

PS206 Homework #3 (Due Tuesday, May 4th)

Listed under Homework #3 on our class page is a tab-delimited text dataset called “ps206hw3.txt”. This data is from a survey conducted just before the 1992 U.S. presidential election, and contains 678 observations on five variables – vote (0 for Bush, 1 for Clinton), education (a 7 point scale with higher numbers meaning more education), age (in years), gender (0 for men, 1 for women), opinions on limiting imports (0 for opposes limits, 1 for favors limits), and party identification (0 for Democrat, 1 for independent, 2 for Republican). Download the data and do the following five problems:

(1) Estimate a probit model estimate a probit model with vote as the dependent variable and all other variables as independent variables. Now define a “hypothetical individual,” and calculate the probability that this hypothetical individual would vote for Clinton. You may define your hypothetical individual any way you like.

Note that the **Stata** command to calculate the CDF of a normal and save it as a variable called `p1` is `generate p1 = normprob(X)`, where `X` is the argument of the function (in this case, the equation with the estimated coefficients and hypothetical values of the independent variables). To call up a coefficient estimated in the last model we can use the “`_b`” command. For example, to get the coefficient on the variable “education” in the last model estimated type `_b[education]`. Thus, in **Stata**, to calculate the value of a 3 on the education scale times the coefficient estimated on education you would type `_b[education]*3`. This should give you enough information to calculate probabilities for hypothetical cases in **Stata**.

In **R** you can create a vector of values for the independent variables for the hypothetical individual, and multiply them by the estimated coefficients from the model. You can do this by modifying the **R** code for this homework.

Describe your hypothetical individual and state the probability you calculated.

(2) Now using the same probit model and technique as in problem 1, calculate the change in the probability of voting for Clinton as an individual moves from the minimum to the maximum education level, holding all else constant for our hypothetical individual. Describe what you find in this case.

(3) Repeat problems 1 and 2, except this time use a logit model rather than a probit model. Note that in **Stata** you can calculate the CDF of a logistic and save it as a variable called `p1` by typing `generate p1 = exp(X)/(1 + exp(X))`, where `X` is the argument of the function (note you need all those parentheses to calculate the formula correctly). To do this in **R** you can use the code from this homework to modify the logit code from Homework #2.

Describe what you find in this case.

(4) Now repeat the calculation in problem 2, but this time obtain standard errors on the probabilities you calculate. To do this in **Stata** you should use *Clarify* – type `net search clarify`, and then click on the links that come up to install the package. Use the help file if you don't remember how to use *Clarify*. To do this in **R**, modify the code for this homework.

Describe what you find in this case.

(5) Repeat problem 4, but this time use a logit model, a different hypothetical individual, and examine the effect of a different variable in the model. Describe what you find in this case.