Preliminary Syllabus for the Course: Econometrics 1

Michael E. Kummer[1]

[1] Georgia Institute of Technology, School of Economics
kummergtteaching@gmail.com

August 17, 2015

1 Course overview

- instructor: Michael E. Kummer
- office: # 222 (Old Civic Engineering Building)
- office hours: Mon 1.30pm. - 3.00pm.(t.b.c.) or after class/by appointment
- Class email: kummergtteaching@gmail.com
- class time Mon.: 5pm to 8pm (we start at 5.05)
- Location: Room 310 (Old Civic Engineering Building)
- Webpage:

2 Goals

This course introduces quantitative models for business decision-making. An important emphasis is put on practical application, and on the use of statistical software packages to analyze real-world datasets.

The course will review basic statistics, and build the ground for the workhorse econometric methods, and cover methods based on (pseudo-) experimental treatments. Special emphasis will be placed on identification and potential pitfalls of econometric methods.

The course should equip you to pursue your own empirical questions by choosing an appropriate research design and being aware of its strengths and limitations. After taking this course you should:

- Know the probability concepts and mathematical statistics that are needed in econometrics
- Distinguish different types of data formats (cross section, panel data, etc.) and how they can be analyzed
- Understand how to distinguish correlation and causation.
- Have hands-on experience with using econometric models to analyze real-world datasets.
- Be able to identify the challenges in empirically answering any given applied question.
- Devise an appropriate research design to answer the applied question understanding your design's limitations.

3 Course Participation Rules

Class meetings will be a combination of lectures on fundamental material, applied practice sessions and student exercises. Classes will mainly consist of lectures. You are asked to actively participate, however! The more questions you ask, the better.

I am currently developing the course and preparing new lecture slides. As these are work in progress, any feedback about typos, comprehensibility etc. is very much appreciated. You are also invited to tell me about any missing materials you would like covered (in more depth).

I would like you to determine a classroom-spokesperson. Any grievances with how the materials are taught, can be shared either with the spokesperson or with me directly. If you prefer to talk to the spokesperson, they can pass on your grievance in an anonymous fashion.

I plan to dedicate a few minutes in every class to small exercises which you should try to solve on your own, before I (or a volunteer among you) will present the solution on the board. After each section of class, I will review some of the material covered. These reviews are good occasions for you to ask any questions that have not been answered, yet.

Reading assignments will cover the core material and I expect that you will be prepared for class discussions. You should attend every class session, and notify me if you plan to miss a part of the lecture. Failure to do so will have a direct impact on class grade. If you think that you do not need to listen to the classes, please come and see me in my office hours, to discuss.

I will check my email once a day during the week (Monday through Friday). Please use the class-email and the special tag [**Econometrics_I_GaTech**] in the email subject header. I use this tag to make sure I process class email first. If you fail to include the special tag, or use a different email-address, I may not read the email for a long time.

4 Grading

The grade breakdown is as follows:

- Individual Class Participation and Quizzes - 10%
- Four Home Assignments and Class Projects - 40% (10 % each)
  - "Standard Assignments," consisting of a theoretical and/or a practical part.
  - "Presentation Assignment," which, again, consists of 2 parts:
    - active role: presenting a paper (8%)
    - passive role: summary of presentation (2%)
- Midterm Exam - 15 %
- Written Final Exam - 35%

Additional remarks:

- The minimum grade for midterm and final exam is 35 out of 100
- All assignments except for exams are to be submitted online.
- Exams will be similar to the regular problem-sets. You should therefore make sure you understand these well.
- Classroom Participation: There are several ways to earn credit:
  - Raising and answering Questions, presenting solutions to small exercises.
  - Participation in After-class Quizzes (1% per Quiz)
  - Presenting your solution to small home assignments.
- I encourage you to come to my office hours with any questions. I will not answer emails with questions on the material.
- If you need any special accommodations for physical or medical reasons, or if you see yourself unable to present the presentation assignment please see me after class or send me an email.

4.1 Participation

You are expected to attend every class session, to arrive on time, or five minutes earlier, and to remain for the entire class. Moreover, you are expected to follow basic classroom etiquette, which includes disconnecting all electronic devices for the duration of the class (unless otherwise noticed).

You are expected to participate in class discussions, have read the reading assignments, and I will assume you understand the material taught in previous lectures.

4.2 Homework

The hands-on tasks will be based on data that I will provide. You will analyze the data to get hands-on experience in formulating problems and using the various techniques discussed in class. You will use these data to build and evaluate policy or business related questions.

The assignments will involve both theoretical questions and practical implementation of the methods discussed in class. Most usefully they will build up to a research paper. For this you may use any of the following (matrix) programming languages: STATA, R, Matlab, Ox. The grade for each assignment will be on a five-grade scale: 10 pts: above 90%; 8 pts: above 70% correct, 6 pts: above 40% correct, 4 pts: below 40% correct (but attempted), 0 pts: not handed in. You will have at least a week to complete the assignments. The assignments and programming code must be submitted online. Answers will be discussed in the exercise sessions.

- Late submissions will suffer 2 point penalty in the first day. In the second and third days the penalty is 1 points each day (0-10 scale). After three days, homework is not accepted anymore.
- You are welcome, and in fact encouraged, to collaborate on any of these assignments (exams excluded). However, every one of you has to produce a separate write-up of your problem-set solutions and summaries. Identical write-ups will receive zero points.
- I encourage you to come to my office hours with any questions. I will not answer emails with questions on the material.
- All assignments except for exams are to be submitted online.
- You must have access to a computer where you can install software. If you do not have such a computer, please see me immediately so we can make alternative arrangements. You should bring your computer to class. We will help you install and configure the software in the first problem solving class.

4.3 Exams

The subject matters covered and the exact dates will be discussed in class.

5 Course Content

5.1 Part I - Review

The first part of this class is going to be a review of probability theory. You are expected to already know this material but I am still going to review it to make sure everybody is on the same page.

1. Applied Probability

(a) Basic concepts of probability: definition of probability, conditional probability, independence

(b) Random variables and their properties: definition, probability distribution, mean, variance, covariance

(c) Specific distributions: normal, chi-squared, F, t, binomial, Poisson

2. Statistical Inference

(a) Sampling distributions and Central Limit Theorem

(b) Point and interval estimation

(c) Hypothesis testing

3. Estimators

(a) Small and large sample properties of estimators

(b) Methods of estimation: methods of least squares, maximum likelihood estimation

5.2 Part II - Identification and Data Handling

After this review, I will discuss causality and identification, as well as different data formats and data handling.

Identification: Identification is about mapping population distributions of observed variables into statements about underlying structures or causal effects. I will introduce the potential outcomes framework, and discuss standard approaches to causal inference (randomized experiments, instrumental variables, controlling for covariates).

- What is identification, why care?
- What are the issues, which issue typically matters (or matters less) in which application?
- What are the general philosophical approaches, and which econometric approach can tackle which issue?
- Why discuss this together with data handling?

Data Handling:

- Getting a DataSet - How?
  - Some Tricks
- Making sense of different data formats:
  1. Cross-Section

2. Panel data
3. Independently Pooled Cross-Section
4. Time Series

- How to handle Raw Data.
  - Descriptive Analysis, Graphs, Raw correlations
  - Cleaning, Outliers
- Introduction to Regression Analysis
  1. Ordinary Least Squares Model
  2. Estimation
  3. Inference
- Building and Interpreting Linear Models
  1. Log-linear, Linear-Log and Log-log models
  2. Dummy variables, Non linear effects and Interactions
  3. Regressions coefficients in practice (causality or correlation)

5.3 Part III - Econometric Methods and Applications

1. Advanced Methods (selective coverage for dealing with endogeneity to search for causality)
   (a) Linear Models (extentions)
   (b) Instrumental Variables
   (c) Treatment Effect Models
   (d) Panel: Fixed Effects and First Differences
   (e) Nonparametric methods
   (f) Limited Dependent Variables

Depending on the direction our class discussion takes, I may not cover all material that is initially planed for any particular session. If the notes and the book are not adequate to explain a topic that I skip, you should ask about it by e-mail. I will be happy to follow up and provide you with additional references.

6 Course Readings

Course Material:

Part 1: **Casella, G. and Berger, R. (2001).** *Statistical inference.* Duxbury Press, chapters 1-4. [**CBS**]

Further Readings/Textbooks:

- **J. Blitzstein and Hwang J (2014).** *Introduction to Probability.* Chapman & Hall (Advanced undergrad text; many exercises) [**BHI**]
- **P. Billingsley (2012).***Probability and Measure.* Wiley (More advanced / mathematical than this class) [**BPP**]

Part 2: Angrist, J. and S. Pischke (2009), "Mostly Harmless Econometrics", Princeton University Press. [**MHE**] (Applied microeconomics perspective)

Further Readings/Textbooks:

- **Manski, C. (2003).** *Partial identification of probability distributions.* Springer Verlag. (Principled treatment of (partial) identification) [**MCP**]
- **Pearl, J. (2000).** *Causality: Models, Reasoning, and Inference.* Cambridge University Press. (Theoretical computer scientist on the notion of causality) [**PJC**]
- **Imbens, G. and D. Rubin (2014).** *Causal Inference in Statistics, Social, and Biomedical Sciences: An Introduction.* Cambridge University Press. (Forthcoming book by Imbens and Rubin, hard to get) [**IRC**]

Part 3: Hansen, B. (2015). "Econometrics", Lecture notes University of Wisconsin. [**HaL**]

For more intuition: Angrist, J. and S. Pischke (2009), "Mostly Harmless Econometrics", Princeton University Press. [**MHE**]

Selected Readings: What's New in Econometrics, G. Imbens and J. Wooldridge. [**WNE**]

Further Readings/Textbooks:

- Cameron, A. and P. Trivedi (2005), "Microeconometrics: Methods and Applications", Cambridge: Cambridge University Press. [**CTM**]
- Jeffrey Wooldridge, Introductory Econometrics: A Modern Approach, 4 ed. Thomson, [**JWB**]
- Mark Saunders, Philip Lewis and Adrian Thornhill, Research Methods for Business Students, 3 ed. Prentice Hall [**SLTB**]

I will complement the readings with discussions of applications, cases, and demonstrations. Whenever relevant, I will hand out lecture notes. Further details about the readings will be provided as we go. I expect that you to ask questions about any material in the notes that is not clear after the corresponding class and after reading the book.

7 Instructor Bio

Michael E. Kummer is a new Assistant Professor at Georgia Tech's School of Economics. He obtained his PhD from University of Mannheim, Germany, and he was a researcher at the Centre for European Economic Research (ZEW, Mannheim). His PhD from the Department of Economics was about ICT, Search and Market Outcomes. His primary fields are Empirical IO and Applied Econometrics, and his research interests are Information Technologies, Networks and Online Markets. Specific research topics have focused on social networks, peer production and collaboration in online settings, such as Wikipedia, and firm behavior in price comparison sites. Michael has published papers in the *European Economic Review* and *Information Economics and Policy*.

8 Office Hours

Office hours are planned to be Mondays 1h30am. to 3h00pm. Potential changes will be detailed in the course. Ideally, send your question per e-mail before you come.