Structural Breaks

Time Series Econometrics

F-test review

- Suppose we want to run the regression

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \epsilon_t$$

- and test the joint hypothesis
  - $H_0 : \beta_2 = 4,\ \beta_3 = -2$
  - $H_a : \beta_2 \neq 4,\ \beta_3 \neq -2$
- One idea: variance-weighted squared distance of $(\widehat{\beta}_2, \widehat{\beta}_3)$ from (4,-2)
- Another idea: compare the variance of ϵ_t when $(\widehat{\beta}_2, \widehat{\beta}_3)$ are forced to be (4,-2) vs. when they are freely estimated.
- These turn out to be the same thing: an F-test.

Structural stability

- What if a parameter (coefficient, intercept, or variance) changes at some point (or points) in time?
- If we know (or have a hypothesis about) the break date(s), we can use an F-test or sometimes called a Chow test in this context.
- If we don't know the break date(s), we may need to loop over break dates.

Structural stability: Chow test

1. Estimate model over entire sample, save sum of squared residuals SSR_0
2. Define a dummy variable equal to 1 after the suspected break date t_1 and 0 before
   - $\delta_{[t>t_1]}$ and $\delta_{[t\leq t_1]} = 1 - \delta_{[t>t_1]}$.
3. Estimate model with different coefficients before and after suspected break t_1:
   - $y_{1t} = \beta_1' \mathbf{x_t} \cdot \delta_{[t\leq t_1]} + \beta_2' \mathbf{x_t} \cdot \delta_{[t>t_1]} + \epsilon_t$
   - save SSR_1
4. Calculate $F(t_1) = \frac{(T-2k)(SSR_0 - SSR_1)}{k \cdot SSR_1}$, compare to $F(k, T-2k)$ critical value, null is that $\beta_1 = \beta_2$ (stability)

Structural stability: Partial break

If your hypothesis is that only a subset of coefficients changed (e.g., some variables **z** have constant coefficients):

$$y_{1t} = \alpha' \mathbf{z_t} + \beta_1' \mathbf{x_t} \cdot \delta_{[t \leq t_1]} + \beta_2' \mathbf{x_t} \cdot \delta_{[t > t_1]} + \epsilon_t$$

- still an F-test for the equality of $\beta_1 = \beta_2$

Structural stability: Unknown break date

Andrews 1993 test:

- Do Chow test for all t_1 between first 15% and last 15% of the sample.
- Compute $k \cdot F(t_1)$ for each
- Compare **largest** value of $k \cdot F(t_1)$ to critical value in Table 1 of Andrews (1993)
  - (nonstandard "Andrews Distribution" when comparing multiple dates - the largest draw from a sequence of Chi-squareds. A type of extreme value distribution).

Structural stability: Multiple break dates

Bai & Perron (1998, 2003)

- Extend to multiple break points, select optimal number and timing of breaks.
  - R package “strucchange”, breakpoints() function.
- Idea:
  - Minimize sum of squared residuals (SSR) (run OLS) for a given break date t_1.
  - Repeat over possible dates t_1
  - Pick the date t_1 with lowest SSR.
  - Record the BIC for the break date with lowest SSR.
  - Repeat with two break dates t_1, t_2, then three, etc.
  - Pick the model/number of break dates with lowest AIC.

Structural stability: Fluctuation tests

CUSUM/MOSUM tests

- Mean-zero residuals should cancel-out when summed.
  - Time series of the cumulative sum of regression residuals should not deviate far from zero.
  - If the true coefficient changes in the middle of the sample, cumulative residuals will deviate from zero for periods of time
  - CUSUM = does cumulative sum of residuals from a model forcing no breaks exceed a threshold?
  - MOSUM = does moving sum of residuals (through a subsample window) exceed a threshold?
- CUSUM will see peaks near the break dates
- MOSUM will see a shift near the break dates

Structural stability: Fluctuation tests

Estimates-based tests: estimate coefficients using different subsample windows:

- Recursive estimates test:
  - Estimate coefficients from increasingly large subsample starting at beginning and widening the window to the end of the sample.
  - Compare to the full sample estimates.
  - Normalized difference in estimates will show a peak near the break date.
- Moving estimates test:
  - Estimate coefficients in a subsample window of fixed size, and move the window through time from beginning of sample to the end.
  - Compare to the full sample estimates.
  - Normalized difference in estimates will show a shift near the break date.