

Economics 381, Fall 2016

Problem Set 4 – Limited Dependent Variables

Due Date: October 25

Instructions: Complete the problem set answers in the spaces given. Keep track of your Stata commands, print them out, and **staple** your Stata commands to this sheet.

Name (printed):

- Q1. Suppose you have a panel dataset named `unemp.dta` that contains the unemployment rates in US states over different years open in Stata. You also have a second dataset, named `interest.dta`, containing the interest rate on US Treasury bonds over the same years. The interest rate on federal bonds is, of course, common to all states. What command should you use to merge these two datasets together in Stata? Be precise!
- Q2. Download and use the `lbw.dta` file available on Canvas. Use the command `desc` to look through the variables and their descriptions. Which variables are of storage type `int`, i.e. integers?
- Q3. Perform an OLS regression of low birthweight on age, mother's pre-pregnancy weight, and whether the mother smoked during pregnancy. What is the coefficient on the `smoke` variable?
- Q4. In terms of probabilities, interpret the marginal effect of the `smoke` variable.
- Q5. Perform a logistic regression of low birthweight on age, mother's pre-pregnancy weight, and whether the mother smoked during pregnancy. What is the z -statistic on the `smoke` variable?

- Q6. Suppose you to wanted to do this analysis again, but instead of assuming the predicted values are restricted to a logististic distribution function, you assumed they were restricted by a standard Normal distribution function. What command would you use to implement this type of maximum likelihood regression?
- Q7. Calculate the **average marginal effects** from this previous regression. What is the coefficient on the marginal effect of smoking?
- Q8. In one sentence, interpret this marginal effect.
- Q9. Now you want to calculate the **marginal effects at the mean**. What command would you use to do this?