The Stata Journal

Editors

H. Joseph Newton
Department of Statistics
Texas A&M University
College Station, Texas
editors@stata-journal.com

Nicholas J. Cox
Department of Geography
Durham University
Durham, UK
editors@stata-journal.com

Associate Editors

Christopher F. Baum, Boston College
Nathaniel Beck, New York University
Rino Bellocco, Karolinska Institutet, Sweden, and University of Milano-Bicocca, Italy
Maarten L. Buis, WZB, Germany
A. Colin Cameron, University of California–Davis
Mario A. Cleves, University of Arkansas for Medical Sciences
William D. Dupont, Vanderbilt University
Philip Ender, University of California–Los Angeles
David Epstein, Columbia University
Allan Gregory, Queen's University
James Hardin, University of South Carolina
Ben Jann, University of Bern, Switzerland
Stephen Jenkins, London School of Economics and Political Science
Ulrich Kohler, University of Potsdam, Germany
Frauke Kreuter, Univ. of Maryland–College Park
Peter A. Lachenbruch, Oregon State University
Jens Lauritsen, Odense University Hospital
Stanley Lemeshow, Ohio State University
J. Scott Long, Indiana University
Roger Newson, Imperial College, London
Austin Nichols, Urban Institute, Washington DC
Marcello Pagano, Harvard School of Public Health
Sophia Rabe-Hesketh, Univ. of California–Berkeley
J. Patrick Royston, MRC Clinical Trials Unit, London
Philip Ryan, University of Adelaide
Mark E. Schaffer, Heriot-Watt Univ., Edinburgh
Jeroen Weesie, Utrecht University
Nicholas J. G. Winter, University of Virginia
Jeffrey Wooldridge, Michigan State University

Stata Press Editorial Manager

Lisa Gilmore

Stata Press Copy Editors

David Culwell and Deirdre Skaggs

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	
1-year subscription	$ 79	1-year subscription	$115
2-year subscription	$155	2-year subscription	$225
3-year subscription	$225	3-year subscription	$329
3-year subscription (electronic only)	$210	3-year subscription (electronic only)	$210
1-year student subscription	$ 48	1-year student subscription	$ 79
1-year university library subscription	$ 99	1-year university library subscription	$135
2-year university library subscription	$195	2-year university library subscription	$265
3-year university library subscription	$289	3-year university library subscription	$395
1-year institutional subscription	$225	1-year institutional subscription	$259
2-year institutional subscription	$445	2-year institutional subscription	$510
3-year institutional subscription	$650	3-year institutional subscription	$750

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.

Copyright © 2012 by StataCorp LP

Copyright Statement: The *Stata Journal* and the contents of the supporting files (programs, datasets, and help files) are copyright © by StataCorp LP. The contents of the supporting files (programs, datasets, and help files) may be copied or reproduced by any means whatsoever, in whole or in part, as long as any copy or reproduction includes attribution to both (1) the author and (2) the *Stata Journal*.

The articles appearing in the *Stata Journal* may be copied or reproduced as printed copies, in whole or in part, as long as any copy or reproduction includes attribution to both (1) the author and (2) the *Stata Journal*.

Written permission must be obtained from StataCorp if you wish to make electronic copies of the insertions. This precludes placing electronic copies of the *Stata Journal*, in whole or in part, on publicly accessible websites, fileservers, or other locations where the copy may be accessed by anyone other than the subscriber.

Users of any of the software, ideas, data, or other materials published in the *Stata Journal* or the supporting files understand that such use is made without warranty of any kind, by either the *Stata Journal*, the author, or StataCorp. In particular, there is no warranty of fitness of purpose or merchantability, nor for special, incidental, or consequential damages such as loss of profits. The purpose of the *Stata Journal* is to promote free communication among Stata users.

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 (2012)
12, Number 4, pp. 571–574

The Stata Journal Editors' Prize 2012: David Roodman

1 Prize announcement

The editors of the *Stata Journal* are delighted to announce the award of the Editors' Prize for 2012 to **David Roodman**. The aim of the prize is to reward contributions to the Stata community in respect of one or more outstanding papers published in the *Journal* in the previous three calendar years. For the original announcement of the prize and its precise terms of reference, see Newton and Cox (2012), which is accessible at http://www.stata-journal.com/sjpdf.html?articlenum=gn0052.

The prize is awarded to David Roodman specifically for two outstanding papers in this journal:

- How to do xtabond2: An introduction to difference and system GMM in Stata (Roodman 2009b)
- Fitting fully observed recursive mixed-process models with cmp (Roodman 2011)

David Roodman is a senior fellow at the Center for Global Development (CGD), a think-tank in Washington, DC, focusing on global development. Born in Indianapolis, Indiana, in 1968, he grew up in Hanover, New Hampshire, and Binghamton, New York. Roodman's formal education ended in 1990 with a Bachelor's degree in theoretical mathematics from Harvard College. After years at the Worldwatch Institute and on a Fulbright in Vietnam, he arrived at CGD in 2002 knowing little about econometrics. He discovered that a great way to learn econometrics is to code it. His contributions to the Stata community since then were motivated by a desire to replicate and scrutinize complex, influential studies in development economics, which led him to write `xtabond2`, `cmp`, and other packages; and motivated by a pedagogic bent, which led him to document

© 2012 StataCorp LP

the packages and their mathematics in the *Stata Journal*. He is the author of *Due Diligence: An Impertinent Inquiry into Microfinance* (Roodman 2012).

In his 2011 paper, Roodman provides an encompassing framework for fitting, by maximum likelihood, multiequation systems in which the equation errors are distributed multivariate normal. This extends Stata's built-in commands to cover many models that until now required specialist software or bespoke Stata programs. Any one of the equations in the system might be a standard linear model (for continuous outcome), a probit model (for binary outcome), an ordered probit (for ordered categorical outcome), a tobit (linear model with censoring), an interval regression (grouped data model), or a truncated-distribution regression model (as it says). But the analyst can now combine these different types of equations within the system (and this means that selection modeling in the Heckman sense can also be encompassed). Multinomial probit models can also be fit. Since writing the article, Roodman has added the rank-ordered probit model and hierarchical random effects and coefficients. So, the researcher has the ability to fit many more models than previously possible. The `cmp` program is explained in excellent pedagogic fashion. Roodman provides background and a literature review placing mixed-process multiequation models in context, discusses computational issues, explains `cmp` itself, and shows examples. There is also a commendably moderate and modest style, with a frank discussion of the advantages and disadvantages of what is offered.

The main issue for fitting is computation of multivariate normal probabilities. Roodman uses simulated likelihood methods based on the well-known Geweke–Hajivassiliou–Keane (GHK) algorithm, also used in Stata's `ghkfast()` Mata function. One innovation was a new GHK function for Mata, `ghk2()`, designed to address some limitations of `ghkfast()`. Separately, Roodman pointed out that there was a "missing evaluator" type in Stata's `ml` command, which if provided would in some cases eliminate a theoretically unnecessary tradeoff between features of two existing types evaluator types (`lf` and `d1`) and significantly increase speed. After Roodman brought this to the attention of StataCorp, the company added the missing type (`lf1`). Extra features that extend the utility of `cmp` include its acceptance of the `svy:` prefix and weights as well as the inclusion of a general `predict` postestimation command, all features much in demand by researchers but certainly not trivial to implement. All in all, `cmp` enables the analyst not only to fit many existing models in Stata and those of several bespoke programs but also to extend the portfolio available to very general classes of multiequation processes, while also addressing real-world complications such as censoring and truncation. Before `cmp`, analysts wishing to fit many complex multiequation systems would have had to use other software.

In his 2009 paper about `xtabond2`, Roodman is concerned with models for repeated measures (panel data) on a continuous outcome in which the predictor variables include a lagged dependent variable. Often in practice with such data, there are many subjects but relatively few repeated measures per subject (in many researchers' jargon, "large N, small T"). In applied economics, these models are very widely used, and there is a long list of literature pointing to endogeneity issues: the lagged dependent variable means that special methods are required to get estimates having the usual desirable properties

(consistency, efficiency, and so forth). `xtabond2`, introduced first in 2003 but updated frequently thereafter, brought such facilities to Stata users for the first time, especially system GMM estimation methods, which were not then available in Stata's `xtabond` command. There were other innovations too: Windmeijer corrections to standard-error estimates, additional specification tests, a tool for differencing the data that preserved sample size in panels with gaps, and the facility to handle weighted data. Some but not all of these features have now been incorporated in Stata's `xtabond` and new commands such as `xtdpd` and `xtdpdsys`.

As with `cmp`, `xtabond2` incorporates sophisticated use of Mata programming. There is similar recognition of potential user needs. Model fitting can optionally be done using ado-code, and so the program's functionality is available to Stata users still running old versions. Roodman's 2009 paper on `xtabond2` remains a very important pedagogic reference for any researcher who wants to fit these types of models. The models are notoriously complicated, and estimates can be fragile if users are not careful in their specifications and testing thereof. The 50-page article is a wonderful resource, providing introduction and background to the models, their specification and estimation issues, and many examples. Note also the companion paper that focuses on some issues in more detail; see Roodman (2009a).

In summary, David Roodman has provided to the Stata community:

- Excellent programs substantially extending the functionality available to users
- Programs that incorporate innovative and sophisticated Mata programming
- Excellent accompanying articles in the *Stata Journal* that not only explain the programs but also are excellent free-standing pedagogic pieces in their own right

As editors, we are indebted to David Roodman for biographical material and to a necessarily anonymous nominator for a singularly lucid and detailed précis of Roodman's work.

H. Joseph Newton and Nicholas J. Cox
Editors, *Stata Journal*

2 References

Newton, H. J., and N. J. Cox. 2012. Announcement of the Stata Journal Editors' Prize 2012. *Stata Journal* 12: 1–2.

Roodman, D. M. 2009a. A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics* 71: 135–158.

———. 2009b. How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal* 9: 86–136.

———. 2011. Fitting fully observed recursive mixed-process models with cmp. *Stata Journal* 11: 159–206.

———. 2012. *Due Diligence: An Impertinent Inquiry Into Microfinance*. Washington, DC: Center for Global Development.