

Part 1

Introduction

EE325

Introductory Econometrics

Revision Aug 2020

A little introduction

- Course description

Application of statistical and economic theories in analyzing economic data, with emphases on parameter estimation techniques and applications of simple and multiple regression models to economic analyses. Use of computer application in practice is also covered.

- Class details

EE325 Section 046401: Tuesday and Thursday, 11:00 – 12.30

Lecturer: Weerawat Phattarasukumjorn

Office hours: Room 437, Tha Prachan Campus, Wed Thu Fri from 13.00 - ...

Email: weerawat@econ.tu.ac.th

- Class rules

Be focused and on time! This class is difficult (including for me to teach). It requires a lot of graph reading, mathematical and statistical skills to get the essence of this subject.

Submit all your homework and assignments.

*Class attendance is **not** taken.*

- Reference

Gujarati, D.N., and D.C. Porter, Basic Econometrics. 5th ed., N.Y., McGraw-Hill, 2009.

Wooldridge, J. M. Introductory Econometrics: A Modern Approach. 6th ed. Thompson: South-Western, 2016.

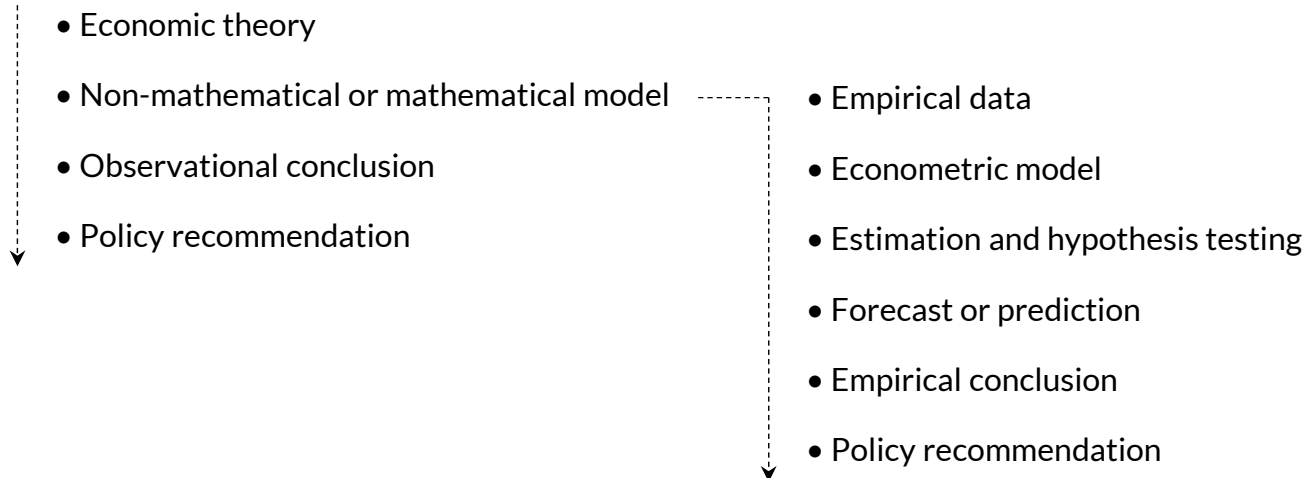
What is the definition of econometrics?

We, first and foremost, try to understand the topic, analyzing morphemes.

- Econometrics = economics + metrics
- Economics: you all know what this means.
- Metrics: a system of measurement.

To conclude, *“econometrics is the application of statistical methods to economic data in order to give empirical content to economic relationships. More precisely, it is the quantitative analysis of actual economic phenomena based on the concurrent development of theory and observation, related by appropriate methods of inference”*.

The flow from economic theory to empirical conclusion



What is the definition of econometrics?

Let's come up with some examples you studied in both EE211 and EE212.

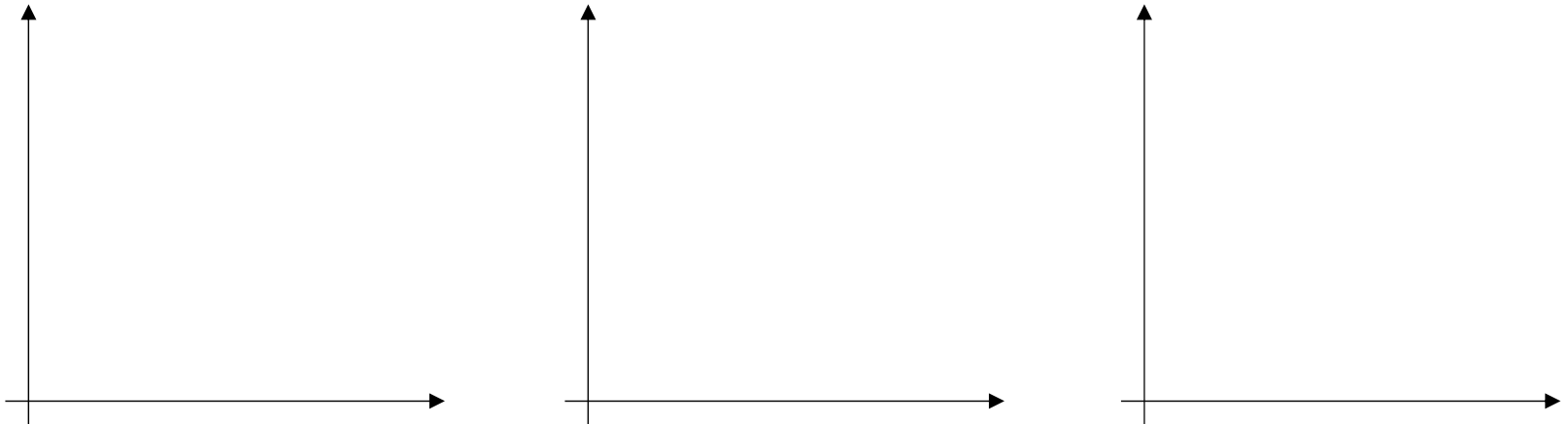
To further enhance our understanding and confirm our theory, econometrics can estimate **magnitude** of dependence and can be **statistically tested**.

What are to be aware of?

It seems like this course will introduce you about how to make an inference of two variables or more. There are 3 types of relationship you should acknowledge.

(1) **Correlation relationship** is any statistical association or the degree of association between two or more variables.

Correlation

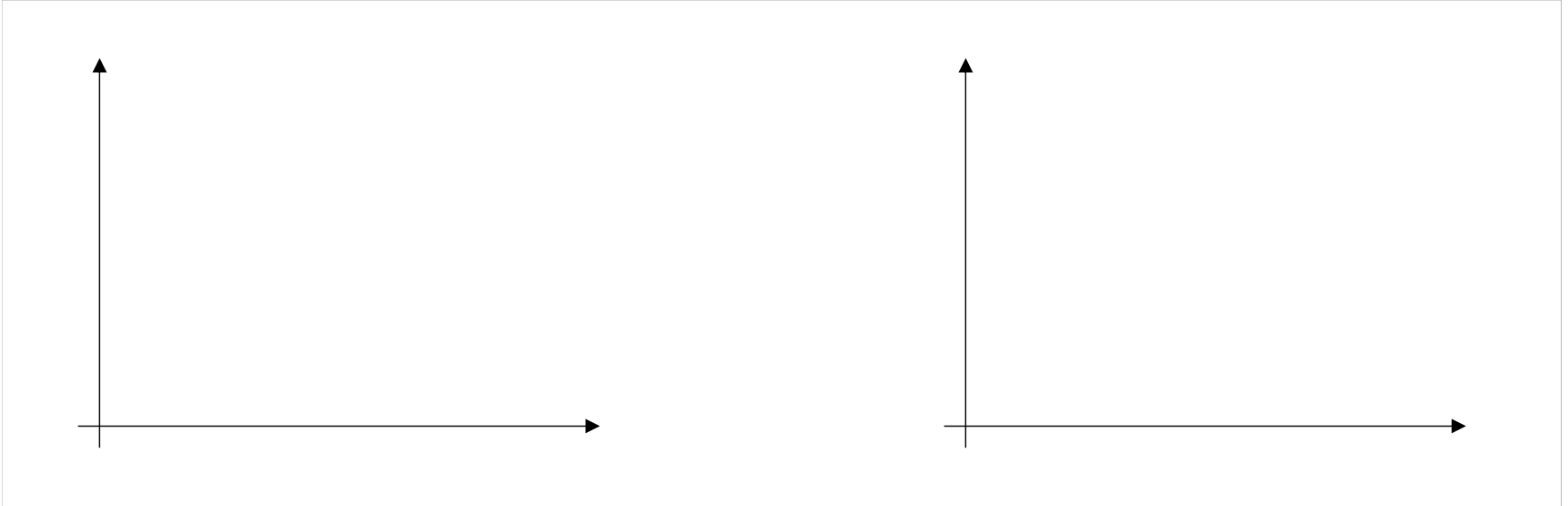


Some examples of correlation here.

What are to be aware of?

(2) **Regression relationship** is a relationship derived from a set of statistical processes for estimating the relationships between a **dependent** and **one or more independent** variables.

Regression



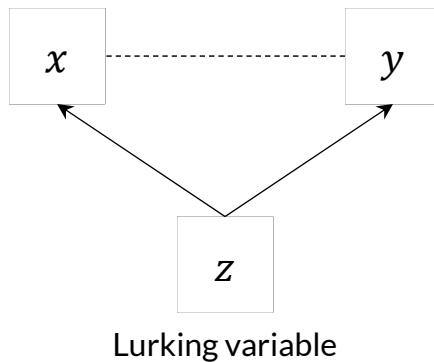
Some examples of regression here.

What are to be aware of?

(3) **Causal relationship** is influence by which one event, process or state (a cause) contributes to the production of another event, process or state (an effect) where the cause is partly responsible for the effect, and the effect is partly dependent on the cause.

Some examples of causality here.

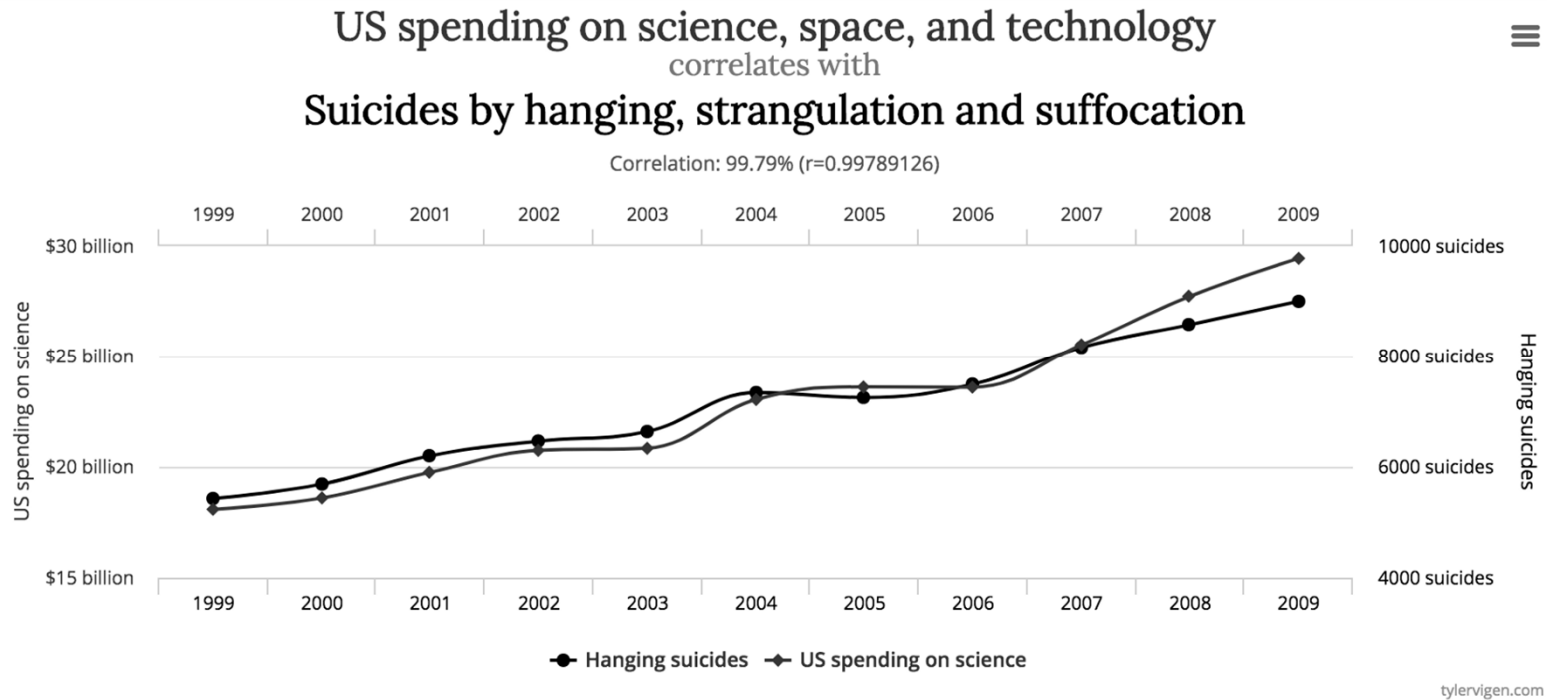
Important note! – Correlation and regression **do not** imply causality.



For example, given that x is ice-cream sales figure and y is crime rate, we may find both correlation and regression relationship in the summer. However, the lurking variable z is temperature.

What are to be aware of?

Another example of correlation without causality.



Data sources: U.S. Office of Management and Budget and Centers for Disease Control & Prevention

tylervigen.com

This can also be called **spurious correlation**.

Types of data and data collection

Data can be used by collecting

- (1) **Primary data** is collected by a researcher from first-hand sources, using methods like surveys, interviews, or experiments.
- (2) **Secondary data** is gathered from studies, surveys, or experiments that have been run by other people or for other research.

Broad types of data are categorized into

- (1) **Cross-sectional data** is a type of data collected by observing many subjects (such as individuals, firms, countries, or regions) at the one point or a period of time. The analysis has no regard to differences in time.
- (2) **Time series** is a series of data points indexed in time order.

	student	GPA
observations	student 1	GPA 1
	student 2	GPA 2
	student 3	GPA 3
	student 4	GPA 4
	student 5	GPA 5
	student 6	GPA 6

	time	GPA
observations	t=1	GPA 1 (t=1)
	t=2	GPA 1 (t=2)
	t=3	GPA 1 (t=3)
	t=4	GPA 1 (t=4)
	t=5	GPA 1 (t=5)
	t=6	GPA 1 (t=6)

Types of data and data collection

(3) **Pooled cross-sectional data** is a multiple cross-sectional data pooled without observing the same subject.

(4) **Panel data** is a set of data collected over time and over the same individuals

student	GPA	time
student 1	GPA 1	t=1
student 2	GPA 2	t=1
student 3	GPA 3	t=1
student 2	GPA 2	t=2
student 3	GPA 3	t=2
student 4	GPA 4	t=2
student 1	GPA 1	t=3
student 3	GPA 3	t=3
student 4	GPA 4	t=3

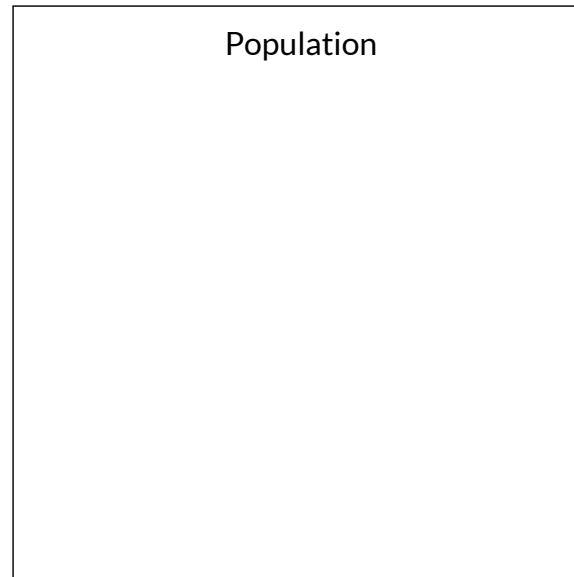
time	student 1	student 2	student 3	student 4
t=1				
t=1				
t=3				

student	GPA	time
student 1	GPA 1	t=1
student 2	GPA 2	t=1
student 3	GPA 3	t=1
student 1	GPA 1	t=2
student 2	GPA 2	t=2
student 3	GPA 3	t=2
student 1	GPA 1	t=3
student 2	GPA 2	t=3
student 3	GPA 3	t=3

time	student 1	student 2	student 3	student 4
t=1				
t=1				
t=3				

Population vs. sample

- (1) **Population** refers to a group of observations of interest. If the data cover all the observations of interest, the set of data is called a **census**.
- (2) **Sample** refers to a subset of population of interest. Most of the time they are statistically random samples which is called **survey**.



Where to find them?

Most of the data used by econometricians are secondary data. Where can we find them in, especially Thai dataset?

(1) Cross-sectional data is not free most of the time. University can request from the National Statistics Office (NSO) for their project. Well-written project proposal must be submitted.

- Household Socio-Economic Survey (SES) – income, expenditure, debt, asset (recently added), etc. Unit of analysis is household.
- Labor Force Survey (LFS) – wage, working hour, occupation, job search. Unit of analysis is individual.
- Health and Welfare Survey (HWS) – health insurance, health benefit, partial utilization, etc. Unit of analysis is individual.
- Office of the National Economic and Social Development Council (NESDC) – the NESDC takes NSO data to calculate important statistics. https://www.nesdc.go.th/more_news.php?cid=74

(2) Time series are widely available, because they are not identity specified.

- Bank of Thailand (BOT) – GDP, financial and monetary statistics, currency exchange, ect. <https://www.bot.or.th/Thai/Statistics/Pages/default.aspx>
- Ministry of Finance (MOF) and Fiscal Policy Office – tax revenue and expenditure, tax disbursement, national debt, etc. <http://www.fpo.go.th/main/Statistic-Database.aspx>
- Others include Ministry of Commerce (MOC), Stock Exchange of Thailand (SET), international organization such as International Monetary Fund (IMF), World Bank, United Nations (UN).

(3) Panel data are very limited in Thailand. The ones that I know of are

- Panel SES (by the NSO) from 2005, 2006, 2007, 2010, 2012 – not free.
- Townsend Thai data (cooperated with UTCC and TRF) from 1997 to 2017 – request needed. <http://riped.utcc.ac.th/panel/data/townsend-thai-data/>

Concluding remarks

Important remarks.

- Statistics and econometrics are mathematical use. They are *not* to find a solid scientific proof.
- Secondary data, especially survey, can normally be biased on their own.
- Correlation and regression cannot imply causation.

List of chapters

(1) Introduction

(2) Statistics revision

(3) Simple linear regression

- Midterm examination (Tue September 29: 12.00-14.00)

(4) Multiple regression

(5) Dummy variable

(6) Regression problems

- Final examination (Mon December 14: 13.30-16.30)