Improving Designs in the Social Sciences

A Book Proposal to Princeton University Press

Graeme Blair   Jasper Cooper   Alexander Coppock   Macartan Humphreys

October 18, 2017

Advances in causal analysis from the graphical models and potential outcomes frameworks have revolutionized empirical social science research. In recent years focus has shifted to more careful definition of causal quantities of interest and a more direct focus on study designs to justify inferences. .

Experimental research in particular has experienced enormous growth in the past two decades. For many years, whole disciplines were viewed as non-experimental sciences; this is no longer the case. Likewise, observational quantitative research has shifted from wholesale modeling of outcomes on the basis of large sets of predictors to explicitly seeking to approximate randomized experiments through the use of matching, regression discontinuity, and instrumental variable designs.

The insights from the "identification revolution" have also illuminated descriptive research practices – not least by clearly delineating how causal questions differ from descriptive questions. More surprisingly perhaps there has been a growth in the use of experimental techniques as measurement devices.

Qualitative research, too, has seen the influence of ideas developed in the causal analysis literature. When qualitative researchers set out to answer a causal question, they typically make use of extremely detailed information about a small number of cases in order to make inferences about what would have happened in each case, had conditions been different. Put in the language of potential outcomes, they "impute the missing potential outcomes."

With our proposed textbook, *Improving Designs in the Social Sciences*, we aim to present a comprehensive view of modern research design. We define a research design as an object with four properties: a Model, an Inquiry, a Data strategy and an Answer strategy. The MIDA framework is flexible enough to accommodate most every empirical research design, including experimental and observational, causal and descriptive, qualitative and quantitative. It is not so flexible, however, as to be without theoretical or analytic bite. Understanding which features of a research design belong to M, I, D, and A often clarifies exactly which design tradeoffs are most important to consider.

Viewing research through the MIDA lens allows us to "declare" research designs and "diagnose" their properties. Declaration means enumerating the theories and procedures

1

that make up the content of a design. Diagnosis entails imagining how well a design accomplishes its goal. We imagine declaration and diagnosis as formal procedures that typically are executed in computer code. To that end, we have written a software package `DeclareDesign` for the popular language `R` that focuses a user on the important parts of improving research designs while handling the less important programming tasks. We hasten to add, however, that the software is only a means to an end. We don't care *how* students and researchers declare and diagnose their designs so long as they do at all.

We envision *Improving Designs in the Social Sciences* in two large parts. The first part will explain the MIDA approach and our unified view of social science research. The second part will declare and diagnose approximately 100 research designs. The discussion of each design will focus on the specific design challenges researchers face.

1 Declaring and Diagnosing Research Designs

Authors and readers of empirical research routinely need to assess the properties of research designs. In doing so, they face two challenges.

First, few tools exist for the comprehensive assessment of the properties of designs. At one extreme, researchers resort to rudimentary power calculators with sometimes hidden assumptions or rely on rules of thumb that may not incorporate important idiosyncratic features of the research setting. At the other extreme, some scholars conduct fully-fledged simulations requiring advanced programming skills beyond the capabilities of many applied researchers. General-use tools for assessing important properties of designs such as statistical power or bias are not available.

Second, surprisingly little attention has been paid to the more fundamental question of what constitutes a design. This lack of clarity carries risks both before and after the implementation of a study. If designs are incompletely specified ex ante, it is difficult for researchers to assess their strengths and improve them. If designs are incompletely specified at the time of analysis, researchers may choose inappropriate analysis procedures or worse, may only report the most attractive model from a set of possible specifications. If designs remain unspecified after analysis, it may be difficult for other scholars to replicate a study or to judge whether a given type of reanalysis is justified.

DDRD describes an approach that addresses these two problems. It introduces a framework, building on a structure outline in King, Keohane, and Verba (1994) to describe the distinct elements of a research design. Then using the framework, it provides a library of common designs, fully characterized. Finally, using the library, DDRD interrogates these designs to elucidate 50 general principles of research design.

1.1 MIDA

The framework we develop conceptualizes a design as including four elements $< M, I, D, A >$:

1. A **model**, M, of how the world works. The model specifies the moving parts – the variables – and how these are causally related to each other. In this sense the model provides the context of a study but also a speculation about the world.

2. An **inquiry**, I, about the distribution of variables, perhaps given interventions on some variables. In many applications I might be thought of as the "estimand." Some inquiries are statements about the values of variables, others about the causal relations between variables. In all cases however the inquiry should be answerable given the model.

3. A **data** strategy, D, generates data on variables. Note that implicitly the data strategy includes case selection, or sampling decisions, but it also represents interventions such as assignment of treatments or measurement strategies. A model M tells you what sort of data you might observe if you employ data strategy D.

4. An **answer** strategy, A, that uses data to generate an answer.

A key feature of this bare specification is that if M, D, and A are sufficiently well described, the answer to question I has a well defined distribution. Moreover moreover one can construct a distribution of comparisons of the answer provided by the answer strategy to the correct answer, under M.

The ability to calculate distributions of answers, given a model, opens multiple avenues for assessment and critique of research strategies. How good is the answer you expect to get from this strategy? Would you do better with a different data strategy? With a different analysis strategy? How good is the strategy if the model is wrong in some way or another?

We call this *design diagnosis*. The key idea of design diagnosis is that if one has a fully specified design—a 'complete' design, in a sense—one can predict what will happen when a design is implemented, at least conditional on the assumptions of the design. From there it is an easy step to assess what would result from multiple implementations of a design and so build up a distribution of quantities of interest—what we term "diagnosands." Researches are familiar with the idea of statistical power—the probability that a design will produce a p value below some specified cutoff. In fact that idea can be extended to many many properties of a design so that, once a design is specified, repeated iteration of the design lets one assess the bias and precision of estimates, and the relative performance of one design relative to another in a given setting.

2 A Computational Approach

3 Contents

The book is organized in three sections as follows:

Contents

[250 pages total, approx; maybe 500]

3.1 Designs, diagnosed

3.2 Principles

4 Relation to existing works

Major research design books:

In short, the "new" causal analysis has changed how social scientists conduct research. However, textbook treatments of research design have not kept apace. The classic test *Designing Social Inquiry* by King, Keohane, and Verba (KKV) considered research design as a general topic with a unifying framework for quantitative and qualitative reasoning. KKV, however, was published in 1994, and does not include many of the new connections across types of research designs generated by the causal identification literature.

King Keohane and Verba

John Gerring. Social Science Methodology: A Unified Framework. Cambridge University Press, New York, second edition, 2012.

Barbara Geddes. Paradigm and Sand Castles: Theory Building and Research Design in Comparative Politics. University of Michigan Press, 2003.

Brady et al. Rethinking Social Inquiry: Diverse Tools, Shared Standards

Jason Seawright. Multi-Method Social Science: Combining Qualitative and Quantitative Tools. Cambridge University Press, 2016.

Dimiter Toshkov. Research Design in Political Science. Palgrave Macmillan, 2016.

Kellstedt, P. M. &?Whitten, G. D. 2013. The Fundamentals of Political Science Research. Cambridge University Press.

Check out this brand NEW/4th edition of Theory Methods in Political Science by Lowndes, Marsh Stoker (eds) methodological pluralism

By topic:

Experiments:
Gerber and Green
Glennester RRE
World Bank ?Impact Evaluation in Practice? by Gertler et al.

Observational:
William R. Shadish, Thomas D. Cook, and Donald T. Campbell. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Houghton-Mifflin, Boston, MA, 2001.

Causal Inference in Statistics: A Primer
Observation and Experiment: An Introduction to Causal Inference
Quasi-Experimentation: Design & Analysis Issues for Field Settings
Causality: Models, Reasoning and Inference
Observational Studies (Springer Series in Statistics)
Design of Observational Studies (Springer Series in Statistics)
Dunning natural experiments

Qualitative:
Bennett and Checkel
George and Bennett
Van Evera 1997

Data science:
Computer Age Statistical Inference: Algorithms, Evidence, and Data Science (Institute of Mathematical Statistics Monographs)

The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics)

An Introduction to Statistical Learning: with Applications in R (Springer Texts in Statistics)

5 Audience

1. General research design courses in masters and Ph.D. courses in political science, public policy, sociology, and business schools in the U.S. and Europe

2. Experiments and causal inference quantitative methods courses in Ph.D. courses in political science, public policy, sociology

3. Summer courses in experimental design and advanced quantitative methods in the social sciences (i.e. XX YY ZZ)

4. Advanced experiments and quantitative methods courses at top undergraduate institutions

5. Uses as modular examples in intro Ph.D. and M.A. substantive courses that often copy parts of texts to explain methodological elements

6. Evaluation practitioners – impact evaluation and measurement and evaluation (M&E) specialists in government, international lending institutions such as the World Bank, bilateral aid agencies, major non-profit aid organizations such as XX YY ZZ, major donors such as the Hewlett Foundation, and political consulting organizations such as the Analyst Institute

6 Timeline

We expect to complete the book over the next 18 months, which corresponds to a grant that is funding research and programming assistance from the Laura and John Arnold Foundation.

By October 2018, we will complete a draft of the book material and hold a book conference (funded by the Arnold Foundation) shortly thereafter. We will then complete the remainder of the manuscript by April 2019.