You will have unlimited time to work on this final. This final is worth 100 points. I will deduct 10 points per day the final is late starting on Tuesday, June 17th, unless prior arrangements are made.

The final will be open book and open notes, but you cannot work with anyone else. You may ask me clarification questions (such as “If I used model X, would you expect me to do Y?”), but I cannot give you any substantial help (such as telling you which model to use).

For each problem, make the appropriate coding decisions, and apply and interpret the correct statistical model. I’m looking for a complete analysis here. Estimate your model, then show how changes in each independent variable affect the dependent variable. Write your results up nicely, as if you were going to submit them to a journal. Put your results into tables, rather than just pasting in your computer output. Write things up carefully, making sure to justify your choice of models. It might help to look at the methods sections of a couple of journal articles to get a feel for what these sections should look like. You should write about 3-4 pages per problem.

There is a tab-delimited text dataset on my website called ps206final.txt. The codebook for this data is on the following page.

For the final you must estimate three different models with this data.

One of your models must be an event count model.

Then pick two models from the following list. Do not estimate more than one of each type (so if you estimate a logit, you can’t also estimate a probit, etc.):

- Logit or probit
- Ordered logit or ordered probit
- Multinomial logit or conditional logit

Additionally, you must do at least one complex hypothesis test for one of these models.

For each model you are free to select whichever dependent and independent variables you like. Be sure to use at least three independent variables in each model. Interpret each model completely (predicted probabilities for a hypothetical individual, examining each independent variable). For each model give me a 3-4 page write-up and your R script file.
Codebook for PS 206 Final

These data come from the 2011 Canadian Election Survey.

- **id**: Respondent ID number.
- **vote**: Vote choice among the four major parties in the election. 1 = Conservatives, 2 = Liberals, 3 = NDP, 5 = Green.
- **feel_Cons, feel_Lib, feel_NDP, feel_Green**: 101 point thermometer ratings for the four parties (0 = greatly dislike, 50 = neutral, 100 = greatly like).
- **natecon**: National economy in the last year (1 = better, 2 = stayed the same, 3 = worse).
- **persecon**: Personal finances in the last year (1 = better, 2 = stayed the same, 3 = worse).
- **education**: An 11 point scale for the respondent’s education (higher = more education).
- **daysnews**: Number of days in the week the respondent spent reading news on the internet.
- **ownhome**: 0 = renter, 1 = home owner.
- **health_spend, ed_spend, unemp_spend, def_spend, old_spend, bus_spend**: A five point scale measuring the respondent’s opinion on how much the government should spend on health, education, unemployment benefits, defense, old age pensions, and supporting business, respectively. (1 = much more, 2 = more, 3 = same, 4 = less, 5 = much less).
- **LR_self, LR_Cons, LR_Lib, LR_NDP, LR_Green**: 11 point scale for respondent’s self-rated ideology and ideological rating of each party. (0 = far left, 5 = moderate, 10 = far right).
- **yearborn**: The year in which the respondent was born.
- **gender**: 0 = male, 1 = female.