Economics 371

Problem Set #9

(1) Stock and Watson, 9.1. *Note:* You will need to look up values of the Normal cumulative density function on pages 642-643 of your book to carry out this exercise.

(2) This problem makes use of the data PNTSPRD.dta, which is available on the course website.

(2a) Suppose that a *linear probability model* is used to estimate the model:

$$favwin_i = \beta_0 + \beta_1 spread_i + u_i,$$

where $favwin$ is a binary variable, equal to 1 if the team that was favored to win by the Las Vegas point spread does, in fact, win. *Spread* equals the number of points by which the favored team is expected to win. Run this regression in STATA.

(2b) Explain why, if *spread* incorporates all relevant information, we would expect that $\beta_0 = .5$. Test this hypothesis. Can you rationalize the result of your hypothesis test?

(2c) What is the predicted probability that the favored team wins if $Spread = 10$?

(2d) Instead of using a linear probability model to fit this relationship, use a probit model where

$$\Pr(Favwin_i = 1|Spread) = \Phi(\beta_0 + \beta_1 Spread_i)$$

and Φ denotes the standard normal cdf. To implement the probit model using STATA, simply type

probit favwin spread

(2e) Show why, in the probit model, if *spread* incorporates all relevant information, we would expect that $\beta_0 = 0$. Test this hypothesis. (Note: Standard Errors, test

statistics and p-values are reported along with the ususal STATA output. You can use these to conduct the test.)

(2f) Using your coefficient estimates, and the normal cdf tables of your book, what is the predicted probability that the favored team wins if $Spread = 10$? Compare this with (2c).

(2g) Calcualte *marginal effects* from the probit model. To produce the marginal effect output instead of the coefficients themselves, simply type

dprobit favwin spread

Interpret your marginal effect estimate.

(3) The eddata.dta data set, contained on the course website, contains three variables: IQ (IQ score ranging from 50 to 145), *meduc*, the number of years of education completed by the respondent's mother, and *edcat*, a categorical education variable taking on three values: 1 if the individual is a high school dropout, 2 if the individual is a high school graduate and 3 if the individual has at least some college education.

(a) Using this data set, fit an ordered probit model containing only an intercept. To do this, type

oprobit edcat

The regression output will include two cutpoints, denoted "cut1" and "cut2." Note that these are slightly different from those discussed in class. Based on our class lectures, we obtained

$$\begin{aligned} \Pr(edcat = 1) &= \Phi(-\beta_0) \\ \Pr(edcat = 2) &= \Phi(\alpha - \beta_0) - \Phi(-\beta_0) \\ \Pr(edcat = 2) &= 1 - \Phi(\alpha - \beta_0). \end{aligned}$$

As the cupoints cut1 and cut2, STATA reports

$$\begin{aligned} cut1 &= -\beta_0 \\ cut2 &= \alpha - \beta_0 \end{aligned}$$

Use $cut1$ and $cut2$ together with the formulas above to calculate the probabilities associated with each of the three education categories. For the sake of comparison, type

tabulate edcat

in STATA. Compare your ordered probit probability estimates to what you see in the output of the tabulate command.

(b) Now, include IQ and *meduc* in the ordered probit model by typing

oprobit edcat IQ meduc

This ordered probit will also provide slope coefficients on the two explanatory variables.Are the slope coefficients on IQ and *meduc* statistically different from zero at the 5 percent level?

(c) What is the probability that someone whose mother had 10 years of schooling, and with an IQ score of 75 will be a high school dropout? A high school graduate? Get some amount of college education? (You will need to use the normal cdf tables of your book to perform these calculations).

(d) What is the probability that someone whose mother had a B.A. (i.e., meduc=16) with an IQ perctile of 120 will be a high school dropout? A high school graduate? Get some amount of college education?

(e) Calcualate a specific marginal effect: the change in the probability that someone will get at least some college education in response to a one-year increase in *meduc*. When calculating this marginal effect, set $IQ = 100$ and $meduc = 12$. In our lecture notes, we showed that this effect could be calculated as

$$\frac{\partial \Pr(y = 3|\widehat{IQ = 100}, meduc = 12)}{\partial meduc} = \phi\left(\widehat{\alpha} - [\widehat{\beta}_0 + \widehat{\beta}_{IQ}100 + \widehat{\beta}_{meduc}12]\right)\widehat{\beta}_{meduc}.$$

To carry out this calculation, compute the quantity $\widehat{\alpha} - [\widehat{\beta}_0 + \widehat{\beta}_{IQ}100 + \widehat{\beta}_{meduc}12]$ by hand. Then, to find the ordinate of the normal density at a particular value c, i.e., to calculate $\phi(c)$, note that you can do this in stata by typing

generate y = normalden(c)

The desired density ordinate is then stored in y. You can see the value of y simply by typing

mean(y)

for example. Use this process and set of commands in stata to calculate the marginal effect above. Interpret this result.