

# THE STATA JOURNAL

## Guest Editor

David M. Drukker  
StataCorp

## Editor

H. Joseph Newton  
Department of Statistics  
Texas A & M University  
College Station, Texas 77843  
979-845-3142; FAX 979-845-3144  
jnewton@stata-journal.com

## Editor

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

## Associate Editors

Christopher Baum  
Boston College

Rino Bellocco  
Karolinska Institutet, Sweden and  
Univ. degli Studi di Milano-Bicocca, Italy

David Clayton  
Cambridge Inst. for Medical Research

Mario A. Cleves  
Univ. of Arkansas for Medical Sciences

William D. Dupont  
Vanderbilt University

Charles Franklin  
University of Wisconsin, Madison

Joanne M. Garrett  
University of North Carolina

Allan Gregory  
Queen's University

James Hardin  
University of South Carolina

Ben Jann  
ETH Zurich, Switzerland

Stephen Jenkins  
University of Essex

Ulrich Kohler  
WZB, Berlin

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

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  
Cornell University

Jeffrey Wooldridge  
Michigan State University

## Stata Press Production Manager

## Stata Press Copy Editor

Lisa Gilmore  
Gabe Waggoner

**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, and Stata is a registered trademark of StataCorp LP.

## Stata tip 31: Scalar or variable? The problem of ambiguous names

Gueorgui I. Kolev  
Universitat Pompeu Fabra  
Barcelona, Spain  
gueorgui.kolev@upf.edu

Stata users often put numeric or string values into scalars, which is as easily done as

```
. scalar answer = 42  
. scalar question = "What is the answer?"
```

The main discussion of scalars is in [P] [scalar](#). Scalars can be used interactively or in programs and are faster and more accurate than local macros for holding values. This advantage would not matter with a number like 42, but it could easily matter with a number that was very small or very large.

Scalars have one major pitfall. It is documented in [P] [scalar](#), but users are often bitten by it, so here is another warning. If a variable and a scalar have the same name, Stata always assumes that you mean the variable, not the scalar. This naming can apply even more strongly than you first guess: recall that variable names can be abbreviated so long as the abbreviation is unambiguous (see [U] [11.2 Abbreviation rules](#)).

Suppose that you are using the `auto` dataset:

```
. sysuse auto  
(1978 Automobile Data)  
. scalar p = 0.7  
. display p  
4099
```

What happened to 0.7? Nothing. It is still there:

```
. display scalar(p)  
.7  
. scalar list p  
      p =          .7
```

What is the 4099 result? The dataset has a variable, `price`, and no other variable names begin with `p`, so `p` is understood to mean `price`. Moreover, `display` assumes that if you specify just a variable name, you want to see its value in the first observation, namely, `price[1]`. (The full explanation for that is another story.)

What should users do to be safe?

1. Use a different name. For example, you might introduce a personal convention about uppercase and lowercase. Many Stata users use only lowercase letters (and possibly numeric digits and underscores too) within variable names. Such users

could distinguish scalars by using at least one uppercase letter in their names. You could also use a prefix such as `sc_`, as in `sc_p`. If you forget your convention or if a user of your programs does not know about this convention, Stata cannot sense the interpretation you want. Thus this method is not totally safe.

2. Use the `scalar()` pseudofunction to spell out that you want a scalar. This method is totally safe, but some people find it awkward.
3. Use a temporary name for a scalar, as in

```
. tempname p  
. scalar 'p' = 0.7
```

Each scalar with a temporary name will be visible while the program in which it occurs is running and only in that program. Temporary names can be used interactively, too. (Your interactive session also counts for this purpose as a program.) In many ways, this method is the best solution, as it ensures that scalars can be seen only locally, which is usually better programming style.