Ch.1 Econometrics & Data

1. What is Econometrics?
2. Steps in empirical economic analysis
3. The structure of economic data
4. Causality and the notion of ceteris paribus in econometric analysis

1.1 What is Econometrics?

- Econometrics is a kind of statistical methods for
  - Estimating economic relationships,
  - Testing economic theories, and
  - Evaluating & implementing policy

  etc…

Cont. What is Econometrics?

- Econometrics has borrowed a lot of techniques from mathematical statistics, but
  - Having experimental data is ***rare*** in economics
  - It is necessary to use non-experimental data to make inferences
  - It is important to be able to apply economic theory to real world data

1.2 Steps in Empirical Analysis

- An empirical analysis uses data to test a theory or to estimate a relationship

1. Constructing *economic model*

   $wage = f(educ, exper, training)$ (1.2)

2. Specifying *econometric model*

   $wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 training + u$ (1.4)

   - u is error term, and βs are parameters.

1.3 Structure of economic data

Cross-sectional data

- Each observation is a new individual, firm, etc. with information at a point in time.

  →e.g. table 1.1 & 1.2

- Cross-sectional data is a ***random sample***.
  - If the data is not a random sample, we have a sample-selection problem.

TABLE 1.1

A Cross-Sectional Data Set on Wages and Other Individual Characteristics

obsno	wage	educ	exper	female	married
1	3.10	11	2	1	0
2	3.24	12	22	1	1
3	3.00	11	2	0	0
4	6.00	8	44	0	1
5	5.30	12	7	0	1
⋮	⋮	⋮	⋮	⋮	⋮
525	11.56	16	5	0	1
526	3.50	14	5	1	0

TABLE 1.2

A Data Set on Economic Growth Rates and Country Characteristics

obsno	country	gpcrgdp	govcons60	second60
1	Argentina	0.89	9	32
2	Austria	3.32	16	50
3	Belgium	2.56	13	69
4	Bolivia	1.24	18	12
⋮	⋮	⋮	⋮	⋮
61	Zimbabwe	2.30	17	6

Special Lecture 7

Cont. Structure of economic data

Time series data

- Time series data has a separate observation for each time period – e.g. table 1.3.
- Since *rarely a random sample*, we have to consider different problems about it.
- It has some category depending on time frequency.

Special Lecture 8

TABLE 1.3

Minimum Wage, Unemployment, and Related Data for Puerto Rico

obsno	year	avgmin	avgcov	unemp	gnp
1	1950	0.20	20.1	15.4	878.7
2	1951	0.21	20.7	16.0	925.0
3	1952	0.23	22.6	14.8	1015.9
⋮	⋮	⋮	⋮	⋮	⋮
37	1986	3.35	58.1	18.9	4281.6
38	1987	3.35	58.2	16.8	4496.7

Special Lecture 9

Cont. Structure of economic data

Pooled cross section

- It can pool random cross sections to account for time differences –e.g. table 1.4.

Panel data

- It consists of the same random individual observations over time –e.g. table 1.5.

Special Lecture 10

TABLE 1.4

Pooled Cross Sections: Two Years of Housing Prices

obsno	year	hprice	proptax	sqrft	bdrms	bthrms
1	1993	85500	42	1600	3	2.0
2	1993	67300	36	1440	3	2.5
3	1993	134000	38	2000	4	2.5
⋮	⋮	⋮	⋮	⋮	⋮	⋮
250	1993	243600	41	2600	4	3.0
251	1995	65000	16	1250	2	1.0
252	1995	182400	20	2200	4	2.0
253	1995	97500	15	1540	3	2.0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
520	1995	57200	16	1100	2	1.5

Special Lecture 11

TABLE 1.5

A Two-Year Panel Data Set on City Crime Statistics

obsno	city	year	murders	population	unem	police
1	1	1986	5	350000	8.7	440
2	1	1990	8	359200	7.2	471
3	2	1986	2	64300	5.4	75
4	2	1990	1	65100	5.5	75
⋮	⋮	⋮	⋮	⋮	⋮	⋮
297	149	1986	10	260700	9.6	286
298	149	1990	6	245000	9.8	334
299	150	1986	25	543000	4.3	520
300	150	1990	32	546200	5.2	493

Special Lecture 12

1.4 Causality & Ceteris Paribus

- Economist's goal is to infer that one variable has a ***causal effect*** on another variable, for testing economic theory or for evaluating policy.
  - **Causal effect:** A *ceteris paribus* change in one variable has an effect on another variable.
  - **Ceteris paribus:** All other relevant factors are held fixed.

Example: Returns to Education

- A model of human capital investment implies getting more education should lead to higher earnings
- In the simplest case, this implies an equation like

$$Earnings = \beta_0 + \beta_1 education + u$$

Example *cont.*

- The error term, u, includes other factors affecting earnings, like experience or job training.
- The estimate of β_1 is the return to education.

$$\beta_1 = \frac{\partial Earnings}{\partial education} \quad s.t. \frac{\partial u}{\partial education} = 0$$

Causality & Ceteris Paribus *cont.*

- Simply establishing a relationship between variables is rarely sufficient.
- If we've truly controlled for enough other variables, then the estimated ceteris paribus effect can often be considered to be causal.
- ***Econometric methods can simulate a ceteris paribus experiment.***