Class 2 – What is Econometrics?

Pedram Jahangiry

JON M.
HUNTSMAN
SCHOOL OF BUSINESS
UtahStateUniversity

What is econometrics?

Econometrics is the branch of economics that:

develops and uses **statistical methods** for **estimating** economic **relationships**

Typical goals of econometric analysis

❑ Estimating relationships between economic variables

❑ Testing economic theories and hypotheses

❑ Forecasting economic variables

Steps in econometric analysis

1) Specifying the models (Economic vs Econometric models)

2) Collecting the data needed to quantify the models

3) Quantfiying the models with the data

❑ **Economic models** (mostly skipped!)
- Often use optimizing behaviour, equilibrium modeling, …
- Establish relationships between economic variables

❑ **Econometric models**
- Quantify / Test relationships between economic variables
- Attempt to quantitatively bridge gap between economic theory and the real world.

Different kinds of economic data sets

Econometric analysis requires data

❑ Cross-sectional data
❑ Time series data
❑ Panel/Longitudinal data
❑ Pooled cross sections

Cross-sectional data sets

- Sample of individuals, households, firms, cities, states, countries, or other units of interest at a given point of time/in a given period

- Cross-sectional observations are more or less independent

- For example, pure random sampling from a population

- **A Key feature:** Ordeding of the data does NOT matter!

Cross-sectional data set on wages and other 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 1.1 A Cross-Sectional Data Set on Wages and Other Individual Characteristics

© Cengage Learning, 20

Indicator variables
(1 = yes, 0 = no)

Observation number

Hourly wage

Years of education

Years of experience

Time series data

- Observations of a variable or several variables over time
- For example, stock prices, consumer price index, gross domestic product, automobile sales, …
- Time series observations are typically serially correlated
- Data frequency: daily, weekly, monthly, quarterly, annually, …
- Typical features of time series: trends and seasonality

- **A Key feature:** Ordeding of the data DOES matter! Past events can influence future events!

Time series data on minimum wages and related variables

_____

TABLE 1.3   Minimum Wage, Unemployment, and Related Data for Puerto Rico

obsno	year	avgmin	avgcov	prunemp	prgnp
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

© Cengage Learning, 2016

Average minimum wage for the given year

Average coverage rate

Unemployment rate

Gross national product

US data, minimum wage vs unemployment!

Panel or longitudinal data

- The **same** cross-sectional units are followed over time

- Panel data have a cross-sectional and a time series dimension

- Ordering in the cross section of a panel data does not matter

Two-year panel data on city crime statistics

TABLE 1.5 A Two-Year Panel Data Set on City Crime Statistics

obsno	city	year	murders	population	unem	police
1	1	1986	5	350,000	8.7	440
2	1	1990	8	359,200	7.2	471
3	2	1986	2	64,300	5.4	75
4	2	1990	1	65,100	5.5	75
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
297	149	1986	10	260,700	9.6	286
298	149	1990	6	245,000	9.8	334
299	150	1986	25	543,000	4.3	520
300	150	1990	32	546,200	5.2	493

Each city has two time series observations

Number of police in 1986

Number of police in 1990

© Cengage Learning, 2016

Pooled cross sections

- Two or more cross sections are combined in one data set
- Cross sections are drawn independently of each other
- Pooled cross sections often used to evaluate policy changes

- Example: Drop in property taxes in 1994
  - Evaluate effect of change in property taxes on house prices
  - Random sample of house prices for the year 1993
  - A new random sample of house prices for the year 1995
  - Compare before/after (1993: before reform, 1995: after reform)

Pooled cross sections on housing prices

Property tax

Size of house in square feet

Number of bedrooms

Number of bathrooms

obsno	year	hprice	proptax	sqrft	bdrms	bthrms
1	1993	85,500	42	1600	3	2.0
2	1993	67,300	36	1440	3	2.5
3	1993	134,000	38	2000	4	2.5
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
250	1993	243,600	41	2600	4	3.0
251	1995	65,000	16	1250	2	1.0
252	1995	182,400	20	2200	4	2.0
253	1995	97,500	15	1540	3	2.0
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
520	1995	57,200	16	1100	2	1.5

TABLE 1.4 Pooled Cross Sections: Two Years of Housing Prices
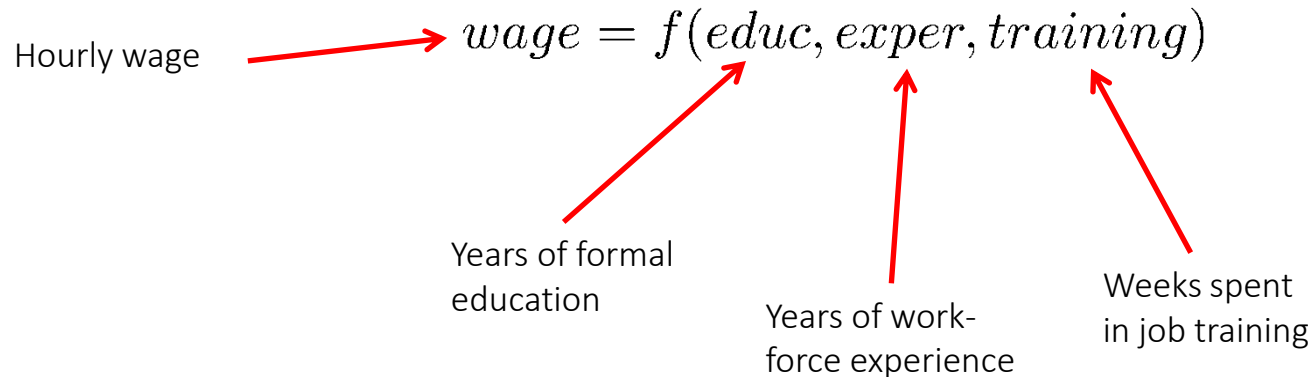
Before reform

After reform

© Cengage Learning, 2016

Model of job training and worker productivity

- What is effect of additional training on worker productivity?
- Formal economic theory not really needed to derive equation:

Hourly wage

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

Years of formal education

Years of work-force experience

Weeks spent in job training

- Other factors may be relevant, but these are the most important

Econometric model of job training and worker productivity

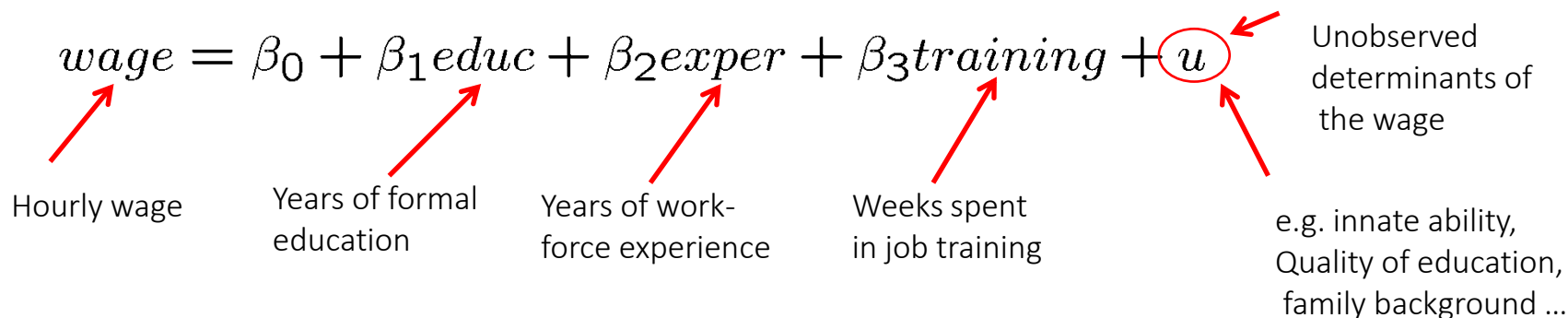$$wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 training + u$$

Hourly wage

Years of formal education

Years of work-force experience

Weeks spent in job training

Unobserved determinants of the wage

e.g. innate ability, Quality of education, family background ...

- Most of econometrics deals with the specification of the error u
- Econometric models may be used for hypothesis testing

  - For example, the parameter β_3 represents "effect of training on wage"
  - How large is this effect? Is it different from zero?

Causality and the notion of ceteris paribus

Definition of causal effect of x on y :

"How does variable y change if variable x is changed
but all other relevant factors are held constant"

❑ Most economic questions are ceteris paribus questions

❑ In reality it will not be possible to literally hold all else equal (Observational vs Experimental data)

❑ Econometrics analysis allows us to address ceteris paribus questions.

Correlation VS. Causation

❑ Regression results cannot prove causality!

❑ For example, if variables Y and X are correlated statistically, then:

- – Y might "cause" X.
- – X might "cause" Y.
- – Some third factor might "cause" both.
- – The relationship might have happened by chance : Spurious correlation

Results from regression analyses always should be viewed with some caution!

Could a taller FED chair mean rates rise?

Source: LPL Research, Bloomberg 10/22/17

We don't actually believe that interest rates are determined by the height of the Fed chair, but it has been an interesting coincidence.