

Process Tracing and the Problem of Missing Data

Ezequiel Gonzalez-Ocantos & Jody LaPorte

University of Oxford

Forthcoming, *Sociological Methods and Research*

Abstract: Scholars who conduct process tracing often face the problem of missing data. The inability to document key steps in their causal chains makes it difficult to validate theoretical models. In this paper we conceptualize “missingness” as it relates to process tracing, describe different scenarios in which it is pervasive, and present three ways of addressing the problem. First, researchers should contextualize the data generation process. This requires characterizing the process whereby the actors that populate models decide whether to leave traces of their actions and motives. Researchers can thus assess whether or not incentives to produce missingness are compatible with the microfoundations of the theory, and consequently, whether or not missingness is disconfirmatory. Second, researchers may invest in indirect tests of causal mechanisms. Generating out-of-context data about microfoundations offers a plausible window into inaccessible mechanisms. Third, specifying the analytical status of steps in the causal chain allows scholars to make up for deficiencies in evidentiary support.

Introduction

Quantitative scholars who work with time-series cross-sectional datasets or surveys often confront the problem of missing data, which arises when “respondents answer some questions but not others (or, in general, scattered cells in a data matrix are missing)” (King et al. 2001: 49). Ignoring the problem with list-wise deletion, or using more or less educated “guessing” techniques, can destroy valuable information or create widely unreliable data, thereby biasing results and increasing inefficiency. Statisticians have therefore developed techniques, such as multiple imputation, in order to deal with “missingness” (King et al. 2001; Honaker and King 2010).

Qualitative scholars who engage in process tracing also have to deal with “missingness.” Process tracing is grounded in the assumption that good social explanations are those that document the existence of plausible causal pathways connecting the independent and dependent variables. Although these causal pathways often include unobservable mechanisms “that make the system tick” (Bunge 2004: 182; Little 2015, George and Bennett 2005, 137, but see Hedstrom 2008), they tend to leave behind residual

evidence of their presence. Process tracing therefore begins with a hypothesized causal chain consisting of various steps leading to the outcome of interest; scholars then proceed to reconstruct it by documenting a sequence of events compatible with those theoretical steps. In addition, they search for evidence of the mechanisms that explain why the actors involved behave in ways that push the story forward along these steps, triggering one event after the other. The goal is to “examine the fingerprints” that the process “should have left in the empirical record” (Beach and Pedersen 2014, 42).

Yet, depending on the type of field site or the theoretical properties of the mechanisms, the constituent parts of these causal chains can be quite opaque. For example, the difficulties of collecting accurate data in conflict zones are well documented (Lyll 2014; Kapiszewski et al 2015). Non-democratic regimes pose similar challenges. Crucial archives may be closed to researchers (Edgars 2006), or evidence simply may not exist to document the more unsavory aspects of authoritarian rule (Way and Levitsky 2006). Even in more open settings, individuals who engage in illegal or socially undesirable activities are unlikely to leave behind comprehensive evidence. This poses a problem for researchers seeking to explicate, for example, the causes of corruption, the effects of electoral manipulation, or the internal dynamics of criminal gangs. On less provocative topics, too, key actors may neglect to record their actions and motivations.

The inability to document any part of a hypothesized causal process is problematic for process tracers for at least two reasons. First, unlike observations in a data set, which are collected following standardized measurement procedures for each variable, qualitative evidence is much more eclectic. Documenting different aspects of the causal story often requires marshaling evidence from radically different sources. For instance, whereas national economic statistics may authenticate a key step in a causal chain, it might also be necessary to conduct interviews with policymakers to account for the mechanisms that led to this

particular state of affairs. Moreover, the observations used to document the same causal process in different cases can vary. For example, social desirability norms can be context specific, rendering similar interview quotes from similarly situated actors more credible in one country than in another. For these reasons, when qualitative researchers confront gaps in the evidentiary record, they cannot rely on the average score of a variable, as quantitative scholars do when they seek to come up with plausible values for an empty cell.

Second, in quantitative research, missing data can increase inefficiency when estimating average effects, and in some cases, introduce bias. But in process tracing, the inability to reconstruct causal chains carries much more serious consequences for causal inference. Evidentiary gaps undermine the most basic goal of the method, which is to show there is indeed an uninterrupted pathway consistent with theory, connecting cause and effect. As George and Bennett (2005, 207) put it, “*all* the intervening steps in a case must be as predicted by a hypothesis [...] or else that hypothesis must be amended [...] It is not sufficient that a hypothesis be consistent with a statistically significant number of intervening steps” (emphasis in the original). Similarly, Beach and Pederson (2014, 39) maintain that we should think of our models as a system: “either all parts of the mechanism are present, or the mechanism itself is not present.” In this sense, process tracing is successful when the evidence allows us to “infer that all of the parts of a hypothesized causal mechanism were actually present in a case” (ibid, 69).

In light of these difficulties, and in the absence of practical tools designed to deal with missingness, researchers risk falsely rejecting a valid hypothesis due to the evidentiary pressures generated by data-poor environments. Scholars may choose to avoid conducting research on these topics or cases altogether, in an attempt to prevent “disconfirmatory” bias. This is obviously a highly unsatisfactory solution because it would leave important political phenomena and cases under-researched. In the natural world, for example, fossilization is

quite unusual, which means that there is fragmentary evidentiary basis for inferences about human evolution. Yet it would be unreasonable to conclude, as a result, that scientists should refrain from studying evolutionary processes.

This paper adopts a different position. Building on existing treatments of the problem of missingness in qualitative research, we contend that deep contextual knowledge, as well as careful theorizing of causal chains, can help researchers calibrate credible empirical narratives amidst missingness. Specifically, we go beyond the important advice to be cognizant and transparent about the “sources of uncertainty” (Waldner 2016, 32; see also Bennet and Checkel 2014, 31).

We contribute to this literature by outlining three practical remedial instruments that can be applied to minimize the inferential damage caused by missingness. First, we argue that when researchers confront missingness they can *contextualize the data generation process*. The actors in our models, in addition to pushing the causal story forward, are also responsible for generating the data we need. And they may have incentives to either leave traces of actions and motives, or to omit evidence in order to cover their tracks. Scholars can assess the microfoundations of the model, which specify the reasons why these actors behave in certain ways along the causal chain, to evaluate whether these microfoundations also provide a coherent justification for missing data. Microfoundations may or may not be compatible with cover-ups, and therefore indicate whether or not missingness is likely to be disconfirmatory. Second, researchers can design *indirect tests* of the observable implications of the hypothesized micro-behavior or psychological processes for which no direct data is available. We contend that by conceptualizing mechanisms at a higher level of abstraction, it is possible to generate out-of-context but plausible data about distant or inaccessible micro-foundations. This is particularly useful for those studying historical or macro-social processes. Third, researchers can *specify the analytical status of steps in the causal chain*, and in doing so draw

on the logical status of necessary and sufficiency bonds to overcome missingness. When qualitative data is missing for a specific step in the causal chain, researchers can explore the hypothesized causal role of this particular step, and see whether missingness matters to affirm the argument. We contend that an argument may still stand if the absence of evidence applies to any step that is logically implied by a preceding step through sufficiency claims.

To be clear, missingness is not a function of the quantity of available data. As Bennett (2008) suggests, one piece of “doubly-decisive” evidence can outweigh lots of less compelling pieces of evidence. Thus, in contrast to some of the existing work on this topic (Lyll 2014), our solutions do not singularly focus on generating *more* data. In principle, researchers could simply work harder to find data that seemingly does not exist or is momentarily inaccessible. They may uncover hidden truths by digging deeper into the archives, by fiercely interrogating political actors for information they are hesitant to provide, or by inventing unusual means of observing backroom deals. But missingness arises precisely when those efforts are unsuccessful. Rather than focusing on how we may become better fieldworkers to correct the problem of missing data (cf. Kapiszewski et al 2015), we propose a series of procedures to reinterpret existing and available data in ways that lend plausibility to the claim that hidden steps in the hypothesized causal chain are present, and/or that the hypothesized mechanisms are actually operative.

The paper proceeds as follows. First, we present an overview of the method of process tracing, and discuss how the problem of missing data relates to current standards of best practice in the discipline. Second, we describe different scenarios in which missingness might be pervasive. Third, we assess the implications of missingness for developing and testing theories using process tracing, and discuss three remedial instruments for drawing causal inferences. We illustrate these points using examples from Political Science. Finally, the paper concludes by highlighting how these procedures advance current understandings of

process tracing and the use of qualitative methods. In particular, we think of our efforts as part of a second generation of essays on process tracing. Whereas the first generation opened new paths by clarifying the assumptions behind the method and set out general guidelines or best practices, a second generation of studies is beginning to think more carefully about practical implications and problem solving. We discuss how this paper advances that agenda, with implications for Bayesian updating, mixed-method research, and transparency in qualitative research.

I. Process Tracing and the Causes of Missing Data

Process tracing is predicated on a mechanistic understanding of causal relations, in which the validity of any causal claim rests on the analyst's ability to establish the means by which X causes Y (Bunge 1997; Waldner 2007; Little 1991, 2015; Mahoney 2016). Over the past decade, qualitative methodologists have engaged in substantial debates about the precise nature of causal mechanisms (King, Keohane and Verba 1994; Mahoney 2001; Gerring 2010; Little 2015). Without denying these different points of view, we follow recent literature in understanding mechanistic explanations as those in which “the focus is not on relationships between variables, but on actors, their relationships, and the intended and unintended outcomes of their actions” (Hedstrom 2008, 320). In line with scientific realism, mechanisms are “physical, social or psychological processes through which agents with causal capacities operate [...] to transfer energy, information, or matter to other entities” (George and Bennett 2005, 137, see also Hedström and Swedberg 1998; Beach and Pedersen 2014; Little 2015).¹

¹ We adopt an actor-centric approach to process tracing here as a matter of pragmatism in order to shed light on the sources and types of missingness in qualitative data. As should be clear from the following discussion, we see this approach as compatible with other process tracing methods, including process tracing tests (Collier 2011), Bayesian process tracing (Bennett and Checkel 2014, Humphreys and Jacobs 2015; Fairfield and Charman 2017), and set theoretic methods (Beach and Rohlfing 2015). For recent advances regarding the compatibility of different process tracing methods, see Trampusch and Palier (2016), Zaks

The fundamental goal of process tracing is to document the presence of these causal mechanisms, and the actions, reactions and events they bring about on the way to the outcome of interest. Doing so requires, first, the specification of the causal chain under scrutiny. This involves outlining the set of variables along the causal pathway in order “to represent, as fully as possible, the set of causal relationships that constitute the process being traced” (Waldner 2015, 131). Second, researchers start to operationalize the model by specifying the relevant events, as well as actions and reactions by the protagonists of the causal story that, when observed in a specific sequence, generate the outcome of interest (*inter alia* Mahoney 2016; Waldner 2014; Beach 2016; Falletti 2016). The researcher also will identify the mechanisms linking these events. This routinely involves discussing the logic of action that governs behavior in each of the theoretically relevant steps, and the means by which those actions push the case closer to the outcome. More specifically, scholars theorize actors’ incentives or motives for behaving in certain ways, and the cognitive processes by which they change their minds about important issues, make sense of new information about their environment, or react to the actions of others (Falletti and Lynch 2009).² Finally, each step of the model in question will be matched with its respective observable implications. The researcher uses her case studies to conduct process tracing tests, searching for pieces of evidence—“causal process observations” (Brady, Collier, and Seawright 2004) or “diagnostic evidence” (Bennett and Checkel 2014)—that would support or contradict claims about the presence of the different steps and the mechanisms connecting them. To make valid causal claims, every step in the causal chain should be supported by empirical evidence (George and Bennett 2005: 207; see also Beach and Pedersen 2014, Beach 2016, Collier 2011).

(2017), and the symposium “Causal Mechanisms and Process Tracing” in the Spring/Fall 2016 issue of *Qualitative & Multi-Method Research*.

² Process tracing is compatible with different social theories of what motivates action. The actor-centric view adopted here should not be read as implying that process tracing is only compatible with methodological individualism, or rational choice models.

However, these evidentiary standards may be difficult to satisfy under many circumstances—and the microfoundational basis of causal mechanisms provides some clues as to why. Specifically, the availability of evidence to conduct process tracing is a function of the resources and willingness of the actors that populate our models to disclose information or leave traces of their activities and motives. Ample evidence of high probative value is likely to be generated when it is costless (or even beneficial) for the actors involved to do so. Politicians working on historically significant policies—e.g., the creation new institutions, or international treaty negotiations—are likely to document their actions, decisions and motives in memoirs, interviews, and other written materials. Similarly, major events are usually covered in the media, making it easier for researchers to recreate causal chains. Additionally, certain political environments are favorable to leaving behind visible traces of decision sequences. This is especially true in settings where documentation needs to be preserved for legal or political reasons. Political actors also are more likely to fully comply with these requirements when they seek to take credit for outcomes. In such contexts, it is the researcher’s responsibility to seek out and document those pieces of evidence that the causal process left behind.

In other contexts and topics, by contrast, the causal processes of interest are unlikely to leave behind significant traces. Data on certain topics simply may not exist because it was not recorded at the time. Perhaps the key actors were not prominent enough to do so. For example, low-level bureaucrats working to implement mundane municipal regulations are less likely to see themselves as being important players, and therefore may not leave behind records of their decisions and motives. Observable traces of the model also may be left out of the evidentiary record on purpose. Individuals engaged in illicit activities have strong incentives to obscure or destroy evidence of their actions. Recently released documents detailing the corrupt practices of the Marcos regime in the Philippines illustrate how

complicated these schemes can be, and the lengths to which dictators go to conceal their wrongdoings (Davies 2016). Even in more open settings, elites may have professional incentives to cast their deeds in a positive light, obscuring important features of the decision-making process (Gonzalez Ocantos 2016). Besides issues of illegality, most people are predisposed to social desirability bias and a desire to avoid embarrassment, and so may take care to conceal undesirable intentions or behaviors.

When confronting missing data about any of the component parts of the causal chain the researcher must grapple with the question of whether “absence of evidence” actually means “evidence of absence” (Bennett and Checkel 2014). That is to say, whether missingness implies rejection of the model, or if alternative avenues can be pursued to try to salvage it. In the following section, we begin to address this question by deriving a simple typology of missingness.

II. Types of Missing Data

Missing data can be of radically different kinds. Policy makers may be unavailable for interviews; a diplomat may have failed to report on key aspects of an international negotiation in the cables he channeled back to his superiors; or bureaucrats may have destroyed evidence of illegal behavior. What amounts to a relevant piece of “smoking gun” evidence varies dramatically, depending on the questions we ask, and the theories we devise to answer them. For this reason, it is difficult to come up with a comprehensive list of the data that could go missing. However, it is possible to think about general types of missingness.

We think it is useful to draw on what Waldner (2014, 128) calls the “completeness standard” when thinking about the types of missingness. In summarizing the evidentiary standard for process tracing, he writes:

“Process tracing yields causal and explanatory adequacy insofar as: 1) it is based on a causal graph whose individual nodes are connected in such a way that they are jointly sufficient for the outcome; 2) it is also based on an event-history map that establishes valid correspondence between the events in each particular case study and the nodes in the causal graph; 3) theoretical statements about causal mechanisms linking the nodes in the causal graph to their descendants and their empirics of the case studies allow us to infer that the events were in actuality generated by the relevant mechanisms; and 4) rival explanations have been credibly eliminated, by direct hypothesis testing or by demonstrating that they cannot satisfy the first three criteria listed above.”

The completeness standard implies that missingness can take at least three forms. First, the causal chain may be incorrect or theoretically incomplete. In other words, we could be guilty of making logical leaps in our causal account and failing to propose a plausible or sufficient sequence of variables leading from cause to outcome. This form of missingness resulting from **model misspecification** does not concern us here, as it has more to do with theory building and identifying the correct explanatory variables, and less with the challenge of empirically validating our hypotheses.

Second, data to document key steps in the event history map may be hard to come by, producing **undocumented steps**. While the causal chain consists of random variables, the event history map consists of the “values” that independent, intervening and dependent variables take in a given case. That is, the event history map outlines the empirical events, actions/reactions, or state of affairs that logically occur between cause and effect. In doing so, it lays out the empirical implications that would be predicted by the causal model. Assuming the model is correctly specified, our inability to document that any of these events, state of

affairs, or actions actually occurred would undermine the continuity of our narrative (Waldner 2014).

Finally, after successfully establishing a continuous causal chain in line with the model, we could still face **hidden mechanisms**. Mechanisms are what makes this system “tick” in the direction it does (Bunge 2004). They typically refer to the *reasons* an actor behaves in certain ways or reacts to environmental cues (George and Bennett 2005; Gerring 2010; Bennett and Checkel 2014; Little 2015). Hidden mechanisms occur when we fail to empirically demonstrate these underlying causal forces. Without these mechanisms, we cannot account for what happens on the path between cause and outcome, and the reasons that led actors to push the causal chain forward in each of the relevant steps (see also George and Bennett 2005, Collier 2011).³ For example, officials investigating a plane crash may successfully reconstruct the trajectory of plane by collecting data on altitude changes from radars and debris from the ocean, but in the absence of the plane’s black box they may never know why the plane followed that trajectory. In that case, the narrative would have successfully accounted for *how* we get from one state of affairs to the next, but not *why*, thus undermining our ability to move from description to explanation (Waldner 2012).

To illustrate this typology, consider the following example. One of us is interested in explaining why some autocrats in post-Soviet Central Asia treat oppositions more harshly than others. The argument proposed to answer this puzzle is as follows. The structure of the local economy generates different opportunities for corruption: where the state owns large oil and gas reserves, the autocrat can extract huge bribes from foreign contractors, whereas if these resources are not present, money from bribes must come from local businesses. In the

³ In order to satisfy the fourth element of the completeness standard, the evidence we collect also must bolster our own explanation vis-à-vis its alternatives. Sometimes, however, the available data is inconclusive. But this is a different type of problem. The absence of data is potentially more problematic than collecting inconclusive evidence. Although in the latter case we cannot adjudicate between competing explanations, we at least have sufficient data to lend support for the model.

former case, the autocrat has more incentives to repress civil society because he doesn't rely on their profits and payments for personal enrichment (AUTHOR 2017). Given this causal chain, the next task was to document the events, actions or states of affairs that relate to each step. Thus, for the case study of Kazakhstan, these included: the characteristics of the local economy (reliant on oil and gas production), the flow of bribes to the autocrat (from foreign oil companies), and the levels of repression (high, because the autocrat did not extract bribes directly from his people). The absence of evidence to document any of these points would have produced **undocumented steps**.

During fieldwork in Kazakhstan, it was possible to document that the domestic economy was reliant on oil and gas production and that repression was high (AUTHOR 2014). However, the intervening event—i.e., the flows of bribes from foreign contractors to the autocrat—remained out of sight, until a series of foreign court proceedings and investigative accounts were discovered that shed light on how the president and his inner circle enriched themselves through the sale of ownership and development rights to foreign companies (e.g., United States of America v. James H. Giffen, 2003). It also was difficult to uncover evidence of the key mechanism linking these money flows to high repression—i.e., evidence that government officials evaluated the pros/cons of violence towards their own people in light of the type of bribes they rely on. Missing this crucial data constituted a case of **hidden mechanisms**, and made it impossible to validate the theoretical argument. But armed with evidence of corruption obtained from the court cases, it was possible to probe mechanisms during interviews with experts on Kazakhstani politics, as well as to collect news articles that revealed how the country's political economy shaped the autocrat's thinking about his political opponents. These inferences relied on secondhand sources and publicly available statements, and thus held lower probative value than if regime insiders had revealed their thinking, but nonetheless offered evidence in support of the argument.

Some might argue that process tracing is simply not an appropriate tool for investigating “data poor” topics and environments such as these. Proponents of a micro-foundational social science suggest that these are precisely the type of research questions that one needs to avoid. We ought to shorten our causal chains, make them tractable, and to the extent possible, avoid arguments with crucial micro-foundational components that cannot be directly tested (Elster 1998). We believe this position reflects an unnecessarily strict interpretation of current best practices. Conversely, other scholars might question the relevance of even thinking about how to approximate completeness. It may not be strictly necessary to understand the full process in order to confidently claim that X causes Y; a single piece of “doubly decisive” evidence might suffice (Collier 2011) Or perhaps, we do not need a fully specified roadmap to conduct process tracing, as induction often plays an important role. We believe, by contrast, that although causal models are rarely fully specified at the beginning of a research project, some type of roadmap is necessary in order to approach the data collection process and evaluate the compatibility of evidence with competing explanations. More importantly, the causal model gives readers the opportunity to evaluate the evidence we present against a bar set by ourselves, facilitating transparency, reasonable debate, and refutation.

With the goal of expanding the applicability of process tracing techniques while maintaining high standards of rigor, in what follows we introduce a series of remedial instruments to help researchers address the problem of missingness. We believe that the remedies for missing data proposed in the remainder of the article are useful even for those researchers who do not aspire to some level of explanatory completeness. Such scholars may still find that some instances of missingness undermine the credibility of their causal narratives; our three remedial instruments can help them think of ways to mitigate the problem. Some of our suggestions are designed to address missingness as it relates to

undocumented steps, while others squarely address the problem of **hidden mechanisms**, i.e. motives for action, reasons why given states of affairs trigger certain reactions, or cognitive processes. Table 1 summarizes the three remedial instruments.

Table 1: Summary of the Remedial Instruments

	<i>Type of Missingness Addressed</i>	<i>Procedures Involved</i>	<i>How it Addresses Missingness</i>	<i>Examples</i>
Contextualize the Data Generation Process	* Undocumented steps * Hidden mechanisms	Consider actors' interests vis-à-vis publication of tangible evidence. Evaluate the likelihood that data would exist if the mechanism specified in the model to explain behavior is valid.	Suggests a theoretically coherent justification for missing data. Mechanisms specify the reasons why actors behave in certain ways. These same reasons may lead actors to destroy evidence. Missingness is therefore compatible with the argument, and not disconfirmatory.	Morrison (2016); Freeland (2000)
Indirect Tests of Mechanisms	* Hidden mechanisms	Recreate mechanism at higher level of abstraction and test observable implications outside original context	Provides an indirect, but concrete empirical basis for establishing the plausibility of the mechanism.	Lupu (2016)
Specify Analytical Status of Steps in the Causal Chain	* Undocumented steps	Make explicit the explanatory status of each node in causal chain using necessity and sufficiency claims	Reduces the need to provide evidence for steps following sufficient steps. It thus minimizes the disconfirmatory effects of missing data.	Blyth (2002)

III. Contextualizing the Data Generation Process

One means by which process tracing can proceed despite missing information is through a careful *contextualization of the data generation process*. When confronting evidentiary gaps, researchers are encouraged to consider the ways in which the actors involved in the causal process might also have an incentive to prevent the creation or

dissemination of traceable evidence of their actions. After all, it is the protagonists themselves that generate observable data. Crucially, a thorough characterization of the data generation process allows us to evaluate whether or not the absence of data is compatible with the microfoundations of the theory, and consequently, whether or not missingness means that the argument should be revised. This is because the nature of mechanisms (e.g. crass motives for action) determines the likelihood of finding certain evidence even if the predictions are correct. Consequently, as we shall see below, the nature of the mechanisms often also accounts for the absence of data.

Contextualizing the process of data generation entails making a thorough evaluation of the setting, the actors, and their individual interests vis-à-vis the potential publication of tangible evidence. Thus, we can evaluate the meaning of these evidentiary gaps by “paying especially close attention to the ways in which institutional and political contexts of choice generate *strategic incentives*” which include “pressures for actors to speak, behave, or keep records in ways that occlude, rather than reveal, the considerations motivating their decisions” (Jacobs 2014, 42). But which aspects of context matter? The data generation process can be assessed along four dimensions. First, the normative status the topic of interest holds within society: Is the actor’s behavior or the outcome that results from it considered positive, negative, or neutral? Are these actions legal or illegal? Second, the nature of the political environment: Is there significant possibility of reprisal by the state or other political forces? Assessment of this factor requires consideration of the state’s interests, as well as the government’s capacity for monitoring and punishment of dissenters. Third, the actor’s own goals: What was the actor’s intention in participating in this interview or chronicling this information? Who is their audience? What are their personal and professional ambitions? These factors influence both the likelihood of the individual sharing information, as well as

the reliability of the information they reveal. Fourth, the role of the researcher: Has the identity or actions of the person collecting data influenced the type of data that is generated?

Historians and political methodologists often encourage this type of contextualization exercise in order to engage in “source criticism” (Milligan 1979; Kreuzer 2014; Bennett and Checkel 2014). The underlying message of this literature is that a healthy dose of skepticism is always recommended when making inferences from the data that we do find. Our contention is slightly different: when probing theoretical models in “data poor” environments we should perhaps be less skeptical, and fiercely interrogate the data in order to re-signify silences in the evidentiary record.

Explicitly analyzing the goals and agendas of relevant actors can help distinguish the meaning behind an absence of empirical evidence. If there is little reason to believe that the protagonists prevented the creation or dissemination of this data, then the researcher can reasonably interpret the lack of evidence to mean that the hypothesis needs revision. However, if there is reason to believe that the likelihood of finding a piece of evidence is close to zero—because some actor had an interest in not leaving observable traces—then this suggests a different interpretation. In that case, the lack of evidence does not necessarily undermine the validity of the hypothesized causal process. Well-specified causal mechanisms are a crucial aid in this contextualization exercise because actors’ reasons for behaving in particular ways, which push the causal chain forward, are usually correlated with their incentives to disclose or hide those deeds as well as the motives behind them.

Consider James Morrison’s (2016) study of Britain’s 1931 suspension of the gold standard. Morrison’s narrative hinges on the dynamics between two key actors: Central Bank Governor Montagu Norman, who favored solving the sterling crisis by raising bank rates, and his Deputy, Ernest Harvey, who advocated austerity. Morrison argues that Britain’s decision to abandon the gold standard can be traced back to the ideas held by these policymakers, as

well as shifts in their relative power. While his argument is extensively supported by primary source material, Morrison also admits with regard to the archival record, “there are many gaps; and there is no way of knowing the size, nature, or importance of those gaps”. To approximate completeness he leverages the likely incentives of these key actors to preserve or destroy their records in order to ascertain what those gaps might mean. He describes his reasoning in the extended appendix (p.2): “*A conspiracy lies at the heart of my account...* As such, my account expects that various figures—such as Ernest Harvey at the Bank—had incentives to ensure that he did not leave behind a paper trail documenting his attempts to oust the Labour Government.”

Morrison’s contextualization exercise is compelling precisely because the incentives he identifies regarding the production of observable traces are a function of the same mechanism that drives the hypothesized causal process. In both cases, the actors’ behavior is driven by partisan loyalties and intra-agency competition. Morrison is reflective about how the availability of records affects his confidence in the argument: “Ultimately, the materials consulted here cannot be said to “prove” the veracity of the particular interpretations that underpin the accompanying article. They can show, however, that those interpretations are not demonstrably wrong.” In cases such as these, the absence of evidence weakens the claim about the actions and motives behind the outcome of interest. But a careful characterization of the motives driving those in a position to produce relevant information suggests that Morrison’s hypothesis is plausible.

Besides overcoming the absence of empirical evidence, contextualizing the data generation process also can bolster the researcher’s confidence in the substandard pieces of data that *are* found. Especially when working on politically sensitive topics, a researcher may find only a few sources that disclose the intimate, socially undesirable, or illegal details of causal processes. But those isolated pieces of evidence may in fact be very reliable and have

a high probative value when considered with an eye toward the context. For example, in her investigation into Russia's economic reforms, Chrystia Freeland (2000, 176) struggled to understand why the government had enacted "obscure and fuzzy laws" that discouraged foreign investment. Freeland cites one interview with a prominent businessman as being pivotal for elucidating the causal processes:

"A few years later, Kagalovsky agreed to explain to me his technique for keeping the foreigners out. The key, he said, was to ensure that the law banning foreign participation was intentionally vague and open to multiple interpretations. [...] But, how, I wondered, had Kagalovsky ensured that the laws were written in so precisely vague a fashion as to make it legally too risky for foreigners to participate? I assumed as a matter of course, that Kagalovsky wouldn't tell me; but out of a kind of journalistic duty, I asked anyway. To my surprise, Kagalovsky gave a frank and rather proud answer: 'Well, of course, I wrote the law myself, and I took special care with it.'" (pg. 176)

Freeland states that, after over a decade of living in Russia and countless of interviews with politicians and businessmen, this unlikely admission singlehandedly changed her understanding of the micro-level dynamics at play: "I don't think [Westerners] imagined that Russia's laws were vague on purpose; that their obscurity was not an accidental barrier to investment, ...but that it was quite intentional" (Freeland 2000, 176). In a different context this quote could be downgraded to circumstantial evidence or even cheap talk. But by explicitly discussing the context in which these details were revealed—noting that this was a highly reliable source who she did not expect to confess to such actions—Freeland contributes a rationale for drawing a bold conclusion based on a single piece of evidence.

To be clear, developing an accurate understanding of actors' likely motives and the contexts in which they operate requires in-depth case knowledge. It also requires researchers

to spend considerable time collecting evidence from a wide range of sources, often using a multitude of methodologies, in order to offer corroboration for their assumptions. But the payoff for this investment of time and resources comes in the transparency it affords. In particular, it encourages clarity and precision in the logic by which we assign probative weight to each piece of data. At the same time, the process allows consumers of the analysis to consider whether the characterization of the context is accurate, in order to evaluate the quality of the researcher's conclusions. Being explicit about the process by which data comes into existence thus contributes to efforts to increase transparency in qualitative research.

IV. Indirect Tests of Mechanisms

Our second remedial instrument addresses missing data that arises when trying to uncover hidden micro-level mechanisms, such as those behind macro-historical processes. Macro-historical causal chains can present important evidentiary challenges for process tracing. As George and Bennett (2005, 141-142) explain, even larger historical phenomenon “must operate through the perceptions and calculations of individuals.” Yet, the individual-level actions, motivations, and cognitive processes that collectively engineered the outcome of interest (for example, revolutions, democratization, party system change) may not have been documented. Average citizens who participate in mass politics—by voting, protesting, or striking—and who collectively contribute to macro-social processes, often do not leave behind detailed narratives to account for their reasons for action. Researchers cannot travel back in time in order to document the type of thinking that motivated the group behavior responsible for key episodes of contentious politics. In the absence of a researcher who at the time of the events took the initiative to gather this data, important mechanisms may remain hidden (Schwartz and Cook 2002).

In these cases, we propose that process tracing may proceed through an *indirect test of mechanisms*, which involves testing the observable implications of micro behavior and psychological processes outside of their original context. Data gathered in a different time period, even involving different actors, can effectively shed light on the mechanism in question, especially when such data is the result of replicating the same fundamental psychological cues and reactions to external stimuli specified in the theory. We suggest that researchers can translate the hypothesized mechanism into a more or less generalizable pattern of human behavior or cognition, by couching it in non-context specific ways, and then test its observable implications among similarly situated actors using experiments, focus groups or other methods.

In proposing this remedial instrument, we contend that the driving principles behind some of the component parts of the causal chain can indeed be recreated outside the historical context of interest. As Falleti and Lynch (2009, 1149) explain, mechanisms “may be cast at different levels of abstraction,” yielding more or less portable concepts. It is possible to conceptualize the logic of cognition, information processing, or decision-making that guides action in the causal model as an instance of a broader, more extensive type for which data can be found or created in the present. Climbing the ladder of abstraction allows the researcher to approximate core micro-foundational aspects of these “unreachable” steps.

One may be skeptical about the existence of these “law like” patterns of action and cognition, and therefore doubt the value of generalizations devoid of contextual sensitivity. When it comes to investigating actors’ motives and incentives, however, researchers constantly rely on more or less unbounded “law like” statements or assumptions about how humans think or behave. This logic governs lab experiments in the social sciences, which are predicated on the assumption that they allow researchers to explore fundamental, portable individual thought processes or logics of action, and that those processes would be difficult to

readily observe in an uncontrolled environment. Collier (2011, 824) similarly suggests in process tracing, researchers inevitably resort to “recurring empirical regularities” that cut across contexts to evaluate the diagnostic power of evidence found in very specific contexts (see also Mahoney 2012, 587). It follows that even the most contextual mechanistic accounts are almost always partly based on generalizable assumptions about what motivates human beings to act in certain ways under certain conditions.⁴ For example, a detective that does not find anything missing from the murder scene may use general knowledge or assumptions about how burglars operate to discard a whole array of suspects (“burglars steal things, so the culprit is unlikely to have been a burglar”). To the extent that the hidden microfoundations of any historical sequence tap into these generalizable forms of behavior, cognition, or information processing, it is possible to document them using data gathered outside the immediate context.⁵

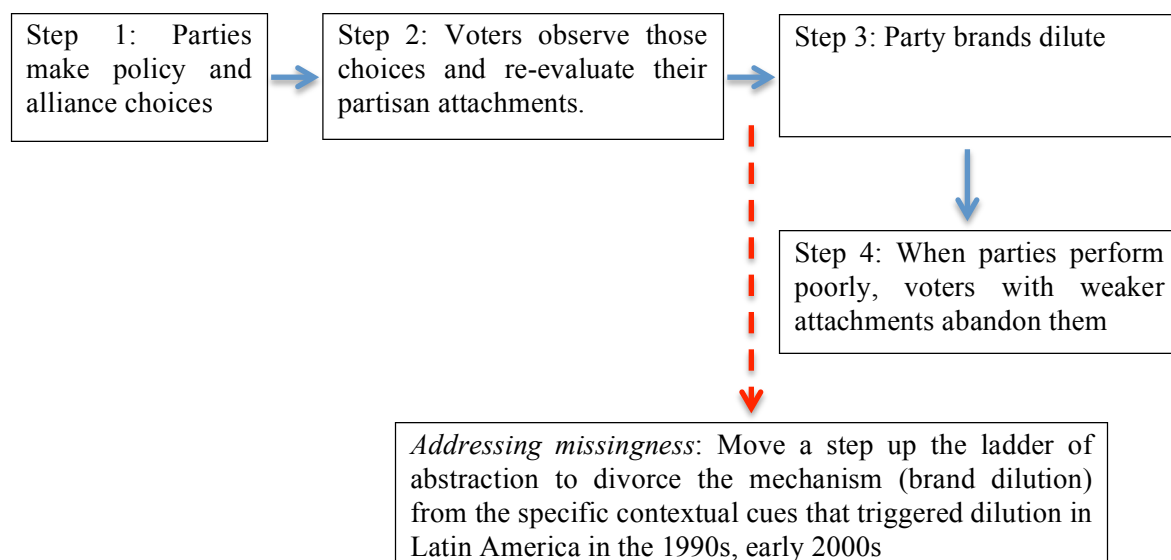
An example illustrates these points. Noam Lupu (2016) is interested in explaining why some established Latin American political parties collapsed in the 1990s and early 2000s. In his view, this process was governed by elite policy choices, society’s perceptions of those choices, and voters’ subsequent electoral responses. The theoretical model of collapse is triggered by parties that implement policies which are largely inconsistent with their traditional programmatic positions (e.g. labor based parties that implemented neoliberal policies), and that forge “strange-bedfellow” alliances with their rivals (Step 1). These choices dilute party brands and erode voters’ partisan attachments (Steps 2 and 3). Without an entrenched partisan core, parties run the risk of being voted out of office on the basis of

⁴ It is important to note that not all mechanisms are “regular” in this way. See Runhardt (2016)

⁵ To use an analogy from the natural world, we may not be able to readily observe how or why our window shattered after being hit by a baseball bat. But given the generalizable properties of glass, we could study similar materials in a laboratory to uncover the ways in which the molecular structure of glass leads to shattering when it is hit by a harder object.

short-term retrospective evaluations. Brands protect parties from the effects of bad performance, but when brands weaken, voters are more likely to jump ship, leading to party collapse (Step 4) (Figure 1).

Figure 1: Brand Dilution and Party Collapse (Lupu 2016)



The mechanism linking the second and third nodes of the causal chain has key politico-psychological microfoundations that Lupu cannot test directly. This would have required the fielding of very specific survey instruments or focus groups at the time of the events in order to gauge voters' reactions to neoliberal policy U-turns or the formation of unprecedented coalitions. But although the mechanism of brand dilution is designed to explain how voters changed their perception of parties in the midst of the political turmoil that characterized Latin American countries in the 1990s and early 2000s, the logic that underpins this part of the argument is hardly historically specific. In fact, Lupu develops a more general theory of how people become or cease to be *partisan*. Armed with these theoretical insights he is able to persuasively argue that the cognitive reaction at the heart of the model should travel well given more or less replicable cues. So while Lupu cannot

observe how levels of partisanship declined as voters saw labor-based parties implementing neoliberal reforms, he can test how partisanship suffers as present day voters receive confusing programmatic cues.

Lupu proceeds to deal with missingness by designing a series of out-of-context survey experiments that allow him to document the fundamental logic of information-processing behind brand dilution.⁶ As expected, voters in the treatment group that received information about alliances and programmatic switching expressed weaker partisan attachments than those who received messages emphasizing strong programmatic distinctions between parties, or a mix of both cues. To be sure, since the survey was administered years after collapse, the environment in which voters were observed reacting to partisan cues differed from the original one in important ways. For example, new parties had emerged. In addition, the cues used in the experiment are necessarily different in content from the specific policy shifts that triggered the causal process under scrutiny. But despite these differences, Lupu plausibly recreates key psychological reactions in part because he uses a nationally representative sample, implements the survey in one of his case studies, and provide voters with information they can easily relate to. All of these features of the design are important because they minimize the contextual differences that inevitably exist between the actual historical process and the setting in which microfoundations were documented.

This way of plugging gaps in incomplete causal chains can lead to productive synergies between within-case analysis and quantitative tools, such as survey experiments, which are especially adept at exploring micro-foundations. But other methods such as focus groups or lab experiments can also prove useful to probe the motives of inaccessible subjects. However, an important caveat follows. The indirect, or “out of context” approach, requires disembedding fundamental aspects of the social process under study from their immediate

⁶ In another excellent study, Seawright (2012) adopts a similar strategy to study party system collapse in Latin America.

context. The credibility of the test will largely depend on its continued sensitivity to the original context, for example, by relying on subjects who live under broadly similar conditions compared to those responsible for historical outcome of interest.

V. Specifying the Analytical Status of Steps in the Causal Chain

Finally, when faced with undocumented steps in the causal chain, scholars can *specify the analytical status of steps in the causal chain*, and use this information to minimize the disconfirmatory effects of missing data. In a nutshell, we contend that in the absence of key data, a careful consideration of the explanatory status of missing steps in the causal chain, and that of those that immediately precede or follow them, can provide researchers with an alternative means to support their argument. Specifically, if a researcher can make a plausible case that one step in the causal chain for which no data is available is logically implied from a prior step, missingness should not be considered disconfirmatory.

Explanatory models in case-oriented social science tend to assume that the underlying causal world that governs political or social phenomena is inherently complex. These models often feature highly conjunctural causes that are only efficient in the company of a multiplicity of others (Ragin 2014; Hall 2003; Goertz and Mahoney 2012). Due to this complexity, scholars in this tradition tend to think hard about the relative importance of the component parts of their theories, and the degree to which they each contribute to the overall explanatory power of the model.

The causal chains that specify the processes whereby each independent variable contributes to the outcome of interest are similarly complex because they are usually made up of multiple steps. Importantly, in any hypothesized causal chain the analytical status of each step need not be equal. As others have argued, each step in a causal chain can be categorized as either necessary or sufficient to obtain the subsequent step (Mahoney and Schensul 2006;

Goertz and Levy 2007; Mahoney 2012). Crucially for our purposes, sufficiency links suggest that the next step in the chain *will* be observed. This can be leveraged to minimize the explanatory damage caused by missingness because it implies that events or actors that follow sufficient steps in the causal chain are less important to demonstrate empirically. In other words, steps in the causal chain that have a probability of being observed close to 1 if the previous step is observed are “overdetermined.” And if missing data corresponds to a step that is theoretically overdetermined by the previous step—and we are able to document that previous step—then the causal argument may still stand. Similarly, if steps that generate a sufficiency link cannot be documented, but their inevitable consequences can, then there is a strong basis for assuming that both steps were present in the case of interest (Mahoney 2012).

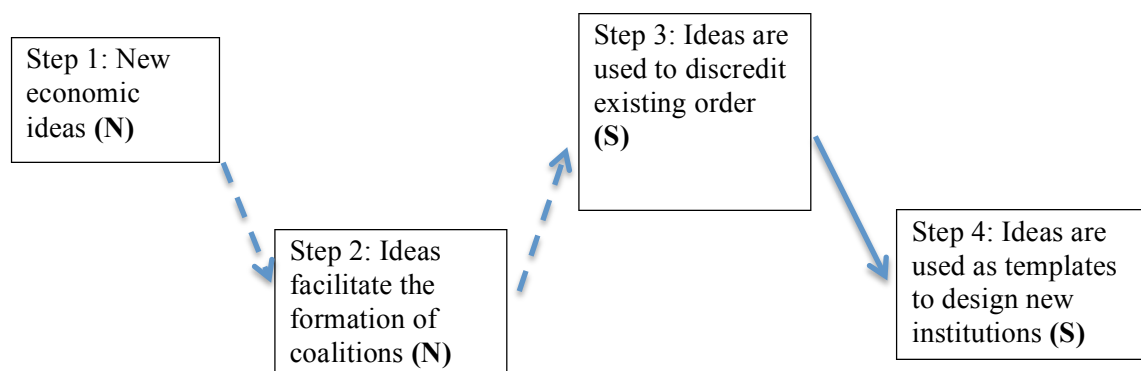
It is important to note, however, that this does not mean that sufficiency links between steps are causally more important than necessity links. After all, the absence of a necessary link interrupts the flow of causal energy and suggests that the outcome will not be produced, while a sufficiency link is fungible. Put differently, counterfactuals are easy to imagine in the case of necessary links (“if no X, no Y”), whereas with sufficiency links, the absence of X does not imply the absence of Y (Goertz and Levy 2007, 16). But when it comes to evaluating the implications of missing data, we argue that sufficiency links can come to the aid of the researcher in ways that necessity links cannot. As Goertz and Levy (2007, 26) explain, a sufficiency link “actually produces the next link in the causal chain, whereas that is not the case for necessary condition links.” In other words, while sufficiency links are enough to guarantee that the following step to materials, necessary links are not.

Mark Blyth’s *Great Transformations* (2002) uses this kind of theorizing. Blyth is interested in understanding how actors make sense of their economic interests during crises. In particular, he traces how the influx of new economic ideas was responsible for important changes in the political economy institutions of advanced industrialized countries during two

different critical junctures. One of the book's main strengths is the careful specification of the steps in the causal chain, and of the status of each step in the model.⁷ The chain has five steps, four of which are relevant for us here (ibid, 34-45).

First, new economic ideas reduce uncertainty in the midst of economic unrest. In order for actors to think of responses to the crisis, it is necessary for them to understand what is going on, what the crisis is about, and why it came about. Second, ideas make possible new coalitions; they facilitate collective action by bringing together individuals and groups in pursuit of institutional changes. Third, ideas provide intellectual and discursive resources that these groups can leverage to defy the economic order. They discredit existing institutions by associating these with economic failure. Finally, ideas become blueprints and dictate what institutional choices will be made to overcome the crisis and avoid a new one (Figure 2).⁸

Figure 2: How Ideas Trigger Institutional Change During Economic Crises (Blyth 2002)



Note: The diagram only reflects parts of Blyth's causal chain.

⁷ Blyth refers to these “steps” as “hypotheses.” Our reconstruction of the causal chain doesn’t fundamentally alter his theoretical framework.

⁸ There is a fifth step in the causal chain: “Following institutional construction, ideas make institutional stability possible” (41).

Whereas the first two steps are necessary for the following one to occur, Blyth (2002, 39) suggests that the last two are sufficient: “While the reduction of uncertainty and the generation of collective action create the necessary conditions for institutional transformation, the sufficient conditions lie in the subsequent roles that ideas play as weapons and blueprints with which ideas can contest and replace existing institutions.” Once ideas are used to discredit existing institutions, their use to argue in favor of specific new institutions is likely over-determined. Let us imagine that the researcher is able to document that agents use ideas to discredit the existing order (Step 3) but unable to show that these same actors use those ideas to argue *in favor* of new institutions (Step 4). Because Step 3 logically implies Step 4, this instance of missingness would not necessarily undermine the overall argument about the role of ideas in institutional change. For example, if the researcher is denied access to records of congressional hearings debating the design of new central banking laws during which these ideas were presumably used as templates for institutional change, he or she might still proceed. Provided that it is possible to show that related ideas were used to argue *against* the previous institutional order, the argument may still stand. This is a plausible assumption: if ideas became weapons against the status quo, it is highly likely they were also actively cited when discussing institutional change.

It goes without saying that we find this solution less satisfying than the others. Logic and theoretical specification are no substitutes for good empirics. One should not use logical shortcuts or theoretical assumptions to deliberately avoid collecting data on intervening steps. The theory’s underlying logic may be air tight, but models with too many assumptions that are not subject to empirical tests are less appealing than those with fewer ones. In the absence of key data, however, a careful consideration of the explanatory status of missing steps in the causal chain, and that of those that immediately precede, or follow them, can provide researchers with an alternative means to support their argument. But because there is little

validation for claims about bonds of necessity or sufficiency between the different component parts of the chain, researchers using this remedial instrument must make their reasoning as transparent as possible, thus opening the door for others to scrutinize the analytic procedures employed to draw inferences from incomplete data.

VI. Discussion

Having outlined these procedures, two further points arise regarding the usage of these remedial instruments. First, how might researchers choose among them? One option is that researchers might select an instrument according to the type of missingness that needs to be addressed. Some of our suggestions offer remediation against undocumented steps in the causal chain. For example, specifying the analytical status of each step in the causal chain involves making explicit (in terms of necessary/sufficiency claims) the explanatory status of each node in causal chain, thus reducing the need to document steps that occur after sufficiency claims. Others rectify the problem of hidden mechanisms. By conducting indirect tests of causal mechanisms, researchers can recreate aspects of the causal chain outside of original context and in order to gauge their plausibility as a motive for action. Finally, researchers can mitigate both undocumented steps and hidden mechanisms by contextualizing the data generation process, which requires them to consider actors' interests vis-à-vis the publication of tangible evidence, in order to evaluate the likelihood that data would exist if the hypothesis were true.

At the same time, we encourage researchers to use our three remedial instruments in tandem, in order to ensure more robust conclusions. For instance, reinterpreting data silences in light of realistic insights about actors' motivations to hide their actions and motives from public view, *and* performing an indirect test of the plausibility of those micro-foundations, would lead to a more credible "imputation" strategy than using any of these two remedies in

isolation. This is because one remedy is used to “check” the untested assumptions that underpin the other. Importantly, when combining remedial instruments, researchers should consider the different epistemic status of each one. Whereas the first two rely on actual data to make judgments about the credibility of causal explanations amidst missingness, the third one relies on theory and logic. As a result, the third remedy should be seen as a technique of last resort. In addition, using data from primary sources to reinterpret silences, as in the case of our first remedy, is in principle preferable to making assumptions about a process that took place in the past using data from another context and time period, as in the case of the second remedy. This is because the latter remedy creates “artificial” data, and therefore makes claims about external validity not needed in the former.

A second point concerns the specter of confirmation bias. Sometimes the absence of evidence is, in fact, evidence of absence. How can researchers avoid using these tools to erroneously validate an incorrect hypothesis? On this front, we advocate a transparent use of the three remedial instruments. Researchers should always preface the use of these methods by admitting that they constitute an effort to address missingness. This should include an account of the type of primary source data that they could not find in the first place. In addition, our remedial instruments all make assumptions that should be discussed at length. For example, characterizing the underlying data generation process requires a thorough account of the unobserved motivations of key actors. Ideally, this account should be supported with ancillary data about the context and the actions/motives of the relevant actors that can actually be collected. When relying on these methods, researchers also can strengthen their claims by drawing from a variety of different source material. Inferences about the context, actors’ interests, and their motives for action will be more reliable when supported by multiple pieces of evidence from different perspectives, rather than by a single data point.

VII. Conclusion

Tracing the causal processes that lead to specific outcomes can be problematic when confronting hidden data and buried mechanisms—that is, when data does not exist or is inaccessible, when subjects lie or remain silent, and when micro-foundations function too deeply to be readily observed. With this in mind, this paper has sought to identify tools to confront these challenges while still maintaining high standards of rigor and transparency in qualitative research.

Our point of departure was an understanding of process tracing that puts emphasis on the actor-centric nature of causal chains and mechanisms, and on the importance of theory building. Prior specification of who the actors are, how they think, and what they want, as well as a clear statement of the key events in the sequence of actions and reactions leading to the outcome, is essential for effective data collection. A rigorous test of the causal argument via process tracing requires the careful reconstruction of each one of those steps. Bearing in mind this actor-centric view of mechanisms and highlighting the centrality of good theory for good process tracing, we developed a series of remedial instruments, as well as research design choices that can help researchers approximate the completeness standard.

The article makes three contributions to the qualitative methods literature. First, the paper brings the literature on mechanisms into dialogue with work being done on the nature and implementation of Bayesian updating. By advocating a specific understanding of what causal mechanisms are—i.e., one that is actor-centric—and by emphasizing the importance of theories about how these actors interact with the world around them, we are able to more clearly interpret the logic behind diagnostic tests. For example, our discussion of the need to qualitatively assess the data generation process in terms of the actors specified in the theory and their context, points to a crucial step prior to the estimation of probabilities in the spirit of the Bayesian framework.

Second, this paper has clear implications for current debates about the nature of transparency in qualitative research. Our remedial instruments have in common that they encourage researchers to clearly outline the thought processes behind their descriptive and causal inferences. Like the literature on diagnostic tests, we believe it is crucial to outline for readers the main assumptions that impact how we assess and weight the evidence. Others may not be able to replicate our interviews, but they should certainly be able to understand why we think this or that quote provides a boost to our theoretical argument. Similarly, instead of glossing over the problem of insufficient evidence, or taking it for granted, our discussion points to the need to openly talk about it, referring to ways in which the theory and the context of data generation might allow us to reduce the “confidence intervals” behind our inherently uncertain claims.

Third, the paper provides an additional rationale for reinforcing current calls for more and better mixed-methods research. If a crucial component of process tracing is indeed the tracing of deep micro-foundations in the form of actors’ thought processes and motives for action, experimental and survey research have a lot to contribute to small-N research designs. There are clear elective affinities. Our discussion of indirect tests in macro-social studies is an example of how to realize these synergies.

VII. Bibliography

- Beach, Derek. 2016. “It’s All About Mechanisms – What Process-Tracing Case Studies Should be Tracing.” *New Political Economy*, 21(5):463-472
- Beach, Derek, and Rasmus Pedersen. 2014. *Process-Tracing Methods: Foundations and Guidelines*. University of Michigan Press.
- Bennett, Andrew. 2008. “Process Tracing: A Bayesian Perspective.” In *The Oxford Handbook of Political Methodology*. Oxford University Press.
- . 2014. “Disciplining Our Conjectures: Systematising Process Tracing with Bayesian Analysis.” In *Process Tracing: From Metaphor to Analytic Tool*, Andrew Bennett and Jeffrey T. Checkel eds. Cambridge University Press.
- Bennett, Andrew, and Jeffrey T. Checkel, eds. 2014. *Process Tracing: From Metaphor to Analytic Tool*. Cambridge University Press.

- Blyth, Mark. 2002. *Great Transformations: Economic Ideas and Institutional Change in the Twentieth Century*. Cambridge University Press.
- Brady, Henry, David Collier, and Jason Seawright. 2004. "Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, Brady, Henry and David Collier eds. Lanham: Rowman & Littlefield
- Bunge, Mario. 1997. "Mechanism and Explanation." *Philosophy of the Social Sciences*, 27 (4):410–65
- . 2004. "How Does It Work? The Search for Explanatory Mechanisms." *Philosophy of the Social Sciences* 34 (2):182–210.
- Collier, David. 2011. "Understanding Process Tracing." *PS: Political Science and Politics* 44(4):823–30.
- Davies, Nick. 2016. "The \$10bn question: What happened to the Marcos millions?" *The Guardian*, 7 May 2016, accessed at <https://www.theguardian.com/world/2016/may/07/10bn-dollar-question-marcos-millions-nick-davies>
- Edgars, Adrienne Lynn. 2006. *Tribal Nation: The Making of Soviet Turkmenistan*. Princeton University Press.
- Elster, Jon. 1998. "A Plea for Mechanisms." In *Social Mechanisms: An Analytic Approach to Social Theory*, Peter Hedstrom and Richard Swedberg eds. Cambridge University Press.
- Fairfield, Tasha, and Andrew Charman. Forthcoming. "Explicit Bayesian Analysis for Process Tracing: Guidelines, Opportunities and Caveats." *Political Analysis*
- Falleti, Tulia. 2016. "Process Tracing of Extensive and Intensive Processes." *New Political Economy*, 21(5):455-462
- Falleti, Tulia, and Julia Lynch. 2009. "Context and Causal Mechanisms in Political Analysis." *Comparative Political Studies* 42(9):1143–66.
- Freeland, Chrystia. 2000. *Sale of the Century: The Inside Story of the Second Russian Revolution*. Little, Brown and Company
- George, Alexander, and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. MIT Press.
- Gerring, John. 2010. "Causal Mechanisms: Yes, But...." *Comparative Political Studies* 43(11):1499–1526.
- Goertz, Gary, and Jack Levy. 2007. "Causal Explanation, Necessary Conditions, and Case Studies." In *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*, Jack Levy and Gary Goertz eds. London: Routledge
- Goertz, Gary, and James Mahoney. 2012. *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Princeton University Press.
- Gonzalez Ocantos, Ezequiel. 2016. *Shifting Legal Visions: Judicial Change and Human Rights Trials in Latin America*. Cambridge University Press.
- Guberek, Tamy and Margaret Hedstrom. Forthcoming. "On or Off the Record? Detecting Patterns of Silence About Death in Guatemala's National Police Archive." *Archival Science*
- Hall, Peter. 2003. "Aligning Ontology and Methodology in Comparative Research." In *Comparative Historical Analysis in the Social Sciences*, James Mahoney and Dietrich Rueschemeyer eds. Cambridge University Press.
- Hedstrom, Peter. 2008. "Studying Mechanisms to Strengthen Causal Inference in Quantitative Research." In *Oxford Handbook of Political Methodology*. Oxford University Press.

- Hedström, Peter, and Richard Swedberg, eds. 1998. *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge University Press.
- Honaker, James and Gary King. 2001. "What To Do About Missing Values in Time Series Cross Sectional Data." *American Journal of Political Science*, 54(2):561-581
- Humphreys, Macartan, and Alan Jacobs. 2015. "Mixing Methods: A Bayesian Approach". Unpublished Manuscript. Columbia University.
- Jacobs, Alan. 2014. "Process tracing the effects of ideas." In *Process Tracing: From Metaphor to Analytic Tool*, Andrew Bennett and Jeffrey Checkel eds. Cambridge University Press
- Kapiszewski, Diana, Lauren MacLean, and Benjamin Read. 2015. *Field Research in Political Science*. Cambridge University Press
- King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press.
- King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review*. 95(1):49-69
- Kreuzer, Marcus. 2014. "The Structure of Description: Elements of and Criteria for Evaluating Historical Analysis." Unpublished Manuscript.
- Lyall, Jason. 2014. "Process Tracing, Causal Inference and Civil War." In *Process Tracing: From Metaphor to Analytic Tool*, Andrew Bennett and Jeffrey Checkel eds. Cambridge University Press
- Little, Daniel. 1991. *Varieties of Social Explanation: An Introduction to the Philosophy of Social Science*. Westview.
- . 2015. "Mechanisms and Method." *Philosophy of the Social Sciences*, 45(4-5):462-480
- Lupu, Noam. 2016. *Party Brands in Crisis*. Cambridge University Press.
- Mahoney, James. 2001. "Beyond Correlational Analysis: Recent Innovations in Theory and Methods." *Sociological Forum*, 16(3):575-593
- . 2012. "The Logic of Process Tracing Tests in the Social Sciences." *Sociological Methods & Research* 41(4):570–97.
- . 2016. "Mechanisms, Bayesianism, and Process Tracing." *New Political Economy*, 21(5):492-499
- Mahoney, James, and Daniel Schensul. 2006. "Historical Context and Path Dependence." In *Oxford Handbook of Contextual Political Analysis*, Robert Goodin and Charles Tilly eds. Oxford: Oxford University Press.
- Milligan, John. 1979. "The Treatment of a Historical Source." *History and Theory*, 18(2):177-196
- Morrison, James. 2016. "Shocking Intellectual Austerity: The Role of Ideas in the Demise of the Gold Standard in Britain." *International Organization* 70(1):175-207.
- Ragin, Charles. 2014. *The Comparative Method*. University of California Press.
- Runhardt, Rosa. 2016. "Tracing the Productive Continuity of Social Mechanisms." *Qualitative & Multi-Method Research*, 14(1/2):22-28
- Seawright, Jason, 2012. *Party-System Collapse: The Roots of Crisis in Peru and Venezuela*. Stanford University Press.
- Schwartz, Joan and Terry Cook, 2002. "Archives, Records, and Power: The Making of Modern Memory," *Archival Science* 2:1-19.
- Tannenwald, Nina. 2005. "Stigmatizing the Bomb: Origins of the Nuclear Taboo." *International Security* 29(4):5–49.
- Trampusch, Christine and Bruno Palier. 2016. "Between X and Y: how process tracing

- contributes to opening the black box of causality," *New Political Economy*, 21(5): 437-454.
- United States of America v. James H. Giffen, US District Court, Southern District of New York, April 2, 2003, accessed at <http://www.justice.gov/criminal-fraud/case/united-states-v-james-h-giffen-et-al>.
- Waldner, David. 2007. "Transforming Inferences into Explanations: Lessons from the Study of Mass Extinctions." In *Theory and Evidence in Comparative Politics and International Relations*, Richard Lebow and Mark Lichbach eds. Palgrave Macmillan.
- . 2012. "Process Tracing and Causal Mechanisms." In *Oxford Handbook of Philosophy of Social Science*, Harold Kincaid ed. Oxford University Press
- . 2014. "What Makes Process Tracing Good? Causal Mechanisms, Causal Inference, and the Completeness Standard in Comparative Politics." In *Process Tracing: From Metaphor to Analytic Tool*, Andrew Bennett and Jeffrey Checkel eds Cambridge University Press.
- . 2016. "Invariant Causal Mechanisms." *Qualitative & Multi-Method Research* 14(1/2):28-33
- Way Lucan and Steven Levitsky. 2006. "The Dynamics of Autocratic Coercive Capacity after the Cold War." *Communist and Post-Communist Studies*, 39(3):387-410