What Is a Case Study and What Is It Good for?

JOHN GERRING *Boston University*

This paper aims to clarify the meaning, and explain the utility, of the case study method, a method often practiced but little understood. A "case study," I argue, is best defined as an intensive study of a single unit with an aim to generalize across a larger set of units. Case studies rely on the same sort of covariational evidence utilized in non-case study research. Thus, the case study method is correctly understood as a particular way of defining cases, not a way of analyzing cases or a way of modeling causal relations. I show that this understanding of the subject illuminates some of the persistent ambiguities of case study work, ambiguities that are, to some extent, intrinsic to the enterprise. The travails of the case study within the discipline of political science are also rooted in an insufficient appreciation of the methodological tradeoffs *that this method calls forth. This paper presents the familiar contrast between case study and non-case study work as a series of characteristic strengths and weaknesses*—affinities—*rather than as antagonistic approaches to the empirical world. In the end, the perceived hostility between case study and non-case study research is largely unjustified and, perhaps, deserves to be regarded as a misconception. Indeed, the strongest conclusion to arise from this methodological examination concerns the complementarity of single-unit and cross-unit research designs.*

The case study occupies a vexed position in the discipline of political science. On the one hand, methodologists generally view the case study method with extreme circumspection (Achen and Snidal 1989; King, Keohane, and Verba 1994; Lieberson [1991] 1992, 1994; Njolstad 1990). A work that focuses its attention on a single example of a broader phenomenon is apt to be described as a "mere" case study.

At the same time, the discipline continues to produce a vast number of case studies, many of which have entered the pantheon of classic works (Allen 1965; Allison 1971; Dahl 1960; Johnson 1983; Kaufman 1960; Lazarsfeld, Berelson, and Gaudet 1948; Lijphart 1968; Pressman and Wildavsky 1973). Judging by recent scholarly output, the case study method retains considerable appeal, even among scholars in research communities not traditionally associated with this style of research—e.g., among political economists and quantitatively inclined political scientists (Acemoglu, Johnson, and Robinson 2003; Bates et al. 1998; Rodrik 2003). By the standard of praxis, therefore, it would appear that the method of the case study is solidly ensconced and, perhaps, even thriving.

Thus, a paradox: Although much of what we know about the empirical world is drawn from case studies and case studies continue to constitute a large proportion of work generated by the discipline, the case study *method* is held in low regard or is simply ignored. 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 that they are doing, methodologically speaking. The case study survives in a curious methodological limbo.

How can we understand the profound disjuncture that exists between the case study's acknowledged contributions to political science and its maligned status within the discipline? If case studies are methodologically flawed, why do they persist?

The paper is divided into two parts. The first part focuses on matters of definition. I argue that for methodological purposes a case study is best defined as an in-depth study of a single unit (a relatively bounded phenomenon) where the scholar's aim is to elucidate features of a larger class of similar phenomena. It is demonstrated that case studies rely on the same sort of covariational evidence utilized in non-case study research. Thus, the case study method is correctly understood as a particular way of defining cases, not a way of analyzing cases or a way of modeling causal relations. I show, finally, that this understanding of the subject illuminates some of the persistent ambiguities of case study work, ambiguities that are, to some extent, intrinsic to the enterprise.

In the second part of the paper I proceed to examine the contrast between case study and non-case study work. The central argument here is that the differences between these two genres are best understood as characteristic strengths and weaknesses—affinities—rather than antagonistic approaches to the empirical world. Tradeoffs, rather than dichotomies, characterize the ongoing case study/non-case study debate.

WHAT IS A CASE STUDY?

What is a case study, and how is it differentiated from other styles of research? Regretfully, the term "case

John Gerring is Associate Professor, Department of Political Science, Boston University, 232 Bay State Road, Boston, MA 02215 (jgerring@bu.edu).

Helpful comments were received from Robert Adcock, Andrew Bennett, Henry Brady, Bear Braumoeller, David Collier, Michael Coppedge, Colin Elman, Peter Hall, Alan Jacobs, Evan Lieberman, Jim Mahoney, Jason Seawright, David Waldner, and three anonymous reviewers for the journal. William Barndt and Joshua Yesnowitz provided crucial assistance at the final stages of manuscript preparation. George and Bennett 2004 reached me in manuscript form as I was developing the ideas for this paper and had a substantial impact on my own understanding of the case study—although some differences of opinion remain. Funding for this research was generously provided by the Frederick S.Pardee Center for the Study of the Longer-Range Future. Work was completed while the author was in residence at the School of Social Science, Institute for Advanced Study.

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

study" is a definitional morass. To refer to a work as a case study might mean (a) that its method is qualitative, small-*N* (Yin 1994); (b) that the research is ethnographic, clinical, participant-observation, or otherwise "in the field" (Yin 1994); (c) that the research is characterized by process-tracing (George and Bennett 2004); (d) that the research investigates the properties of a single case (Campbell and Stanley 1963, 7; Eckstein [1975] 1992); or (e) that the research investigates a single phenomenon, instance, or example (the most common usage). Evidently, researchers have many things in mind when they talk about case study research.[1] As a result of this profusion of 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.

How, then, should the case study be understood? The first three options enumerated above (a–c) seem inappropriate as general definitions of the topic since each implies a substantial shift in meaning relative to established usage. One cannot substitute case study for *qualitative*, *ethnographic*, or *process-tracing* without feeling that something has been lost in translation. These definitions are best understood as describing certain kinds (subtypes) of case studies, rather than the general phenomenon itself. The fourth option (d) equates the case study with the study of a single case, the $N = 1$ research design. This is simply wrong, as argued at length below; case studies always employ more than one case. The fifth option (e), centering on *phenomenon, instance*, or *example* as the key term, is correct as far as it goes but also ambiguous. Imagine asking someone, "What is your instance?" or "What is your phenomenon?" A case study presupposes a relatively bounded phenomenon, an implication that none of these terms captures.

As a substitute for these flawed definitions, I propose to define the case study as *an intensive study of a single unit for the purpose of understanding a larger class of (similar) units*. A unit connotes a spatially bounded phenomenon—e.g., a nation-state, revolution, political party, election, or person—observed at a single point in time or over some delimited period of time. (Although the temporal boundaries of a unit are not always explicit, they are at least implicit.)[2]

To clarify this definition we must establish the relationship of the case study, so defined, to other terms in this crowded semantic field. Following is a set of nested definitions, which should be read carefully. A "population" is comprised of a "sample" (studied cases), as well as unstudied cases. A sample is comprised of several "units," and each unit is observed at discrete points in time, comprising "cases." A case is comprised of several relevant dimensions ("variables"), each of which is built upon an "observation" or observations. For those familiar with the rectangular form of a dataset it may be helpful to conceptualize observations as cells, variables as columns, cases as rows, and units as either groups of cases or individual cases (depending upon the proposition and the analysis).

The most important point is that all these terms are definable only by reference to a particular proposition and a corresponding research design. A country may function as a case, a unit, a population, or a case study. It all depends upon what one is arguing. In a typical cross-country time-series regression analysis (e.g., Przeworski et al. 2000), units are countries, cases are country-years, and observations are collected for each case on a range of variables. However, shifts in the unit of analysis of a proposition change the referential meaning of all terms in the semantic field. If one moves down one level of analysis the new population lies within the old population, the new sample within the old sample, and so forth, such that an observation in the original proposition now becomes a case. Population, unit, case, and observation are nested within each other. Since most social science research occurs at several levels of analysis these terms are generally in flux. Nonetheless, they have distinct meanings within the context of a single proposition, which defines the principal unit of analysis.

I do not issue this somewhat novel definition of case study (an intensive study of a single unit for the purpose of understanding a larger class of units) with any hopes of displacing common usage. Indeed, there is no harm in continuing to refer to a case study in the various ways listed above (options a–e). What is important is that we have recourse to a narrower and clearer definition when methodological confusions arise so that we have a way to arbitrate such confusions. The definition chosen here is useful in this regard. Moreover, it captures the essential features of other extant definitions; it is resonant (Gerring 2001, chap. 3). Finally, as the succeeding portions of this paper show, it clarifies the distinctive features of a broad class of work in the discipline of political science and in neighboring fields of the social sciences. It is theoretically useful.

The Case Study Method Considered as an Empirical Endeavor

The distinctiveness of the case study is most clearly understood when placed within a broader set of methodological options. To understand what a case study is, one must comprehend what it is not.

All empirical evidence of causal relationships is *covariational* in nature. A purported cause and effect must be found to covary. They must appear and disappear, wax and wane, or perform some other transformation in tandem or at some regular, more or less predictable, intervals. Even where this covariation is imagined, as in a counterfactual thought experiment, the evidence we imagine is of a covariational sort. Conversely, the absence of such covariation is taken as disconfirming evidence. If the appearance and disappearance (waxing/waning et al.) of X and Y are *not* associated

[1] In addition to sources cited above, see Brady and Collier 2004, Campbell (1975) 1988, Davidson and Costello 1969, Feagin, Orum, and Sjoberg 1991, George 1979, McKeown 1999, Ragin 1987, 1997, Ragin and Becker 1992, and the symposium, "The Case Study Method in Sociology," in *Current Sociology*, Volume 40, Number 1 (Spring 1992).

[2] Similar understandings of the term "unit" can be found elsewhere (e.g., King, Keohane, and Verba 1994, 76–77).

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

TABLE 1. Research Designs: A Covariational Typology

		Temporal Variation	
		No	Yes
Spatial Variation	None (1 unit)	[Logically impossible]	(a) Case study I
	Within-unit	(b) Case study II	(c) Case study III
	Across-unit	(d) Cross-sectional	(e) Time-series cross-sectional
	Across- and within-unit	(f) Hierarchical	(g) Hierarchical time-series; Comparative-historical

in any way that can be rationally explained, and hence predicted (or postdicted), then the empirical evidence suggests that a causal relationship does not exist.[3]

This provides a useful way of typologizing various research designs. Covariation may be observed (a) in a single unit diachronically, (b) within a single unit synchronically, (c) within a single unit diachronically, (d) across units synchronically, (e) across units synchronically and diachronically, (f) across and within units synchronically, or (g) across and within units synchronically and diachronically, as depicted in Table 1.

It will be seen that the case study occupies one of three possible cells. Type I case studies examine variation in a single unit over time, thus preserving the primary unit of analysis. Other case studies break down this primary unit into subunits, which are then subjected to covariational analysis—either synchronically (type II) or synchronically and diachronically (type III). These are the three logically conceivable approaches to the intensive study of a single unit where that unit is viewed as an instance of some broader phenomenon. Consequently, when one refers to *the* case study method one is in fact referring to three possible methods, each with a different menu of covariational evidence.

The bottom half of Table 1 lays out various across-unit research designs (where some important element of the empirical analysis involves comparisons across units). Here I have listed the methods most commonly identified with these research designs. Across-unit analysis without any explicit temporal component (d) is usually classified as "cross-sectional" (even though a temporal component is usually simulated with independent variables that are assumed to precede the dependent variable under investigation). When a temporal component is included we often refer to the analysis as "time–series cross-sectional" (TSCS) or pooled time–series (e). When one examines variation across- and within units in the same research design one is said to be employing a "hierarchical" model (f). Finally, when all forms of covariation are enlisted in a single research design the resulting method is described as "hierarchical time–series" (if quantitative) or "comparative-historical" (if qualitative) (g). Of all cross-unit research designs the case study is probably closest to the latter, where levels of analysis move up and down more or less simultaneously and where a small number of units are subjected to intensive study. Indeed, the comparative-historical study may be looked upon as a series of case studies combined with explicit cross-unit analysis (Mahoney and Rueschemeyer 2003).

Having placed these standard cross-unit research designs within a covariational typology one must also take note that each of these methods might also be employed *as a case study*. A case study may employ cross-sectional, TSCS, hierarchical, hierarchical time-series, and perhaps even comparative-historical models. It all depends upon the proposition in question. Specifically, it is the purposes to which these analyses are put, and hence the definition of a unit, that determines whether or not they are appropriately referred to as case studies. This will become clearer as we proceed.

The *N* Question

I have argued that what distinguishes the case study method from all other methods is its reliance on covariation demonstrated by a single unit and its attempt, at the same time, to illuminate features of a broader set of units. It follows from this that the number of cases (N) employed by a case study may be either small or large and, consequently, may be evaluated in a qualitative or quantitative fashion.[4]

To see why this must be so let us consider how a case study of a single event—say, the French Revolution—works. Intuitively, such a study provides an N of one (France). If one were to broaden the analysis to include a second revolution (e.g., the American Revolution), it would be common to describe the study as comprising two cases. Yet, as I have argued preliminarily, this is a gross distortion of what is really going on. It would be more correct to describe such a study as comprising two units, rather than two cases, for a case study of a single event generally examines that event over time. France is observed before, during, and after the event to see what changed and what remained the same after this cataclysmic event. These patterns of covariation offer the empirical clues one needs to reach conclusions about causation. They also create multiple cases out of that individual unit. $N = 2$, at the very least (e.g., before and after a revolution), in a case study of type I.

[3] Note that *covariation* (or correlation) refers to the mutual relationship between X and Y; *variation*, to the behavior of a single variable. These words are often used interchangeably. Hume's word for this was *constant conjunction*, and others have been employed as well. I should clarify that although the empirical component of a causal argument is covariational in nature, successful causal arguments depend upon more than just covariation. Among other things, a convincing causal account must identify a causal mechanism (see below).

[4] This section explains and elaborates on a theme first articulated by Campbell (1975) 1988, itself a revision of Campbell's earlier perspective (Campbell and Stanley 1963).

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

If, instead, there is no temporal variation—if, for example, the French Revolution is examined at a single point in time—then the object of investigation will be covariational patterns *within* that unit, a case study of type II. Within-unit cases consist of all cases that lie at a lower level of analysis relative to the inference under investigation. If the primary unit of analysis is the nation-state, then within-unit cases might be constructed from provinces, localities, groups, or individuals. The possibilities for within-unit analysis are, in principle, infinite. Indeed, within-unit *N* often swamps across-unit *N*, particularly where individuals comprise the relevant within-unit case. A single national survey will produce a larger sample than any conceivable cross-country analysis. Thus, in many circumstances case studies of type II comprise a larger *N* than cross-sectional analyses or TSCS analyses.

Evidently, if a case study *combines* both temporal and within-unit variation, as in case studies of type III, then its potential *N* increases accordingly. This is probably the most common genre of case study analysis.

These covariational facts hold true regardless of whether the method is experimental or nonexperimental. It is also true of counterfactual reasoning, which typically consists of four cases—the actual (as it happened) before and after cases and the before and after cases as reconstructed through counterfactual reasoning (i.e., with an imagined intervention). In short, the case study does not preclude high-*N*; it simply precludes *across-unit N* (by definition).

What, then, of the classic $N=1$ research design, which haunts the imaginations of social scientists everywhere? This hypothetical research design occupies the empty cell in Table 1. Its cell is empty because it represents a research design that is not logically feasible. A single unit observed at a single point in time without the addition of within-unit cases offers no evidence whatsoever of a causal proposition. In trying to intuit a causal relationship from this snapshot—a single case without within-unit covariation—we would be engaging in a truly random operation, since an infinite number of lines might be drawn through that one data point.

Ambiguities–Necessary and Unnecessary

The effort in this section has been to clarify what it means to conduct a case study. I have argued that a case study is most usefully defined as the intensive study of a single unit wherever the aim is to shed light on a question pertaining to a broader class of units. Although this definitional exercise does not settle all the ambiguities besetting the case study research design, it does provide a way of *understanding* ambiguities that remain. Six issues deserve emphasis.

The first ambiguity concerns the problem of distinguishing different types of covariational evidence. We have pointed out that case studies may observe a single unit through time (type I), synchronic within-unit variance (type II), or synchronic and diachronic within-unit variance (type III). Notice that types II and III, but *not* type I, involve a change in level of analysis, since cases are drawn from phenomena within the primary unit (as defined by the proposition of interest). Thus, some case studies—but not all—involve a change in the primary unit of analysis. To complicate matters further, case studies often combine observations of the primary unit over time (type I) with synchronic and diachronic observations of within-unit covariation (types II and III). Many case studies are thus hybrids of all three research designs. A final complication is introduced by the fact that it is often difficult to figure out which sort of covariational evidence is being mobilized at a particular juncture. The difficulty owes something to the complexities of within-unit analysis. Although the primary unit of analysis is usually clear, within-unit cases are often multiple and ambiguous.

A second source of ambiguity concerns the blurry line between a unit that is intensively studied—the case study—and other adjacent units that may be brought into the analysis in a less structured manner. Recall that because a case study refers to a set of units broader than the one immediately under study, a writer must have some knowledge of these additional units (a) to choose a unit for special treatment and (b) identify plausible causal hypotheses. Case studies are not immaculately conceived; additional units always loom in the background.

To speak of a case study at all it is helpful to introduce a distinction between *formal* and *informal* units. The formal unit is the unit chosen for intensive analysis—the person, group, organization, county, region, country, or other bounded phenomenon of which the writer has in-depth knowledge. Informal units consist of all other units that are brought into the analysis in a peripheral way, typically in an introductory or concluding chapter. Often, these informal units are studied only through secondary literature; they are always more superficially surveyed than the formal unit under study. Sometimes, the status of informal units is left implicit. This may be warranted in circumstances where the relevant comparison or contrast between the formal unit and other units is obvious or generally accepted. In any case, the distinction between a formal and an informal unit is always a matter of degrees. The more equality of treatment granted to peripheral units, the more a study leans toward a cross-unit style of analysis. The greater the predominance of a single unit, the more it merits the appellation case study.

A third ambiguity occurs whenever a single work *combines* single-unit and across-unit analysis in a formal manner. This would be true of comparative-historical work as well as any work in which an intensively studied unit is "nested" within a broader research design (Coppedge 2002; Lieberman 2003). Indeed, the only thing that distinguishes the single-unit study from a *sample* (which is of course also designed to elucidate the features of some larger phenomenon) is that the latter is generally understood as composed of more than one unit. Case studies, like 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; it is a matter of

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

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.

A fourth ambiguity afflicting case studies is that such works generally partake of two empirical worlds. They are both studies *tout court* and case studies of something more general. As a study, the population is restricted to the unit under investigation. As a case study, the population includes adjacent units—perhaps quite a large number of them. This tension is evident in Graham Allison's (1971) renowned work, whose subtitle, *Explaining the Cuban Missile Crisis*, invokes a narrow topic, whereas the title, *Essence of Decision*, suggests a much larger topic (government decision-making). To complicate matters further, different propositions within the same work commonly apply to different populations. Some may be restricted to the unit under study, whereas others have a wider ambit. This is clearly the case in Allison's study and is noted explicitly in the introduction.

To complicate matters further, the status of a work may *change* as it is digested and appropriated by a community of scholars. "Meta-analyses" are systematic attempts to integrate the results of individual studies into a single quantitative analysis, pooling individual cases drawn from each study into a single dataset (with various weightings and restrictions). The ubiquitous "literature review" often aims at the same objective, albeit in a less synoptic way. Both statistical meta-analyses and narrative literature reviews assimilate a series of studies, treating each of them as case studies in some larger project—whether or not this was the intention of the original authors.

A final ambiguity concerns the sort of argument that a case study is intended to prove or demonstrate. One species of case study examines a loosely defined general topic—war, revolution, gender relations—in a particular setting but offers no specific proposition that might be applied across a larger set of units. E. P. Thompson's *The Making of the English Working Class* (1963) is usually construed as a case study of class formation. This suggests a very general purview, perhaps applicable to all countries in the modern era. Yet Thompson does not proffer a theory of class formation, aside from the rather fuzzy notion of a working class participating in its own development. Thus, his work is probably correctly understood as a study of how a more general phenomenon occurred in one country setting. Virtually any intensive study of a relatively bounded topic qualifies as a case study in this minimal sense, so long as it can be linked with some larger topic via a key word (e.g., class formation). Indeed, the narrowest terrains sometimes claim the broadest extensions. Studies of a war are studies of war, studies of a farming community are studies of farming communities everywhere, studies of individuals are studies of humanity, and so forth.

A very different style of argumentation informs Benjamin Reilly's (2001) study of the role of electoral systems in ethnically divided societies. Reilly argues, on the basis of several case studies, that single–transferable vote (STV) electoral systems have a moderating effect on group conflict relative to first-past-the-post (FPP) electoral systems. Here is a good example of a case study that is more than simply suggestive (for other examples see Eaton 2003, Elman 1997, Lijphart 1968, and Stratmann and Baur 2002). For present purposes, what is significant is that both styles of argumentation—the suggestive and the falsifiable—are legitimately referred to as case studies. Evidently, they have very different methodological implications. But these implications should not be confused with the case study format, which can be implemented in interpretivist as well as positivist modes.

Having flagged these six ambiguities of the case study, the question is begged: Are they necessary? Are they intrinsic to the research design, or might they be avoided?

In many instances, ambiguities can be removed simply by more careful attention to the task of specification (Gerring 2001, 90–99). Writers should be clear about which propositions are intended to describe the unit under study and which are intended to apply to a broader set of units. Regrettably, many studies focused on some element of politics in the United States frame their analysis as a study of politics—by implication, politics *in general* (everywhere and always). One is left to wonder whether the study pertains only to American politics, to all contemporary polities, or, in varying degrees, to both. Indeed, the slippage between study and case study accounts for much of the confusion that we encounter when reading single-unit analyses. To the extent that propositions—and their attendant cases, units, and populations—are stated clearly and explicitly, the author avoids confusion and the work attains a higher degree of falsifiability. This may involve some sacrifice in narrative flow, but it is rightly regarded as the entry price of social science.

However, it hardly seems plausible that the six ambiguities noted above arise solely from the sloppy or unduly belletristic habits of case study practitioners. Indeed, a certain degree of ambiguity is *inherent* in the enterprise of the case study. This pertains, most of all, to the study/case study distinction.

It would be difficult to write a study of a single unit that does not also function as a case study, and vice versa, for reasons already explored. Indeed, it may be difficult to neatly separate the study and case study components of a work (e.g., into different chapters or differently labeled propositions). The reason for this structural ambiguity is that the utility of the single-unit study rests partly on its double functions. One wishes to know both what is particular to that unit *and* what is general about it. It should be kept in mind that case studies often tackle subjects about which little is previously known or about which existing knowledge is fundamentally flawed. The case study typically presents original research of some sort. Indeed, it is the opportunity to study a single unit in great depth that constitutes one of the primary virtues of the case study method (see below). If a writer were to restrict herself only to

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

TABLE 2. Single-Unit Versus Cross-Unit Research Designs: Tradeoffs and Affinities

		Affinity	
		Case Study	Cross-Unit Study
1. Type of inference	(a) Descriptive	+	
	(b) Causal		+
2. Scope of proposition	(a) Depth	+	
	(b) Breadth		+
	(c) Boundedness		+
3. Unit homogeneity	(a) Case comparability (internal)	+	
	(b) Representativeness (external)		+
4. Causal insight	(a) Causal mechanisms	+	
	(b) Causal effect		+
5. Causal relationship	(a) Invariant	+	
	(b) Probabilistic		+
6. Strategy of research	(a) Exploratory (theory generation)	+	
	(b) Confirmatory (theory testing)		+
7. Useful variance	(a) For only a single unit	+	
	(b) For many units		+
8. Ontology		Indeterminate	

elements of the unit that were generalizable (i.e., if she rigorously maintains the "case study" mode of analysis), a reader might justifiably complain. Such rigor would clarify the population of the primary inference, but it would also constitute a considerable waste of scholarly resources. Imagine a study of economic growth that focuses on Mauritius as a case study yet refuses to engage causal questions unless they are clearly applicable to other countries (since this is a case study of a more general phenomenon, growth). No mention of factors specific to the Mauritian case is allowable; all proper nouns are converted into common nouns (Przeworski and Teune 1970). Such a study seems unduly narrow; its conclusions may mislead.

Indeed, it is often *difficult to tell* which of the many features of a given unit are typical of a larger set of units (and hence fodder for generalizable inferences) and which are particular to the unit under study. The appropriate response to such ambiguity is for the writer to report all facts and hypotheses that might be relevant—in short, to overreport. Much of the detail provided by the typical case study may be regarded as "field notes" of possible utility for future researchers—perhaps with a rather different set of inferences in mind. Again, it seems justifiable for case studies to work on two levels simultaneously, the unit itself and some broader class of (perhaps difficult to specify) units.

As a general observation we might say that methods, strictly defined, tend to lose their shape as one looks closer at their innards. A study merges into a case study, a single-unit study merges into a study of a sample, a longitudinal study merges into a latitudinal study, informal cases merge into formal cases, and so forth. Methods that seem quite dissimilar in design bleed into one another when put into practice. There are few "pure" methods. And this is probably a good thing. Chastity is not necessarily an attribute to be cherished in research design.

WHAT IS A CASE STUDY GOOD FOR?

It has been demonstrated that the difference between a case study and a study (tout court) is rarely clear-cut. Indeed, the case study is probably best understood as an ideal-type rather than a method with hard-and-fast rules. Yet the fact that the case study is fuzzy around the edges does not mean that it is lacking in distinctive characteristics. When considered as an ideal type the case study research design, like all research designs, exhibits characteristic strengths and weaknesses relative to its across-unit cousin. These pertain to the type of inference under consideration (descriptive or causal), the scope of the proposition (its depth, breadth, and boundedness), the degree of unit homogeneity found among cases and between the sample and the population, the sort of causal insight desired (causal effect or causal mechanism), the strategy of research (exploratory or confirmatory), and the kind of empirical evidence available. Tradeoffs along these seven dimensions are summarized in Table 2. Ontological presuppositions are also important but of indeterminate import (as indicated in Table 2).

It should be underlined that these tradeoffs represent methodological *affinities*, not invariant laws. Exceptions can be found to each of the general tendencies identified here. Even so, the strengths and weaknesses often noted in case study research, reproduced in many subfields and disciplines over the course of many decades, are not the product of a purely stochastic process. General patterns suggest general interpretations.

I should also emphasize that each of these tradeoffs carries a ceteris paribus caveat. Case studies are more useful for forming descriptive inferences, *all other things being equal*. Since ceteris is not always paribus the reader should not jump to any conclusions about the research design appropriate to a given setting without

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

considering the single-unit/cross-unit options available within that research context.

Finally, readers should note that although many of my examples are drawn from the subfield of comparative politics, with nation-states as the principal unit of concern, these examples could be replicated with other units and in other research settings. The problem of the case study is not limited to a single subfield.

Type of Inference: Descriptive Versus Causal

Descriptive inference remains an important, if undervalued, trope within the social sciences (Gerring 2001, chap. 6; King, Keohane, and Verba 1994, chap. 2). Thus, it is not at all pejorative to observe that there is a methodological affinity between descriptive inference and case study work. When one is examining correlative relationships or proximate causal relationships the case study format seems less problematic and is often highly informative. Indeed, many of the most famous case studies in anthropology, political science, and sociology are primarily descriptive in orientation (e.g., Fenno 1978, Hartz 1955, Lynd and Lynd [1929] 1956, Malinowski [1922] 1984, and Whyte [1943] 1955). How can we understand this affinity?

What? and *How?* questions are easier to answer without recourse to cross-unit analysis than *Why?* questions. The simplest genre of descriptive case study asserts that the unit under study (A) is like, or unlike, other similar units (B and C). A more complicated descriptive case study might assert a classificatory relationship among A, B, and C, such that A falls into a certain typological relationship with B and C. The latter, of course, is more complicated and is more likely to require some explicit cross-unit examination. However, a descriptive inference does not make any assertions about causal relationships (beyond the most proximal) occurring *within* A, B, and C. In this sense, descriptive inference is simpler, methodologically speaking.

To be sure, descriptive case study propositions are implicitly comparative and these comparisons must have a cross-unit reference point. To say "green" is to imply "not blue." However, it is usually fairly easy to make such comparisons without conducting a study of the presumed variation. One knows what blue is without going in search of blue cases. This illustrates something important about the structure of descriptive propositions in social science. They are held together by language—by ordinary or technical terms and their definitions. When describing a phenomenon one is usually comparing it to an ideal-type definition. American political culture is "liberal" or "republican" insofar as it conforms to standard definitions of these two concepts (Smith 1997). To describe is to categorize, and to categorize is to rely on language to divide up the world into identifiable entities. Language, in this sense, provides the "laws" that allow for consistent interpretations of the phenomenal world (Sartori 1984). Chairs are different from tables in predictable ways; labeling entities as one or the other thus allows the use of nomothetic descriptive generalizations without actually studying the phenomenon in question each time one sallies forth with a new pronouncement about the world.[5]

It should be clear that the affinity between case study research and descriptive inference does not denigrate the possibility of causal analysis through case study research, of which one might cite many illustrious examples. Indeed, the discussion that follows is primarily concerned with propositions of a causal nature. My point is simply that it is easier to conduct descriptive work than to investigate causal propositions while working in a case study mode.

Scope of Proposition: Breadth and Boundedness Versus Depth

The variable utility of the case study is also partly a product of the *scope* of the causal argument that a writer wishes to prove or demonstrate. Arguments that strive for great breadth and boundedness are in greater need of cross-unit cases; causal arguments restricted to a small set of units more plausibly subsist on the basis of a single-unit study. The extensive/intensive tradeoff is commonsensical. Insofar as one wishes to make an argument about the universe of nation-states, one is on surer ground if one has studied more than one of those nation-states. A case study of France probably offers better evidence for an argument about Europe than for an argument about the world. Propositional breadth and evidentiary breadth generally go hand in hand. The evidence should be commensurate with the scope of the proposition.

This statement, like all others, has a ceteris paribus caveat. There is a variety of ways in which single-unit 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 or by choosing "crucial" cases (Eckstein [1975] 1992). Even so, a proposition with a narrow scope will be more conducive to case study analysis than a proposition with a broad purview. The breadth of an inference thus constitutes one factor, among many, in determining the utility of the case study mode of analysis.

Similarly, the *boundedness* of an inference is often related to the degree to which it exploits cross-unit variance. Precisely because their focus is so tight, case studies often produce inferences with poorly defined boundaries. It is clear that an inference extends beyond the unit under study, but it is often unclear how far the inference extends. Cross-unit research may also suffer from poorly bounded inferences; however, it is less likely to do so since the research design allows one to test the limits of an inference in an explicit fashion.

[5] Granted, cross-unit examination may be useful, particularly when the terms in question are broad and/or ambiguous. If, for example, one is examining American political culture as an example of a broader class of political cultures deemed "liberal," it will be difficult to reach firm conclusions without a larger sample of nation-states. Even so, it will be easier to describe the subject without cross-unit reference-points than to *explain* it.

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

The sharpness of the boundaries of the population—what is and is not "covered" by an inference—is easier to establish when units in the population also constitute cases under study.

By the same token, one of the primary virtues of the case study method is the *depth* of analysis that it offers. One may think of depth as referring to the detail, richness, completeness, wholeness, or degree of variance that is accounted for by an explanation. The case study researcher's complaint about the thinness of cross-unit analysis is well taken; such studies often have little to say about individual cases. Otherwise stated, cross-unit studies are likely to explain only a small portion of the variance with respect to a given outcome or to approach that outcome at a very general level. A cross-unit study might be satisfied to explain the occurrence of a revolution in Country A, while a case study of Country A might also strive to explain specific features of that event—why it occurred when it did and in the way that it did.

I shall return to the advantages of taking an in-depth, *Gestalt*-oriented look, at a single unit. For present purposes, the conclusion is simple. Research designs invariably face a choice between knowing more about less and knowing less about more. The case study method may be defended, as well as criticized, along these lines (Ragin 2000, 22).

Unit Homogeneity: Case Comparability Versus Representativeness

Single-unit studies provide cases that are likely to be *comparable* to one another. After all, they are all drawn from the same unit (by definition). Cases drawn from different units, in contrast, often force the researcher to make heroic assumptions about the comparability of concepts and causal relationships across the chosen cases. After all, they are different units.

Yet the strength of the case study also suggests a corresponding weakness. Single-unit research designs often fall short in their *representativeness*—the degree to which causal relationships evidenced by that single unit may be assumed to be true for a larger set of (unstudied) units. Are the men chosen by Robert Lane (1962) typical of the American male, white, immigrant, working class? Is Middletown representative of America (Lynd and Lynd [1929] 1956)?

The tradeoff between comparability and representativeness is a general feature of cross-unit sample size. Naturally, there are many ways to overcome the corresponding problems of comparability and representativeness in both case study and non-case study research. Even so, it should be pointed out that the addition of units to a research design can hardly *increase* the case comparability of a sample. Similarly, it is unlikely that the addition of units will *decrease* the representativeness of a sample (though it is certainly possible). Thus, it seems appropriate to regard the tradeoff between comparability and representativeness, like other tradeoffs, as intrinsic to the study/case study choice of research design.

Causal Insight: Causal Effect Versus Causal Mechanisms

A fourth tradeoff concerns the sort of insight into causation that one is able to achieve by examining empirical evidence of a particular X:Y relationship. Traditionally, quantitative researchers have been concerned primarily with the estimation of a causal effect—the effect on Y of a given change in X, taking all background circumstances into account (King, Keohane, and Verba 1994, 81–82)—and, equally important, an estimate of the probability of that effect, captured statistically in the error term. Assuming that the causal relationship is probabilistic in nature (see discussion below), the researcher must examine several instances of this phenomenon to gauge the average causal effect of X on Y and the random element of that variation. The calculation of a causal effect presumes the investigation of cross-unit variation precisely because, for a probabilistic cause, one cannot assume that the behavior of one unit will be indicative of the behavior of other units. Units may behave differently. Thus, the example of a single unit, even if subjected to iterated testing, is not a good way to estimate causal effects and is certainly inadequate to the task of estimating probabilities.

But causal arguments depend not only on measuring causal effects. They also presuppose the identification of a causal mechanism (Gerring, forthcoming; Hedstrom and Swedberg 1998). X must be connected with Y in a plausible fashion; otherwise, it is unclear whether a pattern of covariation is truly causal in nature. The identification of causal mechanisms happens when one puts together general knowledge of the world with empirical knowledge of how X and Y interrelate. It is in the latter task that case studies enjoy a comparative advantage.

Case studies, if well constructed, allow one to peer into the box of causality to the intermediate causes lying between some 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. Clayton Roberts (1996, 66) describes process-tracing as "the minute tracing of the explanatory narrative to the point where the events to be explained are microscopic and the covering laws correspondingly more certain" (see also George and Bennett 2004). Often, the connections between a putative cause and its effect are rendered visible once one has examined the motivations of the actors involved. Intentionality is an integral part of causal analysis, as interpretivists have been claiming for some time (Taylor 1970). Similarly, the investigation of a single unit may allow one to test the causal implications of a theory, thus providing corroborating evidence for a causal argument. This is often referred to as *pattern-matching*. Here, the theory of primary interest "generates predictions or expectations on dozens of other aspects of the culture, and [the writer] does not retain the theory unless most of these are also confirmed. In some sense, he has tested the theory with degrees of freedom coming from the multiple implications of any one theory" (Campbell [1975] 1988, 380).

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

One can readily see why the investigation of causal mechanisms (including both process-tracing and pattern-matching) is commonly associated with the case study research design. The in-depth analysis of a single unit is useful in elucidating causal mechanisms because its characteristic style of evidence-gathering—over-time and within-unit variation—is likely to provide clues into what connects a purported *X* to a particular *Y*. Cross-unit variation, in contrast, is often mute with respect to causal mechanisms. The *X*s and *Y*s may be at a considerable remove from one another; one does not know, or must simply intuit, what connects the dots.

The caveat here is that cross-unit evidence is not always or *necessarily* mute with respect to causal mechanisms, and single-unit analysis is not always revelatory on this score. A cross-unit analysis is opaque if, or insofar as, (a) there is great causal distance between the *X* and the *Y* variables, (b) the intervening variables thought to lie between *X* and *Y* cannot be tested empirically in a cross-unit research design, and (c) the *X:Y* relationship cannot be intuited from common sense or deductive reasoning. There is a general perception—common at least among protagonists of the case study—that cross-unit studies entail large black boxes with no peepholes and that, therefore, they *must* be supplemented by the in-depth analysis of key units. This is not always the case. One can think of plenty of studies in which the relationship is quite clear, and where a case study would be superfluous. And one can point out any number of studies in which intervening variables are in fact investigated with a cross-unit research design.[6] Again, it is important to stress that we are examining typical, not definitional, characteristics of the case study. Ceteris paribus, case studies are more likely to shed light on causal mechanisms and less likely to shed light on true causal effects.

Causal Relationship: Invariant Versus Probabilistic

Causal arguments may be either invariant ("deterministic") or probabilistic. Invariant causal relationships are asserted to be always true, given some set of background circumstances. They take the form of necessary, sufficient, or necessary and sufficient arguments. Probabilistic arguments, in contrast, are true in a probabilistic fashion; a cause increases the likelihood of an outcome and/or the magnitude of a (scalar) outcome.[7]

Ceteris paribus, case study research designs have an easier time addressing invariant causes. Consider that a necessary or sufficient causal proposition can be *dis*proved with a single case study (Dion 1998). *Proving* an invariant causal argument generally requires more cross-unit 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 unit under study carries more weight. Where the causal relationship is assumed to be probabilistic, on the other hand, case study evidence is easier to dismiss; it is, after all, just one example of a general phenomenon assumed to have a stochastic component. Recall that an error term may be incorporated into case study work since the *N* of a case study is indeterminate; however, this error term remains a property of single-unit analysis. In contrast, the error term in a large-*N* cross-unit analysis represents the stochastic nature of an (assumed) probabilistic relationship.

Strategy of Research: Exploratory Versus Confirmatory

Social science research involves a quest for new theories as well as a testing of existing theories, a series of "conjectures and refutations" (Popper 1969). Regrettably, social science methodology has focused almost exclusively on the latter. The former is quickly dismissed as a matter of guesswork, inspiration, or luck—a leap of faith, in any case, and hence a poor subject for methodological reflection. Yet it will readily be granted that many works of social science, including most generally acknowledged classics, are seminal works. Their classic status derives from a new idea, a new perspective, that is subsequently subjected to more rigorous analysis. Indeed, it is difficult to devise a program of falsification the first time a new theory is proposed. Path-breaking research is, by definition, exploratory. Subsequent research on that topic is confirmationist insofar as its primary task is to verify or falsify a preexisting hypothesis or a set of hypotheses. Thus, the world of social science may be usefully divided according to the predominant strategy of research undertaken, *exploratory* or *confirmatory/disconfirmatory* (Gerring 2001, chap. 10). These constitute two moments of empirical research, a generative moment and a skeptical moment, each of which is essential to the progress of a discipline.

Case studies enjoy a natural advantage in research of an exploratory nature. These same advantages,

[6] For example, a vast literature addresses the putative causal connections 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 (1978), is that increased trade openness leads to greater domestic economic vulnerability to external shocks (due, for instance, to changing terms of trade). In subsequent work, writers have attempted to test this hypothesis by examining the correlation between terms of trade and social welfare spending. If Cameron's causal mechanism is operative, one should find a robust correlation between these two variables in a cross-national regression. As it happens, results are equivocal (Alesina, Glaeser, and Sacerdote 2001). The point is, writers can and do exploit cross-unit variation to test assumptions about causal mechanisms.

[7] I avoid the term "determinism" since it has multiple meanings, only one of which—invariance—is relevant here. I assume that to make an invariant causal argument does not commit one to a view that all causes are perfectly determined; some causes may be invariant and others probabilistic. Useful discussions of invariance/determinism can be found in Adcock 2002, Dion 1998, 141, Goertz and Starr 2003, and Waldner 2002. Goertz (2003, 76–94) includes a sample of 150 necessary condition hypotheses deployed in various fields of the social sciences (Goertz 2003, 76–94).

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

however, often serve as impediments in work of a confirmatory nature. Let us explore why this might be so.[8]

Traditionally, scientific methodology has been identified with the segregation of conjecture and refutation; one should not be allowed to contaminate the other. Yet in the real world of social science, inspiration arises from perspiration. "Lightbulb" moments build on a close engagement with the particular facts of a particular case (unit). Ragin (1997) 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. It is a highly circular process. 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 intervention is necessary in the course of defining a case study topic, for there is a great deal of evidentiary leeway. Yet the very "subjectivity" of case study research allows for the generation of a great number of hypotheses, insights that might not be apparent to the cross-unit researcher who works with a thinner set of empirical data across a large number of units and with a more determinate (fixed) definition of cases, variables, and outcomes. It is the very fuzziness of case studies that grant them a strong advantage in research at exploratory stages, for the single-unit study allows one to test a multitude of hypotheses in a rough-and-ready way. Nor is this an entirely conjectural process. The covariational relationships discovered among different elements of a single unit 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 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," and "theory-building" exercises (Eckstein [1975] 1992; Ragin 1992, 1997; Rueschemeyer and Stephens 1997).

A multiple-unit study, in 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 of a confirmatory nature. There is less room for authorial intervention because evidence gathered from a cross-unit research design can only be interpreted in a limited number of ways. Another way of stating the point is to say that whereas case studies lean toward Type 1 errors (falsely rejecting the null hypothesis), cross-unit studies lean toward Type 2 errors (failing to reject the false null hypothesis). Perhaps this explains why case studies are more likely to be theory-generating, whereas cross-unit studies toil in the prosaic but highly structured field of normal science.

[8] For discussion of this tradeoff in the context of economic growth theory see Temple 1999, 120.

I do not mean to suggest that case studies never serve a confirmatory role. As discussed, evidence drawn from a single unit may *dis*confirm a necessary or sufficient hypothesis. Case studies are also often useful in conjunction with a cross-unit study for the purpose of elucidating causal mechanisms, as discussed. However, general theories rarely offer the kind of detailed and determinate predictions on within-unit variation that would allow one to reject a hypothesis through pattern-matching (without additional cross-unit evidence). The point is, theory confirmation/disconfirmation is not the case study's strong suit. The selection of "crucial" cases cannot overcome the fact that cross-unit N is minimal. We are unlikely to reject a hypothesis, or to consider it definitively proved, on the basis of the study of a single unit, particularly if the hypothesis has behind it a corpus of scholarly work. Eckstein himself acknowledges that his argument for case studies as a form of theory confirmation is largely conjectural. At the time of writing, several decades ago, he could not point to any existing study where a crucial case study had performed the heroic role assigned to it (Eckstein [1975] 1992, 172). I suspect that this is still more or less true. Indeed, it is true even of experimental case studies in the natural sciences (Campbell and Stanley 1963, 3). A single case study is still a single-shot, a single piece of evidence lying at the same level of analysis as the proposition itself.

The tradeoff between exploratory and confirmatory research 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. The problem is that work of an exploratory nature, although it may receive praise from the discipline, is unappreciated, and greatly undertheorized, by methodologists.

Useful Variance: Single- Versus Multiple-Unit

The analysis of any causal relationship hinges on the counterfactual assumption—that without X (or with more or less of X), Y would be different. In investigating this assumption the preferred research designs are, in order of preference: (a) laboratory or field experiments, (b) "natural" experiments (where a single unit undergoes unmanipulated change through time that approximates a true experiment), (c) thought experiments (counterfactuals), or (d) statistical controls (a quasi-experimental method of neutralizing irrelevant variables so as to isolate the true causal effects of one or a few factors of theoretical interest). The implication of this hierarchy of research designs is that a laboratory/field experiment, natural experiment, or thought experiment involving a single unit may be more useful than multiple units that attempt to mimic the virtues of the experimental method with purely "statistical" evidence.

For example, in investigating the relationship between campaign efforts and voter turnout one might be more convinced by a field experiment conducted in

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

a single community than by multiple cross-community studies or individual poll data that rely on a host of more or less unsatisfactory quasi-experimental controls (Gerber and Green 2000). Similarly, in investigating the role of electoral systems in conditioning public policy outcomes one might be more convinced by a single natural experiment—a change in a country's electoral system—than by a cross-country study employing statistical controls to examine these complex causal relationships (Horowitz 1985, 602). Even where one is forced to rely solely on counterfactual thought-experiments to evaluate causal claims, one might still prefer the case study mode if useful variance is not available across units. The time-honored question of whether early democratization leads to a quiescent working class and lower levels of social welfare development (Lipset 1963) is difficult to investigate cross-nationally for the simple reason that only one country granted suffrage to the (male) working class prior to industrialization. In this circumstance, a historical study focused on the United States (i.e., a case study) may provide the most compelling evidence of a general proposition.

To clarify, the issue is not whether natural experiments—or real experiments for that matter—are desirable. They are always desirable. The issue is how many experiments are available and how (truly) experimental are they? If many experiments are possible, as they typically are in natural science settings, then a cross-unit research design is probably justified. But this is uncommon in social science settings. Similarly, returning to our previous examples, if nation-states exhibit substantial variance on the independent variables of interest—early democratization and electoral change—then one would probably choose a cross-unit research design (with a time-series dimension). This would not preclude a supplementary case study, but such a case study would undoubtedly carry less weight. In the event, however, one is thrown back on more primitive expedients. Much of the utility of the case study is "contextual" in this special sense. The sort of natural or experimental variation that would make cross-unit analysis fruitful is lacking; such units simply do not exist and cannot be generated.

This helps to explain why case studies often focus on rare ("historical") events. Let me define an *event* provisionally as an instance of substantial and relatively quick change in an independent or dependent variable of theoretical interest. Now imagine a universe of empirical data in which such events occur in only 13 instances. This approximates the situation of scholars whose work focuses on the phenomenon of social revolution (e.g., Skocpol 1979, 287). In principle, the field of empirical evidence is virtually boundless, including all nation-states over the past three centuries. (Let us suppose that this constitutes the domain of the inference.) Yet useful variation is severely limited. The world does not provide many units that have experienced revolutionary upheaval. Arguably, one learns more about this phenomenon through an intensive case study of France, or several case studies focused on the 13 or so social-revolutionary countries (the comparative-historical method), than by a cross-country TSCS research design. The same problem affects the study of many social science questions where only one unit, or a small number of units, undergoes the event that is to be explained: early industrialization (England and the Netherlands), fascism (Germany, Italy), the use of nuclear weapons (United States), world war (WWI, WWII), democratization in Africa (Botswana, Mauritius, South Africa), single-nontransferable vote (SNTV) electoral systems (prereform Japan, Jordan, Taiwan, Vanuatu), and settler societies (Australia, Canada, New Zealand, South Africa, United States).

Ontological Considerations

Thus far, I have looked upon the choice between case study and non-case study methods as a matter of logic (the logic of causal inquiry) and empirics (the investigation of the empirical world). However, this choice rests also on *ontological* presuppositions. 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 (Hall 2003; Kuhn [1962] 1970). As a result of its all-embracing character and ambiguous claims, an ontology cannot be proved or disproved, at least not in the usual (verificationist or falsificationist) sense. Although 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.

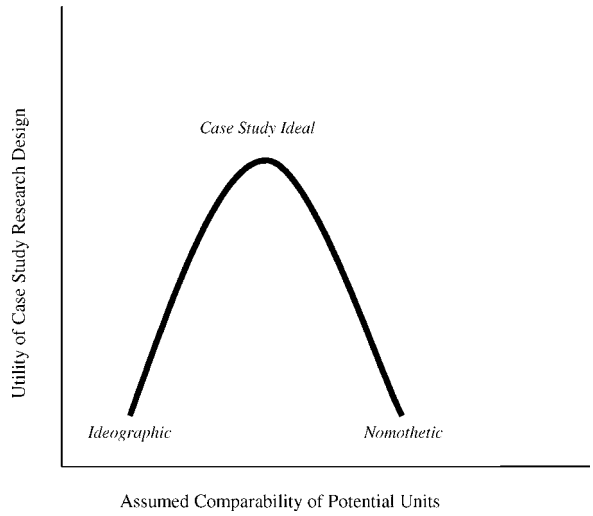
The significance of ontological presuppositions becomes apparent whenever questions of breadth, case comparability, and representativeness are vetted. Consider the possibility that a sample of units will become more and more dissimilar (less comparable) as that sample is enlarged, but with no obvious break-points. Here it evidently is the choice of the researcher how to define the population of a given inference and, hence, what the relevant units of analysis will be. Where do like cases end and unlike cases begin?

If adjacent units are thought to be entirely noncomparable, the case study method is impossible. The perfectly ideographic universe displays such uniqueness among units that absolutely nothing can be learned about one unit by studying another. The notion of a "case study" is nonsensical. At the other extreme, where all units of a given type are perfectly comparable, the case study is equally nonsensical. Why focus on a single unit when other units will do just as well? This is the nomothet's way of looking at things.

Case study researchers are situated between these two extremes. They are dubious about the viability of comparisons drawn across many apparently diverse units. "Variable-oriented" research involves "homogenizing assumptions" (Ragin 2000, 35). Yet they are equally suspicious of the claim, implicit in much historical and anthropological work, that each unit is

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

FIGURE 1. The Ontology of Case Study Research Design



unique. The case study thus occupies a tenuous ontological ground midway between ideographic and nomothetic extremes, as depicted in Figure 1 (see Collier and Collier [1991] 2002, 13–14).

To be sure, there is no profit in dwelling on ontological differences between case study and non-case study researchers. Ontological debates are, by definition, irresolvable. Once one has defended one's position as a matter of ontology, further discussion is superfluous except as it might bear upon matters of logic and coherence. If social science is understood as an evidence-based form of inquiry then matters of ontology are simply not relevant or are only tangentially relevant. Nonetheless, insofar as our ontological presuppositions influence our construction of cases, we had best be cognizant of this fact. Indeed, the middle-range position of case study research on this crucial question may help to account for its ambiguous position in the social sciences. It is neither fish nor fowl, ontologically speaking.

CONCLUSION

This paper has attempted to shed light on the apparent disjuncture between an often-maligned methodology and a heavily practiced method, both of which go by the name of *case study*. 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. The first task of the essay, 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 unit with an aim to generalize across a larger set of units.

In other respects, the predicament of the case study is not merely definitional but rather inheres in the method itself. To study a single unit with intent to shed light upon other units brings in its train six methodological ambiguities that are, to some extent, ineradicable. First, case studies may build upon a variety of covariational evidence; there is no single type of case study evidence, but rather *three* (see Table 1). Second, case studies assume a distinction between formally and informally studied units that is never entirely clear since the latter must be brought into the analysis in some fashion but would compromise the notion of a case study if fully integrated. Third, individual case studies are often grouped together in a single study, thus confounding the distinction between single- and cross-unit analysis. Fourth, case studies usually perform a double function; they are studies (of the unit itself) as well as case studies (of a broader class of units). Finally, the inference(s) pursued by a case study may be either illustrative or falsifiable. These methodological issues bedevil the case study research design. And they shall continue to do so, for they are inherent in the enterprise.

The travails of the case study within the discipline of political science are rooted, finally, in an insufficient appreciation of the methodological tradeoffs that this method calls forth. I have argued that at least seven characteristic strengths and weaknesses must be considered. Ceteris paribus, case studies are generally more useful (1) when inferences are descriptive rather than causal, (2) when propositional depth is prized over breadth and boundedness, (3) when (internal) case comparability is given precedence over (external) case representativeness, (4) when insight into causal mechanisms is more important than insight into causal effects, (5) when the causal proposition at issue is invariant rather than probabilistic, (6) when the strategy of research is exploratory, rather than confirmatory, and (7) when useful variance is available for only a single unit or a small number of units. (Ontological considerations also come into play when one chooses between a single-unit and a cross-unit research design, though the methodological implications of these assumptions are equivocal.)

Of these seven considerations, the last is perhaps the most important. There is little point in pursuing cross-unit analysis if the units in question do not exhibit variation on the dimensions of theoretical interest and/or the researcher cannot manage to hold other, potentially confounding, factors constant. Of course, a preliminary canvassing of these units is necessary to perceive these facts about the sample. But beyond this, cross-unit analysis may be of little consequence. By the same token, individual units may offer useful variance or they may not. In any case, the most important single question researchers should ask themselves as they contemplate various research designs is which one of these options most closely approximates the experimental ideal.

Throughout this discussion I have avoided the discussion of methodological considerations that are *practical* in nature or rooted in specific research contexts. My concern has been with general methodological issues. Before concluding this essay, however, it is important to acknowledge that practical and contextual considerations are often paramount in the choice between a case study and a non-case study research format.

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

The collection of original data is typically more difficult in cross-unit analysis than in case study analysis, involving greater expense, greater difficulties in identifying and coding cases, learning foreign languages, traveling, and so forth.[9] Whatever can be done for a set of units can usually be done more easily for a single unit. Similarly, case studies commonly afford multiple observations of a single case, thus providing firmer evidence of the factual accuracy of a given proposition than would be possible in the analogous cross-unit study.

A second practical consideration concerns the state of research on a given topic. Here one is concerned with the "triangulation" of evidence rather than the ease of evidence-gathering. Social scientists are accustomed to the idea that research occurs within the context of an ongoing tradition. All work is dependent, for the identification of topic, argument, and evidence, on this research tradition. What we need to know, and hence ought to study, is to some extent contingent upon what is already known. It follows from this that the utility of case study research relative to non-case study research is to some extent the product of the state of research within a given field. A field dominated by case studies may have little need for another case study. A field where cross-unit studies are hegemonic may be desperately in need of in-depth studies focused on single units.

Indeed, much of the debate over the utility of the case study method has little to do with the method itself and more to do with the state of current research in that field. If both case study and cross-unit methods have much to recommend them (the implicit argument of this paper), then both ought to be pursued—perhaps not in equal measure but at least with equal diligence and respect. There is no virtue, and potentially great harm, in pursuing one approach to the exclusion of the other or in ghettoizing the practitioners of the minority approach. The triangulation essential to social scientific advance demands the employment of a variety of (viable) methods, including the case study.

This paper is manifestly not a brief *for* the case study. Rather, it is a brief for the better understanding of the case study. We may or may not need more case studies in political science. It is hoped, however, that the foregoing discussion will encourage better case studies and a greater appreciation of their utility within the discipline.

It should also be apparent that the perceived hostility between case study and non-case study research is largely unjustified and, perhaps, deserves to be regarded as a misconception. Case studies may be small- or large-*N*, qualitative or quantitative, experimental or observational, synchronic or diachronic. The case study research design comports with any social-scientific theoretical framework including behavioralism, rational choice, institutionalism, and interpretivism. What differentiates the case study from the cross-unit study is its way of defining cases, not its analysis of those cases or its method of modeling causal relations. The case study research design constructs cases from a single unit while remaining attentive to inferences that span similar units outside the formal scope of investigation. Non-case study research designs construct cases across units exemplifying the principal causal inference. This difference in "casing" has important consequences, as we have shown. However, it does not render the case study epistemologically distinct from the cross-unit analysis. Indeed, the two modes of analysis are interdependent, and this is as it should be. One is at pains to imagine cross-unit research that does not draw upon case study work or case study work that disregards adjacent units. The strongest conclusion to arise from this methodological reflection concerns the complementarity of single-unit and cross-unit research designs.

[9] Granted, a good deal of cross-unit work involves the reanalysis of existing datasets; in this situation the barriers to entry are not so high.

REFERENCES

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2003. "An African Success Story: Botswana." In *In Search of Prosperity: Analytic Narratives on Economic Growth*, ed. Dani Rodrik. Princeton, NJ: Princeton University Press, 80–122.

Achen, Christopher H., and Duncan Snidal. 1989. "Rational Deterrence Theory and Comparative Case Studies." *World Politics* 41 (January): 143–69.

Adcock, Robert. 2002. "Determinism and Comparative-Historical Analysis: Clarifying Concepts and Retrieving Past Insights." Presented at the Annual Meeting of the American Political Science Association, Boston.

Alesina, Alberto, Edward Glaeser, and Bruce Sacerdote. 2001. "Why Doesn't the US Have a European-Style Welfare State?" *Brookings Papers on Economic Activity* 2: 187–277.

Allen, William Sheridan. 1965. *The Nazi Seizure of Power: The Experience of a Single German Town, 1930–1935*. New York: Watts.

Allison, Graham T. 1971. *Essence of Decision: Explaining the Cuban Missile Crisis*. Boston: Little, Brown.

Bates, Robert H., Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry Weingast. 1998. *Analytic Narratives*. Princeton, NJ: Princeton University Press.

Brady, Henry E., and David Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowman & Littlefield.

Cameron, David. 1978. "The Expansion of the Public Economy: A Comparative Analysis." *American Political Science Review* 72 (December): 1243–61.

Campbell, Donald T. [1975] 1988. "'Degrees of Freedom' and the Case Study." In *Methodology and Epistemology for Social Science*, ed. E. Samuel Overman. Chicago: University of Chicago Press.

Campbell, Donald T., and Julian Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin.

Collier, David, and James Mahoney. 1996. "Insights and Pitfalls: Selection Bias in Qualitative Research." *World Politics* 49 (October): 56–91.

Collier, Ruth Berins, and David Collier. [1991] 2002. *Shaping the Political Arena: Critical Junctures, the Labor Movement, and Regime Dynamics in Latin America*. Notre Dame, IN: University of Notre Dame Press.

Coppedge, Michael J. 2002. "Nested Inference: How to Combine the Benefits of Large-Sample Comparisons and Case Studies." Presented at the Annual Meeting of the American Political Science Association, Boston.

Dahl, Robert A. 1961. *Who Governs?: Democracy and Power in an American City*. New Haven, CT: Yale University Press.

Davidson, P. O., and C. G. Costello, eds. 1969. *N = 1: Experimental Studies of Single Cases*. New York: Van Nostrand Reinhold.

Dion, Douglas. 1998. "Evidence and Inference in the Comparative Case Study." *Comparative Politics* 30 (January): 127–45.

Eaton, Kent. 2002. *Politicians and Economic Reform in New Democracies.* University Park: Pennsylvania State University Press.

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press

Eckstein, Harry. [1975] 1992. "Case Studies and Theory in Political Science." In *Regarding Politics: Essays on Political Theory, Stability, and Change*. Berkeley: University of California Press.
Elman, Miriam, ed. 1997. *Paths to Peace: Is Democracy the Answer?* Cambridge: Cambridge University Press.
Feagin, Joe R., Anthony M. Orum, and Gideon Sjoberg. 1991. *A Case for the Case Study*. Chapel Hill: University of North Carolina Press.
Fenno, Richard F., Jr. 1978. *Home Style: House Members in their Districts*. Boston: Little, Brown.
George, Alexander. 1979. "Case Studies and Theory Development: The Method of Structured, Focused Comparison." In *Diplomacy: New Approaches in History, Theory, and Policy*, ed. Paul Gordon Lauren. New York: Free Press.
George, Alexander L., and Andrew Bennett. 2004. *Case Studies and Theory Development*, Cambridge, MA: MIT Press.
Gerber, Alan S., and Donald P. Green. 2000. "The Effects of Canvassing, Direct Mail, and Telephone Contact on Voter Turnout: A Field Experiment." *American Political Science Review* 94: 653–63.
Gerring, John. 2001. *Social Science Methodology: A Criterial Framework*. Cambridge: Cambridge University Press.
Gerring, John. Forthcoming. "Causation: A Unified Framework for the Social Sciences." *Journal of Theoretical Politics*.
Goertz, Gary. 2003. "The Substantive Importance of Necessary Condition Hypotheses." In *Necessary Conditions: Theory, Methodology and Applications*, ed. Gary Goertz and Harvey Starr. New York: Rowman and Littlefield.
Goertz, Gary, and Harvey Starr, eds. 2003. *Necessary Conditions: Theory, Methodology and Applications*. New York: Rowman and Littlefield.
Hall, Peter A. 2003. "Aligning Ontology and Methodology in Comparative Politics." In *Comparative Historical Analysis in the Social Sciences*, ed. James Mahoney and Dietrich Rueschemeyer. Cambridge: Cambridge University Press.
Hartz, Louis. 1955. *The Liberal Tradition in America*. New York: Harcourt, Brace & World.
Hedstrom, Peter, and Richard Swedberg, eds. 1998. *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.
Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.
Johnson, Chalmers. 1983. *MITI and the Japanese Miracle: The Growth of Industrial Policy, 1925–1975*. Stanford, CA: Stanford University Press.
Kaufman, Herbert. 1960. *The Forest Ranger: A Study in Administrative Behavior*. Baltimore: Johns Hopkins University Press.
King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.
Kuhn, Thomas S. [1962] 1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
Lane, Robert. 1962. *Political Ideology: Why the American Common Man Believes What He Does*. New York: Free Press.
Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1948. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. New York, Columbia Univ. Press.
Lieberman, Evan S. 2003. "Nested Analysis as a Mixed-Method Solution to Cross-National Research." Ms. Princeton University.
Lieberson, Stanley. [1991] 1992. "Small N's and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases." In *What Is a Case? Exploring the Foundations of Social Inquiry*, ed. Charles S. Ragin and Howard S. Becker . Cambridge: Cambridge University Press.
Lieberson, Stanley. 1994. "More on the Uneasy Case for Using Mill-Type Methods in Small-N Comparative Studies." *Social Forces* 72 (June): 1225–37.
Lijphart, Arend. 1968. *The Politics of Accommodation: Pluralism and Democracy in the Netherlands*. Berkeley: University of California Press.
Lijphart, Arend. 1975. "The Comparable Cases Strategy in Comparative Research." *Comparative Political Studies* 8 (July): 158–77.
Lipset, Seymour Martin. 1963. *The First New Nation: The United States in Historical and Comparative Perspective*. New York: Basic Books.
Lynd, Robert Staughton, and Helen Merrell Lynd. [1929] 1956. *Middletown: A Study in American Culture*. New York: Harcourt, Brace.
Mahoney, James, and Dietrich Rueschemeyer, eds. 2003. *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press.
Malinowski, Bronislaw. [1922] 1984. *Argonauts of the Western Pacific*. Prospect Heights, IL: Waveland.
McKeown, Timothy. 1999. "Case Studies and the Statistical World View." *International Organization* 53 (Winter): 161–90.
Njolstad, Olav. 1990. "Learning From History? Case Studies and Limits to Theory-Building." In Olav Njolstad, (ed.), *Arms Races: Technological and Political Dynamics* (Thousand Oaks, CA: Sage) 220–46.
Popper, Karl. 1969. *Conjectures and Refutations*. London: Routledge and Kegan Paul.
Pressman, Jeffrey L., and Aron Wildavsky. 1973. *Implementation*. Berkeley: University of California Press.
Przeworski, Adam, and Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley.
Przeworski, Adam, Michael Alvarez, Jose Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Material Well-Being in the World, 1950–1990*. Cambridge: Cambridge University Press.
Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California.
Ragin, Charles C. 1992. "Cases of 'What Is a Case?" In *What Is a Case? Exploring the Foundations of Social Inquiry*, ed. Charles C. Ragin and Howard S. Becker. Cambridge: Cambridge University Press.
Ragin, Charles C. 1997. "Turning the Tables: How Case-Oriented Research Challenges Variable-Oriented Research." *Comparative Social Research* 16: 27–42.
Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.
Ragin, Charles C., and Howard S. Becker, eds. 1992. *What Is a Case? Exploring the Foundations of Social Inquiry*. Cambridge: Cambridge University Press.
Reilly, Benjamin. 2001. *Democracy in Divided Societies*. Cambridge: Cambridge University Press.
Roberts, Clayton. 1996. *The Logic of Historical Explanation*. University Park: Pennsylvania State University Press.
Rodrik, Dani, ed. 2003. *In Search of Prosperity: Analytic Narratives on Economic Growth*. Princeton, NJ: Princeton University Press.
Rueschemeyer, Dietrich, and John D. Stephens. 1997. "Comparing Historical Sequences: A Powerful Tool for Causal Analysis." *Comparative Social Research* 16: 55–72.
Sartori, Giovanni. 1984. "Guidelines for Concept Analysis." In *Social Science Concepts: A Systematic Analysis*. Beverly Hills, CA: Sage, 15–48.
Skocpol, Theda. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.
Smith, Rogers M. 1997. *Civic Ideals: Conflicting Visions of Citizenship in U.S. History*. New Haven, CT: Yale University Press.
Stratmann, Thomas, and Martin Baur. 2002. "Plurality Rule, Proportional Representation, and the German *Bundestag*: How Incentives to Pork-Barrel Differ Across Electoral Systems." *American Journal of Political Science* 46 (July): 506–14.
Taylor, Charles. 1970. "The Explanation of Purposive Behavior." In *Explanation in the Behavioral Sciences*, ed. Robert Borger and Frank Cioffi. Cambridge: Cambridge University Press.
Temple, Jonathan. 1999. "The New Growth Evidence." *Journal of Economic Literature* (March) 37: 1112–56.
Thompson, E. P. 1963. *The Making of the English Working Class*. New York: Vintage Books.
Waldner, David. 2002. "Anti Anti-Determinism: Or What Happens When Schrodinger's Cat and Lorenz's Butterfly Meet Laplace's Demon in the Study of Political and Economic Development." Presented to the Annual Meeting of the American Political Science Association, Boston.
Whyte, William Foote. [1943] 1955. *Street Corner Society: The Social Structure of an Italian Slum*. Chicago: University of Chicago Press.
Yin, Robert K. 1994. *Case Study Research: Design and Methods*. Newbury Park, CA: Sage.

https://doi.org/10.1017/S0003055404001182 Published online by Cambridge University Press