Descriptive Statistics and Graphs in Stata

Ben Shepherd
Principal, Developing Trade Consultants
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 ...`
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 (kdesnity var1) (kdensity var2)
} Twoway (scatter var1 var2) (lfit var1 var2)
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.