The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

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, University of Konstanz, 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
Ian White, MRC Biostatistics Unit, Cambridge
Nicholas J. G. Winter, University of Virginia
Jeffrey Wooldridge, Michigan State University

Stata Press Editorial Manager

Lisa Gilmore

Stata Press Copy Editors

David Culwell, Shelbi Seiner, 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	
Printed & electronic		**Printed & electronic**	
1-year subscription	$115	1-year subscription	$145
2-year subscription	$210	2-year subscription	$270
3-year subscription	$285	3-year subscription	$375
1-year student subscription	$ 85	1-year student subscription	$115
1-year institutional subscription	$345	1-year institutional subscription	$375
2-year institutional subscription	$625	2-year institutional subscription	$685
3-year institutional subscription	$875	3-year institutional subscription	$965
Electronic only		**Electronic only**	
1-year subscription	$ 85	1-year subscription	$ 85
2-year subscription	$155	2-year subscription	$155
3-year subscription	$215	3-year subscription	$215
1-year student subscription	$ 55	1-year student subscription	$ 55

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 © 2014 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 (2014)
14, Number 4, pp. 703–707

The Stata Journal Editors' Prize 2014: Roger Newson

1 Prize announcement

The editors of the *Stata Journal* are delighted to announce the award of the Editors' Prize for 2014 to **Roger Newson**. The aim of the prize is to reward contributions to the Stata community for one or more outstanding articles 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 the following website: http://www.stata-journal.com/sjpdf.html?articlenum=gn0052. The prize recognizes the outstanding article on "Attributable and unattributable risks and fractions and other scenario comparisons" (Newson 2013a) and three other articles by Newson that appeared within the time frame: "From resultssets to resultstables in Stata" (Newson 2012a), "Sensible parameters for univariate and multivariate splines" (Newson 2012b), and "Bonferroni and Holm approximations for Šidák and Holland–Copenhaver q-values" (Newson 2013b).

Roger Benedict Newson was born in Nottingham in 1955 and grew up there. He gained a bachelor's degree from the University of Nottingham, a master's degree from the University of Reading, and a doctoral degree from the University of Sussex. Morphing from a biologist to a statistician, Newson has combined work as a programmer, a statistician, and a lecturer, focusing mostly on biological and medical problems. He is currently a research associate in the Department of Primary Care and Public Health

© 2014 StataCorp LP

at Imperial College London. He has long been a stalwart of the Stata user community, as the author of many frequently downloaded programs, as a frequent speaker at Stata users group meetings (at every London meeting from 2000, in particular), and as a frequent contributor to Statalist. He has been an associate editor of the *Stata Journal* since 2003 and has organized Stata Users Group meetings in London with Stephen Jenkins in 2009, 2011, and 2013.

Newson's article (2013a) "Attributable and unattributable risks and fractions and other scenario comparisons" is an excellent article backed up with excellent software. The `margins` command introduced in Stata 11 and its related commands `pwcompare` and `contrast` added huge power to users' ability to interpret their estimation results. Regression coefficients, especially those for nonlinear models such as `logit` and `probit`, are often difficult to explain to nonspecialists. Results are much more easily understandable if they are expressed using an interpretable metric, say, changes or differences in the outcome of interest, such as a probability for a binary outcome or a mean for a continuous outcome. The great power of `margins` and its related commands is also their potential weakness: going beyond very basic options can be difficult for users. The manual entries are long and complicated, too. Newson's article goes beyond those manual entries.

The article has several notable features. Drawing on his deep understanding of the statistics involved, Newson first provides intuitive background to the problems that need to be solved and how he attacks them: estimation of "scenario means and comparisons", as he puts it. He explains the relevant methods and formulas and introduces his useful wrapper commands—`margprev`, `marglmean`, `regpar`, `punaf`, and `punafcc`—for `margins` and its related commands. Several worked examples illustrate the analytical tools and commands. One important aspect of the calculations of the estimates is their sampling variability, and here the nonlinearities raise issues as well. A further feature of Newson's calculations is careful attention to the automation of symmetric confidence intervals for transformed parameters and asymmetric confidence intervals for the various types of statistics, here making deft use of `nlcom` in the course of normalizing and variance-stabilizing transformations. As in all of Newson's programs, careful attention is paid to detail. All the standard options and extras that you might expect are present, for example, handling of weights, a portfolio of saved results, posting of results, and calculations for subgroups.

Newson considers postestimation calculations not only for standard regression models but also for case–control or survival-data models. In addition, he shows how his approach may be used to compare outcomes when different models are applied to the same scenario, as when standardizing statistics from different subpopulations to a common distribution (for example, by sex and age).

Newson refers to his work in the article as relevant to "applied scientists, especially in the public health sector" (2013a, page 672). Although his work is directed at that audience and terms such as attributable and unattributable risks and fractions may be unfamiliar outside that disciplinary context, Newson is here being too modest. For example, important special cases of his `regpar` and `punaf` commands can produce what

economists would call "average partial effects" and "relative probability ratios". These are produced with appropriate choices of the `atzero()` option. In these and other ways, Newson's work can be interesting and useful beyond his apparent target readership.

We turn now to the other articles cited in recommendation of Newson for the prize. In "From resultssets to resultstables in Stata" (Newson 2012a), Newson focuses on "delivery" aspects, rather than transformation of estimates as in the first article (Newson 2013a). This article takes forward earlier work for which he is deservedly renowned, specifically, Newson (2003a). His earlier programs provide "resultssets", sets of estimation results that are themselves Stata datasets. The issue addressed is how to convert such output into appropriately formatted tables for users who may be using some flavor of TEX or Microsoft Word. Together, those programs encompass most of the Stata user community.

Newson (2012a) provides and illustrates the `listtab` package (which supersedes his earlier `listtex` package). `listtab` calls several other utility packages that Newson has written and that are available from SSC. `listtab` inputs a list of variables and outputs them as a table in one of several formats, including TEX, LATEX, HTML, Microsoft Rich Text Format, or possibly future XML-based formats. This work represents a substantial contribution to, and enhancement of, the medley of output formatting packages available to the Stata user community.

"Sensible parameters for univariate and multivariate splines" (Newson 2012b) returns the emphasis to providing tools that assist end-users' interpretations of estimation results. Splines of various kinds are commonly used in regression modeling to allow for nonlinear relationships between predictors and outcome, but the estimated parameters are tricky to interpret. Newson describes the update of his `bspline` package for creating spline variables (originally published in Newson [2000a]) and the associated utility programs `frencurv` and a new `flexcurv`. He provides extensive illustration using worked examples. There is also helpful technical discussion of different types of splines, useful for those who wish to make connections between the functionality built into Stata and more complicated specifications.

"Bonferroni and Holm approximations for Šidák and Holland–Copenhaver q-values" (Newson 2013b) is a note building on an earlier article on frequentist q-values for multiple-test procedures (Newson 2010b), with its associated `qqvalue` package. Newson identifies a potential problem: there may be issues of precision, especially with q-values corresponding to very small p-values, and proposes a remedy, which he also discusses in relation to computing q-values and discovery sets. This topic is relatively specialized, but the style is all about getting every detail right, thoroughly in the best traditions of Stata.

In summary, we salute Roger Newson for outstanding contributions to the Stata community, based on excellent programs greatly extending the functionality available to users explained in excellent accompanying articles in the *Stata Journal*, all within a laudable emphasis on providing tools delivering interpretable findings to end users. In every case, he takes a focused approach based on sound general statistical principles, and he shows extraordinary care and attention to detail.

As editors, we are indebted to the awardee for biographical material and to a necessarily anonymous nominator for a most helpful appreciation of Newson's work. Below, we give references to the most notable of Newson's publications in the *Stata Technical Bulletin* and *Stata Journal*. Additionally, his software updates flagged in the *Stata Journal* can be found in Stata by typing `search newson, author`.

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

2 References

Newson, R. B. 1999a. dm65: A program for saving a model fit as a dataset. *Stata Technical Bulletin* 49: 2–6. Reprinted in *Stata Technical Bulletin Reprints*, vol. 9, pp. 19–23. College Station, TX: Stata Press.

———. 1999b. sg114: rglm—Robust variance estimates for generalized linear models. *Stata Technical Bulletin* 50: 27–33. Reprinted in *Stata Technical Bulletin Reprints*, vol. 9, pp. 181–190. College Station, TX: Stata Press.

———. 2000a. sg151: B-splines and splines parameterized by their values at reference points on the x-axis. *Stata Technical Bulletin* 57: 20–27. Reprinted in *Stata Technical Bulletin Reprints*, vol. 10, pp. 221–230. College Station, TX: Stata Press.

———. 2000b. snp15: somersd—Confidence intervals for nonparametric statistics and their differences. *Stata Technical Bulletin* 55: 47–55. Reprinted in *Stata Technical Bulletin Reprints*, vol. 10, pp. 312–322. College Station, TX: Stata Press.

———. 2000c. snp16: Robust confidence intervals for median and other percentile differences between groups. *Stata Technical Bulletin* 58: 30–35. Reprinted in *Stata Technical Bulletin Reprints*, vol. 10, pp. 324–331. College Station, TX: Stata Press.

———. 2002. Parameters behind "nonparametric" statistics: Kendall's tau, Somers' D and median differences. *Stata Journal* 2: 45–64.

———. 2003a. Confidence intervals and p-values for delivery to the end user. *Stata Journal* 3: 245–269.

———. 2003b. Stata tip 1: The eform() option of regress. *Stata Journal* 3: 445.

———. 2004a. Generalized power calculations for generalized linear models and more. *Stata Journal* 4: 379–401.

———. 2004b. Stata tip 13: generate and replace use the current sort order. *Stata Journal* 4: 484–485.

———. 2004c. Stata tip 5: Ensuring programs preserve dataset sort order. *Stata Journal* 4: 94.

———. 2006a. Confidence intervals for rank statistics: Percentile slopes, differences, and ratios. *Stata Journal* 6: 497–520.

———. 2006b. Confidence intervals for rank statistics: Somers' *D* and extensions. *Stata Journal* 6: 309–334.

———. 2010a. Comparing the predictive powers of survival models using Harrell's *C* or Somers' *D*. *Stata Journal* 10: 339–358.

———. 2010b. Frequentist *q*-values for multiple-test procedures. *Stata Journal* 10: 568–584.

———. 2012a. From resultssets to resultstables in Stata. *Stata Journal* 12: 191–213.

———. 2012b. Sensible parameters for univariate and multivariate splines. *Stata Journal* 12: 479–504.

———. 2013a. Attributable and unattributable risks and fractions and other scenario comparisons. *Stata Journal* 13: 672–698.

———. 2013b. Bonferroni and Holm approximations for Šidák and Holland–Copenhaver *q*-values. *Stata Journal* 13: 379–381.

Newson, R. B., and the ALSPAC Study Team. 2003. Multiple-test procedures and smile plots. *Stata Journal* 3: 109–132.

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