Teachers

Professor Otto Toivanen
Economics dept., Aalto
Otto.Toivanen@aalto.fi
Office hours: on appointment.

T/As Salla Simola
Economics dept., Aalto
Salla.Simola@aalto.fi
Office hour: on appointment.
Interesting questions

1. How should we price our (new) product(s)?

2. Is our advertising working?

3. Is our new incentive scheme delivering results?

4. What affects the probability of defaulting on a mobile phone credit?
How this course can help

1. Oftentimes, you sit on the answer ➔ data.

2. Oftentimes, getting a roughly-right answer not that difficult ➔ statistics / econometrics.

3. When you don’t sit on the answer, you may get it with a little work ➔ experiments.

4. To do all this intelligently, you need (economic) theory.
What an increasing # firms already do

• Collect data of their own performance – electricity companies.

• Collect data on their customers – retail stores, telecom operators.

• Sometimes, of markets / rivals - Nielsen...

• Analyze these data.

• Increasingly, experiments – Google, airlines, garment stores...
• https://developers.google.com/analytics/solutions/experiments

• https://en.wikipedia.org/wiki/A/B_test

• https://vwo.com/ab-testing/

• https://www.optimizely.com/ab-testing/
Some statistics about statistics / data


• Facebook stores, accesses, and analyzes **30+ Petabytes** of user generated data. (peta = \(10^{15} = 1000\) TB).

• Walmart handles more than 1 million customer transactions **every hour**, which is imported into databases estimated to contain more than 2.5 petabytes of data.
Even images become useful for data-analysis... New York Times 31.12.2017:

How Do You Vote? 50 Million Google Images Give a Clue
By STEVE LOHR DEC. 31, 2017

For the first time, helped by recent advances in artificial intelligence, researchers are able to analyze large quantities of images, pulling out data that can be sorted and mined to predict things like income, political leanings and buying habits. In the Stanford study, computers collected details about cars in the millions of images it processed, including makes and models.

Some statistics about statistics / data

• According to estimates, the volume of business data worldwide, across all companies, doubles every 1.2 years.

• 140,000 to 190,000 deep analytical talent positions and

• 1.5M more data-savvy managers needed in the U.S. by 2018 to “take full advantage of Big Data”.

Bankruptcies in Finland 2016 - 2017

What this course is about

• The ABC of how to

1. understand the answers to
2. evaluate the quality of answers to and
3. provide an answer to

the type of questions just posed.
What this course is about

• Tools: economic theory + statistical tools + data + knowledge.

• In short: econometrics.
What this course is about

• Econometrics:

"A branch of economics in which economic theory and statistical methods are fused in the analysis of numerical and institutional data” Hood and Koopmans (1953, pp. xv.).

What this course is about

• **Economic theory:** modeling economic decisions requires understanding how and why they are made.

• **Statistical methods:** tool to derive numbers from numbers to add to knowledge.

• **Numerical data:** the raw material to be explained, and to be used to explain.

• **Institutional data:** the environment in which the numerical data arises.
What this course is also about

• How to interpret research results.

• How to evaluate research.

• How to conduct (good) research (thesis?).
Approach

1. Practical rather than theoretical, though both covered.
2. Hands-on learning of each step.
4. Own (group) work.
The data

- Source: Telecom operator Elisa.
- Anonymized data on individual customers.
- Customer characteristics.
- Price of gadget; whether or not defaulted on credit.
Contents

• Book: Stock and Watson, An Introduction to Econometrics.

• Lectures follow the book, though we try to make as much use of Elisa’s data as possible.

• Exercises – some surely dull, some hopefully interesting, all useful.

• Capstone work.
Lectures – current plan

• All lectures in U6/U149 10 – 12

2.1 Tu introduction & L1: estimation of the mean

8.1 Mon L2: statistics recap – estimation of the mean

10.1 Wed L3: univariate regression #1 ch4&5

15.1 Mon L4: univariate regression #2 ch5
Lectures – current plan

17.1 Wed   L5: univariate regression c’ed.

22.1 Mon   L6: causal parameters #1: experiments & problems with observational data ch13.1, 13.2, 6.1

25.1 Wed   L7: multiple regression #1: estimation ch6

29.1 Mon   L8: multiple regression #2: interpretation, testing ch7
Lectures – current plan

31.1 Wed    L9: multiple regression #3: problems ch8

5.2 Mon    Elisa guest lecture – date TBC

7.2 Wed    L10: Panel data

15.2    mid-term

19.2 Mon    NO LECTURE
Lectures – current plan

21.2 Wed   NO LECTURE

26.2 Mon   mid-term answer session

28.2 Wed   L11: causal parameters #2: Difference-in-Difference

5.3 Mon    L12: causal parameters #3: Instrumental variable ch12
Lectures – current plan

7.3 Wed  L13: IV #2

12.3 Mon  L14: limited dependent variables ch11

14.3 Wed  L15: time series #1: forecasting ch14

19.3 Mon  L16: time series #2: dynamic causal effects ch15
Lectures – current plan

21.3 Wed  L17: time series #3: VAR ch16

26.3 Wed  L18: time series #4: VAR ch16

28.3  L19: recap

6.4  final exam
Exercises

• This week

• Wed 10 – 12 & Th 17 – 19.

• From then on

• Tu 15 – 17 & Th 17 – 19.
Exercises

• Stuff to be done

1. go through unclear points & questions (model answers provided ahead of lectures),

2. go through empirical analysis & interpretation,

3. cover details not covered in the lectures,

4. learn to use Stata.
Exercises

• Roughly a week to do each exercise set. Deadline for answers the day before the first exercise class (i.e., Mondays).

• Everything through MyCourses.

• Use word-files or pdf, one document / exercise. Remember to include your name and student ID in all documents.

• Group size 1 – 2 students.
Exercises

• We will use the statistical program Stata in the exercises.

• No previous knowledge required.

• Exercise set #1 contains recap of probability.

• There will be an introduction to Stata.

• Follow-up classes provide further information.
Software

• Familiarity with one econometric software package extremely good investment, both for studies and for later.

• The Department of Economics policy is to use Stata in all course-related empirical work.

• Applied Econometrics I and II, Labour Economics, Economics of Science, ...

• Thesis work.
Acquiring Stata

• The department sponsors purchase of Stata.

• Perpetual student license for Stata IC costs 1661SEK ≈ 160€ before sponsoring.

Acquiring Stata

• You need to list (instructions will follow) your student ID and email.

• You will then be able to access a special web shop through a link and purchase the software at a discounted price.
Course evaluation

• Exercises 20%

• Capstone 35%

• Exam 45%