POLS W4291
Advanced Topics in Quantitative Research: Limited and Qualitative Dependent Variables

Fall 2007

Tuesdays and Thursdays, 9:10–10:25am

410 International Affairs Building

Prof. Gregory Wawro 212-854-8540

741 International Affairs Bldg. gjw10@columbia.edu

Office Hours: Wednesdays, 2–4pm and by appt.

Course Description

This course covers methods for models that have dependent variables that are not continuous. These models include dichotomous and polychotomous response models, models for censored and truncated data, sample selection models, duration models, and models for count data.

Prerequisites: Students wishing to take this course should have taken an introductory course in probability, a course on regression analysis, and have basic knowledge of calculus and matrix algebra.

Course Requirements:The grading for the course is based on a mid-term and final exam (30% and 40%, respectively), and homework assignments (30%).

Lecture Notes:
PDF versions of my lecture notes are available from the course web site (the URL is `http://www.columbia.edu/~gjw10/w4291.html`). Students should download and print up a copy of the notes for themselves so that they can follow along with the lectures. Students who do not have a copy of the notes will be at a severe disadvantage.

Texts: I have ordered the following books for this course through Book Culture (the book store formerly known as Labyrinth) on 112th St.

Cameron, A. Colin and Pravin K. Trivedi. 1998. *Regression Analysis of Count Data*, Cambridge University Press (also available in electronic format through CLIO).

Eliason, Scott R. 1993. *Maximum Likelihood Estimation: Logic and Practice*, Sage Publications.

Fox, John. 2002. *An R and S-Plus Companion to Applied Regression*, Sage Publications.

Greene, William H. 2003. *Econometric Analysis*, Fifth Edition, Prentice-Hall.

Lancaster, Tony. 1990. *The Econometric Analysis of Transition Data*, Cambridge University Press.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*, Sage Publications.

Maddala, G.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

All of these books are useful and are strongly recommended. The readings assignments will focus on the Long and Maddala books. Although the material assigned from these two texts overlaps, students should read assignments from both in order to get the most out of the material. The Maddala reading is more difficult and somewhat dated, but will give students a more advanced treatment of the material.

All reading assignments from books are on reserve at Lehman. All reading assignments from journals are available through JSTOR (http://www.jstor.org/).

Computing/Software: We will use `R` and `Stata` for computational exercises in this class, with more emphasis on `R`. The latter is free and available for download from `http://www.r-project.org/`. Students are welcome to use other software packages that they are familiar with, but we will provide support only for `R` and students will be required to program in `R` for problem sets. Be prepared to devote a good deal of time to programming, which is essential to becoming proficient in the methods discussed in this course.

Course Outline

The dates for the sections are subject to change.

I Introduction

1 Course Motivation — Sept. 4, 6

- Long, Chs. 1 and 2.
- Maddala, Ch. 1.

2 Review of Fundamentals of Probability — Sept. 11, 13

- Greene, *Econometric Analysis*, 5th Ed., Appendix B.

II Maximum Likelihood Estimation — Sept. 18, 20

- Greene, *Econometric Analysis*, 5th Ed., Ch. 17
- Eliason, *Maximum Likelihood Estimation*.

III Models derived from a common structure: the normal regression model

5 Introduction — Sept. 25

- Long, Ch. 2.

6 Probit — Sept. 25

- Long, Ch. 3.
- Maddala, Ch. 2.1–2.5.

7 Ordered Probit — Sept. 27

- Long, Ch. 5.

8 Censored Regression — Sept. 27

- Long, Ch. 7.
- Maddala, 6.1–6.6.

9 Truncated Regression — Oct. 2

- Maddala, Ch. 6.9–6.10.

10 Sample Selection Models — Oct. 4, 9

- Maddala, Ch. 9.

Take-home midterm about here

IV Probabilistic Choice Models

11 Introduction — Oct. 11

- Review Long, Ch. 3 for binary logit.

12 The Multinomial Logit Model — Oct. 11

- Long, Ch. 6.
- Maddala, Ch. 2.10–2.12, Ch. 3.

13 The Conditional Logit Model — Oct. 16

- Long, Ch. 6.
- Maddala, Ch. 3.

14 The Nested Logit Model — Oct. 18

- Greene, pp. 865–870.
- Maddala, pp. 67–70.

15 The Multinomial Probit Model — Oct. 23

- Greene, pp. 870–875.
- Alvarez, R. Michael and Jonathan Nagler. 1995 "Economics, Issues, and the Perot Candidacy: Voter Choice in the 1992 Election," *American Journal of Political Science* 39:714–744.

V Duration Models — Oct. 25, 30, Nov. 1, 8

- Box-Steffensmeier, Janet M. and Bradford S. Jones. 1997. "Time is of the Essence: Event History Models in Political Science." *American Journal of Political Science* 41: 1414–1461.
- Keifer, "Economic duration data and hazard functions," *Journal of Economic Literature* (1988) 24: 646–679.
- Lancaster, *The Econometric Analysis of Transition Data*, Chs. 1–3.

VI Event Count Models — Nov. 13, 15

- Long, Ch. 8.

VII Models for Repeated Observations — Nov. 20, 27, 29

- Hsiao, *Analysis of Panel Data*, 2d ed., Ch. 1 and pp. 188–202.
- Baltagi, *Econometric Analysis of Panel Data*, pp. 178–186.
- Beck, Nathaniel, and Jonathan N. Katz. 1995. "What To Do (and Not To Do) with Time-Series Cross-Section Data in Comparative Politics." *American Political Science Review* 89: 634–647.
- Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-Series Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42: 1260–1288.
- Zorn, Christopher J. W. 2001. "Generalized Estimating Equations Models for Correlated Data: A Review With Applications." *American Journal of Political Science* 45: 470–90 (see Errata on p. 748)

Catch-up and Review — Dec. 4, 6