

Mgmt 469

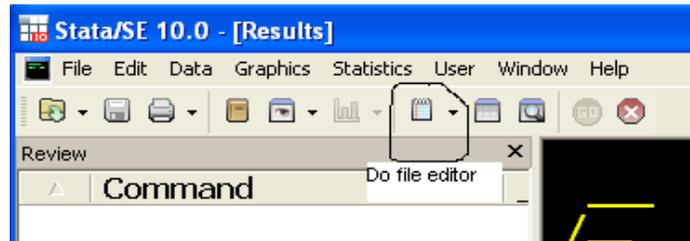
Programming in Stata: Creating do-files

An important feature of any good research project is that the results should be reproducible. One way to make it easy to reproduce your results is to write a set of programs that contain all of your Stata commands. You can even insert comments into the programs to help other researchers (and yourself) follow the thought process.

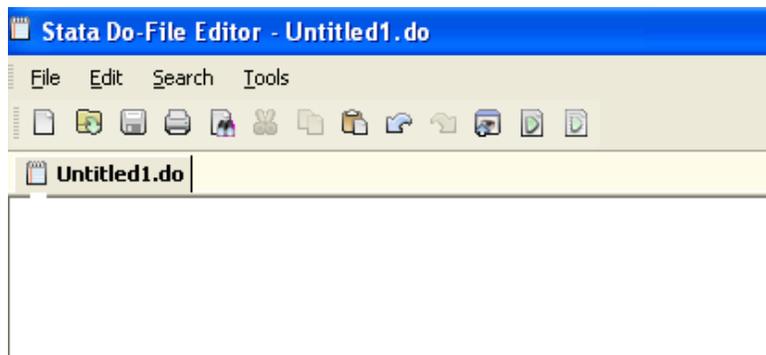
Stata programs are called do-files. They get this name because they have the suffix `.do`. (For example, you might name a program `yogurt.do`.) The do-file contains the Stata commands that you wish to execute. Executing a do-file is the same as executing a series of commands interactively, only you have a permanent record of your commands. This allows you to quickly reproduce work you have already done and go from there. You will never have to make the same mistake twice! Just as important, others can see exactly what you did and build off of your work.

At the end of this note, I will offer advice for moving between interactive work and do-files. Here is how you can create do-files without having to bother with interactive work.

- 1) Open the Stata do-file editor. Click on the button shown below



The do-file editor should open in a new window, with a clean page looking something like this:



Start typing your commands. I suggest starting with:

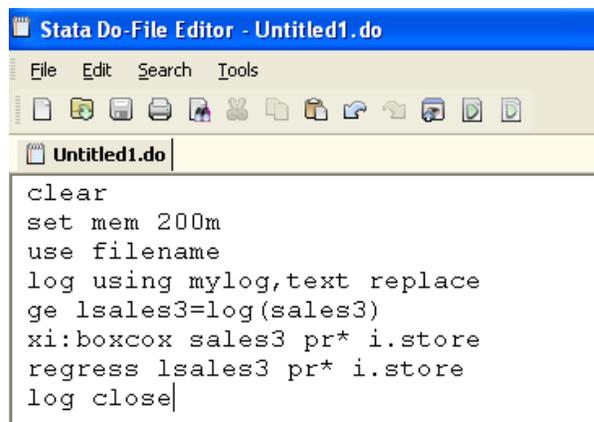
```
clear  
set mem xxm  
use filename  
log using filename, text replace
```

This clears your workspace, frees up memory to speed the calculations, opens your data file, and opens a Stata log. (Be sure to end your do-file with **log close**.)

Now type in more commands (bonus – can you spot the syntax error?¹)

```
clear  
set mem 200m  
use filename  
log using mylog,text replace  
ge lsales3 = log(sales3)  
xi:boxcox sales3 pr* i.store  
regress lsales3 pr* i.store  
log close
```

This will appear as follows in your do-file editor:



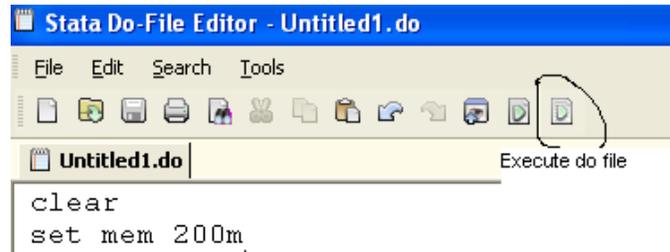
The screenshot shows the Stata Do-File Editor window titled "Stata Do-File Editor - Untitled1.do". The window has a menu bar with "File", "Edit", "Search", and "Tools". Below the menu bar is a toolbar with various icons. The main text area contains the following commands:

```
clear  
set mem 200m  
use filename  
log using mylog,text replace  
ge lsales3=log(sales3)  
xi:boxcox sales3 pr* i.store  
regress lsales3 pr* i.store  
log close|
```

Save your do-file as you would any other file. (Click on the save button and give it a name.) Let's suppose you name the file yogurt.do.

¹ There should be an xi: before regress. Read on to learn how easy it is to fix such mistakes in a do-file. In many applications, this can save a lot of time vis-à-vis working interactively.

You have two options to execute your do-file. First, you can go back to the command window in Stata and type **do yogurt**. Alternatively, you can click on the execute do-file button in the do-file editor (*see picture below*.) Either way, Stata will execute all the commands in the do file. You will see all of these commands executed in the results window. If you create a log file, you can review all of your results at your leisure.



Tips for programming do-files:

- 1) You can continually update your do-file with additional commands. You can try your commands interactively, and if they seem to work, cut and paste them into the do file. (Or use the “save review contents” feature described below.)
- 2) If you can put a * before a line in the do-file, Stata will not execute that line. This serves two different purposes. First, you can rerun your do-file while leaving out certain commands. (Just * the commands you want to skip.) Second, you can annotate your do-file, as shown below.

```
clear
set mem xxm
use filename
log using filename,text replace
*THE NEXT FEW LINES GENERATE LOGS OF VARIABLES
ge lprice3= log(price3)
ge lprice1=log(price1)
*THE NEXT LINE EXECUTES THE BOXCOX TEST
xi:boxcox sales3 pr* i.store
*THE NEXT TWO LINES ARE MY ALTERNATIVE MODELS
xi:regress lsales3 pr* i.store
xi:regress lsales3 price* promo3 i.store
log close
```

3) You can have Stata skip over several lines by using /* and */. The following do-file skips over the boxcox test.

```
clear
set mem xxm
use filename
log using filename,text replace
*THE NEXT FEW LINES GENERATE LOGS OF VARIABLES
ge lprice3= log(price3)
ge lprice1=log(price1)
/*THE NEXT LINE EXECUTES THE BOXCOX TEST
xi:boxcox sales3 pr* i.store*/
*The next two lines are my alternative models
xi:regress lsales3 pr* i.store
xi:regress lsales3 price* promo3 i.store
log close
```

4) If there is a syntax error in your do-file, Stata will stop execution at the point of the error. You can go back to the do-file editor, correct the syntax error, and rerun your program.

5) You may want to create two do-files for any project. The first manipulates the data and creates new variables. At the end of this do-file, be sure to save the resulting data set in a new data file. The second file uses the data set you created in the first file to perform all of your analyses. This will save a lot of time, especially if you have a large data set that requires some time to get into shape prior to analysis.

Getting from Interactive Stata to do-files².

The idea of creating programs is rather daunting and you may prefer to work interactively at first. The good news is that you can save your work into a do-file that is ready for you to use the next time you work in Stata. Here is how:

- 1) Start interactively. Try analyzing your data interactively, as we do in class.
- 2) Right click in the Review window. An intuitive menu will appear. You can use the menu to:
 - Delete commands that you do not want to keep (be sure to highlight these commands before deleting)
 - Highlight the entire contents of the review window (“select all”)
 - Send the highlighted commands to the do-file editor (“send to do-file editor”)
- 3) The do-file editor will open up and contain all the commands that you executed and did not delete
- 4) Edit your do-file Use the Stata do-file editor or even MS Word to edit your file. (If you use MS Word, be sure to save the file as a text file.) You undoubtedly made some mistakes during your interactive session or tried some commands you would rather leave out next time. Now is the time to delete them from your do-file. This is also a good time add comments (with a *).
- 4) Execute your do-file Type **do filename** in the Stata Command window or click on the execute do-file button from the Stata do-file editor.

Common problems with do-files:

- You forget to **clear** at the start of the file (far and away the biggest mistake)
- You forget to close the log file (a strong number two)
- Data sets and do-files are not in the same folder (relatively rare)
- You save a variable into a data set that already has that variable. Be sure to use **replace** rather than save
- You try to merge data sets but do not drop `_merge` before doing so.

² Some of this content is drawn from http://slack.ser.man.ac.uk/progs/stata/do_files.html