The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*), *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
Printed & electronic		**Printed & electronic**	
1-year subscription	$ 98	1-year subscription	$138
2-year subscription	$165	2-year subscription	$245
3-year subscription	$225	3-year subscription	$345
1-year student subscription	$ 75	1-year student subscription	$ 99
1-year institutional subscription	$245	1-year institutional subscription	$285
2-year institutional subscription	$445	2-year institutional subscription	$525
3-year institutional subscription	$645	3-year institutional subscription	$765
Electronic only		**Electronic only**	
1-year subscription	$ 75	1-year subscription	$ 75
2-year subscription	$125	2-year subscription	$125
3-year subscription	$165	3-year subscription	$165
1-year student subscription	$ 45	1-year student subscription	$ 45

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The *Stata Journal* (ISSN 1536-867X) is a publication of Stata Press. Stata, Stata, Stata Press, Mata, Mata, and NetCourse are registered trademarks of StataCorp LP.

The Stata Journal (2014)
14, Number 2, pp. 445–448

Review of Introduction to Time Series Using Stata by Sean Becketti

Dimitriy V. Masterov
eBay Research Labs
San Jose, CA
dmasterov@ebay.com

Abstract. In this article, I review *Introduction to Time Series Using Stata* by Sean Becketti (2013 [Stata Press]).

Keywords: gn0060, book review, time series, forecasting, filtering

1 Introduction

Most statistics and econometrics classes begin with the assumption that observations are conditionally independent. When students are ready to relax this assumption, it is time for summer vacation or even graduation. The instructor covers time-series topics at breakneck speed before mumbling something about spatial and network data on his or her way out of the lecture hall. The coverage of these topics in otherwise excellent introductory econometrics books feels similarly peripheral. Even dedicated time-series courses and textbooks provide students with little help to develop applied skills because it takes them so much time just to navigate through the excess of equations. When students finally begin using a statistics package, their work is further complicated by the inconsistency in command syntax, notation, and even parameterization across the software and books, at least compared with cross-sectional or panel data.

If this describes your own experience with learning time series, or if you are tackling this material for the first time, Becketti's *Introduction to Time Series Using Stata* (ITSUS) is an excellent place to start. Becketti is a financial economist who was an early developer of Stata and an editor of the *Stata Technical Bulletin*, which eventually became this publication. This unique background has allowed him to write a gentle introduction to the core of time-series econometrics with many examples and an emphasis on intuition over theory. ITSUS fills a similar niche to Cameron and Trivedi (2010) and cross-sectional and panel data. The mathematical level of the presentation, as well as the price of the book, is between Hyndman and Athanasopoulos (2012) and Enders (2010). The tone is conversational with jokes that only an economist could love sprinkled throughout the text.

2 Contents and structure

The book opens with a brisk introduction to the Stata language, followed by a quick review of regression and hypothesis testing.[1] The book then continues by distinguishing the four components of a time series that recur throughout the text: the smooth signal composed of trend, seasonality, and cyclicality, and the rough of the noise. After that, the book discusses smoothing techniques that are relatively nonparametric or model agnostic. Becketti starts with the simplest approaches of medians and weighted means, explains when they will work and when they will not, and moves on to more sophisticated smoothers that can handle complications such as seasonality at the expense of added assumptions. At first, the data used in the examples are artificial, but soon Becketti guides the reader through problems using real macroeconomic data with moving averages and various exponential and Holt–Winters techniques to smooth and forecast the data. These methods work well for scalable, automated forecasting in industry and are easy to explain to one's colleagues.[2] The presentation is graphics driven, as it is throughout most of the book.[3] All Stata output is annotated. Each chapter opens with a map of the chapter's contents and closes with a set of bulleted "takeaways", which make the book easy to return to if one puts it down for a few days. There are also periodic pauses to survey the information and connect the dots between chapters.

Once the reader is familiar with smoothers, the following chapter introduces various types of forecasts and uses more intricate smoothers to predict the future. It also covers how to measure forecast accuracy. This section could have been expanded to include discussion of mean (percentage) absolute error and its many derivatives, training and test sets, and time-series cross-validation. Next the book discusses autocorrelated disturbances and how they interact with lagged dependent variables. It considers the possible causes, the problems for inference, testing for disturbances, and the appropriate remedies for either hurdle or both.

The next chapter continues with single-equation time-series models. The reader encounters the flexible workhorse of time-series econometrics: the autoregressive integrated moving-average (ARIMA) model. Becketti shows how to induce stationarity through differencing, how to pick the order of the autoregressive and moving-average parts, and how to perform model diagnostics with the residuals. The ARIMA forecasts are compared with ones from more elaborate smoothing methods, and, surprisingly, the ranking is sometimes unfavorable. This section should have discussed variance-stabilizing transformations: Box–Cox transformations, including the standard use of logarithms, and the problem of back-transformation to the original scale when forecasting are not mentioned at all. The text also fails to consider using information criteria for order selection or for choosing among competing models, and it does not discuss how to deal with outliers (unless it is to revert to compound smoothers).

1. Unfortunately, this is the part of the book where most of the typos are found.
2. These methods are called "filtering" in the text, but they should not be confused with techniques that separate a time series into trend and cyclical components, such as the Hodrick–Prescott filter. These are not covered in this book.
3. Some additional commands for visualizing time-series data are described in Cox (2006, 2009).

The second half of the book takes the reader further into the domains of macroeconometrics, and the mathematical notation becomes more challenging. Becketti first covers the autoregressive conditional heteroskedasticity model and some of its many generalizations. Intermittent volatility is common in financial markets, where any news can cause a period of turmoil followed by a period of calm. Becketti shows how to detect and deal with this special type of heteroskedasticity and even allows the impact of good and bad news to be asymmetric.

In the final part of the book, Becketti focuses on econometrics of multiple-equation models, particularly vector autoregression (VAR) and vector error correction, and on cointegrated nonstationary multivariate time series. In a VAR, all variables are considered endogenous, with each written as a linear function of its own lagged values and the lags of all the other variables. It is also possible to impose restrictions derived from economic theory on the contemporaneous relationship among variables to trace how shocks percolate through the system. The books closes with a list of Stata commands that were not covered and with brief definitions.

3 Gripes

It is much easier to complain about a book than it is to write one, particularly one as good as ITSUS. In my opinion, the second half of the book, except for the VAR material, will be of limited interest to those outside a particular circle of economists. While this challenging subject is explained unusually well, the opportunity cost may have been high. I would have given up the multivariate topics in exchange for more information on additive and multiplicative seasonality and exogenous covariates that turn the ARIMA into a seasonal ARIMA with exogenous variables, a seasonal transfer-function model. There are two reasons. First, seasonality played such a prominent role in the smoothing sections that it cannot be set aside later. My own work experience suggests that this is the case; no one has ever handed me seasonally adjusted data at the office. Second, the advantage of the more sophisticated methods over smoothing is that they provide a framework for counterfactual questions like what would happen to sales if spending for television advertising was cut by 20%. Perhaps a future edition of the book could cover this material in the context of unobserved components models. This would also provide a gradual introduction to state-space models. A future edition could also incorporate the econometric model forecasting commands introduced in Stata 13.

There are also topics that come up repeatedly among time-series users yet are not discussed in any time-series book I am aware of. A salient challenge in the era of big data is modeling high-frequency intraday data, such as the kind a utility company might see. Knowing how to study behavior that has multiple seasonal components (hourly, daily, and so on) would be tremendously helpful. Another increasingly common question, particularly on the web, is how to detect intervention in real time so that data outside the "normal" limits can alert the user that something has gone wrong (and occasionally right). Part of this relates to detecting regime changes; another part concerns dealing with outliers in the data used to build the model. Along these lines, a chapter on

the lore of time series that is not explicitly statistical would be useful. For instance, the chapter could explain how one deals with holidays, especially floating ones such as Easter. It could also demonstrate how one models lead and lag effects of holidays or promotional variables and discuss what form various compositional adjustments should take when time periods include different numbers of paydays. Finally, the chapter could explain how one deals with multiple time series that are hierarchically organized and can be aggregated at several different levels in groups based on products, geography, or some other features and produce consistent forecasts. These are the sorts of specialized, practical topics that I really liked about Cameron and Trivedi (2010).

4 Final thoughts

Stata's arsenal of time-series commands has expanded so much since date functions were added in release 6 that the *Stata Time-Series Reference Manual* currently has over 800 pages. No introductory book can cover this much ground. While there are some blank spots on its map of time-series topics, ITSUS surveys what is truly important and leaves its readers prepared to explore the rest on their own.

5 References

Becketti, S. 2013. *Introduction to Time Series Using Stata*. College Station, TX: Stata Press.

Cameron, A. C., and P. K. Trivedi. 2010. *Microeconometrics Using Stata*. Rev. ed. College Station, TX: Stata Press.

Cox, N. J. 2006. Speaking Stata: Graphs for all seasons. *Stata Journal* 6: 397–419.

———. 2009. Stata tip 76: Separating seasonal time series. *Stata Journal* 9: 321–326.

Enders, W. 2010. *Applied Econometric Time Series*. 3rd ed. Hoboken, NJ: Wiley.

Hyndman, R. J., and G. Athanasopoulos. 2012. *Forecasting: Principles and Practice*. OTexts. https://www.otexts.org/fpp.

About the author

Dimitriy V. Masterov is an economist at eBay Research Labs. His favorite time-series project was forecasting the number of pizza orders placed every five minutes on Superbowl Sunday.