

Economics 571
Problem Set #10

(1) Wooldridge 17.1. This question refers to the “fraction correctly predicted” for various groups. By convention, if $\Pr(\widehat{y_i = 1|x} \equiv \hat{p}_i > .5)$, we would predict $y_i = 1$. So, the fraction correctly predicted for the $y = 1$ subsample would be expressed as

$$\hat{q}_1 = [1/n_1] \sum_{i:y_i=1} I(\hat{p}_i > .5),$$

for example, where n_1 is the number of observations in the sample for which $y_i = 1$ and $I(\cdot)$ denotes the standard indicator function. (Other fractions correctly predicted are defined analogously). This proof is quite simple, so don't be intimidated by the above notation.

(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 usual 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) Calculate *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.