

PS207 Final (Due Monday, December 15th)

You will have unlimited time to work on this final. Open book, open notes, but don't work with anyone else. This final is worth 100 points.

For each problem, apply and interpret the correct statistical model. I'm looking for a complete analysis here. 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 the Stata log file. Write things up carefully and describe the effect of all independent variables. It might help to look at the methods sections of a couple of journal articles (either ones I've handed out or others) to get a feel for what these sections should look like. You should write about 3-4 pages per problem. For each problem, write a separate do file, and save a separate log file. Attach them to your write-ups.

Problem 1: This will be a problem to examine models for time-series data. On my website is a Stata dataset called `Congapp.dta`. This is the replication dataset for the Durr, Gilmore, and Wolbrecht reading from last week. The meaning of the variable names in this dataset should be pretty obvious once you look at the Durr et al. paper. Download the data and do the following problems. One easy way to declare this data to be time-series data is to (1) type `sort year quarter`, (2) type `generate time = _n`, and (3) type `tsset time`.

- Examine the autocorrelation and partial autocorrelation functions for `capp` (Congressional approval) using the `ac` and `pac` commands. What does this reveal about this variable?
- Replicate the Model in Table 2 of the Durr et al. paper **except** leave out the lagged dependent variable. Describe your results.
- Perform a Durbin-Watson test by typing `estat dwatson`. What does this indicate?
- Run the same model as before, but correct for autocorrelation with the `prais` command. Describe the results. What does the new Durbin-Watson statistic indicate?
- Now replicate the same model, but this time include the lagged dependent variable. To get lagged Congressional approval use the lag operation in Stata, and type `generate lagcap = L.capp`. To replicate their test for autocorrelation use the `bgodfrey` command, and to replicate their test for moving average effects use the `archlm` command (see the help file for details).
- To make one point more clear, perform a Durbin-Watson test again. What did you find this time?
- One point of explanation here. Sometimes we see autocorrelation in our error terms because we really should have including a lagged dependent variable – the omitted variable is captured in our error term, resulting in autocorrelation. In this case including a lagged dependent variable will solve the autocorrelation problem. However, if there is still autocorrelation left in the error term after including the lagged dependent variable, this will result in bias. Not a question – just a FYI.

Problems 2 and 3: You can use any two of the other models we have discussed in class for the remaining two problems, as long as you don't simply replicate what we've already done in a homework. I have included three datasets in Stata format for you to use – replication data for my 2005 *PRQ* article, replication data for my coauthored 2006 *AJPS* article, and some data from the 2004 American National Election Study. You may also use any other data that you like.