

PS206 Final (Due Friday, June 12th)

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 Monday, June 15th, unless prior arrangements are made. I don't think the computer labs are open the weekend after finals week, so plan your computer work and writeup times accordingly!

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 – think carefully about possible problems with your data and take appropriate action. 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 referring to 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 (either ones I've handed out or others – I have a few papers linked on my website that might be helpful) 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 PS206Fin, with an accompanying codebook. Download them and do the following three problems. For each problem I want to see your 3-4 page write-up and your Stata or R log file.

Problem 1: Estimate the effect of age, union membership, and two other independent variables of your choice (four variables total) on opinions on personal finances. Interpret this model completely. Then perform the appropriate test to determine if the two independent variables you decided to include in the model are jointly an important determinant of opinions on personal finances. Interpret the results of this test.

Problem 2: Estimate a model determining what effect opinions on the general economy, employment, personal finances, and distances on the 5 issues in the dataset have on vote choice. For each issue there will be 4 distance variables, giving the distance between each voter and each of the 4 parties in the election on that issue – treat each set of issue distances as a single variable that varies across both individuals and parties. Interpret this model completely.

Problem 3: Estimate the effect of opinions on the general economy, employment, personal finances, and distances on the 5 issues in the dataset have on vote choice between the incumbent versus non-incumbent parties. For each issue, you should calculate the mean distance for incumbent parties and the mean distance for non-incumbent parties for each individual, and include this information as a single variable that varies across both individuals and the two sets of parties – in other words, I am looking for one coefficient for each issue. (HINT: Think of this as a special case of Problem 2 with a different number of categories in the dependent variable). Interpret this model completely.