

## From correlation to causation

*Mostly Harmless Econometrics: An Empiricist's Companion.* By Joshua D. Angrist and Jorn-Steffan Pischke, Princeton, NJ, Princeton University Press, 2008, 392 pp., \$35.00/paperback.

“How do you do more than just find statistical relationships?” from the point of view of causality. People such as Steven D. Levitt, co-author of the best-selling books *Freakonomics* and *Superfreakonomics*, have made very nice careers out of this new econometrics, not just academically but also in the popular press. Another such book, *Mostly Harmless Econometrics* is written for people who use econometrics, and who would consider changing their research agenda from finding correlation to showing causation. More specifically, the book is designed for people in the social sciences who may have difficulty setting up their research with the experimental ideal of randomized trials; in other words, social scientists who want to be better empiricists. Before attempting this book, however, readers should be familiar with, among other things, the “core terms and concepts of statistical theory—moments, mathematical expectation, probability limits, and asymptotic distributions.” It is not meant for people brand new to statistics and econometrics.

Joshua D. Angrist is a professor of economics at the Massachusetts Institute of Technology, while Jorn-Steffen Pischke is professor of economics at the London School of Economics and Political Science. Both are labor economists who do applied microeconometrics, so don't expect techniques from other fields. *Mostly Harmless Econometrics* is shorter, chattier, and more humorous

than most other econometrics texts, although the humor unfortunately tends to come only at the beginning of the chapter rather than throughout.

*Mostly Harmless Econometrics* is not designed to replace econometric cookbooks such as those published by Stata Press. Nor is it a general reference such as *Econometric Analysis* by William Greene, or even an in-depth look at one area such as Wooldridge's *Econometric Analysis of Cross Section and Panel Data* or Hamilton's *Time Series Analysis*. There also isn't any model building in either the statistical or economic sense, and big sections of econometrics such as time series analysis are missing. What *Mostly Harmless Econometrics* does do successfully is complement other works in the field of econometrics.

Angrist and Pischke illustrate a wide variety of techniques by referencing the 4-year Tennessee Student Teacher Achievement Ratio (STAR) program, which looked at the importance of class size for educational outcomes. The authors use this example to explain nonstandard error issues and to compare discontinuity design with traditional ordinary least squares. The examples are clear enough that even an inexperienced empirical social scientist or someone in a different field could still say, “Ah, OK I can take that idea and apply it to my own work.” The authors do such a good job of explaining fixed effects and panel data that you can't help but wish they would explain a few other topics as well. If there is a second edition, more examples such as STAR would be great.

Speaking of a second edition, there are a few other areas that could be improved. Sometimes the theorems are not presented as clearly as they could be. Angrist and Pischke explain how the theorems matter, but it would be

nice if they hinted at what to look for. The applied economist doesn't want to wade through a theorem when a sentence or two about which parts are especially important would make the point much clearer. This information could even be added in the form of chapter subtitles: something that lets the reader know what the proof is about in advance. Other theorems need better explanation and better integration with the rest of the text. Otherwise, it's unclear why so many mathematical proofs are in an applied book.

Further, some chapters are very clear while others are very dense. Chapter 8, on nonstandard standard error issues, is the clearest. The theorems in this chapter are explained with just a sentence or two and methods of dealing with nonstandard standard errors are clearly listed. Many STATA users (STATA is a metric program commonly used by academic economists) incorrectly assume that simply typing “, robust” or “cluster” after any regression will solve all problems by presenting robust standard errors. However, if the errors have problems beyond heteroskedasticity, simple solutions won't work. Angrist and Pischke present common standard-error problems that are not simply remedied with “, robust” and discuss ways to manage them.

Taken as a whole, this book is an excellent addition to the applied social scientist's library. *Mostly Harmless Econometrics* complements other books very well and provides much needed support in key areas. I hope for a second edition—not because the first edition is bad, but because it is so good that I want more.

—David Mitchell

Department of Economics and Finance  
University of South Alabama