Submit a single electronic file containing the answers to all questions through canvas. As usual, include your do and log files as appendices.

Question 1. Information Criteria and Cross Validation

a) Run these commands in Stata to simulate data on an AR(3) process.

```
clear
set obs 30
gen t=[_n]
tsset t
gen r=rnormal()
gen y=r if t<4
replace y=0.5+0.5*l.y-0.1*l2.y+0.25*l3.y+r if t>=4
drop r
```

You will wind up with a 30 observation time series for y. There are seven possible autoregressive models for y of order 1, 2, or 3. Obtain the AIC, BIC, 10-Fold CV RMSE, and LOOCV RMSE. Make a table with these four measures for each model.

b) Repeat a, but with the number of observations set to 300 rather than 30. Add the results to your table.

c) Compare model selection using the four criteria. Did they agree in both cases (n=30 and n=300)? For each criterion, are the differences between the "best" model and the next best more or less pronounced with more data?

Note about this exercise: Ideally, we would want to simulate this many times. But I don't want you to spend a lot of time figuring out how to code such a demonstration. The point is just to cement in your mind the relationships between the various model selection criteria and their relationship to sample size.

Question 2. Use the data you used for problem set 1 and 2, prepared for analysis in the same way. Make a table with the model selection measures for each model below. Based on the information in that table, which model do you think is best? Explain why.

- reg d.lnflnonfarm l(1/12)d.lnflnonfarm l(0/12)d.lnfllf l(0/12)d.lnusepr l(0/12)d.lnflbp i.month date
- reg d.lnflnonfarm l(1/12)d.lnflnonfarm l(0/2)d.lnfllf l(0/2)d.lnusepr l(0/2)d.lnflbp i.month date
- reg d.lnflnonfarm l(1/12)d.lnflnonfarm l(0/2,12)d.lnfllf l(0/2,12)d.lnflbp i.month date
- reg d.lnflnonfarm l(1/12,24)d.lnflnonfarm l(1/2,12,24)d.lnfllf l(1/2,12,24)d.lnusepr i.month