Econometrics

Introduction to Econometrics

Lec. **0**

Saeed Tajrishy

Faculty of Economics, University of Tehran

Fall 2023

Slides by **Erfān Rezaei M.N.**

Slides primarily based on *Itroductory Econometrics* by Jeffery Wooldridge (2020)

What is Econometrics?

✔ Econometrics involves the development of statistical methods to analyze nonexperimental economic data

✔ Nonexperimental data (a.k.a. observational or retrospective data) are gathered without controlled experiments on individuals, firms, or segments of the economy

✔ While experimental data is commonly collected in laboratory settings in the natural sciences, it is much more difficult to obtain in the social sciences

✔ The major difference between econometrics and statistics is the use of data:

$$\underbrace{Econometrics}_{\text{Nonexperimental}} \neq \underbrace{Statistics}_{\text{Experimental}}$$

Where Does Econometrics Apply?

Here are a few examples:

✔ Macro- or microeconomic models

    ✗ Macroeconomic model: Forecasting inflation rate

    ✗ Microeconomic model: Illuminating the relationship between prices, quantities, and consumer behavior

✔ Policy evaluation

    ✗ How training programs are effective on hourly wages?

> ## Econometrics
>
> Econometrics is a discipline that focuses on estimating economic relationships, testing economic theories, forecasting economics variables, and evaluating and implementing government and business policies.
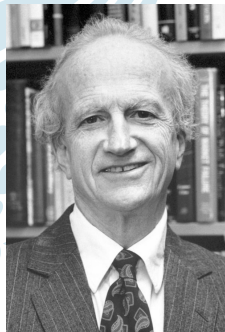
Empirical Analysis - Stages

1. Formulating a question (hypothesis)

2. Outlining an economic framework—this stage is often skipped!

3. Converting economic model to econometric model

4. Gathering data for parameter estimation and hypothesis testing

5. Interpreting results and drawing conclusions

Empirical Analysis - Example

✔ Gary Becker (1992 Nobel Laureate) postulates a utility maximization framework about crime

✔ He tried to explain the framework by specifying economic factors that might affect the economic decision-making process

Source: Nobel Foundation archive

Empirical Analysis - Example (cont'd)

Economic Model of Crime (Becker 1968)

$$y = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$$

Where

- $y =$ hours spent in criminal activities,
- $x_1 =$ "wage" for an hour spent in criminal activity,
- $x_2 =$ hourly wage in legal employment,
- $x_3 =$ income other than from crime or employment,
- $x_4 =$ probability of getting caught,
- $x_5 =$ probability of being convicted if caught,
- $x_6 =$ expected sentence if convicted, and
- $x_7 =$ age.

About Econometrics
oo

Conducting Empirical Analysis
ooo●o

Data Types
oo

Causality
oooooo

Empirical Analysis - Example (cont'd)

✔ The crime model also acknowledges the presence of unobserved factors

✔ The functional form of relationship is somewhat ambiguous, but here a linear relation is assumed

Empirical Analysis - Example (cont'd)

Econometric Model of Crime

$$crime = \beta_0 + \beta_1 wage_m + \beta_2 othinc + \beta_3 freqarr$$
$$+ \beta_4 freqconv + \beta_5 avgsen + \beta_6 age + u$$

Where

- $crime =$ some measure of the frequency of criminal activity,
- $wage_m =$ the wage that can be earned in legal employment,
- $othinc =$ the income from other sources (assets, inheritance, and so on),
- $freqarr =$ the frequency of arrests for prior infractions (to approximate the probability of arrest),
- $freqconv =$ the frequency of conviction,
- $avgsen =$ the average sentence length after conviction, and
- $u =$ every unobserved phenomena. (a.k.a. *error term*, *disturbance*, *residual*, and *noise*)

Economic Data Types

✔ A crucial stage in conducting empirical analysis is gathering data

✔ Thus a good knowledge of different economic data sets is required

✔ To some extent, it's correct to say that there are 4 types of economic data sets:

    ✗ Cross-sectional data
        – Gathered in a single point of time; thus data has no order
        – Somewhat independent and identically distributed (*i.i.d.*)

    ✗ Time series data
        – Mostly serially correlated (autocorrelated)
        – Data is seasonal and has trends; thus data has order

    ✗ Pooled cross sections
        – Combined (pooled) two or more independent cross sections
        – Often used to evaluate policy changes (i.e., the effect of change in property taxes on house prices)

    ✗ Panel/Longitudinal data
        – Just like pooled cross section but with a time dimension
        – Could infer causality better

Economic Data Types - Example

Cross-sectional Data - Example

Pure random sampling from a population

Time Series Data - Example

Apple stock prices (AAPL) in a year with a daily frequency

Pooled Cross Sections - Example

Comparing random sample house prices before and after a reform

Panel/Longitudinal Data - Example

Monitoring crime statistics for each city in two years

About Econometrics
oo

Conducting Empirical Analysis
ooooo

Data Types
oo

Causality
●ooooo

What is Causality?

✔ The causal effect of x on y is defined as: "*How does variable y change if variable x is changed, but all other relevant factors are held constant.*" (or ceteris paribus)

✔ Because of the nonexperimental nature of data in social sciences, usually finding a casual relationship is quite challenging

> ## Ceteris Paribus
>
> Other (relevant) factors being equal.

What is Causality? - Example

Crop Yield - Example

Q: How to find out if fertilizer can increase the production? (ceteris paribus)

A: Choose several equal parts of land and try adding fertilizer randomly; then, compare the yield.

R: The experiment is performed properly since the fertilizer is independent of all other factors influencing crop yield.

Retrun on Education - Example

Q: What's the effect of one more year of education on wage? (ceteris paribus)

A: Random sampling of population is not possible!

R: Education isn't necessarily independent of all factors (e.g. intelligence); thus, other factors might explain it better.

About Econometrics
oo

Conducting Empirical Analysis
ooooo
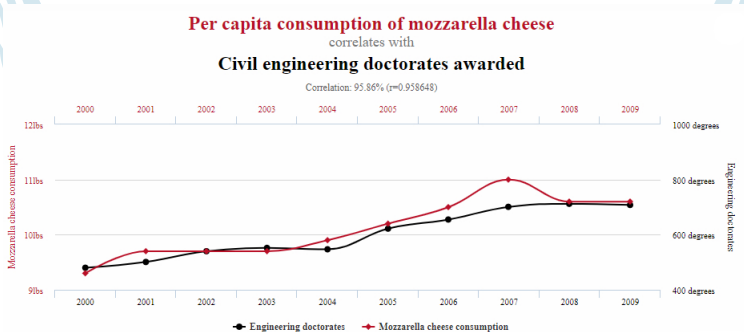
Data Types
oo

Causality
ooo●oo

Causality vs. Correlation

✔ Does correlation necessarily imply causality? The answer is NO! The relation among them is rarely conclusive and more suggestive

✔ The common problem of self-selection bias can lead to the emergence of spurious correlations, which is another prevalent issue

Self-Selection Bias

Occurs when the decision to participate in a study is left entirely up to individuals. This makes the sample non-random.

Causality vs. Correlation - Example

Per capita consumption of mozzarella cheese
correlates with
Civil engineering doctorates awarded
Correlation: 95.86% (r=0.958648)

Source: Tyler Vigen, (2015), *Spurious Correlations*

About Econometrics
○○

Conducting Empirical Analysis
○○○○○

Data Types
○○

Causality
○○○○●○

Is It Always Cause and Effect?

✔ The econometric relation is not always in the form of a cause-and-effect matter, and sometimes a more equation-like relation is apparent

---

Expectations Hypothesis

Long term interest rates equal compounded expected short term interest rates.

$$(1 + r_{lt})^n = \prod_{i=1}^{n} \left(1 + r_{y_i}^e\right)$$

Copyright Disclaimer

These slides were prepared for educational purposes by Erfān Rezaei Mayahi-Nejad, at the Faculty of Economics, University of Tehran. The slide show is licensed under a Creative Commons Attribution - ShareAlike 4.0 International License. Feel free to use any part of it by mentioning where you got it and sharing the result under the same terms. Temporarily, the LaTeX source code is only available upon request via email.