Econ 401: Advanced Econometrics
Roger Klein; Fall 2008: MW 4:30-5:40, SCOTT 101

Econometrics is concerned with formulating, estimating, and testing economic relationships. In this course, we will consider estimators and tests for models that differ from the linear model considered in Econ322. As an example of a standard linear model, assume that the log of wages is specified to depend linearly on such explanatory variables as education, experience,.... In this case, the coefficient on education is termed the return to education and is constant. In particular, the return to education is assumed to be the same when you graduated high school as when you graduate Rutgers. Yet, it would seem that the return to education may not be the same for all education levels. In non-linear models, the return to education will not be constrained to be constant. Accordingly, such models may provide a better description of the behavior in which we are interested. We will consider such models in this course.

In other models of interest, the sample may be selected in some unusual manner that results in a problem termed sample selection. Testing and correcting for this problem will take us out of the linear model context. In still other examples, the dependent variable of interest may be only partially observed. For example, in a model that explains health status, health may only be observed as a categorical variable in that we observe whether an individual’s health is high, average, or low. Again, we will find that a linear model will not be adequate for such data. Departures from the regular linear model will also be required when we seek to jointly explain several dependent variables of interest. For example, in studying murder rates, we might assume that such rates depend on part on some measures of police presence. On the other hand, police may not be randomly assigned to areas and instead our measure of police presence may in turn depend on crime conditions in the area of interest.
OFFICE HOURS:

Office hours will be held at times to be announced, Rm. 311, NJ Hall, CAC. If you have questions and can not come at the announced times, email me we will arrange another time.

Office : Klein@economics.rutgers.edu
Home : Kleins@comcast.net

Since each topic covered depends on previously covered material, I strongly encourage you to ask questions as we go along. I would also encourage you to ask any questions that you have about weekly problem sets before they are due. To do well in this course, you must be willing to ask questions about anything that you do not understand, and in turn I will be very accessible.

TEXT:

There is no required text for this course. The material described below will be covered in handouts.

GRADING:

Weekly Problem Sets 20%
Interm Exam 10%
Midterm Exam 35%
Course Paper 35%

WEEKLY PROBLEM SETS:

Weekly problem sets will be given that will usually be computer intensive and will involve estimating and/or testing various models of interest. We will be using a software language, Gauss, and doing some programming in this language. However, no prior experience with this package is assumed. We will go over in class and cover in handouts all material that you will need to know. While a few students typically become interested in writing programs, programming will not be emphasized in this course. You will be given some exposure to programs so that you will have some understanding
as to how some of the calculations are actually made. In assigning grades, I will drop the weekly problem set with the lowest score.

For students with PC’s, a free copy of Gauss will be available. This free copy is a “light version” and has a limited memory. Moreover, certain useful programs (libraries) are not included in this student version of Gauss. Consequently it will still be necessary to do some work in a lab in the economics department.

You may work together on these weekly problem sets, but each of you should turn in your own write-up. Make sure you understand how to do these problems as approximately 80% of each exam will contain similar problems. You may and are strongly encouraged to ask questions about these weekly problem sets before you turn them in.

EXAMS:

Approximately 80% of the exams will consist of problems similar to those in the weekly problem sets. Accordingly, most of the questions should not be surprising if you have completed and understand the weekly problem sets. Approximately 20% of these problem sets may involve extensions not directly covered in previous problems. Past exams and sample problems will be distributed as a way of studying for the exams. We will then probably have review sessions prior to each exam to go over any questions on such review problems.

PAPER: TO BE DISCUSSED

ATTENDANCE AND MISSED ASSIGNMENTS:

Attendance will not be taken. However, most of the material that we cover will be explained in class and will not be available elsewhere (other than handouts). Consequently, regular attendance will be necessary to do well in this course. If you do miss a class, first get notes from someone else in class and then come to see me to fill in any gaps. As for missed assignments and/or exams, make-ups will not be given. Without a "satisfactory" reason for missing an exam or an assignment, these will count as 0 grades.
COURSE OUTLINE

The following outline gives the topics that we will cover and the approximate dates for each topic. More or less time may be required than as is indicated for several of the listed topics. Accordingly, you should view the dates listed below as tentative. There may also be changes in this outline if a substantial number of students are writing papers involving different areas of econometrics than those in the outline below. Some change in the order in which topics are covered may also be required if a substantial number of students are writing papers on topics currently scheduled towards the end of the course.

- Course Introduction; Review of Expectations (Population Means) and Sample Means: Sept. 3, 8, 10

In econometrics, we are typically interested in conditional expectations. Viewing an expectation as an average in the relevant population, we may be concerned with the average value for wages in the population of individuals with 12 years of education and 6 years of experience. Such a population average is termed a conditional expectation in that it is conditioned on those individuals with 12 years of education and 6 years of experience. A sample mean provides one estimator for such an expectation in that it is "likely" to be close to the corresponding population mean (expectation). Furthermore, in large samples, sample means have a distribution that is approximately normal. In part, this section will serve to review important concepts in statistics that we will need. It will also provide a foundation for many of the methods that we will examine in this course.
• **Simultaneous Equations and Endogenous Variable Bias:** Sept. 15, 17, 22, 24

In estimating a model in which several dependent variables are jointly determined (e.g. demand and supply), the OLS estimator may not be appropriate. Instead, it may be necessary to employ an alternative estimation method termed instrumental variables or two-stage least squares. As a somewhat related issue, we will also briefly discuss panel data issues with unobserved individual specific effects.

• **Nonlinear Models and Least-Squares:** Sept. 29, October 1

In some applications, it is necessary to allow for and/or test for the model being nonlinear. In this section, we will discuss a natural extension to OLS estimation in linear models.

**Interm Exam:** Monday, October 6

• **Nonlinear Models and Least-Squares Continued:** Oct. 8

• **Discrete Choice (binary response) Models:** Oct. 13, 15, 20, 22

In many applications (e.g. labor force participation), the dependent variable is limited to a few values. In the binary case, the dependent variable takes on one of two variables (e.g. one if the individual joins the labor force and zero otherwise). Here, we will describe and implement an alternative estimation method termed maximum likelihood. We will also compare this method with a variant of least-squares.
• Categorical Models: Oct. 27, 29, Nov. 3

In categorical models, the dependent variable takes on a small number of distinct values. For example, individuals may provide self-reported health data in which they state whether their health is excellent, average, or below-average. Even when the original data are not categorized, we will discuss circumstances (model specification tests) when it is desirable to group the data into categories.

• Censored Regression Models: Nov. 5, 10

As an example, suppose that we have demand data that looks continuous when it is positive. In conjunction with such continuous data, suppose that a sizeable fraction of the population has zero demand for the product. A censored regression (Tobit) model is often employed to analyze these two types of data.

Midterm Review : Wednesday, November 12
Midterm Exam : Monday, November 17

• Sample Selection Models: Nov. 19, 24, Dec. 1

Frequently we are interested in estimating an equation of interest that has a very simple structure if we had a random sample from the entire population. For example, the equation may be the same type of linear model previously studied. However, rather than having a random sample from the entire population, the sample may be selected so that we only have data on ”unusual individuals”. Employing the selected sample, in this section we discuss the nature of the sample selection problem and methods for dealing with it.
To illustrate this type of model, suppose that individuals must decide whether or not to participate in a job training program (the "treatment"—discrete choice component). Further suppose that the log-wage (another equation) depends on whether or not the individual participates in a job training program. This two equation model combines aspects of simultaneous equations and discrete choice equations.