THE STATA JOURNAL

Editor
H. Joseph Newton
Department of Statistics
Texas A&M University
College Station, Texas 77843
979-845-8817; fax 979-845-6077
jnewton@stata-journal.com

Editor
Nicholas J. Cox
Department of Geography
Durham University
South Road
Durham City DH1 3LE UK
n.j.cox@stata-journal.com

Associate Editors
Christopher F. Baum
Boston College
Nathaniel Beck
New York University
Rino Bellocco
Karolinska Institutet, Sweden, and
Univ. degli Studi di Milano-Bicocca, Italy
Maarten L. Buis
Vrije Universiteit, Amsterdam
A. Colin Cameron
University of California–Davis
Mario A. Cleves
Univ. of Arkansas for Medical Sciences
William D. Dupont
Vanderbilt University
David Epstein
Columbia University
Allan Gregory
Queen’s University
James Hardin
University of South Carolina
Ben Jann
ETH Zürich, Switzerland
Stephen Jenkins
University of Essex
Ulrich Kohler
WZB, Berlin
Frauke Kreuter
University of Maryland–College Park

Jens Lauritsen
Odense University Hospital
Stanley Lemeshow
Ohio State University
J. Scott Long
Indiana University
Thomas Lumley
University of Washington–Seattle
Roger Newson
Imperial College, London
Austin Nichols
Urban Institute, Washington DC
Marcello Pagano
Harvard School of Public Health
Sophia Rabe-Hesketh
University of California–Berkeley
J. Patrick Royston
MRC Clinical Trials Unit, London
Philip Ryan
University of Adelaide
Mark E. Schaffer
Heriot-Watt University, 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
Jennifer Neve and Deirdre Patterson
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.

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

http://www.stata-journal.com

The *Stata Journal* is indexed and abstracted in the following:

- Science Citation Index Expanded (also known as SciSearch®)
- CompuMath Citation Index®

**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 web sites, file servers, 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*, electronic version (ISSN 1536-8734) is a publication of Stata Press. Stata and Mata are registered trademarks of StataCorp LP.
Stata tip 62: Plotting on reversed scales
Nicholas J. Cox and Natasha L. M. Barlow
Durham University
Durham City, UK
n.j.cox@durham.ac.uk and n.l.m.barlow@durham.ac.uk

Stata has long had options allowing a reversed scale on either the $y$ or the $x$ axis of many of its graph types. Many graph users perhaps never even consider specifying such options. Those who do need to reach for them may wish to see detailed examples of how reversed scales may be exploited to good effect.

The usual Cartesian conventions are that vertical or $y$ scales increase upward, from bottom to top, and horizontal or $x$ scales increase from left to right. Vertical scales increasing downward are needed for graphs with vertical categorical axes following a table-like convention, in which the first (lowest) category is at the top of a graph, just as it would be at the top of a table. Commands such as graph bar and graph dot follow this convention. Indeed, it is likely to be so familiar that you may have to reflect briefly to see that is how such graphs are drawn. Other examples of this principle are discussed elsewhere in this issue (Cox 2008).

Reversed scales are also common in the Earth and environmental sciences. Here, in fields such as pedology, sedimentology, geomorphology, limnology, and oceanography, it is common to take measurements at varying depths within soils, sediments, rocks, and water bodies. Even though depth is not a response variable, it is conventional and convenient to plot depth below surface on the vertical axis; hence, the need for a reversed scale. An extra twist that gives spin to the graph problem is that frequently detailed analyses of materials at each level in a core, bore, or vertical profile yield several response variables, which are all to be plotted on the horizontal axis. Such multiple plotting is easiest when overlay is possible.

Let us look at some specific syntax for an example and then add comments. The data shown here come from work in progress by the second author. Sediment samples at 2 cm intervals down a core from Girdwood, Alaska, were examined for several elements associated with placer mining pollution (LaPerriere, Wagener, and Bjerklie 1985). In this example, the concentrations of gold, cadmium, arsenic, lead, copper, and zinc, measured in parts per million (ppm), are then plotted as a function of depth (cm). See figure 1.

(Continued on next page)
A key detail here is that Stata’s model for scatter and similar plots is asymmetric. One or more $y$ variables are allowed, but only one $x$ variable, in each individual plot. Thus, if you wish to have several $x$ variables, you must either superimpose several variables, as here, or juxtapose several plots horizontally.

Further, it is necessary to spell out what is desired as text in the legend. By default, `twoway` would use information for the $y$ axis variables, which is not what is wanted for this kind of graph.

For these data a logarithmic scale is helpful, indeed essential, for showing several elements that vary greatly in abundance. Gold is reported as less than 0.1 ppm at depth. Such censored values cannot be shown by point symbols.

Even very experienced Stata users will not usually think up an entire graph command like this at the outset. Typically, you start with a fairly simple design and then elaborate it by a series of very small changes (which may well be changes of mind back and forth).
At some point, you are likely to find yourself transferring from the Command window to the Do-file Editor and from an interactive session to a do-file. We also find it helpful, once out of the Command window, to space out a command so that its elements are easier to see and so that it is easier to edit. Spending a few moments doing that saves some fiddly work later on. What is shown above is in fact more compressed than what typically appears within the Do-file Editor in our sessions.

The legend used here is, like all legends, at best a necessary evil, as it obliges careful readers to scan back and forth repeatedly to see what is what. The several vertical traces are fairly distinct, so one possibility is to extend the vertical axis and insert legend text at the top of the graph. `{=Au[1]}`, for example, instructs Stata to evaluate the first value of $Au$ and use its value. The same effect would be achieved by typing in the actual value. Here we are exploiting the shortness of the element names, which remain informative to any reader with minimal chemical knowledge. The legend itself can then be suppressed. The text elements could also be repeated at the bottom of the graph if desired. See figure 2.

```stata
. twoway
   > connect depth Au, `spec´ ms(Oh) cmissing(n) ||
   > connect depth Cd, `spec´ ms(Th) ||
   > connect depth As, `spec´ ms(Sh) ||
   > connect depth Pb, `spec´ ms(O) ||
   > connect depth Cu, `spec´ ms(T) ||
   > connect depth Zn, `spec´ ms(S)
   > yscale(reverse r(46 .)) xscale(log) xscale(titlegap(*10))
   > ylabel(50(10)90, angle(h)) xlabel(0.1 0.3 1 3 10 30 100)
   > ytitle(Depth (cm)) xtitle(Concentration (ppm))
   > text(48 `{=Au[1]}` "Au" 48 `{=Cd[1]}` "Cd" 48 `{=As[1]}` "As" 48 `{=Pb[1]}` "Pb"
   > 48 `{=Cu[1]}` "Cu" 48 `{=Zn[1]}` "Zn")
   > legend(off)
```

(Continued on next page)
Figure 2. Variation in concentrations of various elements with depth, measured at a site in Girdwood, Alaska. The legend has been suppressed and its text elements placed within the graph.

The structure of overlays may suggest that a program be written that takes one $y$ and several $x$ variables and then works out the overlay code for you. Such a program might then be available particularly to colleagues and students less familiar with Stata. That idea is more tempting in this example than in general. In other cases, there might be a desire to mix different `twoway` types, to use different marker colors, or to make any number of other changes to distinguish the different variables. Any code flexible enough to specify any or all of that would lead to commands no easier to write than what is already possible. In computing as in cookery, sometimes you just keep old recipes once you have worked them out, at least as starting points for some later problem.

In the Earth and environmental sciences, reversed horizontal scales are also common for showing time, whenever the units are not calendar time, measured forward, but time before the present, measured backward. Within Stata, such practice typically requires no more than specifying `xscale(reverse)`.

References
