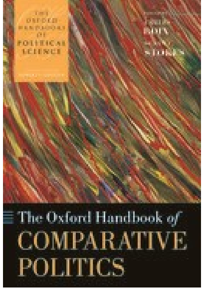

The Oxford Handbook of Comparative Politics
Carles Boix (ed.), Susan C. Stokes (ed.)

https://doi.org/10.1093/oxfordhb/9780199566020.001.0001

Published: 2009 **Online ISBN:** 9780191577482 **Print ISBN:** 9780199566020

CHAPTER

4 The Case Study: What it is and What it Does

John Gerring

https://doi.org/10.1093/oxfordhb/9780199566020.003.0004 Pages 90–122
Published: 02 September 2009

Abstract

This article presents a reconstructed definition of the case study approach to research. This definition emphasizes comparative politics, which has been closely linked to this method since its creation. The article uses this definition as a basis to explore a series of contrasts between cross-case study and case study research. This article attempts to provide better understanding of this persisting methodological debate as a matter of tradeoffs, which may also contribute to destroying the boundaries that have separated these rival genres within the subfield of comparative politics.

Keywords: reconstructed definition, case study approach, comparative politics, contrasts, cross-case study research, case study research, tradeoffs, methodological debate

Subject: Comparative Politics, Political Methodology, Politics

Series: Oxford Handbooks

Two centuries after Le Play's pioneering work, the various disciplines of the social sciences continue to produce a vast number of case studies, many of which have entered the pantheon of classic works. Judging by the large volume of recent scholarly output the case study research design plays a central role in anthropology, archeology, business, education, history, medicine, political science, psychology, social work, and sociology (Gerring 2007*a*, ch. 1). Even in economics and political economy, fields not usually noted for their receptiveness to case-based work, there has been something of a renaissance. Recent studies of economic growth have turned to case studies of unusual countries such as Botswana, Korea, and Mauritius.[1] Debates on the relationship between trade and growth and the IMF and growth have likewise combined cross-national regression evidence with in-depth (quantitative and qualitative) case analysis (Srinivasan and Bhagwati 1999; Vreeland 2003). Work on ethnic politics and ethnic conflict has exploited within-country variation or small-N crosscountry comparisons (Abadie and Gardeazabal 2003; Chandra 2004; Posner 2004). By the standard of praxis, therefore, it would appear that the method of the case study is solidly ensconced, perhaps even thriving. Arguably, we are witnessing a movement away from a variable-centered approach to causality in the social sciences and towards a case-based approach.

p. 91 Indeed, the statistical analysis of cross-case observational data has been subjected to increasing scrutiny in recent years. It no longer seems self-evident, even to nomothetically inclined scholars, that non-experimental data drawn from nation-states, cities, social movements, civil conflicts, or other complex phenomena should be treated in standard regression formats. The complaints are myriad, and oft-reviewed.[2] They include: (a) the problem of arriving at an adequate specification of the causal model, given a plethora of plausible models, and the associated problem of modeling interactions among these covariates; (b) identification problems, which cannot always be corrected by instrumental variable techniques; (c) the problem of "extreme" counterfactuals, i.e. extrapolating or interpolating results from a general model where the extrapolations extend beyond the observable data points; (d) problems posed by influential cases; (e) the arbitrariness of standard significance tests; (f) the misleading precision of point estimates in the context of "curve-fitting" models; (g) the problem of finding an appropriate estimator and modeling temporal autocorrelation in pooled time series; (h) the difficulty of identifying causal mechanisms; and last, but certainly not least, (i) the ubiquitous problem of faulty data drawn from a variety of questionable sources. Most of these difficulties may be understood as the byproduct of causal variables that offer limited variation through time and cases that are extremely heterogeneous.

A principal factor driving the general discontent with cross-case observational research is a new-found interest in experimental models of social scientific research. Following the pioneering work of Donald Campbell (1988; Cook and Campbell 1979) and Donald Rubin (1974), methodologists have taken a hard look at the regression model and discovered something rather obvious but at the same time crucially important: this research bears only a faint relationship to the true experiment, for all the reasons noted above. The current excitement generated by matching estimators, natural experiments, and field experiments may be understood as a move toward a quasi-experimental, and frequently case-based analysis of causal relations. Arguably, this is because the experimental ideal is often better approximated by a small number of cases that are closely related to one another, or by a single case observed over time, than by a large sample of heterogeneous units.

A third factor militating towards case-based analysis is the development of a series of alternatives to the standard linear/additive model of cross-case analysis, thus establishing a more variegated set of tools to capture the complexity of social behavior (see Brady and Collier 2004). Charles Ragin and associates have shown us how to deal with situations where multiple causal paths lead to the same set of outcomes, a series of techniques known as Qualitative Comparative Analysis (QCA) ("Symposium: Qualitative Comparative Analysis" 2004). Andrew Abbott has worked out a method that maps causal sequences across cases, known
p. 92 as optimal sequence matching (Abbott 2001; Abbott and Forrest 1986; Abbott and Tsay 2000). ↳ Bear Braumoeller, Gary Goertz, Jack Levy, and Harvey Starr have defended the importance of necessary-condition arguments in the social sciences, and have shown how these arguments might be analyzed (Braumoeller and Goertz 2000; Goertz 2003; Goertz and Levy forthcoming; Goertz and Starr 2003). James Fearon, Ned Lebow, Philip Tetlock, and others have explored the role of counterfactual thought experiments in the analysis of individual case histories (Fearon 1991; Lebow 2000; Tetlock and Belkin 1996). Colin Elman has developed a typological method of analyzing cases (Elman 2005). David Collier, Jack Goldstone, Peter Hall, James Mahoney, and Dietrich Rueschemeyer have worked to revitalize the comparative and comparative-historical methods (Collier 1993; Goldstone 1997; Hall 2003; Mahoney and Rueschemeyer 2003). And scores of researchers have attacked the problem of how to convert the relevant details of a temporally constructed narrative into standardized formats so that cases can be meaningfully compared (Abell 1987, 2004; Abbott 1992; Buthe 2002; Griffin 1993). While not all of these techniques are, strictly speaking, case study techniques—since they sometimes involve a large number of cases—they do move us closer to a case-based understanding of causation insofar as they preserve the texture and detail of individual cases, features that are often lost in large-N cross-case analysis.

A fourth factor concerns the recent marriage of rational choice tools with case study analysis, sometimes referred to as an "analytic narrative" (Bates et al. 1998). Whether the technique is qualitative or quantitative, scholars equipped with economic models are turning, increasingly, to case studies in order to test the theoretical predictions of a general model, investigate causal mechanisms, and/or explain the features of a key case.

Finally, epistemological shifts in recent decades have enhanced the attractiveness of the case study format. The "positivist" model of explanation, which informed work in the social sciences through most of the twentieth century, tended to downplay the importance of causal mechanisms in the analysis of causal relations. Famously, Milton Friedman (1953) argued that the only criterion of a model was to be found in its accurate prediction of outcomes. The verisimilitude of the model, its accurate depiction of reality, was beside the point. In recent years, this explanatory trope has come under challenge from "realists," who claim (among other things) that causal analysis should pay close attention to causal mechanisms (e.g. Bunge 1997; Little 1998). Within political science and sociology, the identification of a specific mechanism—a causal pathway—has come to be seen as integral to causal analysis, regardless of whether the model in question is formal or informal or whether the evidence is qualitative or quantitative (Achen 2002; Elster 1998; George and Bennett 2005; Hedstrom and Swedberg 1998). Given this new-found (or at least newly self-conscious) interest in mechanisms, it is not surprising that social scientists would turn to case studies as a mode of causal investigation.

For all the reasons stated above, one might intuit that social science is moving towards a case-based understanding of causal relations. Yet, this movement, insofar as it exists, has scarcely been acknowledged, and would certainly be challenged by many close observers—including some of those cited in the foregoing passages.

p. 93 The fact is that the case study research design is still viewed by most methodologists with extreme circumspection. A work that focuses its attention on a single example of a broader phenomenon is apt to be described as a "mere" case study, and is often identified with loosely framed and non-generalizable theories, biased case selection, informal and undisciplined research designs, weak empirical leverage (too many variables and too few cases), subjective conclusions, non-replicability, and causal determinism. To some, the term case study is an ambiguous designation covering a multitude of "inferential felonies."[3]

The quasi-mystical qualities associated with the case study persist to this day. In the field of psychology, a gulf separates "scientists" engaged in cross-case research and "practitioners" engaged in clinical research, usually focused on several cases (Hersen and Barlow 1976, 21). In the fields of political science and sociology, case study researchers are acknowledged to be on the "soft" side of hard disciplines. And across fields, the persisting case study orientations of anthropology, education, law, social work, and various other fields and subfields relegate them to the non-rigorous, non-systematic, non-scientific, non-positivist end of the academic spectrum.

The methodological status of the case study is still, officially, suspect. Even among its defenders there is confusion over the virtues and vices of this ambiguous research design. Practitioners continue to ply their trade but have difficulty articulating what it is they are doing, methodologically speaking. The case study survives in a curious methodological limbo.

This leads to a paradox: although much of what we know about the empirical world has been generated by case studies and case studies continue to constitute a large proportion of work generated by the social science disciplines, the case study *method* is poorly understood.

How can we make sense of the profound disjuncture between the acknowledged contributions of this genre to the various disciplines of social science and its maligned status within these disciplines? If case studies are methodologically flawed, why do they persist? Should they be rehabilitated, or suppressed? How fruitful *is* this style of research?

In this chapter, I provide a reconstructed definition of the case study approach to research with special emphasis on comparative politics, a field that has been closely identified with this method since its birth. Based on this definition, I then explore a series of contrasts between case study and cross-case study research. These contrasts are intended to illuminate the characteristic strengths and weaknesses ("affinities") of these two research designs, not to vindicate one or the other. The effort of this chapter is to understand this persisting methodological debate as a matter of tradeoffs. Case studies and cross-case studies explore the world in different ways. Yet, properly constituted, there is no reason that case study
p. 94 results cannot be ↳ synthesized with results gained from cross-case analysis, and vice versa. My hope, therefore, is that this chapter will contribute to breaking down the boundaries that have separated these rival genres within the subfield of comparative politics.

1 Definitions

The key term of this chapter is, admittedly, a definitional morass. To refer to a work as a "case study" might mean: that its method is qualitative, small-N; that the research is holistic, thick (a more or less comprehensive examination of a phenomenon); that it utilizes a particular type of evidence (e.g. ethnographic, clinical, non-experimental, non-survey based, participant observation, process tracing, historical, textual, or field research); that its method of evidence gathering is naturalistic (a "real-life context"); that the research investigates the properties of a single observation; or that the research investigates the properties of a single phenomenon, instance, or example. Evidently, researchers have many things in mind when they talk about case study research. Confusion is compounded by the existence of a large number of near-synonyms—single unit, single subject, single case, N=1, case based, case control, case history, case method, case record, case work, clinical research, and so forth. As a result of this profusion of terms and meanings, proponents and opponents of the case study marshal a wide range of arguments but do not seem any closer to agreement than when this debate was first broached several decades ago.

Can we reconstruct this concept in a clearer, more productive fashion? In order to do so we must understand how the key terms—case and case study—are situated within a neighborhood of related terms. In this crowded semantic field, each term is defined in relation to others. And in the context of a specific work or research terrain, they all take their meaning from a specific inference. (The reader should bear in mind that any change in the inference, and the meaning of all the key terms will probably change.) My attempt here will be to provide a single, determinate, definition of these key terms. Of course, researchers may choose to define these terms in many different ways. However, for purposes of methodological discussion it is helpful to enforce a uniform vocabulary.

Let us stipulate that a case connotes a spatially delimited phenomenon (a unit) observed at a single point in time or over some period of time. It comprises the sort of phenomena that an inference attempts to explain. Thus, in a study that attempts to explain certain features of nation-states, cases are comprised of nation-states (across some temporal frame). In a study that attempts to explain the behavior of individuals, individuals comprise the cases. And so forth. Each case may provide a single observation or multiple (within-case) observations.

For students of comparative politics, the archetypal case is the dominant political unit of our time, the
p. 95 nation-state. However, the study of smaller social and political ↳ units (regions, cities, villages, communities, social groups, families) or specific institutions (political parties, interest groups, businesses) is equally common in other subfields, and perhaps increasingly so in comparative politics. Whatever the chosen unit, the methodological issues attached to the case study have nothing to do with the size of the individual cases. A case may be created out of any phenomenon so long as it has identifiable boundaries and comprises the primary object of an inference.

Note that the spatial boundaries of a case are often more apparent than its temporal boundaries. We know, more or less, where a country begins and ends, even though we may have difficulty explaining *when* a country begins and ends. Yet, some temporal boundaries must be assumed. This is particularly important when cases consist of discrete events—crises, revolutions, legislative acts, and so forth— within a single unit. Occasionally, the temporal boundaries of a case are more obvious than its spatial boundaries. This is true when the phenomena under study are eventful but the unit undergoing the event is amorphous. For example, if one is studying terrorist attacks it may not be clear how the spatial unit of analysis should be understood, but the events themselves may be well bounded.

A *case study* may be understood as the intensive study of a single case for the purpose of understanding a larger class of cases (a population). Case study research may incorporate several cases. However, at a certain point it will no longer be possible to investigate those cases intensively. 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*. Evidently, the distinction between a case study and cross-case study is a continuum. The fewer cases there are, and the more intensively they are studied, the more a work merits the appellation case study. Even so, this proves to be a useful distinction, for much follows from it.

A few additional terms will now be formally defined.

An *observation* is the most basic element of any empirical endeavor. Conventionally, the number of observations in an analysis is referred to with the letter *N*. (Confusingly, *N* may also be used to designate the number of cases in a study, a usage that I shall try to avoid.) A single observation may be understood as containing several dimensions, each of which may be measured (across disparate observations) as a variable. Where the proposition is causal, these may be subdivided into *dependent* (Y) and *independent* (X) variables. The dependent variable refers to the outcome of an investigation. The independent variable refers to the explanatory (causal) factor, that which the outcome is supposedly dependent on.

Note that a case may consist of a single observation (N=1). This would be true, for example, in a cross-sectional analysis of multiple cases. In a case study, however, the case under study always provides more than one observation. These may be constructed diachronically (by observing the case or some subset of within-case units through time) or synchronically (by observing within-case variation at a single point in time).

This is a clue to the fact that case studies and cross-case usually operate at different levels of analysis. The
p. 96 case study is typically focused on within-case variation (if there ↳ a cross-case component it is probably secondary). The cross-case study, as the name suggests, is typically focused on cross-case variation (if there is also within-case variation, it is secondary in importance). They have the same object in view—the explanation of a population of cases—but they go about this task differently.

A *sample* consists of whatever cases are subjected to formal analysis; they are the immediate subject of a study or case study. (Confusingly, the sample may also refer to the observations under study, and will be so used at various points in this narrative. But at present, we treat the sample as consisting of cases.) Technically, one might say that in a case study the sample consists of the case or cases that are subjected to intensive study. However, usually when one uses the term sample one is implying that the number of cases is rather large. Thus, "sample-based work" will be understood as referring to large-N cross-case methods—the opposite of case study work. Again, the only feature distinguishing the case study format from a sample-based (or "cross-case") research design is the number of cases falling within the sample—one or a few versus many. Case studies, like large-N samples, seek to represent, in all ways relevant to the proposition at hand, a population of cases. A series of case studies might therefore be referred to as a sample if they are relatively brief and relatively numerous; it is a matter of emphasis and of degree. The more case studies one has, the less intensively each one is studied, and the more confident one is in their representativeness (of some broader population), the more likely one is to describe them as a sample rather than a series of case studies. For practical reasons—unless, that is, a study is extraordinarily long—the case study research format is usually limited to a dozen cases or less. A single case is not at all unusual.

The sample rests within a *population* of cases to which a given proposition refers. The population of an inference is thus equivalent to the breadth or scope of a proposition. (I use the terms *proposition*, *hypothesis*, *inference*, and *argument* interchangeably.) Note that most samples are not exhaustive; hence the use of the term sample, referring to *sampling* from a population. Occasionally, however, the sample equals the population of an inference; all potential cases are studied.

For those familiar with the rectangular form of a dataset it may be helpful to conceptualize observations as rows, variables as columns, and cases as either groups of observations or individual observations.

2 What is a Case Study Good For? Case Study versus Cross-Case Analysis

I have argued that the case study approach to research is most usefully defined as the intensive study of a single unit or a small number of units (the cases), for the purpose of understanding a larger class of similar
p. 97 units (a population of cases). This is put ↳ forth as a minimal definition of the topic.[4] I now proceed to discuss the *non*-definitional attributes of the case study—attributes that are often, but not invariably, associated with the case study method. These will be understood as methodological affinities flowing from a minimal definition of the concept.[5]

The case study research design exhibits characteristic strengths and weaknesses relative to its large-N cross-case cousin. These tradeoffs derive, first of all, from basic research goals such as (1) whether the study is oriented toward hypothesis generating or hypothesis testing, (2) whether internal or external validity is prioritized, (3) whether insight into causal mechanisms or causal effects is more valuable, and (4) whether the scope of the causal inference is deep or broad. These tradeoffs also hinge on the shape of the empirical universe, i.e. (5) whether the population of cases under study is heterogeneous or homogeneous, (6) whether the causal relationship of interest is strong or weak, (7) whether useful variation on key parameters within that population is rare or common, and (8) whether available data are concentrated or dispersed.

Along each of these dimensions, case study research has an affinity for the first factor and cross-case research has an affinity for the second, as summarized in Table 4.1. To clarify, these tradeoffs represent methodological *affinities*, not invariant laws. Exceptions can be found to each one. Even so, these general tendencies are often ↳ noted in case study research and have been reproduced in multiple disciplines and subdisciplines over the course of many decades.

p. 98

Table 4.1 Case study and cross-case research designs: affinities and tradeoffs

	Affinity	
	Case study	**Cross-case study**
Research goals		
1. Hypothesis	Generating	Testing
2. Validity	Internal	External
3. Causal insight	Mechanisms	Effects
4. Scope of proposition	Deep	Broad
Empirical factors		
5. Population of cases	Heterogeneous	Homogeneous
6. Causal strength	Strong	Weak
7. Useful variation	Rare	Common
8. Data availability	Concentrated	Dispersed

It should be stressed that each of these tradeoffs carries a ceteris paribus caveat. Case studies are more useful for generating new hypotheses, *all other things being equal*. The reader must bear in mind that many additional factors also rightly influence a writer's choice of research design, and they may lean in the other direction. Ceteris are not always paribus. One should not jump to conclusions about the research design appropriate to a given setting without considering the entire range of issues involved—some of which may be more important than others.

3. Hypothesis: Generating versus Testing

Social science research involves a quest for new theories as well as a testing of existing theories; it is comprised of both "conjectures" and "refutations."[6] Regrettably, social science methodology has focused almost exclusively on the latter. The conjectural element of social science is usually dismissed as a matter of guesswork, inspiration, or luck—a leap of faith, and hence a poor subject for methodological reflection.[7] Yet, it will readily be granted that many works of social science, including most of the acknowledged classics, are seminal rather than definitive. Their classic status derives from the introduction of a new idea or a new perspective that is subsequently subjected to more rigorous (and refutable) analysis. Indeed, it is difficult to devise a program of falsification the first time a new theory is proposed. Path-breaking research, almost by definition, is protean. Subsequent research on that topic tends to be more definitive insofar as its primary task is limited: to verify or falsify a preexisting hypothesis. Thus, the world of social science may be usefully divided according to the predominant goal undertaken in a given study, either hypothesis *generating* or hypothesis *testing*. There are two moments of empirical research, a lightbulb moment and a skeptical moment, each of which is essential to the progress of a discipline.[8]

p. 99 Case studies enjoy a natural advantage in research of an exploratory nature. Several millennia ago, Hippocrates reported what were, arguably, the first case studies ever conducted. They were fourteen in number.[9] Darwin's insights into the process of human evolution came after his travels to a few select locations, notably Easter Island. Freud's revolutionary work on human psychology was constructed from a close observation of fewer than a dozen clinical cases. Piaget formulated his theory of human cognitive development while watching his own two children as they passed from childhood to adulthood. Lévi-Strauss's structuralist theory of human cultures built on the analysis of several North and South American tribes. Douglass North's neo-institutionalist theory of economic development was constructed largely through a close analysis of a handful of early developing states (primarily England, the Netherlands, and the United States).[10] Many other examples might be cited of seminal ideas that derived from the intensive study of a few key cases.

Evidently, the sheer number of examples of a given phenomenon does not, by itself, produce insight. It may only confuse. How many times did Newton observe apples fall before he recognized the nature of gravity? This is an apocryphal example, but it illustrates a central point: case studies may be more useful than cross-case studies when a subject is being encountered for the first time or is being considered in a fundamentally new way. After reviewing the case study approach to medical research, one researcher finds that although case reports are commonly regarded as the lowest or weakest form of evidence, they are nonetheless understood to comprise "the first line of evidence." The hallmark of case reporting, according to Jan Vanden-broucke, "is to recognize the unexpected." This is where discovery begins.[11]

The advantages that case studies offer in work of an exploratory nature may also serve as impediments in work of a confirmatory/disconfirmatory nature. Let us briefly explore why this might be so.[12]

Traditionally, scientific methodology has been defined by a segregation of conjecture and refutation. One should not be allowed to contaminate the other.[13] Yet, in the real world of social science, inspiration is often associated with perspiration. "Light-bulb" moments arise from a close engagement with the particular facts of a particular case. Inspiration is more likely to occur in the laboratory than in the shower.

The circular quality of conjecture and refutation is particularly apparent in case study research. Charles Ragin notes that case study research is all about "casing"— defining the topic, including the hypothesis(es) of primary interest, the outcome, and the set of cases that offer relevant information vis-à-vis the hypothesis.[14] A study of the French Revolution may be conceptualized as a study of revolution, of social revolution, of revolt, of political violence, and so forth. Each of these topics entails a different population and a different set of causal factors. A good deal of authorial p. 100 intervention is necessary in the course of defining a case study topic, for there is a great deal of evidentiary leeway. Yet, the "subjectivity" of case study research allows for the generation of a great number of hypotheses, insights that might not be apparent to the cross-case researcher who works with a thinner set of empirical data across a large number of cases and with a more determinate (fixed) definition of cases, variables, and outcomes. It is the very fuzziness of case studies that grants them an advantage in research at the exploratory stage, for the single-case study allows one to test a multitude of hypotheses in a rough-and-ready way. Nor is this an entirely "conjectural" process. The relationships discovered among different elements of a single case have a prima facie causal connection: they are all at the scene of the crime. This is revelatory when one is at an early stage of analysis, for at that point there is no identifiable suspect and the crime itself may be difficult to discern. The fact that *A*, *B*, and *C* are present at the expected times and places (relative to some outcome of interest) is sufficient to establish them as independent variables. Proximal evidence is all that is required. Hence, the common identification of case studies as "plausibility probes," "pilot studies," "heuristic studies," "exploratory" and "theory-building" exercises.[15]

A large-N cross-study, by contrast, generally allows for the testing of only a few hypotheses but does so with a somewhat greater degree of confidence, as is appropriate to work whose primary purpose is to test an extant theory. There is less room for authorial intervention because evidence gathered from a cross-case research design can be interpreted in a limited number of ways. It is therefore more reliable. Another way of stating the point is to say that while case studies lean toward Type 1 errors (falsely rejecting the null hypothesis), cross-case studies lean toward Type 2 errors (failing to reject the false null hypothesis). This explains why case studies are more likely to be paradigm generating, while cross-case studies toil in the prosaic but highly structured field of normal science.

I do not mean to suggest that case studies never serve to confirm or disconfirm hypotheses. Evidence drawn from a single case may falsify a necessary or sufficient hypothesis, as discussed below. Additionally, case studies are often useful for the purpose of elucidating causal mechanisms, and this obviously affects the plausibility of an X/Y relationship. However, general theories rarely offer the kind of detailed and determinate predictions on within-case variation that would allow one to reject a hypothesis through pattern matching (without additional cross-case evidence). Theory testing is not the case study's strong suit. The selection of "crucial" cases is at pains to overcome the fact that the cross-case *N* is minimal. Thus, one is unlikely to reject a hypothesis, or to consider it definitively proved, on the basis of the study of a single case.

Harry Eckstein himself acknowledges that his argument for case studies as a form of theory confirmation is largely hypothetical. At the time of writing, several decades ago, he could not point to any social science study where a crucial case study had performed the heroic role assigned to it.[16] I suspect that this is still more or less true. p. 101 Indeed, it is true even of experimental case studies in the natural sciences. "We must recognize," note Donald Campbell and Julian Stanley,

> that continuous, multiple experimentation is more typical of science than once-and-for-all definitive experiments. The experiments we do today, if successful, will need replication and cross-validation at other times under other conditions before they can become an established part of science ... [E]ven though we recognize experimentation as the basic language of proof ... we should not expect that "crucial experiments" which pit opposing theories will be likely to have clear-cut outcomes. When one finds, for example, that competent observers advocate strongly divergent points of view, it seems likely on a priori grounds that both have observed something valid about the natural situation, and that both represent a part of the truth. The stronger the controversy, the more likely this is. Thus we might expect in such cases an experimental outcome with mixed results, or with the balance of truth varying subtly from experiment to experiment. The more mature focus...avoids crucial experiments and instead studies dimensional relationships and interactions along many degrees of the experimental variables.[17]

A single case study is still a single shot—a single example of a larger phenomenon.

The tradeoff between hypothesis generating and hypothesis testing helps us to reconcile the enthusiasm of case study researchers and the skepticism of case study critics. They are both right, for the looseness of case study research is a boon to new conceptualizations just as it is a bane to falsification.

4. Validity: Internal versus External

Questions of validity are often distinguished according to those that are *internal* to the sample under study and those that are *external* (i.e. applying to a broader— unstudied—population). Cross-case research is always more representative of the population of interest than case study research, so long as some sensible procedure of case selection is followed (presumably some version of random sampling). Case study research suffers problems of representativeness because it includes, by definition, only a small number of cases of some more general phenomenon. Are the men chosen by Robert Lane typical of white, immigrant, working-class, American males?[18] Is Middletown representative of other cities in America?[19] These sorts of questions forever haunt case study research. This means that case study research is generally weaker with respect to external validity than its cross-case cousin.

The corresponding virtue of case study research is its internal validity. Often, though not invariably, it is easier to establish the veracity of a causal relationship pertaining to a single case (or a small number of cases) than for a larger set of cases. Case study researchers share the bias of experimentalists in this regard: p. 102 they tend to be ↳ more disturbed by threats to within-sample validity than by threats to out-of-sample validity. Thus, it seems appropriate to regard the tradeoff between external and internal validity, like other tradeoffs, as intrinsic to the cross-case/single-case choice of research design.

5. Causal Insight: Causal Mechanisms versus Causal Effects

A third tradeoff concerns the sort of insight into causation that a researcher intends to achieve. Two goals may be usefully distinguished. The first concerns an estimate of the causal *effect*; the second concerns the investigation of a causal *mechanism* (i.e. pathway from X to Y).

By causal effect I refer to two things: (a) the magnitude of a causal relationship (the expected effect on Y of a given change in X across a population of cases) and (b) the relative precision or uncertainty associated with that point estimate. Evidently, it is difficult to arrive at a reliable estimate of causal effects across a population of cases by looking at only a single case or a small number of cases. (The one exception would be an experiment in which a given case can be tested repeatedly, returning to a virgin condition after each test. But here one faces inevitable questions about the representativeness of that much-studied case.)[20] Thus, the estimate of a causal effect is almost always grounded in cross-case evidence.

It is now well established that causal arguments depend not only on measuring causal effects, but also on the identification of a causal mechanism.[21] *X* must be connected with *Y* in a plausible fashion; otherwise, it is unclear whether a pattern of covariation is truly causal in nature, or what the causal interaction might be. Moreover, without a clear understanding of the causal pathway(s) at work in a causal relationship it is impossible to accurately specify the model, to identify possible instruments for the regressor of interest (if there are problems of endogeneity), or to interpret the results.[22] Thus, causal mechanisms are presumed in every estimate of a mean (average) causal effect.

p. 103 In the task of investigating causal mechanisms, cross-case studies are often not so illuminating. It has become a common criticism of large-N cross-national research—e.g. into the causes of growth, democracy, civil war, and other national-level outcomes—that such studies demonstrate correlations between inputs and outputs without clarifying the reasons for those correlations (i.e. clear causal pathways). We learn, for example, that infant mortality is strongly correlated with state failure;[23] but it is quite another matter to interpret this finding, which is consistent with a number of different causal mechanisms. Sudden increases in infant mortality might be the product of famine, of social unrest, of new disease vectors, of government repression, and of countless other factors, some of which might be expected to impact the stability of states, and others of which are more likely to be a result of state instability.

Case studies, if well constructed, may allow one to peer into the box of causality to locate the intermediate factors lying between some structural cause and its purported effect. Ideally, they allow one to "see" *X* and *Y* interact—Hume's billiard ball crossing the table and hitting a second ball.[24] Barney Glaser and Anselm Strauss point out that in fieldwork "general relations are often discovered *in vivo*; that is, the field worker literally sees them occur."[25] When studying decisional behavior case study research may offer insight into the intentions, the reasoning capabilities, and the information-processing procedures of the actors involved in a given setting. Thus, Dennis Chong uses in-depth interviews with a very small sample of respondents in order to better understand the process by which people reach decisions about civil liberties issues. Chong comments:

> One of the advantages of the in-depth interview over the mass survey is that it records more fully how subjects arrive at their opinions. While we cannot actually observe the underlying mental process that gives rise to their responses, we can witness many of its outward manifestations. The way subjects ramble, hesitate, stumble, and meander as they formulate their answers tips us off to how they are thinking and reasoning through political issues.[26]

Similarly, the investigation of a single case may allow one to test the causal implications of a theory, thus providing corroborating evidence for a causal argument. This is sometimes referred to as pattern matching (Campbell 1988).

Dietrich Rueschemeyer and John Stephens offer an example of how an examination of causal mechanisms may call into question a general theory based on cross-case evidence. The thesis of interest concerns the role of British colonialism in fostering democracy among postcolonial regimes. In particular, the authors investigate the diffusion hypothesis, that democracy was enhanced by "the transfer of British governmental
p. 104 and representative institutions and the tutoring of the colonial ↳ people in the ways of British government." On the basis of in-depth analysis of several cases the authors report:

> We did find evidence of this diffusion effect in the British settler colonies of North America and the Antipodes; but in the West Indies, the historical record points to a different connection between British rule and democracy. There the British colonial administration opposed suffrage extension, and only the white elites were "tutored" in the representative institutions. But, critically, we argued on the basis of the contrast with Central America, British colonialism did prevent the local plantation elites from controlling the local state and responding to the labor rebellion of the 1930s with massive repression. Against the adamant opposition of that elite, the British colonial rulers responded with concessions which allowed for the growth of the party—union complexes rooted in the black middle and working classes, which formed the backbone of the later movement for democracy and independence. Thus, the narrative histories of these cases indicate that the robust statistical relation between British colonialism and democracy is produced only in part by diffusion. The interaction of class forces, state power, and colonial policy must be brought in to fully account for the statistical result.[27]

Whether or not Rueschemeyer and Stephens are correct in their conclusions need not concern us here. What is critical, however, is that any attempt to deal with this question of causal mechanisms is heavily reliant on evidence drawn from case studies. In this instance, as in many others, the question of causal pathways is simply too difficult, requiring too many poorly measured or unmeasurable variables, to allow for accurate cross-sectional analysis.[28]

To be sure, causal mechanisms do not always require explicit attention. They may be quite obvious. And in other circumstances, they may be amenable to cross-case investigation. For example, a sizeable literature addresses the causal relationship between trade openness and the welfare state. The usual empirical finding is that more open economies are associated with higher social welfare spending. The question then becomes why such a robust correlation exists. What are the plausible interconnections between trade openness and social welfare spending? One possible causal path, suggested by David Cameron,[29] is that increased trade openness leads to greater domestic economic vulnerability to external shocks (due, for instance, to changing terms of trade). If so, one should find a robust correlation between annual variations in a country's terms of trade (a measure of economic vulnerability) and social welfare spending. As it happens, the correlation is not robust and this leads some commentators to doubt whether the putative causal mechanism proposed by David Cameron and many others is actually at work.[30] Thus, in instances where an intervening variable can be effectively operationalized across a large sample of cases it may be possible to test causal mechanisms without resorting to case study investigation.[31]

p. 105 Even so, the opportunities for investigating causal pathways are generally more apparent in a case study format. Consider the contrast between formulating a standardized survey for a large group of respondents and formulating an in-depth interview with a single subject or a small set of subjects, such as that undertaken by Dennis Chong in the previous example. In the latter situation, the researcher is able to probe into details that would be impossible to delve into, let alone anticipate, in a standardized survey. She may also be in a better position to make judgements as to the veracity and reliability of the respondent. Tracing causal mechanisms is about cultivating sensitivity to a local context. Often, these local contexts are essential to cross-case testing. Yet, the same factors that render case studies useful for micro-level investigation also make them less useful for measuring mean (average) causal effects. It is a classic tradeoff.

6 Scope of Proposition: Deep versus Broad

The utility of a case study mode of analysis is in part a product of the scope of the causal argument that a researcher wishes to prove or demonstrate. Arguments that strive for great breadth are usually in greater need of cross-case evidence; causal arguments restricted to a small set of cases can more plausibly subsist on the basis of a single-case study. The extensive/intensive tradeoff is fairly commonsensical.[32] A case study of France probably offers more useful evidence for an argument about Europe than for an argument about the whole world. Propositional breadth and evidentiary breadth generally go hand in hand.

Granted, there are a variety of ways in which single-case studies can credibly claim to provide evidence for causal propositions of broad reach—e.g. by choosing cases that are especially representative of the phenomenon under study ("typical" cases) or by choosing cases that represent the most difficult scenario for a given proposition and are thus biased against the attainment of certain results ("crucial" cases). Even so, a proposition with a narrow scope is more conducive to case study analysis than a proposition with a broad purview, all other things being equal. The breadth of an inference thus constitutes one factor, among many, in determining the utility of the case study mode of analysis. This is reflected in the hesitancy of many case study researchers to invoke determinate causal propositions with great reach —"covering laws," in the idiom of philosophy of science.

By the same token, one of the primary virtues of the case study method is the depth of analysis that it offers.
p. 106 One may think of depth as referring to the detail, ↳ richness, completeness, wholeness, or the degree of variance in an outcome that is accounted for by an explanation. The case study researcher's complaint about the thinness of cross-case analysis is well taken; such studies often have little to say about individual cases. Otherwise stated, cross-case studies are likely to explain only a small portion of the variance with respect to a given outcome. They approach that outcome at a very general level. Typically, a cross-case study aims only to explain the occurrence/non-occurrence of a revolution, while a case study might also strive to explain specific features of that event—why it occurred when it did and in the way that it did. Case studies are thus rightly identified with "holistic" analysis and with the "thick" description of events.[33]

Whether to strive for breadth or depth is not a question that can be answered in any definitive way. All we can safely conclude is that researchers invariably face a choice between knowing more about less, or less about more. The case study method may be defended, as well as criticized, along these lines.[34] Indeed, arguments about the "contextual sensitivity" of case studies are perhaps more precisely (and fairly) understood as arguments about depth and breadth. The case study researcher who feels that cross-case research on a topic is insensitive to context is usually not arguing that *nothing at all* is consistent across the chosen cases. Rather, the case study researcher's complaint is that much more could be said—accurately—about the phenomenon in question with a reduction in inferential scope.[35]

Indeed, I believe that a number of traditional issues related to case study research can be understood as the product of this basic tradeoff. For example, case study research is often lauded for its holistic approach to the study of social phenomena in which behavior is observed in natural settings. Cross-case research, by contrast, is criticized for its construction of artificial research designs that decontextualize the realm of social behavior by employing abstract variables that seem to bear little relationship to the phenomena of interest.[36] These associated congratulations and critiques may be understood as a conscious choice on the part of case study researchers to privilege depth over breadth.

7 The Population of Cases: Heterogeneous versus Homogeneous

The choice between a case study and cross-case style of analysis is driven not only by the goals of the
p. 107 researcher, as reviewed above, but also by the shape of the empirical ↳ universe that the researcher is attempting to understand. Consider, for starters, that the logic of cross-case analysis is premised on some degree of cross-unit comparability (unit homogeneity). Cases must be similar to each other in whatever respects might affect the causal relationship that the writer is investigating, or such differences must be controlled for. Uncontrolled heterogeneity means that cases are "apples and oranges;" one cannot learn anything about underlying causal processes by comparing their histories. The underlying factors of interest mean different things in different contexts (conceptual stretching) or the *X*/*Y* relationship of interest is different in different contexts (unit heterogeneity).

Case study researchers are often suspicious of large-sample research, which, they suspect, contains heterogeneous cases whose differences cannot easily be modeled. "Variable-oriented" research is said to involve unrealistic "homogenizing as-sumptions."[37] In the field of international relations, for example, it is common to classify cases according to whether they are deterrence failures or deterrence successes. However, Alexander George and Richard Smoke point out that "the separation of the dependent variable into only two subclasses, deterrence success and deterrence failure," neglects the great variety of ways in which deterrence can fail. Deterrence, in their view, has many independent causal paths (causal equifinality), and these paths may be obscured when a study lumps heterogeneous cases into a common sample.[38]

Another example, drawn from clinical work in psychology, concerns heterogeneity among a sample of individuals. Michel Hersen and David Barlow explain:

> Descriptions of results from 50 cases provide a more convincing demonstration of the effectiveness of a given technique than separate descriptions of 50 individual cases. The major difficulty with this approach, however, is that the category in which these clients are classified most always becomes unmanageably heterogeneous. "Neurotics," [for example],...may have less in common than any group of people one would choose randomly. When cases are described individually, however, a clinician stands a better chance of gleaning some important information, since specific problems and specific procedures are usually described in more detail. When one lumps cases together in broadly defined categories, individual case descriptions are lost and the ensuing report of percentage success becomes meaningless.[39]

Under circumstances of extreme case heterogeneity, the researcher may decide that she is better off focusing on a single case or a small number of relatively homogeneous cases. Within-case evidence, or cross-case evidence drawn from a handful of most-similar cases, may be more useful than cross-case evidence, even though the ultimate interest of the investigator is in a broader population of cases. (Suppose one has a population of very heterogeneous cases, one or two of which undergo quasi-experimental transformations. p. 108 Probably, one gains greater insight into causal ↳ patterns throughout the population by examining these cases in detail than by undertaking some large-N cross-case analysis.) By the same token, if the cases available for study are relatively homogeneous, then the methodological argument for cross-case analysis is correspondingly strong. The inclusion of additional cases is unlikely to compromise the results of the investigation because these additional cases are sufficiently similar to provide useful information.

The issue of population heterogeneity/homogeneity may be understood, therefore, as a tradeoff between *N* (observations) and *K* (variables). If, in the quest to explain a particular phenomenon, each potential case offers only one observation and also requires one control variable (to neutralize heterogeneities in the resulting sample), then one loses degrees of freedom with each additional case. There is no point in using cross-case analysis or in extending a two-case study to further cases. If, on the other hand, each additional case is relatively cheap—if no control variables are needed or if the additional case offers more than one useful observation (through time)—then a cross-case research design may be warranted.[40] To put the matter more simply, when adjacent cases are unit homogeneous the addition of more cases is easy, for there is no (or very little) heterogeneity to model. When adjacent cases are heterogeneous additional cases are expensive, for every added heterogeneous element must be correctly modeled, and each modeling adjustment requires a separate (and probably unverifiable) assumption. The more background assumptions are required in order to make a causal inference, the more tenuous that inference is; it is not simply a question of attaining statistical significance. The ceteris paribus assumption at the core of all causal analysis is brought into question. In any case, the argument between case study and cross-case research designs is not about causal complexity per se (in the sense in which this concept is usually employed), but rather about the tradeoff between *N* and *K* in a particular empirical realm, and about the ability to model case heterogeneity through statistical legerdemain.[41]

Before concluding this discussion it is important to point out that researchers' judgements about case comparability are not, strictly speaking, matters that can be empirically verified. To be sure, one can look—and ought to look—for empirical patterns among potential cases. If those patterns are strong then the assumption of case comparability seems reasonably secure, and if they are not then there are grounds for doubt. However, debates about case comparability usually concern borderline instances. Consider that many phenomena of interest to social scientists are not rigidly bounded. If one is studying democracies there is always the question of how to define a democracy, and therefore of determining how high or low the threshold for inclusion in the sample should be. Researchers have different ideas about this, and these ideas p. 109 can hardly be tested in a rigorous fashion. Similarly, there ↳ are long-standing disputes about whether it makes sense to lump poor and rich societies together in a single sample, or whether these constitute distinct populations. Again, the borderline between poor and rich (or "developed" and "undeveloped") is blurry, and the notion of hiving off one from the other for separate analysis questionable, and unresolvable on purely empirical grounds. There is no safe (or "conservative") way to proceed. A final sticking point concerns the cultural/historical component of social phenomena. Many case study researchers feel that to compare societies with vastly different cultures and historical trajectories is meaningless. Yet, many cross-case researchers feel that to restrict one's analytic focus to a single cultural or geographic region is highly arbitrary, and equally meaningless. In these situations, it is evidently the choice of the researcher how to understand case homogeneity/heterogeneity across the potential populations of an inference. Where do like cases end and unlike cases begin?

Because this issue is not, strictly speaking, empirical it may be referred to as an *ontological* element of research design. An ontology is a vision of the world as it really is, a more or less coherent set of assumptions about how the world works, a research *Weltanschauung* analogous to a Kuhnian paradigm.[42] While it seems odd to bring ontological issues into a discussion of social science methodology it may be granted that social science research is not a purely empirical endeavor. What one finds is contingent upon what one looks for, and what one looks for is to some extent contingent upon what one expects to find. Stereotypically, case study researchers tend to have a "lumpy" vision of the world; they see countries, communities, and persons as highly individualized phenomena. Cross-case researchers, by contrast, have a less differentiated vision of the world; they are more likely to believe that things are pretty much the same everywhere, at least as respects basic causal processes. These basic assumptions, or ontologies, drive many of the choices made by researchers when scoping out appropriate ground for research.

8 Causal Strength: Strong versus Weak

Regardless of whether the population is homogeneous or heterogeneous, causal relationships are easier to study if the causal effect is strong, rather than weak. Causal "strength," as I use the term here, refers to the magnitude and consistency of X's effect on Y across a population of cases. (It invokes both the shape of the evidence at hand and whatever priors might be relevant to an interpretation of that evidence.) Where X has a
p. 110 strong effect on Y it will be relatively easy to study this relationship. Weak ↳ relationships, by contrast, are often difficult to discern. This much is commonsensical, and applies to all research designs.

For our purposes, what is significant is that weak causal relationships are particularly opaque when encountered in a case study format. Thus, there is a methodological affinity between weak causal relationships and large-N cross-case analysis, and between strong causal relationships and case study analysis.

This point is clearest at the extremes. The strongest species of causal relationships may be referred to as *deterministic*, where X is assumed to be necessary and/or sufficient for Y's occurrence. A necessary and sufficient cause accounts for all of the variation on Y. A sufficient cause accounts for all of the variation in certain instances of Y. A necessary cause accounts, by itself, for the absence of Y. In all three situations, the relationship is usually assumed to be perfectly consistent, i.e. invariant. There are no exceptions.

It should be clear why case study research designs have an easier time addressing causes of this type. Consider that a deterministic causal proposition can be *dis*proved with a single case.[43] For example, the reigning theory of political stability once stipulated that only in countries that were relatively homogeneous, or where existing heterogeneity was mitigated by cross-cutting cleavages, would social peace endure.[44] Arend Lijphart's case study of the Netherlands, a country with reinforcing social cleavages and very little social conflict, disproved this deterministic theory on the basis of a single case.[45] (One may dispute whether the original theory is correctly understood as deterministic. However, if it *is*, then it has been decisively refuted by a single case study.) *Proving* an invariant causal argument generally requires more cases. However, it is not nearly as complicated as proving a probabilistic argument for the simple reason that one assumes invariant relationships; consequently, the single case under study carries more weight.

Magnitude and consistency—the two components of causal strength—are usually matters of degree. It follows that the more tenuous the connection between X and Y, the more difficult it will be to address in a case study format. This is because the causal mechanisms connecting X with Yare less likely to be detectable in a single case when the total impact is slight or highly irregular. It is no surprise, therefore, that the case study research design has, from the very beginning, been associated with causal arguments that are deterministic, while cross-case research has been associated with causal arguments that are assumed to be minimal in strength and "probabilistic" in consistency.[46] (Strictly speaking, causal magnitude and consistency are independent features of a causal relationship. However, because they tend to covary, and because we tend to conceptualize them in tandem, I treat them as components of a single dimension.)

p. 111 Now, let us now consider an example drawn from the other extreme. There is generally assumed to be a weak relationship between regime type and economic performance. Democracy, if it has any effect on economic growth at all, probably has only a slight effect over the near-to-medium term, and this effect is probably characterized by many exceptions (cases that do not fit the general pattern). This is because many things other than democracy affect a country's growth performance and because there may be a significant stochastic component in economic growth (factors that cannot be modeled in a general way). Because of the diffuse nature of this relationship it will probably be difficult to gain insight by looking at a single case. Weak relationships are difficult to observe in one instance. Note that even if there seems to be a strong relationship between democracy and economic growth in a given country it may be questioned whether this case is actually typical of the larger population of interest, given that we have already stipulated that the typical magnitude of this relationship is diminutive and irregular. Of course, the weakness of democracy's presumed relationship to growth is also a handicap in cross-case analysis. A good deal of criticism has been directed toward studies of this type, where findings are rarely robust.[47] Even so, it seems clear that if there *is* a relationship between democracy and growth it is more likely to be perceptible in a large cross-case setting. The positive hypothesis, as well as the null hypothesis, is better approached in a sample rather than in a case.

9 Useful Variation: Rare versus Common

When analyzing causal relationships we must be concerned not only with the strength of an X/Y relationship but also with the distribution of evidence across available cases. Specifically, we must be concerned with the distribution of *useful variation*—understood as variation (temporal or spatial) on relevant parameters that might yield clues about a causal relationship. It follows that where useful variation is rare—i.e. limited to a few cases—the case study format recommends itself. Where, on the other hand, useful variation is common, a cross-case method of analysis may be more defensible.

Consider a phenomenon like social revolution, an outcome that occurs very rarely. The empirical distribution on this variable, if we count each country-year as an observation, consists of thousands of non-revolutions (0) and just a few revolutions (1). Intuitively, it seems clear that the few "revolutionary" cases are of great interest. We need to know as much as possible about them, for they exemplify all the variation
p. 112 that we have at our disposal. In this circumstance, a case study mode of analysis is ↳ difficult to avoid, though it might be combined with a large-N cross-case analysis. As it happens, many outcomes of interest to social scientists are quite rare, so the issue is by no means trivial.[48]

By way of contrast, consider a phenomenon like turnover, understood as a situation where a ruling party or coalition is voted out of office. Turnover occurs within most democratic countries on a regular basis, so the distribution of observations on this variable (incumbency/turnover) is relatively even across the universe of country-years. There are lots of instances of both outcomes. Under these circumstances a cross-case research design seems plausible, for the variation across cases is regularly distributed.

Another sort of variation concerns that which might occur *within* a given case. Suppose that only one or two cases within a large population exhibit quasi-experimental qualities: the factor of special interest varies, and there is no corresponding change in other factors that might affect the outcome. Clearly, we are likely to learn a great deal from studying this particular case—perhaps a lot more than we might learn from studying hundreds of additional cases that deviate from the experimental ideal. But again, if many cases have this experimental quality, there is little point in restricting ourselves to a single example; a cross-case research design may be justified.

A final sort of variation concerns the characteristics exhibited by a case relative to a particular theory that is under investigation. Suppose that a case provides a "crucial" test for a theory: it fits that theory's predictions so perfectly and so precisely that no other explanation could plausibly account for the performance of the case. If no other crucial cases present themselves, then an intensive study of this particular case is de rigueur. Of course, if many such cases lie within the population then it may be possible to study them all at once (with some sort of numeric reduction of the relevant parameters).

The general point here is that the distribution of useful variation across a population of cases matters a great deal in the choice between case study and cross-case research designs.

10 Data Availability: Concentrated versus Dispersed

I have left the most prosaic factor for last. Sometimes, one's choice of research design is driven by the
p. 113 quality and quantity of information that is currently available, or ↳ could be easily gathered, on a given question. This is a practical matter, and is distinct from the actual (ontological) shape of the world. It concerns, rather, what we know about the former at a given point in time.[49] The question of evidence may be posed as follows: How much do we know about the cases at hand that might be relevant to the causal question of interest, and how precise, certain, and case comparable is that data? An evidence-rich environment is one where all relevant factors are measurable, where these measurements are relatively precise, where they are rendered in comparable terms across cases, and where one can be relatively confident that the information is, indeed, accurate. An evidence-poor environment is the opposite.

The question of available evidence impinges upon choices in research design when one considers its distribution across a population of cases. If relevant information is concentrated in a single case, or if it is contained in incommensurable formats across a population of cases, then a case study mode of analysis is almost unavoidable. If, on the other hand, it is evenly distributed across the population—i.e. we are equally well informed about all cases—and is case comparable, then there is little to recommend a narrow focus. (I employ data, evidence, and information as synonyms in this section.)

Consider the simplest sort of example, where information is truly limited to one or a few cases. Accurate historical data on infant mortality and other indices of human development are currently available for only a handful of countries (these include Chile, Egypt, India, Jamaica, Mauritius, Sri Lanka, the United States, and several European countries).[50] This data problem is not likely to be rectified in future years, as it is exceedingly difficult to measure infant mortality except by public or private records. Consequently, anyone studying this general subject is likely to rely heavily on these cases, where in-depth analysis is possible and profitable. Indeed, it is not clear whether *any* large-N cross-case analysis is possible prior to the twentieth century. Here, a case study format is virtually prescribed, and a cross-case format proscribed.

Other problems of evidence are more subtle. Let us dwell for the moment on the question of data comparability. In their study of social security spending, Mulligan, Gil, and Sala-i-Martin note that

> although our spending and design numbers are of good quality, there are some missing observations and, even with all the observations, it is difficult to reduce the variety of elderly subsidies to one or two numbers. For this reason, case studies are an important part of our analysis, since those studies do not require numbers that are comparable across a large number of countries. Our case study analysis utilizes data from a variety of country-specific sources, so we do not have to reduce "social security" or "democracy" to one single number.[51]

Here, the incommensurability of the evidence militates towards a case study format. In the event that the authors (or subsequent analysts) discover a coding system that provides reasonably valid cross-case
p. 114 measures of social security, democracy, and ↳ other relevant concepts then our state of knowledge about the subject is changed, and a cross-case research design is rendered more plausible.

Importantly, the state of evidence on a topic is never entirely fixed. Investigators may gather additional data, recode existing data, or discover new repositories of data. Thus, when discussing the question of evidence one must consider the quality and quantity of evidence that *could* be gathered on a given question, given sufficient time and resources. Here it is appropriate to observe that collecting new data, and correcting existing data, is usually easier in a case study format than in a large-N cross-case format. It will be difficult to rectify data problems if one's cases number in the hundreds or thousands. There are simply too many data points to allow for this.

One might consider this issue in the context of recent work on democracy. There is general skepticism among scholars with respect to the viability of extant global indicators intended to capture this complex concept (e.g. by Freedom House and by the Polity IV data project).[52] Measurement error, aggregation problems, and questions of conceptual validity are rampant. When dealing with a single country or a single continent it is possible to overcome some of these faults by manually recoding the countries of interest.[53] The case study format often gives the researcher an opportunity to fact-check, to consult multiple sources, to go back to primary materials, and to overcome whatever biases may affect the secondary literature. Needless to say, this is not a feasible approach for an individual investigator if one's project encompasses every country in the world. The best one can usually manage, under the circumstances, is some form of convergent validation (by which different indices of the same concept are compared) or small adjustments in the coding intended to correct for aggregation problems or measurement error.[54]

For the same reason, the collection of original data is typically more difficult in cross-case analysis than in case study analysis, involving greater expense, greater difficulties in identifying and coding cases, learning foreign languages, traveling, and so forth. Whatever can be done for a set of cases can usually be done more easily for a single case.

It should be kept in mind that many of the countries of concern to anthropologists, economists, historians, political scientists, and sociologists are still terra incognita. Outside the OECD, and with the exception of a few large countries that have received careful attention from scholars (e.g. India, Brazil, China), most countries of the world are not well covered by the social science literature. Any statement that one might wish to make about, say, Botswana, will be difficult to verify if one has recourse only to secondary materials. And these—very limited—secondary sources are not necessarily of the most reliable sort. Thus, if one wishes to say something about political patterns obtaining in roughly 90 percent of the world's countries and if one wishes to go beyond matters that can be captured in standard statistics collected by the World
p. 115
Bank and the IMF and other agencies (and these can also be very sketchy ↳ when lesser-studied countries are concerned) one is more or less obliged to conduct a case study. Of course, one could, in principle, gather similar information across all relevant cases. However, such an enterprise faces formidable logistical difficulties. Thus, for practical reasons, case studies are sometimes the most defensible alternative when the researcher is faced with an information-poor environment.

However, this point is easily turned on its head. Datasets are now available to study many problems of concern to the social sciences. Thus, it may not be necessary to collect original information for one's book, article, or dissertation. Sometimes in-depth single-case analysis is more time consuming than cross-case analysis. If so, there is no informational advantage to a case study format. Indeed, it may be easier to utilize existing information for a cross-case analysis, particularly when a case study format imposes hurdles of its own—e.g. travel to distant climes, risk of personal injury, expense, and so forth. It is interesting to note that some observers consider case studies to be "relatively *more* expensive in time and resources."[55]

Whatever the specific logistical hurdles, it is a general truth that the shape of the evidence—that which is currently available and that which might feasibly be collected by an author—often has a strong influence on an investigator's choice of research designs. Where the evidence for particular cases is richer and more accurate there is a strong prima facie argument for a case study format focused on those cases. Where, by contrast, the relevant evidence is equally good for all potential cases, and is comparable across those cases, there is no reason to shy away from cross-case analysis. Indeed, there may be little to gain from case study formats.

11 Conclusions

At the outset, I took note of the severe disjuncture that has opened up between an often-maligned methodology and a heavily practiced method. The case study is disrespected but nonetheless regularly employed. Indeed, it remains the workhorse of most disciplines and subfields in the social sciences. How, then, can one make sense of this schizophrenia between methodological theory and praxis?

The torment of the case study begins with its definitional penumbra. Frequently, this key term is conflated with a set of disparate methodological traits that are not definitionally entailed. My first objective, therefore, was to craft a narrower and more useful concept for purposes of methodological discussion. The case study, I argued, is best defined as an intensive study of a single case with an aim to generalize across a larger set of cases. It follows from this definition that case studies may be small-or large-N, qualitative or quantitative, experimental or observational, synchronic or diachronic. It also follows that the case study research design comports with any ↳ macrotheoretical framework or paradigm—e.g. behavioralism,
p. 116 rational choice, institutionalism, or interpretivism. It is not epistemologically distinct. What differentiates the case study from the cross-case study is simply its way of defining observations, not its analysis of those observations or its method of modeling causal relations. The case study research design constructs its observations from a single case or a small number of cases, while cross-case research designs construct observations across multiple cases. Cross-case and case study research operate, for the most part, at different levels of analysis.

The travails of the case study are not simply definitional. They are also rooted in an insufficient appreciation of the methodological tradeoffs that this method calls forth. At least eight characteristic strengths and weaknesses must be considered. Ceteris paribus, case studies are more useful when the strategy of research is exploratory rather than confirmatory/disconfirmatory, when internal validity is given preference over external validity, when insight into causal mechanisms is prioritized over insight into causal effects, when propositional depth is prized over breadth, when the population of interest is heterogeneous rather than homogeneous, when causal relationships are strong rather than weak, when useful information about key parameters is available only for a few cases, and when the available data are concentrated rather than dispersed.

Although I do not have the space to discuss other issues in this venue, it is worth mentioning that other considerations may also come into play in a researcher's choice between a case study and cross-case study research format. However, these additional issues—e.g. causal complexity and the state of research on a topic—do not appear to have clear methodological affinities. They may augur one way, or the other.

My objective throughout this chapter is to restore a greater sense of meaning, purpose, and integrity to the case study method. It is hoped that by offering a narrower and more carefully bounded definition of this method the case study may be rescued from some of its most persistent ambiguities. And it is hoped that the characteristic strengths of this method, as well as its limitations, will be more apparent to producers and consumers of case study research. The case study is a useful tool for some research objectives, but not all.

References

Abadie, A. and Gardeazabal, J. 2003. The economic costs of conflict: a case study of the asque Country. *American Economic Review*, 93: 113–32. 10.1257/000282803321455188
Google Scholar WorldCat Crossref

Abbott, A. 1990. Conceptions of time and events in social science methods: causal and narrative approaches. *Historical Methods*, 23 (4): 140–50.
Google Scholar WorldCat

—— 1992. From causes to events: notes on narrative positivism. *Sociological Methods and Research*, 20 (4): 428–55. 10.1177/0049124192020004002
WorldCat Crossref

—— 2001. *Time Matters: On Theory and Method*. Chicago: University of Chicago Press.
Google Scholar Google Preview WorldCat COPAC

p. 117 —— and Forrest, J. 1986. Optimal matching methods for historical sequences. *Journal of Interdisciplinary History*, 16 (3): 471–94. 10.2307/204500
Google Scholar WorldCat Crossref

—— and Tsay, A. 2000. Sequence analysis and optimal matching methods in sociology. *Sociological Methods and Research*, 29: 3–33. 10.1177/0049124100029001001
Google Scholar WorldCat Crossref

Abell, P. 1987. *The Syntax of Social Life: The Theory and Method of Comparative Narratives*. Oxford: Clarendon Press.
Google Scholar Google Preview WorldCat COPAC

—— 2004. Narrative explanation: an alternative to variable-centered explanation? *Annual Review of Sociology*, 30: 287–310. 10.1146/annurev.soc.29.010202.100113
WorldCat Crossref

Acemoglu, D. Johnson, S. and Robinson, J. A. 2003. An African success story: Botswana. Pp. 80–122 in *In Search of Prosperity: Analytic Narratives on Economic Growth*, ed. D. Rodrik. Princeton: Princeton University Press.
Google Scholar Google Preview WorldCat COPAC

Achen, C. H. 1986. *The Statistical Analysis of Quasi-Experiments*. Berkeley and Los Angeles: University of California Press.
Google Scholar Google Preview WorldCat COPAC

—— 2002. Toward a new political methodology: microfoundations and ART. *Annual Review of Political Science*, 5: 423–50. 10.1146/annurev.polisci.5.112801.080943
WorldCat Crossref

—— and Snidal, D. 1989. Rational deterrence theory and comparative case studies. *World Politics*, 41: 143–69. 10.2307/2010405
Google Scholar WorldCat Crossref

Alesina, A. Glaeser, E. and Sacerdote, B. 2001. Why doesn't the US have a European-style welfare state? *Brookings Papers on Economic Activity*, 2: 187–277.
Google Scholar WorldCat

Allen, W. S. 1965. *The Nazi Seizure of Power: The Experience of a Single German Town, 1930–1935*. New York: Watts.
Google Scholar Google Preview WorldCat COPAC

Almond, G. A. 1956. Comparative political systems. *Journal of Politics*, 18: 391–409. 10.2307/2127255
Google Scholar WorldCat Crossref

Angrist, J. D. and Krueger, A. B. 2001. Instrumental variables and the search for identification: from supply and demand to natural experiments. *Journal of Economic Perspectives*, 15 (4): 69–85. 10.1257/jep.15.4.69
Google Scholar WorldCat Crossref

Bates, R. H. Greif, A. Levi, M. Rosenthal, J.-L. and Weingast, B. 1998. *Analytic Narratives*. Princeton: Princeton University Press.
Google Scholar Google Preview WorldCat COPAC

Bendix, R. 1963. Concepts and generalizations in comparative sociological studies. *American Sociological Review*, 28: 532–9. 10.2307/2090069
Google Scholar WorldCat Crossref

Bentley, A. 1908/1967. *The Process of Government*. Cambridge, Mass.: Harvard University Press.
Google Scholar Google Preview WorldCat COPAC

Blumer, H. 1969. *Symbolic Interactionism: Perspective and Method*. Berkeley and Los Angeles: University of California Press.
Google Scholar Google Preview WorldCat COPAC

Bollen, K. A. 1993. Liberal democracy: validity and method factors in cross-national meas ures. *American Journal of Political Science*, 37: 1207–30. 10.2307/2111550
Google Scholar WorldCat Crossref

Bonoma, T. V. 1985. Case research in marketing: opportunities, problems, and a process. *Journal of Marketing Research*, 22 (2): 199–208. 10.2307/3151365
Google Scholar WorldCat Crossref

Bowman, K. Lehoucq, F. and Mahoney, J. 2005. Measuring political democracy: case expertise, data adequacy, and Central America. *Comparative Political Studies*, 38 (8): 939–70. 10.1177/0010414005277083
Google Scholar WorldCat Crossref

Brady, H. E. and Collier, D. eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, Md.: Rowman & Littlefield.
Google Scholar Google Preview WorldCat COPAC

Braumoeller, B. F. and Goertz, G. 2000. The methodology of necessary conditions. *American Journal of Political Science*, 44 (3): 844–58. 10.2307/2669285
Google Scholar WorldCat Crossref

Bunge, M. 1997. Mechanism and explanation. *Philosophy of the Social Sciences*, 27: 410–65. 10.1177/004839319702700402
Google Scholar WorldCat Crossref

Buthe, T. 2002. Taking temporality seriously: modeling history and the use of narratives as evidence. *American Political Science Review*, 96 (3): 481–93.
Google Scholar WorldCat

Cameron, D. 1978. The expansion of the public economy: a comparative analysis. *American Political Science Review*, 72 (4): 1243–61. 10.2307/1954537
Google Scholar WorldCat Crossref

p. 118 Campbell, D. T. 1988. *Methodology and Epistemology for Social Science*, ed. E. S. Overman. Chicago: University of Chicago Press.
Google Scholar Google Preview WorldCat COPAC

——and Stanley, J. 1963. *Experimental and Quasi-experimental Designs for Research*. Boston: Houghton Mifflin.
Google Scholar Google Preview WorldCat COPAC

Chandra, K. 2004. *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India*. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Chernoff, B. and Warner, A. 2002. *Sources of fast growth in Mauritius: 1960–2000.* Center for International Development, Harvard University.
Google Scholar Google Preview WorldCat COPAC

Chong, D. 1993. How people think, reason, and feel about rights and liberties. *American Journal of Political Science*, 37 (3): 867–99. 10.2307/2111577
Google Scholar WorldCat Crossref

Coase, R. H. 1959. The Federal Communications Commission. *Journal of Law and Economics*, 2: 1–40. 10.1086/466549
Google Scholar WorldCat Crossref

——2000. The acquisition of Fisher Body by General Motors. *Journal of Law and Economics*, 43: 15–31. 10.1086/467446
WorldCat Crossref

Collier, D. 1993. The comparative method. Pp. 105–19 in *Political Science: The State of the Discipline II*, ed. A. W. Finifter. Washington, DC: American Political Science Association.
Google Scholar Google Preview WorldCat COPAC

Cook, T. and Campbell, D. 1979. *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin.
Google Scholar Google Preview WorldCat COPAC

De Soto, H. 1989. *The Other Path: The Invisible Revolution in the Third World*. New York: Harper & Row.
Google Scholar Google Preview WorldCat COPAC

Dessler, D. 1991. Beyond correlations: toward a causal theory of war. *International Studies Quarterly*, 35: 337–55. 10.2307/2600703
Google Scholar WorldCat Crossref

Dion, D. 1998. Evidence and inference in the comparative case study. *Comparative Politics*, 30: 127–45. 10.2307/422284
Google Scholar WorldCat Crossref

Eckstein, H. 1975. Case studies and theory in political science. Pp. 79–133 in *Handbook of Political Science*, vii: *Political Science: Scope and Theory*, ed. F. I. Greenstein and N. W. Polsby. Reading, Mass.: Addison-Wesley.

—— 1975/1992. Case studies and theory in political science. In *Regarding Politics: Essays on Political Theory, Stability, and Change*, by H. Eckstein. Berkeley and Los Angeles: University of California Press.
Google Scholar Google Preview WorldCat COPAC

Elman, C. 2005. Explanatory typologies in qualitative studies of international politics. *International Organization*, 59 (2): 293–326.
Google Scholar WorldCat

Elster, J. 1998. A plea for mechanisms. Pp. 45–73 in *Social Mechanisms: An Analytical Approach to Social Theory*, ed. P. Hedstrom and R. Swedberg. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Fearon, J. 1991. Counter factuals and hypothesis testing in political science. *World Politics*, 43: 169–95. 10.2307/2010470
Google Scholar WorldCat Crossref

Feng, Y. 2003. *Democracy, Governance, and Economic Performance: Theory and Evidence*. Cambridge, Mass.: MIT Press.
Google Scholar Google Preview WorldCat COPAC

Freedman, D. A. 1991. Statistical models and shoe leather. *Sociological Methodology*, 21: 291–313. 10.2307/270939
Google Scholar WorldCat Crossref

Friedman, M. 1953. The methodology of positive economics. Pp. 3–43 in *Essays in Positive Economics*, by M. Friedman. Chicago: University of Chicago Press.
Google Scholar Google Preview WorldCat COPAC

Geddes, B. 1990. How the cases you choose affect the answers you get: selection bias in comparative politics. Pp. 131–52 in *Political Analysis*, vol. ii, ed. J. A. Stimson. Ann Arbor: University of Michigan Press.
Google Scholar Google Preview WorldCat COPAC

—— 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor: University of Michigan Press.
Google Scholar Google Preview WorldCat COPAC

Geertz, C. 1973. Thick description: toward an interpretive theory of culture. 3 3–30 in *The Interpretation of Cultures*, by C. Geertz. New York: Basic Books.
Google Scholar Google Preview WorldCat COPAC

p. 119 George, A. L. and Bennett, A. 2005. *Case Studies and Theory Development*. Cambridge, Mass.: MIT Press.
Google Scholar Google Preview WorldCat COPAC

—— and Smoke, R. 1974. *Deterrence in American Foreign Policy: Theory and Practice*. New York: Columbia University Press.
Google Scholar Google Preview WorldCat COPAC

Gerring, J. 2001. *Social Science Methodology: A Criterial Framework*. Cambridge: Cambridge University Press.

Google Scholar Google Preview WorldCat COPAC

——2005. Causation: a unified framework for the social sciences. *Journal of Theoretical Politics*, 17 (2): 163–98. 10.1177/0951629805050859
WorldCat Crossref

——2007*a*. *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

——2007*b*. Global justice as an empirical question. *PS: Political Science and Politics* (forth coming).

——and Barresi, P. A. 2003. Putting ordinary language to work: a min-max strategy of concept formation in the social sciences. *Journal of Theoretical Politics*, 15 (2): 201–32. 10.1177/0951629803015002647
Google Scholar WorldCat Crossref

——and Thomas, C. 2005. Comparability: a key issue in research design. MS.

Glaser, B. G. and Strauss, A. L. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine de Gruyter.
Google Scholar Google Preview WorldCat COPAC

Goertz, G. 2003. The substantive importance of necessary condition hypotheses. Ch. 4 in *Necessary Conditions: Theory, Methodology and Applications*, ed. G. Goertz and H. Starr. New York: Rowman and Littlefield.
Google Scholar Google Preview WorldCat COPAC

——and Levy, J. eds. Forthcoming. Causal explanations, necessary conditions, and case studies: World War I and the end of the Cold War. MS.

——and Starr, H. eds. 2003. *Necessary Conditions: Theory, Methodology and Applications*. New York: Rowman and Littlefield.
Google Scholar Google Preview WorldCat COPAC

Goldstone, J. A. 1997. Methodological issues in comparative macrosociology. *Comparative Social Research*, 16: 121–32.
Google Scholar WorldCat

——Gurr, T. R. Harff, B. Levy, M. A. Marshall, M. G. Bates, R. H. Epstein, D. L. Kahl, C. H. Surko, P. T. Ulfelder, J. C., Jr. and Unger, A. N. 2000. State Failure Task Force report: phase III Wndings. Available at www.cidcm.umd.edu/inscr/stfail/SFTF%20Phase%20III%20Report%20Final.pdf
WorldCat

Goldthorpe, J. H. 1997. Current issues in comparative macrosociology: a debate on meth odological issues. *Comparative Social Research*, 16: 121–32.
Google Scholar WorldCat

Griffin, L. J. 1993. Narrative, event-structure analysis, and causal interpretation in historical sociology. *American Journal of Sociology*, 98: 1094–133. 10.1086/230140
Google Scholar WorldCat Crossref

Gutting, G. ed. 1980. *Paradigms and Revolutions: Appraisals and Applications of Thomas Kuhn's Philosophy of Science*. Notre Dame, Ind.: University of Notre Dame Press.
Google Scholar Google Preview WorldCat COPAC

Hall, P. A. 2003. Aligning ontology and methodology in comparative politics. In *Comparative Historical Analysis in the Social Sciences*, ed. J. Mahoney and D. Rueschemeyer. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Hedstrom, P. and Swedberg, R. eds. 1998. *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Hersen, M. and Barlow, D. H. 1976. *Single-Case Experimental Designs: Strategies for Studying Behavior Change*. Oxford: Pergamon Press.
Google Scholar Google Preview WorldCat COPAC

Hochschild, J. L. 1981. *What's Fair? American Beliefs about Distributive Justice*. Cambridge, Mass.: Harvard University Press.

Google Scholar Google Preview WorldCat COPAC

Jervis, R. 1989. Rational deterrence: theory and evidence. *World Politics*, 41 (2): 183–207. 10.2307/2010407
Google Scholar WorldCat Crossref

Kennedy, P. 2003. *A Guide to Econometrics*, 5th edn. Cambridge, Mass.: MIT Press.
Google Scholar Google Preview WorldCat COPAC

King, C. 2004. The micropolitics of social violence. *World Politics*, 56 (3): 431–55. 10.1353/wp.2004.0016
Google Scholar WorldCat Crossref

p. 120 King, G. Keohane, R. O. and Verba, S. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
Google Scholar Google Preview WorldCat COPAC

Kittel, B. 1999. Sense and sensitivity in pooled analysis of political data. *European Journal of Political Research*, 35: 225–53. 10.1111/1475-6765.00448
Google Scholar WorldCat Crossref

——2005. A crazy methodology? On the limits of macroquantitative social science research. Unpublished MS. University of Amsterdam.

Kittel, B., and Winner, H. 2005. How reliable is pooled analysis in political economy? The globalization-welfare state nexus revisited. *European Journal of Political Research*, 44 (2): 269–93. 10.1111/j.1475-6765.2005.00228.x
Google Scholar WorldCat Crossref

Kuhn, T. S. 1962/1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
Google Scholar Google Preview WorldCat COPAC

Lane, R. 1962. *Political Ideology: Why the American Common Man Believes What He Does*. New York: Free Press.
Google Scholar Google Preview WorldCat COPAC

Lebow, R. N. 2000. What's so different about a counterfactual? *World Politics*, 52: 550–85.
Google Scholar WorldCat

Levine, R., and Renelt, D. 1992. A sensitivity analysis of cross-country growth regressions. *American Economic Review*, 82 (4): 942–63.
Google Scholar WorldCat

Libecap, G. D. 1993. *Contracting for Property Rights*. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Lieberson, S. 1985. *Making it Count: The Improvement of Social Research and Theory*. Berkeley and Los Angeles: University of California Press.
Google Scholar Google Preview WorldCat COPAC

——1992. Einstein, Renoir, and Greeley: some thoughts about evidence in sociology: 1991 Presidential Address. *American Sociological Review*, 57 (1): 1–15. 10.2307/2096141
WorldCat Crossref

——1994. More on the uneasy case for using Mill-type methods in small-N comparative studies. *Social Forces*, 72 (4): 1225–37. 10.2307/2580300
WorldCat Crossref

Lijphart, A. 1968. *The Politics of Accommodation: Pluralism and Democracy in the Nether lands*. Berkeley and Los Angeles: University of California Press.
Google Scholar Google Preview WorldCat COPAC

——1969. Consociational democracy. *World Politics*, 21 (2): 207–25. 10.2307/2009820
WorldCat Crossref

——1971. Comparative politics and the comparative method. *American Political Science Review*, 65 (3): 682–93. 10.2307/1955513
WorldCat Crossref

Lipset, S. M. 1960/1963. *Political Man: The Social Bases of Politics*. Garden City, NY: Anchor Books.
Google Scholar Google Preview WorldCat COPAC

—— Trow, M. A., and Coleman, J. S. 1956. *Union Democracy: The Internal Politics of the International Typographical Union*. New York: Free Press.
Google Scholar Google Preview WorldCat COPAC

Little, D. 1998. *Microfoundations, Method, and Causation*. New Brunswick, NJ: Transaction.
Google Scholar Google Preview WorldCat COPAC

Lynd, R. S. and Lynd, H. M. 1929/1956. *Middletown: A Study in American Culture*. New York: Harcourt, Brace.
Google Scholar Google Preview WorldCat COPAC

Mc Keown, T. J. 1983. Hegemonic stability theory and nineteenth-century tariff levels. *Inter national Organization*, 37 (1): 73–91. 10.1017/S0020818300004203
Google Scholar WorldCat Crossref

Mahoney, J. 2001. Beyond correlational analysis: recent innovations in theory and method. *Sociological Forum*, 16 (3): 575–93. 10.1023/A:1011912816997
Google Scholar WorldCat Crossref

—— and Rueschemeyer, D. eds. 2003. *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

—— and Goertz, G. 2004. The possibility principle: choosing negative cases in comparative research. *American Political Science Review*, 98 (4): 653–69.
Google Scholar WorldCat

Manski, C. F. 1993. Identification problems in the social sciences. *Sociological Methodology*, 23: 1–56. 10.2307/271005
Google Scholar WorldCat Crossref

Martin, C. J., and Swank, D. 2004. Does the organization of capital matter? Employers and active labor market policy at the national and firm levels. *American Political Science Review*, 98 (4): 593–612.
Google Scholar WorldCat

p. 121 Martin, L. L. 1992. *Coercive Cooperation: Explaining Multilateral Economic Sanctions*. Princeton: Princeton University Press.
Google Scholar Google Preview WorldCat COPAC

Meehl, P. E. 1954. *Clinical versus Statistical Predictions: ATheoretical Analysis and a Review of the Evidence*. Minneapolis: University of Minnesota Press. 10.1037/11281-000
Google Scholar Google Preview WorldCat COPAC Crossref

Mulligan, C. Gil, R., and Sala-i-Martin, X. 2002. *Social security and democracy.* MS. University of Chicago and Columbia University.
Google Scholar Google Preview WorldCat COPAC

Munck, G. L., and Snyder, R. eds. 2007. *Passion, Craft, and Method in Comparative Politics*. Baltimore: Johns Hopkins University Press.
Google Scholar Google Preview WorldCat COPAC

—— and Verkuilen, J. 2002. Measuring democracy: evaluating alternative indices. *Com parative Political Studies*, 35 (1): 5–34.
Google Scholar WorldCat

Njolstad, O. 1990. Learning from history? Case studies and the limits to theory-building. Pp. 220–46 in *Arms Races: Technological and Political Dynamics*, ed. O. Njolstad. Thousand Oaks, Calif.: Sage.
Google Scholar Google Preview WorldCat COPAC

North, D. C., Anderson, T. L. and Hill, P. J. 1983. *Growth and Welfare in the American Past: A New American History*, 3rd edn. Englewood Cliffs, NJ: Prentice-Hall.
Google Scholar Google Preview WorldCat COPAC

—— and Thomas, R. P. 1973. *The Rise of the Western World*. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

——and Weingast, B. R. 1989. Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century England. *Journal of Economic History*, 49: 803–32. 10.1017/S0022050700009451
Google Scholar WorldCat Crossref

Odell, J. S. 2004. Case study methods in international political economy. Pp. 56–80 in *Models, Numbers and Cases: Methods for Studying International Relations*, ed. D. F. Sprinz and Y. Wolinsky-Nahmias. Ann Arbor: University of Michigan.
Google Scholar Google Preview WorldCat COPAC

Orum, A. M. Feagin, J. R., and Sjoberg, G. 1991. Introduction: the nature of the case study. Pp. 1–26 in *A Case for the Case*, ed. J. R. Feagin A. M. Orum and G. Sjoberg. Chapel Hill: University of North Carolina Press.
Google Scholar Google Preview WorldCat COPAC

Papyrakis, E. and Gerlagh, R. 2003. The resource curse hypothesis and its transmission channels. *Journal of Comparative Economics*, 32: 181–93. 10.1016/j.jce.2003.11.002
Google Scholar WorldCat Crossref

Patton, M. Q. 2002. *Qualitative Evaluation and Research Methods*. Newbury Park, Calif.: Sage.
Google Scholar Google Preview WorldCat COPAC

Popper, K. 1934/1968. *The Logic of Scientific Discovery*. New York: Harper & Row.
Google Scholar Google Preview WorldCat COPAC

——1969. *Conjectures and Refutations*. London: Routledge and Kegan Paul.
Google Scholar Google Preview WorldCat COPAC

Posner, D. 2004. The political salience of cultural difference: why Chewas and Tumbukas are allies in Zambia and adversaries in Malawi. *American Political Science Review*, 98 (4): 529–46.
Google Scholar WorldCat

Przeworski, A. and Teune, H. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley.
Google Scholar Google Preview WorldCat COPAC

Ragin, C. C. 1987. *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley and Los Angeles: University of California Press.
Google Scholar Google Preview WorldCat COPAC

——1992. Cases of “what is a case?” Pp. 1–17 in *What Is a Case? Exploring the Foundations of Social Inquiry*, ed. C. C. Ragin and H. S. Becker. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

——1997. Turning the tables: how case-oriented research challenges variable-oriented re search. *Comparative Social Research*, 16: 27–42.
WorldCat

——2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.
Google Scholar Google Preview WorldCat COPAC

——2004. Turning the tables. Pp. 123–38 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, ed. H. E. Brady and D. Collier. Lanham, Md.: Rowman & Littlefield.
Google Scholar Google Preview WorldCat COPAC

Robinson, W. S. 1951. The logical structure of analytic induction. *American Sociological Review*, 16 (6): 812–18. 10.2307/2087508
Google Scholar WorldCat Crossref

p. 122 Rodrik, D. ed. 2003. *In Search of Prosperity: Analytic Narratives on Economic Growth*. Princeton: Princeton University Press.
Google Scholar Google Preview WorldCat COPAC

Rogowski, R. 1995. The role of theory and anomaly in social-scientific inference. *American Political Science Review*, 89 (2): 467–70. 10.2307/2082443
Google Scholar WorldCat Crossref

Ross, M. 2001. Does oil hinder democracy? *World Politics*, 53: 325–61. 10.1353/wp.2001.0011
Google Scholar WorldCat Crossref

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66: 688–701. 10.1037/h0037350
Google Scholar WorldCat Crossref

Rueschemeyer, D., and Stephens, J. D. 1997. Comparing historical sequences: a powerful tool for causal analysis. *Comparative Social Research*, 16: 55–72.
Google Scholar WorldCat

Sambanis, N. 2004. Using case studies to expand economic models of civil war. *Perspectives on Politics*, 2 (2): 259–79.
Google Scholar WorldCat

Sartori, G. 1976. *Parties and Party Systems*. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Sekhon, J. S. 2004. Quality meets quantity: case studies, conditional probability and counter- factuals. *Perspectives in Politics*, 2 (2): 281–93.
Google Scholar WorldCat

Shalev, M. 1998. Limits of and alternatives to multiple regression in macro-comparative research. Paper prepared for presentation at the second conference on The Welfare State at the Crossroads, Stockholm.

Smelser, N. J. 1973. The methodology of comparative analysis. Pp. 42–86 in *Comparative Research Methods*, ed. D. P. Warwick and S. Osherson. Englewood Cliffs, NJ: Prentice-Hall.
Google Scholar Google Preview WorldCat COPAC

Srinivasan, T. N., and Bhagwati, J. 1999. *Outward-orientation and development: are revisionists right?* Discussion Paper no. 806, Economic Growth Center, Yale University.
Google Scholar Google Preview WorldCat COPAC

Stoecker, R. 1991. Evaluating and rethinking the case study. *Sociological Review*, 39: 88–112.
Google Scholar WorldCat

Symposium: Qualitative Comparative Analysis (QCA). 2004. *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods*, 1 (2): 2–25.

Temple, J. 1999. The new growth evidence. *Journal of Economic Literature*, 37: 112–56.
Google Scholar WorldCat

Tetlock, P. E., and Belkin, A. eds. 1996. *Counterfactual Thought Experiments in World Politics*. Princeton: Princeton University Press.
Google Scholar Google Preview WorldCat COPAC

Thies, M. F. 2001. Keeping tabs on partners: the logic of delegation in coalition governments. *American Journal of Political Science*, 45 (3): 580–98. 10.2307/2669240
Google Scholar WorldCat Crossref

Tilly, C. 2001. Mechanisms in political processes. *Annual Review of Political Science*, 4: 21–41. 10.1146/annurev.polisci.4.1.21
Google Scholar WorldCat Crossref

Treier, S., and Jackman, S. 2005. *Democracy as a latent variable.* Department of Political Science, Stanford University.
Google Scholar Google Preview WorldCat COPAC

Truman, D. B. 1951. *The Governmental Process*. New York: Alfred A. Knopf.
Google Scholar Google Preview WorldCat COPAC

Vandenbroucke, J. P. 2001. In defense of case reports and case series. *Annals of Internal Medicine*, 134 (4): 330–4.
Google Scholar WorldCat

Vreeland, J. R. 2003. *The IMF and Economic Development*. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Ward, M. D. and Bakke, K. 2005. *Predicting civil conflicts: on the utility of empirical research.* MS.
Google Scholar Google Preview WorldCat COPAC

Winship, C. and Morgan, S. L. 1999. The estimation of causal effects of observational data. *Annual Review of Sociology*, 25: 659–707. 10.1146/annurev.soc.25.1.659
Google Scholar WorldCat Crossref

—— and Sobel, M. 2004. Causal inference in sociological studies. Pp. 481–503 in *Handbook of Data Analysis*, ed. M. Hardy and A. Bryman. London: Sage.
Google Scholar Google Preview WorldCat COPAC

Wolin, S. S. 1968. Paradigms and political theories. Pp. 125–52 in *Politics and Experience*, ed. P. King and B. C. Parekh. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Young, O. R. ed. 1999. *The Effectiveness of International Environmental Regimes: Causal Connections and Behavioral Mechanisms*. Cambridge, Mass.: MIT Press.
Google Scholar Google Preview WorldCat COPAC

Znaniecki, F. 1934. *The Method of Sociology*. New York: Rinehart.
Google Scholar Google Preview WorldCat COPAC

Notes

1 Acemoglu, Johnson, and Robinson (2003), Chernoff and Warner (2002), Rodrik (2003). See also studies focused on particular firms or regions, e.g. Coase 1959, 2000.

2 For general discussion of the following points see Achen (1986), Freedman (1991), Kittel (1999, 2005), Kittel and Winner (2005), Manski (1993), Winship and Morgan (1999), Winship and Sobel (2004).

3 Achen and Snidal (1989: 160). See also Geddes (1990, 2003), Goldthorpe (1997), King, Keohane, and Verba (1994), Lieberson (1985: 107–15, 1992, 1994), Lijphart (1971: 683–4), Odell (2004), Sekhon (2004), Smelser (1973: 45, 57). It should be noted that these writers, while critical of the case study format, are not necessarily opposed to case studies per se (that is to say, they should not be classified as *opponents* of the case study).

4 My intention is to include only those attributes commonly associated with the case study method that are *always* implied by our use of the term, excluding those attributes that are sometimes violated by standard usage. Thus, I chose not to include "ethnography" as a defining feature of the case study, since many case studies (so called) are not ethnographic. For further discussion of minimal definitions see Gerring (2001, ch. 4), Gerring and Barresi (2003), Sartori (1976).

5 These additional attributes might also be understood as comprising an ideal-type ("maximal") definition of the topic (Gerring 2001, ch. 4; Gerring and Barresi 2003).

6 Popper (1969).

7 Karl Popper (quoted in King, Keohane, and Verba 1994, 14) writes: "there is no such thing as a logical method of having new ideas … Discovery contains 'an irrational element,' or a 'creative intuition.'" One recent collection of essays and interviews takes new ideas as its special focus (Munck and Snyder 2007), though it may be doubted whether there are generalizable results.

8 Gerring (2001, ch. 10). The tradeoff between these two styles of research is implicit in Achen and Snidal (1989), who criticize the case study for its deficits in the latter genre but also acknowledge the benefits of the case study along the former dimension (1989, 167–8). Reichenbach also distinguished between a "context of discovery," and a "context of justification." Likewise, Peirce's concept of *abduction* recognizes the importance of a generative component in science.

9 Bonoma (1985: 199). Some of the following examples are discussed in Patton (2002, 245).

10 North and Weingast (1989); North and Thomas (1973).

11 Vandenbroucke (2001, 331).

12 For discussion of this tradeoff in the context of economic growth theory see Temple (1999, 120).

13 Geddes (2003), King, Keohane, and Verba (1994), Popper (1934/1968).

14 Ragin (1992).

15 Eckstein (1975), Ragin (1992, 1997), Rueschemeyer and Stephens (1997).

16 Eckstein (1975).

17 Campbell and Stanley (1963: 3).

18 Lane (1962).

19 Lynd and Lynd (1929/1956).

20 Note that the intensive study of a single unit may be a perfectly appropriate way to estimate causal effects *within that unit*. Thus, if one is interested in the relationship between welfare benefits and work effort in the United States one might obtain a more accurate assessment by examining data drawn from the USA alone, rather than crossnationally. However, since the resulting generalization does not extend beyond the unit in question it is not a case study in the usual sense.

21 Achen (2002), Dessler (1991), Elster (1998), George and Bennett (2005), Gerring (2005), Hedstrom and Swedberg (1998), Mahoney (2001), Tilly (2001).

22 In a discussion of instrumental variables in two-stage least-squares analysis, Angrist and Krueger (2001: 8) note that "good instruments often come from detailed knowledge of the economic mechanism, institutions determining the regressor of

interest."

23 Goldstone et al. (2000).

24 This has something to do with the existence of process-tracing evidence, a matter discussed below. But it is not necessarily predicated on this sort of evidence. Sensitive time-series data, another specialty of the case study, is also relevant to the question of causal mechanisms.

25 Glaser and Strauss (1967, 40).

26 Chong (1993, 868). For other examples of in-depth interviewing see Hochschild (1981), Lane (1962).

27 Rueschemeyer and Stephens (1997, 62).

28 Other good examples of within-case research that shed light on a broader theory can be found in Martin (1992); Martin and Swank (2004); Thies (2001); Young (1999).

29 Cameron (1978).

30 Alesina, Glaeser, and Sacerdote (2001).

31 For additional examples of this nature, see Feng (2003); Papyrakis and Gerlagh (2003); Ross (2001).

32 Eckstein (1975, 122).

33 I am using the term "thick" in a somewhat different way than in Geertz (1973).

34 See Ragin (2000, 22).

35 Ragin (1987, ch. 2). Herbert Blumer's (1969, ch 7) complaints, however, are more far-reaching.

36 Orum, Feagin, and Sjoberg (1991, 7).

37 Ragin (2000: 35). See also Abbott (1990); Bendix (1963); Meehl (1954); Przeworski and Teune (1970, 8–9); Ragin (1987; 2004, 124); Znaniecki (1934, 250–1).

38 George and Smoke (1974, 514).

39 Hersen and Barlow (1976, 11).

40 Shalev (1998).

41 To be sure, if adjacent cases are *identical*, the phenomenon of interest is *invariant* then the researcher gains nothing at all by studying more examples of a phenomenon, for the results obtained with the first case will simply be replicated. However, virtually all phenomena of interest to social scientists have some degree of heterogeneity (cases are not identical), some stochastic element. Thus, the theoretical possibility of identical, invariant cases is rarely met in practice.

42 Gutting (1980); Hall (2003); Kuhn (1962/1970); Wolin (1968).

43 Dion (1998).

44 Almond (1956); Bentley (1908/1967); Lipset (1960/1963); Truman (1951).

45 Lijphart (1968); see also Lijphart (1969). For additional examples of case studies disconfirming general propositions of a deterministic nature see Allen (1965); Lipset, Trow, and Coleman (1956); Njolstad (1990); discussion in Rogowski (1995).

46 Znaniecki (1934). See also discussion in Robinson (1951).

47 Kittel (1999, 2005); Kittel and Winner (2005); Levine and Renelt (1992); Temple (1999).

48 Consider the following topics and their—extremely rare—instances of variation: early industrialization (England, the Netherlands), fascism (Germany, Italy), the use of nuclear weapons (United States), world war (WWI, WWII), single non-transferable vote electoral systems (Jordan, Taiwan, Vanuatu, pre-reform Japan), electoral system reforms within established democracies (France, Italy, Japan, New Zealand, Thailand). The problem of "rareness" is less common where parameters are scalar, rather than dichotomous. But there are still plenty of examples of phenomena whose distributions are skewed by a few outliers, e.g. population (China, India), personal wealth (Bill Gates, Warren Buffett), ethnic heterogeneity (Papua New Guinea).

49 Of course, what we know about the potential cases is not independent of the underlying reality; it is, nonetheless, not entirely dependent on that reality.

50 Gerring (2007b).

51 Mulligan, Gil, and Sala-i-Martin (2002, 13).

52 Bollen (1993); Bowman, Lehoucq, and Mahoney (2005); Munck and Verkuilen (2002); Treier and Jackman (2005).

53 Bowman, Lehoucq, and Mahoney (2005).

54 Bollen (1993); Treier and Jackman (2005).

55 Stoecker (1991, 91).