Back to the Future:
Endogenous Institutions and Comparative Politics

Submitted for publication in:
Mark Lichbach and Alan Zuckerman, eds.,
Comparative Politics: Rationality, Culture, and Structure
Second Edition
Cambridge University Press

November 9, 2007

Jonathan Rodden
Associate Professor
Department of Political Science
Stanford University
616 Serra Street
Encina Hall Central, Room 444
Stanford, CA 94305-6044
jrodden@stanford.edu

Like a fashion trend traveling from New York to the heartland, soul-searching about causality has made its way from empirical research in economics to that in political science with the usual lag. Gone are the days when it is enough to have a nice theory, a conditional correlation, and some rhetoric about the implausibility of competing explanations while implying but assiduously avoiding the "c" word. Editors, reviewers, and search committees are beginning to look for more explicit and careful empirical treatments of causality. That is, following a definition of causality tracing back to Mill (1848), researchers are expected to lay out a set of possible outcomes, or counterfactuals, generated by a set of determinants, and demonstrate that holding all possible determinants except one at a constant level, the manipulation of that determinants is associated with a specific change in outcome, which can be deemed a causal effect (Heckman 2005).

It does not take much soul-searching to realize, however, that the observational studies that make up the vast majority of empirical explorations in comparative politics are deeply flawed when held up to the experimental ideal. Where does this leave us? In the extreme view, if we cannot do randomized field experiments or perhaps survey experiments, we should do nothing. The opposite extreme position holds that this would remake comparative politics into an arid sub-field of program evaluation, turning a blind eye to the interesting and important questions that animated the field in its golden era (definitions vary). In this view, there is a steep trade-off between "interesting" and "provable," and the best way forward is to maintain a more casual approach to empirical explorations of causality given that the basic problems are intractable for the questions at the heart of the sub-field. As one colleague put it, it is better for the field to suffer from omitted variables than omitted questions.

Rather than entering an abstract debate about the philosophy of science, this essay opts for an admittedly rather offhand descriptive empiricism. How are empirical researchers in comparative politics grappling with problems of causality in their work? To put it bluntly: is comparative politics getting less interesting? Does our newfound concern with causality turn us away from Katzenstein's (1995) "important and interesting questions" (p. 11) or turn us into Eckstein's (1964) "dullards"? (See Lichbach, this volume.) If we ignore all of the data generated by history save for the rare instances when it bestows natural experiments, or the instances when human subjects review boards allow for truly randomized treatment, how much have we given up?

Moreover, are we making any progress? Or are we merely tilting at windmills, applying new techniques but continuing to fall short in our efforts to discern correlation from causality? Even worse, perhaps we are fooling ourselves by trading some dubious but straightforward assumptions about unit homogeneity and conditional mean independence for equally dubious but less transparent beliefs about instruments or propensity scores.

This essay addresses these questions by focusing on endogenous democratic political institutions. One of the achievements of comparative politics over the last 50 years is that it has discovered some impressive associations between institutions and various political and economic outcomes, and many of these are consistent with rather attractive informal and formal theoretical models focusing on the ways different institutions shape incentives for voters, interest groups, and policy-makers. But there has always been an elephant in the room that is only sporadically acknowledged: the institutions (e.g. democracy itself, federalism, electoral rules) are themselves endogenous, and we know relatively little about the processes by which history assigns countries to institutional categories. It is not implausible that some of our most cherished “findings” are epiphenomenal. Perhaps there is some unobserved historical process that generates both democracy and development. Perhaps some dimly understood cultural characteristics or some aspect of British colonialism push some countries to maintain majoritarian electoral systems and fewer political parties, while some other process drove continental Europe and Latin America toward proportional representation and multiple parties. Or perhaps if parties are key actors in choosing electoral systems, Duverger’s law is turned on its head and electoral rules are driven by party systems.

This class of problems is at the center of the current research agenda in comparative politics and political economics. Since the 1960s, comparativists have been building theories linking institutional rules with various political and economic outcomes and examining them with cross-country data sets—first limited to the OECD and now much larger. Informal theories like those of Duverger (1964) and Lijphart (1999) have been supplemented with formal theories like those of Cox (1997) and Persson and Tabellini (2000). As the theories become more precise, empirical studies are able to move beyond the blunt cross-country correlations and hone in on the causal mechanisms they imply. And in the last few years there has been a concerted effort to confront the pitfalls of causal inference given that institutions are not randomly assigned.

In order to stay focused, this essay follows these developments in a relatively recent literature that links electoral rules with the generosity of the welfare state. A fairly stubborn stylized fact has been established, and a variety of alternative causal mechanisms proposed. The next step—still in progress—is to link the specific mechanisms with appropriate empirical tests. But the most difficult challenge has been to assess whether electoral rules might plausibly “cause” distinct public policy outcomes at all, given the myriad unobservable factors that might drive the selection of both electoral systems and public policy outcomes. Researchers are only beginning to examine this question. First, they have used the standard econometric techniques like selection models and instrumental variables. Second, they have delved into the question of electoral regime choice with a mixture of theory and the analysis of quantitative and qualitative data drawn from the 19^{th} century to the present. Third, they seek out natural experiments, and finally, they are beginning to conduct field experiments.

After critically reviewing this evolution, I broaden the scope and note that the literatures on the effects of other institutions—in particular federalism and democracy itself—have followed a similar trajectory.

In response to the questions posed above, I argue that progress in clarifying the causal pathways linking institutions and outcomes is slow, difficult, and uneven but palpable and worth the effort. Moreover, the shift of attention to matters of causality in comparative politics has done anything but make the field less interesting. On the contrary, it has breathed new life into a set of fascinating, basic questions about politics and history. The study of institutional origins has become one of the most vibrant research areas in comparative politics. This is a research field in which structural, rationalist, and to a lesser extent cultural approaches collide and interact.

I will argue that it constitutes one of Eckstein's (1980) "core difficulties": further progress in the study of institutions cannot take place until this problem is addressed. It also accords with Lichbach's (this volume) description of a "thorny puzzle." Yet researchers are also motivated by what he refers to as "big problems" in the real world. Political scientists have immersed themselves in institutional design and reform debates in countries conquered by the United States as well as at the World Bank, IMF, and the other development banks. But unless they specify convincing theories about the conditions under which specific institutions emerge, evolve, and stabilize, they are selling snake oil.

I conclude that even if they ultimately fall short, comparative social scientists have no choice but to set high standards for themselves in the empirical demonstration of causality. This requires attention not only to empirical methodology, but to theory-building as well. For the questions that seem most important, everything seems to be endogenous to everything, but theory helps cut through the web to develop identification strategies. Given the rapid rise of experimental research in political science, observational researchers risk being marginalized if they are not more honest about the specifics of the biases inherent in their work, and do not explore every possible technique for limiting them. The only other alternative to a stubborn obsolescence seems to be a "barefoot empiricist" insistence on exclusively experimental research that many comparativists would find unattractive.

Starting with a stylized fact: Electoral rules and the welfare state

Countries using proportional representation appear to have significantly larger welfare states and conduct more redistribution than those using majoritarian electoral rules. Though the measurement of the dependent variable is controversial, and requires researchers to subjectively code line items in government budgets as “social” or “welfare” expenditures, the relationship seems to hold up using data collected by the OECD as well as the IMF, in addition to a less subjective measure capturing the difference between income inequality before and after taxes and transfers drawn from the Luxembourg Income Study. The independent variable has been conceived in various ways: generally either a categorical variable or a continuous measure of the proportionality through which votes translate into seats. While most studies focus exclusively on wealthy countries (Crepaz 1998; Lijphart 1999; Huber, Ragin, and Stephens 1993; Iversen and Soskice 2006), the relationship holds up in Persson and Tabellini’s (2003) sample of 60 countries, and appears to hold up in multivariate OLS regressions that include all of the usual covariates in cross-country welfare expenditure models.

Other policy differences have been attributed to electoral rules as well (see below), but it is useful to focus on the welfare state because it has received by far the most empirical attention from sociologists, political scientists, and economists. Moreover, researchers have developed several distinct causal mechanisms, each with its own empirical implications and endogeneity challenges. Let us assume that these scholars are onto something, and examine the progress they’ve made in their efforts to sort out causality.

Causal mechanisms

Contributors to this literature vary in the hubris of their claims to have unambiguously demonstrated causality, but a causal role for electoral institutions is usually at least implied. Iversen and Soskice (2006) describe their empirical results with language like "affect" and "is a result of" but generally avoid "cause," and then promptly published another paper (2007) implying that electoral rules are endogenous and at best an intervening variable. Persson and Tabellini (2003), on the other hand, are quite explicit in claiming that electoral rules cause policy outcomes.

Before examining the thorny endogeneity problems, it is useful to clarify the causal mechanisms suggested in the literature. I break the causal claims into three broad categories.[1] First, electoral rules might shape politicians' incentives to provide specific versus targeted benefits. Second, electoral rules might shape policy by favoring parties of the left or right in the process of government formation after elections, or even before that stage in the translation of votes to seats. Third, in the context of a simple spatial model, electoral rules might shape party platforms, and hence policies, by shaping the identity of the relevant median voter either because of geographic clustering or an impact on turnout.

Persson and Tabellini's (2000) model has quickly become the dominant perspective. The logic is simple. With majoritarian rules and single-member districts, a party needs only to receive a plurality of votes in half the districts plus one in order to win an election and form a government. This provides incentives for re-election seeking incumbents to target expenditures and other policy benefits to the pivotal districts that

[1] For a more complete literature review, see McDonald (2006).

determine election outcomes while ignoring superfluous districts. Under proportional representation (in the ideal case, with one national district), a vote is worth seeking no matter where it is located, and parties consider the welfare of the entire electorate as they attempt to maximize their representation. This provides them with incentives to propose and implement policies, like the welfare state, that are national in scope. Different modeling choices and assumptions generate theoretical results with a similar flavor in Lizzeri and Persico (2001) and Milesi-Feretti et al.(2002).

Iversen and Soskice (2006) envision an entirely different causal path. Their starting point is a party's pre-election inability to commit to the policies it will pursue after the election. In a system with single-member districts and only two political parties, middle-income voters do not want to risk choosing the party of the left—even if it claims to have moderate positions—for fear that it will succumb to its left wing after the election and tax both the rich and middle class in order to redistribute to the poor. The worst that can happen to the middle class if the party of the right reneges on its moderate promises is that public services fall below their desired level. Under proportional representation, the middle class is represented by a centrist party that can bargain with parties on the left by extracting a tax-transfer scheme that soaks the rich, and the party of the left is in a better position to form coalitions with the center than the right. Iversen and Soskice (2006) summarize evidence that left-center governments are far more common under proportional representation than under majoritarian rules in OECD countries, and argue that the causal impact of institutions on redistributive policy flows through a partisan bias in the coalition-formation process.

Building on observations made by Butler (1951) and Gudgin and Taylor (1979), Rodden (2006) argues that the geography of industrialization is such that leftist voters are highly concentrated in cities and mining areas. Thus it is difficult to draw small, single-member electoral districts in industrialized societies that are not biased in favor of the right. Indeed, it appears that a bad geographic distribution of support has been responsible for a systematic electoral bias against the right in OECD countries using majoritarian electoral rules since World War II. Moreover, there is an asymmetric distribution of district-level vote shares such that there are far more districts (typically in large cities) dominated by the left than by the right.

It seems on first glance like this bias might not impact policies in the long-run, since instances in which right-wing parties form governments while receiving fewer votes than their leftist competitors are relatively rare. However, the asymmetric distribution of district-level vote shares (and preferences over policies) might induce a more subtle effect. Either because they fear entry by new parties on their left flank in the cities, or because urban leftist incumbents dominate the choice of electoral platforms, parties of the left under SMD might find it difficult to offer a platform that is sufficiently moderate for the median voter in the median district. Thus the left might lose relatively more often under SMD because their platform is too extreme for voters in the crucial districts. Or even if the party of the left successfully staves off its urban base and moves its policy position to that preferred by the median voter in the median district, given the asymmetric distribution of district medians, this policy will be to the right of that preferred by the median voter in the country as a whole, thus tilting policies to the right under SMD relative to PR.

Finally, it is possible that electoral rules cause policy differences by altering the preferences of the median voter in a different way. Countries using PR have long been known to demonstrate higher levels of electoral participation. Higher turnout is often thought to imply greater participation among the poor, which should move the preferences of the median voter leftward, with obvious implications for social policy (Franseze 2001; Boix 2003).

Empirical attempts to illuminate mechanisms

An attractive feature of this literature is that each causal mechanism contains implications beyond a simple correlation between electoral rules and welfare expenditures. For starters, the authors have sought to differentiate their products with different ways of measuring the dependent variable. Iversen and Soskice (2006) note that their model is about income redistribution, and correctly point out that many "welfare" expenditures in Europe are not redistributive at all, but are in fact subsidies for the middle class, leading them to use data from the Luxembourg Income Study to capture redistribution. This is not a problem for the Persson and Tabellini model described above, since their aim is to use social expenditures not to capture redistribution, but rather, the prevalence of programs that are national in scope. Milesi-Ferretti et al. (2001) have a similar goal, but choose a different set of line items from the same budget data.

Ultimately, these efforts don't help differentiate between the causal mechanisms, though, since these variables are highly correlated, and all seem to demonstrate the same pattern. In fact, the relationship looks fairly similar when overall expenditures as a share of GDP are used as the dependent variable—a relationship that Persson and Tabellini (2003) attribute to a different causal mechanism, subsequently updated by Persson, Roland, and Tabellini (2007). It is worth mentioning that PR is also said to be associated with a host of additional policy outcomes, including but not limited to higher income taxes (Powell 1982), higher deficits and different ways of adjusting to macroeconomic shocks (Persson and Tabellini 2003), more liberal abortion policies (Powell 1982), greater efforts at environmental protection (Lijphart 1999, Fredriksson and Millimet 2004), lower incarceration rates and less use of the death penalty (Lijphart 1999), as well as policies favoring producers over consumers (Rogowski and Kayser 2002). Most of these seem to be policies favored by the political left, though in some cases, one might also tell stories about general versus specific interests (e.g. Fredriksson and Millimet, 2004 on environmental protection).

Another way to bolster or undermine confidence in the various mechanisms would be to find cross-country indicators that capture one causal claim or another and use them as independent variables, subjecting them to various statistical horse-races. For instance, one might try to show that the causal impact of electoral rules runs through such readily observable things as partisanship, coalition government, bias in coalition formation, bias in the vote-seat curve, or turnout. To my knowledge, this has not yet been done with a large data-set, though in Iversen and Soskice's limited OECD data, it appears that the impact of electoral rules on redistribution is not attenuated very much when controls for turnout, partisanship, and electoral bias are added to the model.[2]

[2] Using data for total expenditures, Boix (2003) finds that turnout washes out the impact of electoral rules. Also using data for total expenditures, Persson, Roland, and Tabellini (2007) claim to show, using a two-stage model, that the causal role of electoral rules is in shaping the number of parties, which in turn shapes the prevalence of coalition governments, which spend more than single-party governments.

Perhaps this should not be surprising, though, since the causal mechanisms reviewed above are not mutually exclusive—all could be contributing to the correlation.

Perhaps a more promising way to hone in on the causal mechanisms is to move beyond cross-country regressions. For instance, the dominant perspective implies that under SMD, public expenditures should be targeted at swing districts, while the "core support" districts for the two parties should receive far less. It so happens that there is a huge empirical literature on this question, though the theoretical motivation is generally a contrast between Cox and McCubbins (1986) and Dixit and Londregan (1995). It is difficult to summarize the literature without conducting some kind of meta-analysis, but my impression is that the strong version of the "swing district" logic that drives the Persson-Tabellini model has not fared well in empirical tests (see, e.g. Ansolabehere and Snyder 2006).

It is of crucial importance for Iversen and Soskice (2006) that there is substantial bias in the process of coalition-formation under PR. That is, over a long period of time, cabinets tend to be to the left of legislatures. Admirably, they seek to demonstrate this by contrasting the ideology of the median cabinet member with that of the median legislator, using the average of several "expert survey" scores of party ideology as their proxy. They find a gap indicative of left bias. Unfortunately, however, this has been called into question by McDonald, Mendes, and Budge (2004), who find no such gap when using data from party manifestos rather than expert surveys as their proxy for ideology. Even if one trusts the "experts" more than the coding techniques used in the manifesto project, it would be useful to dig deeper and verify that the bias comes from center parties who, as formateurs, choose to eschew equally attractive parties on the right in favor of the left.

Moreover, it would be useful to demonstrate that in PR systems, there is indeed such an animal as “middle class” parties whose voting support comes primarily from middle-income voters. An alternative possibility is that multi-party systems under PR reflect the presence of multiple issue dimensions rather than a set of finer-grained distinctions along a single economic dimension. Finally, Iversen and Soskice (2006) have generated a fascinating, intuitively appealing, yet to my knowledge untested hypothesis that might be amenable to research using survey experiments: when there are two parties, risk-averse middle-class voters are more fearful of the left than the right, even if the two parties offer the same platform.

Rodden (2006) finds evidence that industrial geography creates the hypothesized asymmetric distribution of district-level vote shares in several countries with majoritarian electoral rules, and demonstrates bias in the translation of votes to seats, but does not (yet) connect any of this with public policy outcomes. An important task in unearthing the causal mechanism here is to get estimates for district-level preferences that are not drawn from election results in order to examine whether in fact there is a gap, as hypothesized, between the preference of the median voter in the country and the median voter in the median district.

Finally, the causal mechanisms suggested by the turnout hypothesis are also a good target for careful empirical research. Do exogenous increases in turnout lead to greater participation among the poor? If so, does this in fact shift the preferences of the median voter? Moreover, it would be nice to know if the lower turnout rates under SMD are merely a product of voter apathy in the “safe seats” that figure so prominently in Persson and Tabellini (2000) and Rodden (2006). It would be straightforward to match on some observable district characteristics and contrast turnout rates in the most competitive districts of SMD systems with those in PR systems.

The endogeneity problem

Hopefully the forgoing discussion conveyed the sense that the exploration of causal mechanisms linking electoral rules and public policy is an active and interesting research program. However, I have not yet mentioned the elephant in the corner. These mechanisms are not causal if they are driven by unobserved characteristics of countries that also drive welfare expenditures and redistribution. Rather than attempting a review of causal inference problems in comparative empirical research in the presence of simultaneity, I direct the reader to Heckman (2005) or Przeworski (2007) and the citations therein.

The basic notion of causality implied in most of the literature reviewed above is explicitly addressed in Persson and Tabellini (2003: 114): "We would thus like to answer counterfactual questions like the following: Suppose we pick a country at random in our sample and, going back in history, change its constitution. How would this alter its current performance?" The argument is that if it were possible to randomly assign the "treatment" of PR to one group of countries and SMD to another, the PR countries would develop larger welfare states (along with greater attention to national public goods, left bias, higher turnout, etc. depending on the theory). The causal effect of the treatment on an individual country is defined as the difference between the outcome variable in that country under PR and SMD, and we are interested in the average of those individual treatment effects.

The problem, of course, is that we cannot observe the relevant counterfactual. For the most part, in the era of the modern welfare state we are only able to observe Sweden under PR and the UK under SMD (more on this below). Thus researchers implicitly make the assumption of unit homogeneity (Holland 1986), which states that if two units have the same values of covariates, they would demonstrate the same states under control and the same states under treatment. The problem of selection is assumed away, and it does not matter which of two units goes into the control group, and which to the treatment group. This must be assumed rather than tested, but the assumption seems reasonable enough when the investigator can randomly assign individuals to the treatment and control groups.

Additionally, the studies reviewed above implicitly assume conditional mean independence. That is, the countries that have not been exposed to the treatment (PR) would (hypothetically) react to it identically to those where the treatment has been received, and the countries receiving the treatment would (hypothetically) not differ in their control state from the group in which the control state (SMD) is observed. Again, this is achieved in experimental studies through randomization.

But in studying national constitutions, of course, researchers rarely have opportunities for random assignment.

Thus the literature on institutions and outcomes is based on very strong assumptions. To the extent that they are not met, the inferences about the causal impact of institutions are biased. While reverse causation and measurement error might also be important, the most obvious problem here is *selection*. First of all, there may be an unobserved difference in the control state between the countries that were exposed to the treatment and those that were not. It is possible, for instance, that proportional representation emerged in countries like Sweden, where a "history of equal opportunity, broad education, and widespread ownership of land… fostered a common culture of equality that has promoted the selection of a proportional electoral system as well as a preference for a welfare state" (Persson and Tabellini 2003: 115). The problem here is that such a "culture" and its associated "preferences," especially to the extent that they are rooted in the 19th century or earlier, are difficult to observe and measure across countries. Moreover, British colonial experience is observable, but it is so highly correlated with electoral rules in OECD data sets that any separate impact of electoral rules would be driven by France and Japan alone.

Given the arguments about partisan bias above, one can also envision a self-selection bias, where the *effect* of the treatment on the treated countries is different from its (unobserved, hypothetical) effect on those who did not receive the treatment. This might happen if, for reasons unobserved by the researcher, countries recruit themselves into the treatment group because they anticipate a desired outcome. For instance, anticipating greater redistribution, the poor and vulnerable might push for proportional representation. Alternatively, if they are aware of the logic sketched by Iversen and Soskice (2006) or Rodden (2006), party leaders on the left should advocate for proportional representation. If the poor, the leftist parties, or both, are more mobilized and powerful in some countries than others, they may be able to maintain their favored institutional framework along with their favored policies. In this event, electoral rules should not be viewed as causing policy outcomes.

Empirical strategies to confront endogeneity

Empirical researchers have developed (at least) three types of strategies for dealing with these problems. First, they reach into an ever-expanding econometric toolkit relying on structural models. Second, these structural models call for the analysis of history from a new perspective. Third, researchers attempt to exploit natural experiments.

Econometric Analysis

In their path-breaking book, Persson and Tabellini (2003) devote considerable attention to exactly this class of problems. After finding a robust relationship between electoral rules and expenditures in a multivariate OLS model, they move on to explore matching estimators. As with OLS, this approach assumes that selection of countries into constitutional rules is essentially random, and using a matrix of observable attributes, develops a propensity score that aims to match each country using a particular constitutional rule with an otherwise similar country using the opposite rule. While this is a potentially useful innovation given concerns about non-linearities in the relationship between constitutions, observable covariates, and outcomes, it still sweeps the endogeneity problem under the rug by assuming conditional independence.

Like most other comparative institutional researchers, Persson and Tabellini rely primarily on the instrumental variables strategy to guard against the biases that might arise if conditional independence and unit homogeneity do not hold. The strategy is to find a vector of instruments that predicts the endogenous regressor (electoral rules), but is orthogonal to the error term, which means that the vector of instruments has no effect on the size of the welfare state other than its impact through electoral rules.[3]

Finding good instruments for institutions like electoral rules is surely more art than science. First, the instruments must be highly correlated with the institution in the first stage regression in order to avoid a cure that is worse than the disease (an even more biased estimator than that caused by the selection problem). Second, the instruments must be exogenous. In this context, controlling for the set of instruments, the remaining random component of constitutional selection must not be correlated with the country-specific, unobserved change in welfare expenditures associated with a change in electoral rules. Sadly, this can only be asserted but not tested. Thus the researcher must tell a convincing story about the exogeneity of the prized instrument, but a leap of faith is always required.

Persson and Tabellini (2003) spill surprisingly little ink convincing the reader to take this leap, and have little to say about the justification of their instruments. First, they include three variables that place the year of adoption of the country's constitution into four periods. The idea is that there appear to have been "waves" during which some constitutional types were popular, and the period during which a country declared independence or adopted its constitution might be correlated with the electoral rules, but

[3] They also use Heckman selection models, but since these rely on the same exclusion restrictions as the instrumental variables approach, much of the discussion that follows is relevant for this approach as well.

Persson and Tabellini argue that there is no reason to believe that the pure timing of constitutional choice should have an impact on contemporary expenditure outcomes. This might be debatable since, as discussed below, many of the countries with the largest welfare states selected institutions of proportional representation during a very brief period after World War I when soldiers returned form the war, the franchise was expanded, and socialist parties were mobilizing—sometimes in the streets and with weapons. In any case, such a debate may be irrelevant for their estimation, because in a reanalysis of their data, Acemoglu (2005) shows that these instruments are hopelessly weak, and do not approach statistical significance by themselves in the first-stage electoral rule equation.

Thus the identification strategy relies solely on the additional instruments, which were adopted from the work of Hall and Jones (1999): latitude, the percentage of the population whose native language is English, and that for whom the mother tongue is another European language. Though Persson and Tabellini are largely silent about the logic behind these instruments, the idea is apparently that these variables reflect the depth of European cultural influence on clusters of institutions,[4] but do not have a direct impact on contemporary fiscal outcomes.

The coefficients in the first-stage regressions suggest that indeed, "proportion English speaking" is associated with majoritarian electoral rules, while "proportion speaking other European languages" is associated with PR. Yet the English-speaking variable is essentially a proxy for British colonial experience, and it is not at all clear why one should believe that electoral rules are the only, or even the most important, aspect of

[4] In the literature from which these instruments are drawn, these institutional clusters are understood as things like "property rights" and "rule of law" rather than specific institutions like electoral rules.

British colonialism that might impact welfare expenditures. For instance, according to La Porta et al. (1998), these countries had more developed financial markets, which may have had an impact on the development of government spending. One can think of many additional unobserved aspects of British, or for that matter French, Belgian, or Portuguese colonial experience that might have an impact on long-run patterns of government expenditure. The Hall and Jones instruments are best suited to a sample of colonies, but many of the countries in the Persson and Tabellini sample in which European languages are spoken are located in Europe of all places, and here again, there are many unmeasured aspects of "Europeanness"—for example culture, preferences, the power of labor unions and the left—that might have an impact on welfare expenditures that does not flow through electoral rules. Acemoglu (2005) makes the important point that even if a matrix of instruments seems plausibly valid for a cluster of inter-related institutions, it is probably not valid for a specific institution like electoral rules if the other associated institutions (common or civil law, bicameralism, etc.) are not entering the equation as control variables.

The point of this discussion is not that Persson and Tabellini have done shoddy work. On the contrary, these results and their connection to the underlying theories established in Persson and Tabellini (2000) represent one of the most important contributions to comparative politics in recent decades. There is clearly a relationship between electoral rules and government expenditures, and it holds up in a wide range of countries. Yet in spite of their strong claims, the issue of causality is far from settled. It relies on exclusion restrictions—ultimately assertions—that are plausible but open for debate.

The same can be said for almost every other attempt to find historical instruments for institutions. Acemoglu (2005) sharply criticizes the validity of the Hall and Jones instruments, and the validity of even the most celebrated instrument for institutions—the settler mortality data mobilized by Acemoglu, Johnson, and Robinson (2002)—has been called into question (Djankov, et al. 2003, Glaeser et al. 2004, Przeworski 2004).

The point is not that the instrumental variables approach holds no promise. Rather, students of endogenous institutions must go well beyond searching for a simple, off-the-shelf magic bullet. Rather, as Acemoglu, Robinson, and Johnson (2004: 7) argue, "To find an instrument one needs a theory of why institutions differ. The theory can act as a guide to find an instrument." Instrumental variables and selection models are not technical fixes that allow researchers to avoid the hard work of delving into serious, theory-guided historical analysis. Rather, in order to come up with believable exclusion restrictions, one must essentially become an analytical historian. This requires more than a passing glance at the secondary literature.

In addition to searching for instruments, such theory-guided historical research might also reveal that things previously deemed "unobservable"—like preferences of voters, the mobilization of labor unions and the left, social arrangements associated with the guild system, the strategic choices of elites—can indeed be theorized and measured. In other words, theory-guided historical research mobilizing quantitative and qualitative data can attempt to build more complex but satisfying structural models where institutions are chosen in the first stage, and have an impact (or not) on contemporary outcomes in the second stage.

In short, it will be difficult to make progress in the comparative institutional literature without doing analytical history. Fortunately, this is exactly where the literature is headed.

Analytical History

Recent years have seen a renaissance in the study of endogenous electoral systems. Most of the analysis focuses on a handful of OECD countries, and much of the empirical action takes place in interwar Europe, while a more recent literature expands the time frame and cases, with special emphasis on the new democracies of Central and Eastern Europe. The first step in this renaissance was the publication of Boix (1999), which revived the classic argument of Braunias (1932) and Rokkan (1970) about electoral choice at the beginning of the 20th century. Especially after World War I, the established elite parties understood that they could not avoid an expansion of the franchise, which they expected would dramatically increase the representation of a new, unified leftist party. When the old parties (often Conservatives and Liberals) feared that it would be too difficult to overcome their differences and coordinate under single-member districts, their best response was to choose proportional representation in hopes of forming a post-election coalition. In the Rokkan and Boix stories, elites are apparently either unaware of the long-term disadvantages of proportional representation for their class interests, or are simply too concerned with their short-term electoral prospects. Rokkan describes right-wing parties who explicitly undermine their class interests and favor PR primarily because they fear being squeezed out of electoral politics by the other right-wing rival. In the terminology used in this volume, this is not a structural argument where social groups and classes play a role. The actors are self-interested politicians seeking to win the next election.

If Rokkan and Boix have it right, perhaps the problem of identifying subsequent causal effects of electoral institutions is not so severe. Perhaps some measure of partisan fragmentation on the right could serve as an instrument if one believes that these rifts on the right—many of them based on religious or town-rural cleavages that subsequently vanished—are essentially random and not plausibly correlated with long-term expenditure patterns.

But no such luck. The next thing to happen in the new literature on endogenous electoral rules was a rather intense focus on criticizing this theory and especially the empirical evidence. The Boix data have been reanalyzed by Cusack, Iversen, and Soskice (1997), Blais, Dobrzynska, and Indridason (2004), and Andrews and Jackman (2005), each of whom attacks the empirical results from a different angle and finds a way to make them disappear. Perhaps this is not surprising since these studies are based on *N* somewhere between 12 and 20 countries, but one critique in particular seems especially compelling. Blais et al. (2004) point out that the move to proportional representation was generally not from the single-member district plurality systems envisioned in the classic argument, but with only the exceptions of Sweden and Denmark, from systems with some form of two-round majority elections. Thus it is difficult to understand why the right would have faced a coordination problem, since even if split in the first round, their voters should have had no problem coordinating on the sole remaining right-wing candidate in the second round. Moreover, it appears that parties of the right were not often in direct competition at the district level in the first place. In fact, most of these countries already had multi-party legislatures and coalition governments prior to the adoption of proportional representation.

Thus Blais et al. (2004) argue that the transition to proportional representation was, in countries with multi-round elections, a fairly uncontroversial move spurred by intense public support for the proportional representation "trend" that was sweeping Europe as part of a larger trend of democratization. Perhaps run-off SMD parliamentary elections, if kept in place, would have generated coalition governments not unlike those witnessed under proportional representation. If this view is correct, for researchers interested in identifying a causal role for institutions, the relevant question is whether there is some reason to believe that there was something unique about countries with multi-round elections that made them more likely to develop large welfare states in the future. To my knowledge this question has not been addressed.

Blais et al (2004) seem to suggest that the enthusiasm for proportional representation was equal across the political spectrum, and that the only skeptics were change-averse sitting parliamentarians who had gained their seats through the old rules. Again, the interests of class groups are not discussed. Like the traditional Braunias/Rokkan perspective, this almost certainly underestimates the affinity of the labor movement and the political left for proportional representation. The adoption of proportional representation was a key demand in the platform of virtually every socialist or workers' party at this time.

This point is emphasized in Alesina and Glaeser (2004), who argue that proportional representation emerged in early 20th century Europe as a direct response to the demands of a mobilized, revolutionary left. Their tour of early 20th century Europe includes a brief paragraph on each country emphasizing the role of leftist movements in pushing for, and ultimately achieving, the expansion of the franchise to the poor and the adoption of proportional representation through strikes, street protests, and the threat of violence. In some cases—interwar Germany, Austria, and Italy in particular—the aristocratic right was severely weakened because of military defeat, and the maintenance of public order was in question. In some cases, like Weimar Germany, the left adopted proportional representation after coming to power. In other cases, like Belgium and Sweden, Alesina and Glaeser suggest that even though proportional representation was implemented by a party of the right, it was the implicit threat of violent revolution that forced their hand.

The treatment by Alesina and Glaeser may over-state the case and over-simplify the complex negotiations leading to electoral reform in the early 20th century. In most countries, the left was pushing simultaneously to abolish property requirements for voting, reserved seats for business, plural voting (upper class voters receiving tens and even hundreds of votes), and powerful appointed upper chambers representing the interests of the elites. Proportional representation was also on the wish list, but the expansion of the franchise was far more important. In Belgium and Sweden, the right-wing parties that made the initial transition to proportional representation did so in an explicit attempt to throw bones to the left and quell the demand for electoral reform while preserving their advantages through franchise restrictions, plural voting, and an undemocratic upper chamber (in Sweden), while maintaining small, badly-apportioned districts that prevented proportional election results. In both countries, the initial transition to proportional representation was essentially a sham (Carstairs 1980, Verney 1957). Yet in spite of remaining franchise restrictions, these electoral reforms did strengthen the parliamentary representation of the left, and in the decades that followed, they were able to achieve their larger agenda to expand the franchise through a combination of parliamentary power and extra-parliamentary agitation.

While the details are sparse, Alesina and Glaeser are clearly onto something. Socialist party leaders understood that proportional representation was in their interest, and since Braunias (1932), the literature has lost sight of this by focusing on the complex handful of cases where right-wing parties disingenuously promoted PR. And as Alesina and Glaeser point out, this creates an identification problem for the literature positing a direct link from electoral rules to the rise of the welfare state later in the century. A precondition for proportional representation was a strong and organized leftist workers' movement with organizational support from labor unions, which is also likely an important part of the story in the rise of the welfare state. According to Alesina and Glaeser, for various reasons the United States and Britain never developed this type of leftist organization, and as a result, developed neither PR nor a large welfare state. Electoral rules in this story are epiphenomenal.

Ticchi and Vindigni (2005) present a model with a similar flavor. Rather than focusing on electoral rules, they draw on Lijphart (1999) and contrast "majoritarian" with "consensus" democracy. In their interpretation of majoritarian democracy, there is a winner-take-all election in one large national district. There are three income groups/parties, and while taxes are proportional, expenditures are highly targeted to the group in power. Thus the wealthy group dominates under majoritarianism because of its preferences for low taxes, a phenomenon they refer to as the "dictatorship of the rich." They model consensus democracy as a process of coalition bargaining between the three groups, and obtain that a less skewed income distribution facilitates a coalition between the poor and middle class, which enhances the size of government, while a highly skewed distribution pulls together the rich and middle class, putting the brakes on government expenditures.

Then moving back a step, they examine the initial selection of institutions, which are chosen by simple majority voting, and every actor "correctly anticipates what their level of utility would be under the two possible constitutions, and vote consequently" (Ticchi and Vindigni 2005: 19). Their model implies that the majority in highly unequal societies would choose majoritarian rules. In a relatively equal society, the majority would choose consensus rules. The rather unusual feature of this model is that the poor rather than the middle class are the "swing" actors. The model implies that if income inequality is sufficiently high, the poor would be frozen out of a coalition-formation process dominated by the center-right under consensus democracy, which would cost them more in taxes than the "dictatorship of the rich," but still provide them with no public goods. Thus for Ticchi and Vindigni (2005), each country's initial income distribution occupies the role played by the organizational power of the left in Alesina and Glaeser's story. It is an exogenous variable that drives both institutional choice and later public spending outcomes.[5] Constitutional rules are once again epiphenomenal.

[5] One should not characterize Ticchi and Vindigni's model as explaining redistribution, however, since redistribution across income groups is explicitly ruled out in the model, which makes very strict and unusual assumptions about targeting of expenditures to class groups and the impossibility of broad social programs.

The same is true for Cusack, Iversen, and Soskice (2007). For them, the omitted variable driving both institutional choice and the development of the welfare state in OECD countries is the structure of the economy in the late 19^{th} century. In this view, everything flows from exogenous "varieties of capitalism." They downplay the role of leftist agitation emphasized by Alesina and Glaeser (2004), and seem to build a theory around the cases of Denmark and the Netherlands, where PR was chosen peacefully and with little controversy as part of a cross-party bargain. As with the traditional Braunias/Rokkan/Boix perspective, proportional representation is a conscious choice by the economic elites who saw it as in their long-term self interest. Proportional representation reflected a class compromise between employers and skilled workers in export-oriented economies who had a common interest in a regulatory and social insurance systems that would protect investments in co-specific assets. In this theory, proportional electoral representation of parties in the legislature seems to be less important than a concomitant system of legislative committees representing employers and workers. Due to their lack of economic coordination or investment in co-specific assets, Britain and its former colonies avoided proportional representation because the economic elites had no incentives for class compromise. If this theory is correct, we are again unable to identify a causal role for electoral institutions in the rise of the welfare state, since a history of local economic coordination through guilds drives the class compromise that gave rise to both PR and the welfare state.

Though quite different, the stories of Alesina and Glaeser (2004), Ticchi and Vindigni (2005), and Cusack et al. (2007) have some features in common. First, while the actors in the Braunias/Rokkan/Boix framework are rational, strategic, self-interested politicians who are primarily interested in getting and retaining power, possibly to the detriment of their long-term class interests, in these new models of endogenous institutions the actors are social classes, income groups, or social groups like factory owners and industrial laborers (farmers are curiously missing). Political parties are mere stand-ins for the interests of these groups. In the terminology of this volume, these works combine elements of rational choice theory and classic structuralism.

While the citizen-candidate framework of Ticchi and Vindigni (2005) is most explicit about it, each of these papers assumes away collective action and agency problems that might be faced by actors on both the right and left. Potential conflicts between the old parties of the right are ignored, as are the serious rifts between Communist revolutionaries and Social Democrats on the left, and the fact that within parties, individually rational behavior does not translate into social rationality. In general, this literature can be described with the same words used by Przeworski (1985: 382) to describe Marxism: "What was important about history happened at the level of forces, structures, collectivities, and constraints, not individuals… Marxism was a theory of history without any theory about the actions of people who made this history."

Second, these class or income groups are far-sighted and have full information about the long-term impact of their choices. There is no veil of ignorance. For Alesina and Glaeser (2004), who discuss endogenous electoral rules directly after presenting some regressions establishing the relationship between electoral rules and social expenditures, it is self-evident that representatives of poor and vulnerable workers would foresee this relationship and advocate for their class interests. For Ticchi and Vindigni (2005), at moments of constitutional choice, each income group is capable of making the complex computations required to fully anticipate the translation of the precise income distribution into its interest in the outcome of a coalition game under consensus rules, and contrast that outcome with the majoritarian rule outcome. For Cusack et al. (2007), the economic elites are able to anticipate the costs of being soaked by the poor under PR and contrast this with the long-term economic benefits of maintaining a skilled and happy workforce.

Third, like Marxism, these theories are animated by a sense of functionalist inevitability. For Alesina and Glaeser (2004), if unions are strong and the left is mobilized, proportional representation can only be delayed but not avoided. For Ticchi and Vindigni (2005), the socially efficient constitution for a given income distribution will emerge naturally over time. For Cusack et al. (2007), proportional representation evolves naturally as a reflection of the society's needs for economic coordination.

This new class of structural-functionalist analytical history is enormously interesting, and constitutes a promising research program in its own right. But there are reasons to hesitate before rejecting the possibility of a causal role for electoral institutions based on analyses that are so lacking in micro-foundations and careful historical analysis. A number of questions present themselves. Why did leftist leaders feel that proportional representation was in their interest? Did leftists have a different interpretation of their interests in Britain and other countries that retained SMD, did they face an internal collective action problem, or were they simply too weak to get their way? Is there any evidence in party documents or personal communications of leaders that the relevant actors understood potential implications of electoral rules for redistribution and expenditures on public goods? How did politicians grapple with situations in which class interests and personal electoral interests may have conflicted? What were the relevant splits within parties, and how did these link up with class and income groups? Is there any evidence that employers' associations in industries with specific skill investments lobbied the parties of the right to promote PR? How important were the exogenous constraints associated with World War I in strengthening the hand of the left?

A general challenge for empirical analysis associated with sweeping structural macrohistorical arguments is to avoid qualitative or quantitative work that is akin to conveniently selected historical anecdotes. It is helpful to have the claims analyzed by disinterested empirical scholars who are not already wedded to them. Only after the work receives considerable attention do other scholars begin to mobilize dissonant historical information, collect additional data, or conduct robustness checks. Some interesting arguments have been placed on the table about endogenous electoral rules, but more work lies ahead. My sense is that the renaissance in the historical analysis of constitutional choice has just begun. As with the renaissance in the study of democratization in the early 1990s, this will likely be characterized by a productive and critical dialogue between these rational-choice structural theories emphasizing the role of class and other social groups as actors, and a different class of rational choice theories that feature office-seeking politicians and focus attention on incentives faced by individuals.

The latter perspective has dominated among political scientists seeking to shed light on constitutional choice in new democracies and recent electoral reform in older democracies, where scholars have much better access to information about beliefs and preferences of key actors, internal party debates, and the role of social groups than in interwar Europe (e.g. Benoit 2004, 2006; Moser 2001, Remington and Smith 1996, Kaminski 2002). The clearest lesson in this literature is about the uncertainty of actors over the political and social outcomes that would flow from institutional choices. Virtually every study of Central and Eastern Europe features an episode in which a set of rules advocated by a party or faction had exactly the opposite impact of what was expected, in many cases marginalizing or ending the careers of its advocates.

It seems likely that similar or even greater levels of uncertainty plagued institutional designers in interwar Europe, given the simultaneous doubling or tripling of the electorate and the lack of polling data (Andrews and Jackman 2005). Taagapera (2002) and Shvetsova (2003) go so far as to argue that uncertainty over potential partisan and policy implications of different constitutional alternatives at moments of reform is so great that institutional endogeneity is simply not an issue for scholars interested in examining the impact of electoral institutions on party systems and policy outcomes. In this view, Socialists could not possibly have anticipated the long-term implications of proportional representation for coalition dynamics and party development. Perhaps PR could have generated a rancorous split on the left, as in Italy, or partisan fragmentation, as in Weimar Germany.

Endogeneity also becomes less troubling when one thinks about the other central observation in this literature: political actors seem to have had surprisingly short time horizons, thinking not in broad terms about the strategic interests of the party or its constituents in the long term, but rather, about the coalition dynamics of the next election. Andrews and Jackman (2005) point out that during a short-lived pact between the Liberals and Labour in 1906 and 1910, these parties enjoyed a temporary improvement in the translation of votes to seats, though they had traditionally been disadvantaged by their inferior geographic distributions of support relative to the Conservatives. Of course the Liberals have suffered in the vote-seat curve ever since, as has Labour to a lesser extent (see Rodden 2006). Proportional representation was very much on the political agenda at the time, but in a series of votes in 1917 and 1918, almost half of Liberal and Labour MPs voted against it. By 1924, the Liberals came to understand their mistake, and strongly advocated PR. But in spite of the supposed long-term interests of their constituents in PR, the majority of Labour MPs supported the retention of SMD in order to squeeze out Liberal competition for the left-wing vote.

This affair demonstrates the difficulty of structural arguments in which parties are viewed as “carriers” of class interests. A similar lesson can be learned when parties are viewed as aggregations of individuals rather than unitary actors. Individual leftist incumbents with carefully cultivated bailiwicks or safe seats, for instance, might have incentives to fight reforms like list-based proportional representation. U.S. Congressional incumbents often successfully squelch efforts by their own co-partisans in the state legislature to improve the party’s position in the reapportionment process because it would do violence to the district boundaries with which they have become comfortable (Butler and Cain 1992).

In short, the recent political science literature makes the move described by Katzlenson (1997) in the previous edition of this volume, reacting to ambitious macroanalytic structural work with a style of historical institutionalism that highlights contingent events and uncertain calculations, though largely from a rational choice perspective. In so doing, it challenges the strongest claims of institutional endogeneity implicit in recent structural theories. The verdict is still out, and perhaps it always will be. But at the very least, the strongest claims of institutional endogeneity seem to be unsubstantiated.

I am relatively certain that leftist leaders in early 20th century Europe preferred proportional representation. My intuition is that the reason for this has not yet been emphasized in the literature, though it was recently suggested by Andrews and Jackman (2005): leftist parties had a relatively skewed geographic distribution of support owing to the uneven geographic spread of industrialization, and stood to gain in the translation of votes to seats under PR. It may well be true that proportional representation was in many cases the result of leftist mobilization.

Yet it is not all that difficult to envision that the same countries, with the same background conditions, could have easily maintained SMD and developed smaller welfare states due to some combination of the causal mechanisms above, if only things had gone slightly differently at the crucial constitutional moment. When reading the historical accounts that are not wedded to a theoretical perspective, one is struck by the contingency owing to the specific strategies selected by key actors. For instance, what if the old parties of the right in Sweden and Belgium could have had more foresight and withstood short-term losses in exchange for long-term gains, like the British Conservatives, by understanding that plural voting and property restrictions were destined for the dustbin of history, but holding fast to SMD? Instead, they tried to hold on to their privileges as long as possible by giving in on proportional representation. At least in Sweden, it is very clear from Verney (1957) that the left would have accepted SMD under the right conditions. It is plausible that this would have locked into place a different set of institutional incentives that could have generated a different tax-transfer regime.

What if Labour in the UK had paid less attention to vanquishing the Liberals, and more attention to competing in the long run with the Conservatives? Or what if the Liberals had faded sooner, or merged with Labour? What if World War I had ended differently, and aristocratic elites regained a firm grip on military and police power in Germany, Austria, and Italy? Perhaps then the constitutional moment would have come and gone without electoral reform in some other countries as well had elites not been so fearful of social disarray.

In historical studies where randomization is not possible, in order to think about causality, we must invoke counterfactuals, and to do so, as Przeworski puts it, "we must be assuming that history could have generated a world different from the one in which we live, that realizations of history other than the actual one are possible" (2007: 9). Given the contingency and uncertainty associated with constitutional decisions, it seems quite plausible that the same countries, with the same social groups and elites, could have developed different electoral rules and welfare regimes.

Natural experiments

These historical counterfactuals are interesting but frustrating. Like the exclusion restrictions in econometric research using instrumental variables or selection models, the researcher tries to be as persuasive as possible, but belief in causality still requires a rather large leap of faith. This is a distant second-best solution in a world where institutions cannot be randomly assigned to subjects, and where it is not possible to observe the same subject under conditions of treatment and control.

But a growing trend in economics and political science is to seek out the rare moments when history bestows an opportunity that resembles randomization. For instance, Banerjee and Iyer (2002) are able to estimate the contemporary impact on productivity and investment of different systems of land taxation set up in Colonial India since these systems depended on the date of conquest rather than any characteristics of the districts. The analysis is especially persuasive because the researchers were able to observe a large sample of geographic neighbors that were similar with respect to background conditions but had been subjected to different institutions of land taxation.

The prospects for this type of research in the area of electoral rules may not seem especially good at first glance, but there are a number of possibilities, especially when one examines some of the specific causal mechanisms outlined above. Above all, it *is* possible to observe the same population under different electoral regimes. These opportunities are provided by such institutional features as state and local governments with different electoral regimes than central governments, bicameralism, and the contrast between district-level results of presidential and legislative elections, and the difference between national and supra-national (European) elections.

The incredible diversity and complexity of electoral institutions in the United States is especially appealing. For instance, in order to test arguments about the importance of the distribution of district-level preferences, it is appealing to contrast the ideological profiles of the senate delegations from large states, who are elected at-large from statewide elections, with those of House members from the same state, who are elected from small single-member districts. Local governments in the United States use a wide variety of electoral institutions including proportional representation. For a long period, the state of Illinois used SMD in one legislative chamber and something akin to list proportional representation in the other.

There are many such opportunities around the world. For instance, the Australian lower chamber uses SMD while the upper chamber uses proportional representation. Even Britain is becoming more diverse as elections to regional assemblies and the European Parliament are conducted with different electoral rules than national parliamentary elections. Alternatively, one might contrast the behavior of individual legislators elected from districts with those elected from party lists in mixed systems like Germany's.

Second, it may be useful to take a closer look at contemporary instances of electoral reform, like those in Italy, France, Japan, and New Zealand. Of course these reforms cannot generally be viewed as exogenous, and there are colorful stories explaining electoral reform as Machiavellian moves by brash incumbents. Yet these moves were often focused on achieving short-term advantage in an upcoming election rather than long-term policy change. The case of New Zealand is especially appealing since electoral reform emerged essentially by mistake when a Labour MP started a landslide by inadvertently promising a referendum on electoral reform during a televised debate, though the leaders of both parties preferred the status quo (Nagel 2004).

Field Experiments

The move toward experimental research in comparative politics is one of the most striking developments of the last decade. The opportunity to randomly assign institutions to subjects would seem to be rare, but such work is possible for researchers who invest in facilitating partnerships with governments, aid agencies, and development banks. Governments in large countries like Mexico and India have been persuaded to randomize policy treatments or even institutions like reserved local government seats for women and minorities at the level of villages or districts, allowing researchers to examine the causal impacts of the policies or institutions in a way that requires far fewer assumptions than observational research. Chattopadhyay and Duflo (2004) examine the effects of female representation by exploiting the random reservation of village council head positions for women in West Bengal and Rajasthan. Olken (2007) randomly assigns institutional mechanisms used for communities to select infrastructure projects in Indonesia, and Humphreys and Weinstein (2007) generate a scorecard rating the performance of members of the Ugandan parliament, and randomize its dissemination across geographic constituencies in order to assess the impact of information and transparency on the performance of elected officials. Most relevant to the discussion in this chapter, Beath, Enidolopov, and Christia (in progress) are able to randomly assign some democratically elected local government councils in Afghanistan with majoritarian election procedures, and the others with "cluster-based elections," while the rest remain without such councils. This type of research will undoubtedly expand tremendously in the years ahead.

Conclusions

By focusing on electoral rules, this essay has identified a pattern that characterizes institutional research more generally. Over several decades, researchers have established some interesting conditional correlations using cross-country data. In addition to the relationship explored above, some highlights include: electoral rules and the number of parties, federalism and welfare expenditures, democracy and economic development. These correlations fueled the creativity of theorists, who developed competing accounts of the underlying causal mechanism. In many cases these accounts led to a wider array of testable empirical propositions. Along the way, questions arose about the direction of causality and the possibility of omitted variables, casting doubt on the causal role of institutions in explaining outcomes. Straightforward applications of the rational choice institutionalist perspective morphed into debates about whether institutions "matter" at all.

To get to the bottom of things, while a smaller strand of literature seeks out natural experiments and more recently, field experiments, two related empirical research strategies have dominated thus far: instrumental variables and analytical history. In fact, I have argued that the successful implementation of the former often requires the latter. In both cases, researchers are looking for truly exogenous variables to help identify causality in a two-stage model where institutions are selected in the first stage and policy or macroeconomic outcomes are explained in the second stage. As a result, we have witnessed a rapid return to fashion of the sweeping structural-historical analysis praised by Katznelson (1997) in the first edition of this volume, although with a rational choice starting point. As scholars attempt to sort out the possible causal role for institutions, these structural analyses contrast with rational choice theories that focus more squarely on the decisions of self-interested individuals at crucial moments of constitutional choice.[6]

Consider one aspect of the literature on federalism. Like single-member districts, a federal constitution appears to be correlated with lower welfare expenditures (for a literature review, see Castles, Obinger, and Leibfried, 2005). Researchers have come up with a number of possible causal mechanisms including status quo bias owing to multiple veto players during the era of welfare state expansion, inter-jurisdictional tax competition, sorting by the wealthy into their own jurisdictions, powerful and conservative courts, and the over-representation of conservative, sparsely populated regions in upper legislative chambers. These arguments have led to interesting empirical studies, but each mechanism generates questions about institutional endogeneity. Perhaps fragmented authority, tax competition, and malapportionment reflect successful schemes by wealthy elites to limit the political power of the poor and undermine redistribution. The current forefront in this literature features theories that make the varieties of federalism endogenous, and empirical analyses that include instrumental variables and analytical history (for a review, see Rodden 2007). Again, there are structural theories in which institutions are driven by factors like the inter-regional and inter-personal distribution of income (e.g. Bolton and Roland 1997, Boix 2003), contrasted with theories focusing on office-seeking politicians (e.g. O'Neill 2003), while some of the work weaves these perspectives together (e.g. Diaz-Cayeros 2006).

A similar pattern characterizes the much larger literatures on democracy and economic development. The correlation has been in the literature for a long time, but the

[6] Though it is often mentioned in passing, analysis of culture as a determinant of electoral institutions has lagged behind, perhaps above all because of the difficulty of measurement, especially in settings where institutional choices were made in the distant past.

1990s saw a renaissance in thinking about the direction and nature of causality, accompanied by a great deal of innovation in both theoretical and empirical analyses, spurred largely by the contributions of Adam Przeworski and his collaborators . Puzzles about the endogeneity of democratic institutions inspired the award-winning work of Boix (2003) and Acemoglu and Robinson (2006). Again, there seems to be a productive friction between these neo-structural-historical works featuring groups like "the poor" and "the elites" as actors, and works focusing on self-interested politicians (e.g. O'Donnell and Schmitter 1986, Przeworski and Limongi 1997).

In conclusion, I return to the questions posed in the introduction. Is the new-found concern with causality adopted from economics turning attention away from the thorny puzzles and big questions that make the field of comparative politics interesting? On the contrary, by turning to historical data in search of natural experiments and instruments, researchers have breathed new life into some of the classic questions of the subfield. In fact, it was a handful of economists puzzling over the endogeneity of political institutions who returned to the sweeping structural approach to analytical history favored by scholars like Barrington Moore, Reinhard Bendix, Gregory Luebbert, and Dietrich Reuschemeyer, now dressed up with game theory.

It is tempting to add Karl Marx to the list. Due to the influence of methodological individualism, theories featuring unitary actors like "the elites" and "the poor" were as unfashionable as Marxism itself in the very recent past. Out were strong assumptions featuring far-sighted, well-informed individuals whose preferences were driven by their place in the income spectrum or social class structure, who could effortlessly overcome collective action problems and translate individual into collective rationality. In were self-interested politicians and interest group leaders whose behavior may or may not link up with the interests of their constituents. It was this rationalist literature focusing on individual politicians, in fact, that generated many of the causal claims about institutions based on the notion that they provided constraints shaping the incentives of office-seeking politicians.

It is not surprising, then, that in the neo-structural literature, where collective action problems and historical contingency play no role and class groups rather than individuals are the movers of history, institutions begin to feel like manifestations of "deeper" forces, as in Marxism. For political scientists steeped in the individual rational choice framework, institutional choice appears to be more contingent, often driven by fleeting coalition dynamics or quixotic efforts to stem vote loss or undermine a rival in the next election by a troubled incumbent with limited information. In this view, the conditions that underpin institutional choice seem unlikely to be correlated with political or economic outcomes 100 years later, and causal claims about institutions seem less problematic.

These literatures generate a creative tension that is driving the institutional literature toward new ways of understanding the emergence and stabilization of institutions. Both perspectives can provide correctives to the excesses of the other. Stories focusing only on highly contingent strategic interactions in specific countries are almost certainly missing some of the "patterns that are hidden and deep" (Lichbach, this volume, page). Institutions reflect the interests of powerful groups in society at the time they were adopted, and institutional reform is likely constrained thereafter by those same groups, as well by a different constellation of interests generated anew by the institutions—e.g.

incumbents, judges, civil servants—as well as other groups that emerge decades after institutions have been adopted. Yet at this point even the best structural-historical theories are extremely simplistic, one-shot stories that are badly in need of micro-foundations. Theoretical and empirical work on endogenous institutions will continue to be at the top of the research agenda in comparative politics for some time.

Finally, are comparative institutional researchers making any progress in solving identification problems? Perhaps the most believable studies are those that unearth natural experiments or are able to craft randomized field experiments. But for the lion's share of observational data on which the field relies, there are no econometric quick-fixes. Like qualitative analytical historical work, all involve fairly strong assumptions and assertions about what is ultimately exogenous. Instead of viewing history as a tangled bundle of "unobservables," the way forward is to delve into theory-guided exploration of the past and grapple with complex questions for which completely satisfying answers will always be elusive.

It is easy to criticize the work of almost any empirical researcher by spinning a story about endogeneity. What seems like a trivial critique to one scholar appears devastating to another, and consensus about these matters is nowhere in sight. Given that natural and field experiments can seem rare, expensive, and unsuitable to the research questions of so many scholars, it is tempting to give up on the whole enterprise and go back to the comfort of correlations and the excitement of big questions, choosing not to worry about the potential biases contaminating inferences.

But advocates of such an approach should be aware that the opposite reaction to the problems of causation and endogeneity is gaining strength—what might be called the extreme experimentalist position. In a provocative paper, Gerber, Green, and Kapler (2003) imply that observational studies should be accorded zero weight (and given zero funding) if they cannot produce believable information about the nature and extent of the bias plaguing their inferences. The idea is that it is better to have a few unbiased, "secure empirical premises" from which to build, even if these are not particularly interesting, than to "grasp for weak tests of expansive and arresting propositions" (p. 25).

An extreme interpretation of this argument is that observational research should simply be abandoned. One might even go on to argue, as statisticians sometimes do, that elaborate theories about causal mechanisms, counterfactuals, and structural models are a waste of time and resources, and political science should only concern itself with a "barefoot empiricist" notion of causality, establishing unadorned facts in the laboratory or through field experiments. The thicket of problems like unobserved heterogeneity, selection, errors-in-variables, and data-mining are simply too daunting to attack with theory and observational research. The solution is to throw up our hands and rely exclusively on randomization.

The most extreme experimentalist position is not likely to win the day in comparative politics. After all, absent some theories and a base of observational research, it is difficult to know what experiments to run or with what kinds of subjects to run them. For some research questions, the threats to external validity of experimental results will simply be too great. It is not clear how much we might learn about the impact of electoral rules on the welfare state by randomly assigning proportional electoral rules to small villages in post-conflict situations where international agencies have the necessary clout. Some of the causal mechanisms laid out above are driven by a kind of geographic heterogeneity that is not possible within a village. Moreover, if institutions are equilibria that must be self-enforcing to survive, what we learn by temporarily imposing them in a contrived way might be illusory. Experiments in comparative politics are likely to emerge as valuable complements rather than substitutes for observational studies.

Yet comparative politics ignores the experimentalist critique at its peril. If the field has nothing to offer but weak tests of expansive propositions, it will quickly become quaint and obsolete. Observational data can be used in conjunction with experimental research in efforts to build the firmest possible base of knowledge, but this requires that observational researchers be as clear as possible about the biases of their claims, and innovate in their use of theory and empirical technique to reduce them. Progress in sorting out causal claims through empirical research is frustrating and slow, and for most institutional researchers, the demonstration of causality will remain a Platonic ideal that cannot be achieved in practice. But imperfect efforts to approach it are far better than the alternatives.

References

Acemoglu, Daron. 2005. "Constitutions, Politics, and Economics: A Review Essay on Persson and Tabellini's *The Economic Effects of Constitutions*," *Journal of Economic Literature* XLIII (December): 1025-48.

Acemoglu, Daron and James Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. Cambridge and New York: Cambridge University Press.

Acemoglu, Daron, Simon Johnson, and James Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117: 1231-94.

_____. 2004. "Economic History and Political Science: Clarifying the Questions, Methods and Answers." *The Political Economist,* Winter.

Alesina, Alberto and Edward Glaeser. 2004. *Fighting Poverty in the U.S. and Europe: A World of Difference*. Oxford: Oxford University Press.

Andrews, Josephine and Robert Jackman. 2005. "Strategic Fools: Electoral Rule Choice under Extreme Uncertainty." *Electoral Studies* 24: 65-84.

Ansolabehere, Stephen and James Snyder. 2006. "Party Control of State Government and the Distribution of Public Expenditures." *Scandinavian Journal of Economics* 108, 4: 547-69.

Banerjee, Abhijit and Lakshmi Iyer. 2002. "History, Institutions and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." Unpublished paper, Department of Economics, MIT.

Benoit, Kenneth. 2004. "Models of Electoral System Change." *Electoral Studies* 23: 363-89.

_____. 2006. "Electoral Laws as Political Consequences: Explaining the Origins and Change of Electoral Institutions." Forthcoming in *Annual Review of Political Science* 10.

Blais, Andre, Agnieszka Dobrzynska, and Indridi Indridason. 2004. "The Adopt or Not to Adopt Proportional Representation: The Politics of Institutional Choice." *British Journal of Political Science* 35: 182-90.

Boix, Carles. 1999. "Setting the Rules of the Game: The Choice of Electoral Systems in Advanced Democracies." *American Political Science Review* 93, 3: 609-624.

_____. 2003. *Democracy and Redistribution*. Cambridge and New York: Cambridge University Press.

Braunias, Karl. 1932. *Das parlamentarische Wahlrecht* (Berlin: de Gruyter).

Butler, David. 1951. “Appendix.” In H. G. Nichols, ed., *The British Election of 1950*. London: Macmillan.

Butler, David and Bruce Cain. 1992. *Congressional Redistricting: Comparative and Theoretical Perspectives*. New York and Toronto: Macmillan.

Carstairs, Andrew. 1980. *A Short History of Electoral Systems in Western Europe*. London: George Allen and Unwin.

Chattopadhyay, Raghabendra and Esther Duflo. 2004. “Women ad Policy-Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica* 72, 5: 1409-43.

Cox, Gary. 1997. *Making Votes Count: Strategic Coordination in the World’s Electoral Systems*. Cambridge and New York: Cambridge University Press.

Cox, Gary and Mathew McCubbins. 1986. “Electoral Politics as a Redistributive Game.” *Journal of Politics* 48, 2: 370-89.

Crepaz, Markus. 1998. “Inclusion versus Exclusion: Political Institutions and the Welfare State.” *Comparative Politics* 31: 61-80.

Cusack, Thomas, Torben Iversen, and David Soskice. 2007. “Economic Interests and the Origins of Electoral Systems.” *American Political Science Review* 101, 3: 373-91.

Diaz-Cayeros, Alberto. 2006. *Federalism, Fiscal Authority, and Centralization in Latin America*. Cambridge and New York: Cambridge University Press.

Dixit and Londregan. 1995. “Redistributive Politics and Economic Efficiency.” *American Political Science Review* 89, 4: 856-66.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2003. “The New Comparative Economics.” *Journal of Comparative Economics* 31: 595-619.

Duverger, Maurice. 1964. *Political Parties: Their Organization and Activity in the Modern State*. London: Methuen.

Eckstein, Harry. 1964. “Introduction: Toward the Theoretical Study of Internal War.” In *Internal War: Problems and Approaches.* Edited by Harry Eckstin, 1-32. New York: Free Press.

_____. 1980. “Theoretical Approaches to Explaining Collective Political Violence.” *Handbook of Political Conflict: Theory and Research.* Edited by Ted Robert Gurr, 135-66. New York: Free Press.

Franseze, Robert. 2001. *Macroeconomic Policies of Developed Democracies*. New York and Cambridge: Cambridge University Press.

Fredriksson, Per and Daniel Millimet. 2004. "Electoral Rules and Environmental Policy." *Economics Letters* 84, 2: 237-44.

Glaeser, Edward, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Schleifer. 2004. "Do Institutions Cause Growth?" *Journal of Economic Growth* 9: 271-303.

Gudgin, Graham and Peter Taylor. 1979. *Seats, Votes, and the Spatial Organisation of Elections*. London: Pion.

Hall, Robert and Charles Jones. 1999. "Why do Some Countries Produce so Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114, 4: 83-116.

Heckman. James. 2005. "The Scientific Model of Causality." *Sociological Methodology* 35, 1: 1-98.

Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81: 945-60.

Huber Evelyne, Charles Ragin, and John Stephens. 1993. "Social Democracy, Christian Democracy, Constitutional Structure, and the Welfare State." *American Journal of Sociology* 99: 711-49.

Humphreys Macartan and Jeremy Weinstein. 2007. "Policing Politicians: Citizen Empowerment and Political Accountability in Africa." Working paper, Stanford University.

Iversen, Torben and David Soskice. 2006. "Electoral Institutions and the Politics of Coalitions: Why Some Democracies Redistribute More than Others." *American Political Science Review* 100: 165-81.

Kaminski, M. 2002. "Do Parties Benefit from Electoral Manipulation? Electoral Laws and Heresthetics in Poland, 1989-1993." *Journal of Theoretical Politics* 14, 3: 325-58.

Katzenstein, Peter. 1995. "The Role of Theory in Comparative Politics: A Symposium." *World Politics 48*: 10-15.

Katznelson, Ira. 1997. "Structure and Configuration in Comparative Politics." In Lichbach, Mark and Alan Zuckerman, eds., *Comparative Politics: Rationality, Culture, and Structure*. Cambridge and New York: Cambridge University Press.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny. 1998. "Law and Finance." *Journal of Political Economy* 106, 6: 1113-55.

Lichbach, Mark. 2007. "Thinking and Working: Discovery, Explanation, and Evidence in Comparative Politics." In this volume, XXX-XXX.

Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven, CT: Yale University Press.

Lizzeri, Alessandro and Nicola Persico. 2001. "The Provision of Public Goods under Alternative Electoral Incentives." *American Economic Review* 91, 1: 225-39.

McDonald, Michael. 2006. "Policy Consequences of SMD versus PR: A Survey." Paper prepared for the conference, *Plurality and Multi-round Elections*, University of California, Irvine, February 19-20, 2006.

McDonald, Michael, Silvia Mendes, and Ian Budge. 2004. "What Are Elections For? Conferring the Median Mandate." *British Journal of Political Science* 34: 1-26.

Milesi-Ferretti, Gian Maria, Roberto Perotti, and Massimo Rostagno. "Electoral Systems and Public Spending." *Quarterly Journal of Economics* 117, 2: 609-57.

Mill, John Stewart. 1848. *Principles of Political Economy with Some of Their Applications to Social Philosophy*. London: J.W. Parker.

Moser, R. G. 2001. *Unexpected Outcomes: Electoral Systems, Political Parties, and Representation in Russia*. Pittsburgh, PA: University of Pittsburgh Press.

Nagel, Jack. 2004. "New Zealand: Reform by (Nearly) Immaculate Design." In *Handbook of Electoral System Choice*, 530-43, edited by Josep Colomer. New York: Palgrave Macmillan.

Castles, Geoffrey, Herbert Obinger, and Stephan Leibfried, eds., *Federalism and the Welfare State: New World and European Experiences*. Cambridge and New York: Cambridge University Press.

O'Donnell, Guillermo and Philippe Schmitter. 1986. *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*. Baltimore, MD: Johns Hopkins University Press.

Olken, Ben. 2007. "Political Institutions and Local Public Goods: Evidence from a Field Experiment in Indonesia." Working paper, Harvard University.

O'Neill, Kathleen. 2003. "Decentralization as an Electoral Strategy." *Comparative Political Studies* 36, 9: 1068-91.

Persson, Torsten and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge, MA and London: MIT Press.

_____. 2003. *The Economic Effects of Constitutions*. Cambridge, MA and London: MIT Press.

Persson, Torsten, Gerard Roland, and Guido Tabellini. 2007. "Electoral Rules and Government Spending in Parliamentary Democracies." Forthcoming, *Quarterly Journal of Political Science.*

Powell, G. Bingham. 1982. *Contemporary Democracies: Participation, Stability, and Violence*. Cambridge, MA: Harvard University Press.

Przeworski, Adam. 1985. "Marxism and Rational Choice." *Politics and Society* 14: 379-409.

_____. 2004. "Economic History and Political Science." *The Political Economist*, Fall.

_____. 2007. "Is the Science of Comparative Politics Possible?" In *Oxford Handbook of Comparative Politics*, Edited by Carles Boix and Susan Stokes.

Przeworski, Adam and Fernando Limongi. 1997. "Modernization: Theories and Facts." *World Politics* 49, 2: 155-183.

Remington, T. F. and S. Smith. 1996. "Political Goals, Institutional Context, and the Choice of an Electoral System: The Russian Parliamentary Election Law." *American Journal of Political Science* 40: 1253-79.

Rodden, Jonathan. 2006. "Political Geography and Electoral Rules: Why Single-Member Districts are Bad for the Left." Unpublished paper, Stanford University.

Rodden, Jonathan. 2007. "Federalism." In *Oxford Handbook of Political Economy*. Edited by Barry Weingast and Donald Wittman. Oxford: Oxford University Press.

Rogowski, Ronald and Mark Kayser. 2002. "Majoritarian Electoral Systems and Consumer Prices: Price-Level Evidence from the OECD Countries." *American Journal of Political Science* 46: 526-39.

Rokkan, Stein. 1970. *Citizens, Elections, Parties: Approaches to the Comparative Study of the Process of Development.* Oslo: Universitetsforlaget.

Shvetsova, Olga. 2003. "Endogenous Selection of Institutions and Their Exogenous Effects." *Constitutional Political Economy* 14: 191-212.

Taagapera, Rein. 2002. "Designing Electoral Rules and Waiting for an Electoral System to Evolve." In *The Architecture of Democracy: Institutional Design, Conflict Management, and Democracy in the Late Twentieth Century*, edited by A. Reynolds. Oxford: Oxford University Press.

Ticchi, Davide and Andrea Vindigni. 2005. "Endogenous Constitutions." Unpublished paper, Princeton University.

Verney, Douglas. 1957. *Parliamentary Reform in Sweden, 1866-1921*. Oxford: Oxford University Press.

Zuckerman, Alan. 2007. "Advancing Explanation in Comparative Politics: Social Mechanisms, Endogenous Processes, and Empirical Rigor." In this volume.