for a quantitative analysis, reducing the diversity of the original information in order to conform to uniform parameters. It is not clear that much would have been gained by doing so. In the event, his study is limited to qualitative forms of analysis.

This issue is treated at length in a later chapter. For the moment, note the fact that case study research often provides a piece of evidence pertaining to A, another piece of evidence pertaining to B, and a third pertaining to C. There may be many observations (in total), and they may all be relevant to a central causal argument, even though they are not directly comparable to one another. These are referred to in Chapter Seven as noncomparable observations.

In summary, large-N cross-case research is quantitative, by definition. This much conforms to usual perceptions. However, case study research may be either qualitative or quantitative, or both, depending upon the sort of within-case evidence that is available and relevant to the question at hand. Consequently, the traditional association of case study work with qualitative methods is correctly regarded as a methodological affinity, not a definitional entailment. It is true sometimes, but not all the time.

3

What Is a Case Study Good For?

Case Study versus Large-N Cross-Case Analysis

In Chapter Two, I argued that the case study approach to research is most usefully defined as an intensive study of a single unit or a small number of units (the cases), for the purpose of understanding a larger class of similar units (a population of cases). This was put forth as a minimal definition of the topic. In this chapter, I proceed to discuss the nondefinitional 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 our minimal definition of the concept.

The case study research design exhibits characteristic strengths and weaknesses relative to its large-N cross-case cousin. These trade-offs 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 trade-offs also hinge on the shape of the empirical universe, that is, on (5) whether the population of cases under study is heterogeneous or homogeneous, (6) whether

1 My intention was 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. For further discussion of minimal definitions, see Gerring (2001: Chapter 4); Gerring and Barresi (2003); and Sartori (1976).

2 These additional attributes might also be understood as comprising an ideal-type ("maximal") definition of the topic (Gerring 2001: Chapter 4; Gerring and Barresi 2003). Recent evaluations of the strengths and weaknesses of case study research can be found in Flyvbjerg (2004); Levy (2002a); and Verschuren (2001).
I. Thinking about Case Studies

### Table 3.1. Case study and cross-case research designs: considerations

<table>
<thead>
<tr>
<th>Affinity</th>
<th>Case Study</th>
<th>Cross-Cause Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Research goals</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Hypothesis</td>
<td>Generating</td>
<td>Testing</td>
</tr>
<tr>
<td>2. Validity</td>
<td>Internal</td>
<td>External</td>
</tr>
<tr>
<td>3. Causal insight</td>
<td>Mechanisms</td>
<td>Effects</td>
</tr>
<tr>
<td>4. Scope of proposition</td>
<td>Deep</td>
<td>Broad</td>
</tr>
<tr>
<td><strong>Empirical factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Population of cases</td>
<td>Heterogeneous</td>
<td>Homogeneous</td>
</tr>
<tr>
<td>6. Causal strength</td>
<td>Strong</td>
<td>Weak</td>
</tr>
<tr>
<td>7. Useful variation</td>
<td>Rare</td>
<td>Common</td>
</tr>
<tr>
<td>8. Data availability</td>
<td>Concentrated</td>
<td>Dispersed</td>
</tr>
<tr>
<td><strong>Additional factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Causal complexity</td>
<td>Indeterminate</td>
<td></td>
</tr>
<tr>
<td>10. State of the field</td>
<td>Indeterminate</td>
<td></td>
</tr>
</tbody>
</table>

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 is 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 3.1.

I argue that other issues impinging upon the research format, such as (9) causal complexity and (10) the state of research in a given field, are indeterminate in their implications. Sometimes these factors militate toward a case study research design; at other times, toward a cross-case research design.

To reiterate, the eight trade-offs depicted in Table 3.1 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.

It should be stressed that each of these trade-offs 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 nine additional factors also rightly influence a writer’s choice of research design, and they may lean in the other direction. Ceteris is 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.

### 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.” 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. 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 protan. 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.

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. Darwin’s insights into the process of human evolution came after his

---

3 Popper (1963).
4 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 2006), though it may be doubted whether there are generalizable results.
5 Gerring (2001: Chapter 10). The trade-off between these two styles of research is implicit in Achen and Snidal (1989); the authors criticize the case study for its deficits in the latter genre but also acknowledge the benefits of the case study along the former dimension (ibid., 167–8). Reichenbach also distinguishes 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.
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. Levi-Strauss's structuralist theory of human cultures built on the analysis of several North and South American tribes. Douglass North's neoinstitutionalist 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). 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 Vandenbroucke, "is to recognize the unexpected." This is where discovery begins.

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.

Traditionally, scientific methodology has been defined by a segregation of conjecture and refutation. One should not be allowed to contaminate the other. Yet in the real world of social science, evidence is often associated with perspiration. "Lightbulb" 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. 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 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.

A large-N cross-case 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.

7 North and Weingast (1989); North and Thomas (1973).
9 For discussion of this trade-off in the context of economic growth theory, see Temple (1999: 120).
10 Geddes (2003); King, Keohane, and Verba (1994); Popper (1934/1968).
11 Ragin (1992b).
12 Eckstein (1975); Ragin (1992a, 1997); Rueschemeyer and Stephens (1997).
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 will be discussed. 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 predic-
tions on within-case variation that would allow one to reject a hypothesis
through pattern matching (without additional cross-case evidence). The-
ory 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 (see
Chapter Five). Thus, one is unlikely to reject a hypothesis, or to consider
it definitively proved, on the basis of a single case.

Harry Eckstein himself acknowledged that his argument for case stud-
ies as a form of theory confirmation was 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.\textsuperscript{13} I suspect that this is still more or less true. Indeed, it is true even of
experimental case studies in the natural sciences. "We must recognize,"
ote 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.\ldots\] Even though we rec-
ognize experimentation as the basic language of proof\ldots\ 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 rep-
resent 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\ldots\ avoids crucial experiments and instead studies dimensional
relationships and interactions along many degrees of the experimental variables.\textsuperscript{14}

A single case study is still a single-shot affair – a single example of a larger
phenomenon.

The trade-off 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

\textsuperscript{13} Eckstein (1975).
\textsuperscript{14} Campbell and Stanley (1963: 3).

What Is a Case Study Good For?

Case study research is a boon to new conceptualizations just as it is a bane
to falsification.

Validity: Internal versus External

Questions of validity are often distinguished according to those that are
\textit{internal} to the sample under study and those that are \textit{external} (i.e., apply-
ing to a broader – unstudied – population). The latter may be conceptual-
ized as a problem of representativeness between sample and 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,
as discussed in Chapter Five). Case study research suffers problems of rep-
resentativeness 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?\textsuperscript{15} Is
Middletown representative of other cities in America?\textsuperscript{16} 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 valid-
ity. 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 respect: 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 trade-off between external and inter-
nal validity, like other trade-offs, as intrinsic to the cross-case/single-case
choice of research design.

Causal Insight: Causal Mechanisms versus Causal Effects

A third trade-off 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 \textit{effect}; the second concerns
the investigation of a causal \textit{mechanism} (i.e., a pathway from $X$ to $Y$).

\textsuperscript{15} Lane (1962).
\textsuperscript{16} Lynd and Lynd (1929/1956).
I. Thinking about Case Studies

When I say "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 of that point estimate.\(^{17}\) 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 possible 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.)\(^{18}\) 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.\(^{19}\) That is, 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 specify the model accurately, to identify possible instruments for the regressor of interest (if there are problems of endogeneity), or to interpret the results.\(^{20}\) Thus, causal mechanisms are presumed in every estimate of a mean (average) causal effect.

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 – for example, 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,\(^{21}\) 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.\(^{22}\) Barney Glaser and Anselm Strauss point out that in field work "general relations are often discovered in vivo; that is, the field worker literally sees them occur."\(^{23}\) 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.\(^{24}\)

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 (see Chapter Seven).

One example of case study evidence calling into question a general theoretical argument on the basis of an investigation of causal mechanisms concerns the theory of rational deterrence. Deterrence theory, as

---

\(^{17}\) The correct estimation of a causal effect rests upon the optimal choice among possible estimators. It therefore follows from the previous discussion that sample-based analyses are also essential for choosing among different estimators – as judged by their relative efficiency and bias, among other desiderata. See Kennedy (2003) for a discussion of these issues.

\(^{18}\) 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 United States alone, rather than cross-nationally. However, since the resulting generalization does not extend beyond the unit in question, this is not a case study in the usual sense.

\(^{19}\) Achen (2002); Dealer (1991); Elster (1988); George and Bennett (2005); Gerring (2005); Hedstrom and Swedberg (1998); Mahoney (2001); Tilly (2001).

\(^{20}\) 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."

\(^{21}\) Goldstone et al. (2000).

\(^{22}\) This has something to do with the existence of process-tracing evidence, a matter to be discussed later. 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.

\(^{23}\) Glaser and Strauss (1967: 40).

it was understood in the 1980s, presupposes a number of key assumptions, namely, that "actors have exogenously given preferences and choice options, and [that] they seek to optimize preferences in light of other actors’ preferences and options . . . . [that] variation in outcomes is to be explained by differences in actors’ opportunities . . . . [that] the state acts as if it were a unitary rational actor." A generation of case studies, however, suggests that, somewhat contrary to theory, (a) international actors often employ "shortcuts" in their decision-making processes (i.e., they do not make decisions de novo, based purely on an analysis of preferences and possible consequences); (b) a strong cognitive bias exists because of "historical analogies to recent important cases that the person or his country has experienced firsthand" (e.g., "Somalia = Vietnam"); (c) "accidents and confusion" are often manifest in international crises; (d) a single important value or goal often trumps other values (in a hasty and ill-considered manner); and (e) actors’ impressions of other actors are strongly influenced by their self-perceptions (information is highly imperfect). In addition to these cognitive biases, there is a series of psychological biases. In sum, while the theory of deterrence may still hold, the causal pathways contained in this theory seem to be considerably more variegated than previous work based on cross-case research had led us to believe. In-depth studies of particular international incidents have been helpful in uncovering these complexities.

Dietrich Rueschemeyer and John Stephens offer a second 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 post-colonial regimes. In particular, the authors investigate the diffusion hypothesis, that democracy was enhanced by "the transfer of British governmental 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

27 George and Smoke (1974: 504). For another example of case study work that tests theories based upon predictions about causal mechanisms, see McKeown (1983).

What Is a Case Study Good For?

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.

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 greater social welfare spending. The question then

29 A third example of case study analysis focused on causal mechanisms concerns policy delegation within coalition governments. Michael Thies (2001) tests two theories about how parties delegate power. The first, known as ministerial government, supposes that parties delegate ministerial portfolios in toto to one of their members (the party whose minister holds the portfolio). The second theory, dubbed managed delegation, supposes that members of a multiparty coalition delegate power, but also actively monitor the activity of ministerial posts held by other parties. The critical piece of evidence used to test these rival theories is the appointment of junior ministers (JMs). If JMs are from the same party as the minister, we can assume that the ministerial government model is in operation. If the JMs are from different parties, Thies infers that a managed delegation model is in operation, where the JM is assumed to perform an oversight function regarding the activity of the bureau in question. This empirical question is explored across four countries – Germany, Italy, Japan, and the Netherlands – providing a series of case studies focused on the internal workings of parliamentary government. (I have simplified the nature of the evidence in this example, which extends not only to the simple presence or absence of cross-partisan JMs but also to a variety of additional process-tracing clues.) Other good examples of within-case research that shed light on a broader theory can be found in Canon (1999); Martin (1992); Martin and Swank (2004); and Young (1999).
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, is that increased trade openness leads to greater domestic economic vulnerability to external shocks (due, for instance, to changing terms of trade). If that is true, 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. 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.

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 judgments 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 trade-off.

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 trade-off is fairly commonsensical. 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—for example, 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), as discussed in Chapter Five. 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. 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/nonoccurrence 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.

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. 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

---

30 Cameron (1978).
32 For additional examples of this nature, see Feng (2003); Papyrakis and Gerlagh (2003); and Ross (2001).
33 Eckstein (1975: 122).
34 My use of the term “thick” is somewhat different from the usage in Geertz (1973).
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.\(^{36}\)

Indeed, I believe that a number of traditional issues related to case study research can be understood as the product of this basic trade-off. 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 slight relationship to the phenomena of interest.\(^{37}\) These associated congratulation and critiques may be understood as a conscious choice on the part of case study researchers to privilege depth over breadth.

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 researcher, as just reviewed, 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 assumptions.”\(^{38}\) 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.\(^{39}\)

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.\(^{40}\)

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. Probably, one gains greater insight into causal patterns throughout the population by examining these cases in detail than by undertaking a 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.

---

36 Ragan (1987: Chapter 2). Herbert Blumer’s (1969: Chapter 7) complaints, however, are more far-reaching.


40 Hersen and Barlow (1976: 11).
I. Thinking about Case Studies

The issue of population heterogeneity/homogeneity may be understood, therefore, as a trade-off 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—i.e., 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. 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. This is not simply a question of attaining statistical significance. The ceteris paribus assumption at the core of all causal analysis is brought into question (see Chapter 6). 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 trade-off between \( N \) and \( K \) in a particular empirical realm, and about the ability to model case-heterogeneity through statistical legerdemain.

Before concluding this discussion, it is important to point out that researchers' judgments 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 a 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 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 is 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. 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.

Causal Strength: Strong versus Weak

Regardless of whether the population is homogeneous or heterogeneous, relationships are easier to study if the true causal effect is strong, rather

---

41 Shalev (1998).
42 To be sure, if adjacent cases are identical, the phenomenon of interest is important. In that case 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 has some degree of heterogeneity (cases are not identical), some stochastic element. Thus, the theoretical possibility of identical, invariant cases is rarely met in practice.
43 Gutting (1980); Hall (2003); Kuhn (1962/1970); Wolin (1968).
than weak. Causal “strength” refers here to the magnitude and consistency of X’s effect on Y across a population of cases. (It involves both the shape of the evidence at hand and whatever priors might be relevant to an interpretation of that evidence.) Where X1 has a 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, that is, 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 disproved with a single case.44 For example, the reigning theory of political instability once stipulated that only in countries that were relatively homogeneous, or where existing heterogeneity was mitigated by cross-cutting cleavages, would social peace endure.45 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.46 (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. Stochastic variation is ruled out.

---

44 Dion (1998).
45 Almond (1956); Bentley (1908/1967); Lipset (1960/1963); Truman (1951).
46 Lijphart (1968); see also Lijphart (1969). For additional examples of case studies disconfirming general propositions of a deterministic nature, see Allen (1963); Lipset, Trow, and Coleman (1936); Njolstad (1990); and the discussion in Rogowski (1995).
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 — that is, 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 nonrevolutions and just a few revolutions. 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 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.49

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 evenly 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 (X)

49 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 (World War I, World War II); single nontransferable 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 Buffet); ethnic heterogeneity (Papua New Guinea).

varies, and there is no corresponding change in other factors that might affect the outcome. (The quasi-experimental qualities of the case may be the result of a manipulated treatment or a treatment that occurs naturally. These issues are explored in Chapter Six.) 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. (Many of the issues discussed in Chapters Five and Six are relevant to this discussion of what constitutes "useful variation." Thus, I have touched upon these issues only briefly in this section.)

Data Availability

I have left the most prosaic factor for last. Sometimes, one's choice of research design is driven by the quality and quantity of information that is currently available, or could easily be gathered, on a given question. This is a practical matter and is separate from the actual shape of the empirical universe. It concerns, rather, what we know about the former at a given point in time.50 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

50 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.
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 incomparable formats across a population of cases, then a case study mode of analysis is almost unavoidable. But if it is evenly distributed across the population—that is, if 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). 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 welfare 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.53

Here, the incommensurability of the evidence militates in favor of a case study format. In the event that the authors (or subsequent analysts) discover a coding system that provides reasonably valid cross-case measures

---

51 Gerring (2006c).
52 Mulligan, Gil, Sala-i-Martin (2002: 13).
53 Bollen (1993); Bowman, Lehoucq, and Mahoney (2005); Munck and Verkuilen (2002); Treier and Jackman (2005).
54 Bowman, Lehoucq, and Mahoney (2005).
55 Bollen (1993); Treier and Jackman (2005).
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 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 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—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."  

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 design. 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.

Causal Complexity

Not all factors that impinge upon the choice of research designs have clear affinities to case study or cross-case study research. Others are indeterminate in their implications. Whether these factors favor the focused analysis of a few cases or a relatively superficial analysis of many cases depends upon issues that are difficult to generalize about.

Let us begin with the vexed question of causal complexity. Case study researchers often laud their favored method for its better grasp of complex causes, while critics claim that the more complex the causal relationship, the more necessary is cross-case evidence. Intuitively, both positions seem plausible, and much evidently depends upon the interpretation of "complexity," which might refer to probabilistic (rather than invariant) causal patterns, necessary and/or sufficient causes, nonlinear relationships, multiple causes ("equipartition"), nonadditive causal interrelationships, causal sequences (where causal order affects the outcome of interest), a large number of plausible causes (the problem of overdetermination), and many other things besides. Indeed, "complexity," as the term is used in social science circles, seems to refer to any feature of a causal problem that does not fit snugly with standard assumptions of linearity, additivity, and independence. As such, it is a red herring, for it has no determinate meaning.

Some kinds of causal complexity, like necessary and sufficient conditions, may militate in favor of a case study research design, as argued earlier in this chapter (see the section on causal strength). Others, I will argue, are indeterminate. That is, sometimes complex causal relationships are rendered visible in case study research, and we are able to parse out the independent causal effects of each factor (which may depend on their position in an extended causal chain). This is what case study research does, if it is done well and if the chosen case is amenable to that style of research. But oftentimes, this is simply not feasible. Similarly, sometimes one is able to model complex causal relationships in a cross-case setting, and sometimes not. In short, it all depends.


57 Abbott (1990); George and Bennett (2003); Ragin (1987: 54; 2000: Chapter 4); Ruef (2003).

58 Goldthorpe (1997); King, Keohane, and Verba (1994); Lieberson (1985).
Let us explore an example. Suppose one is interested in the influence of fiscal pressures on social revolution—the idea that as governments get more strapped for cash, they are likely to seek to raise taxes, which, in turn, may spark revolt. A nice (confirming) case study would show precisely that, without any interfering (confounding) factors. It would be eventful, in a quasi-experimental way (see Chapter Six). An intervention (treatment) would occur—increasing budget deficits, followed by increasing taxes—and the result could be observed. However, a bad (confirming) case would show that lots of things were happening at the same time that could also have caused revolution. As it happens, lots of things do tend to happen together during critical junctures like revolutions, and so it is often quite difficult to tease out real and spurious causal effects. In statistical terms, this may be understood as a problem of collinearity. Now, let us suppose that you have at your disposal 100 countries, with annual measurements of fiscal pressure, tax instruments, as well as various confounders (controls). Collinearity is still a formidable problem. But with a great deal of cross-case evidence, there is at least a fighting chance that it can be overcome, while there is little chance of overcoming it in most case study settings. (Indeed, some statisticians have looked upon the problem of collinearity as a problem of data insufficiency.)

The general point remains. “Complexity,” by itself (keeping in mind that complexity can mean many things), does not favor either a case study or a cross-case approach to causal analysis.

The State of the Field

Another sort of contextual consideration concerns the state of research on a given topic within a field. 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-case studies are hegemonic may be desperately in need of in-depth studies focused on understudied cases.

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 a particular field. If both case study and cross-case methods have much to recommend them (an implicit assumption of this book), 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. But there is little that we can say about this desideratum in general, since it depends on the shape of an individual field or subfield.
In the opening pages of this book, I highlighted the rather 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, as demonstrated in Chapter One. How, then, can one make sense of this discrepancy between methodological theory and methodological praxis? This was the question animating Part One of the book.

The torment of the case study begins with its definitional penumbra, as described in Chapter Two. Frequently, this key term is conflated with a disparate set of methodological traits that are not definitionally entailed. Our first task, 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 (or a small set of cases) with an aim to generalize across a larger set of cases of the same general type. If the inference pertains to nation-states, then a case study would focus on one or several nation-states (while a cross-case study would focus on many nation-states at once). If the inference pertains to individuals, then a case study would focus on one or several individuals (while a cross-case study would focus on many individuals at once). And so forth.

It follows from this definition that case studies may be small- or large-N (since a single case may provide few or many observations), qualitative or quantitative, experimental or observational, synchronic or diachronic. It also follows that the case study research design comports with any macro-theoretical framework or paradigm— for example, behavioralism, rational choice, institutionalism, or interpretivism. It is not epistemologically
II. Doing Case Studies

II. Doing Case Studies

by offering a more carefully bounded definition of the method it might be rescued from some of its ambiguities. It is also 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 and in some research settings, but not all.

In the second section of the book, I turn to practical questions of research design. How does one employ the intensive study of a single case, or a small number of cases, to shed light on a broader class of cases?

Chapter Four addresses preliminary issues pertaining to this quest. Chapter Five examines the problem of case selection. Chapter Six examines the problem of internal validity through the prism of experimental research designs. Chapter Seven approaches the problem of internal validity through the use of a rather different approach called process tracing. The epilogue addresses research design elements of single-outcome studies — where a single outcome, rather than a broader class of outcomes, is of primary interest.

1 Epistemological differences between case study and cross-case work are a theme in Orum, Feagin, and Sjoberg (1991: 22).

1. 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.

In other respects, the predicament of the case study is not merely definitional but inheres in the method itself. To study a single case with intent to shed light upon other cases brings in its train several methodological ambiguities. First, the concept of a case study is dependent upon the particular proposition that one has in mind, a proposition that may change through time (as the study is digested by the academic community) or even within a given study (as the author changes her level of analysis). Second, the boundaries of a case are sometimes — despite the researcher’s best efforts — open-ended. This is particularly true of temporal boundaries, which may extend into the future and into the past in rather indefinite ways. Third, case studies usually build upon a variety of covariational evidence; there is no single type of case study, but rather five (see Table 2.4).

The travails of the case study are rooted, additionally, in an insufficient appreciation of the methodological trade-offs that this method calls forth, as discussed in Chapter Three. At least eight characteristic strengths and weaknesses must be considered (see Table 3.1). Ceteris paribus, case studies are more useful when the purpose of research is hypothesis generating rather than hypothesis testing, 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 is heterogeneous rather than homogeneous, when causal relationships are strong rather than weak, when useful variation on key parameters is rare rather than commonplace, and when good-quality evidence is concentrated rather than dispersed. Causal complexity and the existing state of a field of research may also influence a researcher’s choice to adopt a single-case or cross-case research design, though their methodological implications are equivocal.

The objective of the first section of the book was to restore a sense of meaning, purpose, and integrity to the case study method. It is hoped that