

## Preface to “Endogeneity in Probit Response Models”

I am deeply saddened by the passing of my friend and colleague, David A. Freedman. I feel fortunate that we completed this manuscript before his death in October, 2008. We spoke about which edits were acceptable and which were not, up until the very end.

Freedman was Professor of Statistics at the University of California, Berkeley. He has about 200 papers in the professional literature and was a member of the American Academy of Arts and Sciences. In 2003, he received the John J. Carty Award for the Advancement of Science from the National Academy of Sciences, recognizing him “for his profound contributions to the theory and practice of statistics, including rigorous foundations for Bayesian inference and trenchant analysis of census adjustment.” He advanced the theory of the bootstrap, martingale inequalities, Markov chains, de Finetti’s theorem, hierarchical Bayes models, the use of regression to analyze experimental data, and other topics.

Although Freedman’s early professional reputation was built on a series of influential papers on probability and mathematical statistics, he thought that such work was more like poetry than science: elegant, aesthetically pleasing, but of dubious relevance for scientific practice.

Above all, he objected to the tendency of researchers to substitute intellectual capital for labor. The technology of statistics, created largely by others, is easy to use and promises to open a wide variety of questions to the research effort. However, the appearance of methodological rigor can be deceptive. The models themselves demand critical scrutiny.

Modeling assumptions are made primarily for mathematical convenience, not for verisimilitude. The assumptions can be true or false—usually false. When the assumptions are true, theorems about the methods hold. When the assumptions are false, the theorems do not apply. How well do the methods behave then? When the assumptions are “just a little wrong,” are the results “just a little wrong”? Our article asks these questions about one specific type of estimator.

Freedman’s full view is presented in a new volume entitled *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*, which is edited by David Collier, Philip B. Stark, and myself (Cambridge University Press, December 2009). In this book, Freedman provides the definitive synthesis of his approach to causal inference in the social sciences. The volume brings together 20 articles by Freedman on statistical modeling and causal inference in social science, public policy, law, and epidemiology. They show when, why, and by how much statistical modeling is likely to fail. They show that assumptions are not a good substitute for research design and relevant data. One has to expend the “shoe leather” to find and then exploit natural variation to mitigate confounding and to exhaust other explanations.

On a more personal note, Freedman was a generous colleague and friend to not only myself but to several generations of Political Scientists. He was surprisingly well read in Political Science, and he sought to influence the discipline. He wanted this article to be sent to *Political Analysis* first, a journal he thought understood the gap between statistical theory and scientific application better than most. Our article presents theorems that are new, but the theorems are not its central point.

**Jasjeet S. Sekhon**