Econ 401: Advanced Econometrics: Spring, 2000
Roger Klein; FH A5, MW 4:30-5:50

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 this 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.

Departures from the regular linear model will also be required when we seek to jointly explain several variables of interest. For example, in studying murder rates, we might assume that such rates depend in 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 areas of interest. As discussed in class, the OLS estimator in this context can be severely biased.

In other models of interest, the dependent variable of interest may only be partially observed. For example, in a model that explains health stratus, this variable may only be observed as a categorical variable in that we observe whether an individual's health is high, average, or low. We will require a very different estimation method to estimate this type of model. In still 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.

OFFICE HOURS:

Office hours will be held at times to be announced, Rm. 311, NJ Hall, CAC. After surveying times when you are free, I will try to set a time slot when most of you are available. However, If you have questions and can't come at the announced times, contact me and we will arrange another time to meet. My email address is:

rogerwklein@gmail.com

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 may have about weekly problem sets **before** they are due.

TEXT:

There is no required text for this course. The material described below will be covered in handouts and selected readings that will be distributed.

GOALS AND ASSESSMENT: The purpose of this course is to provide the foundations for doing applied work in economics and to do when the model is not necessarily linear. Grades will be determined as follows:

Weekly Problem Sets	25%
Interm Exam	15%
Midterm Exam	35%
Paper	25%

I will try to teach this course at what I judge to be approximately the median level of the class. If you are having trouble with the material, I strongly encourage you to ask questions in class and to come to office hours. If you would like to see more formal or complete arguments than I am giving in class,, I am glad to provide these outside of class (if I judge the material not appropriate for a general class lecture).

WEEKLY PROBLEM SETS:

Weekly problem sets will be given, which will require use of a computer. You will be using Gauss, a software package. Some programming in Gauss

will be required in these problem sets. 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. Gauss is available without charge on a variety of platforms and can be downloaded to your own computer. If you do not have your own computer, let me know and I will make arrangements to have an account set up for you on a computer in the graduate lab located in the library on the 3rd floor of NJ Hall. **You may work together on these weekly problem sets, but each of you should turn in your own write-up. You may and are encouraged to ask questions about these weekly problem sets before they are due.** In assigning grades, I will drop the weekly problem set with the lowest score.

EXAMS

Approximately 80% of each exam will consist of problems similar to those in the weekly problem sets or that we have emphasized in class. Accordingly, most of the questions should not be surprising if you have completed and understand the problem sets. Approximately 20% of this exam will involve extensions not directly covered previously and I reserve the right to be creative here.. One of the exams will be given very early in the semester so that you will know what my exams are like before a midterm which counts substantially more than the first exam. An early exam also makes it possible for me to catch and usually correct problems related to understanding the material.

PAPER:

There will be no final in this course, but there will be a final paper that is due at the end of the course. We will discuss in class what is required in the paper.

ATTENDANCE AND MISSED ASSIGNMENTS:

Attendance will not be taken. However, much of the material that we cover will be explained in class and will not be available elsewhere (other than handouts and assigned readings). Consequently, regular attendance will be necessary to do well in this course. If you do miss a class, first get notes from someone in the class and then come to see me to fill in any gaps. As for missed assignments and exams, make-ups will not be given. Except in special circumstances, these will count as 0 grades unless there is a very good and documented explanation.

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 is indicated for several of the listed topics. Accordingly, you should view the dates listed below as tentative.

- **Course Introduction: Review of Expectations (Population Means) and Sample Means. : Jan. 22, 27, 29**

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, a sample mean will (under somewhat general conditions) have a distribution that is approximately normal. We will employ (and adapt) these two results in a wide class of models below.

- **Endogenous Variable Bias: Feb. 3,5,10,12**

In estimating a model in which several dependent valuables are jointly determined, the OLS estimator may not be appropriate. Instead, it may be necessary to employ an alternative estimation methods such as instrumental variables , two-stage least-squares, or control methods. These methods can be equivalent in certain contexts, but are not in general the same.

Optional Review Session: Friday, Feb. 14
Interm Exam, Feb. 17

- **Nonlinear Least-Squares: Feb. 19, 24**

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

- **Discrete Choice (binary response) Models: Feb. 26, March 3 ,5**

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 describer and implement an alternative estimation method termed maximum likelihood. We will also compare this method with nonlinear least-squares to provide intuition for the fact that under appropriate conditions the maximum likelihood estimator has a minimal variance. .

- **Endogenous Treatment Models: March 10,12**

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 equation). Further assume that an individual's wage (another equation) depends on whether or not the individual participates in a job training program.

.

SPRING BREAK: March 15-23

- **Categorical Models March 24, 26, 31. April 2**

In categorical models, the dependent variable takes on a small number of distinct values. For example, we will consider a demand application in which it is natural to view actual demand as being unobserved. Instead, we only observe whether demand is "low", "medium", or "high"

Optional Midterm Review April 4
Midterm Exam April 7

- **Estimating Demand with Misreported Survey Data: April 9, 14, 16, 21 (Paper)**

Data are often collected in surveys where individuals report various quantities of interest. In such data sets, misreporting can be a serious issue. This issue, as will be discussed in class, can be particularly serious when survey data are employed to study demand for new products.

- **Sample Selection and Censored Models: April 23, 28, 30, May 5**

Frequently we are interested in estimating an equation of interest that would have 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 will discuss the nature of the sample selection problem and control methods for dealing with

In a related class of models, we may only observe the dependent variable when it is above some threshold. For example, with survey data, we may only obtain "demand" data for individuals that like the product (as only such individuals may complete survey questionnaires).