

Recent Advances in the Field of Trade Theory and Policy Analysis Using Micro-Level Data

July 2012

Bangkok, Thailand

Cosimo Beverelli and Nadia Rocha
(World Trade Organization)

Descriptive statistics and graphs in Stata

Descriptive statistics

- Stata's summarize command produces basic descriptive statistics
 - *summarize var1 var2 var3 ...*
- Stata's tabulate command gives information on categories within categorical variables
 - *tabulate var1*
- Stata's correlate command produces a correlation matrix for the listed variables
 - *correlate var1 var2 var3 ...*
- More commands and examples [here](#)

Basic graphs

- Use the histogram command to produce a histogram of the listed variable
 - *histogram var1*
- Stata's *kdensity* command can be used to produce a kernel density plot of the listed variable
 - *kdensity var1*
- The *twoway* command can be used to produce combinations of graphs of various types
 - *twoway (kdensity var1) (kdensity var2)*
 - *twoway (scatter var1 var2) (lfit var1 var2)*
- See also [this](#) and [this](#)

The *if* command

- Descriptive statistics and graphs can both be made more interesting by using the *if* command to limit the sample:
 - *summarize var1 if dummy1==1*
 - *summarize var1 if dummy1~=1*
 - *twoway (kdensity var1 if dummy1==1) (kdensity var1 if dummy1==0)*
- The regression commands examined in subsequent sessions can also be applied to limited samples using *if*