Problem Set 4

EE426 Econometrics 2

Due May 1, 2015 (by 4.00pm)

Please report the regression results if needed (no need to write estimated equation) and print STATA .do file attached at the end of your answer.

1. The file JTRAIN2.dta contains data on a job training experiment for a group of men during 1976-1977. The idea is to test whether participation in the job training program had an effect on unemployment probabilities and earnings in 1978.

(1.1) The variable *train* is the job training indicator. Run a linear regression of *train* on *unem74, unem75, age, educ, black, hisp*, and *married*. Are these variables jointly significant at the 5% level? Also, calculate the percent of correctly predicted.

(1.2) Estimate a probit version of the linear model in (1.1). From the reported likelihood ratio test for joint significance of all variables, what do you conclude?

(1.3) Based on your answers to (1.1) and (1.2), does it appear that participation in job training can be treated as exogenous for explaining 1978 unemployment status? Explain.

(1.4) Run a simple regression of *unem78* on *train* and other variables from (1.1) as additional controls. What is the estimated effect of participating in the job training program on the probability of being unemployed in 1978? Is it statistically significant? Plot the graph between predicted probability of being unemployed in 1978 and *train*. Do you see any predicted values outside [0,1]?

(1.5) Run a probit of (1.4). Does it make sense to compare the probit coefficient on *train* with the coefficient obtained from the linear model in (1.4)? Estimate the average marginal effect of *train* on the 1978 unemployed probability. How do you interpret this result? How does this average marginal effect compare with the OLS estimate from (1.4)? Are they comparable now?

2. Use the data in FRINGE.dta for this exercise

(2.1) For what percentage of the workers in the sample is *pension* equal to zero? What is the range of *pension* for workers with nonzero pension benefits? Why is a Tobit model appropriate for modeling *pension*?

(2.2) Estimate a Tobit model explaining *pension* in terms of *exper, age, tenure, educ, depends, married, white*, and *male*. Do whites and males have statistically significant higher expected pension benefits?

(2.3) Use the results form (2.2) to estimate the difference in expected pension benefits for a white male and a nonwhite female, both of whom are 35 years old, are single with no dependents, have 16 years of education, and have 10 years of experience.

(2.4) Add *union* to the Tobit model, comment on its significance and interpret its meaning.

3. Please answer the following questions:

(3.1) From Tobit model, we know that $E(y|x) = P(y > 0|x)E[y|y > 0, x]$, and hence $\frac{\partial E(y|x)}{\partial x_j} = P(y > 0|x)\frac{\partial E[y|y>0,x]}{\partial x_j} + \frac{\partial P(y>0|x)}{\partial x_j}E[y|y > 0, x]$. Please show that $\frac{\partial E(y|x)}{\partial x_j} = \beta_j \Phi\left(\frac{x\beta}{\sigma}\right)$.

(3.2) For AR(1), we have $y_t = \rho_1 y_{t-1} + u_t$, and u_t is uncorrelated with y_{t-1}. Please show that $Cov(y_t, y_{t+s}) = \rho_1^s \sigma_y^2$ and $\sigma_y^2 = \frac{\sigma_u^2}{1-\rho_1^2}$.