Integrated Inferences

A Bayesian Integration of Qualitative and Quantitative Approaches to Causal Inference

A Book Proposal To Cambridge University Press (Accepted)

Macartan Humphreys* **Alan Jacobs**†

December 16, 2014

1 Introduction

Social scientists are increasingly pursuing mixed-method research designs. There is a growing consensus around the virtues of research strategies that include both quantitative with qualitative tools of inference (Ahmed and Sil, 2012). A typical mixed-methods study includes the estimation of causal effects using data from many cases as well as a detailed examination of the processes taking place in a few. Prominent examples include Lieberman's study of racial and regional dynamics in tax policy (Lieberman, 2003); Swank's analysis of globalization and the welfare state (Swank, 2002); and Stokes' study of neoliberal reform in Latin America (Stokes, 2001). Major recent methodological texts seem to offer strong intellectual justification of this trend toward mixing, characterizing small-n and large-n analysis as drawing on a single logic of inference and as serving complementary functions (King, Keohane, and Verba, 1994; Brady and Collier, 2004). The American Political Science Association now has an organized section devoted in part to the promotion of multi-method investigations, and the emphasis on multiple strategies of inference research is now

*Department of Political Science, Columbia University. Email: mh2245@columbia.edu
†Department of Political Science, UBC. Email: alan.jacobs@ubc.ca

embedded in guidelines from many research funding agencies (Creswell and Garrett, 2008).

Despite this strong movement toward combining inferential approaches, however, scholars have offered surprisingly little concrete advice about how data of different kinds should be aggregated, when mixing methods is beneficial, or how to choose the optimal mix of quantitative and qualitative evidence for answering a given research question. *Integrated Inferences* is the first book to introduce a unified framework, grounded in Bayesian logic, that provides guidance for the design of mixed-method research projects and analytical tools for estimating causal quantities of interest from mixed (quantitative and qualitative) data. Among the issues that the framework will address are:

- **Design tradeoffs.** Given scarce resources, how should researchers make tradeoffs between extensive analysis of a larger set of cases and a more intensive examination of a small set of cases? The framework introduced here allows researchers to use their background assumptions about the research situation to identify the optimal mix of qualitative and quantitative investments.

- **Case- and evidence-selection.** How should researchers choose their cases and observations? The Bayesian approach advanced in the book offers novel ways of thinking about the case-selection strategies that yield greatest leverage and about the kinds of evidence that have the greatest probative value.

- **Analytical integration.** How can inferences drawn from different types of evidence be combined to arrive at causal conclusions? The integration of inferences from qualitative and quantitative research lies at the heart of any mixed-method approach. The book (and accompanying software) provides analytic procedures for estimating, from any mix of qualitative and quantitative data, the full range of causal quantities typically of interest to social scientists, including average causal effects, case-specific explanations, the prevalence of causal mechanisms, and the distribution of causal effects in a population (the degree of causal heterogeneity). We illustrate these procedures with applications to major empirical debates in the social sciences.

- **Knowledge cumulation.** How do we learn about the world? All causal inference is grounded in assumptions—e.g., about the probative value of evidence or the assignment of cases to causal conditions—that are generally not directly testable in a given study. The book's framework illuminates how knowledge about the foundational beliefs of causal inferences can cumulate across studies. We specifically show how, within our framework, qualitative data can be used to update the assumptions underlying quantitative analysis and vice-versa. The book further leverages the Bayesian approach to demonstrate how the specific sequence of empirical steps in a research program can be consequential for the cumulative learning that occurs.

2 The Approach: Bayesian Integration of Qualitative and Quantitative Approaches to Causal Inference (BIQQ)

In this book, we present a unified analytical framework for adjoining quantitative and qualitative data in order to draw causal inferences. The approach, Bayesian Integration of Quantitative and Qualitative data (BIQQ), uses Bayesian logic to aggregate the separate inferential contributions of correlational and process-based observations while allowing data of each type to inform assumptions underlying the interpretation of the other type. Bayesian analysis has become increasingly common in quantitative social science and, as qualitative scholars have pointed out (Bennett, 2008), also lies at the heart of process tracing. Yet we are aware of no previous attempt to formally unify Bayesian reasoning about both forms of data. From a formal perspective, the approach that we propose amounts to a straightforward application of Bayes' rule. Put briefly, the method draws leverage from asking how likely we would be to observe a given set of quantitative and qualitative observations if a particular causal theory was true, compared to the likelihood of observing those data if the alternatives were true, while taking into account our prior beliefs about the parameters of interest.

Here we briefly characterize BIQQ, the book's central analytic framework. A detailed exposition of the framework, with empirical applications, can be found in Humphreys and Jacobs (2014), enclosed with this proposal.

The basic intuition underlying the BIQQ approach is as follows. At its core, quantitative analysis draws inferences from the observation of X and Y values for one or more cases. Quantitative inferences hinge critically on assumptions about how cases have been assigned to values on X (e.g., was assignment independent of potential outcomes). Qualitative analysis, in contrast, draws its strength from the observation of specific features of the processes unfolding within a case—sometimes referred to as causal process observations—that are thought to be diagnostic of particular causal effects or mechanisms operating in the case. Inferences in process tracing thus depend critically on the researcher's beliefs about how process-level observations—which we term "clues"—are associated with the causal effects or mechanisms of interest (e.g., about the probability, q, that a clue would be observed if a given X had in fact caused Y).

Despite these differences in method, quantitative and qualitative modes of inference can both be understood as problems of updating beliefs based on the data. Each is readily approached within a Bayesian framework, in which the researcher asks how likely we were to see the data if one hypothesis rather than the alternatives was true while conditioning conclusions on the prior degree of confidence in the hypothesis. In elaborating BIQQ, we first present simple formalizations of Bayesian quantitative and qualitative inference, within a potential outcomes framework, and then take the short step of unifying these inferential procedures into a single Bayesian model.

In the integrated model, the researcher begins with three sets of prior probability distributions:

- a prior on causal effects, that is, the probability that a unit is of a particular causal "type." In the simplest setup, a unit may of a type that reacts negatively or positively to treatment (which we refer to as A (adverse) and B (beneficial) types respectively); or of a type that has negative or positive outcomes *independent* of treatment status (which we refer to as C (chronic) types or D (destined) types, respectively).

- a prior distribution over assignment propensities for a given type (in our framework, p_a or p_b, say)

- and a prior distribution over the diagnostic value of process-tracing clues (derived from the probability of observing a "clue" given a type and treatment status, eg q_{a0}, q_{d1}).

Together, these beliefs define a joint prior distribution over both our inferential assumptions (the p's and q's) and our causal estimands. We then carry out a straightforward Bayesian operation over this joint distribution, calculating the likelihood of observing the observed pattern of correlational and process-tracing evidence for each combination of p, q, and the estimand, given the joint priors.

The result is an updated set of beliefs about the world in light of the data, conditional on our starting assumptions.

The BIQQ approach is extremely flexible. First, it can be applied to any mix of quantitative and qualitative evidence. Moreover, the procedure can be used to generate posterior distributions for many causal questions of interest to both quantitative and qualitative researchers. BIQQ can be used to estimate population-level causal effects, to adjudicate among case-specific explanations, to estimate the proportion of cases in a population subject to different treatment effects, and to discriminate among theoretical logics (e.g., theories of mechanism).

Further, because all assumptions enter the model jointly, the BIQQ framework generates learning about the *assumptions* underlying qualitative and quantitative inference. That is, upon observing a set of mixed data, BIQQ allows us to update our beliefs about both assignment propensities and the probative value of the types of process-tracing evidence that have been employed. Here the framework exploits the fact that mixed-method research involves the simultaneous use of two forms of evidence and lines of inferential reasoning that are partially *distinct* from one another. This partial independence allows us to use quantitative data to update our beliefs about how process-level clues are associated with causal effects and, likewise, to use case evidence to update our beliefs about assignment processes. The approach thus allows forms of knowledge cumulation across studies that can improve causal inferences as a research program unfolds.

Finally, by modeling the process of learning flowing from different types of evidence, the framework yields practical guidance on research design. The framework generates advice on the research situations under which additional quantitative observations (more cases for correlation-based analysis) or additional qualitative observations (conducting process-tracing on more cases or more intensively process-tracing the same cases) are likely to generate the greatest leverage. Further, we demonstrate how BIQQ can be used to guide the selection of cases to maximize the probative value of case-study evidence, conditional on the causal question of interest.

While BIQQ requires the formulation of a number of prior beliefs, the book explains how the framework can be employed heuristically when priors are highly diffuse. For instance, researchers can begin by formulating a "rough idea" of whether they believe a causal effect to be very common or not very common; whether confounding is likely or unlikely; and whether a given feature of a process is highly likely to be observed, moderately likely to be observed or very unlikely to observed under a given causal theory. Scholars can then use the model to explore the consequences of different values in the plausible range either for her findings (given a set of collected data) or for her research design choices (prior to collecting data). Critically, researchers can then express their conclusions explicitly as *conditional* on their priors, or report findings for a range of prior values. In this last sense, the BIQQ framework represents an major advance in research transparency for scholars employing mixed methods.

3 Relationship to existing works

Integrated Inferences is consistent with much of the spirit of major existing works on qualitative and mixed methods, but takes a distinctive approach to the challenge of multi-method research. We can, first, distinguish this book from a few of the most-cited books on qualitative methods and research design in political science.

At a high level, King, Keohane, and Verba (1994) offers a unified logic for thinking about qualitative and quantitative methods, grounded in the concept of an observable implication. King, Keohane, and Verba also apply to qualitative research design several lines of reasoning about causal identification, measurement, and case selection drawn from a statistical framework. Two important features distinguish *Integrated Inferences* form *Desigining Social Inquiry* (DSI). First, DSI's understanding of qualitative inference is essentially correlational, while *Integrated Inferences* builds on more recent understandings of qualitative analysis as drawing on non-correlational within-case evidence of the unfolding of causal processes, as captured most prominently by Collier, Brady, and Seawright (2010)'s concept of a "causal process observation." Second, DSI offers no explicit guidance about how the researcher should integrate findings across qualitative and quantitative observations.

Brady and Collier (2010) and George and Bennett (2005), amomg the most widely cited books in the field, draw a key disinction between the correlational logic of quantitative inference and the within-case, process-oriented logic of qualitative analysis. Integrated Inferences builds directly on this distinction. Collier, Brady, and Seawright (2010) also point toward the possibility of integration across the two logics in arguing that much can be learned from adjoining "causal process" to "dataset observations." Brady and Collier (2010)'s volume also includes a chapter by Sidney Tarrow detailing ways of "bridging the qualitative and quantitative divide." Tarrow's chapter suggests, for instance, that qualitative analysis can help uncover causal mechanisms to explain quantitative correlations, that quantitative analysis can be used to test for the representativeness of causal effects found in case studies, and that the two approaches can be used to triangulate findings. Gerring (2012) takes a further integrative step in showing how a variety of research methods can be understood and evaluated within a single criterial framework. The BIQQ framework, in an important sense, builds on the arguments about multi-method research developed in these works. Our book takes these insights further, however, by formalizing a procedure for combining inferences drawn from correlational and process-based observations into a single set of inferences, and by showing precisely how and under what conditions qualitative and quantitative inference can complement one another.

Integrated Inferences can perhaps be most sharply distinguished from Goertz and Mahoney (2012), which unpacks the differences in assumptions, goals, and conceptualizations that tend to underlie qualitative and quantitative work. The BIQQ approach that we present is in fact quite compatible with Goertz and Mahoney's understandings of much quantitative and qualitative research and is sufficiently flexible to encompass many of the goals they associate with each method; Goertz and Mahoney, moreover, see mixed-method causal inference as feasible, if challenging. The two works, nonetheless, point the reader in very different directions, with Goertz and Mahoney emphasizing the crucial differences between the "two cultures" of social science and *Integrated Inferences* presenting a procedure for integration.

We can also more precisely and broadly situate the book's conceptualization of method-mixing relative to that found in current scholarship. We can identify in the literature four general approaches to the relationship between qualitative and quantitative research. These approaches are characterized by: a.) whether they conceive of qualitative and quantitative methods as addressing distinct or common questions; b.) whether they envision the use of diverse forms or a single form of data; and c.) whether they describe integrated or distinct tools for analyzing the two kinds of data. *Integrated Inferences* is unique in providing a comprehensive treatment of how researchers can undertake integrated analyses of diverse forms of data to answer the full range of causal questions of interest to social scientists.

To briefly describe each approach:

1. Distinct questions. Some authors have characterized qualitative and quantitative modes of inference as seeking to generate distinct forms of knowledge. In general, this implies the use of different data and different analyses.

In some accounts, the in-depth exploration of individual cases can contribute to theory-generation, but the testing for or estimation of causal effects can derive only from the analysis of covariation *across* cases (e.g., Beck (2010)). Other accounts suggest that the contribution of qualitative approaches lies in linking quantitative causal estimates to theory. In Paluck (2010)'s treatment, for example, cross-case experimental evidence provides estimates of causal effects, while process tracing illuminates the *mechanism* through which any effects are produced.

In these approaches, each type of evidence answers a different kind of question (see also Collier and Sambanis (2005) and Fearon and Laitin (2008)). These works thus do not identify ways in which qualitative and quantitative data can be adjoined to answer the same question, such as whether one variable has a causal effect on another. Indeed, some accounts, such as that of Goertz and Mahoney (2012), lay emphasis on how very different qualitative and quantitative analyses are—in their premises, goals, and procedures—and on the challenges of integrating inferences across these approaches.

2. Common questions, diverse data, distinct analyses. Many works have argued that qualitative and quantitative approaches can be understood as distinct methods for addressing the same basic questions: the testing of causal theories or claims (Collier, Brady, and Seawright, 2004; Hall, 2003; Gerring, 2012; Lieberman, 2005; Seawright and Gerring, 2008; Coppedge, 2012). While the seminal book by King, Keohane, and Verba (1994) describes differences in the forms of data and analytic tools employed by small- and large-n scholars, these authors understand both approaches as contributing to causal inference and testing for causal effects through the examination of a theory's observable implications. Works such as Collier and Brady (2010) and Hall (2003) have drawn a sharper distinction between the types of observable implications tested in the two approaches: i.e., between the prediction of certain associations between independent and dependent variables *across* cases and predictions relating to observable features of processes unfolding *within* cases. These authors nevertheless view both approaches as feeding into the same causal-theory-testing endeavor.

While this "common questions" orientation has pointed to the possibility of integration, works in this tradition have to date not indicated how such integration should occur. In particular, these works have not provided specific principles or procedures for aggregating inferences—whether mutually reinforcing or contradictory—across different forms of data and distinct modes of analysis.

3. Common questions, single form of data, integrated analysis. A small set of studies—all article-length treatments—have outlined methods of analytically integrating qualitative and quantitative inferences. One approach to integration has involved the use of a single form of data and an analysis of that data that takes explicit account of the ordinal or nominal nature of those data (Barton and Lazarsfeld, 1955; Young, 1981). Glynn and Ichino (2013), for instance, outline a framework in which the researcher draws on qualitative information from in-depth case studies to generate ordinal rankings of cases on particular variables that, in turn, inform the statistical estimation of causal effects.

4. Common questions, diverse data, integrated analyses. Finally, a very small set of studies specifies procedures for aggregating *diverse*—quantitative and qualitative—forms of data into a single inference or set of inferences. In the approach suggested by Gordon and Smith (2004), available expert (possibly imperfect) knowledge regarding the operative causal mechanisms for a small number of cases can be marshalled to anchor the statistical estimation procedure in a large-N study. Relatedly, in Glynn and Quinn (2011), researchers use knowledge about the empirical joint distribution of the treatment variable, the outcome variable, and some post-treatment variable, alongside assumptions on how causal processes operate, in order to tighten bounds on causal effects. Critically, these contributions focus on inferences using both *cross-case* correlations and *within-case* evidence bearing on a causal claim.[1]

Integrated Inferences is most similar in spirit to this fourth group of journal publications. The book's contribution shares with Glynn and Quinn (2011) a focus on combining inferences from correlational and within-case data.

The approach presented in *Integrated Inferences* is a much more encompassing and flexible framework, yielding a broader set of analytical and methodological payoffs. Rather than focusing on the bounds on causal effects, the BIQQ framework can generate a wide range of estimates of substantive and methodological interest, including average causal effects for a population, case-level causal effects, and the validity of rival theories. Further, the approach presented here allows the analyst to update the premises on which the interpretation of evidence is based, systematizing the cumulation of knowledge that can derive from triangulation across diverse forms of evidence. Finally, unlike any approach that we are aware of, the BIQQ framework allows us to derive advice about the mix of qualitative and quantitative data that is likely yield the optimal inferential payoffs under particular research conditions.

In short, *Integrated Inferences* unique contribution lies in its elaboration of a unified framework for reasoning through a wide range of analytical and research-design tasks confronting multi-method researchers.

4 Audience

We expect *Integrated Inferences* to have a broad audience in the social sciences and among policy practitioners. The approach presented is a general one that can be applied to the testing of causal claims in a vast range of domains. The book will be oriented toward teaching as well as elaborating the BIQQ framework. To maximize accessibility, the exposition will pair formal derivations of the methods with intuitive accounts of their basic logic, and will illustrate the use of the BIQQ approach through applications both to simulated data and to the analyses of mixed

[1] See also White and Philips (2012).

datasets addressing a range of social-scientific questions, such as the causes of civil war to the effectiveness of development aid to the design of electoral systems. Chapter appendices will provide the reader with step-by-step instructions for working through the simulations and applications using the accompanying software.

We foresee three main audiences for the book:

- **Faculty in the social and policy sciences**. With the growing use of mixed-method designs, empirically oriented social scientists from political science, sociology, public policy, and economics will take interest in an approach that allows for a principled integration of quantitative and qualitative findings.

- **Graduate students in the social sciences**. *Integrated Inferences* is expected to be widely adopted in graduate courses in research design, qualitative methods, and quantitative methods throughout the social sciences. Given growing appreciation of the power multi-method designs in the policy world, we expect the book also to be assigned in policy evaluation courses in MPP and MPA programs.

- **Policy practitioners**. Beyond academia there is great interest in using mixed methods research to address policy questions. Indeed, policy practitioners often face precisely the kind of problem that the book's framework is intended to address. Making evidence-based decisions frequently requires practitioners to draw on seemingly incompatible research findings generated from seemingly incompatible research strategies. *Integrated Inferences* will provide tools to guide such practitioners both in the commissioning of research and in their learning from disparate research findings.

5 Timeline

We expect to complete the manuscript over the next three years. We have developed the core BIQQ framework, with applications to real data, in a paper initially presented at the Institute for Qualitative Research Methods Authors' Workshop and at the 2013 Annual Meeting of the American Political Science Association. This paper won the Sage Prize of the APSA's Organized Section on Qualitative and Multi-Method Research, given to the best paper on qualitative methods presented at each year's Annual Meeting. The paper is currently under review at a major political science journal. This paper contains core elements of Chapters 2-7 and 10, described below.

We currently have two further papers at the planning stage, roughly corresponding to chapters 11 and 13. We expect to work on these papers over the next two years. We plan to complete the remainder of the manuscript during a joint research leave in the third year.

6 Annotated Table of Contents

Integrated Inferences

Part I: Knowledge

1. Introduction [10 pages]

We begin by characterizing the recent growth of mixed-method research designs, describing prominent examples and highlighting the ways in which they seek to integrate inferences from qualitative and quantitative analyses. We then provide our working definition of qualitative and quantitative methods of causal inference. By quantitative methods, we refer to approaches that draw causal inferences from the analysis of correlations between independent (X) and dependent (Y) variables (i.e., "dataset observations" in the terminology of Collier, Brady, and Seawright (2004)). Qualitative methods are those that seek to draw inferences about whether X affects Y by exploiting information from the causal process connecting X and Y or from the context in which that process is thought to unfold (i.e., "causal process observations" (Collier, Brady, and Seawright, 2004)). The chapter also reviews and identifies limitations of current treatments of mixed-method research, summarizes the intuition behind the BIQQ framework, and provides a roadmap for the book ahead.

2. Causal Effects and Causal Accounts [10 pages]

Here we identify the problems of causal inference that the BIQQ framework addresses. We introduce the now standard "potential outcomes" framework in which the BIQQ framework is embedded and the associated fundamental problem of causal inference. We highlight the metaphysical nature of causal claims: a causal claim is not a statement about observable processes only, but contains statements about *counterfactuals*. Thus causal claims requires identification strategies, not just measurement strategies; we argue that this is as true for qualitative inference as for quantitative inference.

We differentiate among several different causal questions that are commonly of interest to social scientists and that BIQQ can help answer. These include a search for case-specific causal effects/explanations, population-level average causal effects; the distribution of case-specific causal effects in a population; and mechanisms of causation.

We also seek to relate the problem of identifying causal effects to the more general problem of explanation in social science. To do so we introduce ideas developed in the study of causal mediation and in new work on Bayesian networks as well as a new notion of "meta-estimands," or quantities that characterize the domains across which causal findings can travel.

3. Being Bayesian [6 pages]

Chapter 3 is a short chapter that serves as a primer on Bayes rule and Bayesian inference. We distinguish among the different types of prior beliefs entering into Bayesian inference, including prior levels of confidence in the test hypothesis and beliefs about the likelihoods with which evidence will be observed under alternative states of the world. We also discuss the classic Duhem-Quine problem, which describes the difficulty of distinguishing between a false theory and a false set of background assumptions underlying an empirical test.

4. Bayesian Correlational Inference [6 pages]

This short chapter begins to build up the BIQQ framework by setting up quantitative causal inference within a Bayesian framework. Within a simplified binary-variable framework, we show how Bayes rule is used to draw causal inferences from correlations between independent and dependent variables across cases. We emphasize here the importance, for correlational inference, of background assumptions about assignment: the probabilities with which cases are assigned to values on the independent variable. We also discuss the role of priors on sampling in drawing inferences from samples to populations. We illustrate Bayesian correlational inference using a simple substantive example. Like chapter 3, this chapter contains no original analysis but is needed to ensure accessibility for readers unfamiliar with Bayesian inference.

5. Bayesian Process Tracing [20 pages]

In this more extended chapter, we lay another cornerstone of the BIQQ framework by formalizing the logic of process tracing in Bayesian terms (building on Bennett [2010]). We introduce here the concept of a "clue," which is central to our understanding of qualitative research. A clue is a within-case observation of process or context that the researcher believes to be associated with alternative causal effects or mechanisms with particular probabilities. Just as assignment probabilities serve as critical background assumptions for correlational inference, "clue probabilities"—the probabilities linking clues to causal effects or mechanisms—constitute the key premises of process tracing. We also introduce here the key concept of a clue's probative value, which depends on the difference in probabilities with which we expect the clue to be observed under alternative hypotheses. The chapter discusses types of clues that researchers commonly use to draw causal inferences in within-case studies and how different combinations of clue probabilities generate different types of empirical tests (cf. Van Evera 1997). A key theme we highlight is that beliefs about clues *contain* beliefs about causal effects, thus defending claims regarding clues requires an *identification strategy.*

The chapter also unpacks the logic of searching for multiple clues within a case. We hope to develop a new method for ranking, or partially ranking, multiple clues, as a function of the correlation structure between them. We sketch here the approach we are currently developing for ordering clues.

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8
$K_1 = 1$ ("Movement")	F	T	F	T	F	T	F	T
$K_2 = 1$ ("Pulse")	F	F	T	T	F	F	T	T
$j \in B$ ("Alive")	T	T	T	T	F	F	F	F
Prior ($\Pr(\omega_j)$)	$\frac{1}{6}$	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$	0	0	0

Table 1: Two clues, K_1 and K_2 available for assessing the type of a subject (here "alive"); both have probative value, but K_1 has no probative value conditional on the observation of K_2.

Table 1 provides an example of a situtation in which there are two possible clues that might indicate whether a person is alive: movement and pulse. Gathering data on pulse may be more costly as it requires greater proximity. If the person is dead, neither clue will be observed (when sought). If the person is living it is still possible that neither will be observed; however, although it is possible that a pulse will be found, even if there is no movement, it is not possible that there will be movement if there is no pulse. Clearly clue K_2 has greater probative value than clue K_1 here[2] however in this case the probative value of clue 1 is not just weaker but is *dominated* by the probative value of clue 2 in the sense that clue 1 has no probative value, *conditional* on the presence or absence of clue 2, but not *vice versa.*[3]

In principle also, the probative value of a clue might increase with the examination of other clues, which highlights the importance of generating a notion of clue complementarity as well as substitution and dominance.

[2] Since $|Pr(K_2|j \in B) - Pr(K_2|j \notin B)| > |Pr(K_1|j \in B) - Pr(K_1|j \notin B)|$

[3] Specifically if no pulse is observed then there is no point looking for movement (the probability of observing movement is 0 whether the person is alive or dead); but if no movement is found there is still value in checking for a pulse (the *conditional* probability of finding a pulse is .5 if the person is living; but 0 if they are dead).More formally, let $\phi_B(K|D)$ denote the probative value os clue K, given data D for proposition B; that is: $\phi_B(K|D) = |\Pr(K|D\&B) - \Pr(K|D\&\neg B)|$. Then for this data we have: $\phi_B(K_1) = \frac{1}{3}; \phi_B(K_1|K_2 = T) = 0; \phi_B(K_1|K_2 = F) = 0$; thus the probative value of K_1 is extinguished by K_2. In contrast, $\phi_B(K_2) = \frac{2}{3}; \phi_B(K_2|K_1 = T) = 0; \phi_B(K_2|K_1 = F) = \frac{1}{2}$; thus te probative value of K_2 is reduced, but not extinguished by K_1.

Part II: BIQQ

6. Integration [20 pages]

Here we develop and illustrate the integrated BIQQ framework, in which the researcher can employ prior beliefs about assignment and clue probabilities to derive causal inferences from any combination of correlational (quantitative) and clue (qualitative) data.

The chapter will focus on using BIQQ to estimate population-level average causal effects and will likely contain three parts. In the first we provide the broad intuition, working through a simple example and employing point priors. In the second part, we introduce the general approach, with priors expressed as distributions, and demonstrate the simulation procedures used to generate posterior distributions on the quantities of interest. A third section will explore extensions, including the use of controls, continuous variables, interactions, and systems of relations.

7. Illustrations [20 pages]

In chapter 7 we provide substantive applications of the BIQQ framework using mixes of quantitative and qualitative data speaking to major issues in the social sciences. These include the causes of civil wars, the origins of electoral systems, and the effectiveness of development aid. We continue to use these examples in later chapters introducing further uses and implications of the BIQQ approach.

8. Uses [20 pages]

While Chapter 6 focuses on the use of BIQQ to generate inferences on population-level causal effects, here we show how the approach can be used to address other estimands: *case-level* causal explanations, estimates of the distribution of case-level causal effects in a population, and the validity of *alternative theories* of mechanism. Moreover, we show how the use of BIQQ yields updated beliefs about assignment and clue probabilities, i.e., about the background assumptions underlying quantitative and qualitative causal inference. These uses are illustrated through simulated extensions of the substantive applications introduced in Chapter 7.

9. Robustness [8 pages]

In BIQQ, causal inferences can be sensitive to entering assumptions about clue and assignment probabilities. This chapter discusses, and illustrates with substantive examples, how researchers can estimate and report the robustness of their conclusions to different priors on assignment and on the probative value of clues.

We hope in this chapter to introduce a *robustness* metric that expresses the degree to which conclusions depend on priors.

PART III: Implications for Design

10. Wide or Deep [15 pages]

In Chapter 10, we begin to show how BIQQ can be used to generate implications for research design. The present chapter focuses on implications for the researcher's choice, given finite resources, between alternative mixes of extensive, quantitative and intensive, qualitative data. We show how the BIQQ method can be used to estimate the mix of qualitative and quantitative data that can be expected to generate the most accurate causal inferences, conditional on the analyst's beliefs about the research situation. We demonstrate how the framework identifies differing optimal mixes of correlational and clue observations depending on, for instance, the level of causal heterogeneity in the population, the degree of uncertainty about the assignment process, and the degree of uncertainty about the probative value of clues. The chapter also unpacks the tradeoff that arises in purely qualitative research between doing deeper process tracing (collecting more clues) within a given set of cases and doing process tracing of a given depth in a greater number of cases.

11. Case Selection [20 pages]

Chapter 11 explores BIQQs capacity to yield guidance on the problem of case selection for process tracing. The chapter derives several lines of reasoning about case selection that are distinctive from advice in the current literature. The key contribution is a treatment of how case selection ought to turn on the probative value of available clues. To illustrate: frequently, analysts choose cases for process tracing based on the values of X and/or Y, with a common focus on $X = Y = 1$ cases. Whether this strategy is efficient, however, depends on the probabilities linking causal effects (types) to available clues. For instance, it could be the case that the probability of observing the available clues in treated cases is the same for types experiencing a causal effect (B types) and types that would have generated the outcome independent of treatment (D types). In this case, the clues have no probative value within $X = Y = 1$ cases and little can be learned. On the other hand, it could be the case for the available clues that the probabilities of observing them in untreated cases that would have been affected by treatment (B types) are very different from the probability of observing them in untreated cases that would not have experienced the outcome independent of treatment (C types). Under these evidentiary conditions, $X = Y = 0$ cases will in fact be more informative for assessing causal effects. Similar reasoning extends to identifying the optimal *mix* of cases, given a diverse set of clues and clue probabilities.

In addition, in this chapter we seek to address the distinctive contribution that off-the-regression line cases can make to theory-testing; the dependence of case selection on the estimand of interest; procedures for selecting on mediator variables; and a generalization and Bayesian formalization of the most-/least-likely selection strategy.

12. How Do We Learn? [15 pages]

Causal inference depends on a set of background beliefs about the process being analyzed. In BIQQ, the analyst begins with beliefs about causal effects, beliefs about assignment probabilities and beliefs about the probabilities tying qualitative clues to causal effects. It is these latter beliefs that are more critical — conclusions can be robust to errors in the first set of beliefs, and the second set of beliefs can sometimes be under the control of researchers, or otherwise known. Bayesian correlational inference can proceed with flat priors about assignment probabilities because the X, Y data can allow updating over these probabilities. However, causal inference in process tracing cannot proceed with wholly uninformative priors over the probabilities with which clues are associated with causal effects.

This final chapter asks how we can form the beliefs required to draw causal inferences from qualitative data. How do we arrive at knowledge about the probative value of specific types of case-study evidence for drawing causal inferences?

This is an extremely difficult problem, one with which the qualitative methods literature has not wrestled to a great extent. We tentatively explore here the logic, prospects, and limits of three possible foundations for knowledge about clue probabilities: causal theory; crowd-sourcing of subjective probabilities; and a new empirical approach that derives bounds on clue probabilities from X, Y, and clue observations across multiple populations within a specified domain, given assumptions on the invariance of clue probabilities across populations in the domain.

Conclusion [10 pages]

The conclusion summarizes our contribution and points to what we see as the major challenges and open questions in the integration of inferences. We focus especially on the question of sequencing in a research program. Sequencing matters in the framework we examine because optimal research design choices depend on beliefs—e.g., about the probative value of clues or about the heterogeneity of causal effects—that will themselves typically be informed by prior data. We also touch here on a deep problem in the linkage between theory and evidence. In idealized settings, theory is assessed on the basis of evidence that is generated independently of the data that initially gave rise to the theory. In many real-world research situations, however, there are (often unacknowledged) dependencies between the data used for testing and the data used for theory-generation. Such is generally the case, for instance, in studies employing historical data, where any "new" data collected will tend to be correlated with observations of the world that inspired the theory. In retrospective empirical work, it is not obvious how to determine when findings constitute new evidence or "fished" results. We describe in this concluding chapter how a Bayesian framework can allow for an assessment of what is learned from new data on old events.

Appendices [20 pages]

Appendices (which might be online or included in the text) will provide **R** code for the key analytic procedures described in the text.

Page count: 200 pages

References

Barton, Allen H, and Paul Felix Lazarsfeld. 1955. *Some functions of qualitative analysis in social research*. Number 181 Bobbs Merrill.

Beck, Nathaniel. 2010. "Causal Process "Observation": Oxymoron or (Fine) Old Wine." *Political Analysis* 18 (4): 499–505.

Bennett, Andrew. 2008. "Process Tracing. A Bayesian Perspective." In *The Oxford Handbook of Political Methodology*, ed. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier. Oxford, UK: Oxford University Press.

Brady, H.E., and D. Collier. 2010. *Rethinking Social Inquiriy: Diverse Tools, Shared Standards*. Rowman & Littlefield Pub Incorporated.

Collier, David, Henry E Brady, and Jason Seawright. 2004. "Sources of leverage in causal inference: Toward an alternative view of methodology." In *Rethinking social inquiry: Diverse tools, shared standards*, ed. David Collier and Henry E Brady. Lanham MD: Rowman and Littlefield.

Collier, David, Henry E Brady, and Jason Seawright. 2010. "Sources of leverage in causal inference: Toward an alternative view of methodology." In *Rethinking social inquiry: Diverse tools, shared standards*, ed. David Collier and Henry E Brady. Lanham MD: Rowman and Littlefield.

Collier, P., and N. Sambanis. 2005. *Understanding Civil War: Africa*. Stand Alone Series World Bank.

Coppedge, M. 2012. *Democratization and Research Methods*. Strategies for Social Inquiry Cambridge University Press.

Fearon, James, and David Laitin. 2008. "Integrating qualitative and quantitative methods.". Oxford Handbook of Political Methodology Cambridge, UK: Oxford University Press.

George, A.L., and A. Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. A BCSIA book MIT Press.

Gerring, John. 2012. *Social Science Methodology: A Unified Framework*. Strategies for Social Inquiry Cambridge University Press.

Glynn, Adam N., and Kevin M. Quinn. 2011. "Why Process Matters for Causal Inference." *Political Analysis* 19: 273-286.

Glynn, A.N., and N. Ichino. 2013. "Using Qualitative Information to Improve Causal Inference.".

Goertz, G., and J. Mahoney. 2012. *Tale of Two Cultures - Contrasting Qualitative and Quantitative*. University Press Group Limited.

Gordon, Sanford C, and Alastair Smith. 2004. "Quantitative leverage through qualitative knowledge: Augmenting the statistical analysis of complex causes." *Political Analysis* 12 (3): 233–255.

Hall, Peter A. 2003. "Aligning Ontology and Methodology in Comparative Research." In *Comparative Historical Analysis in the Social Sciences*, ed. James Mahoney and Dietrich Rueschemeyer. Cambridge, UK and New York: Cambridge University Press.

King, G., R.O. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press.

Lieberman, Evan S. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* 99 (7): 435–452.

Paluck, Elizabeth Levy. 2010. "The Promising Integration of Qualitative Methods and Field Experiments." *The ANNALS of the American Academy of Political and Social Science* 628 (1): 59-71.

Seawright, Jason, and John Gerring. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61 (2): 294-308.

White, Howard, and Daniel Philips. 2012. "Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework." *International Initiative for Impact Evaluation Working Papers* 15 (6).

Young, Forrest W. 1981. "Quantitative analysis of qualitative data." *Psychometrika* 46 (4): 357–388.