

Original Article

Qualitative and Multi-Method Research

Fall 2020, Vol. 18, No. 2

<https://doi.org/10.5281/zenodo.4046562>

Shadow Cases in Comparative Research¹

Hillel Soifer

Temple University

Scholars conducting small-N research often deploy ancillary or peripheral cases that are intended to evaluate the more general validity of the findings of their core case studies. Yet we lack a clear set of methodological guidelines for these ancillary cases.² Drawing on scholarship in the comparative politics subfield for examples, I identify two broad approaches to ancillary cases—the *case illustration* and the *shadow case study*. The case illustration, which consists of showing that outcomes in additional cases match what we expect given theory generated from the core case and the values of independent variables in those additional cases, is widely used to evaluate the generality of findings. Part One of the paper argues that this common practice is not as analytically valuable as it could be. I show that the case illustration approach faces challenges to both external and internal validity and propose a larger-N alternative to assessing claims of generality instead.

In contrast to the case illustration's exploration of correlation, the shadow case study investigates a case more deeply, drawing on techniques of case study research. This entails the deeper investigation of cases that characterizes a well-conducted case study. It goes beyond the scoring of cases on an outcome of interest, a proposed explanation, and other key variables.³ Part Two provides a concrete and precise definition of the shadow case study that centers on its generation of inference from within-case analysis and its role as an ancillary case principally intended to shed light on external validity of

findings generated in the core cases. It also discusses the constraints under which it is investigated. The emphasis on within-case analysis differentiates it from the case illustration, and the constraints under which it is conducted distinguish it from the core case study that sits at the heart of small-N research, rendering it a diminished subtype (Collier and Levitsky 1997) of the case study.⁴ In Part Three, I show that in addition to assessing scope conditions and providing confirmatory illustrations, shadow case studies can be used to explore observable implications, learn from outliers, or assess alternative explanations. I identify conditions under which each of these tasks might be especially useful for the researcher and briefly suggest how cases for these tasks might be selected.

The “Case Illustration,” its Limitations, and the Advantages of a Large-N Alternative

Case study and small-N researchers often complement their core cases with what Goertz describes as “superficial case studies used to determine generalization” (Goertz 2017, 239). These commonly entail scoring of cases on the independent and dependent variables to show that the patterns of these scores observed in a series of cases align with the expectations of theory developed and tested in the core case. For example, Ziblatt complements his study of conservative parties and democratization in nineteenth-century Europe with glances at sixteen Latin American countries, some Southeast

1 I thank Rachel Riedl, Tasha Fairfield, Adam Ziegfeld, Sean Yom, Marcus Kreuzer, Erica Simmons, and anonymous reviewers for helpful comments. I also benefitted from feedback from audiences at the IQMR Authors' Workshop at Syracuse University, the Princeton Politics Department Qualitative Research Colloquium, and the Northwestern University Comparative Historical Social Science Seminar. I am especially grateful to Cassandra Emmons for research assistance and several iterations of helpful comments.

2 Throughout, I use the terms *ancillary case* or *peripheral case* to refer to cases other than the *core cases* that are the central focus of the small-N analysis. This is, as discussed further below, the broad category within which I situate both *case illustrations* and *shadow case studies*.

3 See Goertz (2017, 1-6 and *passim*) for a similar discussion of within-case analysis. Like Goertz, and Fairfield and Charman (forthcoming), I conceptualize within-case analysis as encompassing, but not limited to, process tracing evidence about causal mechanisms.

4 I thank Daniel Encinas for suggesting this way to position the shadow case study and use the full term (shadow case study) throughout, rather than “shadow case” to highlight that what I propose is a type of case study.

Asian cases, and those embroiled in the Arab Spring, and concludes that (emphasis added) “*the point appears again and again ... ultimately we cannot understand the course of a country’s democratization trajectory without factoring in the role of a democracy’s adversaries*” (Ziblatt 2017, 361). I call these exercises that dominate existing practice *case illustrations* to highlight the fact that they do not entail any detailed within-case analysis of the ancillary cases.

Scholars select these illustrative cases in a range of ways. Many explore other cases in the region or time period covered by the core cases, building on the implicit assertion that region or time period operates as a scope condition on their theoretical claims (Levitsky 2003; Ziblatt 2006). Other scholars (Handlin 2017; Mylonas 2012; Ziblatt 2017) explicitly mention the importance of using out-of-region cases because they seek to evaluate their claims at a more general level.

In a related but distinct analytical exercise, case illustrations can also be used to explore explicitly defined scope conditions, whether via confirming that the predictions of the theory hold in cases that fall within that scope (Truex 2016, 178) or showing that predictions are disconfirmed outside of them (O’Neill 2005, 236). Though they differ in the extent to which generality is asserted and in how explicitly scope is discussed, these exercises are fundamentally common in their use of evidence about predictions and outcomes in ancillary cases to assess the external validity of claims derived and evaluated in the core cases.

Yet the case illustration is characterized by two fundamental limitations. First, the logic of case selection is rarely made fully explicit. For example, Daly (2016) argues that her account should generalize to civil wars that are non-ethnic and multi-party but does not justify how her ancillary investigations of Nicaragua, Guatemala, Peru, Uganda, and Indonesia were chosen from among the many cases that fit that characterization. The result is that inferences about external validity are, by definition, limited.

Some scholars do provide explicit logics of case selection. Most common perhaps is to select cases in terms of the overall population to which the claim is intended to apply, selecting especially salient cases as settings to assess generalizing claim. Pearlman (2011), for example, uses dramatic liberation struggles in South Africa and Northern Ireland as ancillary cases in her study of the Palestinian National Movement, and Singh (2015, 244) glances at “the two most prominent cases of subnationalism in the world” to examine whether dynamics parallel

to those she finds in India can be observed.

Scholars can also select ancillary cases to explore the outcome of interest at the full range of values of the independent variable, as seen in Barnes (2016). Here, to the extent that outcomes are scored as expected in cases that take a wide range of values of her independent variable, the case illustrations can increase our confidence in the generality of her explanation for women’s legislative collaboration. Yet other scholars provide justifications of case selection that provide weaker grounds for generalization. For example, Wilkinson justifies selecting the ancillary cases of nineteenth-century Ireland, post-independence Malaysia, and post-Communist Romania to complement his account of the political incentives underpinning religious violence in India on the grounds that they represent “one case from each of the three great waves of democratization” (Wilkinson 2004, 204). It is unclear, however, what inference we can draw from the fact that cases vary on this dimension, which is not explicitly connected to the theory developed in the core case of India.

Yet even when this sort of explicit discussion of case selection increases the external validity of these exercises, a more fundamental limitation remains. These case studies tend to do little more than score cases on the independent and dependent variable and show that the outcomes take on the values predicted by the theory given the value taken by the independent variable. These exercises, then, principally rely on evidence of correlation between an independent and dependent variable in each of a series of cases. But if we take seriously the ontological and epistemological underpinnings of a process-oriented approach to causation, we ought to discount this pattern-matching evidence very heavily for falling short of the standards for causal claims developed by qualitative methodologists (Lange 2013, 17).

Scholars are aware of this limitation. For example, Boas (2016, 186) notes that his examination of ancillary cases “is generally insufficient to uncover the mechanisms” by which causation occurred. The novelty in what I propose, then, is not in the recognition of the problem, but in the provision of suggested new ways in which ancillary cases can be used. Rather than resigning ourselves to view the use of ancillary cases as providing little inferential leverage, I suggest some ways forward ways in which insights from additional cases might inform the findings of the core case study analyses. I first propose the replacement of the case illustration with a more systematic cross-case assessment of generality—a

move to a mixed-methods approach that complements case studies with a population-wide generalizing analysis. In the remainder of the paper I identify a set of other uses to which ancillary cases might be put that draw on the tools of case study analysis to increase internal validity and thus provide stronger support for inferences about external validity.

Rather than using case illustrations, I suggest that scholars might instead assess generality or evaluate scope conditions through a larger-N exercise. This would entail coding *all* cases to which the theory is expected to apply on the independent and dependent variables and seeing whether the expected relationship holds or is stronger than alternative accounts.⁵ This approach provides a satisfactory assessment of external validity by explicitly defining a universe of cases and evaluating the performance of the theory in all of them. And though it doesn't address concerns about internal validity in any of the cases included—it relies even more baldly on correlation as evidence for claims about the performance of a theory—it makes this limitation transparent to the reader and forces the author to “bite the bullet” of resting the argument precisely on this type of inference and evidence in an explicit manner rather than relying on imprecise statements of the inference that can be drawn from case illustrations.

Some scholars, of course, already rely on a mixed-methods approach of this type. See, for example, Grzymala-Busse (2015), who explicitly states that “the argument relies chiefly on the paired comparisons” of cases (p. 14), while using large-N analysis (in the Appendix) to assess generality. Other scholars (Mylonas 2012, 170; Handlin 2017, 251) reject the large-N approach to generalization for providing no evidence of causal mechanisms and rely instead on case illustrations. Yet I have shown above that the case illustrations they choose instead also tend to fail to establish causal links precisely because they struggle with identical issues of internal validity. The difference between these approaches, then, is not in the standards of evidence for causation that they use. Instead, it is in the extent of quantification and in the breadth of cases covered. Thus, these critiques seem overstated.

Turning to a large-N approach to generalization depends on the feasibility of coding a large number of cas-

es on both the independent and dependent variables to examine the correlation between them. This may be the case either because they are prohibitively hard to identify or code, or because the population of cases to which a theory is proposed to generalize is fairly small. Where identifying and coding a sufficient number of additional cases is possible, scholars might benefit from turning to a population-wide exercise in assessing generality, rather than doing so in a more limited number of case illustrations.

The Shadow Case Study—A Definition

If assessing generality is not the best use for ancillary cases, how might scholars instead deploy them? Departing from the presumption that scholars engaged in small-N research might prefer to explore the generality of their claims in ways that highlight the inferential power of within-case analysis, the next step is to define the *shadow case study*. Despite the massive flowering of guidelines for case study research, we lack a definition in existing methodological scholarship for a shadow case. The closest approximation comes from Gerring, who suggests that the peripheral or ancillary case is part of the cross-case component of an analysis in which “the emphasis of a study shifts from the individual case to a sample of cases” (Gerring 2007, 20). These are “most often ... surveyed through a quick reading of the secondary literature or through a statistical analysis” (ibid, 22).

I suggest that the shadow case study should be seen in terms of three key characteristics. First, it is a type of case study, rather than an illustration based on a particular case—this means that it relies on “intensive” (Gerring 2007, 20) analysis of a case. Second, it is ancillary to a central case study (oriented to external rather than internal validity), and third, it is constrained in its investigation and presentation by its subsidiary place in an overall project. I therefore define the *shadow case study* as a component of small-N research that entails the examination of an ancillary or peripheral case, drawing inference from the within-case analysis of that case to shed light on the generality of claims most centrally evaluated in the core case.

Case studies and case illustrations: The threshold for what qualifies as a case study rather than a case illustration is not a clear-cut one—it is easier to articulate the standards for a maximally good case study than to iden-

5 One could in principle conduct an analogous exercise using QCA, but the goal would be identifying the set of causal configurations that produce an outcome of interest rather than, as in the approach to case illustrations in much existing scholarship, investigating the validity and generality of a single explanation. Thus, the large-N QCA exercise would produce a different kind of more general claim than the type scholars currently tend to assert via their use of case illustrations.

tify the minimal threshold for qualifying as one.⁶ I have suggested that the presentation of a case becomes more of a “study” and less of an “illustration” as it draws more on tools of within-case analysis to generate causal inference and less on correlations between key variables. This line of argument echoes the critique of King, Keohane, and Verba (KKV) articulated in the distinction between causal-process observations and dataset observations in Collier and Brady (2004). Yet as I discuss further below, one need not restrict case studies to investigations of causal process; inference can be generated from other kinds of evidence collected via intensive study of cases, and indeed this is consistent with the original definition of causal-process observations articulated by Collier and Brady. The distinction between case studies and case illustrations, then, derives from the intensity with which a case is investigated and the extent to which the researcher draws on evidence (whether of causal process or of some other aspect) that goes beyond the scoring of the case on key variables.

Ancillary and core cases: The shadow case study is distinct from the *core case study* or central analytical exercise in small-N research. This is the case study that lies at the heart of scholarship in which the shadow case study may play a complementary role. Oriented principally to internal validity, the core case study draws on tools of within-case causal analysis to leverage the strengths of case selection logics and generate causal inference. The *shadow case study* is an ancillary case, oriented to external validity. The goal of examining the shadow case is principally to show that the theory developed elsewhere holds more generally. But unlike the case illustration (and like the core case), the shadow case study seeks to generate some analytical leverage from within-case analysis.

Bringing these first two characteristics together, we can envision a two-by-two table, with the horizontal dimension scoring emphasis on internal versus external validity, and the vertical dimension scoring the extent of within-case analysis. The core case study would be in the upper-right quadrant (internal validity and within-case analysis); the case illustration in the lower-left (external validity and no within-case analysis). The lower-right cell, which refers to exercises oriented to internal validity without within-case analysis, has been shown to be logically impossible by the recent generation of qualitative

methods scholarship. The remaining, top-left cell, which uses within-case analysis to generate external validity, refers to the shadow case study. The challenge is how to do so under a set of constraints that place a ceiling on the extent of within-case analysis possible in an ancillary case and preclude it from becoming a core case study.

Constraints: Three central forms of constraints can be identified; I term these *logistical*, *presentational*, and *research cycle* constraints. Logistical constraints are principally a function of researcher time and resources. Because the examinations of ancillary cases are relatively abbreviated and are not the core focus of scholarly endeavor, the amount of effort that can reasonably be allocated to their analysis is fairly limited. Limits on researcher time, effort, and resources may limit the data available to a researcher conducting an ancillary case study. This can impact case selection, as scholars prioritize data availability over other considerations (Posner 2005, 260). It can also limit scholars to only exploring parts of their arguments in ancillary cases (Daly 2016, 220). Overall, the impact is to place an upper bound on the strength of the inferences that can be drawn from these cases.

A second type of constraint is presentational. References to space as a limitation in the discussion of ancillary cases are common. Scholars refer to them as “illustrations” (Mylonas 2012, 171), “glimpses” (Daly 2016, 230), or “brief examinations” (Pearlman 2011, 215) that “cannot... be thought of as a full-fledged engagement” (Kurtz 2013, 237 n. 15). The fact that the presentation of ancillary cases can be no more than “brief” or “schematic” places limits on their inferential leverage. In other words, even if scholars conduct a full-fledged within-case analysis of these ancillary cases, they cannot *present* that analysis. Thus, like logistical constraints, presentational constraints place an upper bound on the extent of within-case inference that these ancillary cases can provide.⁷

By research cycle constraints, I refer to the fact that these ancillary cases are investigated after the analysis of the core cases, limiting in practice the set of ways they can be of use for researchers. Their place in the research cycle constrains the analytical tasks of ancillary cases to confirmatory exercises: They cannot serve as sites for other analytical tasks like theory building. Were a scholar to find something in examining a case intended for a shadow case study that led to a fundamental rethinking

6 A similar problem exists in empirical coding of ancillary cases as well; at what point does a scholar’s investigation of an ancillary case advance from an “illustration” to a “study”? There’s a clear-cut difference between the paragraph-length illustrations and the chapter-length studies seen for example in Cammett (2014), but where should the boundary be defined for less clear-cut instances? Parallel with the discussion that follows in this paragraph, I suggest that the intensity with which within-case evidence is used should be the deciding factor.

7 Even if word limits are not a concern, the researcher must still limit the space devoted to ancillary cases in the interest of streamlining presentation and making the overall project more digestible or readable. See Fairfield and Charman 2017 for a similar point.

of the core case, that case would likely take on a different role in the overall study rather than remaining in an ancillary role.⁸

Together, these constraints make the use of full within-case analysis in shadow case studies a logical impossibility and make the shadow case study a diminished subtype of case study rather than a full member of that conceptual category. Yet shadow case studies vary in their intensity. At one end of a continuum, we have studies like Cammett (2014), which supplements an analysis of the logic of social provision by sectarian political actors in Lebanon with a chapter-length investigation of the behavior of similar actors in Iraq. To the extent that a shadow case study like this generates some internal validity, it comes close to a core case study. But the modal shadow case study, bound by the constraints discussed above, would fall closer to the boundary separating this category from the case illustration. The purpose of this paper is to explore how a scholar operating within the constraints that limit the number of core cases one can conduct might leverage shadow case studies, and the within-case inference that they allow, to best increase the validity of the inferential claims developed and advanced in the core cases.

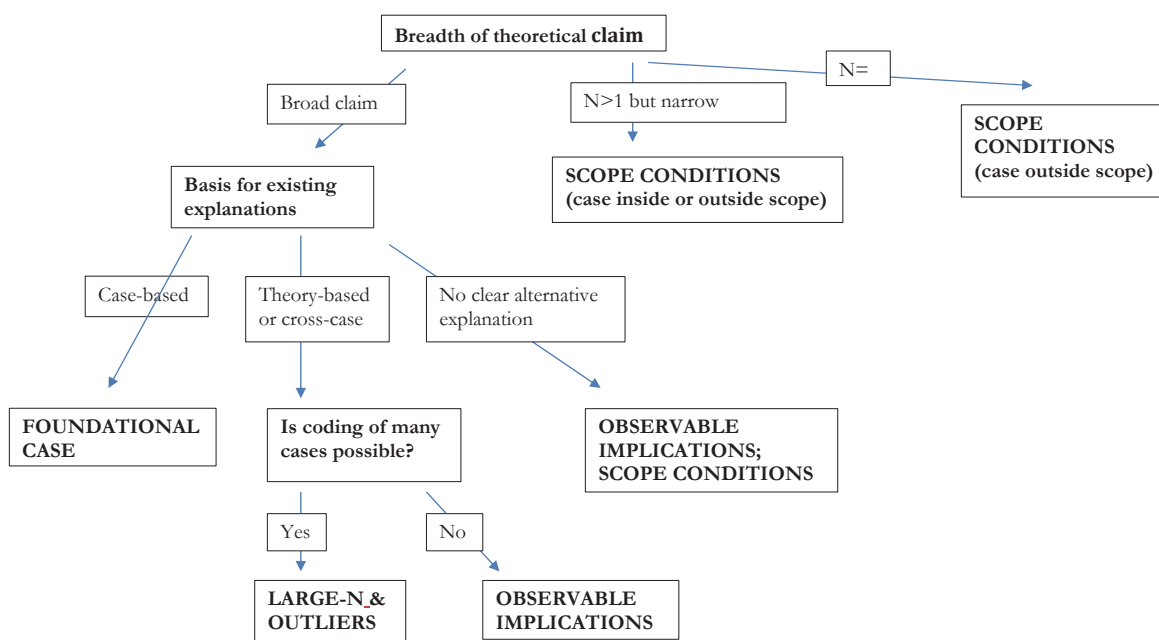
Uses for the Shadow Case Study

As discussed above, existing practice has often deployed ancillary cases to investigate scope by exploring cases within or outside scope conditions. Where the

claims advanced are broad, I suggested that a larger-N alternative might be preferable to the investigation of a relatively small number of cases as a means for assessing generality. But there may be a valuable use of ancillary cases where the breadth of analytical claims is fairly narrow, and a large-N alternative is therefore not feasible.

Yet scholars might also consider using ancillary cases for a set of other possible analytical tasks, described below. This section shows that shadow case studies can be used as a site for a researcher to *address alternative explanations*, to *learn from outliers*, and to *explore additional observable implications*. These inferential tasks are standard fare for case-oriented scholarship, but we lack a systematic discussion of how they can be carried out within the confines of an ancillary case study. I show how each improves the inferences about theory generated and tested in the core cases, while leveraging the distinct advantages of the within-case analytical toolkit within the constraints of the ancillary case setting. I also provide, as outlined in Figure 1, some guidelines for how scholars might choose from among these inferential tasks as well as the exploration of scope conditions (seen on the right side of the figure), as they seek to generate some external validity for their studies. Before proceeding, I note that there may be other valuable uses for shadow case studies beyond those listed here, which were selected because of the obvious ways in which they provide inferential leverage for the broader study.

Figure 1: Choosing a Use for a Shadow Case Study



⁸ I thank Sean Yom for suggesting the research cycle constraint and Erica Simmons for helpful comments relating to this issue, though I expect she will continue to disagree with my views on it.

Assessing alternative explanations: If we think about the state of the literature on a given topic in terms of alternative explanations, we can imagine three situations, as indicated in the bottom half of Figure One. In some areas of research, no clear alternative explanation exists—this is often the case in fairly new currents of scholarship, or where arguments are especially contested. In others, there is a clear conventional wisdom or consensus around the analytical power of an existing explanation. Yet sometimes this consensus is based on theory or broad cross-case patterns—for example, the idea that the level of inequality affects regime outcomes. Finally, in other areas of research, scholarly confidence in the validity of an existing explanation is based on insight from a particularly illuminating case. I call this type of case—a case on which general claims for a particular explanation of an outcome are most centrally grounded—a *foundational case*.⁹ Argentina, for example, is a *foundational case* for the link between peripheral industrialization and democratic breakdown (O'Donnell 1973) in that the confidence scholars have in this explanation is based not on broad cross-case patterns or on theory alone, but on the within-case understanding of the Argentine case advanced by O'Donnell.¹⁰

A first possible use for shadow case studies is to reconsider the validity of these case-based alternative explanations. As Bayesian scholars have argued in particular (Fairfield and Charman 2019), the plausibility of a new argument advanced depends not only on its fit with cases but on its ability to out-perform or subsume alternative explanations. Under some conditions, a shadow case study of a *foundational case* for an alternative explanation can be an especially powerful analytical exercise.

A common existing practice in case study research is to deal with alternative explanations by controlling for them or by dismissing them based on a brief exploration of patterns in the population. Yet the vast majority of small-N scholars have an epistemological commitment to the importance of evidence about causal mechanisms as a means of evaluating hypothesized explanations for outcomes, and a central tenet of this approach is that evidence about causal process is the strongest and most valid way to evaluate proposed causal relationships. It is striking, then, that when dealing with alternative explanations, scholars tend to draw on types of evidence that

they hold to be weaker and use the evidence from cross-case analysis that is often described by advocates of process tracing as providing nothing more than evidence of correlation.

To provide convincing evidence that casts doubt on alternative explanations, we should instead evaluate them using the same within-case analysis that lies at the heart of small-N research, by examining the causal process theorized to link proposed cause to outcome and showing that it does not operate. This is not a novel suggestion: Many scholars, particularly those who take a formally Bayesian approach to process tracing, argue that one must do so in the core cases themselves (Fairfield and Charman forthcoming).¹¹ Moreover, as suggested above, case-based engagement with an alternative explanation might be especially valuable when our confidence in that alternative is itself based on a case study rather than on other kinds of evidence. It might be especially valuable to select as our core case the foundational case for an alternative explanation, and to show that our explanation out-performs the alternative on its strongest terrain.

Yet, as Gerring (2001) insists, scholars must weigh many criteria in making research design choices. This means that scholars will often choose for their core cases a case other than the foundational case for an alternative explanation. An alternative approach is to select the foundational case for a shadow case study. A scholar presenting a new explanation for a broadly occurring outcome for which a prominent alternative explanation exists can use this approach to show that her preferred theory either subsumes that explanation into a more general account or out-performs it in matching the pattern of observations in the case.

This effort relies heavily on the tools of within-case analysis. Analysis of the foundational case requires a deep engagement with process-based evidence for both the author's preferred explanation and the alternative for which a foundational case has been chosen. One example of the payoffs of this approach can be seen in Kurtz (2013), who turns to Prussia as a shadow case study for his argument that rural social structure was the key determinant of state-building outcomes, as elaborated in his core cases drawn from Latin America. He describes a robust current of scholarship that emphasizes external security threats as the determinant of state strength

9 Earlier versions of this paper used the term “paradigmatic case”; I have moved away from that term to highlight that the case is chosen based on the *foundation* on which rests an existing explanation.

10 Note that a foundational case is not defined in terms of an *outcome*, but in terms of an *explanation*. It is therefore distinct from existing definitions of a crucial case, and from the definition of an influential case in Seawright and Gerring (2008).

11 Yet some methodological guidance diverges on this point: Goertz (2017, 78ff.) suggests that cases where both the alternative explanation and the new one being advanced might operate are characterized by over-determination and should not be chosen for process tracing.

in Prussia and describes it as the foundational case for the claim that war made the state. By showing that “external threat did not drive modernization” until a shift in agrarian social structure toward capitalist agriculture after 1850, Kurtz challenges the “bellic” school of state formation on the very grounds where arguments of that nature are given the most credence (Kurtz 2013, 247).

As this discussion suggests and Figure 1 shows, the foundational case might be an especially apposite use for a shadow case study when a scholar is advancing a broadly general claim, when the state of the literature is such that a new explanation for an existing outcome confronts a particularly prominent and widely held alternative, when confidence in that most prominent alternative derives from the dynamics of an existing case rather than from cross-case evidence or from theory, and when other research design criteria preclude the selection of this existing case as the core case for analysis. Under those circumstances, using a shadow case study to confront the alternative in its foundational case may be the best way to assess the relative explanatory power of a newly proposed explanation and a prominent alternative. In Bayesian terms, this is selecting a case in which our priors strongly favor an alternative explanation. If we can conclude after investigating that our case is favored even in such a case, our confidence in its overall explanatory power relative to that alternative is enhanced. Following the terminology of Fairfield and Charman (forthcoming, Chapter 11), then, this kind of shadow case study may be a setting for high information gain.

Learning from outliers: Rather than using shadow case studies to engage an alternative explanation, scholars might instead use them to explore the shortcomings of one’s own proposed explanation. It may be fruitful to investigate cases mis-predicted by the theory, to engage in within-case analysis of these outliers to find what accounts for the poor fit of the proposed explanation, and thus seek to increase confidence in the theory generated in the core case studies.

One possibility is that cases identified as outliers are a product of measurement error. Scholars can check their scoring of key variables in these cases based on a scholarly consensus by case experts. Doing so not only has the effect of showing that the theory fits those cases, it also increases our confidence in the explanation developed in the core cases, because a greater share of cases is shown to be consistent with theory. Another possibility,

which entails more extensive within-case investigation, is that outlier cases result from causal heterogeneity. Investigating outliers to identify why the proposed explanation does not hold can therefore result in theoretical refinement that improves the validity of causal claims.

In his study of variation in how insurgents treat civilians, Weinstein (2007, 311-326) provides examples of both of these kinds of learning from outliers. He examines four outlying cases in which the number of conflict deaths during civil war is higher than his theory would predict given the relatively low access to resources for rebels. Two (Algeria 1962-64 and Lebanon 1975-91) are found to be a product of measurement error. The other two (Colombia’s *La Violencia* 1949-63 and Algeria 1992-2000) are shown to be outliers because of causal heterogeneity, and their investigation prompts theoretical refinement to incorporate the role of state weakness and to add a scope condition on shifts in government strategy.

As the preceding discussion suggests, the exploration of outliers requires three research tasks: (1) scoring a sizable number of cases or a sizable share of the universe of cases to which the theory is said to apply on reasonable proxies for the independent and dependent variable so that outliers can be identified; (2) assessing whether the cases identified as outliers are an artifact of coding error through an examination of case-specific sources to refine scorings on these variables; and (3) drawing on case-specific secondary sources to identify the causal process producing the outcome in the cases remaining as outliers once task (2) has been completed.¹²

These tasks are feasible within the constraints of the shadow case study and potentially valuable, not because of what we learn about these cases in particular, but in terms of the broader research enterprise. As Figure 1 suggests, this might be an especially valuable use of ancillary cases where the theoretical claim being advanced in the core case is intended to have broad applicability, where ancillary cases do not serve to evaluate alternative explanations, and where the coding of many cases on key variables is feasible.

Two final points can be made about this use for the shadow case study. First, of the uses I describe, this one falls closest to the “floor” between the shadow case study and the case illustration discussed above, since it requires less intensive within-case analysis than the other ways to deploy ancillary cases. Second, and more importantly,

12 If a large number of outliers remain as task (3) is confronted, the researcher faces the problem of how to choose among them. Given the analytical payoffs from theoretical refinement that they may generate, and the feasibility of doing so in a fairly efficient manner, scholars should start with the most high-leverage outliers but explore as many as constraints allow. Thus, again, my recommendation for case selection builds on the Bayesian idea of information gain articulated in Fairfield and Charman (forthcoming, Chapter 11)

since identifying these outliers requires that the researcher code a relatively large number of cases on the independent and dependent variable, this use of the shadow case study is complementary with the mixed-methods approach to assessing generality discussed above. I suggest that these two can fruitfully be combined.

Exploring observable implications: A final use for a shadow case study entails the investigation of a theory's observable implications. The exploration of observable implications outside a core case follows logically from either one of two theoretical starting points about the nature of evidence and inference. First, we tend to hold that confidence in the explanatory power of a theory is improved to the extent that we can confirm with empirical evidence a range of observable implications it generates beyond predictions about the central outcome. This point is, I submit, too often rejected in some recent scholarship on process tracing (e.g., Beach and Pedersen 2019) that limits causal inference to the examination of causal process, reifying "mechanism" and "process" at the expense of ignoring other opportunities to use information about cases to arbitrate among alternative explanations. To be provocative, I suggest that there has been too literal an interpretation of the term "causal process observation" articulated by Collier and Brady to distinguish within-case analysis from dataset observations, and that the result has been an overly narrow approach to causal analysis in small-N research. Like Fairfield and Charman (forthcoming, Chapter 3), I argue that we should take a broad view of the kinds of evidence that generate inference about the causal claims we seek to assess, and I suggest that we can use shadow case studies to explore a range of evidence that goes beyond a narrow definition of causal process.

Second, in most traditions of the philosophy of causation (see for example Falleti and Lynch 2009), those observable manifestations of a causal process in a given case are generally taken as evidence in favor of the more general operation of that process.¹³ If we hold either or both of these two points to be true, we can easily imagine situations where a theory generates observable implications that cannot be assessed in the core cases chosen by a researcher. I present two such scenarios as examples.

First, data needed to evaluate observable implications may not exist in the core cases. This might occur, for example, if the core cases are historical but the theory has

implications that can be assessed in contemporary contexts with survey or experimental data. Singh fruitfully uses a brief discussion of ancillary case evidence in her conclusion as a way of "traveling with the core insight of this book all the way down to the individual," drawing largely on survey experimental evidence not available in her core cases (Singh 2015, 255-6).

Second, some observable implications do not operate in the core cases. For example, a theory tested in democratic contexts might have observable implications in non-democratic settings. Posner takes his theory beyond the core case of regime transition and the "relative political salience of tribal and language group identities in Zambia" and explores how "the general logic it articulates extends well beyond these particular independent and dependent variables," using examples of institutional changes different from those seen in his core case of Zambia (Posner 2005, 274).

While the analysis entailed is not a holistic case study, it is also analytically distinct from the pattern-matching exercise of the case illustration in that it focuses on observable implications that go beyond scores on independent and dependent variables. Because this within-case analysis exists principally as confirmatory evidence for the broader theoretical project, rather than as investigation of the particular case in its own right, these analytical exercises are instances of within-case analysis in the service of the broader findings and are properly placed within the shadow case study category. Scholars should be more attuned to the possibility that these analytical opportunities exist, and that shadow case studies might fruitfully be used to seize upon them. This might be an especially valuable use for ancillary cases where opportunities to weigh one's own preferred explanation against evidence outside core cases provides inferential gains.

Conclusion

Scholars whose research focuses on one or a small number of cases often use ancillary cases in an effort to strengthen their claims. Yet we lack any systematic investigation of the research design and methodological choices entailed in these or any guidelines for how to interpret the inferences they generate. This paper has sought to begin to remedy that gap and develop a framework for considering how one might select and deploy ancillary cases to complement small-N research.

13 This suggests that one might use ancillary cases to explore the generality of a causal mechanism beyond the core case. To the extent that this exercise relied on within-case analysis rather than a review of the treatments of the case by other scholars with the inferential problems that entails, and that the scholar articulated a case selection strategy oriented toward external validity, this could be another use for the shadow case study. Note that some scholarship on process tracing and causal mechanisms (e.g., Beach and Pedersen 2019) rejects the possibility of generalizing across cases about causal mechanisms.

Existing practice has tended towards what I call the case illustration, which uses ancillary cases to make claims of generality or to explore the validity of theorized scope conditions. Scholars most commonly code ancillary cases on core variables of interest and engage in a sort of pattern-matching or correlational exercise. This approach may be of value to the extent that case selection is conducted with an explicit logic in ways that maximize variation within the scope being evaluated and to the extent that the researcher is comfortable with correlational evidence as a way to assess theory. Yet one might argue that both of these criteria may be better satisfied with a large-N and universal exercise in assessing generality, especially when the researcher seeks to advance broadly generalizing claims.

I therefore suggest that scholars who seek to generate inference from internal validity might instead turn to *shadow case studies*, or more intensive investigations of ancillary cases with specific analytical tasks in mind that generate increased inferential leverage. These can serve a variety of inferential purposes, including the exploration of outliers, the search for observable implications, and the confrontation of alternative explanations in the foundational case for them. There are, of course, other uses for the shadow case study. The purpose of this paper, then, is not to provide an exhaustive menu of options. Instead, it is intended to de-center the common practice of the case illustration and to call for more thoughtful research design choices about ancillary cases, which can serve to generate inference to a much greater extent than has heretofore been realized.

References

- Barnes, Tiffany. 2016. *Gendering Legislative Behavior: Institutional Constraints and Collaboration*. New York: Cambridge University Press.
- Beach, Derek, and Rasmus Brun Pedersen. 2019. *Process Tracing Methods*. Ann Arbor: University of Michigan Press.
- Boas, Taylor. 2016. *Presidential Campaigns in Latin America: Electoral Strategies and Success Contagion*. New York: Cambridge University Press.
- Cammett, Melani. 2014. *Compassionate Communalism: Welfare and Sectarianism in Lebanon*. Ithaca: Cornell University Press.
- Collier, David, and Henry Brady. 2004. *Rethinking Social Inquiry*. Lanham: Rowman and Littlefield.
- Collier, David, and Steven Levitsky. 1997. "Democracy with Adjectives: Conceptual Innovation in Comparative Research." *World Politics* 49, no. 3 (April): 430-451.
- Daly, Sarah Zukerman. 2016. *Organized Violence after Civil War: The Geography of Recruitment in Latin America*. New York: Cambridge University Press.
- Fairfield, Tasha, and Andrew Charman. 2017. "Explicit Bayesian Analysis for Process Tracing." *Political Analysis* 25, no. 3 (July): 363-380.
- . 2019. "A Dialogue with the Data: The Bayesian Foundations of Iterative Research in Qualitative Social Science." *Perspectives on Politics* 17, no. 1 (March): 154-167.
- . Forthcoming. *Social Inquiry and Bayesian Inference: Rethinking Qualitative Research*. New York: Cambridge University Press.
- Falletti, Tulia, and Julia Lynch. 2009. "Context and Causal Mechanisms in Political Analysis." *Comparative Political Studies*. 42, no. 9 (September): 1143-1166.
- Gerring, John. 2001. *Social Science Methodology: A Criterial Framework*. New York: Cambridge University Press.
- . 2007. *Case Study Research: Principles and Practices*. New York: Cambridge University Press.
- Goertz, Gary. 2017. *Multimethod Research, Causal Mechanisms, and Case Studies: An Integrated Approach*. Princeton: Princeton University Press.
- Grzymala-Busse, Anna. 2015. *Nations Under God: How Churches Use Moral Authority to Influence Policy*. Princeton: Princeton University Press.
- Handlin, Samuel. 2017. *State Crisis in Fragile Democracies: Polarization and Political Regimes in South America*. New York: Cambridge University Press.
- Kurtz, Marcus. 2013. *Latin American State Building in Comparative Perspective: Social Foundations of Institutional Order*. New York: Cambridge University Press.
- Lange, Matthew. 2013. *Comparative-Historical Methods*. Thousand Oaks: Sage.

- Levitsky, Steven. 2003. *Transforming Labor-Based Parties in Latin America: Argentine Peronism in Comparative Perspective*. New York: Cambridge University Press.
- Mylonas, Harris. 2012. *The Politics of Nation-Building: Making Co-Nationals, Refugees, and Minorities*. New York: Cambridge University Press.
- O'Donnell, Guillermo. 1973. *Modernization and Bureaucratic Authoritarianism: Studies in South American Politics*. Berkeley, CA: University of California Press.
- O'Neill, Kathleen. 2005. *Decentralizing the State: Elections, Parties, and Local Power in the Andes*. New York: Cambridge University Press.
- Pearlman, Wendy. 2011. *Violence, Non-violence, and the Palestinian National Movement*. New York: Cambridge University Press.
- Posner, Daniel. 2005. *Institutions and Ethnic Politics in Africa*. New York: Cambridge University Press.
- Seawright, Jason, and John Gerring. 2008. "Case selection techniques in case study research: A menu of qualitative and quantitative options." *Political Research Quarterly* 6, no.2: 294-308.
- Singh, Prerna. 2015. *How Solidarity Works for Welfare: Subnationalism and Social Development in India*. New York: Cambridge University Press.
- Truex, Rory. 2016. *Making Autocracy Work: Representation and Responsiveness in Modern China*. New York: Cambridge University Press.
- Weinstein, Jeremy M. 2007. *Inside Rebellion: The Politics of Insurgent Violence*. New York: Cambridge University Press.
- Wilkinson, Steven. 2004. *Votes and Violence: Electoral Competition and Ethnic Riots in India*. New York: Cambridge University Press.
- Ziblatt, Daniel. 2006. *Structuring the State: The Formation of Italy and Germany and the Puzzle of Federalism*. Princeton: Princeton University Press/
- . 2017. *Conservative Political Parties and the Birth of Modern Democracy in Europe*. New York: Cambridge University Press.