

# Introductory Econometrics (EE325)

Instructor: Assistant Professor Dr. Kaewkwan Tangtipongkul

Office: Room 458, 4th Floor (Prachan Campus)

Office hours: TBA

Email: [kaewkwan@econ.tu.ac.th](mailto:kaewkwan@econ.tu.ac.th)

# Course Goals

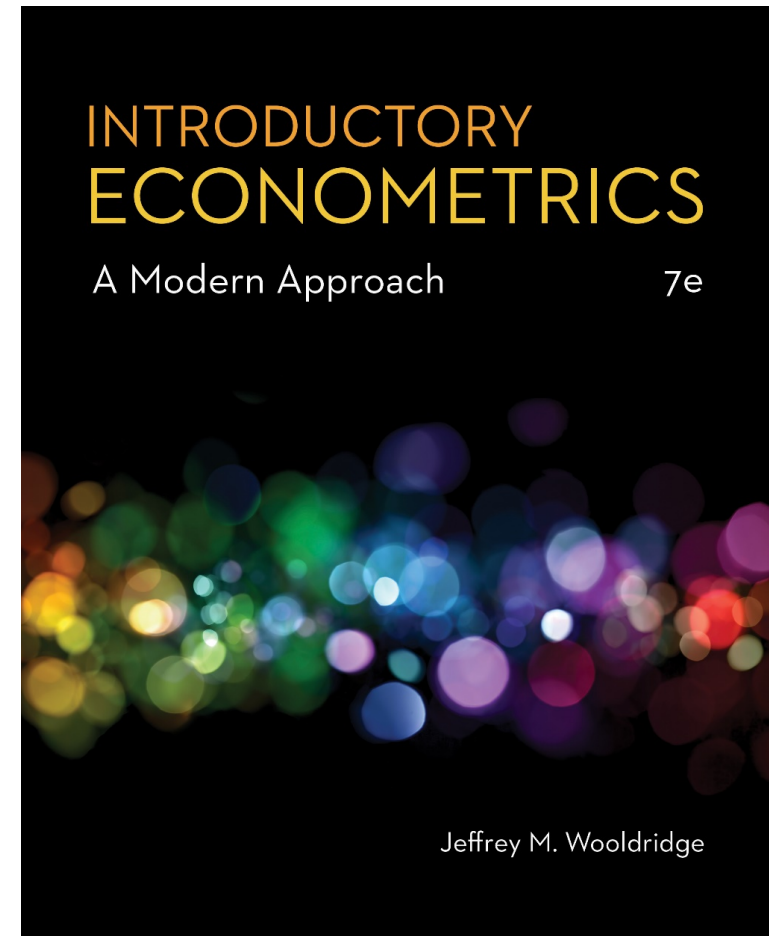
To provide students an understanding of why econometrics is necessary and a working ability with basic econometric tools such that:

- Students can apply basic econometric tools to estimation, inference, and forecasting in the context of real world economic problems
- Students understand how to process information from a sample of economic data
- Students can read critically the results and conclusions
- Students have a foundation for further study of advance econometrics.

The image part with relationship ID r1018 was not found in the file.

# 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**

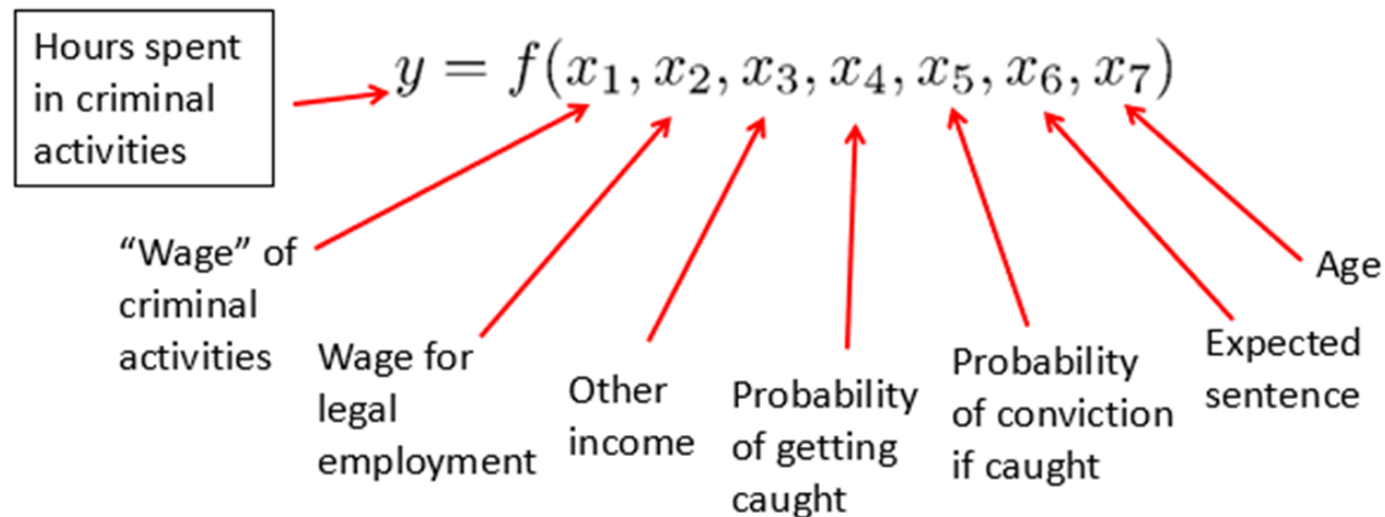
- 1) Economic model (this step is often skipped)
- 2) Econometric model

- **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)

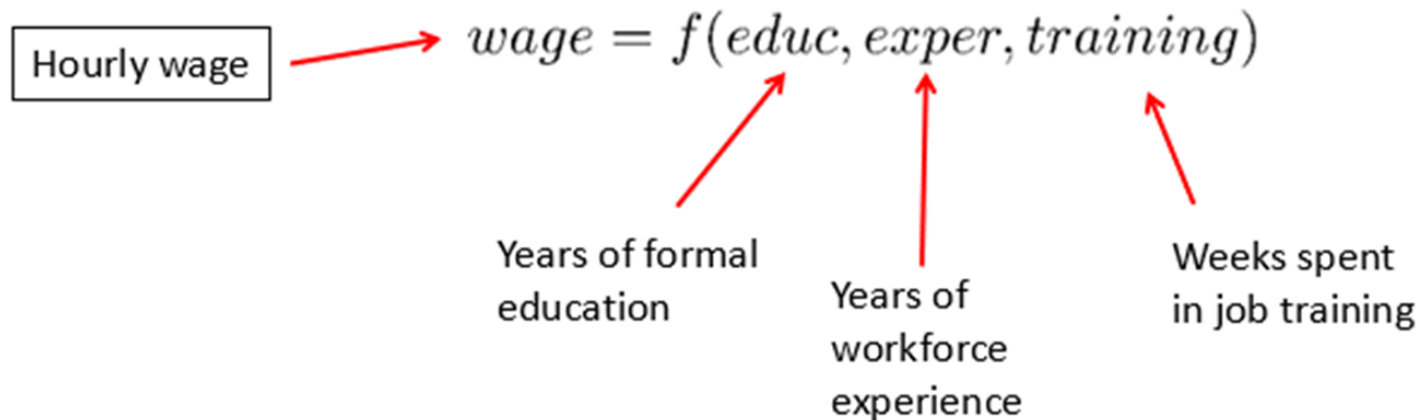
- **Economic model of crime (Becker (1968))**
  - Derives equation for criminal activity based on utility maximization.



- 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.

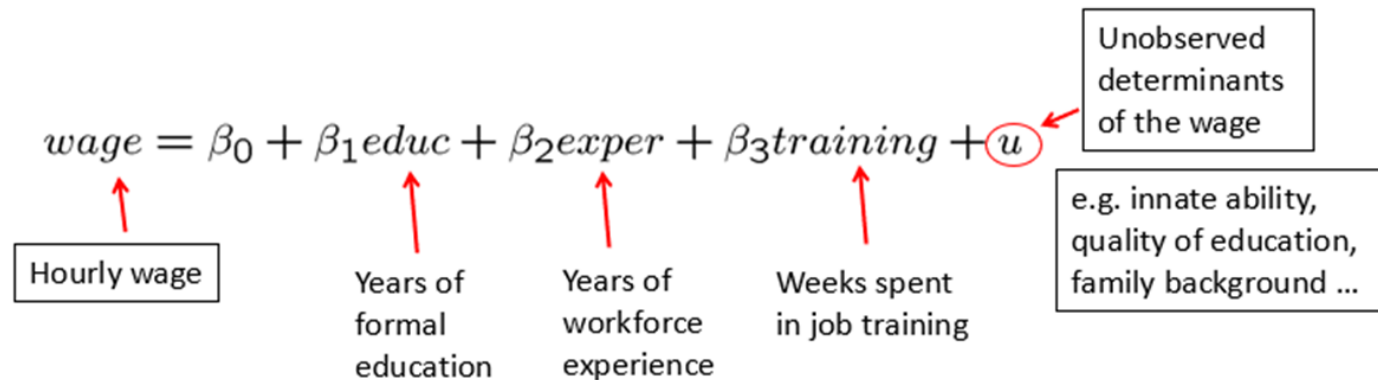
- **Econometric model of criminal activity**

- 
- The diagram illustrates a regression model for criminal activity. The dependent variable is *crime*. The model is represented by the equation:
- $$crime = \beta_0 + \beta_1 wage_m + \beta_2 othinc + \beta_3 freqarr + \beta_4 freqconv + \beta_5 avglsen + \beta_6 age + u$$
- Red arrows point from descriptive labels to the variables in the equation:
- Measure of criminal activity** points to *crime*.
  - Wage for legal employment** points to  $\beta_1 wage_m$ .
  - Other income** points to  $\beta_2 othinc$ .
  - Frequency of prior arrests** points to  $\beta_3 freqarr$ .
  - Frequency of conviction** points to  $\beta_4 freqconv$ .
  - Average sentence length after conviction** points to  $\beta_5 avglsen$ .
  - Age** points to  $\beta_6 age$ .
  - Unobserved determinants of criminal activity** points to the error term *u*.
- A box at the bottom right provides examples for the unobserved determinants: "e.g. moral character, wage in criminal activity, family background ...".



## 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  $\beta_3$  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

© 2020 Cengage

## 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        |

© 2020 Cengage

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

Property tax

Size of house in square feet

Number of bedrooms

Number of bathrooms

Before reform

After reform

| 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    |

© 2020 Cengage

## 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  $\neq$  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  $\neq$  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.

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

- 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.

# Methodology of Econometrics

## **8 Steps in Empirical Economic Analysis**

1. Statement of Theory or Hypothesis
2. Specification of Mathematical Model of the Theory
3. Specification of the statistical, or econometric, model
4. Obtaining the data
5. Estimation of the parameters of the econometric model
6. Hypothesis testing
7. Forecasting or prediction
8. Using the model for control or policy purposes

# Statement of Theory or Hypothesis

## Example

Keynes postulated that the marginal propensity to consume (MPC), the rate of change of consumption for a unit change in income, is greater than zero but less than 1

# Specification of Mathematical Model of the Theory

$$Y = \beta_1 + \beta_2 X \quad 0 < \beta_2 < 1$$

Where      Y = consumption expenditure  
              X = income

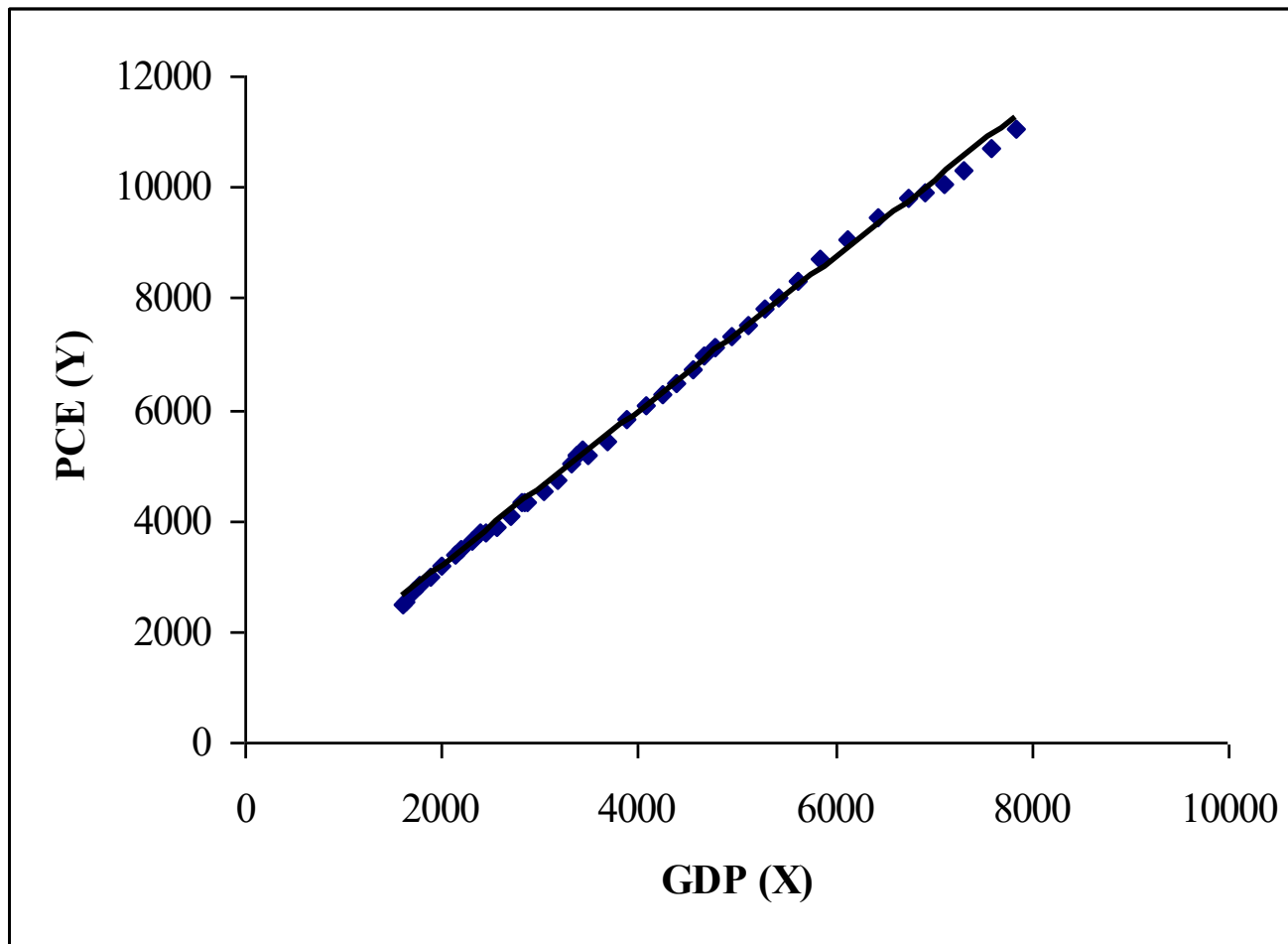
# Specification of the statistical, or econometric, model

$$Y = \beta_1 + \beta_2 X + u$$

$u$  known as the disturbance or error term

# Obtaining the data

- Aggregate personal consumption expenditure and gross domestic product from period 1960-2005 in the United States measured in billion dollars.
- The data are in real terms. They are measured in constant (2000) prices



# Estimation of the parameters of the econometric model

$$\hat{Y}_t = -299.5913 + 0.7218X_t$$

MPC was about 0.72 – For the sample period an increase in real income of one dollar led, on average, to an increase of about 72 cents in real consumption expenditure



# Hypothesis testing (Statistic Inference)

We must enquire whether this estimate is sufficiently below unity to convince us that this is not a chance occurrence of the particular data we have used.

In other words, is 0.72 statistically less than 1?

## Forecasting or prediction

To illustrate, suppose we want to predict the mean consumption expenditure for 2006. The GDP value for 2006 was 11319.4 billion dollars. We obtain:

$$\begin{aligned}\hat{Y}_{2006} &= -299.5913 + 0.7218(11319.4) \\ &= 7870.7516\end{aligned}$$

Thus, given the value of the GDP, the mean, forecast consumption expenditure is about 7870 billion dollars.

The actual value of the consumption expenditure reported in 2006 was 8044 billion dollars. The estimated model thus **under predicted** the actual consumption expenditure by about 174 billion dollars. We could say that **forecast error** is about 174 billion dollars.

Suppose the president decides to propose a reduction in the income tax. What will be the effect of such a policy on income and thereby on consumption expenditure and ultimately on employment?

As macroeconomic theory shows, the change in income, a dollar's worth of change in investment expenditure is given by the income multiplier, which defined as

$$M = \frac{1}{1 - MPC}$$

$$MPC = 0.72, M = 3.57$$

An increase (decrease) of a dollar in investment will eventually lead to more than a threefold increase (decrease) in income.

## Using the model for control or policy purposes

Suppose we have the estimated consumption function given in

$$\hat{Y}_t = -299.5913 + 0.7218X_t$$

Suppose further the government believes that consumer expenditure of about 8750 billion dollars will keep the unemployment rate at its current level of about 4.2% (early 2006). What level of income will guarantee the target amount of consumption expenditure?

$$8750 = -299.5913 + 0.7218(GDP_{2006})$$

An income level of about 12537 billion dollars, given an MPC of about 0.72, will produce an expenditure of about 8750 billion dollars.

By appropriate fiscal and monetary policy mix, the government can manipulate the **control variable X** to produce the desired level of the **target variable Y**