Introductory Econometrics

Introduction

Yaohan Chen

School of Big Data and Statistics, Anhui University
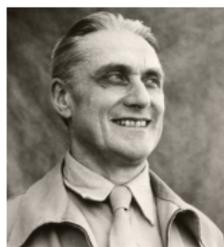
Spring, 2026

Contents

Econometrics

- Econometrics is a subdiscipline of Economics.
- "**Econometrics**" was initially coined by Norwegian economist, Ragnar Frisch in 1926.

Ragnar Frisch   Jan Tinbergen

- The Econometric Society was founded in 1930. *Econometrica* was established in 1933, edited by Ragnar Frisch from 1933 to 1954.
- Ragnar Frisch won the first Nobel Memorial Prize in Economic Sciences in 1969, together with Jan Tinbergen.

What is about Econometrics

- Statistical Method + Economic Theory + Economic Data.
- Paul A. Samuelson and William D. Nordhaus (2004): *To sift through mountains of data to extract simple relationships.*
- Macroeconomics + Microeconomics + Econometrics: Pivotal roles in modern economics.

Econometric Model

- Model: Linguistic, Physics, Biology, ...
- Mathematics Model.
- Economics Model.

$$\text{Porduction Function:} \quad Q = f(T, K, L)$$

- Econometric Model: Put the data to fit the Model.

$$Q = f(T, K, L)\, u$$

where u collects the randomness of data.

Structure of Econometrics

- Preliminary, Intermediate, and Advanced Econometrics.

- Theoretical Econometrics and Applied Econometrics.

- Microeconometrics and Macroeconometrics.

  - Microeconometrics: Individuals and Households.

  - Macroeconometrics: Time Series Data, Structural Estimation, most for empirical Macroeconomic analysis and Financial Economics.

- More refined classification: Financial Econometrics, Spatial Econometrics, Nonparametric Econometrics, Bayesian Econometrics, Panel Data Econometrics...

- Modern Perspectives and Methodologies: Machine Learning and Causal Inference.

Economics and Econometrics

- Economic theory as the foundation, in combination of the Statistical methods and Mathematical methods.

- Economists from different fields have contributed to the development of Econometrics.
  - Robert Solow, Nobel Laureate in 1987.
  - Robert Fogel and Douglass North, Nobel Laureates in 1993.
  - Robert Engle and Clive Granger, Nobel Laureates in 2003.
  - Eugene Fama, Robert Shhiller, and Lars Peter Hansen, Nobel Laureates in 2013.

- Mathematical Statistics, which is fundamental to Econometrics, but different.

Econometric Modeling

- Model Design.
- Data Collection.
- Parameter Estimation.
- Test.

Economic Modeling

- Which variables to include ?
  - Economic theory meaning for the selected variables.
  - Which are dependent variables ? Which are explanatory variables ?
- Which equation(s) to establish ?
  - Economic theory.
  - Intuition is also important.
  - $Q = Ae^{\gamma t} K^{\alpha} L^{\beta}$ or $\ln Q = \ln A + \gamma t + \alpha \ln K + \beta \ln L$?

Data Collection

- Types of data:
  - Time Series Data. $\{x_t\}$
  - Cross-sectional Data. $\{x_i\}$
  - Panel Data. $\{x_{i,t}\}$
  - Continuous Data or Discrete Data ?

- Cleaning data:
  - How to handle missing observations ?
  - How to handle measurement errors ?
  - Comparability of different variables ?

Data Sources (I): USA and other North America

- Panel Study of Income Dynamics (PSID),
  http://psidonline.isr.umich.edu/

- National Longitudinal Surveys of Labor Market Experience
  (NLS),
  http://www.bls.gov/nls/home.htm

- Current Population Survey (CPS),
  http://www.census.gov/cps/

- World Bank's Living Standards Measurement Study (LSMS),
  http://www.worldbank.org/LSMS

- World Bank,
  http://data.worldbank.org

Data Sources (II): Europe

- Eurostat: The European Community Household Panel (ECHP),
  http://ec.europa.eu/eurostat/web/ microdata/european-community-household-panel

- British Household Panel Survey (BHPS),
  https://www.iser.essex.ac.uk/bhps

- Swedish Panel Study of Market and Non-market Activities (HUS),
  http://www.nek.uu.se/faculty/ klevmark/hus.htm

- German Socio-Economic Panel (G-SOEP),
  http://www.diw.de/soep

Data Sources (III): Mainland China

- China Household Finance Survey (CHFS), http://www.chfsdata.org/
- China Health and Nutrition Survey (CHNS), http://www.cpc.unc.edu/projects/china, admin- istered by University of Carolina?s population center
- China Multi-Generational Panel Dataset, http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/265

Parameter Estimation

- What is statistics ? How to use data to pin down the values of parameters we are interested ?
- Which estimation methods to use ?
- How to implement ? Programming.
    - Eviews, Stata
    - MATLAB, R, Python, SAS

Test

- Statistical significance of estimated parameters.
- Model specification test.
- Interpreting the underlying mechanism using Economic theory.
- In-sample fitting goodness and out-of-sample (OOP) explanation power.

Time Series Data

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

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

© Cengage Learning, 2013

Back

Cross-sectional Data

TABLE 1.2 A Data Set on Economic 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

© Cengage Learning, 2013

Back

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

© Cengage Learning, 2013

Back