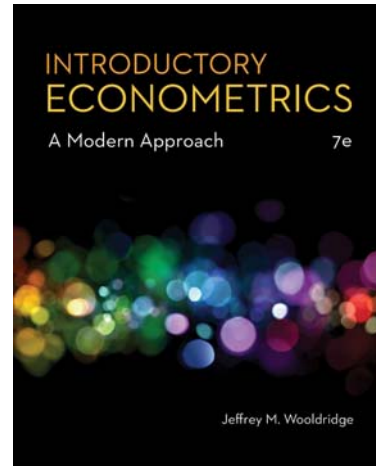


Chapter 1

The Nature of Econometrics and Economic Data



The Nature of Econometrics and Economic Data (1 of 22)

- **What is econometrics?**
 - Econometrics is the use of statistical methods to analyze economic data.
 - Econometricians typically analyze nonexperimental data.
- **Typical goals of econometric analysis:**
 - Estimating relationships between economic variables.
 - Testing economic theories and hypotheses.
 - Evaluating and implementing government and business policy.
- **Common applications**
 - Forecasting macroeconomic variables (interest rates, inflation rates, GDP).
 - Forecasting non-macro variables (less visible).

The Nature of Econometrics and Economic Data (2 of 22)

- **Steps in econometric analysis**
 - 1) Economic model (this step is often skipped)
 - 2) Econometric model

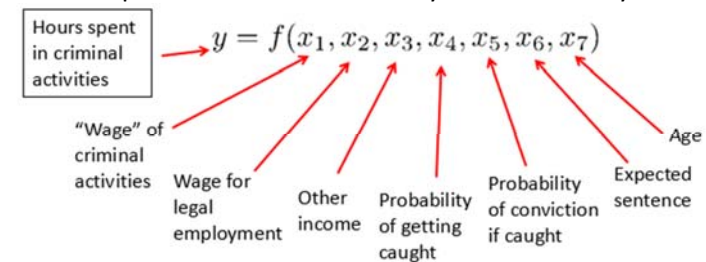
• **Economic models**

- Maybe micro- or macromodels
- Often use optimizing behaviour, equilibrium modeling, ...
- Establish relationships between economic variables
- Examples: demand equations, pricing equations, ...

The Nature of Econometrics and Economic Data (3 of 22)

• **Economic model of crime (Becker (1968))**

- Derives equation for criminal activity based on utility maximization.

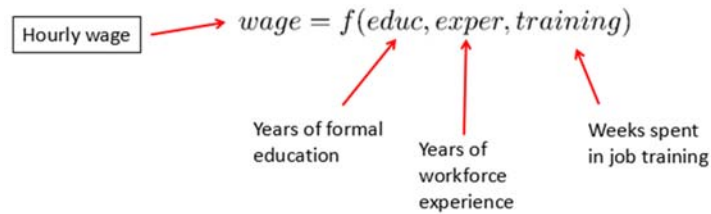


- Functional form of relationship not specified.
- Equation could have been postulated without economic modeling.

The Nature of Econometrics and Economic Data (4 of 22)

- **Model of job training and worker productivity**

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

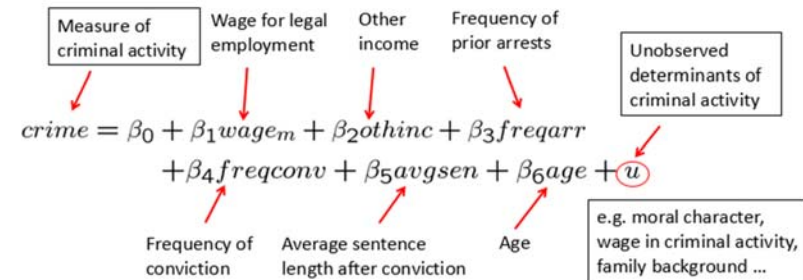


- Other factors may be relevant, but these are the most important.

The Nature of Econometrics and Economic Data (5 of 22)

- **Econometric model of criminal activity**

- The functional form has to be specified.
- Variables may have to be approximated by other quantities.



The Nature of Econometrics and Economic Data (6 of 22)

- **Econometric model of job training and worker productivity**



- 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 the effect of training on wages.
 - How large is this effect? Is it different from zero?

The Nature of Econometrics and Economic Data (7 of 22)

- **Econometric analysis requires data.**

- There are several different kinds of economic data sets:

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

- Econometric methods depend on the nature of the data used.

- Use of inappropriate methods may lead to misleading results.

The Nature of Econometrics and Economic Data (8 of 22)

- **Cross-sectional data sets**

- These may include samples of individuals, households, firms, cities, states, countries, or other units of interest at a given point of time or in a given period.
- Cross-sectional observations are more or less independent.
- An example is pure random sampling from a population.
- Sometimes pure random sampling is violated, for example, people refuse to respond in surveys, or sampling may be characterized by clustering.
- Cross-sectional data is typically encountered in applied microeconomics.

The Nature of Econometrics and Economic Data (9 of 22)

- Table 1.1: 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 |

Indicator variables
(1 = yes, 0 = no)

Observation number Hourly wage Years of education Years of experience

The Nature of Econometrics and Economic Data (10 of 22)

- Table 1.2: Cross-sectional data on 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 |

Average growth rate of real per capita GDP

Government consumption as a percentage of GDP

Adult secondary education rates

The Nature of Econometrics and Economic Data (11 of 22)

- **Time series data**

- This includes observations of a variable or several variables over time.
- Examples include stock prices, money supply, consumer price index, gross domestic product, annual homicide rates, automobile sales, and so on.
- Time series observations are typically serially correlated.
- Ordering of observations conveys important information.
- Data frequency may include daily, weekly, monthly, quarterly, annually, and so on.
- Typical features of time series include trends and seasonality.
- Typical applications include applied macroeconomics and finance.

The Nature of Econometrics and Economic Data (12 of 22)

- Table 1.3: Time series data on 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 |

© 2020 Cengage

Average minimum wage for the given year

Average coverage rate

Unemployment rate

Gross national product

The Nature of Econometrics and Economic Data (13 of 22)

- 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 are often used to evaluate policy changes.
- Example:
 - Evaluating 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).

The Nature of Econometrics and Economic Data (14 of 22)

- Table 1.4: Pooled cross sections on two years of housing prices

| obsno | year | hprice | proptax | sqft | 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 |

© 2020 Cengage

Property tax

Size of house in square feet

Number of bedrooms

Number of bathrooms

Before reform

After reform

The Nature of Econometrics and Economic Data (15 of 22)

- Panel or longitudinal data**

- The same cross-sectional units are followed over time.
- Panel data have a cross-sectional and a time series dimension.
- Panel data can be used to account for time-invariant unobservables.
- Panel data can be used to model lagged responses.
- Example:
 - City crime statistics; each city is observed in two years.
 - Time-invariant unobserved city characteristics may be modeled.
 - Effect of police on crime rates may exhibit time lag.

The Nature of Econometrics and Economic Data (16 of 22)

- Table 1.5: 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

© 2020 Cengage

The Nature of Econometrics and Economic Data (17 of 22)

- **Causality and the notion of ceteris paribus**

Definition of causal effect of x on y :

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

- Ceteris paribus: “other relevant factors being equal.”
- Most economic questions are ceteris paribus questions.
- It is important to define which causal effect one is interested in.
- It is useful to describe how an experiment would have to be designed to infer the causal effect in question.

The Nature of Econometrics and Economic Data (18 of 22)

- **Causal effect of fertilizer on crop yield**

- “By how much will the production of soybeans increase if one increases the amount of fertilizer applied to the ground.”
- Implicit assumption: all other factors that influence crop yield such as quality of land, rainfall, presence of parasites, and so on are held fixed.
- Experiment = Feasible
 - Choose several one-acre plots of land; randomly assign different amounts of fertilizer to the different plots; compare yields.
 - Experiment works because amount of fertilizer applied is unrelated to other factors influencing crop yields.

The Nature of Econometrics and Economic Data (19 of 22)

- **Measuring the return to education**

- “If a person is chosen from the population and given another year of education, by how much will his or her wage increase?”
- Implicit assumption: all other factors that influence wages such as experience, family background, intelligence, and so on are held fixed.
- Experiment \neq Infeasible
 - Choose a group of people; randomly assign different amounts of education to them (infeasible!); compare wage outcomes.
 - Problem without random assignment, amount of education is related to other factors that influence wages (e.g. intelligence).

The Nature of Econometrics and Economic Data (20 of 22)

- **Effect of law enforcement on city crime level**
 - “If a city is randomly chosen and given ten additional police officers, by how much would its crime rate fall?”
 - Alternatively: “If two cities are the same in all respects, except that city A has ten more police officers than city B, by how much would the two cities’ crime rates differ?”
- Experiment ≠ Infeasible
 - Randomly assign number of police officers to a large number of cities (virtually impossible, as no two cities are alike in all respects except size of police force!).
 - More importantly, in reality, the number of police officers occurs contemporaneously with determination of crime rate.

The Nature of Econometrics and Economic Data (21 of 22)

- **Effect of the minimum wage on unemployment**
 - “By how much (if at all) will unemployment increase if the minimum wage is increased by a certain amount (holding other things fixed)?”
- Experiment ≠ Infeasible
 - Government randomly chooses minimum wage each year and observes unemployment outcomes.
 - Experiment would theoretically work because level of minimum wage is unrelated to other factors determining unemployment.
 - In reality, the level of the minimum wage will depend on political and economic factors that also influence unemployment.

The Nature of Econometrics and Economic Data (22 of 22)

- Testing predictions of economic theories
 - Economic theories are not always stated in terms of causal effects.
 - For example, the expectations hypothesis states that long-term interest rates equal compounded expected short-term interest rates.

$$(1+r_{tt})^n = (1+r_{year1}^e)(1+r_{year2}^e)\cdots(1+r_{yearn}^e)$$

- An implication is that the interest rate of a three-month T-bill should be equal to the expected interest rate for the first three months of a six-month T-bill; this can be tested using econometric methods.