

*Counterfactual and Causal Inference:  
Methods and Principles for Social Science Research*

by Stephen L. Morgan and Christopher Winship  
2nd edn, New York: Cambridge University Press, 2015  
ISBN 978-1-107-69416-3  
Softcover, \$43.95, 504 pp.

Reviewed by Peter Messeri  
Mailman School of Public Health. Columbia University (New York)

Stephen L. Morgan and Christopher Winship's 2007 book *Counterfactual and Causal Inference*—along with Shadish, Cook, and Campbell's *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* and Angrist and Pischke's *Almost Harmless Econometrics*—take pride of place in my personal quantitative methods library. A common theme uniting these texts is the embrace of the counterfactual approach to causal inference and analysis.

The attraction of the counterfactual framework for causal inference is that it presents a mathematically elegant definition of causal effects that dispenses with the spatial-temporal constant conjunction definition of causality. The counterfactual framework imagines that individuals may occupy multiple causal states and each has multiple potential outcomes, one for each causal state. The causal effect for an individual is then defined as the difference in his or her potential outcomes for two causal states. Since an individual can occupy only one causal state at a time, only one potential outcome can be observed as the actual realization of the outcome. The potential outcomes for the remaining causal states are the unobserved counterfactuals. Although this definition rules out identification of causal effects for individuals, the simple but rigorous mathematical definition of causal effects, and the clearly articulated identification problem, have proven to be extraordinarily fruitful in motivating the development of new lines of research into innovative methods for estimating various population-averaged causal effects.

In 2007, Morgan and Winship published the first comprehensive textbook on the counterfactual approach. It is therefore of great interest to discover what new material the authors added to the second edition of *Counterfactual and Causal Inference*, published eight years later. Befitting a field of research with an actively growing literature, the second edition page size increased by 29 percent, and the number of pages of text from 290 to 450. The authors have done their due diligence in updating the book's references. Indeed, I counted 231 entries published between 2007 and 2014 in the book's 49-page bibliography.

Under the counterfactual rubric, Morgan and Winship bring together several strands of research, dating from the 1970s. The most prominent is the counterfactual approach to causal inference and matching estimators developed in statistics, and the parallel development of the potential outcomes approach to counterfactual causal inference and instrumental variable estimators in economics. Finally, there is Judea Pearl's (2009) singular contribution of a graph-theoretic approach to causality.

The greatly expanded second edition goes into greater depth on topics found in the original edition, and adds new material, but it maintains the organization of topics covered in the original edition. A short introduction chronicles the history of causal analysis in the social sciences—the rise and overreach of the “Age of Regression” analysis and the subsequent emergence of the counterfactual/potential outcomes approach. The book’s second section presents basic concepts and definition of counterfactuals, potential outcomes, and average causal effects. It continues with a comprehensive introduction to Judea Pearl’s directed acyclic graphs, or DAGs, which the authors wisely choose to refer to, less mysteriously, as *causal graphs*.

The bulk of the remaining book is divided into two sections that focus on methods for estimating causal effects. The first presents matching algorithms and regression methods for estimating causal effects when conditioning on observable variables is sufficient to identify the causal effect of interest. This section incorporates Pearl’s “back-door” criteria for determining when causal effects are identified by observed variables. The next major section covers identification of causal effects when conditioning on observables is ineffective. It begins with a new chapter on potential outcome models, and causal graphs incorporating latent variables for selection and effect heterogeneity, followed by an expanded treatment of instrumental variable methods and natural experiments. The section then turns to Pearl’s front-door criteria for identification of causal effects that specifies an exhaustive and isolated set of intervening causal mechanisms. It concludes with the potential outcomes approach to time series analysis, panel data methods, selection bias models, and sensitivity analysis.

The book concludes with major objections to the counterfactual approach, and the authors’ rejoinders and a brief account of how the counterfactual approach fits within a more general mode of causal inquiry in the social sciences.

The mathematics required to grasp the book’s content are covered in a first semester applied probability and statistics course, as well as familiarity with principles of regression analysis and rigorous program evaluation. Derivation of mathematical formulation and descriptions of the construction of causal graphs are followed by lengthy textual exposition intended to unpack the mathematical formulations and explain graphic representations. Despite Morgan and Winship’s straightforward exposition, the more complex arguments are not always easy to follow and require careful reading in order to pick up the subtleties packed into the text. Many “demonstrations” are scattered throughout the text that present simulations or substantive examples which concretely illustrate the more abstract principles.

The second edition’s most substantial enhancement is the increased prominence given to Pearl’s contribution to the causal inference literature. A central theme of the second edition is to confer co-equal and complementary status on causal graphs and potential outcomes approaches for identifying causal effects and formulating more cogent causal models. Every counterfactual application covered in the text has both a potential outcomes mathematical derivation and causal graph representation. The expanded coverage given to causal graphs makes the second edition a much better how-to manual than the first in constructing causal graphs and their application for diagnosing and solving identification problems.

Other new and expanded topics to be found in the second edition include elaboration of effective methods of regression analysis in the estimation of average treatment effects, a chapter on causal graphs and potential outcome perspectives applied to selection and unmeasured effect heterogeneity, interpretation of instrumental variables in the presence of individual-level heterogeneity, recent developments in estimating local average treatment effects, the virtues of intervening mechanisms for deep explanations, and use of multiple pre-treatment observations to model causal effects.

My reservations are relatively few and mostly address what is missing in the book. Although the first chapter summarizes six studies that we are told are used as examples throughout the book, concrete illustrations are mainly drawn from James Coleman's study of the effect of Catholic school education on learning, and an instrumental variable study of school vouchers and learning. There are good reasons for the heavy reliance on the Coleman study, as he was the acknowledged master in empirical tests of carefully specified theoretical models. The large critical secondary literature and Coleman's rejoinders provide a rich empirical reservoir for concrete illustrations of the plausibility of common model assumptions and causal specifications. However, an understanding of the common and unique challenges confronting the counterfactual approach across disciplines would be better served by more illustrations drawn from the life course, public health, job training, and voting studies summarized in the introduction.

Despite the detailed presentation of casual graphs, I would have liked more causal graphs from actual rather than hypothetical studies. This is particularly true for the exposition of a *collider*, a seemingly innocuous variable that is determined by two or more causally antecedent variables. With the collider, they form part of a blocked back-door pathway. When unrelated determinants—the independent variables, if you will—are conditioned on the collider, the dependent variable, the seemingly counter-intuitive result, is to induce an association between the independent variables and unblock the previously blocked back-door pathway. An analyst that unwittingly adds a collider under the mistaken belief that it will control for latent variable confounders—for example, adding a lagged outcome to the list of observed independent variables—may have the opposite effect of creating new confounding. Now the authors persuasively demonstrate that Pearl's logic is impeccable for diagnosing a potentially flawed strategy to identify casual effects. However, the logic is not easy to grasp, and more concrete illustrations from actual rather than hypothetical situations would have been a more convincing demonstration of the practical rather than the theoretical importance of the collider.

For clarity of theoretical exposition, the text restricts model specification to binary causal variables and interval or continuous outcomes. Discussion of extension of casual variables with ordinal or multiple states is generally limited, although the authors do refer the interested reader to the relevant literature. There was no discussion of the state of the art, if any, of extensions to outcomes that are constructed from binary, count, or survival data.

The second edition of *Counterfactuals and Causal Inference* should be part of the personal library of any social scientist who is engaged in quantitative research. For those with a copy of the first edition, purchase of the second edition is indeed well worth the investment. I would hold on to your copies of the first edition. In my opinion, the second edition entails a more rigorous treatment of similar material covered in the original text; for good reason, it is now an intermediate textbook suitable for advanced graduate seminars on causal analysis, whereas the original edition remains a suitable textbook for an undergraduate or a first-year graduate quantitative methods course that requires a nontechnical introduction to current concepts in causal analysis.

### Additional References

- Angrist, J.D., and Joro-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- Pearl, J. 2009. *Causality: Models, Reasoning and Inference*. 2nd ed. Cambridge (UK): Cambridge University Press.
- Shadish, W.R., T.D. Cook, and D.T. Campbell. 2001. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.