Ec508: Econometrics

Syllabus and General Information
Professor Ani Dasgupta

Spring 2020
Boston University

Instructor: Dr. Ani Dasgupta
Office: Room 317, Economics Department, 270 Bay Street
Telephone: 353-5650
Email: AniruddhaDG@aol.com
Office Hours: TTh 8.00-9.30 PM and also by appointment[1]
Course Website: http://www.bu.edu/tech/services/teaching/lms/blackboard/
(*Navigate to Spring 2019 -> Ec509 site and enter BU userid and password when prompted*)

INTRODUCTION: Ec508 is an entry-level *graduate* econometrics course, focusing mainly on cross-sectional techniques[2]. It is entry-level in the sense that you are not presumed to have any prior acquaintance with econometrics, although you are assumed to have the required statistical background at Ec507 level and coursework in linear algebra and calculus including some knowledge of multivariable optimization. You are also supposed to have basic familiarity with a statistical software such as STATA or R. Though entry level, this is very much a graduate course, which among other things, means that *rigor and understanding of the techniques are very much emphasized as opposed to learning cookbook methods*. It attempts to serve two types of audiences. For those who wish to pursue applied data analysis in the real world, it presents a wide array of problem instances and tools appropriate for those instances. I will expose you to a tool, show you why it works (at least in most cases) and ask you to apply the tool to solve similar problems with new datasets. It also serves as a stepping stone for those interested in knowing the field more intimately and perhaps going on to do a Ph.D. in Economics, which it does by introducing them to a fair amount of theory and by giving them a tour of a small selection of classic and contemporary papers written in Econometrics.

REQUIREMENTS, GRADING, DATES AND GROUND RULES: For most courses, you learn only a small fraction of the material by just listening to the lectures. For econometrics, that fraction could be really small. You are expected to read the lecture

---

[1] On Tuesdays and Thursdays, about 15 minutes after the lecture, I will hold office hour in a classroom with questions asked and answers given publicly for all who care to stay back and attend these sessions. If you wish to speak with me privately, you can meet with me individually by making an appointment. Appointments may be available on Fridays and weekends if need be.

[2] This course serves as the core econometrics requirement of the departmental master's program. Unlike some other sections of the course, this one is fully matrix-algebra based and offers students a research paper writing option.

material and suggested sections of certain texts immediately after every lecture, working through the derivations yourself (yes, we do care about derivations). Most of the time, you will be expected to read preparatory material before a lecture, and you will need to come prepared to the lecture having read that material. When problem sets are distributed, you are expected to try out each and every problem.

Your course grade will be based on 2 quizzes, a final, problem sets, and an optional term project and class performance. The weights for the two tracks (project and non-project) are displayed below:

Item	**Weight (in %)**	
	Project	**Non-Project**
Quizzes	20	30
Final	30	35
Problem Sets	20	25
Project	20	-
Class Performance	10	10

The first quiz is an in-class mostly theory-based exam. The second quiz is a mostly computation-based exam where you will be tested on your software skills. The final is a cumulative two-hour exam and it should prepare Econ MA students for their comprehensives. Doing well in these tests will require that you have mastered the problem sets of which there will be about eight in all. These are to be done in groups, each group having up to three members.

In addition, if you choose the project track, you are required to turn in a group project at the end of the semester with each group again having up to three members (if you will really like to go solo, you can – but talk to me about this first). The project can be of one of two types. You can either choose to read and report on a collection of papers written on a particular area. The second and the preferable option is to embark on an independent empirical analysis on a problem chosen by you. Of course, I will be available for consultation throughout the life of the project; however, the choice of the problem and collection of appropriate data will be your responsibilities. I will at some point during the semester circulate some project reports done by former students so that you can have an idea about what these are supposed to look like. You will have until the middle of the semester to decide on whether to take the project option or not. The final component of grading is a subjective class performance item, where I will take into account such intangibles as attitude and effort.

Here are some important dates. The first quiz will be held on **Thursday, February 27th.** In case you choose the project option, I require a proposal that is due latest by **Thursday, March 17th** (however you are encouraged to submit it earlier). I will need at least a three-page proposal where you have demonstrated having given serious thought to your ideas and perhaps having done some preliminary library/internet research on literature and data sources as well. Please be forewarned that if you miss this deadline, two adverse things will happen. First, you will lose points on your project grade and second, you will have less access to me as a consultant compared to other groups which were on time. The second

quiz, at this time is tentatively scheduled on **Friday April 14th**. The project report is due the last class day on **Tuesday, April 30th.** For the final exam we will follow the slot scheduled by the registrar's office and unless the entire class is agreeable with and keen on shifting the time, I will stick to it. Make-ups for quizzes and exams may be offered only in case of a serious medical emergency (you need to inform me before the exam about it and a doctor's certificate to the effect that you were unwell enough not to be able to make it to the exam is required). The format of the make-up(s) will be at the instructor's discretion. I take honesty and integrity issues very seriously. Any violation of the academic conduct code will be referred to the authorities, and if the individual concerned is found guilty, it will automatically result in an F for the course and perhaps other unpalatable actions.[3] If you wish to know more about what constitutes a violation and what the due process is, please refer to the CAS Academic conduct code.

MEETING TIMES: Please note that we will normally meet Tuesdays and Thursdays 6.30 pm to 7.45 pm (except when officially a class is cancelled). In addition to these lecture periods, each week I will hold a 'study/lab sessions' on Fridays at 5.45 pm (some weeks, instead of me, the TA will hold them). These sessions will be used either to provide software tutorials, or to go over Problem Sets/Exams, and on some rare occasions to wrap up lecture material. You must attend these sessions.

TEXTBOOKS AND STUDY MATERIAL: There is no textbook for the course. I have given up on finding a text suitable for the Masters level and rely on my posted/distributed notes entirely. In the past I have used the text **Econometric Methods with Applications in Business** by **Christian Heij et. al.** (Oxford University Press) for this course. This textbook is a nice introduction to econometrics and fairly modern in its coverage, although it is not quite as rigorous as I would have liked it to be. Its strength lies in the fact that it discusses numerous empirical exercises and provides useful datasets for you to try those exercises on your own.

Here are a few other texts and material that you may consider accessing. A popular and nice undergraduate text is **Jeffrey Wooldridge's Introductory Econometrcs** (SouthWest College Publications). This book does not use matrix notation and does not delve much into theory, but it does a great job of providing intuition and economic context. It also has numerous empirical exercises. Similar in spirit and scope is **James Stock and Mark Watson's Introduction to Econometrics** (Addison Wesley). A wonderfully lucid, matrix-oriented book at this level (though mostly for classical topics) is **Econometric Methods** by **Jack Johnston and John DiNardo** (McGraw-Hill). If you can find it, I recommend the older third edition (written by Johnston alone) in addition to the fourth (which is coauthored with John DiNardo). Finally, if you are looking for a book at undergraduate level, but with

[3] Please be aware of that getting unauthorized help for problem sets (e.g. accessing and using solution sets distributed to previous years' course participants) constitutes violation of academic conduct.

good coverage of time series, you can try **Econometric Modelling with Time Series** by Vance Martin, Stan Hurn and David Harris (Cambridge).

A book that is not recommended as a stand-alone textbook but one that gives you an interesting perspective and overview of the classical topics in the field is **Peter Kennedy**'s **A Guide to Econometrics** (MIT Press). For those of you with an applied bend, I strongly suggest that you take a look at **Ernst Berndt**'s **The Practice of Econometrics: Classic and Contemporary** (Addison Wesley). For each chapter, the book chooses an area of interest to economists and illustrates empirical analysis in that area using a particular technique. I will use some applications from this book to illustrate the theoretical models.

Going into more advanced texts, **William Greene'**s **Econometric Analysis** (currently in its 8th edition) is probably one of the best compendium of modern econometric techniques and if you want to keep one reference econometrics book in your library, Greene will probably the best choice. Unfortunately, its explanations are quite terse and sometime unintuitive, reflecting, over time, the consequence of incorporating newer and newer material without making the book weigh like an iron chest (it already has 1100+ pages!). I will refer you to Greene from time to time though both for study material as well as exercises.

If you want to know classical econometric techniques well, an excellent choice is **Paul Ruud**'s **An Introduction to Classical Econometric Theory** (Oxford), which provides most detailed step-by-step explanations for the classical technique. Its coverage of least squares theory, in particular, is exhaustive. Two other serious Ph.D. level texts in the market are **Jeff Wooldridge's Econometric Analysis of Cross Section and Panel Data** (MIT Press), and **Russell Davidson and James McKinnon's Econometric Theory and Methods** (Oxford)**.** Both these texts make the reader think beyond the theorem-proof approach. Wooldridge, although mainly a theory text, is particularly suited to seeing how theory can be taken to data offering many pertinent examples. Davidson-Mackinnon is very strong in theory, particularly in the geometrical underpinning behind algebraic equations. Another very good PhD level text which focuses on cross-sectional methods is **Microeconometrics: Theory and Applications** by **A. Colin Cameron and Pravin Trivedi.** The latest entrant in the market for PhD level econometrics is **M. Hashem Pesaran's Time Series and Panel Data Econometrics** which should be of interest to students who wish to specialize in Macroeconometrics.

My personal favorite Ph.D. level text, which is more textbookish than Wooldridge and Davidson-Mackinnon and more modern in approach than Ruud while being amazingly thorough and readable with plenty of stimulating exercises, and covering both cross-sectional and time-series techniques more than adequately, is **Fumio Hayashi**'s **Econometrics** (Princeton University Press). It's a hard read, but if you aspire to get into a Ph.D. program someday, you will need to learn the content of this book. In my lectures, I will make use of material from all the above-mentioned texts.

These days, one can learn a lot in almost any field from surfing the web; econometrics is no exception. You can download great set of lecture notes written by noted econometricians and find out what is being taught at other prestigious graduate programs. Professor **Bruce Hansen** (U Wisconsin Madison) has a nice set of lecture notes at http://www.ssc.wisc.edu/~bhansen/econometrics/ while Professor **Herman Bierens** of

Penn State has another great set of notes on various topics at
http://personal.psu.edu/hxb11/LECNOTES.HTM .

SOFTWARE: The course places heavy emphasis on using software to analyze data, so students need to get up to speed with this aspect of the course real soon. The official software for this course is **Stata**; however in the past I have also worked with **R**, **LIMDEP** and **SAS**. Any of these software should be adequate for this course, but note that a) all your quiz and Problem Set related submissions must be done in Stata and b) if you wish to use something other than Stata, please be forewarned that when you get stuck with something, you will be on your own.

Students from Ec507 course should be already quite familiar with **Stata,** which has numerous canned packages plus a fair bit of programming capabilities. If you were not in Ec507 last semester or have not already purchased it then, note that as a BU student, you may buy a perpetual license for the standard version (the "Intercooled version") at an academic price of $198 (go to www.stata.com and look under 'gradplan' for details). They also have an annual license for a reduced price of $89. The current version is Stata 16 but if you happen to have Stata 15 (or even STATA 14), you will be able to do everything you need to do in this course.

Since, most of you have been introduced to Stata via Ec507, I will not waste class time here re-introducing you to the package; however, in case you were not a student in that course, I will post my introductory material on Stata to get you started. There are also numerous (free) tutorials that are available both online with the package and on the net some of which will be posted on our website. In this course, we will take Stata programming to somewhat higher level, and in particular, spend considerable time doing maximum likelihood estimation with Stata. At the beginning of the course, some dedicated sessions will be offered on Stata in the areas of matrix manipulations and simulation as well.

There are several free alternatives to STATA that are downloadable from the net. In particular, I must mention **R**, which is a clone of S/SPLUS, and is also a terrific piece of software with great graphics and programming options. To download a free copy, go to: http://cran.r-project.org/. R like Stata, is very much a live software in the sense that people are all the time contributing their programs for free usage. Also, Springer has published a whole collection of books on R, some of which may be accessible online via BU library.

A fair question at this stage is if several statistical/econometric software are available for free why am I recommending that students purchase STATA? There are several reasons for that. First, not all packages are written with all disciplines in mind (they cannot be). Stata has a definite econometrics audience in mind, and several canned packages specifically needed by econometricians are available in Stata (and nowhere else); in fact some of the folks in charge of maintaining the repository of Stata programs are econometricians. But at the same time, Stata allows you to write your own distributable programs which, as far as I know, among the freely available packages, only R does. Finally, note that one problem with free software is that you do not get technical support if things go wrong. I have had reasonably good

experience with Stata's technical support and I suspect that from time to time you might need it too.

TOPIC LIST: The following is a list of areas I wish to cover in this course (reference to reading material besides lecture notes being provided in parentheses[4]). This is indeed a wish list, for the list is long and hence, tentative; we will almost certainly skip some topics to concentrate on others. Topics 1-8 make up the core part of the course and hence, we will focus on them. Of the remaining topics, time constraints, class interests and relevance to ongoing student projects will determine which ones we will discuss. If you are very interested in a particular topic, please let me know.

Topic 0 (Introduction):

- Introduction to the course and the subject of econometrics

Topic 1 (Linear Algebra):

- Basic Linear algebra and matrices including a thorough treatment of eigenvalues and eigenvectors (G Appendix A, J Chapter 4)

Topic 2 (Statistics):

- Univariate statistics refresher (use your Ec507 text)
- Multivariate Statistics (T Chapter 1, G Appendices B and C, J Chapter 4)

Topic 3 (Classical Linear Model):

- The multiple regression model (T chapter 2 for background, H 3.1, G Chapter 2, H Chapter 1)
- Least squares geometry (T 3.1, G Chapter 3, DM Chapter 2, H Chapter 1, R Chapters2, 3)
- Finite sample properties (T 3.2, G Chapter 4, DM Chapter 3, H Chapter 1, J Chapter 5, R Chapters 6-8)
- Inference (T 3.3-3.4, G Chapter 5, H Chapter 1, J Chapter 5, R Chapters 9-11)
- Application: Returns to scale in electricity (Marc Nerlove's paper, H Chapter 1, B Chapter 3)
- Dummy variables (T 5.3, G Chapter 6, J Chapter 6)
- Parameter consistency and structural stability (T 5.3, G Chapter 6, J Chapter 6, 10)

---

[4] B refers to Berndt, CT refers to Cameron-Trivedi, DM to Davidson-McKinnon, H refers to Hayashi, G refers to Greene (8th ed), J refers to Johnston (3rd ed), JD refers to Johnston-Dinardo (4th ed), and R refers to Ruud, and T refers to the Heij et. al. text. You are only responsible for reading my lecture notes and any original paper I ask you to read.

- Application: Solow's 1957 study of technical change (G Chapter 6)

Topic 4 (Model Specification):

- Consequences of Under and Over Specification (T 5.2, G Chapter 5)
- Model Selection Criteria (CT 8.5)
- The pitfalls of data mining (Lovell Paper)

Topic 5 (Large Sample Theory):

- Convergence Concepts (T 4.1, G Appendix D; H Chapter 2; R Chapter 13, B. Hansen Lecture Notes)
- Classical regression model with large samples (T 4.1, H Chapter 2, G Chapter 4, R Chapter 13)
- Application: Semiparametric inferential techniques (excerpts from Adonis Yatchew's book on this topic)

Topic 6 (Instrumental Variables):

- Causes of Endogeneity (T 5.7, H 3.1, 3.2, G Chapter 8)
- IV and 2SLS Estimation (T 5.7, G Chapter 8)
- Tests with IVs (T 5.7, G Chapter 8, DM 8.5-8.7)
- Weak instruments and too many instruments (CT 4.9)
- Application: Returns to Schooling (Angrist and Krueger paper, Ashenfelter and Krueger paper)

Topic 7 (Nonlinear Models and More General Estimation Frameworks):

- Nonlinear models: NL regression, Maximum Likelihood Estimation and Generalized Method of Moments (T 4.2, H Chapters 4 and 7, DM Chapters 6, 9 and 10, G Chapters 12, 13 and 14, R Chapter 14)
- Testing via the trinity (T 4.3, G Chapter 13, 14, R Chapters 17)
- Application of ML: stochastic frontier model; Aigner, Lovell and Schmidt's study of metal industry (G Chapter 14, original article)
- Application of ML: Who wrote "And Quiet Flows the Don"? (Claeskens and Hjort book on Model selection)
- Application of GMM: Consumption and Asset Pricing Models (Hall, Hansen and Hansen-Singleton Papers)

Topic 8 (Data Problems):

- Generalized least squares (T 5.4, G Chapter 9)
- Heteroskedasticity (T 5.4, G Chapter 9; R Chapter 18)
- Application: credit card expenditures (G Chapter 9)

- Autocorrelation (T 5.5, G Chapter 20; H Chapter 2, 6; R Chapter 19)
- Application: Fama's test of efficient market hypothesis (Original article; H Chapter 2)

Topic 9 (Multiple Equation Systems):

- System of equations and SUR (T 7.7, G Chapter 10, J Chapter 8)
- Application: estimation of interrelated factor demands (B Chapter 9)
- Identification in simultaneous equation modeling (T 7.7, G Chapter 10, J Chapter 11)
- Estimation in simultaneous equation systems (T 7.7, G Chapter 10, J Chapter 11)
- Application: Improving the efficiency of a queuing system; Study of Calcutta port (Dasgupta-Ghosh's original article)

Topic 10 (Discrete Choice and Limited Dependent Variable Models):

- Binary choice models (T 6.1, G Chapter 17)
- Multinomial choice models (T 6.2, G Chapter 18)
- Ordered data and count data (T 6.2, G Chapter 18)
- Censoring, truncation and tobit models (T 6.3, G Chapter 19)
- Sample Selection (T 6.3, G Chapter 19)
- Application: College Choice Decision (Bridgett Long's article)

Topic 11 (Panel Data Models):

- Fixed effects (T 7.7, G Chapter 11, H Chapter 5)
- Random effects (T 7.7, G Chapter 11, H Chapter 5)
- Dynamic panels (G Chapter 11)
- Application: Growth and convergence (Hayashi Chapter 5)