Introductory Econometrics

Overview of the Course

- Economics suggests important relationships:
  - How does another year of education change wages?
  - What is the effect of inflation on housing prices?

- This course is about using data to measure these relationships.

Aim of econometrics

- Econometrics differ from statistics:
  - by the applications (economic variables)
  - by the theories that are used to "model" the relationships between variables

Course Topics

1. Introduction
2. Introduction to Data Analysis
3. Probability and Statistics Review
4. Fundamentals of Regression Analysis
   -Single Regressor
   -Multiple Regressors
5. Extensions
   -Binary Variable
   -Panel Data
   -Instrumental Variables
   -Time Series

First lecture

-What is Econometrics?

-Steps in Empirical Economic Analysis

-The Structure of Economic Data

-Causality and the Notion of Ceteris Paribus in Econometric Analysis

1.1 What is Econometrics?

◆ Econometrics = Econo + metrics

◆ Econometrics, literally "economic measurement," involves quantitative analysis of economic problems.

◆ It is the application of statistical methods to connect theoretical economic models to data.

1.1 What is Econometrics?

- Econometric methods are therefore central to economic research.

Economic Models ⟺ Econometrics ⟺ Economic Data

1.1 What is Econometrics?

◆ *Econometrics*

-employs statistical methods

-to analyze data

-in order to

1. estimate economic relationships,
2. evaluate government and business policies,
3. test economic theories, and
4. make predictions and forecasts.

1.1 What is Econometrics?

◆ *Examples of questions that econometrics is useful for*

-How are hours spent on homework and final grade related?

-Suppose the university imposes a maximum number of problem sets on ECO 332. How does this affect students final grades?

-Is there a causal relationship between the education of the mother and education of her children?

1.2 Steps in Empirical Economic Analysis

1. Formulate the questions of interest
2. Obtain and analyse data
3. Construct an economic model
4. Specify an econometric model : estimation, testing, prediction or forecasting

1.2 Steps in Empirical Economic Analysis

1) Formulate a question of interest

-Does income influence driving habits?

2) Construct an economic model

"**Economic Models** consist of mathematical equations that describe various relationships."

-Driving=f(age, income, training, family, vehicle, location)

1.2 Steps in Empirical Economic Analysis

Economic Models Can Come From Formal Derivations…

Formal Derivations Arise From Economic Assumptions and Models:

-Resources are scarce

-An increase in price causes a decrease in quantity demanded

1.2 Steps in Empirical Economic Analysis

VERY SIMPLE Formal Derivations…

-Brushing one's teeth is a function of inputs…:

$$brushing=f(time,\ toothpaste)$$

-The amount of toothpaste purchased is a function of price, availability, income and price of substitutes (ie: whitening strips)…:

$$toothpaste=f(P_{tp},\ avail,\ I,\ P_{y})$$

1.2 Steps in Empirical Economic Analysis

-Time is a function of income, work, sleep, family status, motivation (laziness)

time=f(I, work, sleep, family, motivation)

1.2 Steps in Empirical Economic Analysis

Economic Models Can Also Arise From Intuition or Observation:

-Tall people don't like Wii video games

-Small businesses are less likely to change prices

-Marks are higher in morning classes than
   afternoon classes

1.2 Steps in Empirical Economic Analysis

Economic Models Can Also Arise From A Mixture of Formal Derivations, Intuition or Observation:

-Tall people don't like Wii video games

And

-Quantity demanded is a function of price

Therefore

Wii game demand=f(height, price)

1.2 Steps in Empirical Economic Analysis

From Economic Model to Econometric Model

Formal economic modeling is sometimes the starting point for empirical studies, but it is more common to arrive at econometric models directly by economic "intuition".

1.2 Steps in Empirical Economic Analysis

3) Specify an Econometric Model

◆ Econometric Models have specific functional forms and OBSERVABLE parameters

◆ A model of human capital investment implies getting more education should lead to higher earnings

◆ In the simplest case, this implies an equation like

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

1.3 The Structure of Economic Data

There exist a variety of types of economic data:

- ➤ **Cross-Sectional Data**
- ➤ **Time Series Data**
- ➤ **Pooled Cross Section Data**
- ➤ **Panel Data**

1.3 The Structure of Economic Data

1) Cross-Sectional Data

-A cross-sectional data set consists of a sample of individuals, households, firms, cities, states, countries or a variety of other units, taken at a given point in time.

1.3 The Structure of Economic Data

Cross-Sectional data set is assumed to be a RANDOM SAMPLE

However, a sample of the population is not random if:

1) A <u>sample selection problem</u> occurs (some categories of respondents are more likely to respond than others)

2) Sample size is too small

3) Sample size is too large

1.3 The Structure of Economic Data

Sample Selection Example:

-Rich households are less likely to report their incomes

1.3 The Structure of Economic Data

Small Sample Size Example:

-Any study with less than 30-40 observations

Large Sample Size Example:

-Asking 80% of this class their opinions on the text and expected grade

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

1.3 The Structure of Economic Data

2) *Time Series Data*

-Time series tracks the movement of (one agent/group's) variables over time

Examples

-Stock

-GDP

-Beirut's vacancy rate

1.3 The Structure of Economic Data

2) *Time Series Data*

-Time series data can vary in *data frequency* (daily, weekly, quarterly, etc.)

-frequent time series data can be *aggregated* to evaluate all data on the same frequency

Time Series Data

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

1.3 The Structure of Economic Data

3) Pooled Cross Sections

-Pooled Cross sections are a combination of RANDOM samples from different years

-the same observation should not be followed over different years

-Analysis is similar to cross sectional data, with the additional consideration of structural changes due to time

Pooled Cross-Sectional Data

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

1.3 The Structure of Economic Data

4) Panel Data

-time series data for EACH cross-sectional agent

-also called longitudinal data

1.3 The Structure of Economic Data

Panel Data Advantages:

-able to control for unobserved characteristics

-able to work with a larger data set

-However, more difficult to work with

Panel Data

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

1.4 Causality and the Notion of Ceteris Paribus in Econometric Analysis

One goal of econometric analysis is to examine the causality of two variables

Why?

-a <u>simple plotting</u> of two variables or <u>calculation of correlation</u> will only see if the two variables move together

  -can't show causation

1.4 Causality and the Notion of Ceteris Paribus in Econometric Analysis

Ceteris paribus

-causality can only be correctly examined *Ceteris Paribus* – with all else held equal

-one variable's impact on another variable can only be isolated if all other variables remain constant

1.4 Causality and the Notion of Ceteris Paribus in Econometric Analysis

Causation in the real world

-in the real world, variables change for a reason

-Is the change in Z due to the change in A, B, X, Y?

Z=f(A)?   Z=f(B)?   Z=f(X)?   Z=f(Y)?

Or   Z=f(A, B, X, Y)?

1.4 Causality and the Notion of Ceteris Paribus in Econometric Analysis

Causation example

Take the statistic: Living together before marriage increases the chance of divorce:

Living Together   ⟶   *Higher Divorce Chance*

1.4 Causality and the Notion of Ceteris Paribus in Econometric Analysis

Causation example

BUT why do two people decide to live together?

Uncertainty about partner

Living Together   ?   *Higher Divorce Chance*

Fear of Commitment   ?

?

What actually affects divorce rates?