

“intensive testing,” which is ideal for case studies, but it has a very different purpose and logic of inference. The goal of intensive testing is to judge which of several competing hypotheses does the best job of explaining a single case. It is therefore what Gerring discusses in the epilogue as “single-outcome studies,” and here and there as “internal validation,” but it doesn’t get the emphasis it deserves, because it constitutes at least half of the justification for doing case studies. Unlike extensive testing, which tests the same propositions in a large number of cases, intensive testing tests a large number of propositions in a single case. The logic is, “if my theory is true, then I would expect to observe these 20 things in this case. If the alternative theory is true, then I would expect to observe these 20 different things. Using Bayesian logic, if the 20 predictions of my theory are confirmed and the 20 alternative predictions are not, there is only a very low probability that my theory is wrong, and it becomes the better explanation for this case.” It is usually impossible to quantify these probabilities, but the logic behind them is very strong, and it makes case studies a very powerful method for explaining single outcomes.

This different emphasis would alter a few of the book’s passages. For example, I endorse Gerring’s conclusion on p. 147 that “Case studies...rest upon an assumed synecdoche: the case should stand for the population. If this is not true, or if there is reason to doubt this assumption, then the utility of the case study is brought severely into question.” I think there are *always* reasons to doubt this assumption, so it is *never* safe to generalize from one or a few cases. That’s why we should use them for theory development and intensive testing rather than for any attempt at extensive testing.

Another example: In his interesting discussion of matching as a promising alternative to specifying control variables in a regression, Gerring states that simply asserting that two cases are more or less the same for the purpose of matching “can be a huge advantage over large-N cross-case methods, where each case must be assigned a specific score on all relevant control variables—often a highly questionable procedure, and one that must impose strong assumptions about the shape of the underlying causal relationship.” (133–34). Yet it is always possible to specify at least a subjective dummy variable as a control, which would be exactly as accurate as asserting that two cases match, and it is often possible to assign more precise scores for regression variables. If assumptions about the linearity of a relationship are false, they can be modified and tested. I come away convinced that matching, which Gerring explains very clearly, is a method worth trying, but I suspect, as I think he does, that it will not be as useful in practice as it sounds in principle.

A final example concerns scope conditions. I love Gerring’s call in chapter 4 (76–85) for making scope conditions explicit and non-arbitrary; this is essential. But its implications are ambiguous unless we make it clear what the scope conditions demarcate. If it is *tested* propositions, there is little room for arbitrariness: the scope of tested propositions is exactly as large as the sample or the case used in the test; we can’t generalize beyond it, unless it was a random sample of sufficient

size, in which case we can generalize to the population. But if we are talking about how far a *hunch* might travel, then the scope of the hypothesis is hypothetical. It is essential to speculate about what the scope conditions may be, but we won’t really know until some extensive testing is done.

I also have one question that is unrelated to any of this. In chapter 6 (with Rose McDermott), which makes a beautiful, concise argument that an experimental logic undergirds all case studies, the most rigorous category, “Dynamic Comparison,” is defined as having both spatial and temporal variation. I wonder whether cross-sectional time-series analysis meets this criterion.

In conclusion, I think that in reality I agree with Gerring on just about everything and he agrees with me. I have quoted some passages in which he seems to have an opinion different from mine, but they are balanced by other passages that sound very close to what I have said on these issues. If we have differences, I believe they are only differences of emphasis.

Moving the Doormat to the Main Menu: Case Study Research Methods in the Social Science Toolkit

<https://doi.org/10.5281/zenodo.997308>

Evan S. Lieberman
Princeton University
esl@princeton.edu

John Gerring’s motivation for his book, *Case Study Research*, is the same as Harry Eckstein’s writing on the same subject three decades ago: He points out that case studies are much maligned—the methodological doormat if you will—despite their recurrence in so many influential works in our field and throughout the social sciences. To address this conundrum, Gerring hopes to “restore a sense of meaning, purpose and integrity to the case study method” (66).

And I think he largely does just that. He gives scholars the *potential* to do case studies in such a way that any social scientist could clearly see the logic through which the analysis could generate strong causal inferences.

It is a vital and lucid work that ought to appear on any graduate research methods syllabus. As much as it is a book about case studies, it is a treatise on research design and logical thinking that updates and integrates many classic and more recent contributions.¹ The book keeps its feet on the ground by examining a rich array of examples of completed work in political science, often with a healthy dose of pragmatism.

In my comments, I will highlight some of the novel insights found within various chapters in the book, and also raise some issues that I think warrant some additional attention, either by Gerring, today, or by him or other scholars in the future.

Definitional Issues

First is the question of defining the case study. If the quest is to dignify case studies, then it is necessary to know

what we are dignifying. There is a lot of ambiguity in the conventional use of the term, and in Gerring-style, we are afforded a careful, well-thought-out definitional discussion in chapter 2.

He writes that the case study is “the intensive study of a single case where the purpose of that study is—at least in part—to shed light on a larger class of cases (a population)... at the point where the emphasis of a study shifts from the individual case to a sample of cases, we shall say that a study is *cross-case*” (20).

By definitional fiat, he declares that case studies should be theoretically oriented, and not purely idiographic. Of course, others, particularly in other disciplines, may use the label “case study” for other purposes, but he makes clear that this is an enterprise that is social scientific. If one cites this book when doing a case study, this is a clear signal: the case study will be used to explore, illuminate, probe, or test broader propositions about specified causes and effects, even as it uses proper names and particularities in the discussion and analysis.

What Can Case Studies Do? How Are They Used?

In Chapter 3, Gerring puts his best foot forward and highlights the opportunities and comparative advantages of this method relative to other methods. In attempting a similar task, Eckstein may have done more harm than good for the method by *over-selling* the inferential possibilities with case studies. Gerring both tempers and expands the use of case studies by customizing the opportunities for case study work within the research cycle, and according to the type of theory one is dealing with, particularly whether the theory is of a more deterministic or probabilistic variety.

In many ways, Gerring’s book shares a vision of what constitutes convincing evidence that is *more* similar to King, Keohane, and Verba’s *Designing Social Inquiry* than many other recent works by qualitative methodologists. While it is true that Gerring has written extensively on the subject, he is clearly *not* arguing that qualitative research or case studies are in any way the “superior” empirical strategies. He is measured about what he promotes. As compared with *Designing Social Inquiry*, however, one does not finish *this* book thinking, “well, really the best thing to do is to avoid case studies, and to redesign a social inquiry such that it will produce many observations leading to a dataset amenable to quantitative analysis.” Gerring stresses that there are often key observations that shed disproportionate light on the connections between causes and effects, and that these must frequently be combined *across* different types of units in order to understand political dynamics. However, these connections are not made within the context of conventional datasets, which, by contrast, *require* unit homogeneity.

In this sense, the book really champions the idea of a larger division of labor—one that might be integrated within studies, or across studies, which is a view that many give lip service to but this book really defends in logical terms. A handful of key interviews or historical records may reveal the plausible links between cause and effect that complement correlational research of different types of observations such

that we can be convinced about the validity of a particular proposition. Brady and Collier refer to this as causal process observations, and Gerring also identifies a complementarity with dataset observations. In any case, this is a book that could be called “Mixed-Method” research as much as it is “Case Study Research.”

Issues of Case Selection

The largest chapter of the book is chapter 5: “Techniques for Choosing Cases.” I think any researcher on the verge of doing case study research ought to go through this list of nine selection strategies and make sure they can self-consciously recognize one or more of these types as characterizing their own, and if not, that should indicate that the study could probably be re-framed in a more crisp manner.

However, I do have some questions to raise. First is the notion of the “pathway” case, as a separate type of case study. A pathway case, according to the author, is one in which the causal effect of one factor can be isolated from other potentially confounding factors. As a pragmatic concern, I don’t understand how one can choose a particular case for its potential to illuminate causal mechanisms *prior* to having done the research. And how is this different from some of the other strategies, such as the “crucial case” strategy or even the “typical” case strategy, which are also used for hypothesis *testing* given particular scores on independent and dependent variables? I suppose the identification of this strategy as just one of nine makes me wonder what we are doing in the other hypothesis-testing case study strategies.

I really appreciate the attention to deliberate case selection, and I know Gerring realizes that many case studies are done with less pre-study consideration than he recommends. But I do think more needs to be said about the role of area expertise and personal experience. Traditionally, a clear justification for the choice of cases has been that some investigators develop internal databases of contextual knowledge and measurement skills that increase the reliability and validity of the study. If true, perhaps this strategy deserves its own place—it is probably the most practiced strategy in any case, and it is hard to imagine this changing.

Relatedly, there are the examples of what I might call the “convenient” case study. This is akin to available non-probability sampling, but it is also the reality of being a social researcher. Sometimes case studies find us, we don’t look for them. We may be working or traveling somewhere or reading something, and our interest is piqued, perhaps because what we observe confirms or contradicts some prevailing theory. One might respond, “well, then you did not use a case ‘selection’ strategy *per se*.” Perhaps, but I am going to guess that a sizeable portion of case studies are generated this way—they are the product of life circumstances and interests, personal and professional. Thus, it may be hyperbole to speak always of case *selection* when we talk about case studies. What we might be able to do is to engage in *case justification* and use the strategies that Gerring identifies to describe what it is we have a case of.

Also, I am not sure where we would put Michael Burawoy’s

“extended case study” into this typology of case studies. In various works, Burawoy has sought to use extended and deep case studies to discover flaws in and then modify existing case studies. While it is true that in some sense the goal is slightly different from the stated one of Gerring’s book, which is to make inferences about a larger population of cases, it does share an orientation toward theoretical conclusions. But again, for Burawoy, the selection of the particular case tends to be based more on the prospects for depth of study. This approach is intended to highlight when variables were simply conceptualized poorly or relationships misunderstood. It strikes me that by depriving this type of case a real label, it is devalued in principle, while in practice, it provides potentially enormous contributions to knowledge. But to my knowledge one could select such a case without any prior knowledge of the case’s place in the distribution of explanatory or outcome variables.

Notwithstanding, the chapter does an excellent job of bringing together the diverse forms of case selection strategies that have been used by scholars, including the logic of inference associated with Mill’s methods, and integrating them within a single, comprehensible framework.

Process Tracing

In chapter 7, Gerring argues that process tracing is *usually* a component of case study research; it usually relies heavily on contextual evidence. He states, “the hallmark of process tracing, in my view, is that multiple types of evidence are employed for the verification of a single inference—bits and pieces of evidence that embody different units of analysis...individual observations are therefore non-comparable” (173). I like this characterization. But I wonder—and I am really putting this out here for discussion—should the properties identified for *process tracing* be *definitional* for what we mean by case studies in political science? Should we reserve the term “case study” for those studies that employ such heterogeneous evidence? Otherwise, one might call, for example, every single lab experiment a case study.

The chapter on process tracing highlights clearly some of the strong empirical findings using this approach. He makes a nice recommendation—that we ought to be able to graphically diagram an argument in a series of steps, if even in the somewhat frighteningly complex manner that Mahoney does in the case of Skocpol’s *States and Social Revolutions*. I agree that diagramming is a good heuristic technique and probably ought to be used as a standard.

But the chapter on process tracing is short. A mere 13 pages, one-fifth the page-matter afforded the chapter on case selection. I don’t want to overinterpret, but this brevity, I think, reveals some of Gerring’s own apprehension with the case study method as it is generally practiced or understood in political science. Gerring writes, “process tracing evidence is, almost by definition, difficult to verify, for it extends to evidence that is nonexperimental and cannot be analyzed in a sample-based format...” (184). He says that the mitigating factors for process tracing are that it is (1) supplemental; and (2) can be vetted by “experts.” He concludes, “despite its apparently mysterious qualities, process tracing has an im-

portant role to play in case-based social science...it deserves an honored place in the toolkit of social science” (185).

Well, I detect a touch of inner conflict in that last sentence. And I think much more needs to be said about process tracing. Doing case studies well is doing this kind of analytical detective work. It is hard. I don’t know if we can develop general rules and strategies, but I think we can try, and the George and Bennett volume on case studies offers some additional discussion of process tracing.

My own suggestion for advancing the technique of process tracing is to identify more tailored sets of guidelines according to theoretical content and the level of analysis under investigation, whether it be the mobilization of collective actors, the making of policy, or the development of institutions. Political scientists would benefit a great deal by breaking down a set of criteria which they believe would establish a reasonable baseline for convincing or at least acceptable process-tracing evidence, including, say, temporal proximity of links in a causal chain; the explicitness of actor intentions; and/or the types of sources used. These benchmarks would not be ironclad rules, but might provide some standard for how we could evaluate the robustness of a qualitative result, just as conventional statistical analysis has measures of statistical significance. Just as a 95 percent confidence interval is arbitrary, so would these standards be, but without them, we have no reference line for discussing the content of evidence, except for completely useless metrics like number of months spent in the field.

Single-Outcome Studies

Finally, there is a concluding chapter on what Gerring calls, “Single-Outcome Studies.” A single-outcome study is when a researcher seeks to explain a single outcome for a single case. This is an incredibly important chapter—it takes on the elephant in the room of much social science research: that research agendas inevitably get driven by real-world, often catastrophic, events, such as the 9/11 terrorist attacks. I offer Gerring the acronym SOS for single-outcome study.

But I must admit, I leave this chapter a bit confused. Are SOS’es case studies, or not? Oftentimes the language of the chapter *contrasts* the SOS with the case study, but in describing the studies associated with nested analysis (e.g., Lieberman 2005) as types of single-outcome studies, well, that seems to me to clearly meet the criteria for a general case study. Indeed, the very terms used to describe those studies were “model-building” and “model-testing,” suggesting a direct engagement with the types of cross-case claims identified as central to Gerring’s ideal of the case study method. More generally, it is hard to imagine a political scientist studying a “single outcome” without some view of a larger universe of cases, so I require greater elaboration of what this type of study actually is, and how it is distinct from others.

Descriptive Inference

Finally, I want to discuss, if briefly, the explicit omission from the book, which is the task of descriptive inference. It is important to recognize that in this book, *most* of the science or

method of case study research comes from the strategies of case selection.

At the outset of his book, Gerring highlights that he focuses on *causal inference* because treatment of the “descriptive task of gathering evidence is well covered by other authors” (9). Actually, I disagree. I think that political scientists have very, very few good references on gathering appropriate data, particularly for the type of enterprise Gerring describes. The collection of observations that come from heterogeneous sources and unit of actors remains fairly ad hoc, and the task of summarizing accounts has received scanty treatment in political science.

Imagine a hypothetical study of ethnic conflict: someone is doing a case study in county X, and a survey reveals no hostile negative attitudes, but five in-depth insider accounts relate a mix of characterizations, and there is a riot in one province in which various ethnic slurs were shouted. As far as I know, the methodological literature tells us very little about how to score this case. And yet, this is the ever-present dilemma for the case study researcher working with multiple sources of data.

I don’t blame Gerring for omitting this type of discussion from the book because it is really a different kettle of fish, but I did want to highlight my belief that the integrity of case study research will rest on principles of descriptive inference at least as much as on principles of causal inference. More generally, I think that the discipline has devalued case studies for the very reason that we have emphasized the value of causal inference to a much greater extent than good measurement and descriptive inference, even though we know you can’t do the former without the latter.

Going Forward

To conclude, *Case Study Research: Principles and Practices* is at its very best in relating the possibilities for case study work in logical juxtaposition to other types of inferential strategies. It is smart, and provides sturdy analytical scaffolding for the development of new case studies. I think it should encourage us to do more case study work with our heads held high. But we will still need to be explicit and self-conscious about how those studies get done so that we can have an even better handle on what it is about intensive study of a case that convinces us of the strength of a general proposition. Gerring has made a major contribution to social science by helping to systematize this genre of research.

Note

¹ Such as Campbell and Stanley (1966) on quasi-experimental research design; Eckstein (1975) on case studies; Collier’s (1991) and Sartori’s (1970) statements on the comparative method and the relationship to statistical methods; and much of the more recent qualitative methods research carried out by other scholars such as Mahoney (1999); George and Bennett (2004); Brady and Collier (2004); and King, Keohane, and Verba (1994).

References

- Brady, Henry E., and David Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowman & Littlefield.
- Burawoy, Michael. 1991. *Ethnography Unbound: Power and Resistance in the Modern Metropolis*. Berkeley: University of California Press.
- Campbell, Donald T. and Julian C. Stanley. 1966. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin Co.
- Collier, David. 1991. “The Comparative Method: Two Decades of Change.” In *Comparative Political Dynamics*. Dankart A. Rustow and Kenneth P. Erickson, eds. New York: Harper Collins.
- Eckstein, Harry. 1975. “Case Study and Theory in Political Science.” In *Handbook of Political Science*. Fred I. Greenstein and Nelson W. Polsby, eds. Reading, MA: Addison-Wesley.
- George, Alexander L. and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.
- Gerring, John. 2007. *Case Study Research: Principles and Practices*. New York: Cambridge.
- King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.
- Lieberman, Evan. 2005. “Nested Analysis as a Mixed-Method Strategy for Comparative Research.” *American Political Science Review* 99:3, 435–52.
- Sartori, Giovanni. 1970. “Concept Misformation in Comparative Politics.” *American Political Science Review* 64:4, 1033–53.

Particularizing Case Studies: A Critical Review of Gerring’s Case Study Research

James Mahoney

Northwestern University

james-mahoney@northwestern.edu

A large methodological literature addresses the topic of case studies. But much of this work focuses on issues pertaining to data collection, including techniques of data retrieval (e.g., ethnography, interviews), coding, and recording. By contrast, John Gerring’s stimulating new book, *Case Study Research: Principles and Practices*, considers the logic of case study research design. Gerring seeks to explicate the meaning, purposes, and payoffs of the case study. Although the book focuses on practices as well as principles, it is not so much a “user’s guide” as a full-blown theory of the inner workings and rationales of the case study method. That said, all scholars who read the book will discover many new ideas for carrying out better case study research in practice.

The book is delimited by the *kind* of case study research on which it focuses. Within the broad field of case study research, Gerring’s interest is very explicitly on work that seeks to make causal inferences, not work whose primary goals fall into the realm of descriptive inference. For example, the methodological issues that arise for case studies that are mainly interpretive or rooted in critical theory are not the focus of the discussion.