

PS206 Homework #1 (Due Tuesday, April 14th)

Listed under Homework #1 on our class page is a tab-delimited text dataset called “ps206hw1.txt”. This is (hypothetical) survey data on 500 individuals, measuring their years of formal education, residence in a “motor-voter” state (1 if living in a state that allows people to register to vote while renewing their driver’s license, 0 if not), and voter registration status (1 if registered, 0 if not).

Load this data into the statistical package of your choice (Stata or R) and do the following two problems. To load this data in Stata you can type:

```
insheet using http://www.polsci.ucsb.edu/faculty/glasgow/ps206hw1.txt
```

(R users already know what to do).

(1) Begin by estimating a logit model to estimate the effects of education and living in a “motor-voter” state on the probability of registering to vote. To do this in Stata you would type `logit registered education motervoter`. To estimate a logit model in R with dependent variable “y” and independent variable “x” you would type:

```
> glm(y ~ x, family = binomial(link=logit), data = data)
```

This code assumes that you named your data set “data”; it should look very similar to what you’re used to doing in order to run a regular linear regression. You can save the output from this model as an object as usual (e.g., `model.1 <- glm(...)`), which you can then feed to the `summary` command.

What does this model tell you about the effect of education and residence in a “motor-voter” state on the probability that someone is registered to vote? Explain in as much detail as possible.

(2) Now estimate a probit model with the same dependent and independent variables. To do this, just replace ‘logit’ in the commands above with ‘probit’. Again, explain the model results in as much detail as possible. How does this probit model compare to the logit model you just estimated? What accounts for the differences you see?