The Stata Journal

The Stata Journal, electronic version (ISSN 1536-8734) is a publication of Stata Press, and Stata is a registered trademark of Stata Corporation.

The Stata Journal (2001)
1, Number 1, pp. 76–85

From the help desk

Allen McDowell
Stata Corporation
amcdowell@stata.com

Abstract. Welcome to *From the help desk*. *From the help desk* is written by the people in Technical Services at StataCorp and deals with issues that they have found to be of concern to a large fraction of Stata users. It is the rare column in this series that deals with sophisticated programming issues because such issues, by definition, are not of concern to a large fraction of Stata users. *From the help desk* discusses the use of sophisticated programs and the use of sophisticated statistics.

Keywords: pr0002, internet, web, ado-files, Stata executable installation, updates, downloading, user-written additions, packages, search, find

1 Updating Stata

It is surprising how many technical-support questions we receive for which the answer is, "Update your Stata and let me tell you how." Updates to Stata are available over the Internet, and Stata itself can find and install those updates. Although we follow no formal schedule for the release of updates, the fact is that we update Stata about once every eleven days. As of the date of this writing, Stata 7 is 263 days old, and since its release we have updated it 25 times. Stata 6 was updated 63 times over its 703 days of life. That amounts to $263/25 = 10.52$ and $703/63 = 11.16$ days on average between updates.

We update Stata for two reasons: to fix bugs and to add new features. In the 25 updates we have released for Stata 7 at the time of this writing, we have fixed 91 bugs and added 53 new features, so the average update includes 3.64 bug fixes and 2.12 new features, and we do that, on average, once every eleven days. Most of the additions—whether bug fixes or new features—are admittedly minor, but the occasional fix or feature can be of great importance and may be important to you, whether or not we here at Stata consider it minor.

Installing updates is easy; all you have to do is

type `update query`

or

pull down **Help**, select **Official Updates**, and click on *www.stata.com*

Do that and then follow the instructions, which will amount to clicking or typing one more command to tell Stata to load and install the update.

Once you have installed an update, to find out what has changed

type `help whatsnew`

or

pull down **Help** and select **What's New**

Official updates are only a part of what is available for Stata over the Internet. There are FAQs that explain, in detail, simple and complicated things, books on Stata that are available for free, and user-written programs that do simple and/or remarkable things. The trick is learning how to find these materials. In the official updates, thanks to a user's suggestion, we have added a new command to make finding these resources easier. The command is `findit`, but, if you have not updated your Stata and you type `findit`, you will see

```
. findit seemingly unrelated regression
unrecognized command:  findit
r(199);
```

Here is what I see using my up-to-date Stata:

```
. findit seemingly unrelated regression
                                                        5 Sep 2001 10:13:12
Keyword search

        Keywords:  seemingly unrelated regression
       Criterion:  Select only entries that have ALL the above words (*)
          Search:  (1) Official help files, FAQs, and STBs
                   (2) Web resources from Stata and from other users

        * To search entries that have ANY of the above words, type
          findit seemingly unrelated regression, or

Search of official help files, FAQs, and STBs

[R]     biprobit . . . . . . . . . . . . . . . . . . . . . .  Bivariate probit
        (help biprobit)

[R]     reg3 . .  Three-stage estimation for systems of simultaneous equations
        (help reg3)

[R]     sureg  . . . . . . . . . . .  Zellner's seemingly unrelated regression
        (help sureg)

FAQ     . . . . . . . . . . . . . . . . . . . Missing R-squared for 2SLS/IV/3SLS
        . . . . . . . . . . . . . . . . . . . . . . . .  W. Sribney and V. Wiggins
        4/99    For two-stage least-squares (2SLS/IV/ivreg) estimates, why
                is the R-squared statistic not printed in some cases?
                For two-stage least-squares (2SLS/IV/ivreg) estimates, why
                is the Model Sum of Squares sometimes negative?
                For two-stage least-squares (3SLS/IV/ivreg) estimates, why
                are the R-squared and Model Sum of Squares sometimes
                negative?
                http://www.stata.com/support/faqs/stat/2sls.html

FAQ     . . . . . . . . . . . . . . .  What is seemingly unrelated regression?
        . . . . . . . . . . . . . . . . . . . .  UCLA Academic Technology Services
        5/01    What is seemingly unrelated regression and how can I
                perform it in Stata?
                http://www.ats.ucla.edu/stat/stata/faq/sureg.htm
```

```
STB-52  sg121  . . Seemingly unrelated est. and cluster-adjusted sandwich est.
        (help suest if installed)  . . . . . . . . . . . . . . . . . J. Weesie
        11/99   STB Reprints Vol 9, pages 231--248
        provides a generalized Hausman test and seemingly unrelated
        estimation (a generalization of seemingly unrelated regression)
        that can work with a combination of many different estimators
        including combinations of discrete and continuous models; two
        or more models may be combined and cross-model hypotheses may
        be tested

Web resources from Stata and other users

(contacting http://www.stata.com)

3 packages found (STB omitted)
------------------------------

madfuller from http://fmwww.bc.edu/RePEc/bocode/m
    'MADFULLER': module to perform Dickey-Fuller test on panel data /
    madfuller performs the multivariate augmented Dickey-Fuller panel / unit
    root test (Sarno and Taylor, 1998; Taylor and Sarno, 1998) / on a variable
    that contains both cross-section and time-series / components.  The test

reg3 from http://fmwww.bc.edu/RePEc/bocode/r
    'REG3': modules to perform three-stage least squares and SURE (version 5)
    / Version 6 users should use the built-in -reg3 command. This is / version
    1.06 (11 June 1998) of a three-stage least squares / command.  reg3
    introduces a new syntax for multiple equation / models. You could estimate

xttest2 from http://fmwww.bc.edu/RePEc/bocode/x
    'XTTEST2': module to Breusch-Pagan LM test for cross-sectional correlation
    in fixed effects model / xttest2 calculates the Breusch-Pagan statistic
    for / cross-sectional independence in the residuals of a fixed effect /
    regression model.  xtreg, fe estimates this model assuming / independence

1 reference found in tables of contents
---------------------------------------

http://www.fss.uu.nl/soc/iscore/stata/
    ICS-Lib Commands for Stata release 6 version 14mar2000 / Jeroen Weesie
    email: J.Weesie @ fss.uu.nl / Department of Sociology URL:
    www.fss.nl/soc/iscore/staff / Utrecht University tel:  (+)31 - 030 - 2 53
    1922 / The Netherlands fax:  (+)31 - 030 - 2 53 4405 / The first version

(end of search)
```

I typed `findit seemingly unrelated regression` and found three references to the manual and on-line help, two FAQs (one at UCLA and one at StataCorp), and four user-written programs (of which one was published in the STB).

The best thing about the above output is that I can click on the underlined references to be taken directly to the source. If the source is on-line help, I'll see the help file. If it is a FAQ, my browser will open and display the FAQ. If it is a program, I will be taken to a description of the program where I am just one click away from automatic installation.

But before you can use `findit`, you must update ...

2 Updates

There are actually two components to Stata: the Stata executable itself and Stata's ado-files, which are programs written in Stata's programming language. Most of Stata's statistical features are implemented as ado-files.

Correspondingly, there are two kinds of updates: updates to Stata's ado-files and updates to the Stata executable. The most common update is the ado-file update, and that is the easier to install. Of the 25 updates so far, 19 are ado-file updates. When an update contains only ado-files, all you need to do is tell Stata to install the new ado-files. Once Stata has done that, you are updated. You do not even have to restart your Stata.

The remaining 6 of the 25 updates have been executable updates. These are more work to install because (1) downloading the new executable takes longer and (2) you have to exit Stata and rename the executable yourself.

Even if you have never updated your Stata, you will not have to install the 25 updates one at a time. You will install one ado-file update and one executable update, and Stata will determine, within each, what files need updating automatically. It all starts by typing `update query` or pulling down **Help**, selecting **Updates**, and clicking on *www.stata.com*.

Here is what I see when I type `update query` on my fully up-to-date Stata:

```
. update query
(contacting http://www.stata.com)
Stata executable
    folder:               c:\stata\
    name of file:         wstata.exe
    currently installed:  08 Aug 2001
    latest available:     08 Aug 2001
Ado-file updates
    folder:               c:\stata\ado\updates\
    names of files:       (various)
    currently installed:  24 Aug 2001
    latest available:     24 Aug 2001
Recommendation
    Do nothing; all files up to date.
```

Here is what I would see if I had never updated my Stata and I had installed it from the original CDROM that shipped when Stata 7 was first released:

```
. update query
(contacting http://www.stata.com)
Stata executable
    folder:               c:\stata\
    name of file:         wstata.exe
    currently installed:  15 Dec 2000
    latest available:     08 Aug 2001
```

```
Ado-file updates
    folder:               c:\stata\ado\updates\
    names of files:       (various)
    currently installed:  15 Dec 2000
    latest available:     24 Aug 2001

Recommendation
    Type -update all-
```

`update query` does nothing more than report on the status of your installation and tell you what you should do next. Under the recommendation, `update query` will suggest

1. Do nothing; all files up to date
2. Type `update ado`
3. Type `update executable`
4. Type `update all`

(1) is what I hope you will see but, when you are out of date, you will see (2), (3), or (4). Mostly, you will see (2). You will hardly ever see (3) because, whenever there is an executable update, we put out an ado-file update to go with it if for no other reason than to update the help files, so in that case you will see (4).

Type what Stata suggests. I admit I cannot promise you that nothing will go wrong, but I can tell you that since we started this in January of 1999, nothing has ever gone wrong and we do not think anything can go wrong. That is because Stata, behind the scenes, goes through an involved process to obtain updates: it copies the files to a staging area, it verifies that the files look right and that everything is consistent, and only then, when no more communication over the web is required and everything is known to be correct, does Stata actually copy the files to where they affect your installation.

If Stata suggests you type `update ado` and you do that, you will get output that looks like the following:

```
. update ado
(contacting http://www.stata.com)

Ado-file update log
    1.  verifying c:\stata\ado\updates\ is writable
    2.  obtaining list of files to be updated
    3.  downloading relevant files to temporary area
        downloading ksm.ado
        downloading net.hlp
          (output omitted)
        downloading xtregar.ado
        downloading whatsnew.hlp
    4.  examining files
    5.  installing files
    6.  setting last date updated

Updates successfully installed.

Recommendation
    See help whatsnew to learn about the new features
```

If Stata suggests you type `update all`, you will see

```
. update all
```

```
> update ado
(contacting http://www.stata.com)
Ado-file update log
    1.  verifying c:\stata\ado\updates\ is writable
    2.  obtaining list of files to be updated
    3.  downloading relevant files to temporary area
        downloading ksm.ado
        downloading net.hlp
        (output omitted)
        downloading xtregar.ado
        downloading whatsnew.hlp
    4.  examining files
    5.  installing files
    6.  setting last date updated
Updates successfully installed.
Recommendation
    See help whatsnew to learn about the new features
```

```
> update executable
(contacting http://www.stata.com)
Executable update log
    1.  verifying "c:\stata\" is writable
    2.  downloading new executable
New executable successfully downloaded
Instructions
    1.  Exit Stata
    2.  Change to c:\stata\
    3.  Copy wstata.exe to wstata.old
    4.  Copy wstata.bin to wstata.exe
    5.  Try Stata
        Later, erase wstata.old if satisfied
        or copy wstata.old back to wstata.exe
```

`update all` is nothing more than an `update ado` followed by `update executable` and, in fact, you could give the commands separately (and in either order) if you desired. The instructions following `update executable` will vary according to your operating system.

By the way, there is nothing wrong with typing `update ado` or `update executable` or `update all`, even when no update needs be done:

```
. update ado
(contacting http://www.stata.com)
ado-files already up to date
. update executable
(contacting http://www.stata.com)
executable already up to date
```

```
. update all
```

```
> update ado
(contacting http://www.stata.com)
ado-files already up to date
```

```
> update executable
(contacting http://www.stata.com)
executable already up to date
```

3 What's new

After installing an update, you can find out about the fixes and new features by typing `help whatsnew` or by pulling down **Help** and selecting **What's New**.

```
. help whatsnew
```

```
help for whatsnew                                    (manual:  [U] 1.3 What's new)

Additions made to Stata since installation of release 7.0

Update history:
        Stata 7.0 base, 15dec2000
        updated to      24aug2001

This file records the additions and fixes made to Stata, and installed in this
copy of Stata, since the 7.0 release.  (To see a list of the features that were
added to the original 7.0 release see help whatsnew6to7.  To see a list of
additions and fixes for version 6.0 see help whatsnew6.)

Updates are available for free over the Internet.  You can click here to obtain
the latest update, or you can see help updates for more detailed instructions
on obtaining updates.
```

There is a lot more to this file. What comes next is the update log, which lists every change we have made to Stata:

```
Update log  (most recent changes listed first)

Note:  Starred (*) items mean the update was made to the executable.

------ update 24aug2001 ------------------------------------------------------

  1.  On-line help and search index brought up to date for STB Reprints Vol.
      10.  Stata related FAQs found at http://www.ats.ucla.edu/stat/stata/
      have also been added to the search index.

  2.  findit now allows the word "for" to be included among the search words.

  3.  ksm now allows the by() option of graph.
```

```
    4.  xtregar would not report results when the time-variable was included in
        the regression and the lbi option was specified.  xtregar now
        explicitly prohibits the inclusion of the time-variable in the
        regression.

------ update 14aug2001 ------------------------------------------------------

    1.  bs, used with certain user-supplied commands, would sample from the
        entire data set even when some of the observations were not originally
        used.  Now bs checks if the user-supplied command is e-class and, in
        that case, resamples only the observations within the estimation
        subsample and, for other commands, bs displays a warning message.

(output omitted)
```

This file continues for another 12 pages, so I am not going to print it all here. You, however, should look at it, and I want to point out things to note.

Note the first item on the 24aug2001 update:

```
    1.  On-line help and search index brought up to date for STB Reprints Vol.
        10.  Stata related FAQs found at http://www.ats.ucla.edu/stat/stata/
        have also been added to the search index.
```

When you search for a feature or statistic (and we will get to how you do that below), part of the search indexes are on your computer and part are located at *www.stata.com*. Its organization actually does not make much sense and that is something the development group is working on changing, but it really does not matter where the indexes are. What is important to understand is that the *results of your searches will only be complete if your Stata is up to date*.

The next item I want to show you is from the 14aug2001 update:

```
    4.  nlogit reported an "unbalanced data" error when the values of the
        group() variable exceeded float precision.  This has been fixed.
```

This is the most common sort of bug fix: someone asked Stata to do something, it should have done that, but it refused. We get a lot of calls on the technical line with these kinds of problems. These are the kinds of bugs that, if you experienced them, cannot have gone unnoticed. The following, however, appeared in the 15jul2001 update:

```
    3.  reshape long could produce incorrect results when (1) variable names
        for the j() identifier were longer than 8 characters, (2) the j()
        identifiers were string variables (string option specified), and (3)
        you were converting the data from wide to long.  This is fixed.
```

This is an out-and-out bug and, when you see something like this, you want to ask yourself whether this bug could have bitten you and gone unnoticed. You will not see many reports like the one above, and I wish I could tell you that you will never see anything like it. What I can tell you is that, once a bug is reported, we fix it quickly (typically within one or two days), and that is one more reason you should update frequently.

Finally, watch for new features. Sometimes we do not say much:

> 2. Stata also has a new **set varlabelpos** command that controls the positioning of variable labels in the Variables window; see help varlabelpos.

and sometimes we say more:

> 2. findit is a new command that finds and lists sources of information on Stata and Stata commands already installed on your computer or available on the web. **findit** is Stata's most thorough, most complete search command. The results include (1) official help-files installed on your computer, (2) FAQs available at the Stata website, (3) material published in the STB and the Stata Journal, and (4) user-written programs and help files available over the web.

4 Searching

I hope you read the What's New text quoted directly above about `findit`, because that is what I want to tell you about next.

Recently, someone on Statalist asked

> Is there a way of getting a simple confidence interval for the median value of a variable? I've found ways of doing similar things but they're all a lot more complicated. I'm working to a deadline so it'd be very handy not to have to code it myself (being new to Stata that might take a little while).

Eventually the person got an answer, but the questioner could have gotten an answer immediately by simply typing

```
. findit confidence interval for median
                                                        6 Sep 2001 10:19:34
Keyword search

       Keywords:  confidence interval for median
      Criterion:  Select only entries that have ALL the above words (*)
         Search:  (1) Official help files, FAQs, and STBs
                  (2) Web resources from Stata and from other users

       * To search entries that have ANY of the above words, type
         findit confidence interval for median, or

Search of official help files, FAQs, and STBs

[R]     centile  . . . . . .  Report centile and confidence interval
        (help centile)

STB-61  snp16.1  . . . . . .  Update to cendif
        (help cendif if installed) . . . . . R. Newson
        5/01    STB Reprints Vol 10, pages 331--332
        updated for compatibility with Stata 7
```

```
STB-58  snp16  Robust conf. int. for median & percentile dif. between 2 groups
        (help cendif if installed) . . . . . . . . . . . . . . . . . R. Newson
        11/00   STB Reprints Vol 10, pages 324--331
        computes confidence intervals for median differences and
        other percentile differences between values of a variable in
        two samples and is robust to the possibility that the two
        population distributions differ in ways other than location

Web resources from Stata and other users

(contacting http://www.stata.com)

3 packages found (STB omitted)
------------------------------

cid from http://fmwww.bc.edu/RePEc/bocode/c
    'CID': module to calculate confidence intervals for means or differences /
    cid calculates confidence intervals (CIs) for means or / differences. In
    the first form, cid calculates a CI for the mean / of varname. In the
    second form, cid produces a CI for mean / difference between varname1 and

predxcon from http://fmwww.bc.edu/RePEc/bocode/p
    'PREDXCON': module to calculate predicted means, medians, or proportions
    for a continuous X variable / predxcon calculates and prints predicted
    values and 95% / confidence intervals for linear, quantile, or logistic
    model / estimates for a continuous X variable, adjusted for covariates. /

somersd from http://fmwww.bc.edu/RePEc/bocode/s
    'SOMERSD': module to calculate confidence intervals for a range of rank
    order statistics / somers calculates values of Somers' or Kendall's tau-a
    for the / first variable of varlist as a predictor of each of the other /
    variables in varlist, with estimates and jackknife variances / stored as

(end of search)
```

`findit` is Stata's most thorough, most complete search command. If you look in the manuals, you will discover two other search commands: `search` and `net search`. The first command searches the indexes stored on your local computer and the second searches the indexes at *www.stata.com*. You have probably already jumped to the conclusion that `search` searches what is available on your computer and `net search` what you could get over the web. You are wrong. As I mentioned, what each index contains actually does not make much sense and that is something the development group is working on. As things are right now, `search` searches what is on your computer and a part of what is on the web, and `net search` searches the rest of web.

But forget all that and use `findit`. It searches everywhere, and that is what you want, anyway.

About the Author

Allen McDowell is Director of Technical Services at Stata Corporation.