What is econometric analysis & why do we do it?

Inspecting & summarizing data
What is econometric analysis & why do we do it?
What is econometrics?

- **Statistical methods for:**
  - Estimating economic relationships
  - Testing economic theories
  - Evaluating policies and programs

- **Why separate discipline from stats?**
  - Non-experimental/observational data
Steps in econometric analysis

1. Research question(s)
2. Economic model or other conceptual/theoretical framework
3. Operationalize #2 → econometric model
4. Specify hypotheses to be tested in #3
5. Collect data; create variables
6. Inspect & summarize data
7. Estimate econometric model
8. Interpret results; hypothesis testing & statistical inference

*We will focus on 6-8*
The structure of economic data

- **4 common data structures:**
  1. Cross-sectional
  2. Time series
  3. Pooled cross sections
  4. Panel or longitudinal data

- **Begin with cross-sectional because:**
  1. Fewest conceptual & technical difficulties (random sampling)
  2. Illustrates most of key themes
Causality & ceteris paribus effects

- **Policy analysis**: measure causal effect of policy change on outcome variable (HH income)

- **Ceteris paribus**
  - If other factors NOT held fixed, cannot measure the causal effect
  - Impossible to hold ALL else equal
  - Held ENOUGH fixed to make case for causality?
  - Challenge with non-experimental data
Inspecting & summarizing data PRIOR to estimation
Why inspect & summarize data BEFORE estimation?

- Familiarize yourself with:
  - Dataset structure
  - Units
  - Level (binary/categorical/ordinal/continuous)
  - Range of values
  - Prices real or nominal? Base year?
  - Rates: proportion or percentage?

- User missing (−9) vs. sysmis (.)

- Check for data problems, outliers
Outliers

• **Reasons for outliers:**
  1. Enumerator or data entry (or respondent!) error
  2. Sampling from small population & some members very different

• **Why worry?**
  - Don’t want to use erroneous data
  - Influen
tial observations: outliers can have big impact on regression results
Examples of summary stats tables

See Summary_stat_tables_examples_nm1.doc
Key commands in Stata

- **describe**: describes the dataset
- **tabulate**: frequency table
  - Crosstabs if **tab** var1 var2
- **tab1**: one-way frequency table(s)
- **list**: lists values of variable(s)
- **summarize**: summary stats
- **tabstat**: summary stats but with more user control over which stats
- **histogram**

SEE help command FOR LIST OF OPTIONS
REFERENCES (**main reference)**


*Session materials prepared by Nicole Mason with input from Bill Burke. November 2011. masonn@msu.edu*