionalists with those of the statistical analysts. "Political economists," they write, "have always been interested in the differences in the economic and political institutions that occur across countries. . . . [C]omparative political economy revolves around the conceptual frameworks used to understand institutional variation across nations" (2001a, 1). Relying on endogenous growth theory, which would have us predict that different national rates of growth are conditioned by the institutional structure of the national economy, Hall and Soskice focus on comparative institutions. Since institutions affect the character of technological progress and rates of economic growth, understanding institutions as equilibria plays a direct role in explaining growth.[22]

The Hall-Soskice approach is based on the new economics of organization, with the firm as the fundamental unit in a capitalist economy adjusting to exogenous shocks. In the model, firms reduce risk by making commitments to their workers and to other firms, and this occurs in several spheres: bargaining over wages and working conditions; securing a skilled labor force; getting finance; coordinating with other firms, e.g., on standard setting; and getting employees to act as agents of the firm. Strategies

22. This move, to see the political foundation of modern markets, was foreshadowed by Ruggie's notion (1983) of "embedded liberalism."

to resolve these commitment-coordination problems are conditioned on the national institutional environment. The "national political economies" (2001a, 8) form the principal units of variation, as "we emphasize variations in corporate strategy evident at the national level. We think this is justified by the fact that so many of the institutional factors conditioning the behavior of firms remain nation-specific" (16). The principal dimension on the dependent variable is nations in which "firms coordinate their activities primarily via hierarchies and competitive market arrangements" (liberal market economies, or LMEs) and those in which "firms depend more heavily on non-market relationships to coordinate their endeavors with other actors and to construct their core competencies" (coordinated market economies, or CMEs) (8). The LMEs are Coasian; that is, firms keep arms length from other firms and engage in formal contracting. The CMEs have much incomplete contracting, widespread sharing of private information, and more collaborative interfirm relations. In equilibrium, LME firms should invest in switchable assets that have value if turned to another purpose; CME firms should be more willing to invest in asset specificity, which depends on the active cooperation of others. The separate components of these political economies are complementary, in the sense that high returns from one component entail high returns for another component in the system. So CME firms that give long-term employment contracts profit when they are in a financial system that doesn't punish short-term losses. Complementarity explains the clustering among the solutions to the commitment problems across the spheres of risk.

The model makes predictions in regard to the exogenous shock of globalization. The microeconomists' assumption of pressures to liberalize everywhere, they argue, is based on the (wrong) view that the key to profitability is lower labor costs; this is true for LMEs but not CMEs. Thus the hypothesis is that under globalization, CME firms might relocate in LME countries in order to get access to the radical innovations, while LME firms might move some activities to CME settings to secure quality control and publicly provided skills for labor. (Here they would make a different prediction from Garrett 1998a, who sees the CMEs as having a superior equilibrium in the face of globalization). A second hypothesis is that conflict between labor and capital in the face of globalization will be low in CMEs, where capital and labor often line up in support of existing regulatory regimes (and where labor unions will remain strong), but high in LMEs, where business is pushing hard for deregulation of labor markets (and where labor unions will weaken). The social cleavages that result will therefore be different in the two political economies.

The Hall-Soskice approach takes account of many of the cross-sectional findings in the comparative political economy field, most importantly the apparent stability of social democracy under a wide range of challenges, but also of the intercorrelations of high government spending, social welfare provisions, and union density that come together as a package. However, once you endogenize politics and economics, new forms of statistical testing (as suggested in Alt and Alesina 1996) are in order and remain on the agenda. This approach also takes into account the principal framework of the early historical institutionalists, who took for granted the equilibrium properties of the different forms of capitalism. It makes sense of why we should expect some degree of cross-national variation in effects of the apparently homogenizing force of globalization. (For a complementary theoretical account of why we should expect greater heterogeneity as a result of globalization, see Rogowski 1998.) And finally, it presents a compelling alternative to price theory, which sees institutions as constraints to efficiency but not as sets of equilibria that are dynamically stable. However, as we will see in the next section, the findings in formal theory diverge somewhat from a new generation of narratives in historical institutionalism.

NARRATIVES

From the 1970s, debates within the historical institutionalist research program have been on the causal factors (sectoral balances, timing of industrialization, level of social partnership among classes, and land/labor/capital ratios) explaining the emergence of the different worlds of capitalism. Because of the methodological orientation of the historical institutionalists, the literature they produced was rich in narrative, where case studies (Zysman 1977) and comparisons (Katzenstein 1985; Gourevitch 1986) were the dominant mode. Even Rogowski's strongly theoretical treatise (1989)—where land/labor/capital ratios explained political coalitions—contained historically based narratives elaborating the theory with real-world cases.

The historical institutionalists have continued to write empirically dense narratives that speak directly to the dependent variables that have defined statistical and formal research, but their impact on the practice of statistical and formal modelers has been limited. Consider Katzenstein (1984, 1985). He sought to explain how political stability in the small European states could be maintained under conditions of enormous economic flexibility. The answer was in corporatist governance. He identified two subtypes of the democratic corporatist form of capitalism, liberal and social. Like many of the historical institutionalists, he provided an historical account for these patterns, highly influenced by the structural factor of smallness making these states "takers" rather than "makers" of international rules. What distinguished the first volume (Katzenstein 1984), however, was the careful sectoral analyses in Austria (the social corporatist example) and Switzerland (the liberal corporatist example). In these narratives, the ideology of social partnership coming from a common sense of vulnerability plays an important role in sustaining countrywide institutions (1985, 87–89). This ideological variable is hard to specify for more general explanations, but it should have paved the way for future cross-sectional and theoretical work that encompassed this factor (as well as other variables identified in the narratives), as a test of the magnitude of their impact on sustaining historical institutions. In general the fuzzy variables that attracted the historical institutionalists as consequential rarely find their way in cross-sectional statistical research or in formal theories of the market.

In the 1990s, with questions turning toward the breakdown of institutional differences, Berger and Dore (1996), having observed industrial processes in Japan and Europe, were convinced that national political economies would retain their institutional integrity. In a commissioned set of narrative studies, their intuitions were in large part confirmed, as were those of the historical institutionalists in regard to the oil shocks. National institutions were retaining their historic peculiarities in the face of globalization. Thelen and Kume (1999) find similarly in regard to labor policy in Germany and Japan. It is notable that these studies have not compelled those who have emphasized the homogenizing impact of globalization to figure out why, at least in the short run, the world isn't conforming to their predictions.

Now consider Pierson (1994). He uses the narrative mode in comparing the dismantling of the welfare state under Thatcher and Reagan. By examining two of what Hall and Soskice call LMEs, one would have predicted that under conservative governments, there would have been significant retrenchment. Pierson finds, instead, grand goals but very limited success. Seeking to explain the failures to cut back programs conservative leaders considered inefficient and evil, Pierson finds institutional structure to be of limited power. For example, consider the institutional variable of veto points. The numerous veto points in the U.S. system as compared to the few such points in the U.K. system would lead to the prediction that given the same goals, Thatcher would be more successful than Reagan. Wrong. Reagan, on the margin, was more successful. Pierson finds, instead, that the relative success of programs to survive the conservative onslaught could be explained by the very features of the programs being dismantled. He calls this policy feedback. This variable would not be easy to explore statistically, since every policy has many dimensions of policy feedback, some allowing for easy dismantling, others blocking any change. Consequently, there is no simple value of policy insulation for social programs. Pierson suggests that, "a more promising strategy is to develop middle-range theories that acknowledge both the complexity of feedback and its context-specific qualities" (1994, 171).

But two more concrete proposals suggest themselves. Given the differences the cross-sectional and theoretical literatures find between LMEs and CMEs, Pierson's study should be replicated across this divide to see if policy feedback is consequentially important in the same way in two different political economies. The Hall-Soskice portrayal of LMEs suggests that they would be far more capable of dismantling welfare state programs (and their theory would predict that by holding steady against the welfare state, Reagan and Thatcher effectively held back its predicted development given growth in GDP).[23] But Pierson's study suggests an alternative—that it isn't institutional structure but the policy complexities of welfare state benefits that make them resistant to exogenous shocks. Pierson (1996) examines four cases that do cross the CME/LME divide (Sweden and Germany are added) and finds in contrast to the Hall-Soskice portrayal, there is no clear evidence of CME relative success in sustaining the welfare state.

Second, Pierson's list of programmatic criteria should be organized in a way to allow exploratory statistical tests on policy feedback. In Pierson (1996), data on relative retrenchment are presented cross-nationally but no attempt is made to capture policy feedback (and a set of other proposed independent variables, 176–78) with cross-national data. (More progress is made in Myles and Pierson 2001.) Narrative uncovered a plausible variable to explain crucial outcomes in political economy; this variable requires attention in the formal and cross-sectional domains.[24]

Still to be assimilated by scholars in the comparative political economy field, Herrigel (1996) in his examination of German industrialization finds that the notion of a national economy with its peculiar institutions to be a sham. Close examination shows that there has long been two intersecting German institutional frameworks, each with its own internal logic. The implication of Herrigel's research is that future cross-sectional studies are making a grave error to the extent that they use OECD tapes that rely on country-level trends. To be sure, if central bank independence or monetary policy is the key independent variable, this may present no problem. But if variables such as Katzenstein's (density of social networks) are being tested, Herrigel's work demands that we disaggregate our economic data to the lowest administrative level. To develop such data (though OECD is beginning to collect some data at the level of region) would be an enormous enterprise. But if the variables pointed to by Katzenstein and Herrigel are seen to be consequential, there can be no alternative than to seek major funding for far more disaggregated economic data than OECD supplies the research community for free.

The theoretically informed narratives of the historical institutionalists present several clear challenges and opportunities to the statistical and formal models in comparative political economy. The overall research program remains vibrant; it lacks only the sense of challenge to reconcile inconsistent findings across the tripartite methodology.

23. Pierson does not perform general equilibrium tests of his model. This may help to explain why he believed that the massive budget deficits incurred by Reagan would long endure and help conservative successors, arguing fiscal necessity, to dismantle other parts of the welfare state.

24. Given Lambert 2000, we see that dismantling the welfare state (in Australia) is not as formidable a task as Pierson's book suggests. This variation can easily be taken advantage of in cross-sectional work.

Conclusion

This review has identified a potential coherence in the comparative politics subdiscipline. The hope here has been to present not only coherence but also a frame within which most comparativists will be able to fit their contributions and the contributions of colleagues whose work they most admire. To be sure, not all dependent variables of consequence have been treated herein and not all possible specifications of the dependent variables have been discussed. Nor have consequential independent variables in the classical tradition, such as class structure, political culture, or the socioeconomic status of the population, gotten their own sections. They were brought into the analysis to the extent that theorizing in the past decade linked them to the three dependent variables that received attention here.

Nor still have the big "isms" that dominate many debates received special attention. But the independent variables these isms privilege have gotten attention, whether they were discourse regimes (postmodernism) or country GDP (modernization theory). Unhooked from their isms, these variables can be assessed as to the magnitude of their effect and their interaction effects with variables from other isms.[25] To be sure, many comparativists have bet their careers on the power of independent variables (often associated with a particular paradigmatic school) to explain a range of outcomes. Others have focused on important changes in parameter values that should affect all past explanations. But comparativists who are engaged in these scholarly endeavors should have no trouble in placing their work in the framework provided herein. Moreover, if they cannot link key independent variables or parameters to consequential outcomes, they should reduce their confidence in the importance of the factors that they are analyzing.

Finally, while this review has not highlighted the substantive progress in our subdiscipline, major findings on each of the dependent variables demand summary. High GDP per capita helps explain the consolidation of democracy but not its initiation. Cultural differences and group grievances have no explanatory power in regard to rebellion. Meanwhile, country-level data such as per capita GDP, population size, terrain, and economic growth are better able to explain civil wars. Each form of capitalism constitutes a robust equilibrium and is far less subject to homogenizing effects that one might predict in looking at globalization and the revolution in electronics. (If this last perspective is correct, my successor writing the re-

25. Those who organize the field by paradigm would segregate Wedeen's book, discussed in the section on order, in a postmodernism ghetto. However, by taking a standard postmodern variable—regime of discourse—and showing its role in accounting for order, this review seeks to incorporate postmodern insights into the standard comparative corpus, alongside explanatory models coming from other isms.

view of comparative politics for the next decade will not have a section on the causes of the demise of social democracy.)

The central concern here has been in the organization of research for the comparative field. My argument is not that all comparativists should have highly cultivated statistical, formal, and narrative skills. But, it would be a great loss to comparative politics if only one of these skill sets were to define the subfield. My fear is a Chomsky-like revolution in comparative politics, where the formal theorists drive the field workers out of the subdiscipline. An equal fear is if the field workers put up barricades separating themselves from the findings in the formal and statistical worlds. Utopia is in the social organization of physics. In that discipline, there is a division of labor between the experimentalists and theorists, and a pervasive sense of mutual hostility between the camps. But the difference is that in physics, it is unimaginable that the experimentalists would ignore the implications of the most recent theoretical findings, if only to blow them out of the water. Meanwhile, theorists grudgingly seek to account for empirical realities that experimentalists report.

Methodologically, this review argues that in comparative politics, interdependence across the tripartite methodological divide, with grudging toleration built on mutual suspicion of practitioners across the divide, is a key to scientific progress. Despite the occasional portrayal of a Manichean world of qualtoids versus quantoids, this review shows specks of evidence that a common focus is emerging. Comparativists are finding that in a division of disciplinary labor, they need to satisfy two audiences. First, they must demonstrate that their work meets standards within their own methodological community. But second, they ought to feel challenged, even threatened, by advances by scholars within the other two methodological traditions and feel pressure to adjust or delimit their claims in light of findings in those traditions.

This hoped for interdependence is far more promise than reality. The promise is common focus on consequential dependent variables and a joint attempt to address variance across polities on these variables, by scholars working within three methodological approaches. Those scholars who see this as a reflexive and interdependent division of labor—and not as a war among paradigms—will be remaking the comparative method.

Barry R. Weingast

Rational-Choice Institutionalism[1]

Introduction

Political science has witnessed a revolution in the study of institutions. As Hall and Taylor (1996) emphasize, three approaches predominate: the sociological (March and Olsen 1989), historical institutionalism (Evans, Rueschemeyer, and Skocpol 1985; Thelen 1999), and rational-choice approaches to political institutions (North 1990; Riker 1982). An impressive aspect of this revolution is that no approach dominates. Each perspective has its comparative advantages.

The purpose of this essay is to explain the comparative advantages of the rational-choice perspective. Rational-choice theory provides a distinctive set of approaches to the study of institutions. It covers three categories of questions: the effects of institutions; why institutions are necessary at all; and the endogenous choice of particular institutions, including their long-term durability and survival. Rooted in the economic theory of the firm (Coase 1934, 1960; Milgrom and Roberts 1991; Williamson 1985, 1996), economic history (North 1981, 1990), and positive political theory (Hinich and Munger 1997; Riker 1982; Shepsle and Bonchek 1997), this approach provides a systematic treatment of institutions. Although it has much in common with other approaches to institutions, rational-choice theory has its distinctive features, most importantly, providing the microfoundations of institutional analysis. Applications range across all political and social problems, from the effects of the major political institutions of the developed West (such as legislatures, courts, elections, and bureaucracies) to more recent studies of developing countries (for example, ethnic conflict, international security, equilibrium traps that prevent development, and democratic stability).

1. The author gratefully acknowledges helpful conversations from Randy Calvert, Rui de Figueiredo, John Huber, James Morrow, Theda Skocpol, and Kathy Thelan.

Rational-choice scholars model institutions as "humanly devised constraints" on action (North 1990). Methodologically, this definition translates into studying how institutions constrain the sequence of interaction among the actors, the choices available to particular actors, the structure of information and hence beliefs of the actors, and the payoffs to individuals and groups.

The rational-choice approach to institutions divides into two separate levels of analysis (Shepsle 1986). In the first, analysts study the effects of institutions. In the second, analysts study why institutions take particular forms, why they are needed, and why they survive. The first approach takes institutions as exogenous; the second as endogenous. Moreover, the first level of analysis is clearly antecedent to the second: A choice theoretic approach to institutions requires that individuals have expectations about the effects of institutions. As this level naturally arose first, it is far more developed than the second.

The second level of analysis, in turn, takes the study of institutions to a deeper level. In combination, these approaches provide both a method for analyzing the effects of institutions and social and political interaction and a means for understanding the long-term evolution and survival of particular institutional forms. The study of endogenous institutions yields a distinctive theory about their stability, form, and survival. In contrast to approaches that take institutions as given, this approach allows scholars to study how actors attempt to affect the institutions as conditions change.

This approach potentially provides the microfoundations for macro–political phenomena such as revolutions and critical elections (see, e.g., North and Thomas 1972; North 1981; Stewart and Weingast 1992; Poole and Rosenthal 1997; Riker 1982, ch. 9; Weingast 1998a). Until recently, macro–historical phenomena remained largely the domain of historical institutionalists. Although applications of rational-choice theory are relatively new to these questions, its approach provides links with microbehavior, potentially affording a new methodology for comparison across cases. Explicit models of discontinuous political change provide an exciting new set of applications of rational-choice theory.

This essay begins with the study of the effects of institutions, then takes up the question of why institutions are needed at all, and ends by considering the endogenous choice and survival of institutions.

■ | The Effects of Institutions

Rational-choice approaches begin with a set of individuals, assumed to have a well-defined set of preferences. This first approach takes institutions as exogenous. Institutions affect individual interaction and choice in a variety of ways: institutions constrain individual choices, how individuals interact, their information and beliefs, and their payoffs.

INSTITUTIONAL SOURCES OF THE BALANCE OF POWER BETWEEN THE EXECUTIVE AND THE LEGISLATURE

The relative powers of the executive and the legislature vary considerably across nations. Presidents in many Latin American countries, such as Argentina and Mexico, are sufficiently powerful relative to their legislatures that these systems are often referred by the appellation *presidentialismo.* The U.S. Congress, in contrast, is one of the most powerful legislatures in the world. Prime ministers in most European parliamentary systems are agents of the majority party (or ruling coalition) without electoral independence of the legislature. These differences reflect a host of institutional details about legislative-executive relations. One of the hallmarks of rational-choice analysis is the demonstration of how microlevel details imply macropolitical differences. This subsection provides a comparative framework for showing how various powers and institutional details affect both the legislative-executive balance of power and policy choice.

UNITED STATES

We begin with the context of the United States, where these techniques were first developed. Although the president can propose legislation, only a member of Congress can officially introduce it into Congress. If legislation is brought to the floor, it is typically considered under an "open rule" where any and all amendments are allowed. The executive holds veto power.[2]

To keep the discussion manageable, we consider a single policy dimension, which might be considered a left-right political dimension or a more specific policy issue, such as degree of protection of the environment or the size of the budget for the armed services. The model represents the executive and members of the legislature as having preferences over the policy dimension. The president and each legislator are assumed to have an *ideal policy* that they prefer over all others and to prefer policies closer to their ideal to those further away. The model also distinguishes a particular policy Q, called the status quo. This represents the policy that will remain in effect if no action is taken. We label the executive's ideal policy E and the legislative median's ideal L.

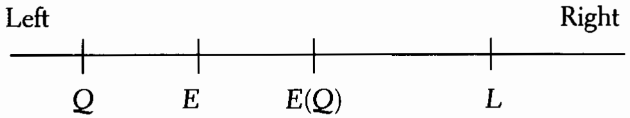
Figure 1 | Executive, Legislature, and Status Quo

2. For simplicity, this analysis ignores the complication of a veto override; it can be easily incorporated (see Brady and Volden 1997; Cameron 2000; Krehbiel 1998).

Consider the scenario in figure 1, where the status quo Q is on the left. The executive's ideal policy, E, is a bit more to the right, while the median legislator prefers policy L considerably more to the right.

If the legislature were free to chose policy without considering the president, it would introduce legislation and, given the open rule, produce a policy L on the right. Yet legislators must take the president's veto into account.[3] To see the efects of the veto power, consider the set of policies that the executive prefers to the status quo, namely, all policies between Q and $E(Q)$ (see figure 2).

Figure 2 | The Effect of a Presidential Veto



If the legislature presents the president with legislation to the right of $E(Q)$, the president will exercise his veto, yielding the status quo. Although the congressional median prefers all policies to the right of Q over Q, only those policies preferred to Q by the president will become legislation. The best the median in Congress can do is pass legislation equal to $E(Q)$. In this configuration of preferences, the presidential veto constrains congressional policymaking.

Of course, the executive veto is not always constraining. This can easily be seen by supposing that the positions of the legislative median and the executive were reversed in the figure. Following Kiewiet and McCubbins (1987, 1991), this discussion shows that the president's veto power is asymmetric. For example, if the above policy issue represents budgetary issues, the analysis shows that the president's veto constrains budgetary decisionmaking only when the president wants a lower budget than Congress. When the president wants more than Congress, the veto is not credible, so Congress can ignore the president's interests.[4]

STRONG PRESIDENTIAL SYSTEMS IN LATIN AMERICA

Among modern legislatures, the U.S. Congress is unusually powerful. In contrast, the strong presidential systems of Latin America typically grant the president greater powers over the legislature than in the United States. As Londregan (2000) suggests for Chile, these systems can be modeled as

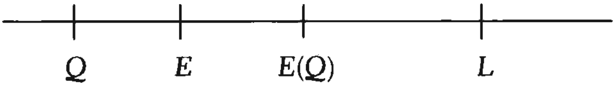
3. Cameron (2000) provides the best analysis of presidential veto power and its effects on the legislative process. See also Kiewiet and McCubbins (1987), Matthews (1989), and McCarty and Poole (1995).

4. The basic model abstracts from a range of important legislative, executive, and bureaucratic institutions that may be incorporated into the model.

granting the president the power to present the legislature with a take-it-or-leave-it choice over policy.[5]

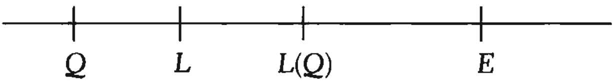
To emphasize the analysis's comparative dimension, I use the same policy configurations of preferences and status quo as used in the analysis of the United States. In the first scenario (figure 3), the president's ability to present the legislature with a take-it-or-leave-it choice allows the president to obtain his ideal policy *E*. Although the legislature might threaten not to pass the president proposal of *E* unless the president moves his proposal closer to *L*, this threat is not credible. Vetoing *E* results in the status quo, which makes a majority in the legislature worse off. Thus, in this political environment, the president can force the legislature to accept his ideal policy. In contrast, the U.S. Congress's greater ability to amend legislation allows it to force the president to accept *E*(*Q*) instead of *E*.

Figure 3 | Policy Choice in Chile



In the second scenario (figure 4), the Chilean president must consider the legislature's veto. To succeed, the president's proposal must make the median legislature better off than the status quo. Hence we consider *L*(*Q*), the policy such that *L* prefers all policies between *Q* and *L*(*Q*) to *Q*. The median legislator, and hence a majority in the legislature, will accept any of these policies. The president can take advantage of his take-it-or-leave-it powers by presenting the legislature with the policy among this set which he most prefers. In this scenario, the best the executive can do is *L*(*Q*).

Figure 4 | The Effect of Legislative Veto in Chile



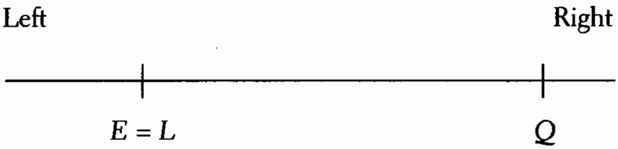
5. Among the many powers listed by Londregan for the case of Chile are: (1) the president can impose time limits in which constitutional legislation must be passed or rejected; (2) constitutional provisions "contain powerful veto provisions that allow the president to have the last word in the legislative debate by introducing amendments along with his or her veto, amendments which must be voted up or down without further change by the Congress"; and (3) as with most Latin American legislatures, Chile's is understaffed and lacks infrastructure to help it fight the president (2000, 66).

Here too, the rules advantage Latin American presidents over the U.S. president. Whereas the U.S. Congress can force the president to accept L, the weaker powers of the Chilean congress allow the president to obtain $L(Q)$ instead of L.

The Chilean constitution contains some critical features designed by the dictator Augusto Pinochet and his supporters to limit the policy flexibility of the democratic governments after the return to democracy. One of the most important of these features is the creation of a set of institutional senators appointed by the president and various enclaves. Pinochet's preferred party could not gain an electoral majority in the presidential election or a majority in the Chamber of Deputies or the Senate. The creation of a set of institutional senators, initially appointed by Pinochet and his faction, allowed his party to maintain a majority in the Senate for many years. This constitution-induced Senate majority thus institutionalized divided government in Chile, and along with it a veto for the opposition on the right against the center-left coalition, the *Concertación*, which typically controls the presidency and the lower chamber.

To show the effect of this constitutional detail on Chilean policy choice, we derive a comparative statics result, showing executive-legislative choice with and without this rule. First, consider the legislature without the institutional senators, which I will call unconstrained democracy. These rules allow the *Concertación* to control the presidency and the congress. Although the election is likely to produce some differences among the medians in the two chambers and the president, these will all be members of the *Concertación*, so for simplicity I ignore these (relatively small) differences. The rules of unconstrained democracy yield a situation with the status quo on the right created under the Pinochet regime and both chambers of the legislature's favored policy located at L and the president's favored policy located at C on the left:

Figure 5 | Postauthoritarian Political Environment In Chile without Institutional Senators



Given congress's legislative powers, the president will introduce legislation of his ideal policy E, and the congress will accept this (see figure 5). Of course, this degree of policy change is a disaster from the perspective of Pinochet and his supporters because it grants the new government the ability to alter the fundamental issues Pinochet and his supporters sought to protect, such as the economic system and the absence of sanctions for human rights violations under the dictatorship.

Institutional senators affect Chilean policy choice by affording the opposition a veto over policy. As in figure 5, the president's preferred policy and the median preferred policy C in the Chamber of Deputies are located on the left. The institutional senators imply that the median preferred policy *S* in the Senate is held by a member of the opposition party, who prefers to maintain the status quo (see figure 6).

Figure 6 | Postauthoritarian Political Environment in Chile with Institutional Senators

Because the constitution creates divided government, the opposition on the right holds the power to veto major changes. Institutional senators produce a major effect on public policymaking. As figure 6 suggests, the degree of policy movement available to the ruling *Concertación* depends on the location of the Senate median. Some issues, such as human rights, appear just like figure 6, allowing the ruling coalition no flexibility. On other issues, such as labor and education, some institutional senators have more moderate preferences. This implies a median senator having a preferred policy between *Q* and *P*, allowing the ruling coalition to change policy.[6]

PARLIAMENTARY SYSTEMS

Space constrains my ability to analyze parliamentary systems at the same level of detail using the same framework. Parliamentary systems without an independent executive produce different policy outcomes than presidential systems. Because the prime minister is an agent of the majority party or ruling coalition, her preferences tend to be closely aligned with the party median. These settings can be analyzed using the same framework, for example, by assuming in a Westminster system that the prime minister's preferences correspond to the party median's position.

6. Londregan's theory and empirical analysis allow us to guage Pinochet's success in rigging the system to preserve the status quo after the period of democratization. The creation of the institutional senators greatly constrains the democratic government's ability to alter the status quo. On policies with a consensus among members of the opposition on the right, the institutional senators allow the Senate to preserve the status quo.

THE ROLE OF THE COURTS: EXPANDING CIVIL RIGHTS

The above discussion left policy abstract. To show the power of these models to yield new and surprising conclusions, I turn to the evolution of an important policy area in the United States, expanding the meaning of civil rights legislation.

The landmark Civil Rights Act of 1964 represented a dramatic departure in the status quo, forcing southern states to end their system of apartheid suppressing African Americans. Beginning in the early 1970s, a series of court cases expanded the meaning of this act.[7] In brief, the Civil Rights Act was an antidiscrimination law, requiring equal opportunity for all individuals regardless of race, creed, or gender. Several critical court decisions in the 1970s expanded the meaning of the act to include a degree of affirmative action.

A major puzzle is why a conservative Supreme Court under Chief Justice Warren Burger expanded civil rights. Surely the conservative majority on the Court preferred the status quo to this outcome.

Eskridge (1992b) suggests that the answer to this puzzle lies in the interaction of Congress and the courts. On constitutional issues, the Supreme Court is just that—supreme. To overrule a Supreme Court constitutional decision, the country must amend the Constitution. As amendments are difficult to pass, the Supreme Court controls constitutional interpretation. In other areas, the Supreme Court's authority is less supreme. On statutory decisions—decisions to interpret the meaning of legislation—the Court does not have the last word: elected officials can overturn a judicial decision interpreting legislation by passing new legislation.

Eskridge argues that the conservative Court acted strategically, forestalling an even larger change in the scope of the law through legislation. His argument draws on an additional piece of theory from the pivotal politics model (Brady and Volden 1997; Krehbiel 1998), the "filibuster pivot." Although based on a seemingly small detail from the Senate, this institution proves critical for legislation in the United States. The Senate allows a minority of senators to defeat a bill by "filibustering," continuing the debate to prevent a measure from coming up for a vote. The Senate can end a filibuster only by a supramajority vote.

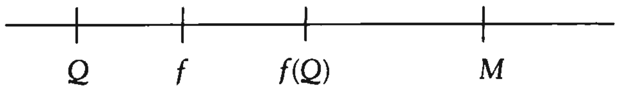
To pass the 1964 legislation required defeating a filibuster by southern Democrats. To do so required obtaining the support of two-thirds of the Senate. The effect of the filibuster is to require that legislation not just make a simple majority better off than the status quo, but two-thirds of the Senate.[8]

7. Notably, *Duke Power* (1971) and *United States Steelworkers v. Weber* (1979).

8. The rule has since been altered to require sixty of one hundred senators instead of sixty-seven.

The following policy setting reveals the effect of the filibuster:

Figure 7 | The Filibuster Pivot



In figure 7, *Q* is the status quo policy, *f* is the ideal policy of the filibuster pivot, and *M* is the median legislator's ideal policy. The filibuster pivot prefers all policies between *Q* and *f*(*Q*) to *Q*. Any policy outside of the range makes the pivot worse off. If the Senate tries to pass the median legislator's preferred policy *M*, senators on the left will filibuster, and the majority favoring *M* will not be able to defeat the filibuster.

The main result of the pivotal politics model in this context is that the majority can move policy from *Q* toward *M* as far as *f*(*Q*) but no further. Moreover, *f*(*Q*) is the equilibrium policy choice that will result under an open rule amendment process. Much of the drama in the passage of the 1964 act concerned the parliamentary maneuvers to defeat the filibuster.[9]

To understand the transformation of civil rights by a conservative Supreme Court in the 1970s, consider the following policy setting:

Figure 8 | Civil Rights Policy

The set of policy alternatives represents the degree of federal support for civil rights, *J* represents the ideal policy of the conservative Supreme Court majority, A represents the policy enacted by the 1964 act, *f* is the ideal policy of the filibuster pivot (a conservative Republican), and *M* is the median senator's ideal policy (see figure 8). As before, the filibuster pivot prefers all policies between A and f(A).

The critical feature of the new political environment of the 1970s is that the median in Congress was far more liberal than the median in the 1964 Congress that passed the Civil Rights Act. Eskridge argues that the more liberal Congress would undoubtedly pass new civil rights legislation moving policy from A to *f*(A).

In this setting, the Supreme Court acted first to preserve as much of the status quo as possible. By acting first, the Supreme Court moved policy

9. See, e.g., Eskridge and Frickey (1988), Graham (1990), and Whalen and Whalen (1985). Rodriguez and Weingast (1995) provide a model and evidence from the perspective developed here.

from A to *f*. This move precluded any further move by Congress: Because any policy change from *f* toward *M* would make the filibuster pivot worse off, no legislation was possible.

This model has several implications. First, it shows the power of the model in specific policy settings to give new answers to important political puzzles.[10] Second, it shows the strategic role of the courts in the United States policymaking process.

VARIANTS

Rational-choice models studying the effects of institutions represent a light industry in political science: there are simply too many variants to mention. Among the topics receiving considerable attention are:

1. Models of cabinet formation and parliamentary systems (see, e.g., Austen-Smith and Banks 1988; Baron 1993; Laver and Shepsle 1996; Moser 2000; Tsebelis 1994)
2. The recent gridlock/pivotal politics model in U.S. politics, that looks more closely at legislative details (Brady and Volden 1997; Krehbiel 1998)
3. The delegation of policymaking authority to the bureaucracy (Fiorina 1981; Epstein and O'Halloran 1999; McCubbins and Schwartz 1984; Moe 1989; Weingast and Moran 1983)
4. Elections (see, e.g., Bawn 1993; Cox 1999)
5. The role of information and uncertainty in politics (see, e.g., Gilligan and Krehbiel 1990; Austen-Smith and Riker 1987; Austen-Smith and Banks 1996; Krehbiel 1991)
6. Constitutional choice (Colomer 1995, ch 6; Hardin 1989; Przeworski 1991; and Riker 1986, chs. 4, 8).

▪ | Why Do Institutions Exist?

The previous discussion took institutions as given and studied their effects. This is by far the dominant mode of analysis in all three brands of institutional analysis. Works of this type, however, beg the question of why institutions are necessary at all. Why can't a society or group do without them?

10. Rational-choice theorists have applied analyses of this type in hundreds of contexts. In addition to the cites for the courts in no. 7, see Brady and Volden (1997, ch. 2), on the minimum wage; Ferejohn and Shipan (1989) on telecommunications policy; Riker (1986, ch. 11), on federal aid to education; Weingast and Moran (1983) on regulation by the Federal Trade Commission; and Weingast (1984) on the Securities Exchange Commission.

The purpose of this section and the next is to study a range of rational-choice models that explain why institutions exist and why they take the specific form they do. In brief, the answer is parties often need institutions to help capture gains from cooperation. In the absence of institutions, individuals often face a social dilemma, that is, a situation where their behavior makes all worse off. The prisoner's dilemma is the classic social dilemma, though there are many others (see, e.g., Elster 1989b). In this section, I use a variant of the prisoner's dilemma derived from Milgrom, North, and Weingast's (1990) model of the "Law Merchant." This model demonstrates why institutions are necessary for human interaction in complex societies.

Consider a group of individuals who face gains from cooperation. These gains may be economic gains from exchange or social gains from cooperation, say, to produce social peace and prosperity. The problem is that members of the society face various types of incentive problems where many or all individuals have short-term temptations not to cooperate. The prisoner's dilemma and coordination dilemmas illustrate this problem. Moreover many individuals are vulnerable to the action of others. Prior to the rise of the nation states, farmers were vulnerable to roving bandits who would expropriate the fruits of their labors. Similarly, once firms have made significant sunk investments, say in a water delivery system, citizens are tempted to force them to provide the water at unremunerative prices. Finally, individuals may fail to cooperate because they cannot agree on a distribution of the gains from cooperation.

Appropriately configured institutions restructure incentives so that individuals have an incentive to cooperate. Moreover, because incentive problems differ greatly across environments, the types of institutions necessary to mitigate these problems also varies. The essence of institutions is to enforce mutually beneficial exchange and cooperation. Oliver Williamson summarizes this logic:

> Transactions that are subject to ex post opportunism will benefit if appropriate safeguards can be devised ex ante. Rather than reply to opportunism in kind, therefore, the wise [bargaining party] is one who seeks both to give and receive "credible commitments." Incentives may be realigned, and/or superior governance structures within which to organize transactions may be devised (1985, 48–49).

Studying the limitations of the familiar repeated prisoner's dilemma reveals the logic of this approach to institutions. The well-known "folk theorem" in game theory, popularized in political science in the context of the prisoner's dilemma by Axelrod (1984) and Taylor (1976), shows that although players have a short-run temptation to cheat, they have long-run incentives to cooperate.

Consider the following prisoner's dilemma.

	C	D
C	1, 1	–1, 2
D	2, –1	0, 0

Clearly, both players are better off cooperating (C) than defecting (D), since each receives 1 from mutual cooperation and 0 from mutual defection. Yet each player is better off defecting: if the row player chooses to cooperate, column receives 1 from cooperation and 2 from defection; if row defects, column receives –1 from cooperation and 0 from defection. If both players follow their short-run incentives to defect, they are both worse off, since mutual defection implies the absence of cooperation.

A different result occurs, however, if the players interact over time. When this game is repeated, and assuming that the players do not discount the future too heavily, they can sustain cooperation. In repeated play, players' choices today reflect not only their payoffs today but how they anticipate that their opponents will react to their behavior. Typically players use trigger strategies, such as, "I'll cooperate with the other player as long as she cooperates, but if ever she defects, I will follow this defection in kind."

An opponent's trigger strategy alters a player's incentives. By tying behavior today with (negative) consequences tomorrow, trigger strategies alter the trade-off between defection and cooperation. An opponent's trigger strategy presents a player with the following choice: cooperate today, which yields cooperation in this and every future period. Defect today, which yields a short-term gain but which causes mutual defection in all future periods.

This analysis yields the following payoffs:

Cooperating forever yields payoffs that are a series of 1s, discounted by δ for each period:

$$1 + \delta 1 + \delta^2 1 + \delta^3 1 + \ldots = 1/(1 - \delta)$$

Defecting today and facing mutual defection thereafter yields:

$$2 + \delta 0 + \delta^2 0 + \delta^3 0 + \ldots = 2$$

For δ sufficiently large, each player will prefer cooperation. The calculation above implies that as long as $\delta > \frac{1}{2}$, the players will cooperate.

This discussion illustrates an important feature of rational-choice institutionalism: cooperation must be *self-enforcing* in the sense that all players in society have incentives to cooperate.

Several stringent assumptions prevent this model from being a general model of social interaction. Its conclusions are wildly optimistic, implying that it cannot explain a range of common breakdowns in cooperation, such as wars, ethnic conflict, government and private opportunism, and other systematic failures to capture gains from cooperation.

The model's flaws derive from its simplicity. Two assumptions implicit in the standard prisoner's dilemma model, both concerning defection, limit its wide applicability to the world. Everything is public in the prisoner's dilemma. First, it is obvious what actions constitute defection; there is a profound absence of moral ambiguity and an inexplicable degree of social consensus about how to cooperate. Second, defection is observable in the sense that anyone who defects is immediately known.

As Greif, Milgrom, and Weingast (1994) and Milgrom, North, and Weingast (1990) show, the failure of either assumption plagues the logic of cooperation described above. First, consider moral ambiguity or a lack of consensus about what constitutes cooperation. In contrast to the prisoner's dilemma, people often have fundamental disagreements about what is best for society. For two reasons, multiple ways to cooperate hinder the ability of a society to cooperate. First, individuals must somehow coordinate on the particular way to cooperate; disagreement often leads to cooperation failure (see Weingast 1997). Second, people often disagree about the best way to cooperate in part because of distributional effects. Some means of organizing society benefit one set of individuals at the expense of others, while another means of cooperation yields the opposite result. In this setting, cooperation may fail in part because the parties fight about which mode of cooperation to implement.

The second assumption causes cooperation to fail in a different way. Consider a large group of actors—whether individuals in a society or nations in the world system. In each round of play, the actors are paired and they interact through a prisoner's dilemma. Yet they do not always interact as pairs sufficiently frequently to sustain cooperation. Axelrod and Keohane (1985) show that in this setting cooperation can nonetheless be sustained if the players use generalized trigger strategies in which all players punish another player's defection, regardless of whether they were the target of that defection. Axelrod and Keohane set their analysis in the international system, where different states must decide whether to honor bilateral treaties with one another. As the players interact, each has an incentive to honor the treaty when defection is punished by the entire community.

Sustaining cooperation in this environment clearly depends on the assumption that defection is observable, that is, known by all members of the community, not just the target of the defection. Yet this rules out a host of problems common to bilateral relationships. Often the interaction in a bilateral relationship is private. Outsiders to the relationship have no way of knowing what really took place. Instead, they typically observe a dispute in which both players present internally consistent accounts of how the other defected.

The following example reveals the problems raised by the lack of observable defection. In the 1970s, the United States claimed that Japanese firms were unfairly dumping television sets on the U.S. market, selling them below costs. The Japanese argued that they were not dumping. Instead, they could sell their televisions at costs below those of the U.S. manufacturers because Japanese firms were more efficient producers and thus had much lower costs. Both claims are plausible; yet both cannot be simultaneously true. So who is right? Were the Japanese dumping? Or was the United States opportunistically claiming the Japanese were dumping?

The inability of outsiders to tell what really happened hinders their ability to punish defection. When community enforcement is required to police defection, some defections may go unpunished because the community cannot verify whether a defection in fact took place. This in turn gives rise to a type of opportunism: if some defections cannot be observed, opportunistic players can masquerade their subterfuge qua defection in a plausible rationale that the other defected. Put simply, in the face of these two problems, cooperation cannot be sustained.

The failure of cooperation in the face of uncertainty about who defected, in turn, provides the rationale for institutions. Consider the problem with observing defection. The law merchant model shows how institutions can help resolve problems that arise when defection is not observable (Milgrom, North, and Weingast 1990). The context is the rise of trade during medieval Europe prior to the development of the modern state with a formal legal system extending over its territory. At this time, merchants found that resolving disputes was a major problem. Because defection was hard to detect, sustaining faithful cooperation and exchange was difficult.

Over time, a system of private law judges emerged, whereby some merchants specialized in resolving merchants' disputes and became known as law merchants. A body of precedents and rules evolved for resolving disputes. These principles enabled the law merchant to resolve a range of problems among merchants.

But how did a legal system without appeal to sanctions by the state work? Put another way, why did merchants abide by the law merchant's judgements?

The answer is that the law merchant as an institution complemented the incentives of repeat play to make cooperation self-enforcing where decentralized interaction (as described above) led to failure of cooperation. By investigating disputes and declaring which merchants had defected, the law merchant made cooperation self-enforcing where only defection could be sustained without this institution. The law merchant institution did this in two ways. First, the law merchant's ruling provided an observable signal to the community announcing whether cooperation or defection had taken place. By publicizing defection, the law merchant's judgment could be used to trigger punishment.

Second, this system had two cost advantages over punishment through trigger strategies. First, in a decentralized system envisioned by Axelrod and Keohane, every player must know the full history of every other player. In contrast, the law merchant system requires only that the players find out which players have judgments against them by the law merchant. Second, trigger strategies require that, in the future, other merchants forgo the benefits of trade with the errant merchant. Creating law, even without the force of the state, provided another alternative: as part of the a judgment against a merchant, the latter was asked to make the plaintiff whole, that is, to pay a fine. Of course, the absence of a nation-state implied that the law merchant could not put the merchant in jail or force him to pay a fine. Yet the law merchant system provided an incentive for the errant merchant to make good on judgments against him: by making the plaintiff whole, the errant merchant could prevent the community from initiating its trigger strategy against him. Because outstanding judgments implied punishment by the community, merchants had incentives to make good on their fines.

The role of institutions in the law merchant system complements the incentives of repeat play. Because defection is unobserved, repeat play alone à la the prisoners' dilemma without institutions cannot sustain cooperation. Adding the law merchant to the system, however, changes the consequences of defection and helps the community police defection. Institutions are thus critical to creating the self-enforcing conditions for mutual cooperation.

This latter conclusion is a general result of rational-choice institutionalism. Institutions arise in part to help create the conditions for self-enforcing cooperation in an environment where there are gains from cooperation but also incentive problems that hinder a community's ability to maintain cooperation.

This model applies to a range of problems. The U.S.-Japan dumping example suggests an explanation for one purpose of the General Agreement on Tariffs and Trade and its successor organization, the World Trade Organization, which help govern international trade. These organizations have no independent force of law in the sense that they have no means of coercing member states to obey their rules. Yet they do have the means of investigating disputes among members and coordinating punishments when judgments are made against particular countries. The law merchant approach provides a model of how this works; that is, how the international institutions help sustain self-enforcing trade agreements even when these institutions have no independent ability to punish its member states.

Calvert (1995a) applies this approach to the study of Congress. A range of possibilities exist. One is that this model helps explain one of the roles played by the congressional leadership: they help police exchange among members. In this setting, many exchanges are bilateral or among small groups and are negotiated in private. They are thus not readily visible to others. As with the law merchant model, outsiders to the exchange cannot easily observe defection; instead, they observe a dispute, typically accompanied by two plausible but contradictory accounts in which each party blames the other. Outsiders have no way of knowing what happened. Because other members of Congress have strong incentives to police defection, they need institutions to help resolve such disputes. The leadership has long been recognized for helping put together deals.

Endogenous Institutions

Having suggested why institutions exist, I now turn to a more in-depth investigation of endogenous aspects of institutions: when will institutions matter and why do they take particular forms instead of others? An important concept throughout this study is credible commitments, self-enforcing incentives provided by institutions for relevant actors to behave in a certain manner.

AN ENDOGENOUS MODEL OF THE INDEPENDENT JUDICIARY

Although nearly all countries have judicial systems, these vary enormously in their independence and authority. In some countries, courts are an independent branch of government capable of overruling decisions by government officials, while in others they are subservient to politics. What determines the degree of independence of the courts? Clearly many factors matter (see, e.g., Shapiro 1981). In this subsection, I review how rational-choice theorists address this question for democracies, emphasizing its comparative implications.

The spatial models of the separation of powers in the previous section took the powers of the various institutional actors as given. This subsection adds a court to address the question of judicial independence. Clearly, the power of the judiciary varies considerably among systems. Although political scientists have characterized these differences well, what they havedone less well is explain them. There is a tendency to assume that the independence of the judiciary is a characteristic inherent in the judiciary itself, such as the norms of judges. In this section we show instead that the power of the judiciary depends critically on the relationship between the judiciary and the other branches of government.

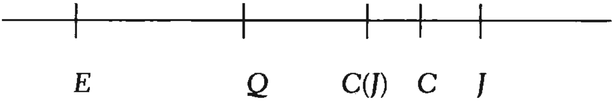
Consider a presidential system with a legislature independent of the president. Courts nominally have the power to interpret legislation, that is, the power to alter the legislation's meaning and hence policy. In what follows, I draw on a range of now standard models in the application known in the legal literature as positive political theory and the law.[11] The approach begins with the one-dimension spatial model and the policy preferences of three actors, the executive's (*E*), the Congress's (*C*), and the courts' (*J*), with the status quo policy (*Q*). Consider a typical political configuration (see figure 9).

Figure 9 | The Role of Courts in Interpreting Statutes

Notice that every policy between *E* and *C* is legislative equilibrium: Any bill preferred by one makes the other worse off, so no legislation can pass. But these status quos are not necessarily equilibria when we allow the judiciary to interpret the meaning of the law. When asked to judge the law and given their preferences and the latitude afforded them by virtue of their role as interpreters, the courts will move policy from *Q* to their ideal point *J*.

Of course, the courts' ability to influence legislation depends on the political configuration. For example, suppose that the court's ideal is outside the interval between *E* and *C* (to the right of *C* in figure 10).

Figure 10 | Constraints on an Extremist Court

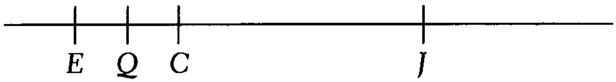


If the court attempts to implement its ideal policy *J*, it will fail. Because the policy *J* is outside the range of *E* and *C*, both Congress and the president are better off moving policy back between their ideals, specifically, inside the region between *C*(*J*) and *C*.

These examples illustrate a general result: in a separation of powers system, the range of discretion and hence independence afforded the courts is a function of the differences between the elected branches. The narrower the range of policies between the branches, the lower judicial discretion. This can be seen using figure 11, a variant of figure 10. In this

11. Marks (1988) provided the pioneering work. See also Epstein and Knight (1999), Eskridge and Ferejohn (1992), Eskridge (1992), Farber and Frickey (1992), Levy and Spiller (1994), McCubbins, Noll, and Weingast (1987, 1989), and Cohen and Spitzer (1994).

Figure 11 | Courts Facing Elected Officials with Very Different Preferences



political setting, *J* holds the same relation to *E* as in the first figure, but *C* is much closer to *E*, perhaps representing united government in which the same party holds the presidency and a majority in Congress. If the courts attempt to implement their ideal policy *J*, they will fail. Both the executive and Congress prefer all points between their ideal policies to *J*. The best the court can do is to implement policy C. This approach shows how judicial independence depends on the political environment.

This perspective yields a range of comparative implications about the independent authority of courts. In political environments where the legislature is subordinate to the president, the latitude of courts is small, so they cannot exert much political independence or authority. In the strong presidential systems of Latin America, for example Argentina or Mexico, the courts typically exercise relatively little independent authority. Yet, as Chavez (2000) shows, the perspective explains the vicissitudes of judicial independence in Argentina since democratization. When government is divided (at least one branch held by the opposition), courts have more latitude. Thus, the courts were relatively independent during the Alfonsin presidency in the 1980s, when the opposite party held one chamber of the legislature. In contrast, the dramatic pliancy of the courts through most of President Carlos Saúl Menem's term in the early 1990s reflected his party's hold of united government. The emergence of a more independent judiciary at the end of Menem's term occurred when he too faced divided government.[12]

This perspective also explains aspects of Chile's judicial independence (Chavez 2000; Levy and Spiller 1994; Londregan 2000). In the postauthoritarian era, Chile has been characterized by both a strong presidency and divided government. As noted above, the system is rigged so that the opposition holds a working majority in the Senate through the institutional senators. As the model predicts, Chile also has an independent judiciary (also initially rigged to a degree to favor Pinochet's interests).

Barros (2000) describes the emergence of a degree of judicial independence during Chile's authoritarian period. Although the range of judicial independence was relatively small in comparison to that in the democratic era, Barros shows that it was nonnegligible after 1980 and grew

12. Chavez (2000) also shows that this logic applies to judicial independence across Argentina's provinces.

over time. Judicial independence arose in part because the branches of the armed services had no mechanism for coordinating their activities. This problem emerged in the first days of the coup, when two different branches issued conflicting curfew decrees. In the 1980s, the courts began to rule on constitutional issues. Although the courts' latitude was relatively small, the regime's preference for a referee allowed the courts to emerge as an interpreter.

The history of the judicial independence in the United States also fits the model. Consider first the partisan era of the first 150 years of the republic. Prior to World War II, the typical pattern was for one party to dominate electoral institutions and the courts: the Federalists prior to 1800, the Jeffersonian and Jacksonian Democrats until 1860; the Republicans until 1930; and the Democrats until the end of World War II. For example, in the seventy years between 1860 and 1930, the Republicans held united government in half the Congresses, while the Democrats held united government in only four.[13] Except for some crucial appointments during the Wilson presidency, the Republicans dominated appointments to the courts during this era.

This pattern has not held during the post–World War II era, in which divided government and swings between the parties have been common. Although Franklin Roosevelt managed to create a Democratic Supreme Court, neither party has dominated appointments to the courts since World War II.

The model explains the varying degree of judicial independence. During the partisan era, courts were instruments of the ruling party and exhibited relatively little independence on statutory issues. In contrast, the courts have been most independent since World War II, having become a major interpreter of legislation and a champion for a host of new rights.

Since World War II, swings between the two parties have been the norm. Although each party has held united government for short periods (e.g., the Democrats for 1961–1968, 1977–1980, 1993–1994; the Republicans for 1953–1954); this is often followed by periods of divided government, particularly with a Republican president and a Democratic Congress or a divided Congress (1955–1960, 1969–1976, 1981–1992). No party has been able to dominate appointments to the Supreme Court.

The substantial political and constitutional differences between the Democrats and the Republicans has granted the Supreme Court considerable discretion. Per predictions of the model, the post–World War II era has had one of the most independent Supreme Courts, one involved in creating a host of new rights and intervening in many policy areas. Although every major Supreme Court decision has been met with elected of-

13. Although divided government was the norm between 1874 and 1896, the Democrats rarely held the Senate or the presidency. This allowed the Republicans to dominate judicial appointments even during these years.

ficials who have decried it, nearly all have also had elected officials who have at supported it. The absence of sustained united government implies the Court has had the political discretion to exercise its authority in many areas.

Similarly, the above perspective explains the two big confrontations between the courts and the elected branches, the first in the late 1860s during Reconstruction and the second in the 1930s during the New Deal. Although these cases differ significantly, in both partisan and ideological change afforded a confrontation between a determined set of elected officials and a judiciary of very different persuasion. In both cases, the Supreme Court suddenly faced a united government with very different preferences, much like that in figure 11. Without any political officials to protect them, they capitulated.

DEMOCRATIC CONSOLIDATION QUA SELF-ENFORCING DEMOCRACY

One of the most exciting topics in comparative politics concerns the question of long-term democratic stability or consolidation. Linz and Stepan (1996, 5) argue that consolidation occurs when democracy is the "only game in town," that is, when no significant groups advocate violating the constitutional rules. Further, constitutionally, citizens and politicians "become habituated to the fact that political conflict will be resolved according to the established norms and that violations of these norms are likely to be both ineffective and costly" (see also Burton, Gunther, and Higley 1992; Diamond 1999a).

Although scholars are clear on the definition of consolidation, there is no consensus on the conditions that produce it. In this section, I review recent rational-choice contributions to the topic, which provide a critical advance to our understanding of the conditions that produce democratic consolidation.[14]

I (Weingast forthcoming) begin by restating the definition of democratic consolidation as requiring three conditions: (1) No significant group out of power advocates the use of force to secede or capture the government. (2) Those in power respect the constitutional rules; in particular, they do not use their power to transgress the rights of their opponents. (3) Citizens are willing to defend the constitutional rules by withdrawing their support from leaders and groups who advocate violating the rules.

Stated in this manner, democratic consolidation is centrally concerned with incentives: consolidation requires that democracy is self-enforcing in the sense that all actors have incentives to adhere to the rules.

14. The discussion draws on the work of Colomer (1995), Fearon (2000), Przeworski (1991, 2001), and Weingast (1997, forthcoming).

Because consolidation is centrally concerned with incentives, rational choice theory provides a necessary input into understanding this question.

The relevance of this approach is demonstrated by considering what I call the "Przeworski moment" (Przeworski 1991, ch. 2): a party in power has just lost an election but retains power until the date of legal transition: Why would it ever give up power? Democratic consolidation clearly necessitates that this party do so. Because the Przeworski moment is a time when many democracies fail, it is not obvious why parties in power adhere to the rules. Przeworski provides the abstract answer: it must be in the interests of those in power to do so.

In what follows, I provide two principles relevant for this question and then summarize two results about the conditions that produce democratic consolidation. The first principle concerns the *rationality of fear* model (de Figueiredo and Weingast 1997). The concept of the rationality of fear is based on the premise that, when citizens or groups are threatened, they take steps to defend themselves. Suppose that a threat involves a certain probability of becoming a reality. The main result is that, the larger the magnitude of the stakes, the lower the probability triggering defensive action. For very large threats—for example, those concerning people's lives and livelihoods—the probability triggering defensive reaction can be quite low, such as one in ten.

The rationality of fear model implies discontinuous political change. When citizens' perceptions of a threat is just below the probability threshold triggering their reaction, they continue to honor the rules. Suppose that an event occurs that increases the probability of the threat; then the probability rises above the threshold. Per the model, the threatened citizens suddenly defect from the regime as a means of protecting themselves. This principle shows why situations involving high stakes can be so explosive: first, the probabilities triggering support for extraconstitutional action, such as violence, can be quite low; second, when the probability of the threat nears the threshold, small increases in the perception of the threat can trigger action.

The second principle follows from an observation of Przeworksi (1991, ch. 2): all successful constitutions limit the stakes of politics. The reason is easily seen in the context of the first principle. When constitutions limit the stakes of politics in ways valued by most citizens, the citizens are far less likely to resort to extraconstitutional means to defend themselves. Constitutional limits that protect what citizens value imply that they are less likely to experience threats that cause them to support defections from the regime.

This argument implies a selection effect: democracies that limit the stakes of political competition are more likely to survive. When democratically elected governments threaten what some citizens consider their fundamental rights, they support leaders who will defend them. Democracies with constitutions that place constraints on government valued by citizens are more likely to survive because they are less likely to threaten their citizens.[15]

This mechanism has dozens of applications. For example, consider Chile in the early 1970s. Many on the political right felt their economic rights threatened by Allende's government, leading them to support the military (see, e.g., Valenzuela 1978). Similarly, during the Second Republic in Spain, many agrarian landholders, the church, and industrialists felt threatened by the regime, leading them to support Francisco Franco (Agüero 1995; Alexander 2002). In the United States, large numbers of Southerners felt threatened by the newly elected Republicans in 1860 (see, e.g., Weingast 1998). In each case, a sufficiently large group of citizens supported extraconstitutional action that a civil war or coup ensued.

In short, the relevance of the rationality of fear model for democratic consolidation is this: constitutions that limit the stakes of power are more likely to survive.

I now suggest how these principles interact, and add three additional results about democratic consolidation. The first result concerns the particular institutions of democratic governance. It is remarkable that students of democratic consolidation typically ignore how the constitution works, in particular, the ways in which the institutions of governance interact with democratic stability.[16] A critical aspect of these institutions follows from the second principle above: constitutional institutions often help lower the stakes of politics by creating self-enforcing limits on politics or by creating procedural limits on democratic decision making that reduce the set of feasible choices of elected officials. Put another way, constitutions that impose widely accepted substantive and procedural limits on the government lower the stakes of competitive elections and thus make it more likely that democracy will survive.

Democratic stability in the United States in the early–nineteenth century illustrates this conclusion. From the beginning of the republic, Southerners worried about the safety of their "property" and their "peculiar institutions," that is, slavery. A necessary condition for them to have remained in the Union was that Northerners provided a credible commit-

15. This argument suggests an explanation for the selection effect finding of Przeworski et al. (2000), that richer democracies are far more likely to remain democratic than poorer ones. The reason is that sustaining a vibrant economy requires secure property rights, which in turn require limits on government. In other words, economies that grow satisfy the second principle above noted in Przeworski (1991), and therefore make democracies more likely to survive.

16. Students of pacts often note that they restrict government decision making (e.g., O'Donnell and Schmitter 1986) but rarely study how or why these work. Although some scholars have studied effects of different election mechanisms (e.g., Linz and Stepan 1996), this is separate from studying the effects of different institutions of governance and public choice. An important exception in this literature is Ordeshook's study (1996) of Russia.

ment to honor Southern rights in slaves (see Weingast 1998). Several mechanisms of the Constitution contributed to this. First, federalism helped devolve many issues, such as property rights and slavery, to the states, outside the purview of the national government. Second, the separation of powers of the national government also limited the scope of democratic policymaking at the national level. Beyond these institutions, however, Southerners required an additional limit. From nearly the beginning of the republic and made explicit in the Missouri Compromise of 1820, a critical aspect of U.S. democratic stability was sectional balance—the notion that both sections would have equal representation and hence veto power in the U.S. Senate. Veto power implied that Southerners could veto any attempt by Northerners to use the national government to attack slavery. Moreover, it is clear that this power was necessary. With some frequency, Northerners used their majority in the House of Representatives to pass several antislavery initiatives. Further, the coming of the Civil War is wrapped up with the demise of section balance in the 1850s (Weingast 1998).

The second result concerns pacts, agreements often between previously warring groups that modify or specify the rules of the political game. Pacts have long been recognized in the literature as critical to democratization and democratic consolidation (see, e.g., Burton, Gunther, and Higley 1992; O'Donnell and Schmitter 1986). The literature is far less successful at explaining why some pacts succeed while others fail. Rational-choice approaches provide an answer: successful pacts must be self-enforcing; that is, they must provide the parties to the pact with incentives to abide by the pact's provisions. How does this work?

Weingast (forthcoming) provides four conditions for self-enforcing pacts. First, the pact must create (or be imbedded in a context that has already created) a set of rules and rights, that is, substantive and procedural limits on the state. Second, the parties agreeing to the pact must believe that they are better off under the pact than without it. If this condition fails for one of the parties, that party will be better off without the pact, so the pact will fail. Third, each party agrees to change its behavior in exchange for the others' simultaneously doing so. Fourth, the parties to the pact must be willing to defend the parts of the pact benefiting the others against transgressions by political leaders. This occurs when each party anticipates that its rights will be defended by the others: that each party is better off under the agreement than not and that if ever one party fails to protect the rights of others, the others will fail to come to its rescue. Put another way, the pact becomes self-enforcing when all parties are better off under the pact and when all realize that unilateral defection from the pact implies that the others will also defect, destroying the pact.

Because there is no external authority to police the democratic rules, these rules must be self-enforcing. Rational-choice institutionalism thus provides important additions to the literature on democratic stability and consolidation.

RATIONAL-CHOICE MODELS OF ETHNIC CONFLICT

The study of ethnic conflict has become a light industry within political science and sociology. A range of approaches have emerged. Perhaps one of the best works on this topic is Horowitz 1985. Horowitz discusses "severely divided societies" in Asia, Africa, and the Caribbean, arguing that ethnic identity elsewhere is different. In the west, "there is an important overarching level of identity. . . . A survey of Switzerland that tapped levels of identity found that, in spite of ethnic differences, about half of all respondents identified themselves as Swiss" (1985, 17–18). This view takes the degree of ethnic conflict as exogenous.

Rational choice provides a different perspective, suggesting that the degree of ethnic and related forms of conflict is endogenous. In particular, this conflict depends in part on whether there exist political institutions that adequately protect groups from being taken advantage of. Adequate protection is modeled in terms of credible commitment (Fearon 1994c; Fearon and Laitin 1996; de Figueiredo and Weingast 1997). In societies where institutions provide credible commitments against ethnic conflict, ethnic identity becomes only one of many bases of group formation and interaction. Moreover, these institutions allow the emergence of many cross-cutting cleavages. In these societies, ethnicity is only one of many bases for the formation of individual identity. In contrast, in societies without adequate institutional protection against ethnic violence, coalitions become based largely or solely on ethnicity, as does identity. Horowitz's observation about the Swiss is thus not an exogenous difference that separates the west from the areas he studies but part of the phenomena to be explained.

In what follows, I draw on Fearon's model (1994c) to suggest the importance of credible commitments and the emergence of ethnic conflict in their absence. A fundamental puzzle about ethnic conflict concerns why it occurs given that it is so costly, and typically for both sides.[17] One answer is that ethnic conflict is simply irrational; the participants are sufficiently emotionally involved in mutual hatreds that their decisions are not rational. Most approaches now reject this view. Nonetheless, these approaches fail to answer this fundamental puzzle.

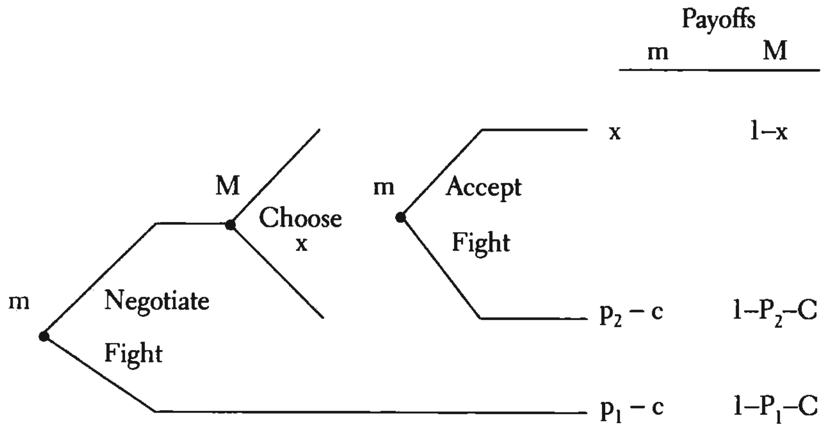
Fearon provides a creative new insight into this question. Consider a society with a majority group M and a minority group m. Facing the possibility of conflict, the two groups have the opportunity to fight or to bargain for a peaceful solution. Both groups have access to arms.

The minority moves first and must decide whether to fight now or to bargain with the majority and accept some accommodation (see figure 12).

17. I call this the *economic puzzle* about ethnic conflict in "Constructing Trust" (Weingast 1998); the *political puzzle* of ethnic conflict concerns how to explain the timing of the outbreak of ethnic conflict, whether among groups that have never experienced such conflict or ones that do so after long periods of peace.

A bargain is assumed to involve a division of the social surplus. For simplicity, the total surplus is normalized to 1, so that a decision about its division is simply a proportion x going to the minority and $1 - x$ retained by the majority. As part of the peace bargain, the minority agrees to disarm. If the minority decides to fight, then with probability p_1 the minority wins and captures all the surplus. With probability $1 - p_1$, it loses, and the majority retains all the social surplus. Fighting costs the minority and majority c and C respectively. Total social surplus when the groups cooperate is 1; and when they fight, $1 - (c + C)$.

Figure 12 | Ethnic Conflict as the Absence of Credible Commitment



If the minority accepts the bargain, it lays down its arms. The majority then decides whether to honor the bargain, and if not, what proportion x of the surplus to give to the minority. The minority must then decide whether to acquiesce and accept x or to fight. If it fights, the minority wins with probability p_2. As before the costs of fighting are c and C respectively. Because accepting the bargain involves the minority's laying down its arms, the probability of winning if the minority then fights is lower than if it decides to fight at the beginning; that is, $p_2 < p_1$.

Payoffs from this interaction are modeled as follows: If the minority decides to fight at the beginning, then with probability p_1, the minority wins the payoff of 1 and with probability $1 - p_1$, it loses and receives nothing. Regardless of whether it wins, the minority must pay the cost of fighting. Thus, the minority's expected payoffs are $p_1 * 1 + (1 - p_1) * 0 - c = p_1 - c$. Similarly, the majority's expected payoffs are: with probability $1 - p_1$, the majority wins the social surplus and with probability p_1, it loses and receives none of the surplus; in both cases it must pay the costs C of war. Thus, the majority's expected payoffs are $p_1 * 0 + (1 - p_1) * 1 - C = 1 - p_1 - C$.

If instead the minority decides to bargain with the majority, it receives an offer of x from the majority. If it decides to accept this, its payoff is x and the game ends. If it decides to fight, then as before, its payoffs are $p_2 * 1 + (1 - p_2) * 0 - c = p_2 - c$. For the majority, its payoffs are x if the minority accepts the majority's proposal. If instead, the minority rejects this proposal, then its payoffs are $p_2 * 0 + (1 - p_2) * 1 - C = 1 - p_2 - C$.

To solve this game for the payoffs and behavior, we work backward through the game tree, first considering the minority's decision assuming it at first decided to bargain with the majority. At this last decision node (in the upper right of the figure 12), the minority will acquiesce if and only if $x \geq p_2 - c$. We then move back a step to calculate the majority's decision about x. Because it wants to maximize its portion of the surplus, and knowing the minority's decision rule about whether to fight, the majority will choose x so that the minority just barely prefers not to fight; that is, it will set $x = p_2 - c$.

Knowing the majority's decision about the allocation of the surplus, the minority can make its decision at the first node (on the left of the figure 12). If it decides to bargain, then, regardless of what the majority promises it initially, once it lays down its arms, the majority will give it only $x = p_2 - c$. If it decides to fight now, it receives $p_1 - c$. Because $p_2 < p_1$, the minority prefers to fight now rather than to accept a bargain.

The real problem in this interaction is the inability for the majority to credibly commit to treating the minority honorably. Both sides would be better off if they could avoid fighting; fighting imposes (often quite large) costs on both. Thus, were they able to bargain credibly, they could save the costs $c + C$ of ethnic conflict. Yet once the minority acquiesces, the majority can take advantage of it. Knowing this, the minority prefers to fight rather than bargain. In the absence of institutions credibly protecting the minority, the majority cannot prevent itself from behaving opportunistically and taking advantage of the minority.

Fearon's analysis does not address the question of timing: if this mechanism was always present in Yugoslavia, why did Yugoslavia not self-destruct during the previous forty years? And if it was not present during the previous years, why did it appear in the late 1980s and early 1990s? The answer is that the institutions of the *ancien Yugoslav regime* protected against this in several ways: they created a balance of ethnic groups, with Croats and Slovenians balancing Serbs; they created institutions that decentralized much decision making in ways that both made the central state and granted each of these ethnic groups a set of resources under their control; and finally, Marshall Tito, as head of the regime, steadfastly punished any attempt to make ethnic appeals. These institutions helped construct the basis for trust among ethnic groups, making ethnic peace self-enforcing.[18]

18. Burg (1983) describes these institutions in greater detail, and Weingast (1998) provides an argument about why they were self-enforcing in "Constructing Trust".

The *ancien regime* broke down in the mid-1980s for several reasons. For example, the fall of the Soviet Union and changes in Europe gave the Slovenians and Croats hope they might defect and join Europe. Further, after Tito's death, Yugoslav leaders had no interest to—or could no longer—restrain ethnic appeals. Once Slovenia defected from Yugoslavia, these institutional changes set the stage for the mechanism identified by Fearon, which applied in two ways within Yugoslavia: first, with the larger Serbian group's potentially taking advantage of the Croats; and within Croatia, the Croatian majority's taking advantage of the Serb minority.

GREIF'S MODEL OF THE PODESTA IN THE RISE OF GENOA

Greif (1998, forthcoming) has studied the rise of medieval Italy in a variety of contexts. An especially interesting idea concerns the role of political institutions in the trading rivalry of Genoa and Pisa.[19] In the early years of the rivalry, Genoese merchants fell behind those from Pisa, in part because the two main clans spent too many resources fighting one another instead of cooperating in their competition with rival Pisa. Greif explains how, although Genoese merchants had a collective interest in cooperation to compete with their rivals, their incentives led them instead to fight one another.

After many years of falling behind Pisa, Genoese merchants adapted an institution from elsewhere in Northern Italy, called the Podesta, to suit their needs. The Podesta was a foreigner whom Genoa invited to be mayor. They endowed him with some resources but not so many that he could vanquish either clan.

The Genoese hoped that the Podesta would help solve their internal feuds so that they could focus their energy and resources on competing with Pisa. To this end, they charged the Podesta with the following behavior: If either clan attacked the other, the Podesta was to side with the defender. Although the resources granted the Podesta were small relative to those of the clans, they were big enough so that the Podesta plus one of the clans implied a large advantage over the other clan. (Because the two clans were relatively evenly matched, these conditions were not hard to meet.)

For our purposes, the principal question is how did the institution of the Podesta make cooperation among the clans self-enforcing? Why not side with one of the clans instead and, subjugate the other clan and divide its resources?

Greif's analysis answers this question by analyzing the incentives facing the Podesta. The Podesta's primary interest is to maximize personal power, influence, and wealth. Consider the Podesta's incentives to side with the defender when one clan attacks the other. If the Podesta joined a coalition with the attacking clan to subjugate the other, then nothing protects the

19. This subsection discusses only one aspect of Greif's especially rich approach.

Podesta once the other clan is subjugated. The succeeding clan will then have no use for the Podesta and will simply get rid of it. If, on the other hand, the Podesta sides with the defender, the coalition is likely to succeed and the Podesta will continue in his position of prestige and power. Following the charge is thus better for the Podesta than the alternatives.

Greif also discusses the larger implications of his model for the problem of the political foundations of economic development. Genoa's political institutions formed a necessary part of the successful economic development. Without an institution that helped Genoese merchants cooperate and focus their energy on competition with Pisan merchants, the Genoese merchants could not cooperate with each other. The institution of the Podesta allowed them to cooperate and, in the long run, gain the upper hand over Pisa.

CRITICAL JUNCTURES IN THE AMERICAN REVOLUTION

Paralleling the literature, this essay has thus far concentrated on rational-choice institutionalism, only occasionally referencing other perspectives. Although political science draws on several rich approaches to institutions, there has been far too little cross-fertilization among these approaches. In this subsection, I provide a rational-choice interpretation for a concept originating in historical institutionalism—the notion of critical junctures or points of dramatic, irreversible, and discontinuous change (Collier and Collier 1991; Pierson and Skocpol 2002). According to Collier and Collier, these are "major watersheds in political life . . . [which] establish certain directions of change and foreclose others in a way that shapes politics for years to come" (1991, 27).[20]

What mechanisms underlie a critical juncture? Why do they happen only sometimes, and why at a particular moment? I address these questions in the context of Rakove, Rutten, and Weingast's (2000) and Schofield's (1999) discussion of the American Revolution.[21]

Historians of the American Revolution emphasize the role of ideas (Bailyn 1967; Greene 1986; Rakove 1996; Reid 1995): the Americans and the British fought because of fundamental disagreements over the nature of the constitution, the structure of the empire, and the mechanisms protecting liberty in the colonies. As Wood suggests, "If the origin of the American Revolution lay not in the usual passions and interests of men, wherein did it lay? . . . Never before in history had a people achieved a 'revolution by reasoning' alone. . . . The Revolution was thus essentially intellectual and declaratory" (1966, 162–63).

20. Katznelson (2000) opens his paper with the critical juncture in which Eve hands Adam the apple.

21. Elsewhere, I provide a model of the mechanism underlying the critical juncture leading to the U.S. Civil War (Weingast 1998).

The Americans, drawing on the constitutional ideas surrounding the English Glorious Revolution of 1688, emphasized the role of precedent: practices enshrined for a generation or more attained constitutional status. With respect to the empire, they argued that it had a federal structure (though they didn't use that term): sovereignty was divided, with the British having control over the empirewide issues of security and trade and the Americans having control over their local affairs, including religious and social regulation, taxation, and the structure of economic and political rights. This system had evolved over the 100 years prior to the revolutionary crisis (1763–1776) and had thus, by the Americans' reasoning, attained constitutional status. This view implied that the British lacked authority to regulate or otherwise control domestic colonial affairs.

The British, in contrast, had begun to evolve different ideas of their constitution, which attained its mature form in the nineteenth century. They argued that the British Parliament had supreme authority in Britain and the empire, including the power to interpret and alter the constitution. The British also believed that the structure of empire was a matter of policy choice, not the constitution. After the Seven Years' War (1756–1763), the British began to exercise their powers, including the imposition of relatively modest domestic taxes on the American colonies.

For our purposes, the actors in American colonial politics can be divided three groups: at one end of the political spectrum stood the small group of radicals seeking immediate remedial action, in the middle was a large group of moderates, and at the other end of the spectrum stood the loyalists. The moderates were politically pivotal: if they remained loyal to the British, revolution had no hope; if they joined the radicals, a revolution would ensue.

The radicals reacted strongly to the new British policies, arguing that the British actions represented a major change in precedent that if unchallenged would allow the British to Intervene in all domestic affairs. For them, this signaled that the British were no longer a benign and remote presence. The radicals' rhetorical argument about precedent was clever: by suggesting this small tax could imply much larger changes, they argued that the British action threatened American liberty. If the British could impose domestic rulings here, they could do so for any policy, including religion, economic regulation, or slavery.

Initially most Americans disagreed with the radicals and did not believe that this minor tax signaled a major change in the British view toward America. Yet a decade later, the moderates had come to side with the radicals to support fighting the British. How did this occur? I argue that a critical juncture took place.

To model this, Rakove, Rutten, and Weingast (2000) draw on the rationality of fear concept discussed above (de Figueiredo and Weingast 1997). The new idea proposed by the radicals was that the British had become a malign presence bent on taking away American liberty. This meant the possibility of significant new taxes, limits on religious freedom, and significant new economic regulatory restrictions. In short, American lives and livelihoods were at stake. Initially most moderates thought the likelihood that the radicals were right was quite low. Rakove-Rutten-Weingast model this by assuming that the moderates had beliefs about the probability P that the radicals were right; implying that with probability $1 - P$, the radicals were wrong, that the British were not malign, and that the status quo would by and large prevail.

As noted above, the rationality of fear mechanism implies several results relevant for this context. First, there exists a threshold probability P^* such that, if American moderates believe that the likelihood of the new idea exceeds P^*, then they would change sides and fight the British. Second, the larger the stakes, the lower is P^*. The high stakes of the crisis implied that P^* was likely to be significantly closer to 0 than to 1.

Rakove-Rutten-Weingast model the initial situation, in which most Americans did not believe the radicals, by assuming that the moderates' belief of the probability P that the radicals were right was too low (i.e., $P < P^*$). The critical juncture occurred because over time events occurred that changed the moderates' beliefs. Explicit in the radicals' view of the world was that the British would significantly alter the structure of American liberty, which depended on the colonial assemblies. In every colony, these assemblies established the laws that protected American liberty.

At several points during the revolutionary controversy, the British decided to show a strong hand to the Americans. When New York refused to agree to quarter British troops in 1766, the British passed the New York Suspending Act, which rendered all acts of the New York Assembly void until it voted funds for troops. A few years later, in 1773, following the Boston merchants' protest of the new duties on tea—known as the Boston Tea Party—the British reacted even more strongly with what became known in America as the Intolerable Acts: they disbanded the Massachusetts Assembly, imposed marshal law, and closed the port of Boston.

To the surprise of most Americans, the British had acted just as the radicals predicted: they had taken away the heart of the institutions protecting American liberty—the colonial assembly. Regardless of British intentions, it became clear to the majority of Americans that the British were no longer a benign, remote presence. British behavior raised most Americans' beliefs about the probability P that the radicals were right. When these beliefs became larger than the critical threshold P^*, the moderates came to support the revolution.

This model has several implications. First, it shows how politics can exhibit discontinuous change: Americans supported the status quo as long as P remained lower than the critical threshold P^*. Once P rose sufficiently so that it was above the threshold, however, Americans suddenly reacted to defend themselves. Second, this model provides a mechanism for why this critical juncture—the point of no return—occurred. Per the rationality of fear model, Americans came to fear the British and acted to defend themselves. It shows how and why the political cleavage underlying the revolutionary crisis changed so as to support a revolution and forever alter American politics. Third, the model suggests the microfoundations of macropolitical, indeed revolutionary, change. Finally, this exercise reveals the complementarity of rational-choice institutionalism and historical institutionalism. It began with a concept devised by historical institutionalists but not by rational-choice institutionalists, and then showed how rational-choice models can incorporate this notion.

CONCLUSIONS

I close the two sections on why institutions are needed and endogenous institutions with four observations. First, the approach shows that, to survive, institutions must be self-enforcing in the sense that relevant actors have incentives to abide by them. Democratic consolidation requires that members of a society adhere to a list of rules: those in power must abide by the results of elections; those out of power must refrain from the use of force to seize the government. The law merchant rules are effective because merchants have an incentive to abide by his rulings. Only in some political environments do political officials have an incentive to adhere to judicial constraints. And to keep peace among the clans, the Podesta must have the appropriate incentives.

Second, many important political phenomena reflect the breakdown of institutions, such as the emergence of ethnic conflict following the destruction of the institutions that promoted trust, including incentives against ethnic conflict. Similarly, the American Revolution occurred when the institutions protecting American liberties changed following the Seven Years' War, and Americans had incentives to defect from the empire.

Third, the rational-choice approach emphasizes the mechanisms explaining why institutions have particular effects and why they survive. Extensive literatures have long existed on each of the endogenous institutions (or their breakdown) surveyed here—the law merchant, the independent judiciary, democratic stability, ethnic conflict, and the American Revolution. The comparative advantage of the rational-choice approach is to add a focus on the mechanisms that promote stability or breakdown.

Fourth, the approach demonstrates that a critical reason for institutions is to support cooperation. As is well known, in many environments repeat play alone is sufficient to support cooperation. Yet the law merchant model shows how in complex environments this is not true and that institutions help make cooperation self-enforcing.

Conclusions

Three approaches to institutions contribute to modern political science, and each makes valuable contributions (Hall and Taylor 1996). The purpose of this paper is to survey rational-choice institutionalism. The paper divides institutional analysis into two separate modes.

The first mode of institutional analysis takes institutions as given and studies their effects. This is by far the dominant form of institutional analysis in all three approaches. This paper surveyed two sets of analyses studying the effects of institutions. The first studied the interaction of the legislature and executive. This discussion illustrated a hallmark of rational-choice institutionalism: the demonstration that seemingly small differences in institutional details often have macropolitical effects. This application emphasized the institutional sources of political differences across types of democratic systems. The second application provided a new explanation for the judicial expansion of civil rights in the United States. The analysis revealed the strategic behavior of the courts, in this case, a conservative Supreme Court expanding civil rights legislation in an effort to forestall a more far reaching revision by the Congress.

The second mode of institutional analysis studies institutions as the endogenous variable, attempting to provide explanations for why particular institutions exist, evolve, and survive. Although the less developed of the two modes of rational-choice institutional analysis, it has larger power in the long run. This paper develops these ideas in two ways. First, it provides an analysis of why institutions exist. Developed in the context of the institutions underpinning the rise of trade in medieval Europe, this analysis showed that in some contexts, the incentives of repeated play are sufficient to enforce long-term social cooperation. The prisoners' dilemma is the best-known example. Yet in more complex environments, repeat play alone fails to provide for long-term cooperation. The analysis shows that, in some contexts, institutions evolve to alter incentives so that cooperative behavior becomes self-enforcing. Put another way, a fundamental aspect of institutions is that they provide the means for the enforcement of cooperation.

The medieval law merchant, for example, issued rulings in disputes among merchants prior to the rise of the nation-state and thus without any means to enforce its judgments. This institution nonetheless proved critical for maintaining cooperation and exchange among merchants because it helped identify defectors. The inability to ascertain who defected implies not only that defection goes unpunished but that opportunistic players exploit this inability. By providing a mechanism for publicizing defectors, the law merchants' rulings provided vital information to the community, enabling it to maintain cooperation.

Second, the essay turned to a series of applications of the approach that explain the evolution of particular types of institutions. The first application provided a segue from the discussion that took institutions as exogenous. That discussion studied the conditions for an independent judiciary capable of placing constraints on elected officials. It applied the conditions for an independent judiciary to a range of democratic systems. The second application raised the issue of democratic consolidation, that is, the conditions for long-term democratic stability. This discussion showed that incentives are fundamental to democratic consolidation and suggested a range of conditions providing for self-enforcing democracy. The third application concerned ethnic conflict. This discussion showed that a fundamental aspect of this type of conflict is the absence of credible commitments to protect the rights of minorities. The fourth application studied the institutions of cooperation in medieval Genoa.

I conclude this survey with three observations about the value of the rational-choice approach. First the discussion on the effect of institutions shows how rational-choice institutionalism helps integrate the study of American politics into the larger study of comparative politics. The research such as that reported in this essay takes an important step toward that goal for institutions in democratic societies.

Second, rational-choice theory provides a variety of mechanisms that afford predictions of discontinuous change. A large range of questions in political science involve sudden change—the emergence of ethnic conflict, wars, the transition to or the failure of democracy, revolutions, major policy swings within particular countries. The various mechanisms of discontinuous change thus hold promise for new insights into these phenomena. An interesting implication of this development is that rational-choice theory has begun to discuss macropolitical phenomena, events that were once largely the domain of historical institutionalism.

Finally, let me touch on a few issues at the frontiers of rational-choice institutionalism. First, endogenous emergence, choice, and survival of institutions are likely to be major topics of the next decade. Second, choice theorists of many stripes have begun an impressive set of investigations into the limits of rationality in combination with the means of extending the theory to cover these more general circumstances. These include Elster (1999) on emotions and Knight and North (1997) on cognitive science, not to mention one of the first provided by Kahneman and Tversky (1978). Third, rational-choice theorists have long included the study of uncertainty and incomplete information in their analysis. The far-reaching results in those areas of study suggest that we have yet to see the full implications and power of these tools. Finally, an important frontier for the discipline, increasingly the subject of research, is the integration of rational-choice institutionalism with other forms of institutional analysis, particularly historical institutionalism. The discussion here about critical junctures in the American Revolution suggests the plausibility of this goal.

Paul Pierson and Theda Skocpol

Historical Institutionalism in Contemporary Political Science

Like the character in Moliere's play who spoke prose all his life without knowing it, contemporary political scientists are familiar with leading examples of historical institutionalist research without necessarily realizing that they exemplify a coherent genre—much as do works in the other two major research approaches practiced in empirical political science, survey-based behavioralism and rational-choice modeling. Historical institutionalists analyze organizational and institutional configurations where others look at particular settings in isolation; and they pay attention to critical junctures and long-term processes where others look only at slices of time or short-term maneuvers. Researching important issues in this way, historical institutionalists make visible and understandable the overarching contexts and interacting processes that shape and reshape states, politics, and public policymaking.

Stephen Skowronek's *The Politics Presidents Make* (1993), for example, reveals recurrent cycles in the nature and success of presidential leadership throughout U.S. history. Another long-term study in American politics, John Mark Hansen's *Gaining Access: Congress and the Farm Lobby, 1919–1981* (1991), develops a model of interest group interaction with government and uses it to explain the emergence, persistence, and ultimate eclipse of the political influence of national farmers' associations. Ranging across nations as well as time, Peter A. Hall's 1986 book *Governing the Economy: The Politics of State Intervention in Britain and France* explains how institutions and organizations intersect to shape not just the policies of governments but also the strategies and alliances of interest groups and public intellectuals. Painting on even grander canvases, *Shaping the Political Arena: Critical Junctures, the Labor Movement, and Regime Dynamics in Latin America* by Ruth Berins Collier and David Collier (1991) and *Birth of the Leviathan: Building States and Regimes in Early Modern Europe* by Thomas Ertman (1997) explain regime dynamics and the varied formations of modern national states. In the study of international relations, *Activists Beyond Borders* by Margaret Keck and Kathryn

Sikkink (1998) analyzes the historical roots and contemporary proliferation of transnational advocacy networks, while Judith Goldstein's *Ideas, Interests, and American Trade Policy* (1993) shows how the institutionalization of policy ideas at an earlier juncture had lasting consequences for subsequent U.S. trade measures.

These are but a few of many possible citations, for historical institutionalist studies have cumulated to provide wide-ranging as well as causally precise understandings of such important matters as transitions to democracy,[1] the emergence and demise of authoritarian regimes,[2] the intersection of domestic and international politics,[3] the origins and development of welfare states,[4] social identities in politics,[5] the political dynamics of gender rights,[6] the development of economic regimes,[7] and the causes and consequences of social movements and revolutions.[8]

Obviously, studies using historical-institutionalist strategies of analysis vary in many important ways. Some are explicitly comparative, while others analyze trends within just one macrocontext. Some offer suggestive interpretations (e.g., Hart 1994), while others develop explicit models framed in general terms (e.g., Hansen 1991; Luong 2002). Some historical-institutionalist studies draw extensively from primary sources (e.g., Gamm 1999), while others synthesize findings from secondary publications (e.g., Skocpol 1979; Downing 1992). And some deploy arguments about strate-

1. See, for example, Anderson 1999; Baloyra 1987; Bratton and Van de Walle 1997; Diamond 1999b; Downing 1992; Gould 1999; Haggard and Kaufman 1995; Rueschemeyer, Stephens, and Stephens 1992; and Yashar 1997. A comprehensive review of hypothesis testing and cumulative theoretical development in this field appears in Mahoney 2002.

2. Examples include Chehabi and Linz 1998; Doyle 1986a; Ekiert 1996; Im 1987; Mahoney 2001; and Snyder 1998. For further review of the literature, see Mahoney 2002.

3. Examples include Friedberg 2000, Gourevitch 1986, Ikenberry 2001, Katzenstein 1978, Krasner 1978, Simmons 1993, and Sparrow 1996.

4. See Esping-Andersen 1990, Flora and Heidenheimer 1981, Hacker 1998, Hicks 1999, Howard 1997, Huber and Stephens 2001, Immergut 1992, Maioni 1998, Pierson 1994, Skocpol 1992, and Steinmo 1996. The striking cumulation of knowledge in this field is reviewed in Amenta 2002 and Pierson 2000d.

5. Examples include Anderson 1991, Hattam 1993, Katznelson and Zolberg 1986, Kryder 2000, Lustick 1993, Marx 1998, Varshney 2001, and Vogel 1978.

6. Examples include Charrad 2001; Htun forthcoming; Jenson 1986; Mettler 1998; and O'Connor, Orloff, and Shaver 1999.

7. For key examples from a wide-ranging literature, see Karl 1997, Richards and Waterbury 1990, Soskice 1999, Streeck 1992, Thelen 1993, 1994, and Zysman 1994.

8. Important examples include Banaszak 1996; Goldstone 1991; Goodwin 2001; McAdam 1982; McAdam, Tarrow, and Tilly 2001; Skocpol 1979; Tarrow 1998; and Wickham-Crowley 1992.

gic choice and the impact of the rules of the game (e.g., Immergut 1992; Pierson 1994), while others adopt culturalist modes of explanation (e.g., Anderson 1991; Hattam 1993). Any vibrant tradition of research encompasses variety and flourishes through internal debates, and historical institutionalism is certainly no exception. In another context, we could elaborate differences among works with historical-institutionalist features—defending our own choices within these debates.

But we have a different goal for this chapter. Despite variety on many key dimensions, historical institutionalists commonly employ distinctive and complementary strategies for framing research and developing explanations. What historical institutionalists broadly share becomes apparent when we juxtapose their ways of asking questions and seeking answers—their research strategy—to the strategies normally used by behavioralists and rational-choice modelers. Without denying variety within major approaches, this essay aims to make distinctive core strategies visible, so that we can see the advantages and limits of historical institutionalism compared to the other research approaches extensively used in empirical political science.

We characterize historical institutionalism and other leading approaches in empirical political science on the basis of "elective affinities" shared in practice by many scholars who do each style of work. We are not saying that everyone in each camp marches to the same tune. And we recognize that many scholars blend styles of research in highly creative ways (as we will suggest in the conclusion to this chapter, "boundary crossers" are often among the most creative scholars in our discipline). Practitioners of major approaches nevertheless share ways of posing questions and developing explanations—giving historical institutionalism, behaviorialism, and rational-choice modeling characteristic features, strengths, and weaknesses. Operating as if we were anthropologists documenting the folkways of neighboring and comingled clans, we explore how historical institutionalists, compared to their cousins doing behavioralism or rational choice, regularly go about defining research agendas and developing explanations.

Three important features characterize historical-institutional scholarship in contemporary political science.[9] Historical institutionalists *address big, substantive questions that are inherently of interest to broad publics as well as to fellow scholars*. To develop explanatory arguments about important outcomes or puzzles, historical institutionalists *take time seriously*, specifying sequences and tracing transformations and processes of varying

9. Although we discuss historical institutionalism as one of three major research tendencies in contemporary political science, we readily acknowledge that the relevant literatures also include contributions from "comparative historical" political sociologists. This is hardly surprising. Leading approaches in the social sciences invariably bring scholars together across disciplinary boundaries; and political scientists of every persuasion have long shared theories and methods with their cousins in sociology and economics.

scale and temporality. Historical institutionalists likewise *analyze macro contexts and hypothesize about the combined effects of institutions and processes* rather than examining just one institution or process at a time. Taken together, these three features—substantive agendas, temporal arguments, and attention to contexts and configurations—add up to a recognizable historical-institutional approach that makes powerful contributions to our discipline's understandings of government, politics, and public policies.

To explain how this approach works and make the case for its fruitfulness, we discuss in turn each of the three aspects of historical-institutionalist scholarship, pausing at appropriate points to explore advantages or limitations compared to strategies used by other families of political scientists. We focus on what creative communities of scholars actually do, paying less attention to what they or others say they do (or claim they should do). At the end of the chapter, we step back to consider some of the broadest issues of empirical research method and strategies of knowledge cumulation: Can historical institutionalists really develop valid arguments from case studies and small-*n* comparisons? How scientifically fruitful are research agendas driven by substantive questions rather than the elaboration of a theory or techniques? And what are the prospects for combining the strengths of historical-institutional analysis with advances in strategic modeling or statistically sophisticated survey research? Historical institutionalism is experiencing a bold new phase of methodological development, yet its focus on substance and its theoretical eclecticism simultaneously open the way for fruitful cross-fertilization with the best of sister research traditions.

■ | Big Questions and Real-World Puzzles

Despite the disparate phenomena they investigate, historical institutionalists formulate their research programs in recognizable and distinctive ways. A historical-institutionalist scholar usually starts by asking about varied, historically situated outcomes of broad interest—perhaps posing a puzzle about why something important happened, or did not happen, or asking why certain structures or patterns take shape in some times and places but not others. Why have revolutions occurred in some times and places but not others? How did the U.S. state develop its specific pattern of institutional features? Why have welfare states emerged and developed along various paths? Why have some countries become stable democracies, while others have not? Under what circumstances do ethnic identities become prominent in national or international politics? The focus is on explaining variations in important or surprising patterns, events, or arrangements—rather than on accounting for human behavior without regard to context or modeling very general processes presumed to apply at all times and places.

Proceeding through a constant movement back and forth among cases, questions, and theories, historical-institutionalist scholars immerse themselves in historical instances not only to test previously formulated hypotheses but also to come up with fruitful new puzzles. The problems that interest historical institutionalists often come from identifying heretofore unexplained real-world variations—or from realizing that empirical patterns run counter to received academic or popular wisdom. For example, several recent studies of the development of U.S. social policies started with the observation that standard treatments of the United States as a lagging and ungenerous welfare state miss extensive systems of private social provision encouraged by public subsidies and regulations. Building from this insight, scholars such as Howard (1997), Gottschalk (2000), and Hacker (forthcoming) have developed broader and richer causal accounts of the development of U.S. policy. Similarly, scholars looking into the global evolution of rights for women noted that breakthroughs often occurred under surprising political circumstances, for example, under militaristic or authoritarian regimes. Such observations soon inspired more extensive and nuanced comparative-historical studies (e.g., Charrad 2001; Htun forthcoming; Jenson 1986) highlighting the impact of various state-building sequences, institutional contexts, and political coalitions on movements for women's rights.

From the points of view of academic marketing and disciplinary recognition, there may be drawbacks to the substantively focused and puzzle-driven nature of historical-institutionalist research. Contributions tend to be clustered in somewhat separate topical literatures and scattered across subfields dealing with different eras or regions of the world. But there are also telling advantages to research focused on big, important substantive problems, including studies that start by trying to resolve anomalies not previously noticed or explained. Social science must ultimately be judged, claims Lewis Coser, "on the basis of the substantive enlightenment . . . it is able to supply about the social structures in which we are enmeshed and which largely condition the course of our lives" (1975, 698). Historical institutionalists are mindful of this supreme test. Grappling as they are with such matters as social movements, the development of the modern state, the rise and fall of civic engagement in democracies, the origins and dynamics of political economies, regime transformations, and patterns of public policies, historical institutionalists avoid academic navel gazing. Grappling with pressing puzzles, historical institutionalists address real-world questions of interest to educated publics and university students—not to mention topics that appeal to book publishers.

Within political science itself, it is also worth noting, historical institutionalists bridge many divides, including the gulf that sometimes separates normative theorists from empirical researchers. Normative dilemmas are frequently apparent in the phenomena explored by historical institutionalists, whose studies thus contribute empirical substance to debates raging among political theorists. What is more, research agendas shaped by historical institutionalists may draw considerable interest from formal theorists and behavioralists. The literature on the development of modern welfare states exemplifies multiple forms of bridging. Normative concerns about equality, democracy, and liberty are frankly acknowledged and explored in this literature. And because research invariably focuses on major, real-world developments and dilemmas—as historical institutionalists prefer—choice theorists and survey researchers have been drawn into academic dialogues remarkably free of sterile paradigm disputes. Theoretical and methodological dialogues have been highly fruitful, and particular studies often combine modes of analysis (see the literature reviews in Amenta 2002; Pierson 2000d). With historical institutionalists playing a central role, the entire research community has used a full range of methods and research designs to develop ever sharper and broader explanations—and richer normative understandings—of the origins and dynamics of national systems of economic regulation and social provision. Everyone seems to realize that theoretical eclecticism, multiple analytic techniques, and a broad comparative and historical purview works best.

■ | Tracing Historical Processes

Historical institutionalists may ask big questions and bridge divides within and beyond academia, but how do they go about developing explanations? We have already alluded to what is perhaps the most distinctive feature of this approach: Whether particular works use comparisons or analyze various aspects of one theoretically justified case, historical institutionalists take history seriously, as something much more than a set of facts located in the past. To understand an interesting outcome or set of arrangements usually means to analyze processes over a substantial stretch of years, maybe even many decades or centuries. Scholars working in this tradition have developed compelling methodological and theoretical justifications for historically grounded investigations, by which they mean investigations that look not just at the past but at *processes over time*.

Some reasons for taking history seriously concern methodology and are already recognized (if not always practiced) by political scientists of many persuasions. Extending the time frame of social inquiry obviously widens the range of experience available for examination. This simultaneously makes more data available and generates greater variation in outcomes. Such widening of the empirical terrain is especially important for political scientists, because many matters of great interest—especially macrophenomena such as revolutions, state building, democratization, the construction of welfare states—occur relatively infrequently, or only partially, within any particular slice of time. Historically informed investigation also sensitizes investigators to temporal boundary conditions or period effects, with respect to claims about causal relationships. By examining a wider range of historical settings, an analyst can consider the possibility that supposedly universal effects in fact only hold under particular circumstances.

Historical investigations can also contribute to causal inference. Because theoretically grounded assertions of causal relationships imply temporal relationships among variables (either that one precedes the other or that both happen at essentially the same time), examining historical sequences is extremely useful for testing such assertions (Rueschemeyer and Stephens 1997; Mahoney 2000b). Optimally, assertions of causality should be borne out not just by a correlation between two variables but also by a theoretical account showing why this linkage should exist and by evidence suggesting support for the theorized linkage. Although social research does not always achieve this ideal, efforts to systematically trace social processes can make an essential contribution to the rigorous assessment of claims about social causation (Bennett and George 1997; Hall 2002). Here the relatively small number of cases in many historical institutional studies can become an advantage, allowing for exactly the sorts of detailed examinations of *processes* needed to evaluate claims about causal mechanisms.

Important as these general methodological advantages are, historical institutionalists are also pushing beyond them to *theorize about historical dimensions of causation*. As we are about to see, historical institutionalists have recently begun to underline the theoretical power of their *dynamic* approach to social change and politics. Without the kind of attentiveness to temporally specified process that is a distinctive hallmark of historical-institutionalist scholarship, important outcomes may go unobserved, causal relationships may be misunderstood, and valuable hypotheses may never receive consideration.

PATH DEPENDENCE, SEQUENCES, AND CONJUNCTURES

A central example of why history may be causally critical involves claims about *path dependence* that are common in historical institutionalist scholarship (see, e.g., Collier and Collier 1991; Ertman 1997; Hacker 1998; Shefter 1977; Huber and Stephens 2001). *Path dependence* can be a faddish term, lacking clear meaning, but in most historical-institutionalist scholarship, it refers to the dynamics of self-reinforcing or positive feedback processes in a political system (Pierson 2000a; cf. Mahoney 2000b). A clear logic is involved in such path-dependent processes: Outcomes at a critical juncture trigger feedback mechanisms that reinforce the recurrence of a particular pattern into the future. Such processes have very interesting characteristics. They can be highly influenced by relatively modest perturbations at early stages. Once actors have ventured far down a particular path, however, they are likely to find it very difficult to reverse course. Political alternatives that were once quite plausible may become irretrievably lost. Thus, events or processes occurring during and immediately following critical junctures emerge as crucial.

There are strong theoretical grounds for believing that self-reinforcing processes are prevalent in political life. Once established, patterns of political mobilization, the institutional rules of the game, and even citizens' basic ways of thinking about the political world will often generate self-reinforcing dynamics. In addition to drawing our attention toward critical junctures or formative moments, arguments about path dependence can thus help us to understand the powerful inertial "stickiness" that characterizes many aspects of political development. These arguments can also reinvigorate the analysis of power in social relations, by showing how inequalities of power, perhaps modest initially, are reinforced and can become deeply embedded in organizations, institutions, and dominant modes of political understanding. Path-dependence arguments also provide a useful and powerful corrective against tendencies to assume functionalist explanations for important social and political outcomes. Perhaps most important, an appreciation of the prevalence of path dependence forces attentiveness to the temporal dimensions of political processes. It highlights the role of what Arthur Stinchcombe (1968) has termed historical causation in which dynamics triggered by an event or process at one point in time reproduce themselves, even in the absence of the recurrence of the original event or process.

An appreciation of increasing returns dynamics is one important justification for the focus on issues of *timing and sequencing* that constitutes a second important theoretical rationale for examining historical processes. In path-dependent processes, the order of events may make a fundamental difference. Historical institutionalists tracing broad patterns of political development across a number of countries often argue that the timing and sequence of particular events or processes can matter a great deal (Anderson 1986; Gerschenkron 1962; Kurth 1979; Shefter 1977; Ertman 1997). In Ertman's analysis of European regime formation, for example, it is the sequence of two processes—the expansion of literacy and the onset of military competition—that is crucial to paths of state building. The prevalence of literacy powerfully affected the bargaining position of those who would collect state revenues. Where countries faced intense military challenges prior to the era when literacy became widespread, the character of the state's fiscal apparatus was likely to be very different than when the order of these two processes was reversed. Crucially, because of path-dependent effects these differences were long lasting; that is to say, a country's bureaucratic structures, established at an early juncture, did not simply adjust once literacy rates rose to a higher level.

Jacob Hacker's study (2002) of the development of U.S. social policy presents a second compelling example of this kind of argument. Hacker explores the distinctive paths of development of public versus private, employer-based systems of health care and pension provision in the United States over the past century. Although both policy sectors have roughly equal mixes of public and private spending, they developed along fundamentally different lines. In health care, public interventions have served as supplements or compensations added to a prior set of institutions featuring private provision. By contrast, in pensions, a public core system of retirement provision developed first, with private arrangements playing a supplementary or complementary role. Through processes of policy feedback, these different sequences of public and private intervention generated quite different interest group environments, shifting both the policy preferences and political resources of crucial actors like employers. Moreover, the different patterns of development produced very different distributional outcomes, profoundly shaping the contours of contemporary political struggles over social provision. Any scholar who merely discusses these contemporary struggles without awareness of the history that shaped the terrain of preferences and actors will miss much of central causal relevance to explaining politics and policymaking today.

Waldner's study (1999) of state building and economic performance in late-developing countries rests on a similar claim. He contrasts Syria and Turkey, where state building and mass incorporation occurred simultaneously, with Korea and Taiwan, where greater elite consensus allowed state building to precede mass incorporation. In the latter cases, state elites were able to institutionalize structures that facilitated long-term economic development before the challenges of bargaining with nonelites had to be confronted. In the former cases, patterns of state building were fundamentally altered by the immediate need to win support from broader social groups. Thus Waldner identifies two distinctive paths of late state development, generated by different sequences of state building and incorporation.

Like these arguments of Ertman, Hacker, and Waldner, most propositions historical institutionalists offer about the causal impact of sequences are grounded in claims about self-reinforcing or positive feedback processes (Pierson 2000c). Relative timing, or sequence, matters because self-reinforcing processes, playing out over time in political and social life, transform the consequences of later developments. Path-dependent arguments about self-reinforcement explain why and when sequencing can matter. Increasing returns processes occurring during particular periods generate irreversibilities, essentially removing certain options from the subsequent menu of political possibilities.

Although many path-dependent arguments stress the institutionalization of particular configurations that subsequently prove difficult to dislodge, the focus on sequencing illuminates how arguments about path-dependent processes can be incorporated into explanations of political change as well as political inertia. For instance, path-dependent processes may operate to institutionalize specific political arrangements that ultimately prove vulnerable to some displacing event or process emerging at a later stage in political development (Collier and Collier 1991; Luebbert 1991). The Colliers' work on labor incorporation in Latin America provides an excellent example. In some of their cases, labor was excluded from political participation at a critical juncture, and positive feedback led to the institutionalization of regimes where organized labor had little access to political influence. Over time, however, key changes in social and economic life made it impossible to maintain these systems of labor exclusion. Eventually, organized labor had to be incorporated. When this took place at a later stage of political development, however, it necessarily occurred on quite different terms and with different consequences than was true for countries experiencing early incorporation. The key point is that options available for early incorporators were not available for late incorporators.

It is highly instructive to contrast these arguments about path-dependent sequences with arguments rational-choice theorists have made about sequences within highly institutionalized settings (Shepsle 1986). Working from Arrow's impossibility theorem, which suggests the prospect of endless cycling in many collective-choice situations, rational-choice theorists have argued persuasively that institutional arrangements governing agenda control and decision-making procedures can produce stable outcomes. By demonstrating the crucial role of sequencing as choices are made, these arguments rest on the equivalent of a path-dependent mechanism: Steps in a sequence are irreversible because institutional rules cause forgone alternatives in early rounds to be dropped from the range of possible later options (as, for example, in committee voting rules that require a sequence of binary choices, with losers eliminated). By showing how such irreversibilities can be generated in a wide variety of social contexts, however, it is possible to expand this crucial finding to a far broader range of social phenomena than those covered by the rational-choice literature stemming from Arrow's work. In comparative historical analyses, analogous arguments are often applied to large-scale social changes such as democratization (Collier and Collier 1991; Collier 1999), industrialization (Gerschenkron 1962; Kurth 1979), state building (Ertman 1997; Shefter 1977), or welfare state development (Huber and Stephens 2001). Sequencing matters not only for collective choices within legislatures but also potentially for *any* sociopolitical process where self-reinforcement means that forsaken alternatives become increasingly unreachable with the passage of time.

Historical institutionalists also employ timing and sequence arguments to highlight *conjunctures*—interaction effects between distinct causal sequences that become joined at particular points in time (Aminzade 1992; Orren and Skowronek 1994; Sewell 1996). The ability to identify and explore such conjunctures is a major advantage of the more macroscopic inclinations of historical institutionalism. For example, Skowronek's account (1993) of U.S. presidential leadership emphasizes the interaction of slowly developing modern institutional capacities and the particular position of an individual president within the sequential rise and decline of a dominant political coalition. The explanatory power of Skowronek's analysis therefore derives from highlighting the intersection of long-term developments with recurrent political cycles. Other historical institutional accounts have focused on the political effects produced when separately unfolding processes conjoin. Thus Ertman (1997) stresses links between military competition and developing social capacities for bureaucratic governance, Shefter (1977) analyzes the interaction of state building and party formation, and Skocpol (1979) attributes social revolutionary outbreaks to conjunctures of domestic and international conflicts. These analysts focus on distinct sociopolitical processes that become linked in different and causally crucial ways depending on relative timing. The causal centrality of such conjunctures, it should be stressed, would never be noted by analyses investigating or theorizing about single processes in isolation.

SLOW-MOVING CAUSAL PROCESSES

Another theoretical justification for focusing on historical process is to draw attention to lengthy, large-scale, but often very slow moving social processes (Pierson 2002). Historical institutionalists seek to be attentive to the unfolding of both causal processes and important political outcomes over extended periods of time. Most political scientists are strongly predisposed to focus on aspects of causal processes and outcomes that unfold very rapidly. Yet many things in the social world take a long time to happen. Some causal processes and outcomes occur slowly because they are incremental; it simply takes a long time for them to add up to anything. Changes in pension systems, for example, are not fully translated into levels of public spending for a half-century or more. A second possibility is the presence of threshold effects; many social processes may have little significance until they attain a critical mass, which may then trigger major change (McAdam 1982; Goldstone 1991; Baumgartner and Jones 1993). Alternatively, slow-moving processes may involve transformations that are probabilistic during any particular period and therefore several periods may be necessary before the transformation occurs. Under such circumstances, the social outcome of interest may not take place until well after the appearance of key causal factors. Particularly when it focuses on macroscopic processes, historical-institutionalist research is often primarily interested in such structural preconditions for particular outcomes, rather than the specific timing of those outcomes (Collier 1999; Moore 1966; Rueschemeyer, Stephens, and Stephens 1992). When either structural causes or threshold effects are at work, analysts adopting a short time frame are likely to focus erroneously on the more idiosyncratic or precipitating factors that trigger outcomes. Because some crucial social conditions may change only slowly, analysts studying a narrow time frame will be strongly inclined to take them as fixed and therefore irrelevant to their causal accounts (Rueschemeyer, Stephens, and Stephens 1992; Kitschelt 1991).

Another possibility is that causal processes involve chains with several links, which require some time to work themselves out. To the extent that causal chains of this sort are at work, analyses must frame their studies on a broad time scale. Collier and Collier's influential work on labor incorporation in Latin America presents arguments of this kind, in which the ultimate outcomes of interest reflect a sequence of key developments over extended periods of time (Collier and Collier 1991). Indeed, this type of claim about long-term, multistage causal processes is often invoked in work on state building (Flora 1999a, b) or democratization (Luebbert 1991; Collier 1999).

Swank (2001a) offers an instructive recent example. In assessing the impact of political institutions on welfare state retrenchment, he criticizes the view that fragmented institutions will limit cutbacks by increasing the number of veto points available to defenders of the status quo. Swank argues that this is true as far as it goes but notes that the long-term, indirect effects of institutional fragmentation run largely in the other direction. Not only does institutional fragmentation limit the initial expansion of the welfare state, but it also reinforces social heterogeneity, inhibits the growth of encompassing interest groups, and weakens cultural commitments to universalist social policies. All of these long-term effects strengthen the welfare state's opponents and weaken its advocates. Thus many of the most important effects of institutional fragmentation work themselves out only indirectly and over extended periods of time. Ahistorical analyses typically seek to consider the effects of institutions while holding constant other variables, but these variables are in part the long-term consequences of institutional structures. Ahistorical investigations are therefore likely to systematically misread the impact of institutional structures on the politics of the welfare state.

Analysts who fail to be attentive to these slow-moving dimensions of social life may ignore potentially powerful hypotheses. They are particularly likely to miss the role of many sociological variables, like demography (Goldstone 1991), literacy (Ertman 1997), or technology (Kurth 1979), as well as the impact of other slowly building pressures such as international military competition and fiscal overload (Skocpol 1979). Their explanations may focus on triggering or precipitating factors rather than deeper causes (Kitschelt 1991). Perhaps most fundamental of all, they may fail to even identify some important questions about politics because the relevant outcomes happen too slowly and are therefore simply off their radar screens.

History as Process, Not Just Illustration

In the ways we have just surveyed, theoretical attentiveness to historical processes represents a formidable comparative advantage of historical institutionalism, especially since this attentiveness is linked to macroscopic analysis focusing on institutions and organizations in addition to aggregates of people. Most research in the behavioral tradition uses surveys that offer a snapshot in time. And even when surveys are repeated to offer a longitudinal series, it is rare indeed for behavioral analysts to consider changing institutional contexts, critical conjunctures, or path-dependent large-scale processes as causally relevant to the changing modes of individual behavior they probe. There are exceptions, certainly, such as Carmines and Stimson's dissection of the processes by which race became a transforming issue in U.S. partisan politics, or Putnam's consideration of the impact of World War II and other watershed historical events on "life course" developments for late-twentieth-century U.S. adults (Carmines and Stimson 1989; Putnam 2000). But for the most part, survey analysis relies on one-time data about attitudes and self-reported behaviors to explore individual-level hypotheses about mass patterns. Change over time often enters the discussion only speculatively, as when a researcher reports differences between working women and housewives at a moment in time, and then go on to speculate that the changing proportions of such persons in the population may signify a continuous social change.

To be sure, some rational-choice scholars have turned to historical case studies in recent years. But most of the analytic advantages we have outlined are absent in this work, because the past enters only in a highly restricted sense, as what might be termed illustrative history, the mining of the historical record for outcomes that can be "explained" by particular rational-choice models. Previously established models may be applied in interesting ways to examples in the past (cf. Bates et al. 1998), but the tools of game theory turn out to be poorly suited to analyzing conjunctures or exploring slow-moving macroprocesses. Rational-choice theory faces considerable difficulties in moving from the micro- to the meso- or macrolevels of analysis that are typically featured in works that analyze processes over long stretches of time (Elster 2000; Munck 2001b). Rational choice's reliance on theory-driven agendas and on the identification of empirical terrain favorable for its preferred methods leads relentlessly back to an emphasis on the micro. And the results of game theory quickly become indeterminate or unmanageably complex as one increases the number of actors involved (indeed, in game theory the problem of indeterminacy is often rife even at the microlevel). The fact that many macroprocesses take considerable time to play out presents a further difficulty, since game theory generally requires that all the relevant actors, preferences, and payoffs be established and fixed simultaneously at the beginning of a game. In short, there are real obstacles in rational-choice theory to serious consideration of many key aspects of historical processes.[10]

Analyzing Institutions in Context

Historical institutionalism is characterized by the second part of its label as well as the first. But what does *institutionalism* mean for this family of scholars? We can say of much political science today what Richard Nixon once said of Keynesianism: We are all institutionalists now. In political science today many scholars analyze how institutions influence political behavior and shape processes ranging from legislative decision making to social movements (Hall and Taylor 1996). As Thelen (1999) has ably elaborated, both rational-choice institutionalists and historical institutionalists presume that organizationally embodied routines play a crucial role in allocating resources and structuring the incentives, options, and constraints faced by political participants. In this sense, institutionalism is indeed a broadly shared approach in contemporary political science.

But even though institutionalists of different stripes have converged on some complementary questions and findings (Thelen 1999, 372–81), important differences remain. Rational-choice scholars tend to focus on rules of the game that provide equilibrium "solutions" to collective action dilemmas. Historical institutionalists, meanwhile, probe uneasy balances of power and resources, and see institutions as the developing products of struggle among unequal actors. Rational-choice scholars often focus on one set of rules at a time. Historical institutionalists, by contrast, typically do meso- or macrolevel analyses that examine multiple institutions in interaction, operating in, and influenced by, broader contexts. They pay close attention to ways in which multiple institutional realms and processes intersect with one another, often creating unintended openings for actors who trigger changes. Historical institutionalists investigate the rise and decline of institutions over time, probing the origins, impact, and stability or instability of specific institutions as well as broader institutional configurations. Sometimes the principal goal is to explain the institutional arrangements themselves, at other times to use variables referring to institutional configurations to explain other outcomes of interest.

10. Even where rational-choice theory seems better equipped, it has made very limited moves in this direction. For instance, despite Douglas North's important work (1990) on path dependence, his strong focus on historical process has been largely ignored by rational-choice theorists (for partial exceptions, see Aldrich 1994; Harvey 1998).

INSTITUTIONAL EFFECTS

Much research in historical institutionalism adopts a mesolevel focus, concentrating, for example, on policy developments in a particular issue area (e.g., Hacker 1998; Immergut 1992; Weir 1992b) or changes in organizational fields (e.g., Skocpol, Ganz, and Munson 2000). Historical institutionalists may also tackle the most macrolevel developments, such as modernizing intellectual transformations (Wuthnow 1989) or state formation (e.g., Anderson 1986; Doyle 1986a; Tilly 1975; Downing 1992; Ertman 1997; Skowronek 1982). In either case, analyses tend to highlight and explore causes operating at the interorganizational or interinstitutional level. Certainly, historical institutionalists accept the principle that causes should ultimately be consistent with plausible accounts of individual motivation and behavior (Little 1991). But they also believe that the patterns of resources and relationships in which individuals find themselves have powerful channeling and delimiting effects and that many of these effects are expressed through the conjoint impact of multiple institutions. As Ronald Jepperson puts it, these processes occurring "within and between institutions . . . are of course produced via the behavior of people, but . . . the causal linkages involved in these collective processes are far removed from the aggregation of simple social behavior" (2001, 5–6). So historical institutionalists aim to make those patterns visible and trace their causal impacts.

Historical institutionalists rarely focus on a single institutional or organizational site of contestation, as rational choice scholars often do. There is a strong tendency to doubt the power of many claims about institutional effects that rest solely on an analysis of that institution in isolation. Research on U.S. policy development, for instance, typically focuses on the interplay among multiple organizational actors in multiple institutional settings (Baumgartner and Jones 1993; Skocpol 1992; Weir 1992a). Melnick's analysis (1994) of the "rights revolution" that has fueled regulatory expansion in the United States provides a good example. Tracing this revolution through multiple venues and over an extended stretch of time, he demonstrates that it must be understood as the result of an interaction between the courts and Congress, with newly emergent citizens' organizations playing a crucial role in coupling these distinct institutional sites. Advocates of policy activism in the federal courts and congressional committees have, through interplay between the two branches, been able to advance their agendas beyond what a majority in Congress would have been likely to produce on its own.

Similarly, recent work in both comparative political economy (Kitschelt, Lange, Marks, and Stephens 1999; Hall and Soskice 2001) and social policy (Esping-Andersen 1990; Huber and Stephens 2001) has focused on how *configurations* of policies, formal institutions, and organizational structures generate distinctive welfare state regimes or varieties of capitalism that operate in fundamentally different manners. Outcomes are generated not by some universal operating principles characteristic of a given type of actor or realm of activity but by intersections of organized practices (such as labor markets, firm structures, or crucial policy arrangements). These practices will often have originated at different times and then formed configurations that advantage key actors (such as employer associations and labor unions). These actors, in turn, will generally work to maintain the configuration with only incremental adjustments even when economic, cultural, and geopolitical circumstances shift.

Institutional Development

Issues of long-term institutional development are also central to historical institutional research agendas (Thelen 1999). Historical institutionalists are typically suspicious of functional explanations, in which institutional outcomes are explained by their consequences. In such functional accounts, institutions develop because of their capacity to solve certain collective problems. The implicit or explicit claim is that rational actors produced these outcomes in order to solve these problems.

As suggested above, concern with issues of institutional development within historical institutionalism is strongly linked to theorizing about the causal relevance or origins, sequences, and temporal processes. Functional accounts of institutions seem most plausible when investigations take a one-time snapshot, because long-term effects and inconsistent "layers" of institutional development (see Schickler 2001) are not so readily noticed in such studies. In most cases, synchronic analysts simply probe an extant institution to uncover benefits for particular actors. Analysts then infer that these benefits explain the institution, implying that the actors who are presently advantaged (or their forebears) created the institution to produce the benefits. This is a plausible hypothesis, but it is only a hypothesis and over-time investigation of institutional origins and dynamics often does not bear out such lines of reasoning.

By examining issues of institutional origins and change over an extended time frame, historical institutionalists have been able to highlight a number of potential problems for functional accounts (see Thelen 1993, 1994, 1999; Pierson 2000b for relevant findings and arguments). Functional interpretations of politics are often suspect because of the sizable temporal gap between actors' actions and the long-term consequences of those actions. Political actors, facing the pressures of the immediate or skeptical about their capacity to engineer long-term effects, may pay limited attention to the long run. Thus the long-term effects of institutional choices, which are frequently the most profound and interesting ones, should often be seen as the *by-products* of social processes rather than the realization ("congealed preferences" as Riker put it) of actors' goals.

A second issue concerns unintended consequences. Even where actors may be greatly concerned about the future in their efforts to design institutions, they operate in settings of tremendous complexity. As a consequence, they will often make mistakes. Thus institutions may not be functional even in a context of far-sighted actors, because they do not operate as intended. Although widely acknowledged to be significant in actual politics, political scientists often treat unintended consequences as an error term or simply ignore them by failing to investigate institutions over time. In cross-sectional studies of institutions, the issue of unintended consequences vanishes from view, since either the long-term consequences of institutional choices or the original factors generating the institutional choice will be outside the scope of the analysis. By contrast, historical institutionalists examining institutional development often stress the surprising long-term consequences of earlier political choices and conflicts (Anderson 1986; Luebbert 1991; Skocpol 1992; Thelen 1999, 2002).

Historical institutionalists, finally, emphasize the ways in which institutions are remade over time (Thelen 1999, 2002). Because of strong path-dependent effects, institutions are not easily scrapped when conditions change. Instead, institutions will often have a highly "layered" quality (Schickler 2001; Stark and Bruszt 1998). New initiatives are introduced to address contemporary demands, but they add to rather than replace preexisting institutional forms. Alternatively, old institutions may persist but be turned to different uses by newly ascendant groups. In either case, the original choices are likely to figure heavily in the current functioning of the institution. Thus institutions will rarely look like optimal solutions to present collective action problems.

Clearly, attentiveness to history and the investigation of meso- or macrolevel institutional configurations are highly complementary strategies of analysis. Tracing politics through time is very helpful for identifying the boundary conditions for particular theoretical claims. Even more significant, the emphasis historical institutionalists place on conjunctures and sequencing draws attention to the temporal connections among social processes, and highlights the importance of meso- or macrolevel analysis of institutional configurations. Furthermore, while path-dependent or increasing returns processes can play out at a microlevel (e.g., in the way individuals develop and reinforce particular mental maps of the social world), they are often most significant at the meso- or macrolevel. Particular sets of institutions and organizations are often mutually reinforcing or complementary—the presence of each enhances tendencies for the development of others—in a manner that can be seen as a kind of coevolutionary or selection process playing out over considerable stretches of time. A good example is the recent research in comparative political economy focusing on varieties of capitalism (Hall and Soskice 2001b; Estevez-Abe, Iversen, and Soskice 1999; Soskice 1999). This work stresses how different political economies have developed along quite different lines because of the coevolution of mutually reinforcing institutional and organizational structures. Patterns of firm organization, systems of interest intermediation and key social and economic policy structures contain institutional complementarities that place national political economies on distinct, path-dependent courses of development.

One can also see the advantages of a combined focus on institutions and temporal processes in the development of extensive historical institutionalist work on *policy feedbacks* (Heclo 1974; Weir and Skocpol 1985; Pierson 1993). Many of the works cited here emphasize that government policies establish some of the most important structuring rules of modern societies. The specific design of policies can have an enormous effect on the resources and strategies subsequently available to political actors—indeed, in many contexts, policies can be as important as formal political institutions in shaping political processes and outcomes (Esping-Andersen 1990; Skocpol 1992; Weir 1992a; Pierson 1994; Huber and Stephens 2001; Campbell forthcoming; Hacker 2002). The success of this line of argument demonstrates very clearly how the particular vantage points encouraged by different analytical strategies can either highlight or obscure important aspects of social reality. Because rational-choice scholarship is so focused on the pursuit of broad generalizations, it has concentrated almost entirely on recurrent, carefully delimited *formal* institutions such as legislatures or central banks.

Only a few rational-choice scholars (such as Bates 1988b) focus on the state overall and highlight the ulterior effects of state policies. In formal institutional accounts, policies are usually treated as dependent variables, with their further political consequences through feedbacks falling beyond the scope of analysis, either because these consequences are seen as too idiosyncratic or because analysis of policy feedbacks would require attentiveness to processes unfolding over time. But this omission will not do, because the expanding activities of states—and the policy feedbacks flowing from them—have been a central fact of modern life. Fortunately, state expansion and policy feedbacks have been at the center of historical-institutionalist analyses, even as these phenomena have largely vanished off the radar screen of the more universalist studies of formal institutions. The distinctive foci and contributions of rational-choice and historical-institutionalist work on institutions thus provide a clear demonstration of the advantages of pluralist political science, employing multiple strategies of inquiry. Without competitive pluralism in social science theorizing and research, important issues can too readily fall off the agenda of scholarship.

CAUSAL CONFIGURATIONS AND CONTEXT EFFECTS

The strong emphasis on interaction effects in historical institutionalism reflects some core working assumptions about how most sociopolitical processes operate. Analysts are strongly affected by their basic presumptions about how social processes work; and they tend to frame their problems, generate hypotheses, and employ methods of analysis and research designs that fit such presumptions (Abbott 1988; Hall 2002). Behavioralists, for example, are happy to use statistical techniques to analyze data from as many "cases" as possible—often data from surveys of thousands of individuals—because they are prepared to assume that very general variables operating independently of one another come together to account for the patterns of behavior they are trying to explain. Historical institutionalists, by contrast, assume that operative variables may not be independent of each other at all. When it comes to analyzing the origins and impact of institutions, causally important variables are often bundled together in the real world; and there may be alternative causal paths to similar outcomes (Ragin 1987; Shalev 1998). Historical institutionalists tend to suspect from the get-go that causal variables of interest will be strongly influenced by overarching cultural, institutional, or epochal contexts. This is not a matter of getting mired in thick description. As Andrew Abbott puts it: "Context has two senses. . . . The strict sense . . . denotes those things that environ and thereby define a thing of interest. The loose sense simply denotes detail. The acute reader will note that these correspond nicely to the two judgments of the scientific worth of contextual information. If decontextualization is merely the removal of excess detail, then it's a fine thing, scientifically. If, on the other hand, it is the removal of defining locational information, it is a scientific disaster" (1997, n. 10). Because historical institutionalists share the latter set of expectations, researchers in this tradition tend to move up from single institutions to broader contexts—*they look at forests as well as trees.* And they almost always seek to discover and explicate the impact of configurations of organizations and institutions on outcomes of interest (Katznelson 1997).

In addition to presuming—and concentrating attention on—causal configurations conceptualized at the organizational and institutional level, there is another way in which historical institutionalists highlight interaction effects. They often do this by pointing to overarching contexts—types of regimes, eras, regions, and cultures—that place bounds around the theorizing being done in any given study. Historical institutionalists rarely aim to write about all of humankind through all of global history. Scan the titles of major historical-institutionalist works, for example, and you will often find the proper names of regions of the world or the beginning and ending dates of the specific period that the argument is to cover. This is not done because historical institutionalists in political science are trying to be historians; they do not aspire to just tell stories about a time and place. Nor are dates given only to describe critical junctures when a change occurs that the analyst seeks to explain (e.g., Clemens 1997; Skowronek 1982). Beyond such an obvious use of dates and place, historical institutionalists often set limits to the applicability of their causal arguments, arguing in theoretically explicit ways why variables appear and combine in characteristic ways in one era but might not exist or combine in the same way in other eras.

For example, in his book about the growth (or not) and success or failure of guerrilla-led revolutionary movements, Timothy Wickham-Crowley (1992) develops a rigorous causal analysis and suggests (in his conclusion) that his model links up with theoretical explanations of revolutions in many regions and eras. But Wickham-Crowley carefully bounds his own argument, restricting it to Latin America since 1956. In analytical terms, he explains that specific developments and events created background conditions to generate many guerrilla movements with similar aims and methods, thus setting the stage for the variables he explores to account for movement growth and success or failure. In the end, Wickham-Crowley's careful delineation—and explanation—of the overarching context within which his variable-based analysis applies reinforces its theoretical power. As we move to other continents and periods, we can ask about changes that might influence the variables in play and the likely relationships among them.

Another way in which historical-institutionalist works highlight overarching contexts is by deliberately juxtaposing two or more contexts, to show how variable configurations already analyzed (across multiple cases) may play out in new ways when the overarching context changes. In the revolutions literature, Goodwin (2001) does this by, first, developing explanations for revolutionary successes and failures within Central America and East Asia, and then highlighting the somewhat different actors and conditions that come into play within each region. Addressing a very different problem—Pierson (1994) identifies variables about institutions and established policy characteristics that can explain why British and U.S. conservative politicians succeeded or failed with attempted cutbacks in a number of social policy areas. Yet he also steps back and stresses that the key causal relationships governing policy determination are quite different in the recent period of austerity than they were when welfare states were greatly expanded.

Scholars in other major political science traditions are often much less attentive to overarching contexts than historical institutionalists, in part because they prefer to focus on individual-level behavior or microprocesses but also because they are reluctant to write like historians. Ironically, the result can be less theoretically powerful work. Behaviorialists happily rely on types of data only available at one point in time or just for short time spans but do not necessarily think through what that means. As a result, significant overarching contexts can go unnoticed, unconceptualized. For example, behavioralists sometimes fail to notice that very similar individual-level patterns—joining more or fewer voluntary organizations, voting at different rates—can end up having very different meanings, depending on the kinds of organizations or institutions that predominate in a given nation or era. Theories invoked in behavioralist studies can end up being profoundly underspecified.

Rational-choice scholars, meanwhile, often write as if the models they present were infinitely generalizeable, even as they smuggle all kinds of institutional, cultural, and epochal specifics into their empirical operationalizations. This may seem an optimal way of arguing; isn't social science supposed to generalize? But, in practice, this kind of approach means that we leave important variables implicit and we fail to see how changing background conditions might cause the same variables to play out in very different ways. Again, the preference of rational-choice theorists for examining microsettings strongly reinforces this tendency.

Research Strategies and the Accumulation of Knowledge

Tackling big, real-world questions; tracing processes through time; and analyzing institutional configurations and contexts—these are the features that define historical institutionalism as a major strategy of research in contemporary political science. We have stressed the comparative advantages these approaches afford to all who would better understand government and politics. But we must also acknowledge the claims of critics who dismiss historical institutionalism as a valid approach to doing cumulative social science. Critics of historical institutionalism sometimes pen manifestos announcing that case studies and small-*n* comparisons cannot generate valid knowledge, because cases are not randomly selected and there may not be enough statistical degrees of freedom to test all conceivable hypotheses rigorously (cf. Geddes 1990; Goldthorpe 1991; Lieberson 1991). Statistical methodologists (such as King, Keohane, and Verba 1994) also worry that the historical-institutionalist proclivity for tackling significant issues predisposes them toward "selection on the dependent variable," that is, choosing cases where a phenomenon of interest has occurred, while ignoring the instances where it has not occurred.

Historical institutionalist books and articles are sometimes denounced in ways such as these one at a time, in isolation from one another—an approach that surely would be laughed out of court if applied to isolated works in any other tradition. In all research traditions, individual works build on one another, often extending lines of analysis, retesting arguments, and correcting for earlier limitations. The real issue is whether historical institutionalists in general have headed up blind alleys, because their studies, considered collectively as well as individually, are improperly designed. This is not the place for an exhaustive review of the profusion of recent methodological reflections in this genre, but the major signposts on the road can be mentioned.

Methodological challenges have, in our view, been good for historical institutionalism. Not only have these critiques been testimony to the visibility and intellectual impact of the studies they have dissected. Challenges have had a bracing impact, prompting historical institutionalists to spell out their metatheoretical presumptions and sharpen rationales and tools for doing valid macroscopic and historical studies. A number of scholars, for example, have asked how studies can best be designed to take intuitions about configurational causation and temporal process seriously. In other words, if historical institutionalists doubt that single, highly general variables have uniform effects across contexts, and regardless of interactions with other factors, how can research designs allow for adequate empirical exploration of hypotheses about contexts and configurations (see Abbott 2001; Mahoney 1999; Hall 2002; Ragin 2000)? If historical institutionalists stress mechanisms rather than simple association in causal arguments, how can process tracing be done in a rigorous way? And what criteria must be met to demonstrate valid arguments about path dependence, critical conjunctures, or sequence effects? Although Pierson (2000a, 2002), Thelen (1999, 2000a), and others have primarily focused on the theoretical characteristics of arguments about temporal causality, their reflections also set empirical standards to be met in case analyses and comparative studies that aim to establish the presence of causally relevant processes or events.

Reflecting on the strengths and limits of hundreds of studies in various important literatures, methodologically oriented historical institutionalists have also made the case for circumstances under which in-depth case studies and small- to medium-*n* comparisons are an optimal research strategy (Collier 1993; Mahoney 2000a; Munck 1998; Ragin 1987; Rueschemeyer and Stephens 1997; Tilly 1984). Random selection of cases—say, of particular regimes or nations from hundreds across the globe—is often not appropriate; and it is far from the only way to test hypotheses rigorously. Whatever the risks or drawbacks, hypotheses can be rigorously tested even when scholars cannot sample from large numbers of truly independent cases (for an excellent overview of strategies of inference in small-*n* research, see Mahoney 2000a). Alternative strategies of causal inference have been developed and applied, because there are important intellectual and practical advantages to focusing agendas of research on a small to intermediate number of cases, including instances of substantively compelling outcomes and arrangements we want to understand.

For example, a scholar who wants to understand revolutions is unlikely to be willing to start with a purely random set of times and places; she will want to make sure that clear-cut revolutions are included in her research. But in the best past studies of revolutions—as in many other historical-institutionalist literatures—it has long been standard to test hypotheses with comparisons between positive cases, where phenomena of interest occur, and closely matched negative cases, where the phenomena do not occur (Skocpol 1979; Goodwin 2001). Speaking about comparative studies in many topic literatures—analyses have juxtaposed time periods, regions, and policy sectors, turning what appear to be one or a few national instances into settings for many carefully compared cases. And even within what appear to be single case studies, empirical observations have often been multiplied by formulating and testing hypotheses about the *mechanisms* that connect causes to effects (Bennett and George 1997). All of these strategies of analysis have been used to good effect by scientifically sound historical-institutional studies, and the rapidly developing methodological literatures will make it easier for future scholars to recognize and apply the standards and rationales for best practices in case-based and small-*n* research.

No matter what theories or research methods are deployed, individual studies in isolation never do more than move the scholarly enterprise a step or two forward. Scholarship is an inherently communal enterprise, and so it is appropriate to reflect on how well clusters of scholars do at *accumulating* valid and important findings. Here is where historical institutionalists do quite well, in our view, because substantively compelling, problem-driven research facilitates exactly the sort of intellectual cumulation that allows a community of researchers to make clear progress over time.

Because historical institutionalists tend to tackle big, humanly important questions, scholars are likely to come back to the issues again and again over several academic generations. Within each generation as well, smart people jump in, willing to argue about how best to frame the questions and describe patterns worth explaining, as well as about how to formulate and test hypotheses. Researchers reexamine cases and hypotheses and extend established hypotheses to new sets of cases. A good example is spelled out in James Mahoney's review of several decades of historical-institutionalist research on the origins and dynamics of democratic and authoritarian regimes. In this literature, research took off from pioneering agenda-setting works (i.e, Moore 1966; O'Donnell 1973; Linz 1978), with waves of succeeding scholars executing dozens of historical and comparative-historical studies that served to refute some arguments, refine others, discover new lines of causal argument, and extend findings across eras and continents. Cumulative research now offers a convincing picture of when, why, and how different types of regimes have emerged in various continents over much of modern world history. "Most of our knowledge about the relationship between democracy and various other factors (social classes, international conditions, elite behavior) has been derived through such step-by-step accumulation . . ." Mahoney concludes. "Plainly, if one were to strike all comparative-historical research from the record, most of what we currently know about the causes of democracy and authoritarianism would be lost" (Mahoney 2002). What Mahoney demonstrates about the literature on democracy and authoritarianism holds as well for literatures on revolutions, state building, and varieties of public policymaking, to name just some areas of clear-cut progress in recent decades. In each of these literatures, historical institutionalists have persistently addressed big questions, while collectively examining and reexamining a large variety of cases.

Of course, scholars in other major traditions also think of themselves as savvy practitioners of cumulative social science, and we would hardly suggest that steady progress can occur only in areas where historical institutionalists predominate. But it is worth noting that sets of scholars who define cumulation in theoretically orthodox terms or who let themselves be captured by their methodological tools can lead us up blind alleys.

Rational-choice scholars often focus research agendas on puzzles internally generated by their overarching theory (Green and Shapiro 1994). There is no doubt that this can at times generate significant scientific progress, highlighting new questions, identifying or explaining new phenomena, and generating linkages among subject matters and research communities that had not been visible before. Research on collective action, growing out of Mancur Olson's seminal work (1971) on public goods, probably provides the clearest example. Outside observers have, however, had more doubts about other prominent examples of theory-derived problems, such as the "paradox of voting" or the relative stability (as opposed to Arrow-style "cycling" dynamics) exhibited in Congress. Each of these "problems" has generated a massive amount of published work, and there is no doubt that the answers provided have become more sophisticated over time. What is less clear is whether these agendas have generated knowledge that speaks to major social problems or would be of significant interest to anyone other than enthusiasts of rational-choice theory. How might these bodies of work fare if we asked James Rule's question (1997) for evaluating research programs: Would someone standing outside your research community be likely to acknowledge the usefulness of the knowledge you are providing?

Methods of investigation rigidly tied to theoretical presumptions can also prompt rational-choice research agendas into overly constricted channels. With game theory generally providing the central analytical tool, studies too easily confine themselves to modeling strategic action at the microlevel. Rational choice practitioners prefer to focus on political settings, such as legislatures, that offer favorable conditions for a particular analytical strategy. They look for coherent strategic actors operating in well-bounded contexts where choices are clearly identifiable and payoffs relatively transparent. Efforts to deal with broader social aggregates, whether interrelated organizations or looser social groupings, are often avoided. Or such challenges are handled by simply treating these groups as themselves being coherent strategic actors—a highly questionable move given rational choice's own assumptions. What is more, rational-choice studies typically assume that all major actors and their preferences are present at the outset of the setting to be analyzed, when in the real world, new actors and changed preferences often emerge in later stages of a linked process. Politics ends up sliced and frozen into artificial moments on the slide of a powerful but tightly focused microscope.

The use of one grand theoretical framework and the decision to focus intensively on a restricted range of political phenomena are often justified, at least implicitly, by what could be called the "lego" rationale. This rationale holds that social scientists should focus on developing solid blocks of findings within a presumed whole. Although the findings may in themselves appear small or distant from pressing social concerns, these blocks can then be pieced together to produce robust answers to important questions. Triviality is acceptable if it provides something to build on. But there are strong reasons to doubt that we should place such a heavy bet on the lego approach. As rational-choice scholars, above all, should know, partial actions do not invariably add up to optimal outcomes. More to the point, where is the evidence that a quarter-century of lego-style research has generated ever-broader findings? In certain areas of political science, arguably, instead of the intellectual problems faced by rational choice becoming bigger, the universe of politics deemed suitable for scrutiny gets redefined in ever more diminutive terms. Thus the study of American politics becomes the study of Congress (or, at its most expansive, the study of Congress and administrative agencies). And the study of comparative politics becomes the study of parliaments and government coalitions. Big questions, broader contexts, and long-term transformations recede ever farther from view, and political science risks cutting itself off from concerns important to broad audiences.

Obsession with a single theory for its own sake is not the only way that cumulative research agendas can go awry. Another danger is capture by technique, or overreliance on a single kind of data. Like historical institutionalists, behavioralists are usually more problem focused. But they risk becoming exclusively enamored with social surveys and the statistical manipulations that can be performed on the answers of random samples of individuals. This can lead behavioralists to neglect other kinds of data that might be relevant to the substantive questions at hand. Tendencies in this direction plague some current work on civic engagement, for example. Why should we presume that everything we need to know about civic participation can be discovered from social surveys—usually confined to one moment in time—that ask vast aggregates of people about their attitudes and reported behaviors? Heavy reliance on social surveys has channeled much of the current debate about civic engagement in U.S. democracy toward exploring attitudes of social trust and indications of whether individuals choose to participate in various ways. But data of this sort is available only for the 1970s and after, and key changes in behavior and institutions happened well before that. What is more, we cannot understand the impact of rising or declining individual memberships in voluntary groups unless we know what kinds of groups have increased or decreased and unless we know how groups interact with institutional centers of decision making.

Questions about an important topic like civic engagement and democracy can rarely be adequately addressed with just one type of data or one technique of empirical analysis.

When it comes to avoiding the dangers of capture by theory, technique, or single data sources, historical institutionalists may have an easier time than scholars in the other major political science traditions. Precisely because historical institutionalists are so riveted by big, compelling real-world problems with clear substantive import to audiences beyond fellow academics, they almost never forget that keeping the eye on the substantive prize—remaining determined to understand important phenomena in the real world—is the formula for enduring scholarly achievement. Because the substantive issues at stake are important in their own right, historical institutionalists are often willing to combine theoretical insights, mine various data sources, and stretch the limits of methodological creativity to gain leverage on those issues (for some nice examples, see McAdam 1982; Rueschemeyer, Stephens, and Stephens 1992; and Schickler 2001). Each scholar knows that his or her interim answers may soon be subject to reconsideration by other scholars (or members of educated audiences) who care just as much about explaining the outcome or solving the puzzle. And real-world developments may intervene to change the definition of important issues or bring new dimensions of them into view. Literatures leavened by many historical-institutionalist participants remain, in short, committed to broad conversations and to theoretical and methodological pluralism. This introduces dynamism and reality checks. If the legos are not fitting together to make at least the foundations of a beautiful structure, historical institutionalists won't keep playing with them forever.

■ | Historical Institutionalism in a Pluralistic Political Science

Addressing questions of interest to educated publics as well as academics, historical institutionalists analyze institutional configurations and develop explanations that highlight conjunctures and path-dependent temporal processes. In this overview, we have stressed and illustrated the many comparative advantages of historical-institutionalist scholarship, drawing on recent clarifications of the distinctive theoretical underpinnings and methodological proclivities of this genre of political science. Still, like other analytic strategies, those employed by historical institutionalists also entail certain costs. Critics often characterize historical institutionalism as mired in mere description and marred by intractable methodological limits. In our view, these criticisms are often misdirected or overdrawn, yet it is true that the focus of historical institutionalists on diverse substantive questions can entail intellectual fragmentation. Even though they share theoretical and methodological proclivities, historical-institutionalist scholars in different subfields may not regularly interact—a situation that (at least until recently) has limited theoretical clarity and cross-fertilization of substantive fields, frustrating steady methodological development. What is more, if historical institutionalists themselves are insufficiently clear about shared strategies, then scholars using other approaches may also fail to notice the theoretically and methodologically relevant commonalities among historical institutionalist contributions addressing so many distinct topics and questions. Important modes of analysis, such as ways of conceptualizing and measuring causal processes that play out over a long stretch of time, may thereby recede from view within the discipline. In this respect, recent critiques have had the salutary effect of prodding historical institutionalists to clarify and develop the methodological and theoretical underpinnings of their work.

Each of the leading approaches in contemporary empirical political science—behaviorialism, rational choice, and historical institutionalism—has proven its value. Political science—indeed social science as a whole—benefits from the coexistence and competition of varied approaches to theory and research. And it benefits even more from dialogue that crosses distinct traditions. Not only can external critiques point to blind spots, prompting improvements in scholarship to the degree that practitioners of each approach are prepared to listen and respond—as historical institutionalists have listened and responded of late. In addition to this back and forth of criticism, multiple approaches can set the stage for creative new blends of methodology and theorizing, especially as new generations of young scholars pick and choose and combine ideas from their elders. Breakthrough studies can combine lines of analysis, generating powerful synergies from the complementary strengths of alternative traditions. Indeed, many breakthroughs have already happened, and while we certainly cannot catalogue them here, we can illustrate some of the possibilities through three excellent examples.

Advancing the agenda for legislative studies suggested in "Legislatures as Political Institutions" (Gamm and Huber this volume), Eric Schickler's *Disjointed Pluralism: Institutional Innovation and the Development of the U.S. Congress* (2001) represents a pathbreaking study of institutional innovation in Congress over the past century. In two respects, Schickler's analysis synthesizes contributions from major research approaches. Theoretically, he tests various functional hypotheses about institutional design drawn from rational-choice scholarship, sorting them out and combining them with the aid of historical-institutionalist arguments about the layering of institutional arrangements over time, and about institutional choices as stemming from interacting processes rather than single processes operating in isolation. Methodologically, Schickler combines the quantitative analysis of roll call votes typical of congressional studies with sophisticated process tracing of a large number of institutional innovations drawn from four time periods (each time period lasting a decade or more). The result is the most serious effort yet to test the strength of a number of prominent theories of institutional choice over an extended time period, and the development of some strikingly original insights into the sources of change and continuity in legislative rules.

Across the globe, new democracies are emerging from previously entrenched authoritarian polities. Explaining variations in new institutional rules and in the processes by which they take shape represents a major challenge for students of comparative politics. In her new book on *Institutional Change and Political Continuity in Post-Soviet Central Asia*, Pauline Luong (2002) compares the establishment of electoral systems in Kazakhstan, Kyrgyzstan, and Uzbekistan and, in the process, develops of new and potentially generalizeable theory of "institutions designed under transitional circumstances," blending insights from both rational-choice institutionalism and historical institutionalism. Although historical institutionalists have directed our attention to persistent legacies from the past, new rules of the game can and do emerge from strategic bargains among elites, especially in a period of crisis and uncertainty. Yet rational-choice approaches too easily fall into the trap of assuming that elite bargaining over new arrangements occurs on a tabula rasa, without regard to entrenched understandings and power relationships. In Jones's model, actors change goals and perceptions in response to uncertainty and bargain in a dynamic way, producing different outcomes in three central Asian polities with many prior structural similarities. But elites work from power positions and understandings embedded in inherited arrangements—indeed they try to encode those older meanings and power relationships into seemingly new structures. Luong's work would not have been possible without prior breakthroughs in game theory. At the same time, she is inspired by the theoretical agendas set out by such historical institutionalists as Kathleen Thelen, calling for careful analyses not just of institutional outcomes but of the temporally and structurally embedded *processes* by which actors entrenched in previous institutions maneuver to create modified arrangements that retain many continuities from the past. Methodologically, moreover, Luong makes a powerful case for close comparison of kindred polities within a geocultural region, as a valuable laboratory for working out explanations of institutional change that may have much broader theoretical application.

In different ways, both Schickler and Luong work at the intersection of rational-choice theorizing and historical-institutional analysis. Our third instance of creative synergy, Andrea Campbell's *Shaping Policy, Shaping Citizens: Senior Citizen Activism and the American Welfare State* (forthcoming) exemplifies possibilities for combining strong points of survey-based behavioral analysis with analytic strategies that have been developed by historical institutionalists. Campbell's subject is a traditional focus of behavioralism: citizen participation. Specifically, she is interested in the changing patterns of political behavior of the U.S. elderly, and she ably uses survey results to document aggregate behaviors and attitudes.

At the same time, Campbell draws two major insights from historical institutionalism. First, she stresses the importance of studying participation over time, both to increase variation in her outcome and to identify and assess the contribution of relatively slow moving factors to changing patterns of participation. Second, she highlights the contributions of policy feedbacks—specifically from Social Security and Medicare—to both the level and style of citizen participation among the elderly. Through the painstaking assembly of longitudinal microdata on participation among the elderly and the thoughtful combination of multiple traditions, Campbell demonstrates how major federal programs both raised the level of elderly participation and produced an unusual pattern of participation. Among working-aged U.S. adults, participation is heavily skewed toward better-educated and higher-income groups, but this "upward" bias is largely absent among the elderly, where concerns about Social Security and Medicare have encouraged middle- and lower-strata elders to participate much more fully. Through its creative combination of behavioral and historical-institutional analysis, Campbell's study speaks to major issues of our time. Its findings are highly relevant to ongoing public as well as scholarly debates about entitlement reforms and about civic engagement in U.S. democracy.

None of the exemplary new studies we have just described could have been produced without imaginative efforts to draw on and combine the best of multiple analytical strategies. By featuring these studies, we mean to underline that it has not been our purpose to pit major approaches against one another in any kind of zero-sum game. We have simply sought to clarify the distinctive contributions and advantages of historical-institutionalist scholarship. Indeed, a good reason to clarify distinctions among approaches and highlight potential comparative advantages of each is to facilitate exchanges that combine the strengths of different approaches to maximum effect. We have made the case for historical institutionalism as one of the three research pillars in contemporary political science, and our task has been made easy by the outpouring of so much excellent work in recent years. Historical institutionalism has become, of late, more self-conscious both theoretically and methodologically, in ways we have indicated and illustrated throughout this review essay. Readers who wonder what political science would be like without self-conscious historical institutionalism need only ask what we would have lost if all the works in our bibliography, and many others like them, had not been published. We think the answer is obvious: without historical institutionalism, our discipline would be shorn of much of its ability to tackle major agendas of concern to all political scientists. And without historical institutionalists, political science would have much less to say about questions of great import to people beyond as well as within the ivory tower.

KAREN ORREN AND STEPHEN SKOWRONEK

The Study of American Political Development

The study of American political development (APD) is an inquiry into the temporal attributes of governance. The United States provides the primary setting for this inquiry; U. S. history provides the primary data; and understanding past politics as accurately as possible is a primary concern. But finding out "what happened" in American political history is not the ultimate objective. Rather, practitioners of APD seek knowledge of how governance changes over time; they are interested in specifying the processes by which political innovations are negotiated and new political relationships generated. The aim is to understand these things conceptually, that is to say, in terms that illuminate general features of the polity, over a range of historical instances. Of special interest are dynamics of governance in the past as they affect practices of governance in the present or shed light on its future prospects.

Intellectual currents affecting disciplines closely related to APD research have served in recent years to highlight the field's distinctive concerns. It is often noted, for instance, that APD came into its own as a subfield in political science departments during the 1980s and 1990s, when the study of American political history had lost much of its intellectual currency within history departments (Leuchtenberg 1986).[1] Government and politics, long a centerpiece of the historian's understanding of the United States's past, found itself in the position of a foil against which alternative interests gained ascendancy. Political history was first challenged by social historians who refocused attention on economic power, class relationships, group formation, and demographic trends. Then came the cultural turn in historical writing with its call for narratives of Ameri-

1. Of course, political history never disappeared and there are today encouraging signs of its resurgence. For the purposes of this essay, however, we have concentrated on the work of political scientists. Readers should not read our essay to imply that political scientists have been unaffected by the ongoing work of political historians or to deny the vibrant exchange that has been established between politically minded historians and historically minded political scientists.

cans who were not white, male, or otherwise favored by the state authority and office.

In political science, where government and its public ramifications are more firmly lodged as objects of study, the rise to prominence of alternative perspectives on the past had a different effect. Rather than crowd out interest in high politics and formal institutions, they opened it up. The discovery and exploration of alternative sources of power and authority spurred political scientists to rethink the operations of government and the dynamics of political change. Moreover, when political scientists turned back to history, they did so with the methodological debates of their own discipline in mind. Studies of American political development called attention to interactions over time that were difficult to square with the prevailing assumptions of pluralist theory and behaviorial research. Historical approaches, which had been mainstays of the study of constitutional law and political philosophy since the inception of political science as a discipline, were reworked and redeployed to challenge received conceptions of the U.S. state (Skowronek 1982), of the relationship between governing institutions and political action (Huntington 1981) and of the proper angles of research into state-society relations (March and Olsen 1989).

By joining traditional interests in high politics and formal institutions to the wider array of concerns now understood to attend questions of power and authority, APD research has begun to illuminate the distinctive political problems that arise within the categorical realm of rules and operations to which individuals are expected to conform under threat of legal sanction. The study of changes in governance now ranges from investigations of Congress and the Supreme Court to investigations of workplace and family relations; problems of political development are situated in a variety of arenas, public and private, and within institutions that exhibit varying degrees of formal organization. In each domain, analytic attention has focused on the operation of rules that are enforced by socially legitimate authority designated to perform that task. Recognition of the diversity of these rules and their disparate institutional foundations underlies some of the most important questions now being asked, questions about relations between the whole and parts and how transformations in one arena of governance bears on the enforcement of rules elsewhere.

A second intellectual current that has highlighted the distinctiveness of APD's agenda is the collapse of faith in the sufficiency of development as a premise for historical inquiry in the social sciences. Before this, what conceptual apparatus there was for thinking about American politics over time was drawn from theories of world historical forces acting to modernize Western nations. Hegelian idealists, Marxist materialists, Weberian institutionalists, American progressives, modernization theorists—all assumed that progress onward and upward, toward one state of affairs or another, was history's predominant attribute. The rejection of this assumption and the questions that have come to surround the tradition of social theorizing built on it made American political history a field ripe for revision.

That APD came into its own at a time when the central premise it names had been widely discarded in part explains the intensity of the search that has been set in motion for theories and interpretive frameworks that might relate past and present in more compelling ways. The subfield continues to attract scholars convinced of the enduring value of historical strategies for situating American governance comparatively, for identifying common features and particular configurations and for considering the interaction over time of shared elements combined differently in different locales (Katzenelson 1997; Pierson 1994; Evans, Reueschemeyer, and Skocpol 1985; Bright and Harding 1984; Sheingate 2001). What these scholars have carried over from the tradition of theorizing devoted to uncovering natural laws of development is the intuition that one essential, constitutive dimension of governance is the passage of time itself and an accompanying commitment to rethink its implications. But APD research also evidences the challenges that now confront theorizing along these lines and the spadework that needs to be done if we are to formulate compelling explanations of how a nation's past bears on its subsequent politics.

The broadside critique of development unhinged settled notions across the board, not only about politics in the United States but also about politics in the states to which it had traditionally been compared. Relieved of teleological baggage, the search for meaning in change over time has moved closer to the ground and become more attentive to nuance and complication. It has become a more-contested inquiry as well, a forum for debate over whether anything worthy of the term *development* in fact occurs in politics, over what development might be composed of, and over the conditions under which it might happen. With the *d* word under scrutiny, APD has become as the seedbed of a general reconsideration of the character of political history itself.

A final current distinguishing the APD agenda today has been the rise to prominence of rational-choice theory within political science. Rational choice is a deductive theory whose main advances to this point are seen in the elaboration of microfoundations for analyzing political events and in the modeling of comparative statics. Explaining change is a central objective, but change is normally stylized by practitioners for game theoretic purposes, appearing as an exogenous shock to a stable state that induces the search for a new equilibrium of preferences. In this respect at least, the rise of rational-choice theory in political science may be regarded as of a piece with contemporary disillusionment with historical forces as a basis for political theorizing. As a testing ground for the theory, however, history remains inescapable, and increasing numbers of scholars working within the rational-choice tradition have turned to the past to demonstrate and refine the power of their models.

As we shall see below, the historical turn in rational-choice theorizing has opened new channels for collaboration in APD research. The grounds for such an exchange are clear: Political development necessarily concerns the acts of individuals, and APD and rational-choice research share certain assumptions about individual behavior—that individuals are on the whole rational in pursuing goals, that they act within structures and according to rules they understand, and that they have limited responses to coercion. But notwithstanding the potential for mutual advancement on this ground, a chief product of the exchange thus far has been a clarification of the different agendas that political scientists bring to the study the past. Rational-choice theorists approach political history with a set of propositions assumed to be independent of time and place and thus portable in their application to discrete instances. In this way, the past becomes a reservoir of examples. Research in APD assumes, in contrast, that time is constitutive of politics and that direct examination of dynamics emblematic of political history is one of the prime prerequisites for theory building. Other differences follow. Rational-choice theory offers snapshots, episodes of change in exemplary arenas; APD unreels a "motion picture" of governance, its parts' repositioning with diverse rhythms and mechanisms of coordination (Pierson 2000a, b). Rational-choice applications have sought to engage historians in debates about discrete political episodes (Bates et al. 1998); APD scholarship tends to draw historians into a dialogue about how periods relate to one another and how discrete episodes might be connected.

More will be said about each of these currents as we proceed. Suffice it to say here that APD is pursuing its own lines of inquiry among them. The claim that governance in its various aspects exhibits distinctive dynamics over time and that these are best examined over the long term touches a wide range of subject matters and engages scholarship in a number of specialties. Rather than assume a priori one set of dynamics everywhere or extrapolate from some related discipline to the realm of governance, APD has begun a thoroughgoing reassessment of how politics is arranged empirically at given points in time and how politics moves through time in a place. With long-standing assumptions about these matters exposed and interrogated, APD has emerged today as a field of considerable intellectual ferment and creativity.

▪ | Culture

One important line of APD research pivots on Louis Hartz's argument (1955) that the United States "has always been a place where the common issues of the West have taken a strange and singular shape." Prior to Hartz, discussion of APD had followed the assumptions of progressive historians in directing attention to the common pressures that economic interest, industrialization, and class conflict exerted on political development throughout the West. Hartz challenged that approach, citing insufficient attention to the peculiarities of American political culture and to the distinctive ways that common pressures played out in particular settings. Recent scholarship has shown that this line of criticism is not easily contained. As contemporary scholarship has broadened and deepened inquiry into the "strange and singular" features of American political culture, it has come to find Hartz's own work wanting, too caught up in the primacy of common influences and the elusive standards of development that lurk behind them.

In fact, Hartz built on what orthodox theories of class conflict had said about developmental problems and prospects. His innovation was to present the United States as an exception to these rules, one that he explained with reference to Tocqueville's insight that Americans had never endured a bourgeois revolution against feudalism but had instead been "born equal." Hartz called attention to the fact that the American puritans had removed themselves from the struggle of liberal principles against the reigning doctrines of the English aristocracy. In so doing, he conjectured, they had bypassed a critical stage in the political development of liberalism itself. Hartz theorized that America's middle-class fragment, having been transplanted from its origins in a larger contest of worldviews but already aligned with the political sympathies of their coreligionists, gave rise to a peculiar political culture, one in which the values of individualism, limited government, and equality were so widely shared as to become irresistible to all serious contenders for political power. A "genuinely revolutionary tradition" never took hold in this environment, he argued, nor did a "genuine tradition of reaction," for never was there a need for American liberals to dislodge an antithetical establishment, to replace one set of ideas with another, or defend their own creation against alternatives. Hartz depicted a polity that acted without any consciousness of conflict as a battle between class antagonists over fundamental principles, and he concluded that, in the absence of class consciousness, the economic antagonisms routinely produced by capitalist development elsewhere were shorn of their potential to generate radical change. Having skipped the middle-class revolution against feudalism, Americans became impervious to the idea of a social revolution against capitalism.

Hartz's "consensus" thesis identified American "exceptionalism" with radically truncated cultural boundaries that constrained political thought, political action, and ultimately political development. In fact, Hartz's analysis went far toward rendering the whole idea of American political development an oxymoron. While he acknowledged pervasive and persistent conflict in American politics, he found no juncture at which these conflicts conjured political alternatives or aspirations robust enough to produce anything genuinely new or different; the United States's political struggles appeared to him mere "shadows" of the momentous struggles played out elsewhere on the world stage. Writing at the height of the cold war, Hartz expressed concern about the anachronistic character of American politics, in particular about problems Americans would face in grappling with the political alternatives with which they were now forced to engage. The limits of their capacity to understand anything beyond the narrow confines of their own experience might, he thought, lead them to exaggerate the threat posed by internal dissent and to underestimate the appeal that alternatives held elsewhere in the world.

With the end of the cold war and the emergence of United States unrivaled on the world stage, the concerns which Hartz expressed about American political development appear about as relevant as the predictions of Karl Marx (Foner 1984b). That said, the new tone in APD research has yet to strike a triumphalist chord. If anything, recent work raises even more serious questions about the proposition that the United States is the vanguard state of political development in the world. Consensus theory has been challenged in the very same terms that consensus theorists themselves challenged prior thinking about APD: it is charged with being too Euro-centric, too economistic, and insufficiently attentive to what makes American politics tick (Smith 1993). Scholars are far less willing now to leap from the relative absence of feudalism and socialism in the United States to the insignificance of ideological differences here or to ignore the choices that have divided the nation and shaped their political history. In place of Hartz's portrait of the United States's liberal naïveté is a new appreciation of the multiple dimensions of political culture and of the pivotal character of the struggles that ensue from the different political visions it has supported. In looking for standards by which to evaluate American political development, today's scholars are less preoccupied by the alternatives generated elsewhere and more attentive to those that have been generated from within (Norton 1986; Ellis 1993).

The immediate concern of the literature on the United States's alternatives is to reclaim the political ideas put forth by the United States's diverse social movements from the dismissive presumptions of American political consensus. To this end, it has undertaken a thoroughgoing reassessment of the programs that ultimately lost out in contests to redirect the course of American politics. The revised narrative of political change in the United States now begins boldly with insurgents who bore robust programs for change and were fully conscious of the fundamental choices being adjudicated in the political decisions at hand. Its conclusions are all the more sober, detailing the sidelining of the insurgent's signal claims and the refortication of their opponents. We have gained from this literature a sharper picture of the Progressive's moral vision and how it was lost (Eisenach 1994); of the "maternalist" ideal of U.S. welfare state and why it was scrapped (Skocpol 1992); of the agrarian theory of the modern political economy and how it was coopted (Sanders 1999; Ritter 1997; James 2000), of the indigenous radicalism of labor's commonwealth ideals and why they were aborted (Hattam 1993), and of class demands as they came to be expressed and organized under conditions of expanding manhood suffrage (Bridges 1984).

Beyond its vigorous reassertion of the generative capacities of American political culture, this work compels attention for two general propositions advanced against prior thinking about American political development. One is that the alternatives that lost out promised to move the nation in a starkly different direction from the ones that actually took hold; the other is that these were plausible alternatives for the United States. As the first proposition dispels the illusion of an overarching political consensus within which choices about development are relatively inconsequential, the second challenges the aura of inevitability and progress that history's winners tend to attribute to their own victories. Indeed, by refusing to diminish these alternative tracts as intellectually shallow or hopelessly backward looking—by treating them as coherent, intelligent, and practical—this literature has called into question all claims of special insight into which ideas, groups, or movements represent the true forces of history (Berk 1994). The effect is not only to displace the consensus view of American political history with fundamental conflict but also to strip the conflicting visions in play of the teleological projections that previous work on APD had brought to the study of change.

There remains the irony that by reclaiming the radicalism and practicality of alternatives that were ultimately frustrated, this literature leaves us in the end with a compromised path of development not all that different from the one Hartz described. As we shall see in a moment, this problem has prompted others to push inquiries into cultural multiplicity in the United States in other directions. But the response found within the literature reclaiming lost alternatives is worth considering on its own terms, for notwithstanding Hartz-like results, the explanations provided here are quite different. With politics from below found to be more authentic and coherent than previously thought, scholars have had to look to arrangements above to account for why these aspirations failed to materialize. Debunking Hartz's charge of a culturally constricted vision, this literature has substituted an analyses of political defeat and, in effect, projected a gaping disjunction between the products of the U.S. state and its full-blown political opposition. Each instance points to a gross imbalance in the distribution of political resources and stiff institutional hurdles that all but foreclose certain courses of action. In other words, it is not consensus that stifles political alternatives but governmental arrangements and elite advantages. American political development now appears to have been constrained not by ideology but by a state apparatus which, for all its divisions and internal conflicts, has absorbed, dissolved, and deflected alternative visions with remarkable consistency.

As suggested, not all efforts to push beyond the Hartzian consensus have proceeded in this way. A different approach to studying the relationship between political culture and political development was suggested by the late J. David Greenstone (1986). Greenstone dismissed the lost-alternatives critique of Hartz, accepted the notion of a pervasive consensus on liberal values in the United States, and urged instead a more direct consideration of possibilities for development within that value system. The central deficiency in Hartz's analysis, Greenstone charged, lay in its inability to account for the emergence of any new ideas affecting the operations of American government. By extension, he saw Hartz's theory as prone to understate the real transformative potential of the ideas Americans had at their disposal. Greenstone pointed out that Hartz had not provided a causal theory of change, that the liberal consensus merely specified a peculiar political universe, a set of boundary conditions within which change in American politics had occurred. To explain change, he recommended an investigation of the ways in which new ideas are constructed from the discursive categories of a common culture, that is to say, from its linguistic tools.

To demonstrate, Greenstone examined the debate over slavery in the northern states of the union during the middle of the nineteenth century, a debate about which Hartz had said next to nothing. In *The Lincoln Persuasion: Remaking American Liberalism* (1993) he examined the words *liberty* and *Union* as these were used by politicians representing all the contending positions. He showed not only that these concepts admitted multiple meanings in different relations to one another but that fundamental differences among those meanings were elaborated in the very process of debating them. Some usages drew upon a secular humanist strand in U.S. liberal thought, employing a negative conception of freedom as an absence of constraint to conceptualize the relationship between these values; others drew on a Puritan, pietistic strand with strong moralistic and positive entailments.

It was Lincoln, Greenstone argued, who found in the United States's storehouse of ideals and aspirations a combination of meanings that focused the fundamental conflict of values gestating within U.S. society at the time, a conflict *within* the United States's liberal culture that, once articulated, could not be deflected, absorbed, or dissolved. Greenstone readily granted that when Lincoln conceptualized the Union as a moral force whose ultimate purpose was to facilitate individual self-improvement, he was working wholly within the boundaries of the established culture. At the same time, he insisted that Lincoln was articulating a radical political alternative, one categorically different from, and far more expansive than, that produced either by Jacksonian understandings (in which the Union was merely an instrument for maximizing the freedom of white men to do as they chose) or Whig understanding (in which liberty derived from the Union and had to be regulated in ways consistent with its preservation). In fact, Lincoln's alternative produced something wholly unaccounted for in Hartz's analysis: a normative revolution that carried a triumphant political revolution in its train.

The claim that the United States's liberal discourse has sufficient nuance and dimension to develop of its own accord has an intuitive appeal. Others have argued in a similar vein that liberalism has not merely been rearranged or readapted from time to time but substantively reinvented and conceptually expanded by people driven to think anew about the meaning of common values (Rodgers 1987; Ericson 2000). But as Greenstone's meticulous argumentation suggests, we have come a long way from any easy invocation of that premise, and on inspection, the claims he made on behalf of these prospects were quite modest. His argument was about a toolbox of legitimate concepts and ideals. He ventured that the United States's stock of liberal ideas can be and has been augmented through interactive processes of debate, that we have not been stuck with the ideas with which we began, and that new ideas, once brought to power, are available for elaboration in the future. Little was said, however, about how closely liberalism's conceptual expansion informs the course taken in American politics, about whether these expanded liberal ideas, even when they seize control of the state and change the Constitution, actually determined the subsequent course of public policy.

In fact, while Greenstone's amendment to Hartz allows for liberalism's advance, the argument for linguistic inventiveness can as easily cut the other way. In a culture that is robust enough to conceptualize alternatives and carry them to power, no alternative is really secure; an expansive idea may dominate at one moment only to be countered and undermined by a newly constrictive conception in the next. The fate of the Lincoln persuasion appears to be a prime case in point. As it happened, the vanguard commitments of the Civil War era on behalf of civil and political rights were progressively abandoned as energies for reform waned and the South was reintegrated into national politics. Apparently, even empowered alternatives can be lost.

This is precisely the point made by Rogers Smith in *Civic Ideals: Conflicting Visions of Citizenship in U.S. History* (1997). While Greenstone criticized Hartz for giving too little credit to American liberalism's developmental potential, Smith criticized Hartz for giving too much. Smith's challenge took direct aim at the proposition that liberalism has been hegemonic in determining the course of American politics. In his analysis, the United States's most potent alternatives have neither been lost nor liberal; he demonstrated that nonliberal and downright illiberal traditions have been prominently displayed in American politics and occasionally ascendant in the high affairs of state. *Civic Ideals* documented the persistence in American politics of cultural values supportive of patriarchy and racism, and it showed that advocates of these values have been politically effective in competing with the advocates of expanded rights for the allegiances of the American people. Smith's book detailed counterarguments promulgated on behalf of these alternatives to pace liberalism's advances and call its guiding assumptions into question, and it identified important domains of state action dominated by these contrary claims for extended periods of time. By dubbing these alternatives traditions in their own right, Smith meant to attribute to them a degree of coherence, distinctiveness, resonance, and staying power comparable to the liberal tradition which Hartz had identified with the whole of the American consciousness.

Though assaults on liberalism in our day are never far from view in his work, Smith directed attention to the late–nineteenth century to illustrate just how potent these alternative traditions can become in redirecting American politics. The about-face negotiated after the Reconstruction on matters relating to the civil and political rights cannot, in Smith's reading, be dismissed as an aberrant loss of cultural bearings, as a momentary hiatus in liberalism's march of progress, or as a time-lag in perceptions of the disparity between ideals and reality; it was rather an emblematic moment in which an alternative set of ideals gained the upper hand in American government. Documenting the success of prominent elites, inside the government and out, in reworking counterthemes with powerful claims on the hearts and minds of the American people, and ultimately, in legitimating a substantial reversal of rights formally achieved and exercised in the 1870s, Smith upended the proposition that American political development has been a series of incremental advances on liberal premises.

By exposing unwarranted assumptions of uniformity and linearity in American political development, Smith accelerated the unraveling of developmental premises that Hartz's consensus thesis had already done much to weaken. The discovery of multiple traditions that are fundamentally at cross-purposes shifts the burden of developmental thinking off the stifling hegemony of liberalism in American politics and onto the tentative status of any liberal, or for that matter nonliberal, change that might happen to occur. The specification of illiberal traditions that can drive potent political movements and control significant sectors of public policy presents the United States as a polity susceptible to swing back and forth between competing standards of fundamental value, a polity prone to move over time in opposite directions and to repeal liberal advances once attained. It is difficult to imagine traveling much further away from the developmental conception of American politics without calling into question the enduring significance of liberal ideals altogether.

Periodization

Another line of research in APD may be traced back to V. O. Key's proposition that American politics is periodically punctuated by "sharp and durable" changes. "A Theory of Critical Elections" (Key 1955) was published the same year that Hartz published *The Liberal Tradition in America*. Beyond that, however, the two works might seem to have little in common. While Hartz drew attention to the uniformity of the American political experience and explained why the New Deal had changed so little, Key pointed to large-scale discontinuities like that which ushered in the New Deal and suggested that American politics has not been all of a piece. Specifically, Key argued that some elections in American political history were far more important than others in bringing about political change, and he suggested that these "critical elections" might be the missing link between the everyday operations of democracy in the United States and larger dynamics at work in the transformation of the political system as a whole.

Walter Dean Burnham, who studied at Harvard while Hartz and Key taught there, saw a connection between their two perspectives. Burnham proposed that critical elections were the characteristic product of the United States's exceptional political culture, and with that he set out to resolve a number of riddles implicit in the consensus thesis. If, as Hartz theorized, American politics did not develop in any meaningful sense, what exactly did it do? By registering only the polity's most timeless and encompassing features, consensus theory failed to convey any clear sense of what distinguished one period from the next. Furthermore, the theory failed to explain how a government fit for nothing so much as the preservation of its founding principles could successfully adapt to the rapid pace of change in its economy and society. Hartz had used Marx to identify crucial elements of class consciousness missing in American political culture, but he had failed to contend with Marx's understanding of capitalism itself as an unrelenting engine of social and economic transformation.

Burnham's theory of critical elections addressed these questions by positing a systemic, causal connection between persistence and change in American politics. *The American Party Systems: Stages of Development* (Burnham and Chambers 1967) and *Critical Elections and the Mainsprings of American Democracy* (Burnham 1970) cast electoral realignments as political surrogates for social revolution in a polity culturally impervious to social revolution. At the heart of Burnham's analysis was a Hartzian gloss on the United States's "Lockean cultural monolith," evidenced first and foremost by a Constitution designed to preserve prior consensus and inhibit the power of particular groups to make changes. Burnham proposed that in such a system the new demands on government routinely produced by social and economic change were bound to fester, that as the Constitution intercepted and delayed adaptations to the changing conditions of governance, relations between state and society would drift toward moments of extreme national stress. The characteristic products of this stress were major electoral uprisings, critical elections or election sequences that realigned the coalition base of the contending parties and compelled American government, one way or another, to accommodate the new realities. Such disruptions were likely to be rare and episodic, for the mobilization required would be difficult to sustain. The return to normalcy would feature the new alignment of interests secured by constitutional divisions, and the beginnings of a new cycle.

The orthodox version of this theory highlighted five such transformations coming at intervals of roughly thirty years and centering around the elections of 1800, 1828, 1860, 1896, and 1932. These framed five extended periods of political order, or regimes—the Jeffersonian, the Jacksonian, the Republican, the Progressive, and the New Deal. The theory affirmed that each of these orderings reflected broad cultural commitments to democracy and capitalism but suggested that each also defined a distinctive universe of political and institutional action. They featured different party systems with economic, sectional, and ethnoreligious interests differently divided into national political coalitions, they each brought a different mix of policy debates into government, and each offered economic interests different kinds of institutional services and supports.

Significantly, Burnham himself did not find any particular political trajectory in the movement from one political regime to the next. While he allowed for constitutional changes of the first order and described a mechanism that would periodically reconstruct the relationship between democracy and capitalism in the United States, he was even more explicit than Hartz in raising questions about progress. He dwelled on the realignment of 1896 in particular because he saw in that shift to a more starkly sectional electoral alignment the demobilization of a once-vibrant democratic politics and the insulation of corporate control over processes of industrial consolidation. "When the conflict between industrial capitalism and pre-existing democratic structures came into the open," he charged, "it eventuated in the displacement of democracy, not industrial capitalism" (Burnham 1970, 187, 1965). *Realignment* referred not to development but to change; it identified a tension release mechanism through which the constituent elements of American political culture periodically reconstituted themselves.

Realignment theory transformed the discussion of state-society relations in the United States, introducing new conceptions of time, change, sequence, and regime. Its periodization scheme was widely recognized in the 1970s as the "dominant conceptual picture" of American political history (Bogue 1980), and it inspired literatures not only on electoral politics but on public policy and governmental structure as well (Brady 1988; Ginsberg 1976; Shefter 1978). Its impact is still evident today in a wide range of studies—on party formation (Aldrich 1995), regime formation

(Plotke 1996; Polsky 1997), presidential politics (Skowronek 1993), American political thought (Eisenach 1990), labor politics (Mink 1986), and constitutional law (Ackerman 1991). And yet, recent research, often the very same research that is informed by the periodization scheme, has also become circumspect about limitations of these demarcations and attentive to the value of alternative formulations.

No doubt, the major catalyst to new thinking about periodization has been the elusiveness of realignment since 1932, in particular, the difficulty of squaring classic realignment theory with the twisted and halting course of American politics since 1968 (Shafer 1991). But there are other factors at work as well. The critique of Hartz has played an important role. With more attention's being paid the ideological divisions that have been expressed within American politics, scholars have discovered continuities and breakpoints that don't fit the standard electoral periods. John Gerring, for example, finds that competing party ideologies in the United States have not changed in lockstep with electoral realignments; his periodization of the alternatives offered by mass political organizations in the United States not only crosscuts the realignment divides but also suggests distinct dynamics as work within the different party organizations themselves. Another recent study has taken aim at the claim that the Republican victory in 1896 resolved the most contentious questions about American industrialization by foreclosing effective political challenges to corporate consolidation. Elizabeth Sanders (1999) finds that the agrarian reform agenda not only survived the realignment but refortified itself thereafter. Documenting the implementation of that agenda during the Wilson administration, she argues for extending the moment of the populist insurgency from the 1890s through the 1910s.

Other researchers have been impressed by the changing role of parties themselves, in particular by the secular decline in their relative importance as governing instruments over much of the twentieth century (J. J. Coleman 1996; Milkis 1993). Some work along this line has suggested that the major divide to be drawn in American political history is between the nineteenth century, or party period—when parties were the dominant organizing instruments of governance—and the twentieth century, or bureaucratic period—when administrative instruments began to assume prominence (McCormick 1986; Silbey 1991; Skowronek 1982; Balogh 1991). By calling attention to a twentieth-century shift in the underlying mode of governmental operations, this periodization points to important continuities in American politics before and after the Civil War and before and after the New Deal. It also helps explain why realignments of the sort found in the nineteenth century have been so rare in the twentieth.

Still other challenges have targeted the transformative political significance of seemingly unrelated events—wars, for example—that occurred between realignments (Mayhew 2000; Kryder 2000; Sparrow 1996) or called attention to the critical choices that party realignments left unresolved and contested (Katznelson and Pietrykowski 1991; Katznelson, Geiger, and Kryder 1993; Kryder 2000). At the heart of all of this new thinking about breakpoints and continuities is a recognition of the multiple dimensions of political history, of the difficulties raised by trying to fit everything of significance into a single periodization scheme, of the dangers of simply dichotomizing "normal" politics against transformative "moments." With the explication of alternative schemes, the idea that there is one best periodization has begun to dissipate, and the competition to find out which of those before us synthesizes the most information has lost much of its appeal (but see Burnham, 1986, 1994). In place of the search for a new synthesis, an agenda of sorting out different temporalities at work in political history has emerged (Jillson 1994). By specifying the institutional foundation of each dynamic identified and by considering its operation in light of others, APD scholars have begun to grapple with political interactions among elements of governance differently constituted by time and to employ more complex models of political change in advancing their analytic program. The emergent insight of overriding significance is that different temporalities operate simultaneously in the political construction of governance.

Consider in this regard two recent works that both reflect and delimit the significance of critical elections. In *Presidents, Parties and the State,* Scott James (2000) asked why it was that the central issue contested in the realignment of 1896, the legitimacy of corporate capitalism, was resolved by the losing party to that conflict some two decades after the event. One part of his explanation is that, rather than change in lockstep, the different institutions of American government moved by their own rhythms and at their own pace. In this instance, the Supreme Court, operating with members previously seated, issued a ruling in 1897 that, under the Sherman Anti-Trust Act of 1890, all restraints on trade, not just unreasonable restraints, were unlawful. This decision opened a glaring incongruity between national law and the corporatist commitments of the victorious Republican Party on the central issue which the realignment had presumably resolved. It was not until 1911 that a Court newly constituted to reflect dominant Republican thinking on trusts reestablished the common law standard of reasonableness, but by that time, the issue of how to respond to the radical position staked out by the Court on corporate power had so divided the Republican ranks that they had lost their majority in the political branches to the Democrats, a party committed to a radically anticorporate program of action (see also Sklar 1988).

The second part of James's explanation turns on the operations of the electoral college, and the plight of a minority party in power as it attempted to implement its opposition program under threat of imminent displacement by the reunification of the party system's dominant member. James argues that the national leadership of the Democratic Party, eyeing with trepidation the presidential election of 1916 and desperate to attract coalition allies, gradually pulled its rank and file in Congress toward a position on the trusts that they avowedly opposed. This dynamic ultimately ratified the legitimacy of corporate capitalism and with it the logic of the system of '96, but, James notes, not before new antitrust laws had been enacted and U.S. corporate law had been stamped with a competing set of values.

In *Party and State in America's New Deal*, Kenneth Finegold and Theda Skocpol (1995) compared two emergency programs enacted in 1933, that is, directly in the aftermath of another critical electoral upheaval. As the programs were backed by the same coalition in Congress and aspired to the same aim of raising prices and incomes in their respective sectors, the comparison would seem to be calculated to hold historical variance at bay; and yet, the authors observed that the Agricultural Adjustment Administration (AAA) and the National Industrial Relations Board (NIRB) experienced results that were not only disparate but also in the opposite direction from what strictly sectoral—that is, nonhistorical—factors would have predicted. The farmers should have presented a more difficult problem of coordination than the occupants of workplaces, but in practice, the AAA achieved far higher levels of successful mobilization than the NIRA. To explain this outcome, the authors pointed to the contrasting administrative histories of the two policy domains. The AAA operated within a policy network with origins in the Civil War era and progressive reform. Veterans of earlier federal initiatives in agriculture already possessed a good grasp on the problems to be addressed and the available solutions, and the government officials dealing with these issues were connected by education and experience to their counterparts in private cooperatives and land grant colleges. In contrast, NIRA administrators proceeded on a ground pitted during the past years by private suspicion, public irresolution, and an adverse tradition of antitrust. Understood in this way, the electoral realignment of the 1930s was a multifaceted event, one whose impact varied in accordance with the prior history of reform in the different policy domains it brought up for reexamination.

The question of periodization underlies all assessments of continuity and change in politics. Martin Sklar (1991) has argued that periodization is the historian's method of specifying testable hypotheses about order in human relations and the causes of their transformation, that it is the foundation of the science of history. In taking up the challenge of periodization, recent APD research has added a significant twist. Rather than simply replace one set of breakpoints with another, it has begun to question the underlying conception of American politics as an ordered whole punctuated over time by moments of radical change (Orren and Skowronek 1994). Reckoning with alternative periodizations and the competing temporalities of governance behind them points to a multiple-orders hypothesis, wherein change proceeds through the push and pull of differently constituted elements simultaneously engaged. Considered next to Smith's multiple-traditions thesis, this multiple-orders hypothesis is suggestive of some overarching themes emerging in APD research today. Before turning to those, however, other lines of research need to be brought into view. The notion of multiple orders rests on a more decidedly institutional view of politics than multiple traditions, and as it turns out, the study of institutions claims the lion's share of attention in APD research today.

■ | Institutions I: Interest Mediation

The current preoccupation with institutions in APD research follows directly from the field's interest in reexamining the temporal attributes of governance. Institutions are the premier instruments of governance, durability is one of their leading features, and this tendency to persist engenders distinctive political relationships over time. Moreover when institutions do change, they do so in ways that invite rigorous empirical analysis. They are built, dismantled, and rearranged in ways that are readily perceived and easily specified, thus offering researchers a reliable standard of reference in plotting patterns of continuity and change.

If there is a single unifying question behind institutional research in APD today it is about how institutions construct politics—how they shape action, conflict, order, change, and meaning. But this broad concern admits several different lines of investigation. At least three versions of the constructivist thesis can be distinguished in the contemporary literature, each of which elaborates a different conception of the political significance of institutions. We take these up in turn, beginning with the more familiar ones and proceeding to the less so.

Most familiar of all is the idea that governing institutions construct politics by mediating conflicts of interest within society. This notion hews closely to the traditional association of government with institutionally induced political order and stability, an association expounded on at length by practical statesmen like James Madison and Martin Van Buren during some of the United States's most important institution-building episodes. In revisiting this idea, APD scholars have paid special attention to the precise mechanisms through which order has been created and sustained and to the political contingencies that underlie it.

One study along these lines explains sectional comity during the antebellum years with reference to Congress's balance rule for the admission of slave states and free states to the Union (Weingast 1998). Documenting the potential for the empowerment of a purely Northern, free-labor majority long before one actually appeared, Weingast observes that "nothing inherent in the antebellum era inevitably preserved rights in slaves." More to the point, he argues that with the more rapid expansion of the North in both population and territory, neither the arrangements of the Constitution (the Madisonian solution) nor the arrangements of the party organizations (the

Van Buren solution) sufficed in themselves to induce sectional comity institutionally. Central to this argument is the fact that the convention of sectional balance in state admissions had taken hold early on; it was the product of a time when regional economic prospects were more evenly matched and representatives of the North and South perceived themselves as equally vulnerable to one another in matters of material interest. Incorporated into the Missouri compromise of 1820—before the institutionalization of national two-party competition—adherence to the rule became over time an increasingly important signal to Southern slaveholders of a credible commitment by the North not to press its growing political advantages. Weingast's analysis explains the antebellum political order with reference to Northern legislators' willingness to acquiesce to the rule. It was by granting the South an institutional veto over national programs in this way, Weingast argues, that sectional comity and national stability were sustained long after the balance of economic interests that had given rise to it had changed.

As Congress is the branch of American government least insulated from the direct influence of social and economic pressures, congressional rules, norms, and structures bear a heavy burden in explaining how order is constructed out of conflicting interests. But mediation effects are not limited to the Congress, and the Constitution's structure of coequal branches suggests processes that are far more intricate and broad ranging. A recent study of the political architecture of U.S. industrialization has shown just how elaborate these processes can get. In *The Political Economy of American Industrialization* (2000), Richard Bensel asks how the nineteenth century's most-advanced democracy could sustain a course of rapid industrialization when the policies essential to it were fiercely and persistently contested by powerful political interests mobilized in opposition. His answer identifies rapid industrialization as an institutional construct, one produced by interactions among all the various branches and levels of the constitutional system.

Bensel explains this construct with reference to a contingent alignment of triplets: three government policies were essential to the outcome observed—maintenance of an unencumbered national market, maintenance of the gold standard, and maintenance of the protective tariff; contests over these policies were structured by three principle interests—cotton exporters, yeoman settlers, and financial capitalists; and the antithetical designs of these interests were grounded in and politically bolstered by their geographic concentration in three distinct regions—the South, the West, and the Northeast respectively. His model explains how the peculiar tripartite structure of American government interacted with the peculiar tripartite structure of the late-nineteenth-century economy to foster rapid industrialization when that outcome by itself was supported by no social or political consensus, let alone a natural foundation.

As it happened, each branch of American government operated to support one of the three critical policies and to thwart alternatives supported in the others. The Court worked to secure the interests of both national party coalitions in the maintenance of the open market against regulation-prone majorities from the South and West and the incursions of the more populist state legislatures. Congress worked to secure support for tariff protection against Grover Cleveland, a Democratic president committed to freer trade: the legislature logrolled benefits in such a way as to divide support in the South and West for lower rates and to tie the interests of northern labor and industrial capital to the program of finance capital. Finally, the electoral college system of presidential coalition building worked to produce chief executives friendly to New York financial interests and willing to protect those interests in the gold standard against inflationary interests in the South and West which claimed congressional majorities at critical junctures.

Bensel presents a powerful case for the socially constructive impact of a far-flung institutional politics. Rapid industrialization appears in his analysis as a wholly mediated effect, the synthetic product of interactions within and among institutions each of which was motivated in a different way to deal with interests pressing in on it from the outside. A central issue in the study of the institutional construction of politics today, however, concerns the limitations of this outside-in perspective itself. In both the Weingast and Bensel studies, interests press in on government, and government officials fashion responses in accordance with their own rules and motivations; but in neither of these analyses do institutions appear to alter the interests themselves. The concept of mediation contemplates the institutional construction of policy responses to interests, not an institutional construction of interests themselves.

A study by Terry Moe (1987) of the institutional construction of order in relations between business and labor in the middle of the twentieth century breaches this boundary. Moe tied the equilibrium observed in this period to two factors. First was the passing of the electoral volatility of the 1930s and 1940s and the emergence in the 1950s of a new strategic environment facing labor and business interests. Partisan alignments within and across the institutions of the national government made it unlikely that either could prevail in any further attempt to alter the regulatory regime that had been established during the 1930s and 1940s; it also permitted institutional actors in the three branches of government to turn their attention to other priorities. The other factor Moe identified was the emergence at this time of political stalemate of a highly professionalized regulatory agency. The National Labor Relations Board (NLRB), which was charged to execute the mandates of the 1930s and 1940s, staffed itself with lawyers specially trained in the new law of labor and industrial relations, and it was able, by virtue of its professionalism and the relative indifference of the other institutional actors, to deliver a product that assured the contending interests of the integrity and predictability of the regulatory process itself.

Moe argues that by the time the midcentury political impasse gave way, both business and labor had come to value the underlying stability and predictability that NLRB norms and routines afforded. He points out that when, in 1980, political control of the presidency and the Senate aligned for a major offensive against the interests of organized labor at the NLRB, business showed little interest in casting off the NLRB regimen. In fact, the board nominees suggested by the business clientele of the NLRB were passed over by President Ronald Reagan. Rather than select members who might be expected to support business interests within established norms, the president chose appointees who promised to disrupt and displace those norms. Moe's analysis suggests not only that the institutional arrangements of government had over time affected the preferences of the contending economic interests acting within them but that a government actor, eyeing the strategic environment with his own purposes in view, pressed radical action independent of the expressed interests of his economic allies.

Like Weingast and Bensel, Moe is interested in how institutions stabilize relations among antagonistic interests and how the motivations of institutional actors are reflected in the kind of political order constructed. But Moe's analysis maps movement and interaction on all sides of the various relationships arrayed around the NLRB. His analysis challenges not only the Congress-centered view of the institutional construction of order out of interest conflict but also the larger conception of change modeled as comparative statics. In its place, he presents a picture of systemwide dynamics in which there are multiple points of control none of which is fully competent to produce the result observed and each of which depends for its effective operation on all the others. It is a picture in which interest organizations and institutional actors push and pull one another through time, and one in which the causal arrows point in different directions—from economic interests to the institutions of politics and government and from the institutions of government and politics back to economic interests. The presidency, the Senate, and the bureaucracy appear both as a set of constraints that filter the preferences of economic interests and as locations for alternative purposes and aspirations that create and change interests of their own accord. Government actors appear as much interest makers as interest takers.

None of this is to suggest that the mediation of social interests is less significant in governance than previously thought, only that it has become one factor among others that is being considered in the institutional construction of politics. Notwithstanding its more elaborate and precise specification in the current literature, the concept of interest mediation appears limited by its unidirectional conception of relations between society and the state and by an accompanying narrowness in considering the roles and aspirations of institutional actors themselves. Evidence of a wider array of institutional constructions can now be found in a range of historical studies investigating the development of bureaucratic cultures (Carpenter 2000b), of judicial prerogatives and duties (Gilman 1993) and of political entrepreneurship (Milkis 1993). Investigation of the distinctive temporalities of governance move in very different directions once images of a system built from the bottom up by economic interests contending for the use of state power give way to interactions among relatively autonomous streams of institutional activity, each with its own rhythms, purposes, and repertoires.

▪ | Institutions II: Policy Selection and Political Feedback

Much of the institutional research in APD today self-consciously seeks to avoid an a priori assignment of privilege to societal interests or state interests in explaining patterns of order and change over time. The polity is understood rather as a realm of mutual influence and dialectical exchange, a realm in which the reciprocities of control and voluntarism, structure and agency, and authority and resistance hold sway. This "polity-centered" perspective (Skocpol 1992) assumes that governing arrangements are always up and running, that no matter where one cuts into politics for analytic purposes, such arrangements are already at work fashioning, to a greater or lesser extent, the interests, motives, and movements that seek to use (or challenge) government authority.

This more-encompassing view of institutional construction links explanations of order and stability more directly to explanations of movement and change. The equilibria momentarily sustained through mediation effects are subsumed in an examination of more-extended sequences of action and reaction. Questions of periodization loom large here, for politics is understood analytically to be propelled through time by the dynamics of institutional selection and interest feedback, and wider temporal horizons are necessary to illuminate the path-dependent qualities of change (Pierson 1993).

Anyone who cares to look back far enough in time will likely find that government has a strong hand in creating the interests it appears to be mediating later on. The point is implicit in Bensel's larger corpus. Prior to his analysis of late-nineteenth-century industrialization, Bensel (1990) did a study that traced the consolidation of a powerful class of financiers concentrated in New York and motivated by far-reaching interests in national politics. That process began, he argued, with national banking legislation passed in the early 1860s by a Republican Congress anxious to meet the demands of fighting a war against Southern secession. Considered in conjunction with this earlier work, Bensel's study of the political economy of the 1890s shows American government under the influence of an economic interest that it had itself brought into being for very different purposes decades before.

Those pursuing this line of research in earnest draw inspiration from the proposition that government policy creates interest politics. The idea is almost as familiar to students of American government as that embodied in the proposition that government mediates conflicts of interest within society (Schattschneider 1935; Lowi 1964, 1972). *Policy* in this proposition refers to the precise terms of control selected by government officials to address particular classes of situations. The categories employed, the forms of control chosen, the procedures set up—all of these are to be understood as formative constructs that intrude more or less radically on existing relations of authority and that have important consequences for politics over the long haul. The effects of governmental interventions on a massive scale, like those associated with war and its aftermath, are staples of the study of political development for the reason that policies adopted at these junctures tend to cut widely and deeply through established relations between state and society, thereby altering relations that are possible in the future.

But exactly how widely and deeply they cut is the critical question. Margaret Weir (1992b) has shown that even in wake of depression and world war, American employment policy maintained sharp divisions both conceptually and institutionally between the social problems and the economic problems posed by idle workers, that these divisions came to delimit American development in both domains, and that they continue to distinguish American policy from its counterparts in other countries. Along similar lines, it has been argued that nonintervention by government may prove to be as consequential for development as decisions to enact sweeping new controls. In a comparative study of health care policy, for example, Jacob Hacker (1998) relates the political power of private health care providers to the timing of government intervention into the development of policy networks. In the United States, the failure of comprehensive reform initiatives in the 1930s, 1940s, and 1960s left private interests largely to their own devices in developing elaborate service networks. These, Hacker argues, came to pose ever more-formidable obstacles to comprehensive governmental action as time went on and compounded the difficulties of implementing a national health insurance scheme in the United States in the 1990s. Even more pointed is Marie Gottschalk's analysis (2000) of how government policy in the 1970s entangled labor in private health care networks and business alliances in the 1980s and ultimately compromised its long-standing commitment to national health insurance in the 1990s.

The effects of government policy, whether it entails positive action or de facto inaction, are of interest in the study of APD not only because they determine who gets what in the here and now but also because they classify the groups, impart the identities, forge the divisions, and strike the alliances that channel future political action. Politics has a correspondingly broad construction in this literature, referring both to the institutional creation of policy clienteles—the interests that come to depend on the government's largesse and thus lend support to it—and to the concomitant construction of political dissonance and interests in opposition. In a recent study of welfare politics, Robert Lieberman (1998) analyzed the dismantling of Aid to Families with Dependent Children (AFDC) in 1996 in just this way. He related the effective assault on AFDC back to strategic policy choices made in the 1930s, showing that care had been taken at the outset to separate poor relief from other forms of welfare and to decentralize its administration. Over time, he argued, these arrangements produced a program clientele that was not only more politically isolated than those produced by other welfare policies but also more provocative of local opposition.

Though interests and institutions are understood to be mutually constituted in this literature, characteristic modalities of politics have been discerned on each side. With regard to the institutional construction of interests, it has been shown, for example, that government officials enact particular conceptions of U.S. society simply by incorporating into their policies the norms of the larger political culture of which they are a part. In this way, prevailing prejudices may enter into the demarcation of spheres of legitimate governmental intervention without any external pressure at all and still serve as formative forces delimiting major transformations of the polity. In particular, it has been observed that social programs framed and implemented during the New Deal to aid the forgotten "man" at the bottom of the economic ladder generally failed to address, or even recognize, the distinctive social and economic interests of women. Their effect was to introduce a new gender divide into social provision, liberalizing, nationalizing, and standardizing programs for males while leaving females to appeal their interests to the state and local institutions of an earlier era (Mettler 1998).

At issue, however, are not just the biases carried around in the heads of governmental officials; politics is shaped just as powerfully by their institutional expertise and on-the-job experience. Recent studies of the origins of affirmative action policy in the 1960s and 1970s, for example, show bureaucrats moving far out ahead of the demands being placed on government by the civil rights insurgency. When, in the late-1960s, civil rights organizations were demanding equal treatment in the pursuit of employment opportunities, government administrators were learning that colorblind hiring practices would not produce the concrete signs of progress desired. Pressured by their superiors to show results, administrators began to experiment with rules that would provide for special treatment of minorities. Affirmative action became a cornerstone of African American political interest in this way; that is, as a product of administrative ingenuity and government policy. The political consequences of this institutionally constructed interest would not be fully revealed until affirmative action was taken up and given wider application by a Republican president alert to its potential for setting key Democratic constituencies—blue-collar workers and African Americans—at each others' throats (Graham 1990; Skrentny 1996).

As policy is a specification of governmental purposes, it almost always engages questions of ideology. An expanded view of official motivations has thus also led to more careful distinctions between those aspects of policy which represent government responsiveness to outside pressure and those which reflect the programmatic ideas of officeholders and their determination to secure control in a way consistent with their own political visions. President Reagan's assault on the NLRB is a case in point. Another is federal recognition of the rights of unions to organize and bargain collectively in the first place. Most agree that the disruptions brought on by labor insurgency of the 1930s made some government response imperative, but the extent to which business or labor dictated the government's response has been hotly debated (Goldfield 1990). Several scholars have argued that the Wagner Act was primarily the work of progressive officeholders in power at the time, politicians who had some clear ideas of their own about how best to secure a new order in industrial relations (Plotke 1996; Skocpol and Finegold 1990). The act accorded unions official recognition and it was vigorously opposed by corporate interests for that reason, but it also made the government the ultimate arbiter of the appropriate units of labor bargaining and implicated labor in broader government concerns with securing political order and economic stability. The American Federation of Labor which had long fought against making government a partner in industrial relations found its unions more dependent on the state after they had secured these rights than they had been before (Tomlins 1985).

Turning the perspective around and examining things from the point of view of political interests, the synergistic effects look a bit different. From this angle, the proposition that policy creates politics has illuminated an institutional construction of access and fit. Interests obtain access to officials in ways that are compatible with officials' goals. By this means, government policies can be expected, as a matter of course, to organize some interests into politics and others out. A politics of access is framed by the correspondence or fit between some interests and prevailing government arrangements and by the mobilization of "misfits" for institutional reform.

Access, seen through the lens of fit, likens the polity to an ecological system where institutions establish an environment, and interests thrive, or not, to the extent that they fill niches within it or discover channels for action made available by it (Pierson 1993; Howard 1995). In *Protecting Soldiers and Mothers,* Theda Skocpol (1992) pointed out that while the nineteenth-century United States's patronage-hungry party state proved indifferent or hostile to the welfare demands of laborers, women, and farmers, it reached out to Union veterans of the Civil War with an expansive pension system to support them and their families. In more recent work, Skocpol, Ganz, and Munson (2000) have shown more generally how American government affected the early formation of civic associations in the United States, in particular how the more successful national interests mimicked the government's federated and representative structure.

Narrowly employed, the ecological metaphor evokes the traditional association of the institutions of government with the establishment of political order: the system is self-reinforcing as interests that don't fit get ignored or crowded out. But even in ecological systems, there are other possibilities. Elements that don't fit may die out, but they may also adapt and on occasion they may alter, even radically transform, the environment on which they intrude. The distinctive thing about institutional politics in this regard is that it produces misfits as a matter of course, that transformative elements are introduced as a natural by-product of the particular selections of policy elites. Incongruities of access may intensify over time as institutions and their clienteles form networks of reciprocal support and perpetuate themselves in circumstances often very different from those in which they first took hold. Policy feedback may thus appear in two very different forms. One is mutually reinforcing: interests able to thrive in the niches and channels provided by the existing institutional environment will serve over time to bolster the governing arrangements that sustain them and sustain them over time. The other is mutually threatening as elements excluded or repressed develop their own sense of interest with reference to institutional limits of the existing political space. In this way, institutions construct order and discord simultaneously.

The barriers to entry into politics are formidable, and the mere existence of incongruous interests is no guarantee of political transformation. Explaining why some new interests gain the leverage to change things while others do not is a primary concern in this literature (Clemens 1997). In general, however, it may safely be ventured that the transformational potential of interests that don't fit is far greater than that of those which do. Interests that fit will characteristically seek to minimize alterations in established arrangements; they can be expected to press only those adaptations which promise to maintain the current course or path. Skocpol (1992) shows that pensions for Union veterans filled a niche in the party politics of the late–nineteenth century and that the incipient institutions of a materalist welfare state filled a niche in the antiparty politics of the Progressive Era, but she also shows that each of these developments was bypassed later on, that these niches disappeared when the political climate changed, and that the further elaboration of these incipient welfare ideas was cut short. Incongruous elements, on the other hand, if they can organize at all, are likely to do so in ways that threaten to change more rather than less of the existing state of affairs; their interests are likely to abrade and challenge received institutional arrangements, to target the vulnerabilities of those arrangements and foster rearrangements that fundamentally alter the mix of interests receiving attention.

Historical institutionalists working on these relationships between policy and politics have begun to close the circle of these dialectics of political reconstruction. They have shown how incongruous interests come to define themselves on a site over a span of time: how interest identities are created by the specific institutional arrangements confronted, how demands are shaped by the limitations of those arrangements, and how organizational strategies are informed by the location of interests within the institutional setting and the immediate challenges of dislodging extant controls. Political conditions around the turn of the twentieth century have provided an especially fertile field for illustrating these formative effects, for at that time a state apparatus organized around courts and parties and geared to the assumptions of a dispersed and decentralized society strained to cope with the wrenching social effects of rapid industrialization and rise of economic interdependence. Under these conditions incompatible elements proliferated, and state indifference to their demands spurred diverse actors to experiment with alternative modes of organization and action.

Scholars have found that those who had the greatest transformative effects were those whose demands fell somewhere between a fit with the interests of institutional elites and a frontal assault on their positions. These interests gained leverage by virtue of their development of innovative techniques of lobbying and networking, techniques that by-passed courts and parties while still compelling the attention of office holders (Hansen 1991; Skocpol 1992; Clemens 1997; Harvey 1998). The characteristic government innovations of the Progressive Era—interest-based representation, administrative independence, issue networks, candidate-centered parties—have, in this way, come to be seen not as natural by-products of modernization but as the strategic inventions of new group forms gotten up in a struggle for access to a government which had been arranged and operated with very different interests and purposes in view.

Path dependence, the overriding theme of this version of institutional construction, reintroduces a sense of trajectory into the study of APD. All the various implications of institutional outreach—selection and division, access and exclusion, and clientelism and opposition—anticipate that politics will be propelled along its own course by frictions and abrasions endogenously generated. Natural disasters, foreign disruptions, economic crises, and the like, may intrude from the outside, but their impact is likely to be processed through identities, cleavages, and alliances that have been institutionally constructed. At the same time, this literature has in its own way elaborated the contemporary theme of multiplicity. Gone from this view of development is a single trajectory for the whole system or stages in the unfolding of a singular ideal. Many paths are being blazed simultaneously within the same polity; each according to its own policy-specific political formations. Today's scholars are less likely to portray institutions and political paths as components of larger historical schemes or destinations than as fragments of polity, individual conduits of action, that will in degrees be coordinated or at odds with one another. A final proposition to be found in the literature on institutional construction draws out these implications even further.

Institutions III: Intercurrence

When scholars observe that public policy has subjected women to different standards of control than men or African Americans to different standards than European Americans, something more is hinted at than the interests and choices of the policymakers. These are observations about the temporal makeup or historical constitution of government itself. They suggest that prior political changes were targeted, partial, or incomplete and that, as a consequence, the institutions of government are carrying forward and enforcing a variety of different ordering principles simultaneously.

The simultaneous operation of different, often contrary, orderings of authority is a third feature of the institutional construction of politics, and the most novel of those currently being elaborated in the study of APD today. Its conceptual underpinnings in the nonsimultaneity of institutional origins have led to the discovery of still-more intricate and encompassing temporalities of governance. The proposition is that governments regularly juxtapose different forms of authority because even sweeping political changes are unlikely to change all extant relations between state and society at once in accordance with the same organizing principles. Whether political initiatives and institutional innovations are limited deliberately, unconsciously, or out of practical necessity, their effect is to introduce incongruities into the construction and exercise of government authority over time. The devices that order politics at any given moment are going to be a mixed bag of instruments and are likely to weave political contention into their asymmetries and mutual impingements. We have given this dynamic the label of intercurrence (Orren and Skowronek 1996, 1998).

The idea that intercurrence underlies the institutional construction of politics follows on findings throughout the APD literature, but its revisionist entailments are only now under exploration. In documenting the effects of intercurrence, scholars have begun to present institutions less as sites of conflict resolution or dialectical exchange than as sites of ingrained political contestation, where the successive waves of reform have compounded the difficulties of achieving order over time. Students of environmental politics, for example, have pointed out that the different ideas animating reform over the decades have institutionalized conflicting principles of resource management. Nineteenth-century institutions were designed to distribute federal resources and were geared for their local use and private exploitation. In the Progressive Era, new institutions were created to regulate the use of natural resources and were geared for centralized bureaucratic management and planning, but these never entirely displaced the institutions of distribution and private exploitation. Another wave of reform in the 1970s brought institutions designed for environmental protection and geared toward court-enforced protections of whole ecological systems; these challenged but did not entirely displace the institutions of regulated use or local exploitation. Today, disillusioned reformers are promoting community-based conservation, cooperative public-private initiatives at the local level. Another set of institutions geared to these purposes is entering the field with no sign that previously established institutions will be dislodged (Klyza 1996; Moseley 1999).

In a historical-institutional nexus such as this, authorities collide and standards of legitimacy abrade. Purposes accumulate and crowd in on each other. Institutions fall short of their goals, new goals are specified, and more institutions are created. It is easy to see intercurrence in this light as an indictment of the very possibility of effective government, and indeed scholars have implicated it directly in the persistence of institutions that fail to perform effectively by any standards at all.

In one such study Andrew Polsky (1989) follows the "odyssey of the juvenile court," an institution spawned by the professionalization of social work around the turn of the twentieth century. Social work professionals insisted that juvenile offenders be treated in a less adversarial and more therapeutic manner than common criminals. They promoted the creation of special tribunals in which the traditional court proceeding aimed at an impartial rendering of individual guilt or innocence would be modified to accommodate new concerns for correcting the effects of social injustice and protecting children from any further damage from their environment. The result, Polsky argues, was an institution perpetually torn by the competing standards and purposes that its own officers brought to the table. He finds that involving lawyers in social work and social workers in law enforcement kept the juvenile courts suspended in a state of internal turmoil and turned them into an endless source of political controversy within the larger community. And yet, he also shows that these internal divisions proved critical to the institution's tenacity, as every charge of failure was met by an adjustment and reformulation of the different norms competing for dominance within it. The juvenile court, in this analysis, was from the outset an institution at cross-purposes, in constant and ingrained disequilibrium, but it survived, even thrived, in the face of failure by recycling goals that it could not reconcile.

In another study, this one examining the presidency, Jeffrey Tulis (1987) points to the layering of new rules of presidential rhetoric on older structures and considers the consequences of the resultant combination in undermining political legitimacy. Tulis's empirical analysis of rhetorical standards relates modes of speech, audience, and political purposes and reveals two distinct frameworks. One, a nineteenth-century pattern largely congruent with the arrangements of the Constitution, has incumbents speaking in public mainly on official occasions, according to highly formalized modes of address; policy communications were sent in writing primarily to the Congress and were limited to the president's constitutionally prescribed duties. The other, a twentieth-century pattern associated with the progressive critique of the Constitution, has presidential incumbents speaking often and fluently on a variety of policy issues. Talking frequently over the heads of congressional representatives directly to the public at large, modern presidents aim to mobilize national opinion independently behind their own policy goals. Had these new rules of rhetoric followed on a displacement of the institutional structures on which the old rules had been premised, a coherent standard of legitimacy for the modern presidency might have emerged. Instead, the new rules were layered over the old forms. As a result, twentieth-century presidents have been found to promise far more than they can deliver, Congresses have continued to insist on the vitality of powers anchored in the Constitution (like the war power) when presidents act subject only to their standing in public opinion, and a rising tide of public cynicism has come to engulf officials who invoke different and antithetical standards at will to suit the circumstances of the moment (see also Lowi 1985).

The notion of governance as an intercurrence of different systems of authority relations opens several avenues of exploration. One concerns the internal structure of institutions themselves. In a recent study of Congress, Eric Schickler (2001) challenges the notion that there is a single or even dominant institutional set of rules or incentives ordering legislative politics in the United States over time. Documenting the historical proliferation of disjoint mechanisms that can be invoked by actors at will for their different purposes, he presents Congress as a complex of institutions, multilayered and in tension with one another. A related concern has been to figure out how tensions and contradictions within institutions may work to propel change in patterned and productive ways even as they preclude stable equibibria. Again consider the presidency. It is commonly observed that the U.S. president's constitutional duty—to execute the independent powers of his office and to preserve, protect, and defend the whole—clashes with his role as the leader of a national political party—one who has been nominated by a party convention, elected by a party coalition, and tied to an agenda that depends on cross-institutional cooperation for its implementation.

The tensions between these demands stem from the historical fact that the president's role as party leader was not fixed institutionally until some fifty years after the constitutional duties of the office had been established. But what of the practical effects of these tensions? Beyond the common observations that presidents do not make very responsible party leaders and that national leadership without party support is severely handicapped, this intercurrence has been shown to implicate presidential action in a continual transformation of national government and politics. A recent study points out that when established party coalitions have been at their most robust, the independent actions of presidents have repeatedly sent schisms through their own ranks and thrown national political alignments into disarray but that when political coalitions are already in an advanced state of disarray, the independent actions of presidents have repeatedly galvanized new alliances and secured them in power. A sturdy pattern of systemic political change is to be discerned in this interaction, one that has periodically regenerated American government (Skowronek 1993).

A third implication of intercurrence is that wherever periods of stability are observed, government officials must be actively managing the boundaries and jurisdictions of the different institutions they inhabit. That is to say, multiple orders may work in tandem so long as the endemic problems of coordination they present are effectively addressed. More often than not this means holding divergent principles of government within separate institutional domains and policing their interfaces. The balance rule for the admission of slave and free states in the antebellum era was a managerial device of this sort, a convention negotiated by government officials to hold together in the same orbit whole systems of control that would otherwise have been at loggerheads.

A study of labor in American political development examines stable interactions among multiple orderings explicitly (Orren 1991). The analysis begins, contra Hartz, by pointing to the incorporation into the original Constitution of the common law of master and servant, a body of strictures governing workplace relations between employers and employees that had been passed down with remarkably little change from late-medieval England. The common law arranged workplace relations in sharply hierarchical terms. Contract rights and organizational activities of workers were strictly limited and employer prerogatives were upheld through the authority of the courts. The study details the efforts of judges to sustain this authority in a new constitutional setting where provisions for electoral democracy made it possible, in principle, to implement very different sorts of workplace relations through legislation. For nearly a century, the courts' active patrol and enforcement of these boundaries were successful in holding labor relations safe from encroachments by state legislatures and, generally, in keeping this separate sphere of governance separate.

This same study points to a final implication of intercurrence: the changes most likely to be decisive in political development are those which most thoroughly dismantle and replace established authority. The same institutional activism that held labor governance so effectively within its own domain stymied the restoration of order as unions gained greater power to disrupt the national economy. When organized labor's pivotal position in the new industrial economy became manifest and its challenge to the old rules of workplace became a persistent feature of society and politics, the courts responded by pressing their ancient prerogatives more aggressively. Judges continued to void all efforts to legislate an alternative set of rules, and Congress itself proved reluctant to intrude directly on what had long been the court's domain. The Interstate Commerce Act of 1887, for example, explicitly prohibited the railroads from raising labor issues in rate disputes, even though the pressure on rates which caused Congress to act was driven by demands that railroad unions were pressing ever more forcefully on their employers. It was not until 1937, when the Court upheld the National Labor Relations Act of 1935, that the dismantling and replacement of the old rules of the workplace was complete. For six decades, the court's active management of labor relations implicated it in a disintegration of order and precluded the restoration of stability in the industrial United States (see also O'Brien 1998).

Governance and History

Recent work on political culture, periodization, and institutions points out multiple orders, ingrained incompatibilities, and colliding principles of organization as new premises for the study of American politics. These in turn conjure distinctive conceptions of history. In an essay that is already elaborate, we will limit our conclusion to a brief sketch of two of these emerging conceptions of history, indicating some of the implications for theory building that follow from them.

First, recent APD scholarship has insisted on history as the *site* of governance. The idea conveyed here is something different from, say, history as origins. Whereas origins refer to background and reaches behind to another time, sites occupy the foreground; the reference is to the full spread of extant authority relations that any new state of political affairs must, for better or worse, negotiate. In this view, all political change proceeds on political arrangements, rules, leaders, ideas, practices, attitudes, and so on, that are already in existence. These sites are pitched above the horizon of any discrete actor, at the macrolevel, with the actions of individuals engaging and altering more or less of the larger array of authority relations. The premise of history as a site leads to a conception of political action as an impingement on the authority of others, and it directs attention to those elements in the larger array that are challenged, displaced, transformed, reformed, or unaffected by new political efforts.

Consider, by way of contrast, an ahistorical rendition of the site theme, John Locke's (1960 [1689], 343) statement in *The Second Treatise on Government:* "In the beginning, all the world was America." Locke began by imagining a site "of wild woods and uncultivated waste," a situation before kings and parliaments, a place populated sparsely by individuals who lived without money or property—in other words, an empty place that the word *America* would conjure up in the minds of eighteenth-century English readers. Much of what he said about politics proceeded from this particular proposition about the past. But once actual history becomes the site of action, things look very different. The terrain is filled from the get go, and what already exists defines the problems and substance of change. The proposition "in the beginning all the world was downtown Tokyo" comes closer to the essence of governance as a historical construct.

This image elaborates the notion that governance is always present, even if not in forms we are used to. Wherever we look—at a modern capital like Tokyo, a contested stretch of Arabian desert, an indigenous village in early America, shipboard on the Mayflower—political authority occupies the available space with rules the inhabitants expect will be enforced, and persons assigned to do the job. Vacant lots are few and far between. Constructing something new usually means dislodging something else, and even minor reorganizations are likely to affect impinging activities and installations. This is the essential insight of all those who have detailed the challenges of political development in terms of dismantling and rearrangement (Orren and Skowronek 1998): this includes those who have analyzed antebellum political development as an assault on institutions created by the "regime of local notables" already in place (Shefter 1978; Bridges 1984; Crenson 1975; Wiebe 1995), those who have analyzed the expansion of national administrative capacities in the early–twentieth century with reference to the "state of courts and parties" that had formed in their absence (Skowronek 1982; Clemens 1997), and those who have explored the limitations which U.S. apartheid placed on progressive economic reform (Milkis and Tichenor, 1994, Katznelson, Geiger and Kryder 1993).

Once it is agreed that politics always entails a preexisting where, the question becomes what and after that how and why. In fact, the what question is itself complicated by the difficulty of knowing when a given moment in a course of events constitutes "real" change, and when essentially more of the same. Does the election of a new U.S. president represent "real" change? Only when both houses of Congress change their majority party to his? Always when this happens? To ask such questions is to acknowledge varying magnitudes of change and its combination of different elements. Here a second premise is pertinent, that of history as a matrix. The matrix conveys myriad past experiences pertinent to governance; it presents history as a vessel composed of disparate moving parts out of which all political action emerges and by which it is, in crucial respects, formed.

As we have seen, analysts do not wade into this broth nnaided. They begin with topics, problems, and time frames that provide some semblance of coherence, however provisional. Their next move is to identify regularities in behavior or events that are recognizable as patterns. Unlike sites, which signal containment in time and place, patterns express continuities that are, to the degree indicated, impervious to time's passage. Some patterns are smooth, with no internal breaks: Hartz's liberal polity; some that are cyclical or recurrent: Burnham's realignments; and some that oscillate between opposing extremes like hierarchical and democratic traditions (R. M. Smith 1997; Huntington 1981; Morone 1990). Of course, allowing patterns to stand out from time does not in itself recommend thinking about history as a matrix, nor does the search for patterns distinguish APD's research agenda from many others. But just as history as the site conveys the distinct understanding of a world already fully governed, history as a matrix conveys a distinct understanding of patterns (North 1990).

The image conveyed is one of several patterns moving alongside one another, each with its own rhythms and different points of inception and termination. The interconnections that characterize governance and that are built into the structure of all politics are observed in the parallels and collisions, the push and pull, and the alignments and asymmetries of separate patterns. Likewise, the periodization of various starts and stops, whether staggered or coinciding, becomes a standard of the significance of change, distinguishing that which is narrowly contained from that which is likely to have broader political significance. Research in APD has shown that patterns that suggest nothing controversial when regarded singly become problematic when juxtaposed; this is to say they ask for an explanation of how they are reconciled in practice. The juxtaposition in the nineteenth century of liberal principles of representation against hierarchical common law principles of labor relations is one example (Orren 1991). Another is the juxtaposition of expanding welfare services in the mid–twentieth century amidst continued repression along race and gender lines (McDonagh 1993). Another is the expansion in the twentieth century of a national administrative apparatus that has concentrated discretionary power in bureaucracies that are relatively insulated from the people alongside recurrent social movements, fired by the idea of returning power to the people and demonstrably influential in electoral politics (Morone 1990).

Besides the sheer fun of drawing out the implications of a matrix, these examples indicate an important departure in APD research with implications for all empirical studies of politics that also engaged history. In contrast to the methodical isolation of politics in game theoretic models, for instance, APD research has posited the priority of configurations of authority where persons and agencies exchange authority, influence, resources, and ideas with others outside their sphere or jurisdiction. These sometimes incongruous relations may themselves be patterned in ways that distinguish politics in different countries (Katznelson 1997; Pierson 1994; Sheingate 2001). Politics in the United States stands out in this way for the distinctive bundles of patterns it presents: democratization before bureaucratization, slavery within liberalism, a secular market infused by puritan moralism, and a self-adjusting economy segmented by gender and race. Like other patterns, configurations are sometimes analyzed negatively, as in Greenstone's boundary conditions: the absence of feudalism, for instance, associated with the absence of class politics and with judicial review.

Approaching political change through a matrix affects the essential components of explanation. If governance is composed of different elements juxtaposed as they proceed through time, then causation in political affairs will be difficult to capture as variables ordered separately as cause and effect. Causation in the realm of governance takes place in a vortex of mutual influences and reciprocal effects. At any particular moment, outcomes will turn on questions of timing, that is, on the intersection or conjunction of patterns. As several patterns are likely to be at work at every significant juncture, political change will be characterized generally by weak or diffuse causation. At the same time the weak causation found in a matrix is likely to be joined to a strong sense of agency, as all political change comes down to the matter of negotiating contending elements in motion. It is as agents, not causes, that political actions matter.

Reckoning with politics through patterns in a matrix is APD's signal accomplishment to date. In this, it has presented a fundamental challenge to the rest of political science: to elaborate and adopt an idea of time that is appropriate to the organization of the political universe and the study of governance. The matrix image alerts us to a time different from the $t - 1$, $t - 2$ lockstep characteristic of history neatly periodized into separate eras or politics modeled as games. Sites are too diverse in their composition to travel in so linear a fashion. As a substitute, APD poses a time composed of progressions, sequences, and coincidences of governance, as these are discerned individually and assayed for their mutual and cumulative effects.

ROBERT POWELL

Game Theory, International Relations Theory, and the Hobbesian Stylization[1]

Introduction

Strategic interaction is at the center of much of the work that has been done in international relations theory since at least the end of World War II.[2] At that time, a consensus emerged among international relations scholars on at least two broad points. The first was on the domain of the field, and the second was that this domain made the field of international relations theoretically distinct from other fields, including other areas of political science.[3] Writing in 1949, Frederick Dunn summarized the scope of international relations (IR) as it then seemed "to be taking form in the work of leading scholars in the field. . . . The distinguishing characteristic of IR as a separate branch of learning is found in the nature of the questions with which it deals. IR is concerned with the questions that arise in the relations between autonomous political groups in a world system in which power is not centered at one point" (1949, 142, 144). That is, international relations focuses on understanding the relations among actors or groups that interact in an anarchic political system whereas other fields of political science concentrate on the political interaction that takes place in and is conditioned by a more hierarchically ordered system.

This consensus on the scope and distinctiveness of the field defined the challenge that confronted it. As William T. R. Fox concluded from his survey of interwar research in international relations, "A body of political

1. I am grateful to James Alt, James Fearon, Robert Jervis, Helen Milner, James Morrow, and Ira Katznelson for helpful comments and criticisms.

2. See Knutsen 1997 for a general history of international relations theory, Schmidt 1998 for an emphasis on the century between 1850 and World War II, and Kahler 1997 and Milner 1997 for excellent discussions of the period after 1945. Jervis 1997 provides a conceptual discussion of the importance of strategic interaction in international relations theory.

3. See Milner 1998 on the latter point.

theory dealing with a system characterized by an absence of a central authority is yet to be developed. . . ." (1949, 79). Indeed, we can see much of the work that would be done over the next fifty years as an effort to devise just such a theory.

One very important and influential approach to developing this theory begins by trying to abstract away from many of the details of international politics and foreign policy. This approach focuses instead on attempting to understand the strategic logic of a simple, stylized model of the international system. In this stylization, a small number of states interact in a Hobbesian state of nature in which there is no supranational Leviathan to impose order and to protect the states from each other. There is nothing to stop one state from trying to further its interests by using its military means against another state if the former believes that doing so is its best interest.

This approach has been used to investigate problems and questions that have dominated much of the field over the last half-century: Does the anarchy of the Hobbesian state of nature induce states to try to maximize their power (Morgenthau 1948; Herz 1950; Wolfers 1962; Waltz 1979; Mearsheimer 2001)? Can a system composed entirely of security-seeking states that do not want war still break down in war (Butterfield 1950; Schelling 1960; Jervis 1977, 1978; Waltz 1979; Glaser 1992, 1996; Schweller 1996; Kydd 1997a)? Do offensive advantages make the security dilemma more intense and war more likely (Quester 1977; Jervis 1978; Glaser 1992, 1997; Kydd 1997a; Van Evera 1999)? Do secure, second-strike nuclear forces make war less likely (Brodie 1946, 1959; Jervis 1984, 1989; Waltz 1990). Do states balance or at least act in ways that tend to produce balances of power (Morgenthau 1948; Gulick 1955; Wolfers 1962; Waltz 1979; Walt 1987; Schweller 1994)? Is war less likely if there is an even distribution of power or if one state has a preponderance of power (Claude 1962; Wolfers 1962; Organski 1968; Blainey 1973; Organski and Kugler 1980; Kugler and Lemke 1996)?

Do shifts in the distribution of power make war more likely (Organski 1968; Organski and Kugler 1980; Gilpin 1981)? Do the constraints of the international system induce states to be concerned about their relative gains, and, if so, to what extent do these concerns impede international cooperation (Waltz 1979; Grieco 1988, 1990, 1993; Powell 1991, 1999; Snidal 1991; Gowa and Mansfield 1993; Keohane 1993; Morrow 1997)? To what extent can repeated interaction and reputational concerns mitigate states' short-run incentives to exploit each other (Axelrod 1984; Oye 1986b; Fearon 1998a)? Do international regimes and institutions have an independent effect on states' behavior and, in particular, can they facilitate international cooperation (Krasner 1983; Keohane 1984; Krasner 1991; Mearsheimer 1994–95; Morrow 1994a; Chayes and Chayes 1995; Keohane and Martin 1995; Downs, Rocke, and Barsoom 1996)?

Each of these particular questions shares the same general form. Each begins with a setting in which one actor's optimal behavior depends on what the other actors do. It then asks how different environmental features—the state of the offense-defense balance; the distribution of power; the existence of secure, second-strike nuclear forces; the presence of international institutions—affect the way that the states or, more generally, the units interact. Each question is fundamentally about strategic interaction.

Game theory is a tool for analyzing strategic interaction. Not surprisingly, then, international relations theorists have tried at various times over the last half-century to use game theory to advance their understanding of these problems. This essay focuses broadly on how game theory fits into the continuing intellectual development of international relations theory, especially that branch of the field that has tried to understand certain aspects of international politics by examining the strategic logic of very spare stylizations of the international system based on a Hobbesian state of nature. The essay does not review the specific contributions game theoretic analyses have made to international relations theory.[4] Rather, it considers three more general questions: What does game theory have to offer international relations theory? To what extent is what game theory has to offer needed? And, what are some of the effects on international relations theory of adopting a more game theoretic approach likely to be?

To foreshadow the discussion of these questions, this essay emphasizes that game theory is not a theory of international relations. Nor is it a substitute for deep ideas and good intuitions about the workings of international politics. Nor is game theory a substitute for careful, systematic empirical analysis. Game theory is a research tool which complements systematic empirical work and ideally interacts synergistically with it. Game theory provides a formal method of analyzing strategic interaction.

Formalization offers two advantages to the study of strategic interaction. First, specifying or closing a formal model usually requires one to describe the actors and the strategic environment in which they find themselves more clearly and more precisely than ordinary-language arguments typically do.

The relative importance of what game theory has to offer to international relations theory depends on the state of the field and varies as the field evolves. Formalization is less important when a field has already developed clear and coherent theories that make different empirical predictions. In these circumstances, careful empirical evaluation is likely to offer a higher payoff than greater formalization. By contrast, formalization is more important when assumptions and the conclusions that follow from them are unclear, for then there is nothing to test empirically. A review of some of the major debates that have dominated much of international relations theory in recent years suggests that the value added by greater formalization is large at this stage in the development of the field.

4. See Morrow 1999b as well as Morrow's contribution to this volume for reviews of some of this work.

Adopting a more game theoretic approach seems likely to have several effects on international relations theory. One of them is what Paul Krugman (1995) calls the "evolution of ignorance." Some theoretical deductions that have been derived from ordinary-language arguments will be lost at least temporarily because they cannot be derived formally. Another effect is that game theory will promote a reintegration of international relations with the rest of political science by showing that the strategic problems that states and other actors face in international politics often have close parallels in other areas of political science like U.S. and comparative politics.

The next section of this essay describes some of the general properties of research methods and some of the particular advantages that game theory has to offer in the study of strategic interaction. The subsequent section examines some of the debates that have dominated international relations theory in recent years and shows that much of the debate arises from weakly or poorly specified connections among the actors' preferences, the strategic setting in which the actors pursue their goals, and the resulting outcomes. The final section centers on some of the consequences of adopting a more formal approach to international relations.

Game Theory as a Research Tool

Game theory is not a theory of international politics any more than calculus is a theory of Newtonian dynamics. Newton developed calculus as a tool for helping him analyze the problems that his theory of mechanics posed (Christianson 1984). Similarly, game theory is one of many research tools or methods that can be used to study the problems that international relations theories pose.[5] This section briefly discusses the origins of the most recent efforts to use game theory to understand international politics, the role of research methods in general and of game theory in particular, and some of the advantages and disadvantages of formalization.

DYNAMIC INTERACTIONS, GAME THEORY, AND INTERNATIONAL RELATIONS

The latest concerted attempt to use game theory to advance our understanding of international politics dates from the mid-1980s and draws

5. Game theory, of course, was not developed in order to solve problems in international relations theory. But there are some close connections. The classic example is Thomas Schelling's *The Strategy of Conflict* (1960). Less well known is the fact that the U.S. Arms Control and Disarmament Agency also sponsored work in the 1960s that helped pioneer the study of repeated games with asymmetric information (Aumann and Maschler 1995). O'Neill (1994) reviews the relation between game theory and deterrence theory.

on two technical innovations in game theory that were then beginning to revolutionize economics (Kreps 1990b, 1). The first was due to John Harsanyi. In many situations in international politics, actors are uncertain about the motivations, intentions, or resolve of other actors. For example, Britain in the 1930s was unsure of Hitler's motivations and the extent of his ambitions. As Alexander Cadogan, the permanent undersecretary in the Foreign Office put it shortly after Germany annexed Austria: "I am quite prepared to believe that the incorporation in the Reich of Austrian and Sudentendeutch may only be the first step in a German expansion eastwards. But I do not accept that this is necessarily so, and that we should rush to any hasty conclusions" (quoted in Parker 1993, 135). Harsanyi (1967–68) showed how to model formally situations in which actors are uncertain of each other's intentions, motivation, or resolve.[6]

The second technical innovation was a set of tools designed for analyzing dynamic interactions and the credibility issues inherent in them. A static interaction is one in which no one can respond to anyone else's actions because everyone acts simultaneously or because they are unable to observe what other actors have already done. The prisoner's dilemma or other 2 × 2 games like chicken or stag hunt are all static games. Each player makes a single decision, and neither actor can react to what the other does. A dynamic interaction, by contrast, is one in which at least some actors can observe the actions of others and react to those actions. The bidding in poker is an example of a dynamic interaction.

Many of the interactions we want to understand better in international politics are fundamentally dynamic. If, for example, one defines a crisis, as Snyder and Diesing (1977, 13) do, as a situation in which one state challenges another state and the latter resists, then crises are dynamic interactions. But the formal analysis of dynamic interactions requires more tools than the analysis of static interactions does, for the latter require us to address credibility problems.

These problems arise in dynamic but not static interactions, because actors can make threats and promises in the former but not the later. By their very nature, threats and promises presuppose a situation in which an actor can react to the behavior of others. A threat, for example, is fundamentally a conditional statement. If one actor behaves in certain ways, then the threatener will react in certain other ways. Thus threats and promises can only arise in dynamic interactions.

Once threats and promises are at issue, credibility considerations arise. Intuitively, actors should respond to credible threats and promises differently than they react to patently incredible threats and promises. Thus in order to be able to analyze dynamic interactions formally, we need to

6. More generally, Harsanyi showed how to model situations in which actors have asymmetric information. For example, one actor knows how much risk it is willing to run but not how much risk another is willing to run.

be able to describe formally what it means for a threat or promise to be credible.

In the late 1970s and early 1980s, game theorists devised a set of formal tools for the analyzing the credibility issues inherent in dynamic interactions. The combination of these tools and Harsanyi's earlier work on asymmetric-information games triggered an explosion of work that transformed economics over the next decade.[7] By the mid-1980s the effects of these tools were also beginning to be felt in political science.[8]

RESEARCH METHODS

Good ideas and sharp insights about the various ways that international politics work can come from many places. Sometimes they originate in a deep historical knowledge of particular cases, at other times they arise from spare stylizations, and often they develop through an iterative dialogue that moves back and forth between thick descriptions and thin conceptualizations.[9] Research methods in general and game theory in particular are not a substitute for good ideas.

Research methods have less to do with the origins of good ideas and more to do with evaluating those ideas logically and empirically. Sometimes promising ideas that initially seem deep and insightful turn out to be so, and sometimes they do not. One of the most important functions of research methods—be they formal, statistical, or qualitative—is to help us discipline our thinking about our ideas and intuitions and about the conclusions that seem to follow from them. This discipline plays an essential part in translating those ideas and insights into clear theoretical claims and, ultimately, empirically substantiated findings.

One of the ways that methods provide this kind of discipline is by establishing a set of standards for assessing arguments. To draw an analogy with statistics, these standards ideally reduce the chances of making both type I and type II errors. The latter occur when false claims are mistaken for true ones. The discipline of good research methods reduces the likelihood of this kind of mistake by making it harder to see what, because of motivated or unmotivated biases, we want or expect to see in the data or case studies. For example, Alexander George's method (1979) of "structured, focused comparison" makes it more difficult to accept a causal

7. The 1994 Nobel Prize in economics recognized this work along with that of John Nash. Kreps (1990) provides an accessible and nontechnical survey of some of these developments.

8. Among the earliest work to draw on these tools and appear in the *American Political Science Review* are Brito and Intriligator 1985, Palfrey and Rosenthal 1985, Austen-Smith and Riker 1987, Bendor and Mookerjee 1987, Ordeshook and Schwarz 1987, Powell 1987, and Shepsle and Weingast 1987.

9. See Myerson 1992 and Powell 1999 for a discussion of this interative process.

claim that seems to fit a particular case, because it requires a researcher to make a more-controlled comparison across multiple cases in a way that "defines and standardizes the data collection" (George and McKeown 1985, 41).[10] Similarly, research methods can also reduce the chances of making type I errors which occur when true claims are mistaken for false ones. By providing a set of common standards by which arguments are judged, methods make it more difficult to dismiss or discount a finding because one dislikes it or its implications.

The Discipline Wrought by Formalization

Formalization brings a particular kind of discipline. Mathematical models give us "a clear and precise language for communicating insights and notions" in the sense that they

> show that certain precise assumptions lead to other precise conclusions. It also allows us to stretch our analyses and to unify them; once we have worked our way through the logic that assumptions A imply conclusions X, we may see how assumptions A' lead to conclusions X' by the "same basic argument." It allows us to appreciate how critical are certain (often implicit) assumptions: If A leads to X, but a slight change in A to A' leads to not X, then we can appreciate that X or not X depends on the seemingly slight differences between A and A'; hence X is not a very robust conclusion. Taking logical deductions back to the real world, where the satisfaction of assumptions A or A' is a matter of some controversy, our developed intuition concerning what assumptions lead to which conclusions, together with a sense of how closely the real world conforms to A or A', gives us the courage to assert that X will or will not pertain with very high probability. (Kreps 1997, 63–64)

In essence, formal models provide a kind of accounting mechanism that helps us think through some issues more carefully than ordinary-language models can. Accounting schemes make a firm's financial situation more transparent both to those inside the firm and to those outside it. Formal models make arguments more transparent both to those making them and to those to whom the arguments are made.

This improved transparency comes from two sources. First, models must be fully specified or closed before they can be analyzed. Closing a model often reveals that important but previously unappreciated assumptions have to be made in order to support an argument. Models help make critical assumptions more explicit. Second, the links from assumptions to conclusions are clearer in formal models. Indeed, the derivation of conclusions frequently takes the form of mathematical proofs or demonstra-

10. On the comparative case study method, also see Lijphart 1971; Eckstein 1975; Collier 1993; and King, Keohane, and Verba 1994.

tions. These clearer linkages make it easier to trace the effects of changing one or more assumptions.

Game theory is a particular kind of mathematical modeling which disciplines our thinking about strategic interaction in at least two important ways. The first results from defining a game. Specifying a game tree requires us to describe the strategic environment, that is, who the actors are, the order in which they make decisions, what alternatives each actor has to choose from when deciding what to do, and, finally, what each actor knows when it has to make a decision. We also have to specify the actor's preferences over the possible outcomes of their interaction, that is, their ranking of the terminal nodes of the tree. These specifications make the assumptions being made about the actors' strategic environment more transparent.

Second, games are generally analyzed in terms of their perfect equilibria. Solving a game for its perfect equilibria disciplines our predictions about how the game will be played just as defining a game disciplines our thinking about the strategic setting. Equilibrium analysis forces us to look at the situation being modeled from the perspective of each and every actor and to ensure that the prediction makes sense from all of these perspectives.

A perfect equilibrium is a set of strategies—one for each actor—that satisfy two conditions, and meeting these two requirements is what effectively forces us to look at the situation from each actor's position. The first condition is that the set of strategies must be self-reinforcing. That is, no actor can benefit by deviating from its strategy given that that actor believes that all of the other actors are playing according to their strategies. If this condition did not hold, then at least one actor would want to do something other than what he was predicted to do and the prediction as a whole would not make sense. Strategies that satisfy this condition are called *Nash equilibria.*

The second condition is what makes a Nash equilibrium in the modest language of game theory *perfect.* This requirement is important because self-reinforcing strategies beg a prior question. A set of strategies is self-reinforcing if no actor can increase its payoff by altering its strategy *given* that the other actors follow their strategies. But is it reasonable in the first place for an actor to believe that the other actors will play according to the posited strategies? One situation in which it is unreasonable is if the threats and promises implicit in another actor's strategy are inherently incredible. Suppose, for example, that the strategy an actor is presumed to follow relies on a threat which would not be in that actor's own self-interest to carry out if the time came to do so. If other actors know this, then it no longer makes sense for them to assume that the first actor will follow its posited strategy and carry out its threat.

Insisting that a set of self-reinforcing strategies also be perfect helps to resolve this issue formally. Perfection requires that following through on the threats and promises implicit in each actor's strategy be in that actor's self-interest. Thus, no actor has any reason to doubt that any other actor will not play according to its posited strategy.

We can think of the relative advantages and disadvantages of formalization compared to ordinary-language arguments in terms of a trade-off between type I and type II errors. If an ordinary-language argument can be formalized, that formalization generally forges tighter, more transparent, and deductively rigorous links between the initial ideas and insights and the conclusions that follow from them. These clearer, more precise links often reduce the chances of incorrectly accepting a false claim and thereby making a type II error. For example, much of the recent formal work in international relations theory shows that many widely accepted ordinary-language arguments do not go through when they are formalized.[11]

But, moving toward a standard that requires more formalization may increase the chances of making a type I error at least in the short run. It initially may prove to be impossible to study some existing ordinary-language arguments which are fundamentally correct because no one can figure out how to model them formally. Until some one can figure out how to formalize these arguments, they will remain interesting ideas and conjectures but they will not be accepted because they do not "measure up" (to the formal standard). If, however, these ordinary-language analyses are describing a causal argument that is fundamentally correct, then the shift to a more formal standard will have induced a type I mistake at least temporarily.

This kind of mistake is part of what Paul Krugman calls the "evolution of ignorance" that accompanies modeling. Because what we "know" is partly a function of the standards by which we evaluate arguments, imposing a different set of standards may mean that at least at the outset we "know" less that we thought we did.

> Model-building, especially in its early stages, involves the evolution of ignorance as well as knowledge; and someone with powerful intuition, with a deep sense of the complexities of reality, may well feel that from his point of view more is lost than is gained. . . . The cycle of knowledge lost before it can be regained seems to be an inevitable part of formal model building. (Krugman 1995, 79)

In sum, using formalization to filter arguments is likely to screen out ordinary-language analyses that are fundamentally wrong or seriously incomplete as well as those analyses that are fundamentally correct but presently cannot be modeled. Filtering the former reduces the chances of making type II errors, but filtering the latter raises the chances of making type I errors.

11. See, for instance, Fearon 1994b, 1995b and Powell 1999.

However, this trade-off goes beyond the assessment of existing arguments, and this may be the most important effect of a shift from one set of standards to another. What counts in a research community as interesting and important new work is often judged as such relative to some standard for assessing arguments. A shift in these standards may therefore implicitly redefine what constitutes important new work and thereby lead to a redirection or reorientation of the field.

Assessing the Trade-off

Greater formalization entails a trade-off because there are both costs and benefits to applying a more formal standard. Whether the costs outweigh the benefits depends in part on the state of the field. As emphasized above, much of international relations theory centers on strategic interaction. This work proceeds by making assumptions about states and the strategic arena in which they interact and then by tracing the implications of these assumptions. Formalization promotes a clearer and more precise specification of these assumptions and helps to forge tighter and more transparent links between these assumptions and the conclusions claimed to follow from them. But the value of this contribution will be relatively less important if the field already has several clear and internally coherent arguments. In these circumstances, empirical evaluation and testing would be more valuable. If, however, existing formulations are vague or poorly specified so that it is not clear what is being assumed or what follows from what, then the relative risk of making type II errors is large and the value added of formalization is greater.

This section examines three issues in international relations theory. The first is an essential element of the theoretical infrastructure of the field; it is the very notion of what a structural theory is and what structurally induced preferences are. The second issue is representative of many of the controversies that have dominated the field over the years. It is the emerging debate between offensive and defensive realism. The third issue is the theoretical usefulness and adequacy of one of the fundamental assumptions of much of international relations theory—the assumption that states seek to survive. A review of these issues shows that many of the misunderstandings surrounding them can be traced to the failure to make key ideas and concepts clear. These misunderstandings have impeded the development of the field, and this suggests that game theory and formalization have much to offer at this stage in the development of international relations theory.

STRUCTURAL THEORIES AND STRUCTURALLY INDUCED PREFERENCES

The quest for systemic or structural theories has a long history in international relations theory.[12] The self-conscious search for structural theories can be traced back at least as far as Kaplan's *System and Process in International Politics* (1957) and Waltz's discussion of the third image in *Man, the State, and War* (1959). The key insight underlying the third image (which can also be found in Butterfield's brief discussion [1950] of what we now call the security dilemma) is that one of the important effects of strategic interaction is that it may divorce desired outcomes from realized outcomes. To wit, war may occur even though each state is only seeking to ensure its security and none prefer war to peace.

Man, the State, and War provided a typology of existing theories and arguments but did not present a theory. For example, the defining feature of the third image was anarchy (1959, 159). But Waltz never traced an explicit path from anarchy to outcomes. Rather, anarchy served as a permissive cause: wars occurs in anarchic systems because "there is nothing to prevent them" (232). Twenty years later, Waltz proposed such a theory.

Theory of International Politics was one of the most influential books in international relations theory published since World War II. Its emphasis on international structure and its spare definition of that structure sparked an enormous debate about the proper definition of structure and about the ability to infer a state's preferences and behavior from its structural position. Since the degree to which a state's preferences can be derived from its structural position would seem to be inversely related to the importance of domestic politics in explaining international relations, the debate over the power and usefulness of structural explanations also fueled the continuing dispute about the significance of domestic politics. Although debates about structure no longer dominate the field as they once did, the usefulness of structural theories remains important. For example, the ability to infer state preferences from its structural position lies at the center of the growing controversy about the prospects of constructing neorealist or structural theories of foreign policy.[13]

Despite this long quest for and debate about structural explanations, the very notions of what it means for a theory to be "structural" or of what it means for preferences to be induced or inferred from structure remain uncertain. To illustrate the point, suppose we think of structure as defining the strategic arena in which the units interact and of preferences as the ends that the units pursue. Then, both structure and preferences are prim-

12. See Jervis 1997 for a recent overview of systemic approaches in international relations.

13. Elman (1996) and Rose (1998) survey this debate, and Fearon (1998c) offers a critical assessment of the distinction between international politics and foreign policy.

itives in the analysis of the units' interaction; neither entails the other.[14] But if this is so, what does it mean to say that we have a *structural* theory or that we can infer a unit's preferences from its structural position?

Surely, it does not mean that these primitive preferences can be inferred from the structure since these preferences are prior to any deductions. If preferences can be inferred from structure, then these "induced preferences" must be of a different kind. Here it is useful to distinguish between preferences over outcomes and preferences over strategies. The former is an actor's ranking over the terminal nodes of a game tree or over the cells of a game in payoff matrix (i.e., strategic) form. The latter kind of preference is the way that a player ranks its strategies. Because of strategic interdependence, this ranking depends on what strategies the other players intend to follow. For example, the 2 × 2 game of chicken in figure 1 can end in four ways: mutual compromise if both states submit, mutual disaster if both stand firm, and with one state or the other prevailing if it stands firm while the other submits. Then, the numbers in the cells of the matrix refer to the primitive assumption about the way the players rank the four possible outcomes. However, a player's preferred course of action, that is, whether it stands firm, depends on what it expects the other player to do. In particular, a state prefers to stand firm if it expects its adversary

Figure 1 | Preferences over Outcomes and Strategies in Chicken

	Firm	Submit
Firm	−10, −10	5, −1
Submit	−1, 5	0, 0

14. In a game theoretic model, for example, both the game tree and the actor's preferences, which are defined over the terminal nodes of the tree, are coequal elements that go into the specification of a model.

to submit, and it prefers to submit if it expects its adversary to stand firm.[15]

Preferences over strategies, therefore, are not a primitive of the analysis. At least in principle, they can be derived from the primitives describing the strategic environment and the actors' preferences (over outcomes). Indeed, this is precisely what game theory tries to do. Consequently, efforts to infer "preferences" from structure would seem to be referring to this kind of preference.

What, then, is a structural theory and a structural explanation? James Fearon (1998c) provides a point of departure. In an effort to examine the relation between theories of international politics and of foreign policy, he offers two broad notions of what one might mean by a structural theory. The first sees "states as unitary actors who consider what other states will or might do when they choose their foreign policies" (1998c, 261). That is, a theory is structural by this definition if it treats states as purposive, unitary actors who perceive themselves to be in a strategic environment. The second type of a structural theory is a subset of the first. A theory is structural if it satisfies the first definition by treating states as purposive unitary actors and if it also "adds conditions on which explanatory variables can operate or how they operate. . . ." (1998c, 261).

Waltz, for example, restricts a systemic explanation to be one that relies on only three variables (two of which are moot in international relations): the ordering principle of the system, the functional differentiation or nondifferentiation of the units, and the distribution of capabilities among the units (1979, 79–101). Consequently, explanations of international political outcomes that rely on, say, the advent of nuclear weapons or the offense-defense balance are not structural explanations (for him), because they appeal to variables that have been *defined* to be nonstructural.

Of course, there is nothing sacrosanct about this definition, and one could readily add other variables to the list.[16] To illustrate the point, suppose we follow Glaser (2000b) by adding the offense-defense balance to the list of "acceptable" structural variables. Then Christensen and Snyder's (1990) explanation of states' alliance behavior would not be structural according to Waltz but would be according to this amended definition.[17]

15. See Powell 1994, 318–21, for a discussion of preferences over outcomes and strategies. Frieden (1999) discusses different approaches to specifying preferences over outcomes.

16. Indeed, Waltz himself offers a pragmatic defense of his emphasis on the distribution of power and exclusion of other things that can be cast in distributive terms, e.g., "differences in ideology, in internal structure of property relations, or in governmental form." His justification of this emphasis is that "state behavior varies more with differences in power" than with differences in these other things (1986, 329).

17. Presumably, the empirical effects of the offense-defense balance on alliance behavior would not depend on whether we defined this balance as a structural or nonstructural variable.

Fearon's distinction between two types of structural theory makes what constitutes structure entirely a matter of definition. A theory is *defined* to be structural if it treats states as strategically interdependent unitary actors or if, in addition to this, the theory only considers how some predefined set of variables, for example, the distribution of power, affect this interaction. This purely definitional approach is in keeping with much of the literature in international relations theory and does provide a useful way of categorizing both structural and nonstructural theories.[18]

But there is another way to conceive of a structural explanation which is partly a matter of definition and partly a matter of deduction. This other conception is moreover what many international relations theorists seem to have in mind when they discuss structural explanations and the ability to infer state preferences from structure. To develop this alternative, consider why a powerful structural theory would be a useful thing to have in the first place. Structural theories, if they can be devised, "explain why states similarly placed [in the system's structure] behave similarly despite their internal differences" (Waltz 1996, 54). That is, a structural theory *explains* why we can account for at least "a small number of important things" (Waltz 1986, 329) on the basis of the states' structural positions without having to know anything about, say, the states' domestic politics or their internal characteristics except, possibly, that they seek to survive. To draw a loose analogy with statistics again, structural position is akin to a *sufficient statistic* for predicting outcomes; that is, once one knows the units' structural positions, knowing other things about their unitary attributes would not improve on one's ability to predict outcomes.

However, explaining why variation in a set of (possibly arbitrarily defined) structural variables significantly affects state interactions *and* why variation in the other nonstructural variables does not have a significant effect (and therefore can be ignored) is a matter of deduction and not solely one of definition. A theory that does not allow for variation among nonstructural characteristics cannot *explain* why such variation would or would not have any consequential effects. In order to demonstrate theoretically that variation in, say, some attributes of the units has no little or no effect on the outcome, a theory must represent those attributes in some way, allow them to vary, and show *deductively* that this variation has no significant effects. Unless it does these things, then the theory cannot "explain why similarly placed states behave similarly despite" their different unitary tributes.[19]

To put the point more concretely, suppose that a theorist is trying to create a structural theory that accounts for the likelihood of war. She be-

18. The definition of a structural explanation also implicitly defines what constitutes a nonstructural or domestic politics explanation; this is the point Fearon is emphasizing.

19. Of course, all theories leave some things out and therefore are deductively silent on the effects that variation in these features would have.

gins by treating states as purposive, unitary actors and by defining the distribution of power to be the sole structural variable. That is, explanations which link changes in the likelihood of war to changes in the distribution of power are the only kinds of explanation that count as structural. This is in keeping with Fearon's second notion of a structural explanation.

After much toil and trouble, she succeeds in devising a theory that deductively links variations in the distribution of power to changes in the likelihood of war. Should this count as a structural explanation? It surely satisfies the definition above in that it only appeals to variation in what has been defined to be a structural variable. In this purely definitional sense, then, this is a structural explanation. But would this theory support the claim that "states similarly placed behave similarly despite their internal differences" (Waltz 1996, 54)? Would, for example, the theory support the claim that states' internal structures, for example, whether they are democratic or not, have no significant effect on the likelihood of war?

The answer is clearly no, and there are two cases to consider. If, first, there is no way to represent the distinction between democratic and nondemocratic states in the theory—albeit even at a high level of abstraction—then one cannot ask whether this difference has a significant effect on states' interaction in the context of this theory. The question is simply beyond the theory's scope.

If, by contrast, there is a way to distinguish between democratic and nondemocratic states in the theory, one can pose the question. But in order to answer it, that is, to see whether variation in the states' internal attributes affects their interaction, then one must vary this nonstructural variable. If this variation has no significant effect on states' interaction, the theory would provide a structural account of outcomes. That is, knowing states' structural positions would be sufficient to account for the outcome of their interaction. But if this variation in nonstructural elements led to significant variation in the outcomes, then the theory would have undercut a structural explanation. According to the theory, states similarly placed in the structure would not behave similarly depsite their internal differences.

In these circumstances, we would have a structural theory in the purely definitional sense; it links changes in the distribution of power to unambiguous changes in the likelihood of war *as long as one does not consider the effects of variables that have been defined to be nonstructural.* But the theory would not be structural in the sense that it showed that variation in nonstructural variables had no significant effect on the outcomes. And, it is the latter notion of a structural explanation that makes such explanations useful.

In sum, structural explanation is more than a matter of definition. What one takes to constitute *structure* is a definitional issue. But a *structural explanation*, however structure is defined, is also a matter of deduction, for it shows that variation in these structural variables significantly affects outcomes whereas variation in nonstructural variables does not.

Although the importance of structural explanations has been debated and discussed since Waltz published *Theory of International Politics* more than two decades ago, the relation between definition and deduction in the development of structural theories and explanations is still not widely appreciated nor is this lack of appreciation limited to neorealists. Consider, for example, Andrew Moravcsik's recent effort to provide a "liberal theory of international politics" that takes "preferences seriously." He claims that realism and institutionalism believe that variation in state preferences has little or no effect on the outcomes of state interaction (1997, 522). Variation "in state preferences should be treated as if they were irrelevant, secondary, or endogenous. . . . What states do is primarily determined by strategic calculations . . . which in turn reflect their international political environment" (Moravcsik 1997, 522). Liberalism, by contrasts, "reverses this *assumption*" (522, emphasis added) by taking variation in state preferences to be paramount in explaining outcomes. Variation in "interstate political and strategic circumstances" is unimportant (522–23). Thus, whether structural or nonstructural features are consequential is a matter of assumption for Moravscik and not of theoretical deduction.

How is all of this related to game theory and to the value added of forging tighter links between assumptions about actors and their environments and the conclusions that follow from these assumptions? The real issue underlying the debate about structural explanations is not whether some variables are defined to be structural and others are not. The issue is a theoretical effort to trace the effects of changes in certain factors such as the distribution of power, the offense-defense balance, or whether the states are democratic or nondemocratic. Whether or not some of these factors are called structural is theoretically irrelevant. Whether variation in a variable in the context of a given theory significantly affects the outcome is a deductive property of the theory and does not depend on whether this variable is called structural. Whether a statement can be derived from a set of assumptions does not depend on what those assumptions are called. *The appellation* structure *does no theoretical work.*

If the links between assumptions and conclusions were clearer, all of this would be immediately evident. One would not debate whether this or that factor, say, the existence of nuclear weapons or the offense-defense balance, was structural in an effort to piggy back on the "well-known" implications of structure. Rather, one would simply try to incorporate these elements into a formulation and then try to trace their implications. Clearer links between assumptions and conclusions would have saved an enormous amount of scholarly time and energy, time and energy that could have be devoted to testing theoretical implications.

OFFENSIVE VERSUS DEFENSIVE REALISM AND THE PROBLEM OF INFERENCE

As noted above, much of international relations theory has tried to understand certain aspects of international politics by analyzing the strategic logic of a simple stylized system based on a Hobbesian state of nature. One of the latest efforts revolves around offensive and defensive realism.[20] A brief review of this debate shows again the importance of forging tighter, more transparent links between assumptions and the conclusions claimed to follow from them. This, in turn, underscores the high value added that greater formalization offers international relations theory given the current state of development.

At the center of the controversy between offensive and defensive realism is the question: What do certain assumptions about the international system imply about state interactions? In particular, do these assumptions imply that states attempt to maximize their power?

Defensive and offensive realism as well as Waltz's version of structural realism seem to share a consistent set of core assumptions about state preferences and about the strategic arena in which states interact. Recall that Waltz assumes that states are functionally nondifferentiated, purposive unitary actors seeking security and that these states interact in an anarchically ordered system. Call these two core assumptions A_{SR} where the subscript denotes structural realism. The school of defensive realism is still in flux and the core assumptions are still unsettled. Snyder (1991, 10–13), who coined the term, and Grieco (1997, 164–67) seem to posit the same set of core assumptions that structural realism does. Call this set of assumptions A_{DR}, where DR denotes defensive realism, and note that $A_{SR} = A_{DR}$. Glaser (2000b, 6–12), by contrast, explicitly adds the security dilemma and the offense-defense balance to defensive realism's characterization of the states' strategic setting.[21] Call this set of assumptions $A_{OFF\text{-}DEF}$, and observe that $A_{OFF\text{-}DEF}$ is really A_{SR} along with offense-defense assumptions; that is, $A_{OFF\text{-}DEF} = A_{DR} \cup \Omega$, where Ω represents assumptions about the offense-defense balance. Finally, offensive realism assumes that states are purposive unitary actors seeking security in a strategic environment characterized by anarchy, that the presence of some offensive capabilities gives the states "the wherewithal to hurt and possibly destroy each other," and that states are uncertain about other's motivations (Mearsheimer 1994–95, 2000). Call this set of assumptions A_{OR}.

20. Glaser (2000b) provides an excellent overview of this debate. Other contributions and analyses include Herz 1950, Mearsheimer 1990, Snyder 1991, Zakaria 1992, 1998, Grieco 1997, Labs 1997, Lynn-Jones 1998, and Rose 1998.

21. Lynn-Jones (1995), Van Evera (1998, 1999), and others call this offense-defense theory. However, as Glaser (2000b) points out, offense-defense theory often adds many more things than the offense-defense balance, and these additions have led to a great deal of confusion.

Although they do not emphasize it, both structural and defensive realism also seem to assume at least implicitly that the actors have some ability to hurt each other. This is, after all, an important part of what makes the strategic environment *strategic,* that is, why states have to be concerned about what other states are going to do. Similarly, both structural and defensive realism seem to presume some uncertainty about motivations (e.g., Glaser 1996, 127, 1992; Waltz 1979, 165) although they do not always emphasize this assumption. If we make these two stipulations, then structural and offensive as well as the minimalist version of defensive realism appear to be based on the same set of assumptions, which to simplify the subscripts will simply be called A. This set, in turn, is contained in the version of defensive realism that includes offense-defense variables. That is,

$$A = A_{SR} = A_{DR} = A_{OR} \subset A_{OFF\text{-}DEF} = A \cup \Omega$$

But if two or more theories make the same set of assumptions, what does it mean for a field like international relations theory to be debating whether those assumptions imply a particular conclusion? More specifically, what does it mean to debate whether the set of assumptions A implies that states try to maximize their power as offensive realism claims (Mearsheimer 1994–95, 2001; Herz 1950) or that they do not as structural realism and the minimal version of defensive realism claim (Waltz 1979, 126; Greico 1997)?

One possible interpretation of such a debate is that at present no theorist has been able to link these assumptions *deductively* to power maximization or the absence of it. In this account, the debate would be between contending *speculations* about what will eventually be shown to follow from a set of assumptions. This kind of speculation can play an important role in the development of a field because it motivates theorists to work on a problem.

However, this interpretation of the debate as contending speculation about what ultimately may be shown to follow from a set of assumptions does not accurately characterize the growing controversy between structural realism, defensive realism, and offensive realism. Mearsheimer (1994–95, 2001), for example, claims that one can show that power maximization follows from A. By contrast, Waltz (1979) and Grieco (1997) claim that A implies that states do not maximize power. Indeed, A implies that "states balance power rather than maximize it" (Waltz 1979, 127).

Two conclusions do seem to follow from this brief review. First, a consistent set of assumptions cannot imply both X and not X.[22] Assuming,

22. A basic result in logic is that if a set of assumptions implies one conclusion, say Y, and its opposite, not Y, then *any* other proposition can be deduced from this set of assumptions. Thus any conclusion can be inferred from a theory built on an inconsistent set of assumptions, and this makes empirical testing of the theory meaningless.

therefore, that A is consistent (something which has not been challenged so far as I know), then A cannot imply that states do maximize power and that they do not.[23] Set A implies either that states maximize power or that they do not, or set A by itself does not imply either of these conclusions. If the latter is the case, then some other assumptions have to be added to A in order to support any deduction about power maximization or the lack of it. And, given the importance attached to the idea of power maximization or balancing, these other assumptions—whatever they might be—would seem to constitute a key part of theory and are too important to be left unstated. Indeed, this is precisely the tack Glaser (1996, 2000b) takes by adding offense-defense assumptions to A.

The second conclusion that follows from this review of offensive and defensive realism is more important for present purposes. The only way that a debate like the one brewing over offensive and defensive realism can exist is if the links between assumptions and the conclusions that do deductively follow from them are very loose and opaque. Once again, the state of the field indicates that there would be a very high value added to being able to forge tighter, more transparent links from assumptions to conclusions.

On Seeking Survival

If a theory takes actors to be purposive, then some assumption must be made about the purposes they pursue. A fundamental assumption underlying much of contemporary international relations theory is that states seek to survive. Waltz, most notably, assumes that "states seek to ensure their survival" (1979, 91), and most subsequent theories or applications have adopted this assumption.[24] But the assumption was commonplace before Waltz made it the foundation for his theory of structural realism (e.g., Herz 1950, 1959; Jervis 1978; Wolfers 1951, 1962). Indeed, the basic idea can be traced back to Hobbes if not Thucydides. Hobbes (1991, esp. 117–21) argued that the fundamental quid pro quo between the sovereign and subject is that the ruler provides security in return for the subject's obedience, whereas Thucydides (1954) believed that the fear and insecurity engendered by the growth of Athenian power was the primary cause of the Peloponnesian War.

23. It is important to separate deductive claims from empirical issues. Whether A implies that states maximize power is a matter of deduction and not of empirical evidence, as Glaser (2000b) points out. Empirically, determining whether states maximize power is relevant for assessing a theory once we know what it implies, but determining what a set of assumptions implies is not an empirical issue.

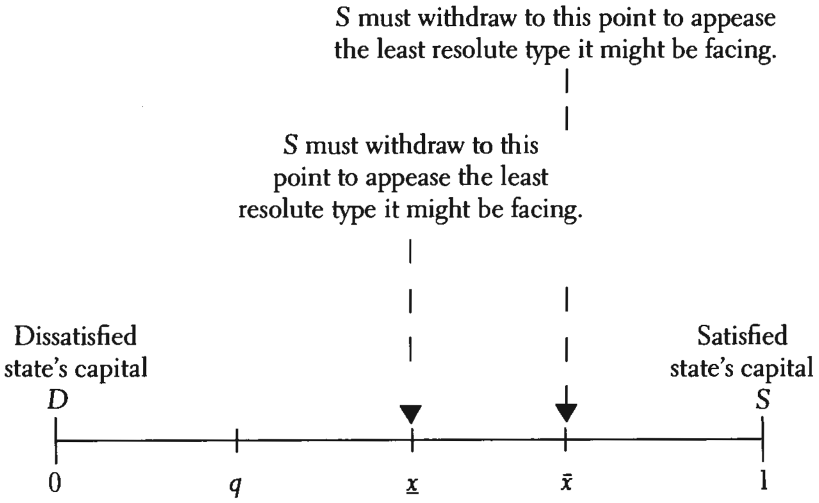
24. See, for example, Brooks 1997, Elman 1996b, Glaser 1992, 1996, 1997, Grieco 1988, 1990, Lynn-Jones 1995, Mearsheimer 1994–95, 2000, Rose 1998, Sndyer 1991, and Van Evera 1998, 1999. Schweller (1996) discusses this assumption in the context of the historical development of realism.

However, the idea that states seek to survive in their anarchic environment is too poorly specified to serve as a theory's core assumption about preferences. Although this assumption is thought to be important, it generally does not do much theoretical work and little if anything follows from it. The reason it appears to be an important assumption and one sufficient for many theories is that the links from this and other assumptions to the conclusions that are claimed to follow from them are too opaque.

To develop these points, suppose that a theorist uses a game to model some interaction over an issue that a theory purports to explain. As noted above, one of the fundamental primitives of this analysis is the actors' preferences, that is, how the actors rank the terminal nodes of the game tree. A theoretically useful assumption about preferences should provide clear guidance as to what to assume about the way that the actors rank the possible outcomes of the game. In economics, for example, the assumption that firms try to maximize profits is useful because it specifies how firms rank outcomes in a wide variety of circumstances. That is, when defining the firm's preferences over the terminal nodes of a game that models some aspect of their interaction, the firms should rank the possible outcomes in terms of the levels of profits associated with those outcomes.

To see that the assumption that states seek to ensure their survival is too vague to provide clear guidance in ranking outcomes and therefore cannot play a role parallel to that of the assumption of profit maximization, consider the following simple bargaining game. A satisfied state S and a dissatisfied state D are bargaining about revising the territorial status quo. The situation is depicted in figure 2 where the total territory is represented

Figure 2 | Bargaining over Territory



as the interval [0, 1]. State D currently controls all of the territory to the left of the status quo border q, and state S controls all of the territory to the right of q.

State D, however, claims to be dissatisfied with the status quo and threatens to go to war with state S unless S makes a territorial concession. To keep things simple, suppose that S can deal with this situation by making a single offer of x to D. If D accepts this offer, it acquires control over the territory to the left of x and S retains the territory to the right of x. If D rejects this offer, then D attacks and the interaction ends in war. To specify the payoffs to fighting, assume that S pays a cost c of fighting and that the war can end in one of two ways. Either S prevails and thereby acquires all of the territory, or D prevails and takes all of the territory. The probability that D prevails is p.

Suppose further that S is unsure of D's willingness to fight. At one extreme, D could be very dissatisfied; this means more formally that S would have to make a large concession to D in order to induce it not to attack. Suppose, in particular, that S would have to withdraw to $\bar{x}$ in order to induce this most-dissatisfied type not to attack. At the other extreme, D could be less willing to use force and would be satisfied if it were to control the territory to the left of $\underline{x}$. In this case, the satisfied state has to make a smaller concession in order to "buy" peace. Finally, suppose that S is sure that D's willingness to fight is between these two extremes but that no point between these extremes is any more likely than another (i.e., D's willingness to fight is uniformly distributed over $[x, \bar{x}]$).[25]

The outcomes of this interaction can be described simply in terms of three elements. If the satisfied state offers x, this offer will be rejected with probability $w(x)$. If this offer is subsequently accepted, the territorial division will be x for D and $1 - x$ for S. If, however, the dissatisfied state rejects x, war follows and the satisfied state prevails with probability p. Thus, the set of outcomes of this interaction can be characterized as $\{x, w(x), p\}$, where x is any number between 0 and 1 and represents what the satisfied state offers; $w(x)$ is the probability of war or, equivalently, the probability that the satisfied state rejects x; and p is the probability that the satisfied state prevails in the event of war.[26]

These outcomes reflect a number of possible trade-offs that may affect the way that a state ranks the outcomes. For example, the state can *ensure* that it will survive as an independent entity by offering $x = \bar{x}$ (or anything more). Conversely, *any* offer less than $\bar{x}$ entails some risk that the state will be eliminated and not survive. This risk is the probability that the dissatis-

25. For elaborations of this basic model and a complete specification, see Powell 1999. Fearon (1995b) discusses a similar model.

26. Given that the dissatisfied state's minimal acceptable offers are evenly between $\underline{x}$ and $\bar{x}$, the probability of war is $w(x) = 1$ for $x \le \underline{x}$; $w(x) = (\bar{x} - x)/(\bar{x} - \underline{x})$ for $\underline{x} \le x \le \bar{x}$; and $w(x) = 0$ for $x \ge \bar{x}$.

fied state will reject x and subsequently defeat the satisfied state, which is $w(x)(1-p)$. The risk is inversely related to the size of the concession; that is, the larger x, the lower the probability of war $w(x)$ and of elimination. Thus, there is a trade-off between territorial concessions and the probability of being eliminated. The more the satisfied state offers, that is, the larger x, the lower the risk of war and elimination, but the less territory the satisfied state will have if its offer is accepted.

In light of these trade-offs, how should S rank these outcomes if it seeks to ensure its survival? This assumption has been variously interpreted to mean that states "give priority to ensuring their security" (Glaser 1996, 127), that states are "security maximizers" (Schweller 1996, 99, 114), or that states have "lexicographic preferences" over security (Schweller 1996, 103). Yet, these different interpretations can have very different implications even in the context of the very simple model described here.

To see this variation, suppose for the moment that security refers solely to the probability of surviving as an independent political entity. Waltz seems to have something like this in mind when he says, "Survival is a prerequisite to achieving any goals that states may have, other than the goal of promoting their own disappearance as a political entity" (1979, 92). This construction suggests that the larger the probability that the satisfied state will be eliminated, the less secure it is.[27] Accordingly, a security maximizer would offer $x = \bar{x}$ and thereby avoid any risk of war. Similarly, a state with lexicographic preferences would always prefer one outcome to another if the former offered a higher level of security. A state with these preferences would therefore offer $\bar{x}$ as well.[28]

But defining *security* solely in terms of the probability of surviving regardless of the territorial sacrifice is only one interpretation of this term, and it is an extremely narrow one. An equally narrow but opposite notion would define *security* solely in terms of territorial integrity. In this context, a security seeker tries to preserve its territorial integrity but has no desire to expand. This would seem to be a plausible interpretation of the idea that the "first concern of states is not to maximize their power but to maintain

27. Kydd (1997b, 121) and Lynn-Jones (1995) interpret Waltz this way and explicitly define security in terms of the probability of surviving. Lynn-Jones also observes that "most definitions of security remain vague" (664).

28. As an aside, one implication of this that has generally not been appreciated in the work on realism is worth noting. Even if a security maximizer (or a state with lexicographic preferences) is uncertain of its adversaries willingness to use force, this state effectively makes a worst-case assumption about its adversary, that is, that its adversary is willing to fight unless it receives $\bar{x}$. This is in keeping with some discussions of structural realism which associate worst-case planning with it (e.g., Brooks 1997; Glaser 2000b). But the security-maximizing state then makes a concession sufficient to appease this worst-case adversary (and therefore sufficient to appease any other adversary) and consequently eliminates the risk of war. *Worst-case planning therefore leads to a lower risk of war.*

their positions in the system" (Waltz 1979, 126). This formulation is also keeping with treating security-seeking states as status quo states.[29]

To trace the implications of this interpretation, suppose that we make the extreme assumption that a security-seeking state is willing to fight in order to preserve its territorial integrity but will not accept any risk of war in order to expand its territory.[30] This state's optimal offer is q. Note further that a state that defines *security* in terms of territorial integrity ranks the outcomes of the game in exactly the opposite way than does a security-seeking state that defines *security* in terms of the probability of survival. In the model, a larger territorial concession "buys" a lower probability of war; that is, if $x > y$, then $w(x)(1 - p) < w(y)(1 - p)$. Thus, a security-seeking state (defined in terms of territory) prefers offering y to x whereas a security-seeking state (defined in terms of the probability of survival) prefers x to y!

Of course, *security* need not be defined solely in terms of either of these extremes. A state's security might be related to both the probability of survival and its territorial integrity. Indeed, this is one plausible reading of what it means for a security-seeking state to give the highest priority to its survival. Unfortunately, this reading of the meaning of security seeking is likely to give rise to yet another ranking of the possible outcomes. If a state is concerned with both the probability of survival and the size of any territorial concessions, then it will offer the concession that balances or equates the marginal benefit of offering slightly more (a benefit which is measured in terms of a slightly higher probability of survival) to the marginal cost of offering slightly more (which is measured in terms of the loss of slightly more territory). The resolution of this trade-off typically results in an offer strictly between $\underline{x}$ and $\bar{x}$. This offer is more than a state that defines security in terms of territorial integrity would offer and less than a state that defines security in terms of the probability of survival would offer.

To summarize this discussion of the assumption that states seek survival, there is no consensus on what the key assumption that "states seek to ensure their survival" means. Different theories seem to be based on different interpretations even though they typically claim to be making the same assumption that Waltz makes. Yet, as the simple model above makes clear, different interpretations lead to widely different preference orderings and to widely different predictions. Effectively, then, these theories are not making the same assumptions, and, once again, the assumption claimed to be doing the heavy theoretical work is not. Rather, the specific formulation

29. See Schweller's (1996) discussion of the relation between neorealism and status quo states.

30. As Glaser (1992) and others observe, sometimes a state has to expand in order to improve the chances that it will not lose what it already has, and this possibility has been central to much recent work on expansion (e.g., Snyder 1991; Zakaria 1992, 1998). However, this complexity does not arise in the simple formulation above.

of what it means to try to ensure one's survival is doing the theoretical work. But because ordinary-language formulations frequently do not force a theorist to specify the range of possible outcomes and precisely how the actors rank them, the fact that "seeking to ensure survival" is too vague to provide the needed guidance in specifying these preferences has generally not been appreciated.

The benefits of formalization are that it requires clearer, more precise assumptions about the actors and their strategic environment and tighter, more transparent links between these assumptions and the conclusions claimed to follow from them. Clear links are moreover critically important to the empirical evaluation of theories. Unless we can be confident about what our theories predict, there is really nothing to test.

But, as noted above, the benefits of formalization also bring costs. The resolution of the cost-benefit trade-off depends at least in part on the state of the field and may change as the field does. In general, the more coherent and clearer existing arguments, the less formalization has to offer and the relatively more important empirical assessment and evaluation may be.

The discussion of structural theories and structurally induced preferences, the nascent debate between offensive and defensive realism, and the assumption that states seek to survive shows that key ambiguities lie at the center of these issues. Indeed, one suspects that the debate surrounding these issues would completely disappear if the assumptions and arguments were more transparent. Either it would be clear that the conclusions did not follow from the stated assumptions, in which case more theoretical work would need to be done. Or, clear but different assumptions would have led to different predictions, at which point the issue becomes a matter of empirical assessment. This current lack of clarity indicates that that branch of international relations theory that tries to understand international politics through an explication of simple stylizations of the international system has much to gain from formalization. At this stage in the development of this branch of international relations theory, the potential for making type II errors by incorrectly accepting fundamentally flawed arguments seems much larger than the potential of making type I errors by incorrectly rejecting sound arguments. The benefits of greater formalization appear to overwhelm the costs.

■ | Some Consequences of Formalization

Although moving toward greater formalization appears to have a high value added, this move still entails costs and consequences. As noted above, one of the likely consequences is the evolution of ignorance. Some conclusions that can be supported with ordinary-language arguments because those arguments *sound* convincing may prove impossible to support formally. Another, more important long-run consequence is the likely reintegration of international relations theory with the rest of political science.

The assumption of anarchy has long served as an analytic wall separating international relations theory from the rest of political science. Recall Dunn's comment in 1949 that the "distinguishing characteristic of IR as a separate branch of learning is found in the nature of the questions with which it deals. IR is concerned with the questions that arise in the relations between autonomous political groups in a world system in which power is not centered at one point" (142, 144). Anarchy is also at the heart of Fox's observation and the challenge implicit in it: "A body of political theory dealing with a system characterized by an absence of a central authority is yet to be developed. . . ." (1949, 79). Most succinctly, "Anarchy is the characteristic that distinguishes international politics from ordinary politics" (Wight 1979, 102). Half a century later, this analytical wall has been breached, and exploiting this breach seems likely to lead to a reintegration of international relations with other fields in political science.

In a thoughtful review appearing in the fiftieth anniversary issue of *International Organization*, Helen Milner (1998) discusses what she sees as an "emerging synthesis of international, American, and comparative politics" based on a rational-institutionalist approach.[31] Each of these fields is too diverse to characterize in terms of a single approach. But, "a strain of such rational institutional work now exists in all three fields, and the scholars who are applying it are asking many of the same questions and using the same analytic tools" (762).

The thrust of her argument is twofold. First, the focus on anarchy and the absence of a higher authority leads to a very strong tendency to treat states as purposive *unitary* actors, which in turn has led to the separation and intellectual isolation of international relations. That is, the emphasis on anarchy renders international relations theory a "narrower specialty, one concerned with a specific type of power relationship among particular kinds of units" (764). The unitary-actor assumption "has cut the field off from other areas of political science" (767). Second, relaxing the assumption that states are unitary actors has important consequences for the field. One of the most important of these is that breaking down the unitary-actor assumption allows international relations theory to focus *explicitly* on the effects that different domestic institutions or different substate-actors' preferences have on international outcomes.[32] These issues, in turn, have

31. Also see her excellent discussion of why the field of international relations diverged from political science (Milner 1998, 762–67).

32. In light of the discussion of structural theories above, it should be clear that taking states to be unitary actors does not preclude the study of the effects of differential *state* preferences. One can, for example, consider the effects of different attitudes toward risk or different discount factors or interests in openness and international trade. However, taking states to be unitary actors does prevent the development of an explicit theoretical link from the preferences of domestic actors or

direct parallels in American and comparative politics, and this provides the basis of the emerging synthesis among at least large and significant parts of these three fields (772–79).

I share Milner's view about the possibility of a synthesis and see this as one of the most exciting developments in international relations theory and political science. But the basis of this potential synthesis seems to me to be different from the one she emphasizes. Milner does see game theory as playing a significant part in this emerging synthesis, because it provides a means for analyzing the strategic interaction that is at the center of the rational institutional approach (783). However, the driving forces behind this emerging synthesis for her are efforts to relax the assumptions that states are unitary actors and that states are the most important or sole actors.

In my view, game theory is the primary carrier of this prospective synthesis. Formalization brings greater clarity and transparency, and this has revealed striking parallels between the strategic problems that states face in international politics and the strategic problems that other actors face in American and comparative politics. These parallels will be the foundation of any synthesis that ultimately emerges, and these parallels exist *even when states are treated as unitary actors.* It is not the assumption that states are unitary actors that has led to the isolation of international relations but the inability to see the strategic parallels that transcend the analytic divide of anarchy.

To illustrate these parallels, consider the bargaining that occurs in the shadow of force in the *anarchic* realm of international politics and the bargaining that occurs between potential litigants in the "shadow of the law" (Cooter, Marks, Mnookin 1982) in the *hierarchical* realm of courts and enforced contracts. If the anarchy-hierarchy distinction had much bite, one would expect it to appear here. Yet, research on bargaining in these two domains has led to strikingly similar questions and hypotheses.

A central question in the work on legal disputes is why do cases go to court?[33] Why, that is, does bargaining sometimes break down in costly litigation and result in a court-imposed settlement? Is there, as Priest and Klein (1984) ask, a relationship between the chances that a dispute will go to court and the strength of the case? Are, for example, clear-cut cases in which the court is very likely to find in favor of one of the parties more likely to go to court than a case in which the ultimate verdict is much less clear? These questions have direct analogs in international relations theory. Why do disputes between states sometimes end in war and costly fighting (Fearon 1995b)? What is the relationship between the risk of war and the distribution of power? Is war less likely if there is a preponderance

from the institutions in which they interact to preferences assumed in the states-as-unitary-actor models. As Milner emphasizes, forging these links requires one to relax the assumption that states are unitary actors.

33. See Cooter and Rubinfeld 1989 for a review of the work on legal disputes.

of power and the victor seems assured, as the preponderance of power school maintains (Organski and Kugler 1980; Kugler and Lemke 1996). Or, is an even distribution of power more stable as the balance of power school argues (Claude 1962; Morgenthau 1948; Mearsheimer 1990; Wright 1965; Wolfers 1962)?

Parallel questions have led to parallel hypothesis. One explanation for why cases go to trial is that the parties disagree about the expected outcome. At least one of the litigants is too optimistic about its prospects of prevailing or about the size of its award (Cooter, Marks, Mnookin 1982; Gould 1973; Landes 1971; Schweizer 1989). For instance, Priest and Klein (1984) argue that excessive optimism is less likely in clear-cut cases. Thus, the probability of litigation is smallest when one party has a preponderance of legal power on its side. Similarly, Blainey claims that a primary cause of war is that states are uncertain of the distribution of power between them and of the likely verdict of fighting (1973, 122). Such disagreements are moreover less likely when one state preponderates. Consequently, a preponderance of power is more peaceful.[34]

A second example of what may be an emerging parallel centers on the effects of shifts in the distribution of power. Ever since Thucydides ascribed the cause of the Peloponnesian War to the rise of Athenian power, shifts in the distribution of power have been seen as a source of tension and conflict in international politics (e.g., Organski and Kugler 1980; Gilpin 1981; Kennedy 1987). Powell develops a model of the way that (unitary) states cope with shifts in the distribution of power and finds that states will make compromises and avoid war if the distribution of power does not shift too rapidly. However, rapid shifts in the distribution of power are doomed to breakdown in war (Powell 1999, 115–48, esp. 128–31).

The relation between political stability and the rate at which the distribution of power shifts appears to be an important mechanism that underlies problems in American and comparative politics as well as international relations. Fearon (1998b), for example, argues that a shift in the distribution of power between a minority and a majority group may be a significant cause of ethnic conflict. Recent work on democratic transitions also suggests that the ability to consolidate a democratic transition depends on the speed at which the distribution of power is shifting among the factions (Acemoglu and Robinson 2001). And, theoretical as well as empirical work links shifts in the distribution of power to the internal structure of political organizations (de Figueiredo 2000a, b).

In sum, one of the most important and exciting consequences of moving toward a more formal standard is that it will make it easier to see the existence of strong parallels between the strategic problems that unitary actors—be they states or not—face in international relations and the

34. For a more extensive discussion of this and other parallels, see Lake and Powell 1999b, 27–9, and Powell 1999, 217–22.

strategic problems that different actors face in comparative and American politics. These parallels presumably exist regardless of the tools one uses to study the problems.[35] In principle, then, a synthesis could emerge without formalization. In practice, however, these parallels have turned out to be much more difficult to see when the analyses of these problems are cast in ordinary-language arguments. Formalization forces the basic mechanism underlying the interaction to the fore and thereby facilities a more-integrated approach to the study of politics. Although not a substitute for good ideas and deep insights, formalization can serve as an important "intuition pump."

Conclusion

Just as game theory is not a substitute for good ideas and deep insights about international politics, good ideas and deep insights are not a substitute for game theory or, more generally, good research methods. Research methods and substantive ideas complement each other, and theoretical progress requires both.

Good theoretical ideas suggest causal relations between a set of assumptions and some conclusions. Game theory helps forge tighter deductive links between these assumptions and the conclusions believed to follow from them. Game theory tests the internal logic of the basic idea. In so doing, it sometimes shows that other important assumptions must be added in order to sustain some conclusions. It also sometimes reveals a fundamental problem with what initially sounded like a very promising idea. Empirical methods in turn provide a way of assessing the empirical power or content of the deductive links.

However, methods and good substantive ideas are more than complements. They are more intimately and subtlety related, for methods do or at least should partly define what a good idea is. In the fiftieth anniversary issue of *International Organization*, Peter Katzenstein, Robert Keohane, and Stephen Krasner describe the evolution and development of the field of international political economy in terms of competing analytic orientations, paradigms, and "isms". This competition has produced lively debates; "substantive findings, however, remain meager" (1998, 683). Barry Eichengreen in the same volume attributes this state of affairs to the lack of "close connections between theory and empirical work" and observes that "the field needs to move in the direction of formulating more parsimonious models and clearly refutable null hypotheses, and developing em-

35. This is largely correct but not entirely so. Looking at different issues through the same analytic lens—a lens which emphasizes some things, like credibility and commitment problems, and deemphasizes other things—will tend to highlight these parallels but not others.

pirical techniques that will allow those theories to be confronted by the data" (1998, 1012). This description applies equally well to international relations theory as a whole. It also suggests that part of what makes an idea interesting is its ability to be formulated in a way that it can be evaluated theoretically and, especially, empirically. Good ideas, therefore, are at least partly defined by the methods we have to develop and evaluate them.

Until the 1980s, ordinary-language arguments were arguably the best analytic tool available for analyzing strategic interaction in a wide range of situations. But better tools are now available, and their application has shown that many widely accepted, ordinary-language arguments are at best incomplete. Formalization has shown too often that the conclusions said to follow from a set of assumptions do not. This, moreover, is not simply a kind of methodological incommensurability, for after seeing the formal result, one can often go back to the ordinary-language argument and find key assumptions that were not previously noticed and that once made explicit greatly weaken the argument. Nevertheless, the extent to which the field of international relations theory will move toward greater formalization over the next decade or so remains uncertain. What is more certain is that the extent to which international relations theory does move in this direction will have an important effect on what will be written about the field in *The State of the Discipline* in 2020.

Charles M. Cameron and
Rebecca Morton

Formal Theory Meets Data[1]

Political Methodology: Where Do We Stand?

We begin by providing the reader the context for this essay and that by Donald Green and Alan Gerber in this volume. Our device for doing so is a recent article by Nobel laureate James Heckman (2000) in which he discusses the current state of statistical methodology in our sister discipline, economics. Heckman's astute insights into econometric practice help illuminate the current role of statistical methods in political science, both by underscoring what is similar to economics and by revealing what is different.

Causal Analysis and the SEM Approach

As Heckman relates the recent history of econometrics, he assigns a central role to the linear simultaneous equations model (SEM), developed during the middle years of the twentieth century. The SEM allowed researchers to translate economic theory into well-posed, causally oriented empirical models. The concepts created by these researchers—exogenous variables, endogenous variables, causal effects, misspecification, omitted variable bias, the identification problem—after some delay entered political science, where they continue to provide the bread and butter of introductory methods training.

But, as Heckman tells the story, by the mid-1960s the SEM "was widely perceived to be an intellectual success but an empirical failure" (2000, 48). A simple example will help explain why. Consider the path

1. We thank Henry Brady, Alan Gerber, Don Green, Ira Katznelson, Gary King, Lisa Martin, Helen Milner, and participants in the 2000 Political Methodology Meetings, 2000 APSA meetings, and the "State of the Discipline" miniconference for very helpful comments on an earlier version. We also thank the more than 200 political scientists and economists who assisted in the literature survey that accompanies this chapter as a web page. Of course, the usual caveat applies.

Figure 1 | A Structural Model

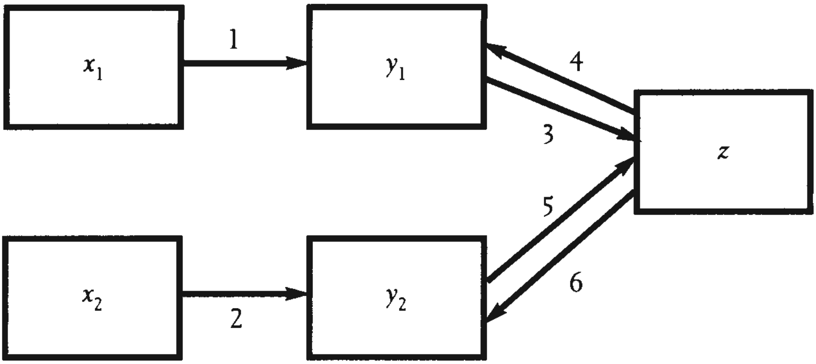


diagram shown in figure 1, which corresponds, to a three-equation model. The path diagram illustrates a structural model, detailing the causal relations between all the variables in the system; it shows how the "shin bone is connected to the knee bone, the knee bone is connected to the thigh bone," and so on. In the diagram, the variables x_1 and x_2 are external, or exogenous, because their values are determined outside the system; the variables y_1, y_2, and z are internal, or endogenous, because their values are determined within the system. Variables like y_1 and y_2 are sometimes called intermediate or intervening variables since their values are determined within the system but in turn determine the value of another variable z. The arrows labeled with numbers indicate the fundamental causal effects of one variable on another. For example, x_1 has a direct causal effect on y_1 via path 1, an indirect effect on z via paths 1 and 3, and an even more indirect effect on y_2 via paths 1, 3, and 6. Direct causal effects are sometimes called structural causal effects. Within the traditional SEM approach, all the structural causal effects in the figure are assumed to be linear, and the task of the empirical analyst is to use data to estimate each of them. Alternatively, if the interest were just variable z, the analyst might use the structural model to justify estimating a reduced form equation $z = z(x_1, x_2)$, relating z only to the two exogenous variables. This is a valid practice since the structural model shows that z ultimately is a function only of the external, or "forcing," variables.

Heckman suggests that the style of thinking illustrated in figure 1 was an intellectual triumph of twentieth-century social science. It's not hard to see why. The concept of stable causal effects is indispensable in most approaches to social science, and the structural approach offers a rigorous yet tractable way to think concretely about causal effects. Because the approach is so grounded in causal thinking, structural parameters are often social-scientifically transparent; they are readily interpretable in terms of theory. Hence, empirical estimates of them can be used to test theory. In addition, empirical estimates of causal parameters can be used for forecasting or performing "what if" policy experiments. Finally, the SEM approach clarifies the limits of purely empirical evidence, by focusing attention on the necessity of identifying assumptions. It is for these reasons that Heckman calls the approach an intellectual success. Most political methodologists would agree.

Yet, as Heckman indicates, the SEM approach soon ran into empirical difficulties, at least in economics. To put the matter simply, only rarely did estimations uncover stable causal relationships (the direct path effects in figure 1). The problem was most notorious in empirical macroeconometrics but, as Heckman documents, widespread in other fields like empirical labor economics. Absent stable structural parameters, the framework illustrated in figure 1 implodes. Heckman suggests that most of what occurred in the methodology of empirical economics since the mid-1960s represents a response to the perceived failure of the SEM program to uncover stable causal parameters.

POST-SEM DEVELOPMENTS IN POLITICAL SCIENCE

Political scientists did not suffer the demoralizing failures of the empirical macroeconomists. Nonetheless, many of the new developments in political methodology have followed or even imitated the post-SEM moves in economics, though sometimes the intellectual origins in the SEM approach are overlooked.

Broadly speaking, there are four such moves. The first (and most conservative) locates the failure of the SEM approach in *tools*. Proponents of this approach argue that simple linear models were too rigid or otherwise inappropriate for political data. Their approach emphasizes the development and use of more appropriate, powerful, or flexible tools—duration models, event count models, and general additive models. It also emphasizes more powerful methods for estimation, including computer intensive methods like bootstrapping or Markov Chain Monte Carlo (MCMC) methods from Bayesian statistics. Not surprisingly, the tools approach is extremely popular among political methodologists. This volume does not contain a tools-oriented essay—they tend to be rather technical—but many are available for interested readers.[2]

The second and third approaches are somewhat more radical. The second locates the failure of the SEM approach in *theory*. It argues that the social scientific theories underlying early efforts were too weak or ill

2. Here is a selection of recent essays and monographs for different models: duration models—Box-Steffensmeier 1998, Gordon forthcoming, and Therneau and Grambsch 2000; count models—King 1989 and Cameron and Trivedi 1998; bootstrapping—Davison and Hinckley 1997; general additive models—Hastie and Tibshirani 1990 and Beck and Jackman 1998; and MCMCs—Congdon 2001, Gelman et al. 1995, and S. Jackman 2000.

formed and the link between the empirical analysis and the theories was too tenuous to sustain well-grounded analysis. We pursue this response in the remainder of this chapter. The third approach locates the failure in *data*. It argues that structural estimation with nonexperimental data is often doomed to failure. Consequently, it focuses on acquiring much better data—whether from laboratory experiments, field experiments, or so-called natural experiments. Stronger data can allow one to establish clear causal effects, often without deploying much social scientific theory. The essay by Green and Gerber in this volume pursues this line of thought.

The fourth approach is perhaps the most radical. It takes a step away from the causal thinking at the center of the structural approach. It argues that identifying assumptions making a strong distinction between exogenous and endogenous variables are untenable. Hence, the best one can do is to stay close descriptively to the data and make short-term forecasts. This line of thinking leads to vector autoregression (VAR) approaches in time series and neural net or other highly black-box approaches in cross-sectional data. The fourth approach is not represented in this volume, but interested readers may consult, for example, Beck, King, and Zeng 2000.

The first three approaches tend to have different adherents, who sometimes disagree intensely. But all three approaches are broadly complementary. Few of their adherents would argue against deep theory, closely tied to empirical analyses employing appropriate tools and strong, on-point data. In our view, the fourth approach runs the risk of throwing out the social scientific baby with the methodological bathwater, but data description and forecasting have a place too. In sum, political methodology is more heterogeneous and less naive than it once was. Yet almost all of its practitioners remain strongly committed to the ideal of *causal inference in service to causal reasoning.*

THIS ESSAY

The remainder of this essay explores new efforts to link theory and data in what we call (for want of a better term) formal empirical (FE) work. In addition, we have constructed a web page with supplementary material (a link may be found at http://www.columbia.edu/~cmc1). This web page lists hundreds of FE articles, books, and working papers, classified by topical subject in political science. The topics range across voting and elections, international political economy, war studies and international relations, legislative studies, executive studies, judicial politics, democratization, a grab bag of topics in comparative politics, and many, many more. Collectively, these studies give the lie to the claim that formal models are never tested in political science. This canard was an overstated but nonetheless plausible description of the state of the discipline in 1985 or even 1990. It grossly misrepresents the practice of political science ten years later.

The essay has the following organization. In the next section, we explore the empirical content of formal models, using a simplistic model as an expository device. The following section provides thumbnail sketches of some interesting real examples. The last section presents an incomplete and prejudiced overview of the substantive accomplishments of FE work.[3]

The Empirical Content of Formal Models

When we say "social scientific theories underlying early efforts were too weak or ill formed, and the link between the empirical analysis and the theories too tenuous, to sustain well-grounded analysis," what do we mean? In this section, as a pedagogic device we construct a rather old-fashioned, tinker-toy formal model and show how to use it to structure empirical work. The model and empirical implementation illustrate in bare bones form the SEM approach, thereby providing a benchmark (or, perhaps, straw man). We then discuss why contemporary theorists view models of this kind as inadequate and how they are moving beyond them, and why some FE analysts see this style of empirical work as unsatisfactory and how they are attempting to forge stronger data-theory links. But we also discuss the virtues of inadequate models for structuring empirical work, for at some level, all models are inadequate.

A Motivating Example: Political Outcomes in a Democracy

What is the relationship between voters, interest groups, politicians, and political outcomes in a democracy? This is a central question in modern political science. Even a cursory summary of the many relevant literatures is beyond our scope. However, for purely pedagogic purposes, pluralist "theory" of the 1950s and 1960s affords a starting point. A caricature of pluralist notions is the parallelogram of forces: government policy reflects the vector of forces created by different pressure groups (Truman 1971 [1951]).

How might one formalize this parallelogram and use it to structure empirical work? The Chicago political economy models of pressure group politics represented an early effort to do so (Stigler 1971; Peltzman 1976; Becker 1983).[4] Most formal political theorists now see these models as in-

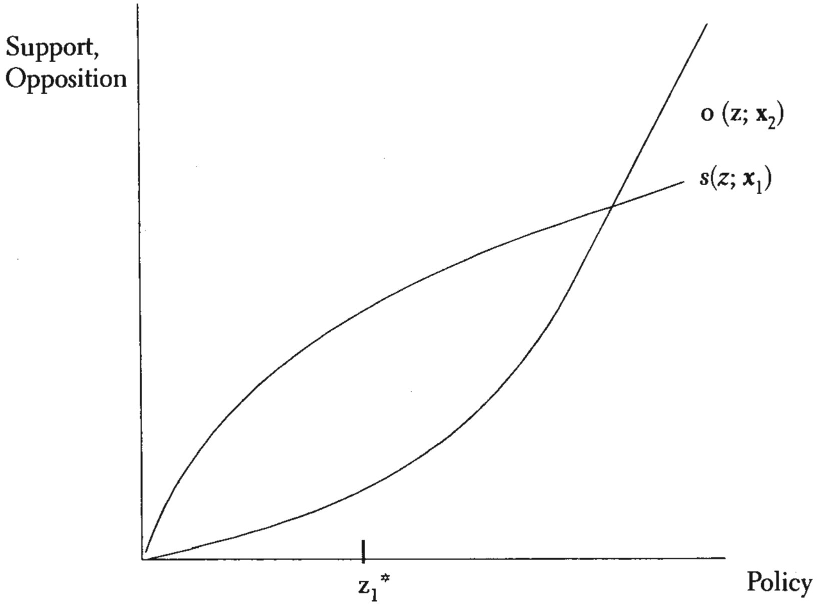
3. Morton's *Methods and Models* (1999) is a book-length exposition of many of the ideas in this chapter. Bates et al.'s *Analytic Narratives* (1998) provides a somewhat different take on FE work, stressing qualitative historical data.

4. Becker's version is rather different from Stigler's and Peltzman's. In particular, in his game theoretic formulation, interest groups are the actors. But both approaches, as well as more recent interest group models in political economy (e.g., Magee, Brock, and Young 1989) and rent-seeking models in public choice economics adopt a broadly similar, black-box approach to elections and political institutions.

adequate. But they afford a relatively painless entree to FE work, because their extremely simple structure allows a clear demonstration of the SEM approach.

In simplest and most schematic form, we imagine a single actor, the government, facing a political support function $s(z; \mathbf{x}_1)$ and a political opposition function $o(z; \mathbf{x}_2)$, where z represents a government policy (e.g., a tax, tariff, subsidy rate, or a liberalism-conservatism score for some complex policy). In addition, both support and opposition are functions of many other variables too ($\mathbf{x}_1$ and $\mathbf{x}_2$ respectively).[5] In the Chicago tradition, these functions are assumed to be everywhere twice continuously differentiable. In this sketch model, we imagine the government setting policy to maximize its net political support, that is, to maximize $n(z; \mathbf{x}_1, \mathbf{x}_2) = s(z; \mathbf{x}_1) - o(z; \mathbf{x}_2)$. Models in this tradition typically assume political support and opposition both increase with the level of the policy, but benefits do so at a decreasing rate while costs do so at an increasing one. Denoting the marginal change in s with respect to z by s' and the marginal change in that change with respect to z as s', and similarly for similar changes in o, it is as-

Figure 2 | Equilibrium in the Expository Example. $z_1{}^*$ indicates the policy leading to maximum net support (that is, the greatest difference in "support" and "opposition").



5. The bold notation, e.g., $\mathbf{x}_1$, denotes a vector, that is, a group of variables. In contrast, x_1 denotes a single variable.

sumed that $s' > 0$, $o' > 0$, $s'' < 0$, and $o'' > 0$. Standard techniques from calculus show that the solution to the government's policy-setting problem is characterized by choosing a level of policy z^* that equates marginal political support and marginal political opposition (that is, $s' = o'$) provided $s'' < o'$ at z^* (which is true by assumption). This answer has an intuitive plausibility: if the government set policy lower than z^*, it could increase its net political support by increasing the level of the policy, while if it set it higher than z^* it could increase net political support by lowering the level of the policy. Figure 2 illustrates the solution graphically.

What is the empirical content of this tinker-toy model? On the one hand, it appears to have a great deal of content in the form of extremely strong and highly contestable assumptions. Among these: there is some entity that can be meaningful considered a policy-setting actor, and it seeks to maximize its net political support; there is something that can be meaningfully thought of as political support, which increases in the level of the policy; and so on. From this perspective, a test of the model could come through direct empirical evaluation of the basic assumptions. On some occasions, as discussed in Morton 1999, *assumption evaluation* of formal models is quite feasible, as when basic assumptions are amenable to straightforward empirical evaluation. When practical, assumption evaluation is appealing because of its directness and simplicity. Unfortunately, many assumptions are not amenable to simple empirical evaluation. Moreover, all theorizing involves abstract concepts and maintained assumptions. Direct assaults on these frequently bring the response "it all depends on how you think about it," or "you are being too literal," or even "you are missing the point." Further discussion then assumes an unproductive, quasi-theological quality. For these reasons, assumption evaluation only rarely proves decisive in practice.

An alternative approach concedes the model's maintained assumptions and then asks, *granting these*, what is the model's empirical content? In some sense this is a charity principle. But it is also affords an even tougher test than assumption evaluation, for if the model says nothing useful even *after* we grant it its underlying assumptions, then it has little to recommend it. Because this approach focuses on the model's empirically observable predictions, Morton (1999) calls this *prediction evaluation.*

If we grant the simplistic pluralist model its maintained assumptions, what is its empirical content in terms of empirically testable predictions? The answer is *nothing*, at least so far. Directly evaluating whether policy has been set to equate marginal support and marginal opposition is an impossible task. So the model has no real empirical content yet.

To imbue the model with empirical content, we must think harder about the impact of *observable* variables on other *observable* variables. Because the model is constructed to explain the level of policy—presumably an observable quantity—z is an obvious candidate for one variable. Observable exogenous variables (or observable but logically prior intervening ones) supply the other candidates. Hence, the question the model must answer is, what is the impact of (observable) exogenous variable x on endogenous variable z? More specifically, in this simplistic pressure group model, the logical chain of inference will run: effect of (observable) exogenous variable x on either (unobservable) intervening variable political opposition o or (unobservable) political support s; effect of (unobservable) support s or opposition o on (observable) policy level z; hence, effect of (observable) exogenous variable x on (observable) endogenous variable z.

There is a direct link between this reasoning and the logic of the SEM approach. In fact, the pluralism model has exactly the form shown in figure 1, with s corresponding to y_1 and o corresponding to y_2. A full-scale empirical implementation of the structural model would use data to estimate the direct effects on support and opposition of the exogenous variables and the policy level z, and the direct effects of support and opposition on effect z. However, *given* the theory embodied in the structural model, one is justified in moving to the reduced form policy equation $z^* = z(x_1, x_2)$. This is exactly the logic of the last part of the preceding paragraph.

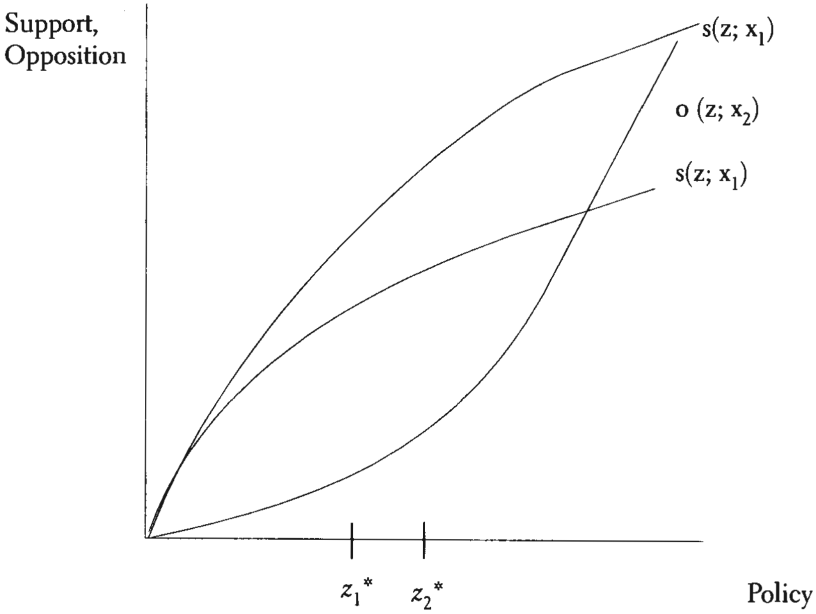
There are many plausible candidates for exogenous variables that enter the political support and opposition functions. Examples include the size of the groups that support or oppose the policy; the wealth of the members of those groups; the geographic dispersion of group members; the organization and procedural rules of the government, including the control of key proposal and veto points by supporters or opponents of the policy; the electoral rules that select politicians and allow supporters or opponents to reward or punish them for their actions; laws that control the use of money in politics; and so on. The silence of Chicago-style models on these points is their Achilles heel; momentarily, we discuss the consequences for theory and empirics of thinking hard about these matters. But to pursue the immediate pedagogic point, assume for the moment that somehow we specify that variable x_1 increases the government's political support at all positive levels of z while leaving political opposition unaffected.

Figure 3 illustrates the effect of x_1 on the support curve: as x_1 increases, the support curve rotates upward. In turn, this rotation changes the location of the point of maximum difference between the support and opposition curves. Consequently, z^* increases. Thus the theoretical result $\frac{\partial z^*}{\partial x_1} > 0$. (One reads this quotient as, "a marginal increase in x_1 strictly increases policy.") This hypothesis, and others like it, constitutes the empirical content of the model.

The figure illustrates the logic of the result but does not constitute a valid proof. A set of techniques called *comparative statics* supplies mathematical tools for carefully deriving such hypotheses from underlying assumptions.[6] In the interest of economy, we forgo a demonstration.

6. Explication of these tools may be found in most textbooks on mathematical economics and in many microeconomics texts. No political science textbook that we know exposits these techniques—a telling omission, perhaps.

Figure 3 | A Simple Comparative Static in the Expository Model. An increase in x_1 leads to a shift upward in the support curve, raising the level of the optimum policy from z_1^* to z_2^*.



However, it is important to note that the formal comparative static result $\frac{\partial z^*}{\partial x_b} > 0$ is not like the "hypotheses" discussed in statistics texts, which might better be described as empirical conjectures or, bluntly, hunches about data. Rather, within FE work, a hypothesis *is a formal mathematical result derived from the basic assumptions of the model*. In our view, the epistemological status of a formally derived hypothesis is quite distinct from an informal hunch about data, however acute.

How would one use the simplistic pressure model to structure empirical work? Working within the SEM tradition, first one would show that the reduced form policy equation $z^* = z(x_1, x_2, \ldots, x_m)$ is linear in the x's (some might simply assume this). Hence, the theory specifies $z^* = b_0 + b_1x_1 + b_2x_2 + \ldots + b_mx_m$. To turn this equation into a stochastic relationship that one could take to actual data, one would add a convenient white noise error term (say, u) to create $z = b_0 + b_1x_1 + b_2x_2 + \ldots + b_mx_m + u$. The error term would be rationalized as reflecting omitted variables, measurement error in z, or inherent randomness in society. Given a modest quantity of data, the parameters in this stochastic relationship can be estimated via multiple linear regressions, yielding actual numerical estimates, with standard errors, for each of the parameters, for example, $\hat{b}_1$ for b_1. It is important to note that the formally derived comparative static hypotheses correspond *exactly* to the parameters in the linear reduced form equation; for example, in the linear reduced form equation $\frac{\partial z^*}{\partial x_1} = b_1$. Consequently, the estimated coefficients, for example, $\hat{b}_1$, and their standard errors allow direct statistical tests of the model's formally derived empirical content!

THE VIRTUES OF INADEQUATE MODELS

Before examining the shortcomings of this style of formal model as theory and this methodology as a template for FE work, it is worth pausing briefly to consider the virtues of even inadequate formal models *as devices for structuring empirical work.*[7] These virtues are almost nascent in the simple parallelogram model, which after all is just a pedagogic device. Yet one can still discern their outlines. Three virtues stand out: clarity, rigor, and unity.

1. Testable hypotheses formally derived from explicit assumptions have the advantage of *clarity*. An overt casual mechanism generates the predictions in a transparent way. This transparent chain of logic may lead to surprises. It may uncover unexpected ambiguities. It may even reveal inadequacies in the underlying assumptions and provoke a reformulation of the theory. But in all cases, everything is laid on the table for inspection.

2. Testable hypotheses formally derived from explicit assumptions have the advantage of *rigor*. Reasoning from abstract assumptions to concrete empirical predictions can be extremely difficult, particularly about situations in which strategic interactions, expectations, beliefs, and communication work in subtle ways. The mathematics of comparative statics makes it possible for an analyst to check rigorously for mistakes in his or her own reasoning. It also makes it possible for others to confirm that the analyst's reasoning is correct—the theoretical equivalent of replication in empirical studies.

3. Testable hypotheses formally derived from explicit assumptions have the advantage of *unity*. In nonformal empirical work, analysts frequently sketch the reasoning behind multiple empirical conjectures. But the reasoning behind one conjecture may have little in common with that behind another. The two may even employ contradictory assumptions. Using a formal model to explicitly derive empirical hypotheses prevents this absurd situation.[8] It may also become clear which hypotheses can only be

7. Morton (1999) addresses the issues in this section more carefully, supplying many concrete illustrations.

8. An illustration: Lax and Cameron (2001) present a formal model of opinion assignment on the Supreme Court and formally derive a series of hypotheses. Many of these hypotheses already exist in the nonformal empirical literature but are motivated by a series of ad hoc and sometimes contradictory assumptions. In contrast, Lax and Cameron show that all can be derived from a single set of assumptions (assigners are effort-constrained maximizers of policy) and the same causal mechanisms.

derived from a different set of assumptions, opening the way to a test of competing models.[9]

Paradoxically, many political scientists seem to see the clarity, rigor, and unity of formal modeling as its principal *disadvantage* for empirical work. For some researchers, a fuzzy, possibly erroneous prediction from a nonformal model with unknown assumptions is somehow preferable. The unspoken belief seems to be that if you keep your assumptions implicit or hidden (perhaps even from yourself), you haven't made any assumptions! Of course, this position is utterly mistaken. Behind the selection of facts and narrative strategy of every case study and behind every regression in every quantitative analysis, lie a multitude of assumptions. Refusing to face those assumptions and their logical consequences is no virtue.

TOWARD BETTER THEORY

In our exposition of SEM-like methods using the parallelogram model, we alluded several times to its inadequacy as theory. What's the problem, and what are the implications for FE work of building better theory?

The parallelogram model shoves most of the interesting politics in policymaking into the support and opposition functions, which are severely undertheorized. We concur with theorist David Austen-Smith's evaluation of the result:

> Questions about how and when influence is effective in majoritarian legislatures, about why some groups have "intrinsically more influence" than others, about how and why resources devoted by groups should map deterministically into a legislative decision and subsequent bureaucratic execution, about why groups adopt different patterns of activity (campaign contributions, informational lobbying, grass-roots activism, etc.), and so on, simply cannot be posed with the aggregate framework [i.e., an approach like the parallelogram model]. (1997, 299)

Austen-Smith concludes that better theory must be "micro-oriented" and more firmly grounded institutionally. Absent this kind of theory, it is hardly surprising that empirical estimates of parameters are unstable.

Austen-Smith's conception of better theory is widely shared among formal political theorists. It is exactly the direction contemporary formal theory is developing in American politics, comparative politics, and international relations. Within political economy, some researchers have preferred to retain very austere depictions of politics but embed them within quite elaborate models of economies, for example, of international trade. But others are moving to build models with strong micro- and institutional foundations.

9. Morton (1999) discusses alternative model evaluation at some length.

Models of this kind raise new issues for FE work. They are inevitably game theoretic. Comparative statics becomes subtler and more difficult.[10] The smooth, linear, continuous functions assumed in our expository example often go out the window. For example, in models using spatial theory, there can be distinct regimes. The behavior of political agents may differ dramatically across these regimes, but the conditions separating them may be difficult to observe empirically.[11] Problems like these raise a host of new and difficult methodological issues that need to be addressed in the years ahead.

FORGING STRONGER LINKS BETWEEN THEORY AND EMPIRICAL WORK

Our review of the current state of FE work in political science indicates that the average study consists of a simple formal model, used to gain insight into a phenomenon, followed by some empirical work (a case study or quantitative data analysis) loosely suggested by the model. The empirical work is inspired by the theory because the researcher uses the formal model to suggest relationships to look for and variables to employ. But the logical connection between theory and empirics is rarely closer than that. For example, the analyst may fail to derive formal comparative statics, consider the model's structural form, or ponder restrictions imposed on reduced forms. This inspired-by methodology gains the analyst the power of formal models for theorizing, which is no small matter. But it throws away the virtues of formal models for structuring empirical work. We predict FE work in political science will increasingly move away from the inspired-by methodology toward the SEM-like methods exposited above.

At the cutting edge of FE work, however, analysts are challenging certain elements of the SEM approach. We cannot hope to resolve the issues being raised, especially since many remain unsettled. But we can give the reader at least some sense of the debates.

Broadly speaking, two issues are at play:

- Is the formal model to be viewed as a *complete* model of the data generating process or as a *partial* model of the data in hand? The answer to this question determines the legitimacy of control variables and arbitrary assumptions about functional forms.

- What is the true nature of the stochastic element in the estimating equations, and what is its relationship to the formal model? The an-

10. A selection of essays addressing this point includes Dixit 1986, Fudenberg and Tirole 1984, and Hirshleifer and Rasmusen 1992.

11. A selection of essays addressing or illustrating this point includes Cameron, Segal, and Songer 2000, Lee and Porter 1984, Moraski and Shipan 1999, Segal and Wang 2001, and Spiller and Gely 1992.

swer to this question determines how the empirical evaluation of the formal model is to be interpreted.

Viewing a formal model as, at best, a partial picture of reality seems like common sense. Thus, it may seem natural to include control variables in the empirical analysis (control in the sense of multiple regression, not experimental manipulations), variables missing from the formal model. It may seem equally straightforward to specify tractable functional forms for estimating equations, so one can use standard statistical packages or results, even though one is unsure whether the specified form is compatible with the model's assumptions. But each such move drives a wedge between the formal model and the empirical analysis. At some point, the wedge becomes so large that an analyst is no longer using the formal model and certainly isn't testing it.

These problems have led some researchers to treat formal models *as if* they specify the complete data-generating process. In this view, if a variable isn't important enough to include in the formal model, it shouldn't be in the empirical work. Similarly, estimating equations should be strictly derived from the formal model, whose assumptions must be explicit enough and tractable enough to allow such derivations.

Similar issues arise concerning the stochastic elements in formal models versus estimating equations. Adding white noise error terms to deterministic models seems at best arbitrary and may lead to logical absurdities. For example, if the actors in a political situation understand that their world has a degree of randomness, this understanding is apt to affect their behavior. But if so, a deterministic formal model of the situation is simply wrong, and sprinkling white noise error terms in estimating equations won't fix it. This line of thought leads to incorporating stochastic elements directly in formal models (usually as games of imperfect or incomplete information) and carrying those stochastic elements through into the empirical estimation.[12]

These efforts may seem like an extraordinary effort simply to achieve logical consistency. Whether they are truly necessary remains controversial within the FE community.[13] In addition, such efforts often require considerable technical prowess. But they promise a payoff that political scientists have been slow to grasp. If full-blown, ultrarigorous structural models actually uncover stable causal parameters, *it becomes possible to perform theory-driven, data-sensitive policy experiments.*

12. Signorino (2000) provides a helpful analysis of these and related issues. In addition, Signorino has written software that estimates structural parameters for several commonly encountered strategic models, incorporating explicit, theoretically plausible stochastic elements. This software is currently available at http://www.rochester.edu/College/PSC/signorino/.

13. They have become quite common in economics, however. Illustrative examples are Donald and Paarsch 1996 and Rust 1994.

To see the point, refer again to figure 1. Let x_1 be a novel policy intervention (e.g., a change in a political institution, like voter registration requirements) with uncertain effect on y_1. If the *remaining* parameters in the model are truly stable, then one can use them to predict the effect of the policy intervention on outcome z, under different assumptions about the effect of x_1 on y_1. Theory-driven, data-grounded "what if" experiments about political and institutional reforms might well deserve the attention of citizens and policymakers. Where data and theory allow such "what if" experiments, the possibilities are exciting.

CONCLUSION

Nonformal empirical work can be informative and revealing. However, the move from nonformal empirical work to inspired-by studies lets the analyst tap into the power of formal models for the purposes of theorizing. The move from inspired-by work to SEM-like work gains the analyst clarity, rigor, and unity in empirical hypotheses. The move from SEM-like work to rigorous structural estimation opens the door to theory-driven, data-grounded policy and political analysis. In our view, each of these moves is valuable.

Illustrative Examples

Space prohibits extensive consideration of real examples, which are invariably far more complex than the previous section's simplistic expository device. Instead, we supply some thumbnail sketches of illustrative examples that interested readers might wish to pursue. We group these into three categories. First are works that take seriously Austen-Smith's call for micro-oriented, institutionally rich theory seriously, and then match these models with SEM-like empirical methods. Second are works that retain somewhat summary models of politics but move toward rigorous structural estimation. Third are works that meld micro-oriented, institutionally rich theory with rigorous structural estimation. The selection of these examples is necessarily arbitrary; many more examples in all substantive fields of political science are compiled on the accompanying web page.

MICRO/INSTITUTIONALIST THEORY, SEM-LIKE METHODS

AMERICAN POLITICS EXAMPLE

Filer, Kenny, and Morton (1993) examine the effect of income on voter turnout. They propose a game theoretic model in which highly motivated elites within social networks in turn motivate others to vote. The authors solve the elites' strategic turnout game and then formally derive comparative static predictions about group turnout by income class. They test the formally derived predictions against country-level data from presidential elections.

COMPARATIVE POLITICS EXAMPLE

Huber and Shipan (2001) examine the efforts of legislators to control bureaucrats through the design of statutes, especially their specificity. The authors construct a formal model of statute design, incorporating substantial variation in institutional arrangements. They formally derive a series of comparative statics, focusing on issues like the extent of policy conflict between legislators and bureaucrats, the internal capacity of the legislature, conflict across chambers in bicameral legislatures, and the availability of nonstatutory means for controlling bureaucrats (e.g., legislative vetoes). They test the formally derived hypotheses with remarkable original data on the specificity of statutes. An interesting element of this analysis is its use of comparative data from the U.S. states simultaneously with cross-national data.

INTERNATIONAL RELATIONS EXAMPLE

Schultz (2001) uses FE work to examine information, threats, and war fighting—issues related to the Democratic Peace. The author builds a formal model of a unitary democratic state engaged in crisis bargaining. In the model, the monopoly party's actions convey information to domestic voters and a foreign bargaining partner. The critical comparative static comes from adding a domestic opposition party, whose support or opposition to the governing party also transmits information. Schultz shows that this institutional difference decreases the probability of war, decreases the likelihood that the democratic state makes threatening moves, and increases the effectiveness of the threatening moves it makes. Schultz then tests these and other propositions empirically, focusing on threatening moves and their effectiveness.[14]

INTERNATIONAL POLITICAL ECONOMY EXAMPLE

Mansfield, Milner, and Rosendorff (2000) examine the effect of regime type (democracy, autocracy) on the likelihood of concluding trade agreements. They construct a formal game theoretic model in which a government undertakes international agreements, which strategically convey information to voters about the government. The authors solve the game between the government and voters, allowing variation in the electoral

14. This choice is of methodological interest. The direct empirical evidence on the Democratic Peace appears inherently too weak to make a definitive determination of the existence and origins of this phenomenon (see, e.g., Gartzke 1998). Schultz's solution is to build a formal model that can accommodate the Democratic Peace but can itself be tested with different, less ambiguous data—an excellent demonstration of the power of joining formal models and empirical evaluation.

power of voters (hence, regime type.) They formally derive empirically testable hypotheses using comparative statics, which they test with data from recent trade agreements.

AUSTERE POLITICAL THEORY, STRUCTURAL ESTIMATION

INTERNATIONAL POLITICAL ECONOMY EXAMPLE

Goldberg and Maggi (1999) test one of the best-known models of endogenous tariff formation, the Grossman and Helpman model (1994), a lineal descendant of the early Chicago models. In this model, two interest groups simultaneously offer a government actor a schedule of campaign contributions, in exchange for trade protection. The government then chooses a level of tariffs. This somewhat austere model of politics is embedded in a fairly elaborate model of an economy. Goldberg and Maggi undertake a structural estimation of the Grossman and Helpman model. By solving the game between the groups and the government, they derive a specific functional form for the relationship between protection and key variables in the model. The stochastic element in the model is explicitly rationalized as measurement error. The authors estimate this equation using cross-sectional data from the United States for 1983, collected at the level of 3 digit SIC codes. The extensive use of sensitivity tests is a particularly interesting feature of the analysis.

MICRO-INSTITUTIONALIST THEORY, STRUCTURAL ESTIMATION

AMERICAN POLITICS EXAMPLE

Schachar and Nalebuff (1999) use a structural approach to empirically evaluate a group-voting model with some similarities to Filer, Kenney, and Morton's model. Shachar and Nalebuff use their formal model to construct a maximum likelihood function for estimating turnout levels by state in presidential elections. The maximum likelihood function's form is directly derived from the formal model's equations. It is quite complex, incorporating a multitude of parameters in the formal model. An interesting feature of the analysis is the use of a structural model to conduct a modest "what if" experiment, concerning higher levels of turnout. They argue that if turnout levels had been 100 percent, Republicans would never have gained the presidency in the post–World War II era.

COMPARATIVE POLITICS EXAMPLE

Diermeier and Merlo (1999) consider a formal, stochastic bargaining model of government formation and duration in parliamentary democracies. The authors derive a parametric structural estimating equation that is compatible with the formal model. They estimate this with maximum likelihood. The data come from 236 governments in 9 countries over the period 1947 to 1997. Using the estimated structural causal parameters, the authors undertake a series of policy experiments that examine the effects of different institutional rules (the investiture rule and the constructive vote of no confidence) on negotiation duration, government duration, and government size.

INTERNATIONAL RELATIONS EXAMPLE

Signorino and Tarar (2001) examine a straightforward incomplete information model of extended deterrence (that is, whether a country can prevent others from making war on its allies). They derive an estimating equation compatible with the formal strategic model, including the assumed form of incomplete information. They then structurally estimate the model using standard data sets on war. An interesting element of the paper is the use of the estimated structural causal parameters to interpret specific cases in recent history, for example, the Berlin blockade of 1948.

■ | Substantive Contributions of Formal Empirical Work

What are the substantive, as opposed to the methodological, contributions of FE work to political science?—for such is the basis on which FE work must ultimately be judged. We can only offer illustrations, for FE work has become so widespread and diverse that it extends far beyond our ability to evaluate knowledgeably. Nonetheless, these illustrations suggest that substantive accomplishments are relatively abundant and likely to grow.

In evaluating FE work, we use four standards:

1. *Understanding political phenomena and solving empirical puzzles.* Did the empirical work show that a particular formal model affords powerful analytic leverage over an important political phenomenon? In other words, did the FE work show the formal model has empirical punch?
2. *Advancing rich theory and stimulating new theory.* Did the empirical work lend support to a formal model or class of models or lead to the formulation or refinement of new theories or models? In other words, does the FE work lend credence or stimulate formal models with theoretical punch?
3. *Rejecting theory.* Conversely, did the empirical work allow the rejection of a plausible formal model or class of models, as offering relatively little empirical or theoretical leverage over an important political phenomenon?
4. *Improving public policy.* Did the FE work offer convincing grounds for better public policy? Did it work as applied political science?

SOLVING PUZZLES AND UNDERSTANDING THE WORLD OF POLITICS

- *The activity puzzle in congressional studies.* In legislative studies, a puzzle of the 1980s was statistical evidence indicating that congressmen's activities (e.g., constituency service) did not affect their reelection prospects. The FE work in Rivers and Fiorina 1989, employing structural estimation, solved the puzzle.
- *Outlier committees in Congress.* Determining the extent and distribution of outlier committees (ones in which the ideological or policy preferences of members differ significantly from those of the average member) was a central thrust of congressional scholarship in the 1990s. This work was stimulated by Gilligan and Krehbiel's formal models (1990) of legislative organization and Krehbiel's subsequent FE work (1990). The work led to a better understanding of the phenomenon.
- *Duverger's law.* Gary Cox's FE work on party structures remains a touchstone in comparative politics.
- *Veto politics.* Recent FE work on vetoes and veto threats arguably advanced the understanding of the presidential veto and interbranch bargaining in separation of powers systems (C. M. Cameron 2000; Groseclose and McCarty 2001).
- *Gridlock in separation of powers systems.* Krehbiel's FE work (1998) on policymaking in the United States establishes the standard for future work in this area.
- *Legislative delegation to agencies.* Epstein and O'Halloran's FE work (1999) established a coherent theoretical framework, demonstrated the feasibility of systematic empirical study, and documented plausible patterns in legislative delegation to agencies.
- *Referenda and legislative constraints.* E. R. Gerber's FE work (1999) on state referenda shows that the presence of this device forces legislators to remain more proximal to the desires of the median voter in the state.

ADVANCING RICH THEORY

- *The spatial model of elections.* The intellectual elegance of the spatial theory is clear. But FE work helped establish its practical usefulness and in turn further stimulated theorists (Enelow and Hinich 1984). But see "Rejecting Theory," below.
- *Duvergerian conceptions of party structure.* Again, primarily via Cox's landmark work. These theories have become a staple and an intellectual highlight of comparative politics.

- *The monopoly agenda setter model.* Work in FE employing and extending Romer and Rosenthal's monopoly agenda setter model confirms that it is a fundamental tool for thinking about separation of powers systems.
- *Regulatory politics.* The FE work of Chicago political economists fundamentally altered discussions of the subject and exerts considerable influence in IPE studies of tariffs and trade barriers.
- *Informational foundations of war.* Drawing on simple formal theory, Fearon (1995b) argued that war between states is most likely to result from informational, rather than material, factors. This mode of thinking represented a considerable departure from realist views (including ones in early formal models). Ongoing FE work has considerably elaborated and empirically tested these ideas.

REJECTING THEORY

- *Downsian models of party politics.* Work in FE has consistently found a substantial degree of policy divergence between candidates, platforms, or policy proposals, even in circumstances that seem to approximate the conditions for the median voter theorem.
- *The separation-of-powers model of judicial independence.* The case is hardly closed, but accumulating evidence suggests that federal courts in the United States are not as responsive to Congress as simple separation-of-powers models seem to suggest (Segal 1997).
- *Political-business-cycle models.* This is a very complex area, but evidence seems to reject early, naïve theories, though the jury is still out on more-sophisticated models.
- *Commitment model of veto threats.* The evidence on veto threats clearly rejects the simplest version of commitment-type veto threats as a general explanation of how veto threats work.

IMPROVING PUBLIC POLICY

We know of few areas in which contemporary academic political science of any kind has demonstrably affected public policy. Consequently, we simply note a few FE research programs that have the clear *potential* to improve public policy.

- *Juries.* Formal work has suggested that changing jury voting rules might lead to different and perhaps better decision making. Experimental evaluation of different rules is under way (Guarnaschelli et al. 2000). If any of the modifications were adopted, the impact on millions of jurors as well as litigants might be enormous.

- *Campaign finance.* Ongoing FE work (e.g., McCarty and Rothenberg 1996) has obvious policy implications.
- *Sequential elections versus simultaneous elections.* Investigations in FE of sequential elections (e.g., as in presidential primaries sequenced over many states) versus simultaneous elections (e.g., one large national primary) have obvious policy implications (Morton and Williams 2001).
- *Redistricting.* Recent FE work on the policy consequences of racial redistricting has substantial policy implications (Shotts 2001).
- *Extended deterrence in international relations.* Signorino's use (1999) of a structural model to interpret historical cases of extended deterrence raises the possibility of using this and similar structural models for predictive and policy analytic purposes.

Summary: The Trajectory of FE Work

We believe the trajectory of FE work is illustrated by advances in an area one of us knows well, turnout in elections. Turnout is a notoriously difficult topic for formal theory; hence, the growth of FE work in such stony soil provides a kind of critical test.

In this area, theorists first formulated a simple, deterministic model of turnout.[15] This model made stark point predictions, allowing evaluation through simple observation of errant observations. These were notoriously abundant. Yet, the initial, brutal collision between theory and data did not lead theorists to abandon their general framework. Instead, most formal theorists temporarily bracketed their failure to explain turnout and used the same general framework to study electoral choices, where it worked much better. Meanwhile, other theorists developed strategic models of the calculus of voting, incorporating the role of groups and the probabilistic nature of leader influence (see the discussion of Filer, Kenney, and Morton 1993, above). Empirical estimations in the SEM style supported the models' predictions. However, the initial rounds of evaluation of the second-generation models stopped short of rigorous structural estimation. Finally, structural estimation of a formal model—using empirical equations directly derived from the formal model's equations—showed support for the formal model as an explanation of voter turnout decisions, explaining nearly 90 percent of the variation in turnout (see Schachar and Nalebuff 1999, above). The structurally estimated models allowed analysts to make theoretically driven, empirically grounded political and policy predictions. Needless to say, this work is not the last word on the subject, but the trajectory in this difficult area is impressive.

15. For a review of the early turnout literature, see Morton 1991.

Conclusion

We are both assertive and modest about the value of FE work. We are assertive in claiming that a thorough, ongoing confrontation between formal theory and data will improve both theory and empiricism in political science and take the discipline a step closer to a cumulative, sophisticated social science of politics. It will improve formal theory by providing feedback on which models have empirical bite and which do not, and thereby stimulate new theorizing along productive lines. It will improve empirical work by allowing analysts to exploit the clarity, rigor, and unity of formal models as devices for generating hypotheses and help them escape the degenerating inductivism so well parodied by the cognitive scientist Allen Newell in his essay, "You Can't Play 20 Questions with Mother Nature and Win." Finally, by providing the profession with a core of clear, deductive models of proven empirical bite, it will advance the discipline as a whole and gradually open the door to theoretically driven, empirically grounded political analysis.

We are assertive in calling for more (and better) FE work in empirical political science, but we also wish to be modest, for two reasons. First—simply to be clear—we do not claim, or for a moment believe, that FE work is the only way to go. The well-established, older style of descriptive and inductive empirical work—quantitative and qualitative—will continue to be essential for providing knowledge about the world of politics. In our view, the only sensible position is that nonformal and formal empirical analyses both contribute to a more robust political science. Similarly, pure formal theory—theory unaccompanied by data and statistical tests—is a precious commodity. Investing in formal theory is investing in our discipline's basic intellectual infrastructure. In fact, formal theory assumes an even greater importance when political scientists are serious about confronting theory with data. But a second and equally important cause for modesty is the difficulty of the enterprise. Combining theory and data involves more than mastering two different skill sets. It requires new ways of thinking and solutions to new methodological problems, problems that political scientists have hardly begun to face. Solving these problems will be a major challenge for the new century of political science.

Donald P. Green and Alan S. Gerber

Reclaiming the Experimental Tradition in Political Science[1]

The history of political science in the twentieth century is typically recounted as the saga of how behavioralism, statistics, and formal modeling freed the discipline from the clutches of storytellers and preachers. The protagonists are the founders of survey research organizations and research programs that emphasize conceptual rigor. Much of the storyline concerns the rise of various theoretical approaches—modernization, dependency, pluralism, rational choice. A lesser, though still important theme is the diffusion and refinement of quantitative methodology, spurred by the development of new statistical methods and easy availability of high-speed computers. Curiously, the most prominent accounts of political science's ascent from darkness into light scarcely mention the sine qua non of science, experiments (Waldo 1975; Easton 1968; Almond 1996).

Before the advent of surveys, formal models, regression analysis, and many other accouterments of modern political science, there existed a brand of political science that was based on field experimentation, the study of controlled interventions into the political world. An early example of such work was Harold Gosnell's study (1927) of voter registration and turnout in Chicago prior to the 1924 and 1925 elections. Gosnell gathered the names, addresses, and background information of thousands of voting age adults living in various Chicago neighborhoods. He then divided these neighborhoods into blocks, assigning certain blocks to the treatment condition of his experiment, which consisted of a letter urging adults to register to vote. Tabulating the registration and voting rates in his treatment and control group, Gosnell found his letter campaign to have produced a noticeable increase in political participation across a variety of ethnic and

1. The authors would like to thank the Institution for Social and Policy Studies, which helped support this research. We are grateful to Henry Brady, Nancy Burns, Gary King, Rebecca Morton, Karen Orren, and the editors for their helpful comments and suggestions. We are also grateful to Matthew Green, Barry McMillion, David Nickerson, and Jennifer Smith for their assistance.

demographic groups. Similarly, in 1935 George Hartmann conducted a controlled experiment in Allentown, Pennsylvania, in which he distributed 10,000 leaflets bearing either rational or emotional appeals for the Socialist Party. Examining ballot returns, Hartmann (1936–1937) found Socialist voting to be somewhat more common in wards that received emotional leaflets. Underhill Moore and Charles Callahan (1943), seeking to establish a "behavioristic jurisprudence," examined the effects of varying New Haven's parking regulations, traffic controls, and police enforcement in an effort to plot a "behavioral response function" for compliance with the law.

These early studies might be characterized as controlled field experiments, as distinct from randomized field experiments. Using certain decision rules, Gosnell, Hartmann, and Moore determined which blocks or wards were to receive their solicitations; they did not assign observations to treatment and control conditions on a purely random basis. In subsequent decades, as the statistical insights of R. A. Fisher (1935) took root in social science, experimentation became synonymous with randomized experimentation.[2] For example, Hovland, Lumsdaine, and Sheffield (1949), working in the Experimental Section of the Research Division of the War Department during World War II, conducted a series of randomized experiments examining the effectiveness of various training films designed to indoctrinate army personnel. While this type of research became more common in psychology than in political science and, at that, more common in the laboratory than in the field, experimentation in naturalistic settings was not unknown to political scientists. Eldersveld's classic study (1956) of voter mobilization in the Ann Arbor elections of 1953 and 1954 built randomization into the basic design of the Gosnell study. Assigning voters to receive phone calls, mail, or personal contact prior to election day, Eldersveld examined the marginal effects of different types of appeals, both separately and in combination with one another.

Although Eldersveld's research was widely admired, it was seldom imitated.[3] To the limited extent that political scientists thought at all about experiments, their prevailing impression was that field experiments typically involve local samples, very specific types of interventions, and little attention to the psychological mechanisms that mediate cause and effect. Each new development in data analysis, sampling theory, and computing seemed to make nonexperimental research more promising and experimentation less so. Once the principles of probability sampling took root in

2. Rubin offers the following definition of an experiment: "If Treatments E and C were assigned to the 2*N* units randomly, that is, using some mechanism that assured each unit was equally likely to be exposed to E as to C, then the study is called a randomized experiment or more simply an experiment; otherwise, the study is called a nonrandomized study, a quasi-experiment, or an observational study" (1974, p 689).

3. Blydenburgh (1971), Gertzog (1970), and Reback (1971) follow Eldersveld's lead but do not appear to have randomized their campaign interventions.

the early 1950s, surveys offered an inexpensive means by which to gather information from nationally representative samples. They could inquire whether the respondent had been contacted by parties or campaigns; indeed, they could examine the psychological mechanisms that might explain why canvassing leads to higher rates of political participation. Moreover, survey data could be mined again and again by researchers interested in an array of different questions, not just the causal question that animated a particular experiment. Surveys seemed superior not only as instruments of measurement and description but also as vehicles for causal analysis.

The narrow purview of experiments also ran afoul of the grand ambitions that animated the behavioral revolution in social science. The aims of science were often construed as the complete explanation of particular phenomena, hence the fascination with the *R*-squared statistic. To students of political behavior, surveys seemed well suited to the task of arranging explanatory variables—economic, demographic, social-psychological—within a "funnel of causality," to borrow a memorable phrase from *The American Voter* (Campbell et al. 1976 [1960]). Experiments, by contrast, could speak to causal questions a few variables at a time. And there could be little hope of using experiments to investigate the big variables that had captured the discipline's imagination—civic culture, identification with political parties, modernization, diffuse support for the political system.

Political science marched on, and field experiments all but fell out of the discipline's methodological kit bag. Riecken et al.'s monograph *Social Experimentation* (1974) mentioned only one field experiment conducted in political science after 1960, Robertson et al.'s study (1974) of the effects of televised public service announcements on behavior. The *Handbook of Political Science* (Brody and Brownstein 1975) devoted a chapter to "Experiments and Simulations." Although the authors praised field experiments, they could point to few examples. Most of these appeared in the young journal *Experimental Study of Politics*, which expired a few years later.

The overwhelming preponderance of empirical work in political science continues to rely on nonexperimental data. To be sure, recent years have witnessed a resurgence of interest in laboratory experiments dealing with topics ranging from media exposure (Iyengar and Kinder 1987; Ansolabehere and Iyengar 1995) to collective action (Dawes et al. 1986) to legislative bargaining (McKelvey and Ordeshook 1990). Surveys with randomized question content and wording have become increasingly common in the study of public opinion, particularly racial attitudes (Sniderman and Grob 1996; Hurwitz and Peffley 1998). Yet, the increasing number and sophistication of such studies has done little to generate interest in field experimentation. Bartels and Brady (1983) make no mention of field experimentation in their synopsis of the discipline's data collection methods. In Donald Kinder and Thomas Palfrey's edited volume *Experi-*

mental Foundations of Political Science (1993), only one of the twenty research essays may be described as a field experiment, Cover and Brumberg 1982. From our canvass of the *American Political Science Review, American Journal of Political Science, Journal of Politics,* and *Legislative Studies Quarterly,* it appears that field experiments were altogether absent from political science journals during the 1990s.

Field experimentation is in our view far more valuable than this meager accumulation of scholarship would suggest. Conversely, nonexperimental research is far more limited in scientific value than its preponderant market share would seem to imply. Nonexperimental analysis, after all, is only informative to the extent that nature performs a convenient randomization on our behalf or that statistical correctives enable us to approximate the characteristics of an actual experiment. Ironically, whether nonexperimental research succeeds in any particular application is difficult to discern empirically without a true experiment against which to compare it.

Our aim is to call attention to the untapped potential of field experiments. Randomized intervention into real-world settings should be a prominent methodology in political science, not a curio from the discipline's distant past. This essay begins by explicating the essential features of experiments, showing why well-conducted experiments are typically more persuasive arbiters of causality than comparable nonexperimental research. Next, we discuss certain important considerations that arise in the design and analysis of experiments, in both the field and the laboratory. We then review the experimental literature in political science, giving special attention to field experiments as distinct from those conducted in artificial settings. After suggesting some potential applications of field experimentation, we review the leading criticisms of experimental social science that have been advanced in other disciplines. Finding these criticisms wanting, we argue that field experimentation should occupy a central place in political science, particularly in subfields where randomized interventions confront fewer practical difficulties.

Why Experiments Are Valuable

Empirical investigation comprises an array of activities, and by focusing here on the investigation of cause and effect, we do not mean to disparage other activities of science that have to do with measurement, description, and classification. Case study, formal modeling, and large-*n* statistical analysis may also be fruitful ways to develop hypotheses. Our focus in this essay, however, is causal inference. How can one gauge the extent to which a change in one variable causes change in some other variable?

For concreteness, suppose one were interested in whether contributing money to a candidate's campaign increases an individual's access to a candidate after he or she is elected. One might formulate the hypothesis that, ceteris paribus, the more a person contributes, the greater the likelihood that he or she will be permitted to schedule an extended face-to-face conversation with the recipient of the donation. Suppose that we had data on how much, if anything, people contributed to candidates and whether their request for an extended meeting was granted. What must we assume about these data if we are to draw inferences about the *causal* influence of donations on access?

One key assumption concerns measurement. The amount of money that each donor contributes—the independent variable—must be measured precisely. Simply asking people to recall the size of their political donations might well generate unreliable data that could lead to biased inferences. A second and often more problematic assumption is that the size of the donation is uncorrelated with other factors that affect whether a meeting request is granted. Intuitively, we sense that nonexperimental data on campaign contributions are unlikely to satisfy this requirement. A candidate's friends and relatives are both more likely to donate to the campaign and to be granted a meeting. Even if donations per se had no effect, we might nonetheless observe a correlation between donations and access.

These threats to inference may be handled through statistical fixes of various sorts.[4] To summarize these technical results briefly, the ill effects of errors in variables (Greene 1997) may be corrected if researchers have redundant indicators of the constructs they seek to measure. The problem of spuriousness may be offset by one of two regression techniques. The *multivariate regression approach* introduces control variables, such that donations become uncorrelated with any remaining unmeasured determinants of the dependent variable. The *instrumental variables approach* makes use of variables that have the special property of being correlated with donations but uncorrelated with unobserved causes of the dependent variable. Instrumental variables remedy both spuriousness and measurement error. With the right statistical adjustments, experimentation is unnecessary.

Unfortunately, making the right statistical adjustments requires additional data and knowledge of the true model by which the data were generated. Much as one might like to use instrumental variables regression, finding valid instrumental variables is no mean feat: one must locate a variable that predicts donations but is known to be unrelated to omitted determinants of access. What this variable might be is far from obvious, nor would one necessarily know when one had found it. And much as one might like to build an ordinary least squares regression model that controls for all threats to unbiased inference, one must first know what the threats are and how to measure them. That will not be easy, particularly when what one knows is derived from prior nonexperimental research.

4. These techniques are discussed in any number of texts, such as Greene 1997.

Nonexperimental data analysis rests on stipulations about which statistical models are adequate. No clear stopping rules exist to tell researchers when their statistical manipulations have succeeded. Researchers analyze their data until they feel that the results are presentable. More to the point, no well-defined *procedure* exists by which causality is inferred from nonexperimental data. At best, researchers work to show that their conclusions are robust across a range of alternative models and hope that the true model was among those which they considered.

A quite different approach is to gather data in such a way as to satisfy the requirements for valid inference. The procedure for drawing valid causal inferences involves random assignment of observations to treatment and control groups.[5] To study the effects of campaign donations, one would randomly dole out contributions of varying size, then ask for meetings.[6] One might, for example, ask people who intend to make contributions totaling $50,000 to produce a list of 100 candidates that they might like to support, with the intention of contributing money to a random subset of the list. Randomization ensures that the those in the various experimental conditions differ only due to random chance, the properties of which are well understood in statistical theory. The larger the number of observations, the less likely it is that observed differences between treatment and control groups could be ascribed to chance correlation between the treatment condition and factors outside the researcher's control.

The beauty of experimentation is that one need not possess a complete theoretical model of the phenomenon under study. By assigning the level of contributions at random, one guarantees that the only factors that could give rise to a correlation between donations and other determinants of access occur by chance alone. The experimenter in effect interrupts a causal process of mutual adjustment and interaction among politicians and donors; this greatly simplifies the task of interpreting the experimentally induced correlation between donations and politicians' behavior. Rather than launch a complex multivariate analysis of the flow to and from donations and access, the researcher may obtain an unbiased assessment of the average treatment effect merely by cross-tabulating access by the size of contribution. Rudimentary data analysis replaces scores of regressions,

5. In some instances, random or near-random interventions occur naturally, as when individuals are chosen for jury duty or military service by lottery (Angrist 1988), required to comply with court-ordered desegregation on the basis of the first letter of their last names (Green and Cowden 1992), or assigned a ballot order in a near random sequence (Miller and Krosnick 1998). For a critique of this literature in economics, see Rosenzweig and Wolpin 2000.

6. Whether experiments such as this are ethical is a question to which we return below. To what extent does the research harm the subjects or the society? Of what benefit is the knowledge acquired from the experiment? How do the risks of this and other experiments compare with experiments that are deemed acceptable?

freeing the researcher from the scientific and moral hazards of data mining.

Granted, the design and interpretation of experiments benefit enormously from sound theoretical knowledge. To the extent that we begin the process with intuitions about the conditions under which money buys access, we may design experiments to isolate high-effect and low-effect circumstances. For example, it may be that donations to well-funded candidates buy little gratitude and hence access, whereas donations to candidates who are strapped for cash have greater effect. Theoretical reflection on the conditions under which a treatment is likely to be effective may prevent us from drawing inferences that do not generalize to other situations. The point remains, however, that the experimental method could discover these nuances merely by trial and error, as researchers make sense of why a sequence of experiments produces different results. By clarifying both what experimental results mean and what kinds of experiments are likely to yield the most fruitful parameter estimates, good theories provide a more efficient path to knowledge about the underlying causal process (see Rosenzweig and Wolpin 2000).

Experimentation resolves certain measurement problems, since the experimenter sets the level of contributions. In addition, experimental procedures can rescue the researcher from certain mishaps in the execution of the study. It sometimes happens that some experimental subjects do not receive the treatment (e.g., some letters containing campaign donations are returned due to incorrect addresses). The experimental data remain informative because the *intended* donation is correlated with the actual donation but uncorrelated with other determinants of access. For this reason, intended donations represent an ideal instrumental variable for purposes of regression analysis. Indeed, random assignment may be characterized as a procedure for generating the instrumental variables necessary for the unbiased estimation of causal parameters (Angrist, Imbens, and Rubin 1996).

Randomization, of course, does not rule out threats to inference. Threats to *internal validity* arise when the treatment does not exactly correspond to the construct that is envisioned as the independent variable. In the campaign donation example, the independent variable is the level of financial support from a person or group. If donors present their checks to the candidates in person, while nondonors have no contact with the candidates, the effects of the treatment will be an unknown combination of the effect of the money and the effect of the greeting. Issues of *external validity* concern the connection between the experimental results and the causal processes that operate in other times and places. Laboratory experiments that examine how undergraduate subjects behave in contrived situations may produce results that are deficient in this respect, depending on whether the causal process in question varies greatly across people and context.

The attractiveness of field experimentation stems from the fact that randomization is performed within a naturalistic setting. The field experiment described above, though fanciful, is very different from a corresponding laboratory study or simulation exercise. We are not asking subjects to envision themselves as elected officials deciding how to allocate their time. Nor are we proposing to study how actual officials answer requests for meetings that they know to be fictitious (cf. Chin, Bond, and Geva 2000). Field experiments involve real contributors and real candidates. While the result of this field experiment cannot tell us the influence of money on political access for all times and places, it speaks to the question of causality with far greater clarity and precision than alternative experimental or nonexperimental methods.

Field Experimentation In Practice

To this point, we have considered only a hypothetical case of field experimentation, constructed in order to illustrate the distinctiveness of this methodology. To appreciate the relative merits of experimental and nonexperimental research in practice, it is instructive to examine a research literature that features both types of investigation. One such literature concerns voter mobilization. As noted above, this line of research dates back to Gosnell's efforts (1927) to register voters and impel them to vote by mailing them leaflets. The experimental studies of Gosnell (1927) and Eldersveld (1956), as well as the lesser-known works of Adams and Smith (1980) and Miller, Bositis, and Baer (1981), have been overshadowed by a plethora of nonexperimental studies (Caldeira, Clausen, and Patterson 1990; Kramer 1970; Rosenstone and Hansen 1993; Wielhouwer and Lockerbie 1994; Huckfeldt and Sprague 1992). These studies use survey data to gauge whether citizens recall having been contacted by political campaigns. In a typical statistical analysis, voter turnout is regressed on reported campaign contact, controlling for background variables such as partisanship, interest in politics, and demographic characteristics.

This approach has certain inherent drawbacks: the researcher must take the respondent's word that reported contacts actually occurred. In addition, the vague way in which campaign contacts are defined and measured raises further concerns. Little attempt is made to distinguish between personal and phone contacts or between single and multiple contacts. Serious though these measurement concerns may be, they are in some sense eclipsed by even more serious doubts about spuriousness. If parties direct their appeals disproportionately to committed partisans and frequent voters, those most likely to vote will be most likely to receive political contact, and the apparent link between contact and turnout may be spurious.

As noted above, these problems may be corrected statistically. For example, one may use instrumental variables regression to offset biases associated with endogeneity and misreporting. None of the aforementioned studies uses instrumental variables regression. This gap in the literature is not hard to understand. Valid instrumental variables are difficult to think of, let alone find in existing surveys. To the wayward scholar in search of knowledge, the Oracle at Delphi commands: Before you may unlock the secret of mobilization's influence, you must find a survey measure that predicts campaign contact but is uncorrelated with unmodeled determinants of voting. Only very gifted scholars can hope to solve this inscrutable puzzle, and, even then, they will seldom do so to the satisfaction of their critics. The validity of any proposed instrumental variable ultimately rests on theoretical stipulations of one sort or another. Only randomization provides a *procedure* for generating instrumental variables that are valid on their face.

A more common but nonetheless problematic strategy for avoiding bias is to formulate an exhaustive model of voting, hoping that the control variables in the model eliminate any correlation between campaign contacts and the disturbance term. Unlike the instrumental variables approach, this methodological tack fails to address biases arising from mismeasurement. Given the seriousness of the measurement problem in most survey-based studies, this defect may altogether invalidate the results. Holding measurement concerns in abeyance, this statistical approach nonetheless requires a leap of faith. How can one tell whether a given set of control variables performs its appointed task? Sometimes scholars seem prepared to defend their models on theoretical grounds, but their choice of control variables often bears a more-than-coincidental resemblance to the measures that happened to be available in the survey they analyze. In the end, reasonable people may differ about whether a regression model is adequate to ward off bias. In the absence of such Delphic certitudes about the completeness of the model and the integrity of the available measures, the only way to determine the success of one's regression model is to conduct a randomized experiment to use as a benchmark.

In the case of voter mobilization, it would be very difficult to predict ex ante whether experimental and nonexperimental results would coincide. Random errors in survey measures of campaign contacts might lead to an underestimate of their effects; on the other hand, if voters are more likely than nonvoters to report campaign contact when none occurred, regression analysis might overestimate the effects of mobilization. If parties and candidates target likely voters, the effects will also be overestimated. Under these conditions, the sign and magnitude of the biases in nonexperimental research are knotty functions of variances and covariances of observed and unobserved determinants of the vote. Even if by some accident of fate the positive biases were to offset the negative, our ability to compare experimental and nonexperimental findings is hampered by the fact that most surveys fail to gauge the precise nature and frequency of campaign contact. It would be a stretch to call this exercise parameter estimation.

As it turns out, experimental studies have found that the effectiveness of mobilization varies markedly depending on whether voters are contacted by phone, mail, or face-to-face. Building on the earlier experiments of Miller et al. (1981) and Eldersveld (1956), Gerber and Green (2000b) conducted a randomized field experiment using a population of 30,000 registered voters in New Haven. These results indicate that personal canvassing boosts turnout by approximately 9 percentage points, whereas phone calls have no discernible effect. The intensity of the mobilization effort matters as well. Registered voters who were sent a series of mailings urging them to vote were more likely to go to the polls than those who were sent just one piece of mail. It could be said that these results corroborate earlier nonexperimental findings, in that both find statistically significant effects of campaign activities on voter turnout. Experimental studies, however, speak to the issue of causality with far greater clarity and nuance, because the researchers had control over the content of the campaign stimulus. It is fair to say that absent the various field experiments that have been conducted over the years, we simply would not know whether the nonexperimental results were trustworthy or artifactual.

These experiments, to be sure, have important limitations. Unlike survey-based research, which uses nationally representative samples, these field experiments involve just one region. And in contrast to many of the nonexperimental studies, which examine a variety of elections, the extant experimental literature examines certain municipal, primary, and midterm elections. Finally, the experimental intervention in this study was a nonpartisan get-out-the-vote drive, which leaves open the question of whether partisan mobilization efforts work differently.

Notice, however, that these limitations are different in kind from those associated with nonexperimental work. In the latter case, doubts about measurement and endogeneity leave us uncertain about whether correlation implies causation. Mobilization might cause voters to go to the polls, or its apparent influence might be artifactual. The limitations of the field experiments invite us to ask whether the demonstrated causation generalizes to other campaigns, times, and places. These limitations represent testable hypotheses about the conditions under which mobilization campaigns succeed. Thus, the solution to each of these concerns is to conduct additional experiments, gradually broadening the external validity of the experimental findings. Accordingly, Gerber, Green, and Green (2000) report the results of seven follow-up experiments examining the effects of partisan mobilization in off-year legislative and mayoral elections in various states, and follow-up studies involving millions of registered voters were conducted during the 2000 campaign. Through replication and extension, experimentation gradually resolves uncertainties about the scope of external validity. Eventually, this line of investigation will lead to a precise understanding of when and where mobilization works.

Accumulation of Data and Knowledge

Certain crucial field experiments have the capacity to break deadlocks that arise when nonexperimental analyses produce conflicting results. Consider, for instance, the large research literature on campaign spending and election outcomes. By political science standards, this literature addresses a particularly well defined question (How much does campaign spending influence the election outcome?) using relatively straightforward independent and dependent variables (money and votes). This literature spans decades of elections. It focuses not only on House and Senate elections but also on state and local races. Most of the literature examines U.S. legislative races, but some attention is devoted to referenda and elections outside the United States. One noteworthy feature of this literature is that replication seems to do little to advance the ongoing methodological debate over the proper way to estimate the influence of spending on election outcomes. For example, Gary Jacobson (1990) has shown that his ordinary least squares regression model produces more or less the same results regardless of the year or type of election studied. According to Jacobson, challenger spending has a profound effect on the election outcome, while incumbent spending has little influence. Green and Krasno (1990) argue that Jacobson's regressions are not to be trusted because spending is endogenous—incumbents spend more as they become more vulnerable to defeat. Green and Krasno estimate an instrumental variables model instead and find that in election after election both challenger spending and incumbent spending matter greatly. Replicability is not at issue here. Depending on one's perspective, the data either reveal the same lawlike parameters over time or the same debilitating biases. Every election cycle brings a harvest of new data, yet this literature remains at a standstill.

One way to proceed is to argue about the validity of various modeling assumptions. Green and Krasno's answer (1990) to the problem of endogeneous spending is to estimate an instrumental variables regression in which past spending by the incumbent is the instrument. Abramowitz (1991) criticizes this model on the grounds that past incumbent spending might be correlated with omitted determinants of the vote, arguing instead that vulnerability can be measured by expert judgments and controlled statistically. Erikson and Palfrey (1998) argue that total spending by both candidates corrects problems caused by unobserved heterogeneity. Levitt (1994) claims that the best way to deal with endogenous candidate selection and spending proclivities is to focus on repeat races in which the same challenger and incumbent square off in successive elections. Gerber (1998) contends that candidates' personal wealth provides a valid instrument with which to assess the effectiveness of their campaign spending. Theoretical pluralism is reflected in the diversity of empirical findings. Levitt (1994) finds that expenditures by incumbents and challengers have little effect; Abramowitz (1991) says that only challenger spending matters;

Gerber (1998) and Erikson and Palfrey (1998) find both to be influential. These discrepant findings reflect the fact that each researcher makes different assumptions when modeling the relationship between money and votes. One senses that scholars have reached the limits of what they can learn from more modeling exercises of this kind.

The alternative to econometric analysis of nonexperimental data is to estimate the marginal effectiveness of spending through experimental intervention. Find a campaign that is willing to allow its campaign tactics to be randomized, and examine the effectiveness of its tactics on a dollar-per-vote basis. This resulting efficiency estimate may be compared to the econometric estimates derived from nonexperimental data. For example, Gerber (2000) notes that the regression models endorsed by Jacobson (1985) put the rate of return at approximately one vote for every $300 that an incumbent spends. The estimates of Green and Krasno (1988, 1990) and Gerber (2000) are closer to $15. The experimental literature on the effects of personal canvassing suggests a figure of approximately $20 per vote. Furthermore, Gerber's own experiment gauging the cost-effectiveness of an actual direct mail campaign on behalf of incumbents in New Jersey legislative elections estimates a return of one additional vote for each $40 spent. Assuming that the marginal returns of direct mail are not atypical of returns to campaign spending in general (a testable proposition), these results strongly suggest that Jacobson's regression models underestimate the efficacy of incumbent expenditures, while the Green and Krasno results overestimate it. Thus, experimentation offers a way out of deadlock that occurs when nonexperimental analyses produce contradictory results.

The literatures on campaign spending and voting may seem unrelated because they involve very different sorts of data. Studies of campaign spending tend to focus on aggregate data, evaluating the effects of incumbent and challenger expenditures on vote outcomes. Studies of mobilization, on the other hand, use survey data to ascertain whether voters are more likely to go to the polls in the wake of a campaign solicitation. The two literatures seldom cite each other. Yet, as one reflects on the experimental logic underlying the two literatures, it becomes apparent that they are closely intertwined. In some sense, this illustrates the value of experimental results. Once a solid fact is generated through an experiment, one must then reevaluate every theory with which it is inconsistent.

■ | The Downstream Benefits of Experiments

For decades, political scientists have argued that, ceteris paribus, acquiring education raises the probability of voting. This causal proposition is confirmed by a wide variety of surveys, most of which are cross-sectional in design. At a given point in time, well-educated people are more likely to vote than those with less education. As with any nonexperimental result, this empirical regularity could be spurious. It could be that unobserved factors that cause one to vote are also correlated with educational attainment. Education could, for example, simply be a marker of one's socioeconomic status, and critics of this proposition have noted that rising levels of education have failed to produce higher levels of voter participation. In an effort to guard against spuriousness, scholars who seek to gauge the effects of education sometimes build elaborate and exhaustive models of voter participation. Have these models expunged all threats to valid inference? Absent some kind of randomized intervention, it would be impossible to know.

An alternative approach would be to examine the consequences of experimental interventions that randomly alter educational attainment. This is not as outlandish as it might at first seem. Scholars outside political science regularly investigate the determinants of educational attainment. For them, educational attainment is a dependent variable of great importance, and from time to time, they launch randomized experiments to see whether they can lower dropout rates, improve college performance, and the like. For political scientists, educational attainment is an independent variable. Any experiment that alters educational attainment provides the wherewithal to assess its effects on voting behavior. Our shared interest in educational attainment creates an opportunity for significant cross-disciplinary collaboration. Returning to these studies, political scientists could graft new information about the subjects' voting behavior and examine whether the treatment and control groups vote at different rates.[7] One such experiment would arguably contribute more to the understanding of education's effects than scores of survey-based regressions.

The larger implications of the previous example bear emphasis. Political scientists would be well advised to take stock of experimentation in other disciplines. Unlike most calls for interdisciplinary collaboration, this one does not rest on the conviction that social sciences should become more theoretically or methodologically ecumenical. The intellectual forces that impel sociologists, psychologists, and economists to conduct randomized experiments are in some sense irrelevant. Any randomized intervention that genuinely produces different outcomes in treatment and control conditions is potentially useful to political scientists. Fortunately for us, other social science disciplines are less benighted about the benefits of randomized field experiments.

This argument also suggests a very different way of looking at the value

7. The estimator used to analyze these data would be an instrumental variables regression of vote on education, using random assignment as the instrument. Note that the key assumption of this model is that nothing about the experimental treatment per se directly affects turnout. This assumption would be violated if the education program in question explicitly advocated civic participation.

of laboratory experiments within political science. Studies that examine the influence of television news on political opinions are frequently criticized on the grounds that the subjects are not watching television in a "naturalistic" setting. (The experimental subjects might not ordinarily watch televised news; in the lab, they watch television news in groups and may have some sense that the experimenter expects them to attend to the content of the shows.) These criticisms lose much of their force when political attitudes, not media viewing, become the independent variable. Consider, for example, the frequently advanced proposition that favorable assessments of the country's economic performance increase the likelihood of voting for and identifying with the incumbent president's party (Kiewiet 1983; MacKuen, Erikson, and Stimson 1989). When a randomized media intervention alters subjects' economic optimism, do their vote intentions and party attachments change also (Cowden and McDermott 2000)? Although some questions remain about the representativeness of the sample of subjects used in laboratory experiments, randomization provides a powerful instrument for gauging the effects of opinion change. *The dependent variables in lab experiments make for highly informative independent variables in subsequent research.*

This way of thinking about experiments greatly expands the range of feasible experiments, because it also corrects for the fact that certain subjects may be unwilling or unable to receive the treatment. The central principle is that if one can intervene in such a way as to alter (randomly) the independent variable, one can study its effects on some dependent variable. Thus, for example, if one were interested in the effects of face-to-face canvassing on voter turnout, one could randomly determine whether a canvasser *attempted* to contact a possible voter. Only some proportion of these attempted contacts would succeed in producing an actual contact.[8] Attempted contact serves as the instrument in an instrumental variables regression of turnout on actual contact. Notice that this estimator provides unbiased estimates even when subjects in the treatment group decide whether to allow a contact to occur.

Having dispelled the commonly held view that experiments are feasible only insofar as the independent variable is under complete control of the experimenter, we briefly describe some potential experiments that might be inspired by recent nonexperimental work.

1. *Ballot structure and passage of school bond measures.* Glaser (2002) argues, in part on the basis of survey experiments, that the passage of ballot measures designed to raise money for schools is facilitated by breaking up an omnibus bond request into a series of smaller bond measures. Rather

8. One can then determine the effects of actual contact on voter turnout by comparing turnout rates among those canvassers attempted to contact (T) and those canvassers did not attempt to contact (C), and dividing by the proportion (p) of attempts that led to actual contact: $(T - C)/p$ (see Gerber and Green 2000b).

than ask voters to support a bond package to fund air conditioning, science facilities, and a new gymnasium, present voters with a series of bond measures, each of which funds a different component. A field experiment to test this hypothesis begins with a sample of school districts planning to present a bond measure to voters. This group is then divided randomly into treatment and control groups. School boards in the treatment group are informed of the potential benefits of structuring their bond request to voters in a disaggregated fashion. So long as some positive proportion of the treatment group acts on this information, an experimental assessment is possible.

2. *Deliberation and opinion change.* Fishkin (1995) argues that extended political deliberation over issues causes participants to become more circumspect and knowledgeable, which in turn changes their policy opinions. An experimental design would be to divide those who show up for a deliberative session into groups that discuss different issues. Another possibility, which makes use of the downstream randomization idea, is to select a list of potential invitees and divide it randomly into treatment and control groups. The treatment group receives invitations, while the control group does not. Only some of those who were invited will actually show up. After the deliberative session, survey both the treatment and control groups (disregarding whether someone actually showed up). The differences between the two groups, when divided by the rate at which people acted on the invitations, indicates the effect of deliberation.

3. *Interpersonal influence.* Huckfeldt and Sprague (1995) contend on the basis of survey evidence that friends influence each other's opinions. This proposition could be tested experimentally by drawing a sample of dyadic friendships. In the treatment condition, one member of each dyad is contacted by a political campaign urging them to vote in a particular way. No attempted contacts occur in the control group. After this intervention, the member of each treatment dyad is interviewed, as are members of the control group. By comparing the opinions of those who were contacted directly, those whose friends were contacted, and those who received neither direct nor indirect contact, we can ascertain the extent to which campaign appeals are transmitted through these personal networks.

4. *The consequences of civic engagement.* Putnam (2000) argues that participation in civic organizations increases the trust one places in other people and in political institutions more generally. This proposition could be tested by means of a field experiment in which civic groups (e.g., parent-teacher associations, the League of Women Voters, Rotary clubs) randomize the effort that they devote to recruiting certain prospective members. Only some of those who are recruited will become active participants, but the effects of participation may be gauged nonetheless by surveying the control group and prospective recruits. If Putnam's thesis is correct, after some period of time the treatment group should show higher levels of trust and political involvement.

The unifying idea behind each of these experiments is the notion that randomized interventions create ripple effects. So long as political scientists have at their disposal interventions that really work, such experiments are possible.[9] Thus, for example, one year after randomly increasing voter turnout through a nonpartisan mobilization campaign, Gerber, Green, and Shachar (2000) found that the treatment group continued to vote at higher rates than the control group, suggesting that the act of voting per se increases the likelihood of subsequent participation. Because experimentation is rare, downstream experimentation is very rare. But as the experimental tradition is rekindled, opportunities for downstream analysis will grow.

The Case against Experimentation

Experiments are not without their limitations. Any fair assessment of their strengths must also take into account the many ways that experiments are thought to fall short as a research methodology in political science. Here we summarize the leading criticisms and offer our rejoinders to them.

Interesting Manipulations Lie beyond Our Reach

The drawback that most readily comes to mind is the inability to manipulate key political variables of interest. It is difficult to imagine how one could randomly assign presidential and parliamentary regimes for the purpose of evaluating their relative strengths and weaknesses. Surely, world leaders cannot be persuaded to allow political scientists to randomize their foreign policies, systems of patronage, or prospects for retaining power. The really big social science variables—culture, economic development, ethnic heterogeneity—probably could not be manipulated even if political scientists were permitted to try. For this reason, it is commonly thought that political science can never hope to become an experimental science. And that is where the discussion of experimentation typically ends.

Before dismissing experimentation as impractical or trivial, one should bear in mind that our discipline lacks a clear sense of what kinds of experiments are feasible. Even in the field of political behavior, field experiments are so seldom contemplated, let alone proposed and executed, that the discipline has little track record with randomized interventions in naturalistic settings. The authors' own experience in this regard may be instructive. In

9. If, due to sampling error or threats to internal validity, the experimental intervention only seems to work but does not actually work, the downstream inferences will be misleading.

1998, we conducted a large-scale mobilization experiment using nonpartisan get-out-the-vote appeals. After examining the results, we wondered whether partisan appeals also stimulate voter turnout. On a whim, we contacted a political consulting firm and asked whether they would be willing to randomize a small portion of their mailing lists. Rather than send each household on their mailing list four mailers, the campaign would randomly divide its list so that some households would receive nine pieces of mail, others would receive four pieces, and a small group would receive none at all. On its face, this sounds like an unsuitable proposal. Why would anyone allow political scientists to meddle in this way? The answer is that this consulting firm was curious about how their mailings affected the election outcome, particularly the campaign mail that was negative in tone. Neither campaign managers nor political scientists have the slightest idea whether the most efficient use of their budget is to send four mailers to fellow partisans, nine mailers to a small set of ardent partisan supporters, or two mailers to everyone.

Political scientists have something to offer those who command discretionary resources and take an interest in cause and effect: expertise in designing randomized interventions and analyzing the statistical results. In general, we would argue that, at a minimum, *experimentation is possible whenever decision makers face constrained resources and are indifferent between competing ways of allocating them.* In such cases, decision makers may as well randomize their allocations, since they could then derive some useful information about the consequences of their actions.

Alternatively, they can roll out their intended intervention in stages or regions, again on a random schedule. Suppose that legislators wish to gauge the influence of a new policy, such as the elimination of a national maximum speed limit. One way to craft an experiment would be to phase in the implementation of the new speed limits, starting with randomly selected stretches of roadway. These treatment roads would otherwise be identical to the roads in the control group, and one could examine the consequence of short-term changes in speed limits on fatality rates, congestion, and other dependent variables. Again, randomized phase-ins can be justified on the grounds that administrative capacity does not allow for the simultaneous introduction of new policies everywhere.

The same principle applies to situations in which political actors work their way through lists. Suppose a political campaign seeks to canvass by phone 10,000 registered voters prior to an election. Given constraints of time and money, they do not know ahead of time whether they will be able to call all the numbers on the list. If the list is sorted in random order ahead of time, then any numbers that are not called constitute a randomized control group. If the campaign makes its way through the entire list, no experimental analysis is possible. But if some names are left over, the data become quite valuable. Indeed, even experiments in which only a small number of observations are assigned to the treatment or control group may produce more knowledge than a conventional large-*n* observational study.[10]

Experiment-minded political scientists, however, must be on hand to propose randomization to the powerful but indifferent. Otherwise, these actors will experiment in the colloquial sense of the term: they will try different courses of action in an effort to guess which one works best. Absent randomization, neither they nor we will be able to draw reliable inferences. So long as political scientists remain oblivious to field experimentation, opportunities for conducting this type of research will be missed.

Sometimes we can do better than merely make decision makers indifferent to our experimental intrusions. Political scientists from time to time acquire knowledge that decision makers would find useful. A good example is Glaser's findings (2002) concerning the disaggregation of school bond referenda. According to Glaser, public officials in Jackson, Mississippi, may have discovered a way to increase the chances of obtaining funds for schools. School boards desperate to win passage of bond issues may welcome the suggestion that they imitate the successful strategy deployed elsewhere. Outside of political science, field experiments routinely grow out of hypotheses about how to reduce domestic violence (Sherman and Berk 1984), increase tax compliance (Perng 1985), or decrease school dropout rates (Crain, Heebner, and Si 1992).

Such experiments are small by political science standards, but they can be used to bolster the investigation of large questions. Big theories hinge on a subsidiary propositions, some of which may be susceptible to experimentation. The microlevel phenomena (e.g., shirking, obedience, conformity) that are thought to undergird aggregate phenomena can be examined experimentally. Big theories may also be tested using institutions or collectivities of different scope and size. Theories that are crafted with an eye toward large units of analysis (e.g., the effects of committee jurisdictions in national legislatures) may be tested using analogous smaller units (e.g., local legislative bodies).[11] To the extent that political science strives to produce theories of general applicability, it makes sense to disaggregate theories and test their constituent parts.

In political science, there is a tendency to dismiss this line of reasoning as reductionist or naive. Our profession encourages students to go after big problems directly, rather than work up to them slowly via bench science.

10. This point may be demonstrated more rigorously by way of a mean-squared error comparison between two estimators, one based on a randomized experiment of size n_1 and the other a potentially spurious regression based on a nonexperimental sample of size n_2. As intuition would suggest, the more serious the expected biases of the latter, the greater the advantages of randomized experimentation.

11. To date, this line of argument has been used to justify decades of laboratory research on legislative behavior. This logic becomes all the more compelling as research moves from the lab into actual legislative arenas.

In fairness, both sides of this question base their methodological strategy on hunches about which type of science works best. It could be that the most fruitful way to learn about the causes of revolution, democracy, or economic development is to study them historically or comparatively. Such studies might produce a dramatic surge in our understanding of these phenomena. Another possibility is that the causes of these phenomena are too complex and intertwined to permit secure causal inference. More reliable inferences may be built up from studies of more basic aspects of these broad-gauged phenomena.

To draw an analogy to biological science, we observe that some leukemia researchers try to ascertain its causes by studying epidemiologic data, linking incidence of this disease to various environmental or genetic risk factors, while others focus their attention on cell biology. The former provides descriptive information about the scope of the problem and hints about its causes. The latter leads down many blind alleys but ultimately leads to an accumulation of secure knowledge about what leukemia is and how it comes about. Both, of course, are important and complementary endeavors, and neither has solved the mysteries of leukemia. But arguably bench science has contributed more to its understanding, as well as to the understanding of cognate phenomena. By investing so much of its resources in broad-gauge research, political science may be missing out on the rewards of basic science.

Experiments Are Expensive and Difficult to Execute

Harold Gosnell's voting experiments (1927) required a complete census of several neighborhoods, the creation and dissemination of direct mail, and the careful examination of voter turnout records. This sort of primary research goes beyond the experience of the typical political scientist. And if one were to contemplate the kind of randomized television campaigns that would test propositions about negative television advertising and opinion change or randomized campaign contributions and access to the legislative agenda, the costs would be astronomical. The benefits of such studies, meanwhile, must be discounted by the probability that something goes awry in the execution of the experimental manipulations. By this view, a cost-benefit analysis of nonexperimental and experimental research comes down in favor of the former.

It is difficult to say in advance what the payoff from field experimentation will be, but we strongly suspect it to be greater than for nonexperimental research, particularly in fields such as political behavior where large-scale randomized interventions are feasible. On the cost side, the overwhelming preponderance of funds flowing to political science supports nonexperimental research or lab experimentation. The situation in our discipline is in no way analogous to that in economics, where large-scale field experiments have in recent decades consumed immense sums of research money. Thus, even if one thought field experimentation promising only in certain subfields of the discipline, one would be forced to conclude that some reallocation of resources in its direction would be desirable.

Beyond this, we hasten to point out that field experiments may not be as expensive as one might think. The very forces that make possible large-*n* studies of voting behavior—computers, specialized databases, and campaign professionals—also make large scale experimentation easier now than in Gosnell's day. In the weeks leading up to the 1998 election, we conducted randomized experiments in which the people we hired distributed thousands of leaflets, sent tens of thousands of pieces of direct mail, made tens of thousands of phone calls, and visited thousands of addresses. These projects, which spanned three neighboring cities, cost a total of $50,000. In 1999, we worked with political campaigns that used randomized quantities of direct mail advocating the election of certain candidates. In order to gauge the effects of these mailers on vote preference, postelection interviews were conducted with members of the treatment and control groups in two different jurisdictions, again at a cost of $50,000. In 2000, we randomized the mail, phone, and canvassing campaigns of organizations working in more than a dozen states. The cost of these mobilization campaigns was largely borne by these organizations, which in turn hired us to conduct randomized experiments evaluating their efforts. The sums of money involved are not trivial, but taken together, they are less than a typical budget request for political surveys funded by the National Science Foundation.

External Validity

The fact that field experiments are often confined to local areas raises important concerns about external validity. The experiments often purport to examine the consequences of a policy intervention (e.g., a nationwide campaign of mailings urging people to vote), but there may be many reasons why the results of a given experiment fail to generalize across time and place. Appeals to civic duty may seem hollow when elections are uncompetitive or in places where that theme has become hackneyed from overuse. Indeed, the fact that direct mail mobilizes voters once is no assurance that it will do so when repeated over a series of elections. Finally, one must be attentive to issues of scale. If a nationwide mobilization campaign were conducted, the campaign itself might draw attention (either positive or negative) that might augment or undermine the effectiveness of the intervention.[12]

12. Note that some of the nuisances that threaten the assessment of average treatment effects may represent research topics in themselves. For example, if those who are mobilized by the experimenter in turn mobilize others in the control group, these spillover effects will lead to biased inference. On the other hand, the relative size and proximity of those in the treatment and control groups may be varied, in an effort to detect the role of interpersonal influence.

These concerns are properly viewed as empirical questions. As is true for all empirical research projects, experimental studies raise questions that lead to new projects and modifications of the experimental procedures. Rather than assume the portability of experimental findings, researchers should replicate their studies in a variety of settings, devoting special attention to the consequences of repeating the same intervention over time. Before endorsing nationwide implementation, researchers should study interventions of varying magnitudes. The accumulation of knowledge from experiments proceeds by increments. In some sense, impatience for answers and reservations about extrapolation constitute arguments on behalf of experimentation far and wide.

EXPERIMENTAL RESULTS ARE DISJOINTED

There is no guarantee that experiments will generate solid facts. Experiments may generate contradictory results, either due to sampling error or to idiosyncracies in their design (Heckman and Smith 1995). As one looks to the field of social psychology, which has invested heavily in laboratory-based experimental research, one may be concerned about creating a tower of babel in which secure generalizations are reputed to be in short supply. Whether experiments will produce a knot of contradictions in political science is an empirical question that cannot be resolved until more of them are performed. In the meantime, it should be remembered that the reputation of social psychology is shaped in large part by the fact that in recent decades its experiments have been confined to the lab.

At the very least, field experiments offer the prospect of cumulative knowledge about the specific circumstances under investigation. For example, Cover and Brumberg (1982) use a randomized field experiment to demonstrate that baby books sent by members of Congress to new parents improved the standing of these representatives in the eyes of the recipients. Whether this finding speaks solely to the effectiveness of baby books or generalizes to franked mail or to all communications from elected officials, the point remains that this finding shows something important about a particular political tactic. By the same token, Miller, Krosnick, and Lowe's randomized study (2000) of direct mail appeals for campaign contributions shows that one raises the most money by mail that strikes an alarmist tone, claiming that the respondent's values are in jeopardy. Whether this finding generalizes to other instances of political fund-raising or political mobilization remains to be seen, but again it provides useful practical knowledge.

INTERNAL VALIDITY

Every experiment raises questions about the true causal agent at work. Indeed, an infinite supply of challenges might be directed at each experiment. Consider an experiment in which some voters are encouraged to vote through a face-to-face conversation while others are contacted by phone. Even if the scripts that are read to subjects are identical, they could be read in different ways, at different times of day, by different canvassers. Rather than chase down every possibility empirically, theory-based decisions must be made to winnow the range of possible hypotheses to a smaller number of plausible ones. Experimental inference is never entirely inductive, and therefore the contrast between experimental and nonexperimental inference is one of degree rather than kind.

This is a fair criticism. Generalizing from any particular experiment to some broader causal law requires an inherent theoretical leap. But like concerns about external validity, this concern suggests the need for more experimentation. The more ways in which a given experimental result is reproduced, the more implausible internal validity critiques become.

Much the same may be said about problems stemming from publication bias. To the extent that researchers report or journals publish only those results that achieve statistical significance, the research literature will become distorted by unrepresentative findings. Consistent with suspicions of publication bias, Gerber, Green, and Nickerson (2001) find that the larger the sample size, the smaller the reported effect of canvassing on voter turnout, presumably because smaller studies had to report larger effects in order to reach statistical significance. Certainly, when one reads that randomized prayer by outsiders on behalf of (unknowing) coronary patients improved their medical outcomes (Harris et al. 1999), one would like to see the result replicated before generalizations are advanced about the curative powers of prayer. The hypothesis of publication bias would suggest that the replication will find smaller effects than the published study.

Black-Box Causality

One common complaint about experimental research is that it often fails to generate a clear sense of why the intervention produced its effects. Canvassing leads to higher turnout, but why? Is it because people would otherwise forget to vote? Does it pique their interest in the election outcome? Does it evoke a sense of civic obligation? Or something else? None of the published experimental work on voter mobilization addresses these questions.

Although existing research tends to be deficient in this respect, experimentation need not involve black-box causality. Having demonstrated a causal connection between an intervention and an outcome, the researcher may take one of two approaches to figuring out why the effect occurs. The first approach is to vary the stimulus so as to isolate particular mechanisms. For example, if canvassing works because voters would otherwise forget Election Day, phone reminders should produce similar effects to messages delivered face-to-face. The audit experiments designed to measure the conditions under which employers, realtors, and commercial operations discriminate on the basis of race frequently take this approach (Yinger 1995).

An alternative approach is to measure the variables that are thought to mediate the relationship between the intervention and the dependent variable. For example, suppose it were thought that face-to-face canvassing increased voter turnout by fostering an interest in politics among those canvassed. For this proposition to be true, it must be the case that subjects in the treatment group show higher levels of political interest after the canvassing. One way to detect this change is to conduct a survey of those in the treatment and control conditions, examining whether the two groups differ with respect to political interest. If political interest does not differ across treatment and control conditions, it cannot be regarded as a mediating variable that explains why canvassing works. Although one would not invest in this kind of research until one were reasonably sure that interventions such as canvassing really work, in principle nothing prevents experimental researchers from investigating causal mechanisms.[13]

The Scope of Experimental Political Science

Randomized field studies threaten to narrow the range of research questions posed in the discipline. Research will be tailored to coincide with opportunities for experimental intervention, and the discipline will lose sight of the big questions that make the discipline vibrant and interesting. Hazardous though it may be to draw causal inferences from nonexperimental data, we typically have no alternative when studying international affairs, the presidency, or the federal budgetary process. To denigrate nonexperimental political science is tantamount to suggesting that the discipline move away from these important objects of inquiry.

In some sense, this criticism posits a trade-off between the significance of the research question and the reliability with which it can be answered. This distinction is misleading insofar as the discipline has relied almost entirely on nonexperimental research to answer all of its questions, both big and small. Field experiments are rare even in circumstances where they are feasible and well suited to the questions at hand.

Even if we grant the existence of this trade-off, it is far from clear how the discipline's effort ought to be allocated. Reflecting on several decades of scholarship, one might ask what we really *know* as a discipline? What sorts of causal propositions have been established with a high degree of

13. From a statistical standpoint, the investigation of causal mechanisms raises an identification problem when the number of randomized interventions is smaller than the number of potential intervening variables. It is useful, therefore, for an experiment to include a range of different treatments.

certainty? We have a great many useful descriptive generalizations, but what do we know about cause and effect? As we reflect on the triumphs that are likely to be reported in other chapters of this volume, it occurs to us that each of them is potentially subject to withering methodological critique. Democracy contributes to international peace. Globalization undermines the welfare state. Closed rules in legislative chambers encourage committee members to gather costly information. Each of these propositions may be true, but it is far from clear that nonexperimental analysis has powerfully shaped our prior intuitions about them.

To put it somewhat differently, suppose you strongly believe one of these propositions to be true. One day, you are presented with a regression based on nonexperimental evidence or the results of a laboratory experiment showing this proposition to be false. Naturally, your first inclination is to find fault with the regression model or the verisimilitude between the laboratory and the real world. Indeed, even a series of such studies might well leave you unpersuaded, since methodological uncertainties cloud them all. By contrast, it would be difficult to maintain a prior view in the face of disconfirmation by a succession of field experiments.

As we noted earlier, political science may be well advised to start small. Using reliable scientific methods such as randomized intervention, first answer basic research questions. Build gradually to larger propositions, as experimental evidence accumulates. We recognize that this vision of progress may be unappealing to a discipline that has long put its faith in two Stachanovite models of science. The first emphasizes the great leaps forward that come as the result of theory-driven breakthroughs. Each generation of political scientists holds out the hope that the right way of looking at politics will lay bare its inner workings. The second emphasizes the importance of asking big questions and throws researchers into the breach armed with nonexperimental data. It requires a combination of genius and good fortune to extract secure causal inferences from theoretical deduction or nonexperimental data. Few rules and procedures guide the scholar through these modes of scientific exploration. Experimentation, on the other hand, is a form of science bounded by procedures that greatly simplify the task of drawing inferences from statistical relationships. Experimentation benefits from genius and good fortune but does not require them.

ETHICS

The previous concerns focused on the capacity of experiments to produce secure knowledge about the political world. A quite different kind of concern surrounds the ethics of field experimentation. Because experimentation involves intervention rather than passive observation, experimental researchers bear an ethical burden that nonexperimental researchers do not. Even if we accept that experimental analysis is a superior method for securing valid causal inferences, there remains the question of whether researchers are justified in altering the world in ways that may adversely affect those connected with the experiment.

This question cannot be answered in the abstract or by analogy to medical ethics, since political science experiments vary widely in terms of the risks that they present and the value of the knowledge that they might generate. Unlike medical experimentation, where a consensus of opinion understands risks in terms of health, political science experiments may involve contested values. Suppose one were considering an intervention thought to raise voter turnout. Do higher levels of voter turnout lead to more desirable political outcomes? Does going to the polls make voters themselves better off?[14]

A related concern focuses on the moral culpability of the researcher. Intuitively, we sense that it is acceptable to raise voter turnout but unacceptable to lower it. But suppose one honestly believes that desirable political outcomes follow from *lower* voter turnout, can one ethically perform experiments designed to reduce voter participation? And what about circumstances where one genuinely has no prior sense of whether an intervention is likely to be harmful or beneficial? The same question arises from ambivalent prior beliefs. Can one justify an experiment on the grounds that decision makers seem to be of two minds about the benefits of a particular course of action?

Like visitors to nature preserves, social scientists who seek to conduct randomized experiments in real-world settings should aspire to leave no footprints. Sometimes there is no practical alternative to intrusive intervention, but where possible, they should employ what we would term "minimally intrusive randomized designs." The hallmark of these designs is that randomization is conducted in ways that have no effect on social outcomes. As suggested earlier, such designs emerge naturally in situations where the scope of an intervention confronts resource constraints. Suppose an environmental organization seeks to recruit members from a list of local residents but lacks the resources to call everyone on the list. If the ordering of the names on the list is randomized before they work their way through it sequentially, the names they do not get to constitute a random control group. Similar designs may be devised for the randomization of waiting lists or the sequence in which a policy change is implemented. Experimental protocols still dictate which individuals or political units receive the treatment without affecting the overall flow of resources or risks.

Perhaps because the values associated with the objects of study are often controversial, ethical questions concerning experimentation with human subjects are typically reduced to procedural questions. For example,

14. Paradoxically, what harms are likely to befall subjects is itself an empirical question, possibly one that can only be answered satisfactorily through experimentation.

human subject committees generally require that investigators obtain consent waivers from participants in experiments. But unlike laboratory studies, field experiments in political science do not lend themselves to these kinds of contractual arrangements. As a practical matter, it is often impossible to obtain informed consent without irreparably damaging the study itself. It would be hard to envision a real-world voter mobilization experiment that first asked subjects to agree to participate in a study. The problem of informed consent grows even more unwieldy as one considers the possibility that an experiment could sway an election, an outcome that impinges on people outside the treatment and control groups.

One of the many complexities of characterizing ethical problems in procedural terms is that the interventions in some political science experiments (e.g., political expression and voter mobilization) may involve constitutionally protected political activity.[15] As citizens, political scientists have the right to engage in this kind of conduct, regardless of their motives. Yet there remains a lurking sense that scientific motivations render this kind of conduct inauthentic and suspect. Suppose, for example, that a political scientist were to run for public office. Suppose that she made ten or twenty such speeches at various churches in attempt to curry public support. We ordinarily would not question the rectitude of this behavior even if the churches were chosen at random. Now suppose that the political scientist were interested not simply in winning the election but also in figuring out how much these speeches affected the election outcome. Should we now regard this behavior in a fundamentally different moral light? Does it make sense to draw an ethical distinction between sincere interventions and experimental interventions?

Rather than focus on motives, one might contend that the standard for ethical responsibility comes down to the social cost and benefits of the research. Here, the discipline should reflect not only the ethical ramifications of experimentation but also the implications of continuing the current practice of attempting to acquire knowledge through nonexperimental observation. We earlier discussed the criticism that field experiments are incapable of tackling the big questions that most interest political scientists. In some sense, this limitation arises because, as Campbell (1969) laments, neither academics nor policymakers have embraced the ethic of an experimenting society in which proposed innovations are introduced and evaluated systematically. By resisting the idea that political institutions should be changed so that we can study their consequences, we preserve a conception of an organic polity untrammeled by science, but we neither reap the benefits of this potential knowledge nor suffer the social costs of obtaining it. In the meantime, the political world changes in ways

15. Another wrinkle is that, in some cases, political scientists themselves do not administer the intervention but rather suggest to other groups how they may structure their activities in ways that make for experimental evaluation.

that are neither informed by rigorous scientific inquiry nor readily interpretable through post hoc investigation.

■ | Conclusion

Political science is bristling with intriguing claims about cause and effect. Participation in church and civic organizations increases one's trust in other people (Putnam 2000). The more education one has, the more likely one is to vote (Wolfinger and Rosenstone 1980). As individuals' finances improve, they become more likely to vote for the incumbent president (Tufte 1978). Propositions such as these lie at the heart of some of the largest research literatures in the discipline. Yet, so far as we know, none of these well-known and extensively researched claims has been tested using field experimentation, and much the same could be said for the bulk of received wisdom in political science. Knowledge about the determinants of political behavior derives almost entirely from nonexperimental investigation.

Nonexperimental research is not valueless. The distinction between experimental and nonexperimental evidence should be viewed as a continuum. At one extreme lie observations that make for clear and direct causal interpretation. Experiments fall in this class when randomization is executed faithfully and precautions are taken to prevent threats to internal validity. Nearby are certain nonexperimental investigations that focus on the consequences of near-random variation, such as unexpected events or policies that divide groups arbitrarily. Toward the center of the spectrum are studies that make use of nonexperimental data but rely on robust statistical analysis. As we move toward the other end of the spectrum, the causal systems become more complex and statistical correctives less robust. Here the state of the discipline resembles monocrop agriculture, efficiently generating prodigious quantities of nonexperimental results but deeply vulnerable to an experimental intrusion that could consume the stock of received wisdom.

By arguing on behalf of field experimentation, we are recommending a fundamental change in the way that political scientists look at research. Not enough attention is paid to the prospects for randomized interventions or to the natural occurrence of near-random variation. At a minimum, political scientists should consider what kind of experiments would *in principle* test the causal propositions they advance. Even in those instances where such experiments are altogether infeasible, this exercise can prove extremely useful, as it clarifies the empirical claims while illustrating how the underlying concepts might be operationalized. Whether or not they are aware of it, social scientists rely on the logic of experimentation even when analyzing nonexperimental data.

Experimentation may also lead political scientists to rethink the way in which they interact with the outside world. Political scientists seek to devise and deploy procedures that lead to valid causal inference; political activists seek to bring about social change; interest groups and those who administer the state seek to understand how policies and programs operate. Each group has an interest in acquiring knowledge (although not necessarily in sharing it). Through collaboration, these groups in some sense strengthen each other. If scholars can demonstrate the practical benefits of science, those who have the discretion and resources to effect change will learn to seize opportunities to acquire knowledge.

The value of experimentation to political science as a profession should not be underestimated. As one reflects on the history of medicine during the last century, it is clear that the status and credibility of the profession soared once doctors began to prevent and cure disease. The demonstrable practical benefits of medicine made it reasonable to instruct an ill person to call a doctor. "Call a political scientist," on the other hand, sounds more like a gag line. Seldom do people in politics look to political scientists for practical advice, and to the extent that we are consulted, political actors have a nagging sense that our advice is scarcely better than their own intuitions. This is not an unreasonable suspicion, because we have seldom deigned to analyze the narrow questions that concern those engaged in day-to-day politics, nor have we attacked these or other problems with sufficient rigor to make the outside world sit up and take notice. Experimentation is not only a secure path to knowledge, it is a secure path to the kind of knowledge that is in great demand.