Intermediate Stata
February 5, 2007

Presented by the
Center for Family and
Demographic Research

Overview

- CFDR Resources
- Log file
- Variable creation
- Appending/merging data sets
- Examples

- Note: In this presentation, “Handout” refers to a document that can be accessed at

http://www.bgsu.edu/organizations/cfdr/dataresearch/programming_research/workshop_handouts/STATA%20Introduction06.pdf

What resources are available from the CFDR?

- CFDR Stata User Guide
  - http://www.bgsu.edu/organizations/cfdr/dataresearch/programming_research/stata.htm
- Stata annotated output
  - http://www.bgsu.edu/organizations/cfdr/dataresearch/programming_research.htm

Log File

- How do I keep a record of my analyses?

Log File (pg. 2 handout)

- The log file is like SAS's log and output windows combined
- Information can be copied from the log and pasted into other programs (Word, Excel, etc.)
- If you want to save your log, you must start the log file at the beginning of your session

Log File (pg. 2 handout)

To start a log…

- Drop down menus
  - File → Log → Begin

- Do-file editor/Command window
  - log using "path\filename.smcl"

Log File (pg. 2 handout)

To “pause” your log…

- Drop down menus
  - File → Log → Suspend

- Do-file editor/Command window
  - log off

Log File (pg. 2 handout)

To resume your log

- Drop down menus
  - File → Log → Resume

- Do-file editor/Command window
  - log on

Log File (pg. 2 handout)

To turn off log completely

- Drop down menus
  - File → Log → Close

- Do-file editor/Command window
  - log close

Variable Creation

- How do I create new variables, recode, etc.?

Variable Creation (pgs. 2-3 handout)

Generate or generate, replace

- generate newvar1 = iv1*iv2

- gen male=0
- replace male=1 if sex==1

Variable Creation (pgs. 2-3 handout)

Macro variables

- global ivset1 “nhblack nhother hisp male ses educ”
- reg dv $ivset
- reg dv $ivset delinq depress effic

Variable Creation (pgs. 2-3 handout)

"If…then" statements

- Unlike SAS, in Stata, the "then" condition comes first…
  - generate newvar =1 if sex==1
  - replace mysample =1 if age>=16

Variable Creation (pg. 5 handout)

Interaction expansion (from Stata help menu):

- xi i.*var*
  - Takes categorical variable *var*, and creates dummy variables
  - _Ivar_2, _Ivar_3, etc.
  - Lowest value is omitted category (unless you change default)

Variable Creation (pg. 5 handout)

Interaction expansion (from Stata help menu):

i.varname	Creates dummies for *categorical* variable varname.
i.varname1*i.varname2 and	Creates dummies for *categorical* variables varname1 varname2: and all interactions and main effects
i.varname1*varname3	Creates dummies for *categorical* variable varname1 and *continuous* variable varname3: all interactions and main effects.
i.varname1|varname3	Creates dummies for *categorical* variable varname1 and *continuous* variable varname3: all interactions and main effect of varname3, but no main effect of varname1.

Variable Creation (pg. 5 handout)

Interaction expansion (from Stata help menu):

- xi: *any stata command var i.var*
  - e.g. xi: mean var1 i.var2
  - executes the specified command with the expanded terms

Appending/Merging

- How do I combine data sets?

Appending/Merging

- Append

var1 var2 var3

R1

R2

R3

Adding cases ↓

- Merge

var1 var2 var3

R1

R2

R3

Adding variables →

Appending data (pg. 4 handout)

- Adding more observations is just like stacking two datasets together
  - Example: Same survey questions asked of males and females, but they are kept in two separate files. To “add” these data sets together…

```
use male
append using female
```

Merging data (pg. 4 handout)

- Need common variable
- Sort both data sets by common variable
- _merge variable