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](http://www.stata-journal.com)

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

- CompuMath Citation Index®
- RePEc: Research Papers in Economics
- Science Citation Index Expanded (also known as SciSearch®)

**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.
Grids of horizontal or vertical lines within graphs were greatly used in the past, but more recently have been greatly disapproved. The need for grids as guidelines in preparing graphs has disappeared as computers have displaced people as graph constructors. Their use for precise look-up of particular values has diminished as more datasets become electronically accessible. To present tastes, past compendiums of graph types, such as Brinton (1914) and Karsten (1923), groan with the weight of heavy grids in example after example. The low point for grids came with the dismissal by Tufte (2001) of dark grid lines as “chartjunk”, but he also emphasized that light grid lines could be helpful. The arguments of Tufte, Cleveland (1994), Kosslyn (2006), and others, and the flexibility of modern graphic technology, imply that grids should and can be subtle and subdued. Ideally, grids will be just noticeable so that they can be tuned in and out of attention by graph readers.

The grounds for grids are pragmatic and aesthetic—they can be useful and they can be pleasing. There is some room for disagreement on the first ground and much room on the second. The aim of this tip is not to change your mind on how you should prepare your own graphs, but more to underline some of the possibilities offered by Stata.

Stata provides support for grids, although you may easily not have noticed. For example, with the auto.dta dataset, Stata’s default s2color graph scheme, and the canonical graph

```
.scatter mpg weight
```

subtle grid lines appear for mpg values 10, 20, 30, and 40. That also is true with the sj scheme used in the Stata Journal. These grid lines are associated with the corresponding axis labels. If you do not want them, you can turn them off with an option, such as `ylabel(, nogrid)`. Grid lines can also be associated with axis ticks: more usually, you would need to turn those on with an option, such as `ytick(15(10)35, grid)`—not that this is an especially good idea. Grid lines in the informal wider sense may also be added through options such as `yline()`. What goes for the y axis also goes for the x axis, at least as far as `twoway` is concerned.

To use grids effectively and tastefully, it is essential to be able to tune line width, color, and style, as well as horizontal or vertical position. Stata makes this easier by providing, at least in principle, `grid`, `major grid`, and `minor grid lifestyles`; see `[G] linestyle`. In practice, these three need not be distinct, depending precisely on the graph scheme in use.Independently of that, you can tweak grid lines as desired through standard line options.
As a first example, let us do what Arbuthnott (1710) did not do: plot his data on the ratio of males and females christened in London. The main statistical point for these data is that the average ratio is definitely not 1, as researchers would now typically flag with a significance test or confidence interval. The fuller history is enlightening and entertaining and can be found in many sources, including Stigler (1986) and Hacking (2006). Plotting these data was inspired by Friendly (2007). A data file is provided with the electronic media for this issue as arbuthnott.dta.

```
. use arbuthnott
```

The obvious reference lines for a time-series plot of the ratio are the line of equality, \( y = 1 \), and the line of the mean, \( y = 1.066918 \). Some experimenting indicates that `yline()` by itself produces rather stark lines, whereas `yline(, grid)` produces very subdued ones. One compromise is to specify `lcolor(gs12)`, thus making use of gray scale (Cox 2009). Adding these lines would make the graph appear a little busy given the grid lines provided by default to match the `ylabel` s at 1, 1.05, 1.1, and 1.15, so we use `nogrid` to turn three of those off. Figure 1 is the result.

```
. line ratio year, yline(1.066918 1, lstyle(grid) lcolor(gs12))
> ylabel(, nogrid angle(h)) xtitle(""")
```

![Figure 1. Arbuthnott's data on sex ratio in London christenings. The use of reference lines is shown in grid style.](image)

Grids can be especially useful on scatterplots. Consider these data on city temperatures in the United States:

```
. sysuse citytemp
```
First, we create Celsius versions of each temperature variable:

```
. clonevar tempjanC = tempjan
. replace tempjanC = (5/9) * (tempjanC - 32)
. clonevar tempjulC = tempjuly
. replace tempjulC = (5/9) * (tempjulC - 32)
```

A natural reference line, which we make thicker than the default, is at the freezing point \(0^\circ\text{C} (32^\circ\text{F})\). We add more grid lines at other temperatures so that both axes are gridded. A major twist is that lighter and darker are inverted: the grid lines are white and the plot region and axes are set to a light gray, `gs14`. Figure 2 shows what this produces.

```
. scatter tempjulC tempjanC, ms(oh)
> xline(0, lstyle(grid) lcolor(white) lwidth(*1.5))
> xlabel(, grid glcolor(white)) ylabel(, angle(h) glcolor(white))
> plotregion(color(gs14)) graphregion(color(white)) xscale(lcolor(gs14))
> yscale(lcolor(gs14)) note({c 176}Celsius) ms(oh)
```

![Figure 2. City temperature data for the United States. Grid lines are shown for both variables, and lighter and darker are inverted.](image)

Such a look to graphs is in conscious imitation of a look very popular in the R community, particularly using its `ggplot2` package as devised and documented by Wickham (2009). Aside from that, the provision of grid lines is thus a natural way of emphasizing zeros or other key levels on one or both variables. Stata users sometimes want to move the axes so that they intersect at the origin \((0, 0)\), just as they may well have been taught to do when young. Stata will not do that if either variable is ever negative, but clear grid lines provide a way to emphasize such levels while also ensuring that the data are not obscured by the axes or their associated labels, ticks, and titles.
References


