Applied Econometrics: an introduction

This version: May 2018

PLEASE CONFIRM YOUR PARTICIPATION BY EMAIL TO hanna.hottenrott@tum.de NO LATER THAN July 15th 2018.

Credits

This course is a modified and extended version of a class previously taught by Prof. Dr. Oliver Alexy in 2014.

General information

Instructor: **Prof. Dr. Hanna Hottenrott**
Contact: Room 2518 (Building 0505, Luisenstr.), hanna.hottenrott@tum.de
Course time: **19 – 25 of July 2018, 9am-4pm**
Location: tba (Central Campus)

Course content

What this course is
This doctoral course aims at giving doctoral students an overview of the essentials quantitative methods. We will not focus on the econometrics, but aim at developing the students' understanding of what the methods are all about and how they work in practice. The course is taught over 5 days. We expect all students to read all the required readings and be prepared to discuss as the class will be interactive. Attendance is mandatory! We have combined lectures that provide an overview of the different concepts as well as more applied exercises that employ the techniques we will discuss. All examples will be illustrated using STATA. Please bring your own laptops – if you do not have STATA, please apply for a 30-day test version by the end of September here: http://www.dpc-software.de/index.php?id=488.

What this course is not

- A fully-fledged theoretical econometrics class – we do cover some of the basics, however, we will speak very little about the maths (only as much as is necessary to get the gist)
- A fully-fledged applied econometrics class – we only cover some of the most basic techniques used in empirical analysis; for more advanced techniques, including the analysis of panel data, you will need to take an additional class for which, however, you should find this course to be an excellent starting point.
- A quantitative research course – again, we will cover key issues, but your training should not stop here.

Course objectives

Knowledge Objectives
In this course, we strive to familiarize you with key building blocks of applied econometrics: where data comes from and how it can be analyzed. We will do so in an applied fashion, using actual data and software tools.

Skills Objectives
- Improve diagnostic and analytical skills
- Enhance verbal skills via class and group discussions
- Build up critical thinking and interpretation skills
- Build up quantitative reviewing and research design skills

Learning Objectives
At the end of this course, students will be able to demonstrate understanding, critical assessment and application of the following:
- Generic design of a quantitative research question or PhD proposal
- Their own, as well as others' quantitative research method and design choice
- Fundamentals of econometrics so as to be able to join more advanced econometrics courses
- Ability to execute standard econometric analyses using STATA

Course Procedures

Have Fun
Our challenge is to make sure that you learn about the fundamental importance of econometrics—even if you do not intend to engage in quantitative research! While this topic may seem dusty you need to have at least some understanding of it to do good research. To illustrate this point to you, we will do our best to make this course as applied as possible and to design it in a way that can show direct benefits for your PhD. Thus, our goal is to enlarge your toolset and sharpen the tools that you have as a researcher in your own projects as well as a reviewer of others' work. Having these skills will make it easier for you to join any academic debate, in particular those outside your area of specialization – who does not like talking about other people's work? And of course, knowing that you have the skills to tackle a research topic and making it 'yours' is great fun!

Prepare and Participate Actively
To make this class a success for everyone, we need you to prepare for class and participate in class. There is a lot of material that we want to cover, and this material is (1) ordered consecutively (so you need the early sessions to really get the later ones) and (2) the very foundation of any empirical-quantitative work you may do in your PhD. Thus, in order to get the most out of this course, we ask you to prepare for each class, so that you will be capable of discussing the assigned readings, which we will discuss in class. If for some reason you are not prepared let me know before the start of class to save us both the embarrassment of my calling on you.

Assessment

There will be no mandatory assessment. Nonetheless, note that attendance and participation are mandatory,

and regular unexcused absence or ill preparation **will** lead you to being banned from the course. In order to receive a certificate, you can miss a maximum of 1 session. However, you are required to submit a written report on the session you missed **the next day**.

In addition, there exists the **option for voluntary assessment** (i.e., everyone who may need a grade should choose this option)**.** Specifically, students will have the opportunity to use this course to develop, elaborate, or scrutinize the research methods needed for their own PhD endeavors their own research questions as laid out below. These are also the tasks we assign to students who are not members of the TUM Graduate School and need an assessment or even a grade to get credit for this course in their respective PhD programs. If you want to take part in these assignments voluntarily, or in case you need to be assessed or even graded, please let me know.

Course Assignments (70% of the course grade)

1. **Progress Report: Current status** (2-3 pages, typed, double-spaced). Sketch your PhD research question, and how you intend to study it quantitatively – that includes both the data and the methods you intend to use. **Due before Session 1. 10% of course grade**
2. **In-class presentation**. Give a *short* (10 minutes max.) presentation in class about your current research, data, and the methods you are applying, or intending to apply. **Due the day before Session 10. 30% of course grade**
3. **Data and methods section** (5-7 pages, typed, double-spaced). Incorporating the feedback you received in your presentation, write a fully-fledged data and methods section building on your in-class presentation. **Due two weeks after Session 10. 30% of course grade**

Participation (30% of the course grade)

It is important to appreciate that every member of the class is a co-producer of the class discussion, listen carefully to one another and attempt to build on or constructively critique prior comments. An effective participant:

- Is a good listener
- Makes points relevant to the ongoing discussion
- Makes comments that add to our understanding of the reading or article
- Is willing to challenge ideas that are being expressed
- Integrates material from past classes, other courses

If you need to receive a grade for this course, I will ask each of the other students to submit a one-page evaluation and grade out of 20, based on these five criteria, for you. These scores will be averaged with one assigned by me to determine your 30% participation grade.

Preliminary course outline

The first session in a day will start at 9:30am and last to about 11:45am, after a short lunch break it will continue from about 12:30pm to 2:30pm. The second session in a day will last from 2:45pm to about 3.45pm. All session will start *sharply* on time (s.t.).

Day 1:

Session 1: Introducing the core concepts
Topics covered:

- What you can address with quantitative studies versus other types of methods
- Types of variables
- Different models
- Collecting data using different techniques
- Hypothesis formulation – what is a good hypothesis?
- How to find help online

Voluntary reading

- ASR [Singleton and Straits (newest edition). *Approaches to Social Research*, Oxford University Press], Chapters 5, 15
- Angrist, J. and Pischke, J.-S. 2008. Mostly Harmless Econometrics, Princeton University Press. Chapters 1 and 2.

Session 2: Introducing STATA – how to work with datasets
Topics covered:

- Importing and exporting data
- Cleaning data
- Dealing with strings
- Reshaping, merging, appending, collapsing, duplicates
- Exporting results
- Do-files and logs
- Loops

Day 2:

Session 1: Survey design
Topics covered:

- Research paradigm
- Sampling techniques
- Survey methods

- Questionnaire and question design
- Response rates

Required readings:

- Krosnick, John A. 1999. Survey Research. ***Annual Reviews of Psychology***, 50: 537-567.
- Berk, Richard A. 1983. An Introduction to Sample Selection Bias in Sociological Data. ***American Sociological Review***, 48: 386-398.
- Schaeffer, N.C., & Presser, S. 2003. The science of asking questions. ***Annual Review of Sociology***, 29: 65-88.

Assignment questions

- What is a good survey? What is a good survey question?

Voluntary readings

- ASR, Chapters 6, 9, 10

Session 3: Problems arising with survey data and alternative data sources

Topics covered:

- Internal and external validity
- Reliability
- Common method biases
- Secondary data
- Limitations to secondary data

Required readings:

- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. ***Journal of Applied Psychology***, 88(5): 879-903.

Day 3:

Session 1: OLS and hypothesis testing

Topics covered:

- Review of basic concepts
- OLS regressions
- Omitted variable bias
- Significance of parameters
- Goodness of fit

Required readings:

- Verbeek, M. (2004), A guide to Modern Econometrics, Wiley. Chapters 2 and 3.

Assignment questions:

- *What type of research questions can be addressed using OLS and which type of data does this require?*

Voluntary readings:

- ASR, Chapter 16
- Angrist, J. and Pischke, J.-S. 2008. Mostly Harmless Econometrics, Princeton University Press. Chapter 3.

Session 3: Dummy variables and interaction effects

Topics covered:

- Assumptions
- How to specify the models
- Interpretation of coefficients
- Illustrating interaction effects

Required readings:

- Brambor, T., Clark, W. R., & Golder, M. 2006. Understanding interaction models: Improving empirical analyses. ***Political Analysis***, 14(1): 63-82.
- Mehra, A., Kilduff, M., & Brass, D. J. 2001. The social networks of high and low self-monitors: Implications for workplace performance. ***Administrative Science Quarterly***, 46(1): 121-146.

Day 4:

Session 1: Logit/Probit and Ordered Models

Topics covered:

- Assumptions, Interpretation of coefficients, interaction effects, and measure of model fit for the following models:
- Logit/probit
- Ordered logit/probit

Required readings:

- Haas, M. R. & Hansen, M. T. 2005. When using knowledge can hurt performance: The value of organizational capabilities in a management consulting company. ***Strategic Management Journal***, 26(1): 1-24.
- Hoetker, G. 2007. The use of logit and probit models in strategic management research: Critical issues. ***Strategic Management Journal***, 28(4): 331-343.
- Zelner, B. A. 2009. Using simulation to interpret results from logit, probit, and other nonlinear models. ***Strategic Management Journal***, 30(12): 1335-1348.
- Verbeek, M. 2004. A Guide to Modern Econometrics, Wiley. Chapter 7.

Assignment questions

- Please skim the papers, but read the methods sections really carefully. What is being done there and why? Given what you know about these approaches, are the methods executed properly?

Session 2: Multinomial, Censored and Count-data Models

Topics covered:

- Multinomial Model
- Tobit models
- Negative Binomial & Poisson

Required readings:

- Shimizutani, S. & Todo, Y. 2008. What determines overseas R&D activities? The case of Japanese multinational firms. ***Research Policy***, 37(3): 530-544.
- Song, J., Almeida, P., & Wu, G. 2003. Learning-by-hiring: When is mobility more likely to facilitate interfirm knowledge transfer? ***Management Science***, 49(4): 351-365.
- Verbeek, M. 2004. A Guide to Modern Econometrics, Wiley. Chapter 7.

Assignment questions

- Please skim the papers, but read the methods sections really carefully. What is being done there and why? Given what you know about these approaches, are the methods executed properly?

Day 5:

Session 1: Resolving some common problems in quantitative studies

Topics covered:

- Multicollinearity
- Heteroskedacity
- Endogeneity
- Reverse causality
- Selection bias
- Instrumental variable models
- Non and semi-parametric matching methods (propensity score matching, difference-in-differences)

Required readings:

- Echambadi, R., Campbell, B., & Agarwal, R. 2006. Encouraging best practice in quantitative management research: An incomplete list of opportunities. ***Journal of Management Studies***, 43(8): 1801-1820.
- Hamilton, B. H. & Nickerson, J. A. 2003. Correcting for endogeneity in strategic management research. ***Strategic Organization***, 1(1): 51-78.
- Verbeek, M. 2004. A Guide to Modern Econometrics, Wiley. Chapter 5.
- Angrist, J. and Pischke, J.-S. 2008. Mostly Harmless Econometrics, Princeton University Press. Chapters 4 and 5.

Assignment questions

- What are the assumptions of OLS? What is endogeneity? What is selection bias? When and why do these issues appear, and how may they be addressed?

Session 2: Student presentations

Please see “Assessment” for more information.