

AUTOCORRELATION FUNCTION IN STATA

Original author: Elizabeth Garrett

No function exists in STATA that makes the autocorrelation scatterplot matrix of residuals, the autocorrelation matrix, or the autocorrelation function. However, I have made a command that does these three things. It is a bit tedious getting the command into STATA, so bear with me...

Setting up the personal ado directory

Some of STATA's commands are called "ado" commands. STATA has two kinds of directories for these commands: a built-in ado directory and a personal ado directory. When STATA is installed on a computer, the built-in directory is created and many STATA commands are put there. The personal ado directory is created by you if/when you want to write your own commands. This directory does NOT exist unless you create it or someone else has created it on your computer (e.g. If you are using a network computer, someone else may have created an ado directory on the c drive. In that case, it would already be there.). So, the first thing you need to do when you get into STATA is type

```
adopath
```

in the command window to see where your personal ado directory should be. The entry next to the [2] is where STATA thinks that your personal ado directory is. This does NOT mean that the directory has been created, it just means that is where STATA will look for it (i.e. it is the "search path"). So, check to see if the directory exists. If it is there, great. If not, create it. For example, on the network computers, the personal ado directory is c:\ado. So, I went into DOS and got to the c:\ prompt. At the c:\ prompt I typed

```
mkdir ado
```

to create my personal ado directory. (To read more about this and about changing where your personal ado file resides, see STATA 5.0 User's Manual Chapter 23.)

Getting the autocorrelation function

The command that has been created for examining autocorrelation is an ado command to be put in your ado directory. The name of the command file is autocor.ado (which is on the LDA class website). You need to copy this file into your personal ado directory (e.g. on the network, this is c:\ado). In addition, it is a good idea to copy it onto a diskette for later use.

Okay, so now everybody has created a personal ado directory in the right place and copied autocor.ado into that directory. Assuming that this is true....

Using the autocorrelation function

Now that autocor.ado is in STATA's personal ado search path, you have the command autocor available to you. The syntax of the command is as follows:

```
autocor y t id
```

where y is the response, t is time, and id is the person/individual identifier. Say, for example, we are interested in the CD4 data and our variable names are cd4cells, year, and person, representing CD4 cell number, years since seroconversion, and person identifier, respectively. To investigate the autocorrelation, we would type

```
autocor cd4cells year person
```

STATA will then:

```
-graph the scatterplot matrix and save it as c:\autoc.gph
-graph the autocorrelation function and save it as c:\acf.gph
-calculate the autocorrelation matrix and print it in the results window
-calculate the autocorrelation under the stationarity assumption and print it in the results window.
```

You will see two items in the results window. The first is the autocorrelation matrix. The second is the correlations under the stationarity assumption, where the row number corresponds to the time lag. In the graph window, two plots will come up at the end. The one on the left is the autocorrelation scatterplot matrix and the one on the right is the autocorrelation function under the stationarity assumption. To look at either graph by itself, in the commands window type

```
graph using c:\autoc
```

or

```
graph using c:\acf
```

This brings graphs back into the STATA graph window after they have been saved.

As an example, I made up a dataset with several missing values where I had response, day and index as the variables. After typing autocor response day index, the following results appeared:

```
. autocor response day index
file temp.dta saved

      | time1   time2   time3   time4
-----+-----
time1 | 1.0000
time2 | 0.9129   1.0000
time3 | -1.0000   0.8783   1.0000
time4 | 0.4697   0.5828   0.8866   1.0000

      acf
1. .8076394
2. .5276556
3. .4696682
```