The Stata Journal (2023)
23, Number 3, pp. 897–900 DOI: 10.1177/1536867X231196497

Review of Environmental Econometrics Using Stata, by Christopher F. Baum and Stan Hurn

Ryan P. Thombs
Department of Sociology
Boston College
Chestnut Hill, MA
thombs@bc.edu

Abstract. In this article, I review *Environmental Econometrics Using Stata*, by Christopher F. Baum and Stan Hurn (2021, Stata Press).

Keywords: gn0094, book review, environmental econometrics, quantitative analysis

1 Introduction

It is no secret that the world faces a series of severe environmental problems, ranging from the climate crisis to the sixth mass extinction. These issues are studied across several disciplines in the natural and social sciences and by researchers in government agencies, the private sector, and other research groups. Many researchers use statistics to model environmental data, but the interdisciplinary nature of environmental issues means that the modeling approaches used are not widely shared or integrated together. Christopher F. Baum and Stan Hurn's *Environmental Econometrics Using Stata* does an excellent job with a daunting task, bringing together econometrics techniques commonly used by researchers into one text. *Environmental Econometrics Using Stata* is an invaluable resource to environmental researchers because it offers an easy-to-follow applied introduction to many of the field's most widely used econometric techniques. Baum and Hurn also do a great job throughout the text emphasizing the need to conduct well-done, reproducible research and outlining several diagnostic tests to substantiate results in each chapter.

2 Contents

The book is broadly divided into two parts. The first part comprises chapters 1 to 8 and covers introductory material. The second part of the book (chapters 9 to 15) focuses on particular methods that readers may be interested in. Below, I overview each chapter and provide comments.

The first two chapters are the most basic and fundamental chapters in the book. Chapter 1 outlines the data types that are common in environmental research, and chapter 2 discusses basic linear regression and explains the `regress` command and how to perform residual diagnostics.

Chapter 3 introduces readers to maximum likelihood estimation and the method of moments (and generalized method of moments) estimator. Of all the chapters, this may be the most mathematically demanding given there are many equations, so the content here may need to be supplemented depending on the audience. With that said, Baum and Hurn effectively and straightforwardly describe the ins and outs of maximum likelihood and method of moments estimation. I really appreciated this chapter because not many resources do a good job describing them and how they relate to ordinary least squares, which Baum and Hurn do well.

Chapter 4 introduces dynamics and univariate time-series modeling. I like that Baum and Hurn bring in community-contributed commands in this chapter (and in other chapters). The Stata community has produced very many useful commands, and they often do not get their due praise. Chapter 5 extends this discussion to multivariate time-series models and introduces vector autoregressive models and structural vector autoregressive models. Chapter 6 introduces common tests used to test for nonstationarity.

Chapter 7 discusses modeling nonstationary variables and cointegration. A major strength of this chapter is that the authors provide one of the clearer explanations of cointegration that I have read in a textbook. One addition I would like to see in this chapter is a discussion of the bounds-testing approach (Pesaran, Shin, and Smith 2001), which is the most common way to test for cointegration in single equations. Kripfganz and Schneider's (Forthcoming) `ardl` package can be used to implement the test, as can Jordan and Phillips' (2018) `dynamac` package, which also allows users to visualize long-run relationships.

Chapter 8 walks readers through how to perform forecasts using Stata's powerful `forecast` command and how to evaluate them. This chapter also discusses how to deal with retransformation bias when dealing with logarithmic dependent variables. This is excellent to see because, despite how common this issue is, researchers often do not have popular resources they can turn to on how to deal with it.

Chapters 9 to 11 deal with more advanced time-series topics. Chapter 9 discusses state-space and unobserved-components models, and chapter 10 discusses nonlinear time-series models. This chapter might be my personal favorite. The authors discuss Stata's `threshold` and `mswitch` commands in depth. They also walk readers through how to estimate smooth transition models in a simple, approachable manner. Chapter 11 examines how to model time-varying variance and introduces readers to the generalized autoregressive conditional heteroskedasticity model and extensions of the generalized autoregressive conditional heteroskedasticity model.

Chapter 12 transitions to panel-data analysis. The authors clearly and succinctly walk readers through the problems tied to using pooled ordinary least squares with panel data and provide a nice discussion of fixed-effects, random-effects, correlated random-effects, and hybrid models. The authors end the chapter with a short discussion on dynamic models. If there is one area in the book I would like to see expanded, it is this section. I think dynamic panel-data modeling deserves its own chapter. Dynamic panel-data modeling is pervasive across the environmental literature, and dedicating a chapter to discussing different approaches to fixed T and large T data analysis would be beneficial. However, I recognize that this is no easy task because this is a large and still-growing body of literature. It is also my experience that a major strength of Stata is the community-contributed commands that are available in this area (Ditzen 2018, 2021; Kripfganz 2016; Williams, Allison, and Moral-Benito 2018; De Vos, Everaert, and Ruyssen 2015; Roodman 2009; Kripfganz and Sarafidis 2021), and I think further developing this section of the book would be a nice addition to any subsequent editions.

The last three chapters of the book deal with spatial modeling, discrete dependent variables, and fractional integration. Chapter 13 provides readers with an introduction to spatial modeling in the cross-sectional context. This chapter effectively introduces readers to spatial weight matrices and common spatial models like the spatial lag and spatial error models. Chapter 14 is the lone chapter on discrete dependent variables, which walks readers through how to model binary dependent variables, ordered dependent variables, and censored dependent variables. Chapter 15, the last chapter, focuses on fractional integration, which is an extension of the previous time-series chapters.

3 Conclusion

Environmental Econometrics Using Stata is a must-have text for anyone interested in the environment and anyone who uses econometrics. The interdisciplinary and applied nature of this book allows it to serve as a core text in a range of graduate-level courses (and potentially upper-level undergraduate courses) that focus on the environment and quantitative analysis. I plan to adopt it in future related courses I teach, and I believe others should do the same.

4 References

Baum, C. F., and S. Hurn. 2021. *Environmental Econometrics Using Stata*. College Station, TX: Stata Press.

De Vos, I., G. Everaert, and I. Ruyssen. 2015. Bootstrap-based bias correction and inference for dynamic panels with fixed effects. *Stata Journal* 15: 986–1018. https://doi.org/10.1177/1536867X1501500404.

Ditzen, J. 2018. Estimating dynamic common-correlated effects in Stata. *Stata Journal* 18: 585–617. https://doi.org/10.1177/1536867X1801800306.

———. 2021. Estimating long-run effects and the exponent of cross-sectional dependence: An update to xtdcce2. *Stata Journal* 21: 687–707. https://doi.org/10.1177/1536867X211045560.

Jordan, S., and A. Q. Philips. 2018. Cointegration testing and dynamic simulations of autoregressive distributed lag models. *Stata Journal* 18: 902–923. https://doi.org/10.1177/1536867X1801800409.

Kripfganz, S. 2016. Quasi–maximum likelihood estimation of linear dynamic short-T panel-data models. *Stata Journal* 16: 1013–1038. https://doi.org/10.1177/1536867X1601600411.

Kripfganz, S., and V. Sarafidis. 2021. Instrumental-variable estimation of large-T panel-data models with common factors. *Stata Journal* 21: 659–686. https://doi.org/10.1177/1536867X211045558.

Kripfganz, S., and D. C. Schneider. Forthcoming. ardl: Estimating autoregressive distributed lag and equilibrium correction models. *Stata Journal*.

Pesaran, M. H., Y. Shin, and R. J. Smith. 2001. Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics* 16: 289–326. https://doi.org/10.1002/jae.616.

Roodman, D. 2009. How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal* 9: 86–136. https://doi.org/10.1177/1536867X0900900106.

Williams, R., P. D. Allison, and E. Moral-Benito. 2018. Linear dynamic panel-data estimation using maximum likelihood and structural equation modeling. *Stata Journal* 18: 293–326. https://doi.org/10.1177/1536867X1801800201.

About the author

Ryan P. Thombs is a PhD candidate in the sociology department at Boston College. His research examines the human dimensions of global environmental change, the drivers of health disparities and population health outcomes, and quantitative methodology, particularly related to panel data and time-series modeling. His work appears in journals such as the *American Sociological Review*, *Sociological Methodology*, the *Journal of Health and Social Behavior*, *Global Environmental Change*, *Environmental Research Letters*, *Science of the Total Environment*, *Climatic Change*, and *Energy Research and Social Science*.