Problem set 7
607: Applied Macroeconometrics
Fall 2016
Jon Faust

The following is due at the beginning of next class. You can turn in any paper in my mailbox or in class; email me and requested computer work. You may work in groups; hand in a single submission for the group. The submission should list those who contributed.

Note 1: This is a new (well, a greatly changed) problem set. Please help me proofread it and let me know if there are confusing or ridiculous bits.

Note 2: This problem raises breaks tests, which we have not talked about and is related to the issue of testing for breaks in correlation, which we started on last time. Two papers you will want to look are these:

Bruce Hansen's paper introduces the concepts in this area.

My paper with Brian Doyle on breaks in the correlation of GDP growth across countries is an application of what I think of as fairly careful econometrics. Well-motivated problem, careful treatment of several subtle econometric problems, bootstrap inference, and a Monte Carlo to check whether the inference approach is plausibly reliable.

Bruce E. Hansen; The New Econometrics of Structural Change: Dating Breaks in U.S. Labour Productivity; JOURNAL OF ECONOMIC PERSPECTIVES; VOL. 15, NO. 4, FALL 2001; (pp. 117-128)

https://www.aeaweb.org/articles?id=10.1257/jep.15.4.117

Doyle, B., and Faust J. 2005. Breaks in the variability and co-movement of G-7 economic growth; Review of Economics and Statistics, 7(4), Nov. 721-740.

https://www.jstor.org/stable/40042889

1. Terms

   (a) Chow test for structural break

   (b) sup $-F$-style Chow test for structural break at unknown date

2. Structural breaks. You have a regression model:

$$y_t = x_t'\beta + \varepsilon_t$$

You want to test for a structural break in β occurring between observations t_b and $t_b + 1$.

   (a) Calling the values in the two subsamples β_1 and β_2, write a single regression such that the test of $H_0 : \beta_1 = \beta_2$ is a simple coefficient restriction on the coefficients of a single regression.

   (b) What is the natrual Wald test of H_0? How is it distirbuted? What is its relation to the F-test described in your $\sup -F$ answer.

   (c) You don't know the break date, t_b. You look at the data and see a likely break date and run a nominal 5 percent Chow test for a break at that date. Why will the test of this procedure tend do have size greater than 5 percent? (Note: View the procedure as: look at the data, pick what looks to be a breakdate, then run the test using that break date.)

   (d) How does the procedure in the $\sup -F$ test compare to the procedure in the previous part?

   (e) Breaks in a slightly different environoment. Now the model is:

   $$y_t = x_t'\beta + z_t'\gamma + \varepsilon_t$$

   The maintained assumption is that there is definitely one and only one break. And in particular, at least one element of β changes at that date. The parameter γ may or may not have a break.

   You use some sensible approach for picking a likely break date. For example, you run the $\sup -F$ test allowing for breaks in any of the βs or γ and take the date of the largest test statistic as t_b. Then you use that breakdate for the date in a standard nominal 5 percent Chow test for constancy of γ only (in the notation of the previous problem: $H_0 : \gamma_1 = \gamma_2$).

   Make an argument that this test of the γ will have proper size asymptotically.

3. Classroom presentation. This continues from our sample correlation Monte Carlo in ps6 (presentation should cover both).

Add the three standard bootstrap confidence intervals to the Monte Carlo from last time.

The three confidence interval procedures are the percentile, other percentile, and percentile-t. The bootstrap DGP is the XY-resampling scheme (resample rows of the data matrix with replacement).

Note. Creating one draw of N observtions from a $(T \times 2)$ data matrix can be done very easily:

```
T=100;
N=100;
XY = randn(T,2);     % fake data for illustration

newXY = XY[ ceil(T*rand(N,1), :];  % the new sample.
```

This relies on the fact that if A is a matrix in Matlab, $A[[25]', :]$ is a matrix composed of the second and fifth row of A.

For various sample sizes and values of the population correlation, and for each of the 3 confidence interval procedures, report the Monte Carlo estimate of the coverage, and the share of draws in which the population value lies to the left and to the right of the confidence interval. These three numbers obviously add to 1.