

# **Ch01: Introduction**

## **Financial Econometrics**

Prof. Huei-Wen Teng, 202402

1. Econometrics is about how we can use theory and data from economics, business, and the social sciences, along with tools from statistics, to answer “how much” questions.
2. Economists and other social scientists work in a complex world in which data on variables are “observed” and rarely obtained from a controlled experiment.
3. In **economics**, we express our ideas about relationships between economic variables using the mathematical concept of a function

$$Q^d = f(P, P^s, P^c, INC) .$$

# An econometric model

- An econometric model consists of a systematic part and a random and unpredictable component  $e$  that we will call a *random error*

$$Q^d = f(P, P^s, P^C, INC) + e,$$

where

$$f(P, P^s, P^C, INC) = \beta_1 + \beta_2 P + \beta_3 P^s + \beta_4 P^c + \beta_5 INC.$$

1. The systematic portion is the part we obtain from economic theory, and includes an assumption about the functional form. The functional form represents a hypothesis about the relationship between the variable.
2. The coefficients  $\beta_1, \beta_2, \dots, \beta_5$  are unknown parameters of the model that we estimate using economic data and an econometric technique.
3. The random component represents “noise” component, which obscures our understanding of the relationship among variables, and which we represent using the random variable  $e$ .

# Economic data types

Economic data comes in a variety of “flavors”.

1. Data may be collected at various levels of aggregation: Micro or Macro.
2. Data may also represent a flow or a stock.
  - Flow: measured over a period of time.
  - Stock: measured at a particular point in time.
3. Data may be quantitative or qualitative:
  - Quantitative: expressed as numbers.
  - Qualitative: expressed as an “either-or” situations.

# Types of data

1. A cross-section of data is collected across sample units in a particular time period. The “sample units” are individual entities and may be firms, persons, households, states, or countries.
2. A time series is data collected over discrete intervals of time. The key feature of time-series data is that the same economic quantity is recorded at a regular time interval.
3. A “panel” if data, also known as “longitudinal” data, has observations on individual micro-units who are followed over time.

# Steps in the research process

1. Use economic theory to think about the problem.
2. Develop a working economic model leading to an econometric model.
3. Obtain sample data and choose a desirable method of statistical analysis based on initial assumptions and an understanding of how the data were collected.
4. Estimate the unknown parameters with the help of a statistical software package, make predictions, and test hypotheses.
5. Perform *model diagnostics* to check the validity of assumptions.
6. Analyse and evaluate the economic consequences and implications of the empirical results.

# R demo i

You need to:

1. Download R program (<https://cloud.r-project.org/>) (Required!)
2. Use **PoEdata** package (<https://github.com/ccolonescu/PoEdata>) (Required!)



# R demo i

```
rm(list=ls()) # Caution: this clears the Environment

install.packages("devtools") # install the "PoEdata" package
library(devtools)
install_github("ccolonescu/PoEdata") # Fixed by TA
library(PoEdata)

data("andy") # retrieve the data
?andy # provides info for the dataset

## Read data from a specified file
andy = read.table('dat/andy.dat', header = FALSE);
andy

head(andy)
```

# R demo i

```
tail(andy)
nrow(andy)

# andy is a matrix.
v1 = andy[,1];
v2 = andy[,2];
v3 = andy[,3];

plot(v1); plot(v2); plot(v3);

boxplot(andy)
hist(v1) # histogram
pairs(andy) # pairwise scatterplots
```

# End of Chapter 1!