Syllabus for Econ 281 - Introduction to Applied Econometrics

Instructor: Richard Walker

Winter 2010

Email: richard-walker@northwestern.edu

Office hours: Tuesdays, 9-11am in my office (AAH 335).

We will meet three times a week: Mondays, Wednesdays and Fridays, at 11am in Leverone Auditorium.

1 Course Overview

This is an introductory course in econometrics. The aim is to equip students with the basic tools needed to estimate and analyse empirical relationships betweeen economic variables. It will be a mixture of theory and practice. The theory is essential if you are to understand why you're doing what you're doing. The practice will help you get a feel for the type of software used in empirical economics, and for the correct interpretation of the results such software generates. So, there will necessarily be a significant theoretical component to the course, but you will occasionally get your hands dirty with real data.

Prerequisites for the course are 201, 202, MATH 220, STAT 210 or equivalent. Of these STAT 210 is most important; get in touch with me if you have not done this course already.

2 Evaluation

- Three assignments (due on Fridays of January 22nd, February 12th and March 5th). (10% each).
- Two midterms on Friday January 29th and Friday February 19th. (20% each)
- Final exam on Monday March 15th at 9am. (30%)

Note that while the final exam will focus on the final third of the course, it will also require students to draw on their knowledge of the earlier material.

3 Sections

Discussion sections are currently scheduled for Tuesdays and Thursday mornings. There is a possibility that one of the Tuesday sections will be moved to an alternative time, but this remains to be arranged with the Registrar. We'll talk about this, and about allocation of students to sections, when we first meet. Blackboard has a signup feature which we will use to allocate students to sections.

The TAs are Erik Loualiche, Cristina Tealdi and Thiago Teixeira Ferreira. They will announce their office hours separately.

Some students have asked me if the sections are mandatory. They are not, at least insofar as you will not be directly penalised for not attending. However, the sections will be very useful. There will be exercises for each section that I **strongly** advise you attempt beforehand. Econometrics is one of the subjects where practice is essential for understanding. You will get a lot more from the course, and find the midterms and final much easier, if you try the exercises before you see the answers. Also: answers to any section discussion questions from the textbook will **not** be published, so if you don't attend sections you should make sure you get the notes of someone who does.

4 Course materials

The textbook will be '*Introduction to Econometrics*' [3rd edition] by Christopher Dougherty. We will be following this book very closely, and it will serve as the blueprint for the course. Norris bookstore should have the book in stock, but it's also available on Amazon. It is absolutely essential that you get your hands on a copy. It is worth talking about the book and the supporting material in some detail.

This is an excellent textbook. Dougherty has taught the 281 equivalent at the London School of Economics since before I was your age, and essentially the book comprises his lecture notes. It is also used by the University of London External Programme, which means it is the textbook for overseas students doing 'distance learning'. As such it has some features that I think are valuable.

The first is that it is one of the few econometrics textbooks that is explicitly designed to be read. By that I mean 'read' in an almost novelistic sense; you can actually sit down and read a chapter without your head hurting (too much). Note that this doesn't mean there are no equations. There are plenty. However, Dougherty talks the reader through the equations in a more explicit manner than in other books, and provides the underlying intuition wherever possible.

The second feature is the body of supporting material available at Dougherty's website, the link to which is provided below. This includes a study guide which has an overview of what you should have learned from each chapter, answers to some of the textbook exercises and further exercises (with answers). Even more valuable are the Powerpoint slideshows and video lectures that Dougherty provides. You should check these out. I may incorporate parts of his slideshows into those I use in class.

Dougherty also provides downloadable datasets to be used in some of the textbook exercises. I'll talk more about this and the software you'll use (STATA) in class when appropriate.

Finally, this book is inexpensive relative to its competitors, which is always nice.

4.1 textbook website

The URL for the website with supporting material is http://econ.lse.ac.uk/courses/ec220. There you will find the following resources:

- downloadable slideshows
- study guide for the textbook
- datasets that will be used for some computer exercises

There are other materials available at Dougherty's website, for example the syllabus for his LSE course. These are obviously not relevant for 281. In general, refer to Blackboard for administrative material/announcements. You should treat Dougherty's site as an invaluable supplement.

4.2 other materials

I will post on Blackboard the slides I'll project in class. You will have the option of printing out copies and bringing them with you. **Note that reading the slides is not a perfect substitute for reading the book, which you must do.**

There will be three versions of the slides available: one identical to that I use in class; one in 'handout' form that omits the 'dynamic overlays' but otherwise resembles the version used in class; and one in 'article' form that's even more compact. You might also/instead want to bring the textbook to class, and leaf through that as I go through material. It's up to you. I will keep close to the texbook in terms of notation and labelling of equations.

5 Lecture topics

We are not going to get through the whole book. The provisional plan is that we cover, in order, all the chapters up to and including Chapter 9 (simultaneous equations). If it seems we are falling behind schedule, my preference is to drop material from the syllabus rather than speed up. I will alert you as soon as this looks likely.

We will spend 2-3 classes on each chapter. The outline below is meant to provide an explicit list of all the mathematical formulae and proofs that you are expected to know for the midterms and final. In general, the emphasis for these exams is not on proofs but on intuitive understanding of the material. I will be more explicit about what is required on the midterms and final closer to the time, and reserve the right to deviate from the guidelines below.

Review chapter: Random variables and sampling theory

Probability distribution of a random variable. Expected value of a random variable. Expected value of a function of a random variable. Population variance of a discrete random variable and alternative expression for it. Expected value rules. Independence of two random variables. Population covariance, covariance and variance rules, and correlation. Sampling and estimators. Unbiasedness. Efficiency. Loss functions and mean square error. Estimators of variance, covariance and correlation. Probability limits and their rules. Consistency. The central limit theorem.

Formulae and proofs for exams: This chapter is concerned with statistics, not econometrics, and should be mostly revision for you. You are not expected to be able to prove any of the results in the chapter. However, you are expected to know these results and to be able to use them.

Chapter 1: Simple regression analysis

Simple regression model. Derivation of linear regression coefficients. Interpretation of a regression equation. Goodness of fit.

Formulae and proofs for exams: You are expected to know the formulae for the regression coefficients and to understand in principle how they are derived, but you are not expected to be able to reproduce the actual mathematical derivations. You are expected to know the definition of R^2 and how it is related to the residual sum of squares. You are expected to know the relationship between R^2 and the correlation between the actual and fitted values of the dependent variable, but not to be able to prove it.

Chapter 2: Properties of the regression coefficients

Types of data and regression model. Assumptions for Model A. Regression coefficients as random variables. Unbiasedness of the regression coefficients. Precision of the regression coefficients. Gauss-Markov theorem. t test of a hypothesis relating to a regression coefficient. Type I error and Type II error. Confidence intervals. One-sided tests. F test of goodness of fit.

Formulae and proofs for exams: You are expected to know the regression model assumptions for Model A; you are expected to know, though not be able to prove, that, in the case of a simple regression model, an F test on the goodness of fit is equivalent to a two-sided t test on the slope coefficient. You are expected to know how to make a theoretical decomposition of an estimator and hence how to investigate whether or not it is biased. You are expected to know the expression for the variance of the slope coefficient in a simple regression model and how to estimate the variance of the disturbance term, given the residuals.

Chapter 3: Multiple regression analysis

Multiple regression with two explanatory variables. Graphical representation of a relationship in a multiple regression model. Properties of the multiple regression coefficients. Population variance of the regression coefficients. Decomposition of their standard errors. Multicollinearity. F tests in a multiple regression model.

Formulae and proofs for exams: You are expected to know how, in principle, the multiple regression coefficients are derived, but you do not have to remember the expressions, nor do you have to be able to derive them mathematically; you are expected to know the expressions for the population variance of a slope coefficient and its standard error in a model with two explanatory variables. You are expected to be able to perform F tests on the goodness of fit of the model as a whole and for the improvement in fit when a group of explanatory variables is added to the model.

Chapter 4: Transformation of variables

Linearity and nonlinearity. Elasticities and double-logarithmic models. Semilogarithmic models. The disturbance term in nonlinear models. Nonlinear regression. .

Formulae and proofs for exams: none.

Chapter 5: Dummy variables

Dummy variables. Dummy classification with more than two categories. The effects of changing the reference category. Multiple sets of dummy variables. Slope dummy variables.

Formulae and proofs for exams: none.

Chapter 6: Specification of regression variables: a preliminary skirmish

Omitted variable bias. Consequences of the inclusion of an irrelevant variable. Proxy variables. F test of a linear restriction. Reparameterization of a regression model. t test of a restriction. Tests of multiple restrictions. Tests of zero restrictions.

Formulae and proofs for exams: You are expected to be able to derive the expression for omitted variable bias when the true model has two explanatory variables and the fitted model omits one of them. You are expected to know how to perform an F test on the validity of a linear restriction, given appropriate data on the residual sum of squares. You are expected to understand the logic behind the t test of a linear restriction and to be able to reparameterize a regression specification to perform such a test in a simple context.

Chapter 7: Heteroscedasticity

Meaning of heteroscedasticity. Consequences of heteroscedasticity. Goldfeld-Quandt and White tests for heteroscedasticity. Elimination of heteroscedasticity using weighted or logarithmic regressions. Use of heteroscedasticity-consistent standard errors.

Formulae and proofs for exams: You are expected to know how to perform the Goldfeld-Quandt test for heteroscedasticity.

Chapter 8: Stochastic regressors and measurement errors

Stochastic regressors. Assumptions for models with stochastic regressors. Finite sample and asymptotic properties of the regression coefficients in models with stochastic regressors. Measurement error and its consequences. Friedman's Permanent Income Hypothesis. Instrumental variables.

Formulae and proofs for exams: You are expected to be able to demonstrate that, in a simple regression model, the OLS estimator of the slope coefficient is inconsistent when there is measurement error in the explanatory variable. You should know the expression for the bias. You should know the expression for an instrumental variable estimator of the slope coefficient in a simple regression model and be able to demonstrate that it yields consistent estimates, provided that certain assumptions are satisfied. You should also know the expression for the population variance of an instrumental variable estimator in a simple regression model.

Chapter 9: Simultaneous equations estimation

Definitions of endogenous variables, exogenous variables, structural equations and reduced form. Inconsistency of OLS. Use of instrumental variables. Exact identification, underidentification, and overidentification. Two-stage least squares.

Formulae and proofs for exams: You are expected to be able to know the expression for simultaneous equations bias of OLS in a simple regression equation.

6 Comments, policies

- The midterm and final dates/times are not negotiable. There will be no early finals to accommodate family holidays, unfortunately-timed job interviews, internships, etc.
- Similarly for assignments. They must be handed in on time. Failure to do so, especially (but not only) with no prior warning, will result in points deducted.
- If you think you were unfairly denied marks on an assignment or exam, do the following. First, recheck the answers (which might not be published, but if not will be covered in a section). If you still think it's unfair then take the matter up with your TA to see if s/he agrees. If you still think it's unfair then (and only then) bring it to me; I will remark the entire assignment/exam, and reserve the right to deduct points as well as add them.
- Will I be grading on a curve? Yes and no. There is always the chance that I make a mistake and ask a question that is too easy or too difficult. If it turns out that people do **much** better or worse than I expected, I may re-examine the difficulty of the assignment/exam and perhaps shift the grade distribution accordingly. To the extent that this happens, it will tend to be in your favour. In general, though, I do not have any preconceived idea of some particular grade distribution I want to see.
- Further to the grading: don't panic if you see you got a 70 on a test and think you failed. Check the letter-grade/number-grade correspondence, which may differ from that you are accustomed to.
- I expect that students will work together when solving the assignments. This is fine. However, I encourage you to at least write up the answers separately, as this is a good way to check you really understand the problem. Such understanding will be useful in the midterms and final.
- If you don't understand something as we're going through it, that's fine. Just be sure to see me or the TA very quickly to sort things out. The nature of the course means that falling behind is a very bad thing indeed.
- There will be no *ex post* reweighting of the relative contributions of assignments and exams to the final grade. Please don't ask, as I find saying no painful.
- Please ask questions in class if something is not clear. I budget plenty of time for questions.