

1 Research methods and democratization

Why are some countries democracies while others are not? Why do some democracies survive while others fail? These are core questions in political science, sociology, history, and economics. Scholars have been trying to answer them ever since Aristotle. After 2,500 years of research, what do we know? How do we know it? This book answers the latter two questions.

Rationale

After so much has been written on these matters, I am obliged to explain why we need this particular book on democracy. The answer is simple: to distill this vast literature down to a comprehensible critical survey! Democratization has been studied for so long and in so many different ways that the literature is overwhelming and bewildering. It encompasses ideas borrowed from the study of early modern Europe, strategic bargaining games, Asian cultures, measurement theory, public opinion about economic trends, legacies of colonialism, geographic networks, theories of class struggle, demographic trends, war and peace, and many other phenomena. There are dozens of hypotheses about the causes of democracy (Chapter 4 summarizes fifty-five of them). The first step toward digesting this mountain of scholarship is taking a comprehensive inventory of all the possible explanations.

This book is not just about democratization, however; it is equally devoted to evaluating research methods in comparative politics. This is another area in which it might seem that there is already a superabundance of reading material (Brady and Collier 2004; Geddes 2003; George and Bennett 2005; Gerring 2001; Goertz 2006; Green and Shapiro 1994; King et al. 1994; Landman 2008; Lave and March 1993; Lichbach and Zuckerman 2009; Peters 1998; Przeworski and Teune 1970; Ragin 1987, 2000; Shively 1998). The past fifteen years have been especially embroiled in debates about political science methodology. There is a growing feeling, however, that in the

latter part of this decade, the polarization is abating in favor of a hopeful methodological pluralism. I believe that the time is right to articulate a unified vision of research methods that recognizes the strengths and weaknesses of each approach with respect to the shared fundamental aims of all social science research.¹ This book proposes such a vision and uses its standards to evaluate three basic approaches in comparative politics: case studies and comparative history, formal models and theories, and quantitative testing.

The principal virtue of this book, however, is not its grouping of research methods into a single framework. There are many alternatives that offer their own classification schemes and others that go much deeper into each method. Neither is its virtue the breadth or depth of its survey of democratization research; there are worthy alternatives for this purpose as well (Diamond 1999; Geddes 2007; Haerpfer et al. 2009; Teorell 2009; Whitehead 2002). Rather, the uniqueness of this book is that it combines a survey of democratization research with a critique of research methods.

These two tasks can be better accomplished together than separately, for two reasons. First, it is often difficult to appreciate the implications of research methods until they are applied to a substantive theme. Examples of actual research demonstrate how choices about research methods can decisively affect substantive conclusions. Democratization is the ideal application for this purpose. Because it has been studied for decades, if not centuries, and studied with almost every method employed in comparative politics, it can supply examples that illustrate and illuminate every conceivable approach in comparative politics. It also has the advantage of being interesting to most people who do research in comparative politics.

Second, a survey of this bewildering literature demands explicit attention to methodology. Precisely because such diverse methods have been used, the literature is a hodgepodge of disparate findings that are difficult to integrate into a comprehensive summary of what we know. Students of democratization have employed class analysis, structural functionalism, case studies, conceptual analysis, game theory, survey research, advanced statistical analysis, and the occasional lab or field experiment. How can we extract any meaningful conclusions from such disparate types of knowledge?

¹ I use the term *approach* loosely and sometimes interchangeably with *method* or even *school of thought*. When a distinction matters, however, I think of an approach as a path to a goal or a means to an end. In this context, methods are the means and the ends are the goals of a research project or program, whether they are as narrow as explaining a particular coup or as broad as integrating models into a theory or testing the general truth of hypotheses.

The key that makes a comprehensive assessment possible is attention to the strengths and weaknesses of different methodological approaches relative to absolute standards for good research. If the assessments in this book sometimes sound harsh, it is because the book uses standards of evaluation that are very high – close to perfection. To evaluate research in other contexts, scholars use *relative* (movable) standards, which are appropriate for those contexts. To judge which articles deserve to be published, we judge them relative to other publications; to judge whether colleagues deserve tenure or promotion, we judge them relative to other scholars at an equivalent stage of their careers; to judge whether to be impressed by a research project, we judge what it accomplished relative to the difficulty of carrying it out. But if we need to judge how much we have explained about democratization, we must judge our explanations by absolute standards for what a true, complete, and certain explanation would be. Relative standards can tell us whether we are making progress, but only an absolute standard can tell us how far from perfection we are and what we must improve to get there.

Although democratization research has made great progress, it still falls far short of perfection. Too often incautious students and scholars develop exaggerated confidence in the conclusions of published research, and authors, understandably enough, tend to downplay the limitations of their own findings. This leaves the impression that our knowledge is more complete and certain than it really is. A careful, comprehensive methodological critique of this literature encourages a more mature and modest appreciation of what we know about democratization and how well we know it. Explicit acknowledgment of what a method is supposed to accomplish creates a standard for evaluating how well it works. Having a unified, comprehensive set of standards makes it possible also to evaluate methods against one another, thus revealing the kinds of insights each one reveals and the kinds it conceals. This book uses these standards in the final chapter to lay out an agenda for future research on democratization. Thus, readers of this book can expect to come away with balanced judgment about what each method can teach us, a comprehensive synthesis of what is known about democratization, and a sober appreciation of the limitations of that knowledge.

Although democratization research is probably the best application for this critique of research methods, it has a few limitations. First, it does not lend itself to experiments. It would be neither feasible nor ethical to split countries or populations into treatment and control groups to examine the impact of economic development, religious traditions, or other supposed causes of democracy (although various split or merged states may provide

opportunities for “natural experiments” – such as Germany, Czechoslovakia, Yemen, Vietnam, and Korea).² Therefore, almost all the research discussed in this book is based on observational data, and my arguments about the challenges to inference are correspondingly more cautionary than they would be about experimental data.

Second, because democracy is a feature of national states, most of the research reviewed here pertains to the national level of analysis. Research on subnational phenomena – movements, parties, classes, regions, individuals – would encounter some different issues. Chapters 4, 5, 6, and especially Chapter 8 do touch on some of these issues, but I have not attempted to address them systematically.

Third, formal theorizing about democratization is relatively recent. As I note in Chapter 6, it is too early to draw conclusions about the long-term potential of this approach for understanding democratization.

Fourth, most research on democratization, including the quantitative research, performs frequentist hypothesis testing, which tells us how likely it is that a hypothesis is true, given the evidence. There is growing support in the social sciences for the Bayesian approach, which seeks conclusions about how evidence changes our prior beliefs about hypotheses.

Finally, as noted in Chapter 4, there is too little research on democratization using Boolean or fuzzy-set Boolean methods to judge how well it works in practice. In all of these respects, however, the limitations of democratization research are also found in research on other topics, from political economy to institutions to ethnic conflict to state building. Democratization research is more the rule than the exception.

Overview

I ground all of my judgments in an exacting philosophy of the social sciences. I start from a commonsense belief that politics is extremely complex. It is complex in that many forces are in play, constantly interacting, and varying by time and place; and often unique events have powerful consequences, further undermining our ability to build useful theories. All social sciences must grapple with this complexity, but the subfield that must grapple with it most

² Susan Hyde’s field experiment on the impact of election observation is one example of experimental research that helps understand an aspect of democratization (Hyde 2007).

fully is comparative politics, which takes the entire world and all of human history as its domain. No parsimonious theory can possibly encompass all of the relevant phenomena or explain any piece of it completely. For this reason, any student of comparative politics must learn to be comfortable with theories that are probabilistic, partial, conditional, and provisional. Our theories are necessarily probabilistic because there are always exceptions, due to unknown causes that we cannot model systematically. Our theories are partial because no political outcome worth explaining has just one cause: many causes help explain it, and none determines it alone. Our theories are conditional because causal relationships often vary according to time, place, and other conditions. Finally, comparative theories are provisional, because we have no choice but to build on probable, partial, conditional knowledge that will change when a better theory comes along.

My stance could be called sadder-but-wiser positivism. I believe in the scientific method for understanding society and politics, but I also understand the challenges it faces. I am not a naive positivist who believes that everything is reducible to a set of eternal equations whose truth can be objectively demonstrated and replicated. Our reality is complex and ever changing, and it appears very different to different observers. As social scientists, we will never understand any phenomenon fully, and we will always have our own subjective biases. I can understand why some in my profession would conclude that scientific methods are inappropriate for understanding politics and society and that the search for theory is futile. Nevertheless, I am convinced that rather than giving up, we must do the best we can; that rather than giving in to our biases, we must struggle against them. We cannot explain everything we observe, but we can hope to explain big pieces of it; we cannot perceive the world objectively, but we can test our perceptions against others' and against evidence, and adjust them. In the process, we will create social and political science that responds to antipositivist critiques of naive positivism by developing more nuanced concepts and theories. In my view, the most useful corrective for poor science is better science.

In comparative politics, the fundamental division of labor is between what I call thick and thin approaches. Thick approaches entertain many intricately intertwined causes, they seek to explain multifaceted outcomes, and they rely on elaborate theoretical assumptions. Anthropological, interpretive, thick description is the extreme case of such an approach; in comparative politics, it approximated by case studies, area studies, comparative history, and some approaches to understanding political culture. At the other extreme,

thin approaches consist of simple, theoretically neutral propositions considered in isolation from the surrounding complex reality. Good examples are econometric analyses of the impact of per capita gross domestic product or presidential constitutions on the age of democratic regimes. Thick approaches lend themselves to rich understanding of specific events; thin approaches lend themselves to hypothesis testing and generalization. In principle, we can strive to combine the best of both approaches by thickening thin concepts and theories (Coppedge 1999). In practice, however, this is difficult, so there is a trade-off. I illustrate this trade-off in Chapter 2 by demonstrating the advantages and disadvantages of thick definitions of democracy and thin quantitative indicators of democracy. I argue that existing quantitative indicators are adequate for some purposes but that we cannot measure democracy much better until we thicken the concept that we are operationalizing to take multiple dimensions of democracy into account (Coppedge 2002). Recognizing the distinct dimensions of democracy would also help us measure each one more accurately.

The three subsequent chapters focus on theory building. I argue in Chapter 3 that, ideally, theory is thick, general, and integrated. Thick theory is rich, descriptively accurate, and sensitive to local and historical variation in concepts and causal relationships. General theory applies to as many times and places as possible; it approaches universality. A truly universal theory of politics does not exist and, if it did, it would not be a set of simple propositions. Rather, a general theory would be a set of interlocking middle-range theories knit together by more general propositions that identify the circumstances in which each middle-range theory is relevant. Integrated propositions are clear, logical, formal, and systematic, which aids the accumulation of theoretical knowledge and creates a fecund mechanism for generating hypotheses. We have no theory that possesses all three qualities; these are merely characteristics of an ideal theory that would explain practically everything in politics well. Although this is an unattainable goal, it points us toward better theory. I then argue that there are three major sources of theory in comparative politics: case studies, large-sample comparisons, and formal theories. However, each source supplies a kind of theory that meets one of these three criteria well but the other two only poorly. Case studies suggest theory that is thick but not general or integrated. Large-sample comparisons, although we tend to regard them as tests rather than sources of theory, do supply a kind of theory: empirically confirmed generalizations that are thin and undertheorized. Formal theory is clear, logical, and systematic but neither thick nor truly general.

The following three chapters illustrate these arguments. Chapter 4 compiles a master checklist of proposed causes of democratization. It defines *democratization* generously to include any process in which countries become democracies or not, become more democratic or less so, survive as democracies, or break down. There are many conventional wisdoms (the plural is deliberate) about the factors that prepare countries for democracy, spur transitions, and help democracies survive. Explanations have suggested dozens of causes related to mass political culture, leadership, the economy, society, the state, various political institutions, and the international system. If we, foxlike, took all of these suggested factors into account, we could make surprisingly reliable predictions about which countries are democracies and which are not, but we would not be able to predict the intermediate or mixed cases well, and we would not be able to say which factors mattered, how much they mattered, why they mattered, or in which combinations. These limitations of checklists justify political scientists' emphasis on developing theory and testing hypotheses. Chapter 4 also describes and evaluates the Linz (1978) breakdowns framework; the O'Donnell and Schmitter (1986) transitions framework; and the Linz and Stepan (1996) consolidation framework, which take steps toward integrated theory. This chapter concludes with a brief critique of Boolean analysis, which can be used to test checklist-type arguments (Ragin 1987).

Chapter 5 begins by contrasting histories and case studies. It praises case studies as the thickest method for generating and intensively testing competing explanations of specific events but notes that the voluminous knowledge they produce is difficult to integrate into general theory and tends to overemphasize dynamic, idiosyncratic factors and underemphasize static, structural factors. Most of the chapter critiques comparative histories, which are intended to represent a compromise between case studies and large-sample comparisons. The result, I argue, is an approach that is probably the best way to generate more general explanations that are likely to survive testing but also one of the worst ways to test them. Examples include books by Moore (1966); Skocpol (1979); Rueschemeyer, Stephens, and Stephens (1992); Ruth and David Collier (1991); Luebbert (1991); Downing (1992); and Ertman (1997).

Chapter 6 surveys formal models and theories of political liberalization, transition, and survival. It distinguishes positional models, in which regimes are ends in themselves, from economic models, in which regimes are a means to economic ends. This approach has the potential to generate a large, logically

connected body of well-integrated theory. It has, however, integrated these models far less than one would expect, although the economic models are much more integrated than the positional ones. Nevertheless, these models have placed new puzzles on the research agenda. Why would authoritarian rulers ever ease repression if this might snowball into a loss of power to the democratic opposition? Why would the opposition ever take the risk of challenging a dictator? How can the rich assure the poor that democracy will be preserved, and how can the poor assure the rich that a democratic government would not confiscate their assets? The downside of this approach is that it works with extremely thin concepts, and its predictions are largely untested, and perhaps untestable. The chapter discusses models developed by Przeworski (1986, 1991), Marks (1992), Colomer (2000), Boix (2003), Acemoglu and Robinson (2001, 2006), and others.

Theories should meet the requirements laid out in Chapter 3, but they must also survive rigorous testing. Chapter 7 proposes comprehensive criteria for rigorous testing and uses them to explain why rigorous testing of theories of democratization is so challenging. The core principle is that we can “prove” causal relationships only indirectly, by disconfirming any alternative hypothesis that other scholars consider plausible. Unfortunately, democratization is presumably so complex that the number of alternative hypotheses is huge. The chapter develops a fundamental distinction between extensive testing, which evaluates whether we can generalize propositions about average causal effects, and intensive testing, which evaluates which of several alternative stories best explains a specific outcome in a single case. Ideally, our theories should survive both extensive and intensive testing.

Chapters 8 and 9 survey the results of testing, mostly extensive. (Because case-study conclusions cannot be summarized without taking them out of context, this book largely ignores the findings of their intensive tests.) Chapter 8 evaluates research on political culture and democratization. It first explains why the method of studying culture that is derived from psychology – survey research – is more useful for developing general theory than methods derived from anthropology. Survey research findings are riddled with paradoxes. Support for democracy as the best form of government remains high even though citizens in the West have lost trust in democratic institutions; in fact, critical citizens tend to support democracy more than their complacent peers (Norris 1999a). Average levels of support for democracy have almost no association with how democratic a country is or how long democracy has survived in a country. And there are striking cross-national differences in certain syndromes, such as institutional trust and belief in self-expression, that

do not exist at the level of individual attitudes (Davis and Davenport 1999; Inglehart and Welzel 2005; Muller and Seligson 1994). A major difficulty in this area is the lack of a theory that could link individual-level attitudes to both behavior and system-level outcomes such as democracy.

Chapter 9 surveys quantitative testing of democratization hypotheses, which has accelerated and deepened in technical sophistication in the past two decades as democracy data have become more abundant. Because I value confirmed generalizations more than untested theories, I tend to be more positive about the large-sample statistical approach. However, it is no exception to my central claim that each approach has one strength and two weaknesses, as each tends to be general but thin and poorly integrated. Quantitative research has unearthed dozens of general empirical regularities, but because the concepts and theories tend to be thin, the findings tend to be consistent with many possible theoretical interpretations. The most robust discoveries, such as geographic diffusion and the logarithmic association between per capita gross domestic product and levels of democracy, have been followed by lengthy debates about what the causal mechanisms (if any) might be. Quantitative researchers tend to crunch first and ask questions later. It has become clear that there are striking and stable cross-national differences in levels of democracy that constitute much of the variance to be explained; identifying their causes will probably have to be a task for case studies and comparative histories. Over the decades, this approach has proliferated new dependent variables that force us to be more precise in our definitions of democratization. It has also given us a detailed and varied empirical basis for judging which explanations are intuitively plausible and which are not.

The concluding chapter, Chapter 10, pulls all the preceding analyses together in an agenda for future research on democratization and for future methodological development. Each approach could, of course, continue to do what it does best and strive to do it better. Case studies and comparative histories can become ever more nuanced, formal models could integrate themselves into a more coherent body of theory, survey research could cover more countries and more questions more frequently, and quantitative testing could bring ever more sophisticated data and estimation techniques to bear on democratization hypotheses. The most fruitful advances, however, will come from efforts to transcend the typical limitations of each approach. Those who immerse themselves in cases could do more to contribute to, and be guided by, broader theoretical debates. Those who develop formal models could try to make them more realistic and testable. And those who do

statistical testing could rely on theory to resist the curve-fitting temptation and collaborate with area specialists to develop thicker indicators of key concepts such as democracy. All of these tasks are difficult, but I believe that our understanding of democracy will deepen only to the degree that we accomplish them.