Lecture 1: Economic Questions and Data

Zheng Tian

Contents

1 What is Econometrics about?

1.1 Definition of Econometrics

Econometricians may give you very different answers for the question of *What is Econometrics*. The following answers are all right from their respective point of views:

- econometrics is the science of testing economic theories;
- it is the set of tools used to forecasting future values of economic variables;
- it is the process of fitting mathematical economic model to real-world data;
- it is the science and art of using historical data to make quantitative policy recommendations in government and business.

Stock and Watson (2015) define Econometrics as

> At a broad level, econometrics is the science and art of using economic theory and statistical techniques to analyze economic data.

Science or art?

Let us dissect the above definition a little bit. First, why is econometrics the science AND art?

- Econometrics is a science because it essentially complies with the principle of **falsifiability**[1] of scientific research, as Karl Popper defined.

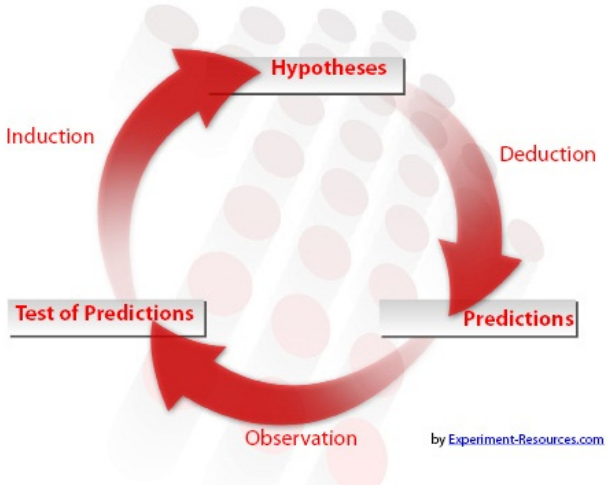

Figure 1: A reasoning cycle of scientific research

  Econometricians propose a hypothesis based on either existing economic theories or their own economic reasoning, and then collect data to test the hypothesis that can be rejected or fail to be rejected. Even though an economic theory is not rejected by one set of data at a time period, it can be very likely to be rejected using another set of data at another time period. Then, a new theory or hypothesis will be brought up.

- Econometrics is an art because the data are usually incomplete and unobserved to validate a hypothesis, so we need to use human creativity to reach a balance between scientific rigor and realistic approximation.

The following quote captures the dual nature of econometrics as both science and art:

> Econometrics is alchemy since econometricians can create nearly any result desired, but it is also science because econometricians also know how to reject and avoid spurious models. – Hansen (1996)

Economic theory, statistics, and data

A complete process of econometric research inevitably consists of three components: economic theory, statistical techniques, and economic data. When we have a research question, we first need to find or formulate an economic theory that can be either a formal mathematical model or a logical economic reasoning. Guided with this economic theory, we build an econometric model to characterize the relationship between various variables

[1] Read this article about falsifiability `https://explorable.com/falsifiability`.

involved in the theory. Then we collect data to measure these variables, and use statistical techniques to estimate and model and test hypotheses raised from the theory.

TODO Use ditta to generate a flowchart Let's look at a real example to get a first impression of what is econometrics.

2 Economic Questions We Examine

2.1 Four practical questions

Question 1 does reducing class size improve elementary school education?

Question 2 is there racial discrimination in the market for home loan?

Question 3 how much do cigarette taxes reduce smoking?

Question 4 what will the rate of inflation be next year?

2.2 How does an econometrician formulate such questions?

Ideally, an econometrician would design his/her econometric modeling by the following steps

1. Establish a theoretical model to qualitatively describe how X could cause Y, holding other factors constant. From the theoretical model, put forward some hypotheses to validate the theory.
2. Find actual data to measure the variables in the theory
3. Set up an empirical model to test the theoretical model, using available data.
4. Choose a suitable estimation method to estimate the empirical model.
5. Perform some hypothesis tests and model specification test to validate the estimation results.
6. Based on the estimation and test results, determine whether the theory is internally and externally valid.

3 Causal Effects and Idealized Experiments

The success of an econometric analysis relies on whether the causal effects between X and Y can be accurately identified, excluding the influences of other factors.

3.1 Randomized controlled experiment

Controlled experiment

Control group (no treatment) versus treatment group (with treatment)

Randomized experiment

the treamtment is assigned randomly

Advantages and disadvantages

Advantages eliminate the possibility of a systematic relationship that could blur the causal effects of the treatment

Disadvantages it is difficult to implement, especially for social science

4 Data Sources and Types

4.1 TODO Experimental versus observational data

4.2 Cross-sectional data

- heights of all 30 students in a class
- total population of each province in China in 2014

4.3 Time series data

- stock price of Company A by hour over the last month
- consumer price index of China by month from 1990 to 2014

4.4 Panel data

- annual wage of a fixed group of respondents in a survey conducted by a statistic agency in 1990, 1995, 2000, 2005, and 2010
- GDP per capita of each country in Asia from 1990 to 2014