Julie Schaffner

Office: Mugar 252A

Julie.Schaffner@tufts.edu

Fall 2016

**EIB E213**

**Econometrics**

*Course Description.* This course introduces students to econometric methods, which are the primary tools of quantitative data analysis employed in the study of economic, social and political relationships, and in policy analysis and program evaluation. The course equips students with the facts, intuition, and experience necessary for critical reading of econometric research produced by others and for independent econometric research. It introduces students to the wide range of models that may be estimated using Ordinary Least Squares and probit estimation methods, and highlights the challenges of obtaining estimates that are unbiased and precise. It pays special attention to the approaches researchers use to obtain estimates that are free of omitted variables (or endogeneity) bias, including the use of randomized control trials and difference-in-differences, fixed effects and instrumental variables estimation. It also addresses problems associated with measurement error, heteroskedasticity, autocorrelation, and multicollinearity.

*Textbooks.* Students are required to read at least one of the following two textbooks. One option for the required text is:

Wooldridge, Jeffrey, *Introductory Econometrics: A Modern Approach*. South-Western College Publishing, 5th edition (2012). [I provide page numbers below for the 5th edition, but using the 4th or 6th edition is also fine.]

The second option for the required text is:

Bailey, Michael A., *Real Econometrics: The Right Tools to Answer Important Questions*. Oxford University Press (2017).

In choosing which of these to read (or whether to read both), students should keep in mind the following strengths and weaknesses. The Wooldridge text, which has been the main textbook for this course for many years, is widely used, is more complete, and digs deeper. It serves as a better stepping stone to more advanced econometrics courses, and would serve as an excellent reference in the future, offering introductions to many topics beyond what we will have time to cover in this class. The downside for the Wooldridge text is that the writing is pretty terse. Some readings are only a few pages long, but can take significant time and effort to read.

The Baily text is somewhat easier and more entertaining to read, and offers many interesting examples that are spelled out in practical detail. The downside is that, while it covers most of the material I will cover in class, it does not cover quite all of the material (so that there are some topics or important points within topics that are covered in lecture but not covered in this text).

It is absolutely essential that you make the effort to read one of these textbooks, even if you think you are following the lecture material adequately. This is your opportunity to learn the written language of econometrics. As with learning any language, reading goes slowly at first, as you learn the vocabulary and symbols. If you make consistent effort, by the end of the semester you will find that your reading speed and comprehension have improved greatly. This will empower you to learn more from the econometric research papers that you read for other classes or in your own research.

In the course outline below, I also provide references for these supplementary textbooks:

Wheelan, Charles, *Naked Statistics: Stripping the Dread from the Data*. W.W. Norton and Company, 2013.

Angrist, Joshua D. and Jörn-Steffen Pischke, *Mastering ‘Metrics: The Path from Cause to Effect,* Princeton University Press, 2014.

The readings in these books (which are on reserve in Ginn library) are recommended rather than required, but I strongly that suggest you check them out. Some students find these books very helpful.

*Materials on Trunk.* My lecture slides, as well as handouts, problem sets, data, study questions, video presentations, and quizzes, will be posted on Trunk. Please check the Trunk home page frequently for announcements and click on the “Study Guide and Course Materials” link to find out how to prepare for each class (and how to make sure you understand the material after class). If you have any difficulty accessing the Trunk materials, please send me an email.

*Assignments.* Course work includes online quizzes, problem sets, and a group “mini-project,” as well as in-class midterm and final examinations. Online quizzes must be completed individually. They are open book but timed, so it is important to study well before taking them. In most cases I provide practice quizzes that you can use to check your level of preparation before taking the quiz.

The problem sets, which students complete in study groups of 4 or 5 students, require students to use the popular statistical software package, STATA. I intend the problem sets to be good vehicles for learning. If you and your study group get stuck or run into questions while working on a problem set, please come to office hours or ask me a question via email.

During the second half of the semester, study groups will complete mini-projects, which require them to identify an econometric research question and specify a related econometric model, obtain relevant data, prepare the data for econometric analysis, perform estimation and testing, and interpret the results. I will help your study groups carve out feasible and interesting project topics. *Please note that you must attend an extra session at which your group will present its project findings on Friday, December 10, or Saturday, December 11.*

Many students have excellent study group experiences and find that their conversations in study groups greatly enrich their learning in this course. Study group experiences are excellent, however, only when all members consistently participate with good will and groups work at maintaining good work and communication habits. I will encourage you to have important conversations and establish good habits early on. *I will also ask each student to submit a self and peer evaluation of study group contributions both midway through the semester and at the end of the semester.*  I hope that students’ knowledge of my interest in the health of their study groups will serve as a friendly nudge, encouraging them to follow through on their study group commitments, even when the semester gets very busy!

*Exams.* The midterm examination is scheduled for Wednesday, October 26. The final exam is scheduled for Tuesday, December 20. *Please note these dates now and plan your travel around them.* The midterm and final examinations will be closed-book, in-class exams, but students will be allowed to prepare and use one 3" by 5" index card of formulas and notes for each exam.

*Grading*. Calculation of final course grades will employ roughly these weights:

Quizzes, problem sets, and *class participation* 20

Mini-project 15

Midterm exam 30

Final exam 35

*Pre-requisites.* Previous study of introductory statistics (at the level of EIB B205) and basic multivariable calculus (at the level of EIB E210m) is required. Introductory economics is strongly recommended. All relevant statistical concepts will be reviewed as they become necessary, but the reviews will be brief (and not adequate for people who have not studied the material carefully in the past). If you would like to assess whether your math and statistics background is adequate for this course, I suggest that you read through Appendices A, B and C in the Wooldridge textbook. If you have studied this material before and you are comfortable with it, then your background is adequate.

I have broken the pre-requisite material down into four sets of concepts and have posted review materials and mandatory quizzes under four headings on Trunk: Stats Reviews 1, 2 and 3, and a Math Review. *Please note that the deadline for completing the Stats Review 1 Quiz is Monday, September 12, before class.* You should feel welcome to complete all of the Stats Review material before the semester begins. If you would like access to these materials in Trunk before the semester begins, please send me an email and I will register you for the Trunk site.

*The learning process and work load expectations.* Students enter this class with diverse backgrounds and strengths. I provide a range of study resources with the aim of meeting the needs of diverse students. All students are required to attend class and participate in class, read a textbook, view required videos, and review their notes while preparing for quizzes or completing problem sets. These are the bare minimum activities, and represent the work that students with strong backgrounds in math, statistics and economics must do to gain command of the concepts and techniques. Students who wish to push further are encouraged also to read some of the academic journal articles I have posted, from which I have drawn class examples.

Students who enter with somewhat weaker backgrounds must plan to work harder and devote more time to the course, using some of the recommended and optional study materials as well. These include study questions for each topic (which help you review the main points of the lectures), a number of recommended videos (which you may want to view both before and after the relevant classes), and references to readings in the recommended textbooks. I think it is wonderful that many students stretch themselves to take this important class, despite some trepidation about quantitative endeavors! *I invite and urge you to come to office hours with questions, or even just to check in to make sure you are on top of the material.*

*A cautionary note:* Students sometimes get lulled into complacency because they come away from lectures feeling that it all makes good sense. Unfortunately, feeling this way is not the same as understanding the material and being able to apply it. If you are attending lecture but are not also reading the textbook, reviewing the material, and working through the problem sets on your own before discussing them with your study group, you are probably not learning the material well and will not get the grade you hope for on the midterm!

*Accommodations.* In accordance with federal and state law, Tufts University provides for reasonable accommodation to students with documented disabilities. If you believe you have a disability requiring an accommodation, please contact the Registrar, Goddard 212, (617) 627-2405.

## Course Outline

W = Wooldridge, *Introductory Econometrics: A Modern Approach,* 5th edition

B = Baily, *Real Econometrics*

NS = *Naked Statistics* by Wheelan

MM = *Mastering ‘Metrics* by Angrist and Pischke

You are required to read either W or B. NS and NM are recommended supplementary readings.

Please refer to ***the Study Guide and Course Materials Page on Trunk*** for the dates on which we will discuss each of the following topics, and for links to videos, lecture slides, study questions and other course materials.

1. **Introduction: What is econometrics? What is it good for?**

###### W, Chapter 1

B, Chapters 1 and 2

NS, Introduction

**Statistics review 1: Probability, random variables and their distributions, samples and estimators**

W, Appendix B, p.722-745

NS, Chapter 2, Chapter 4, and pages 68-82 of Chapter 5

(B does not provide a thorough review of this material)

**Statistics review 2: Desirable properties for estimators: unbiasedness, efficiency and precision**

W, Appendix C, p.755-770

NS, Chapter 8

(B does not provide a thorough review of this material)

**II. Simple (or two-variable) linear regression models**

### Introduction to simple linear regression models and to Ordinary Least Squares (OLS) as a method for fitting the models to data

#### Pindyck, R. and D. Rubinfeld, *Econometric Models and Economic Forecasts (4th edition)*, Irwin McGraw-Hill, 1998, Chapter 1. (on trunk)

W, Chapter 2, p. 22-43

W, Appendix A, p.703-710

B, Chapter 3, p.45-53

1. The R-squared goodness of fit measure

W, Chapter 2, p. 38-39

B, Chapter 3, p.70-73

C. Simple linear regression model with a dummy variable on the right hand side

W, Chapter 7, p.227-230 [Note: Reading these pages is recommended, rather than required, at this point in the course. We will study the use of dummy variables at greater length later, after we have learned a lot more econometrics. You will find these pages easier to read then.]

B, Chapter 6, p. 167-178 [Note: Same as for the Wooldridge reading.]

D. Ordinary Least Squares estimators as random variables

B, Chapter 3, p. 53-56 (stop at top of page)

E. Conditions under which OLS gives unbiased, efficient and precise estimates of causal effects

W, Chapter 2, p.45-57

B, Chapter 3, p.57-70 (This covers conditions for unbiasedness and precision, but not in as structured a way. It does not deal with efficiency.)

**Math review: Functions and Derivatives**

W, Appendix A, p. 705-719

(B does not have a thorough review of this material)

**III. Multiple (or K-variable) linear regression models**

### Introduction to multiple linear regression models

W, Chapter 3, p.69-72; Chapter 6, p.205-206

B, Chapter 5, p. 128-138

B. Ordinary Least Squares and goodness of fit in multiple regression models

W, Chapter 3, p.72-80

B, Chapter 5, p.150-151

NS, Chapter 11

1. Conditions under which OLS gives unbiased, efficient and precise estimates of causal effects

W, Chapter 3, p.83-103

B, Chapter 5, p. 147-150 (This covers conditions for precision, but not unbiasedness and efficiency.)

1. Fitting non-linear models using functional transformations of left and right hand side variables

W, Chapter 2, p.41-44; Chapter 6, p.191-200

B, Chapter 7, p. 207-226 (This doesn’t cover models with interaction terms.)

1. Dichotomous (dummy) independent variables

W, Chapter 7, p.227-248

B, Chapter 6

1. Units of Measurement and re-scaling

W, Chapter 2, p.39-41; Chapter 6, p.186-191

##### Statistics Review 3: Common families of statistical distributions and hypothesis testing

W, Appendix B, p.745-752, and Appendix C, p.770-790

NS, Chapter 9

(This would also be a good time to read NS, Chapter 10, but it is not required.)

##### IV. Interval estimation and hypothesis testing

1. Motivation for interval estimation
2. OLS under the normality assumption

W, Chapter 4, p.118-121

C. OLS in “large samples”

W, Chapter 5, p. 168-178 [This is only recommended, not required.]

D. Confidence intervals/ interval estimation (assessment of precision)

W, Chapter 4, p. 138-140

B, Chapter 4, p. 117-119

E. Testing hypotheses about a single parameter: the t test and statistical significance

W, Chapter 4, p.121-138

B, Chapter 4, p. 91-111

F. Appreciating the distinction between statistical significance and economic importance

W, Chapter 4, p.135-138

B, Chapter 4, p. 115-116

G. Testing hypotheses involving several parameters: the F test

W, Chapter 4, p.143-154

B, Chapter 7, p.226-234

H. Presentation of regression results

W, Chapter 4, p.154- 156

Handout on the construction of good tables

**V. Models for dummy dependent variables**

1. OLS when the dependent variable is dichotomous

W, Chapter 7, p. 248-253

B, Chapter 12, p. 401-406

1. Probit and logit regression models for dummy dependent variables

W, Chapter 17, p.584-588

B, Chapter 12, 406-417

1. Interpreting coefficients in probit models

W, Chapter 17, p.589-596

NS, p. 29 (on percentage point change versus percentage change in rate)

[B, Chapter 12, p. 418-427. Bailey takes a different approach here than we will take in class. The method we emphasize is discussed in a note to Chapter 12 starting on the bottom of page 546.]

1. Interval estimation and hypothesis testing in probit models

W, Chapter 17, p.587-588

B, Chapter 12, p. 427-434

**VI. Introduction to model specification concerns**

1. Overview of specification choices

B. Overview of specification errors

C. The importance of developing a conceptual framework to guide model specification (*before* you start analyzing the data)

**VII. Omitted variables bias and approaches to identifying causal effects**

A. Nature of the problem and description of its consequences

W, Chapter 3, p.88-93

B, Chapter 5, p. 138-144

MM, Chapter 2

B. The Program Evaluation Problem as a special case

W, Chapter 7, p.253-256

NS, Chapter 13

C. Using randomized control trials to eliminate bias

B, Chapter 1, p. 18-22 (optionally, you can read more about RCTs in Chapter 10)

MM, Introduction and Chapter 1

D. Using difference-in-differences estimation in repeated cross section data to eliminate bias

W, Chapter 13, p. 449-459

B, Chapter 8, p. 268-274

MM, Chapter 5, esp. 178-191

E. Using fixed effects methods in panel and pseudo-panel data to eliminate bias

W, Chapter 13, p.459-477; Chapter 14, p.484-492

B, Chapter 8, p. 247-262 (Note that Bailey takes a somewhat different approach to introducing these models than we will take in class.)

F. Using instrumental variables techniques to eliminate bias

W, Chapter 15, p. 528-532

B, Chapter 9, p.287-306

MM, Chapter 3

G. Using proxies, narrowing samples, checking related empirical patterns, and other ways of reducing or ruling out large omitted variables biases

W, Chapter 9, p.308-313

XIII. **Other problems with dependent and independent variables**

A. Including irrelevant variables

W, Chapter 3, p.88

B, Chapter 5, p. 151

B. Measurement error

W, Chapter 9, p.317-323; Chapter 15, p.532-534

B, Chapter 5, p. 144-146

C. Implications of the problems studied thus far for model specification in practice

**IX. Errors that are not independently and identically distributed**

A. Heteroskedasticity

W, Chapter 8

B, Chapter 3, p.67-68 (this covers only a small fraction of the material we will discuss in class)

B. Autocorrelation in time series data and clustered samples

W, Chapter 10, p.344-348, Chapter 12, p.412-416,431-434

B, Chapter 13, p.449-465 (This covers autocorrelation in time series but not clustered samples.)

**X. Multicollinearity and small samples (and how they might limit what we can learn from particular datasets)**

A. Perfect multicollinearity

B. Near perfect multicollinearity and small samples

W, Chapter 3, p.94-98

B, Chapter 5, p. 148-150