



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.

Stata tip 5: Ensuring programs preserve dataset sort order

Roger Newson, King's College London, UK
roger.newson@kcl.ac.uk

Did you know about **sortpreserve**? If you are writing a Stata program that temporarily changes the order of the data and you want the data to be sorted in its original order at the end of execution, you can save a bit of programming by including **sortpreserve** on your **program** statement. If your program is called **myprogram**, you can start it with

```
program myprogram, sortpreserve
```

If you do this, you can change the order of observations in the dataset in **myprogram**, and Stata will automatically sort it in its original order at the end of execution. Stata does this by creating a temporary variable whose name is stored in a macro named **_sortindex**, which is discussed in the manuals under [P] **sortpreserve**. (Note, however, that there is a typo in the manual; the underscore in **_sortindex** is missing.) The temporary variable '**_sortindex**' contains the original sort order of the data, and the dataset is sorted automatically by '**_sortindex**' at the end of the program's execution.

If you know about temporary variables, you might think that **sortpreserve** is unnecessary because you can always include two lines at the beginning, such as

```
tempvar order  
generate long 'order' = _n
```

and a single line at the end such as

```
sort 'order'
```

and do the job of **sortpreserve** in 3 lines. However, **sortpreserve** does more than that. It restores the result of the macro extended function **sortedby** to the value that it would have had before your program executed. (See [P] **macro** for a description of **sortedby**.) Also, it restores the "Sorted by:" variable list reported by the **describe** command to the variable list that would have been reported before your program executed. For example, in the **auto** dataset shipped with official Stata, the output of **describe** ends with the message

```
Sorted by:  foreign
```

This will not be changed if you execute a program defined with **sortpreserve**.