14.384 Time Series Analysis, Fall 2007
Professor Anna Mikusheva
Paul Schrimpf, scribe
October 30, 2007
revised November 3, 2009

Lecture 19

Breaks and Cointegration

The goal of this lecture is to analyze and test for another type of non-stationarity- break in parameters of a process. Empirical Processes Theory will be our working horse. Estimation of break date (once break is detected) is usually done by OLS, and is not discussed here.

Breaks

Suppose $y_t = \beta_t' x_{t-1} + \epsilon_t$, where

$$\beta_t = \begin{cases} \beta & t \leq t_0 \\ \beta + \gamma & t > t_0 \end{cases},$$

x_t and $E[\epsilon_t | I_t] = 0$ are iid variables. We want to test whether there is a break.

Break at a known date. Assume that the date of the break t_0 is known. Then we can run a regression $y_t = \beta_t' x_{t-1} + \gamma' \mathbb{I}_{t>t_0} x_{t-1} \epsilon_t$ and test for a break is a test of hypothesis $H_0 : \gamma = 0$. We could just do an F-test (which in this context is called Chow test).

$$F_T(\frac{t_0}{T}) = \frac{(SSR_{1,T} - (SSR_{1,t_0} + SSR_{t_0+1,T}))/k}{(SSR_{1,t_0} + SSR_{t_0+1,T})/(T-2k)}$$

where $SSR_{t,s}$ is the sum of squared residuals from OLS using the sample from time t to time s. Under the null and assuming that $t_0 = [\delta T]$ (there are increasing number of observations on both sides of the break), we have $F_T \Rightarrow \chi^2_k / k$, where k is the number of restrictions (the dimension of γ). This test is valid when t_0 is known.

Break at an unknown date.

When t_0 is not known, a test as above (testing for a break at a single specific date) is not powerful. Note that t_0 is a nuisance parameter, which is identified only under the alternative, but not under the null hypothesis. One test-statistic is the Quandt statistic that aims at a maximum value of a set of F-statistics for different break times:

$$Q = \sup_{[\delta T] \leq t_0 \leq [(1-\delta)T]} F_T(\frac{t_0}{T}) \quad (1)$$

The statistics seemed to be attractive and was known for a long time, but the asymptotic of it was unknown until Andrews' (1993) paper. Other test statistics are the Mann-Wald (average F-statistic):

$$MW = \frac{1}{T-2r} \sum_{t_0=r}^{T-r} F_T(\frac{t_0}{T}) \ , \ r = [\delta T] \quad (2)$$

and the Andrews-Ploberger (1994) (geometric average):

$$AP = \ln\left[\frac{1}{T-2r}\sum_{t_0=r}^{T-r}\exp\left(\frac{1}{2}F_T(\frac{t_0}{T})\right)\right] \tag{3}$$

To make the statistics above usable, we need to answer an important question: what are their asymptotic distributions? For this we'll use empirical process theory (Functional CLT). We make the following assumptions:

1. Uniform Law of Large Numbers: the following convergence hold uniformly in τ such that $0 < \delta < \tau < 1 - \delta < 1$:

$$\Psi_T(\tau) = \frac{1}{T}\sum_{t=1}^{[T\tau]} x_{t-1}x'_{t-1} \xrightarrow{p} \tau\Sigma_{xx}$$

   The above is a generalization of a law of large numbers $\frac{1}{T}\sum_{t=1}^{T} x_{t-1}x'_{t-1} \xrightarrow{p} \Sigma_{xx}$.

2. Functional central limit theorem (for which we need some additional assumptions, e.g. sufficient conditions would be that x and ϵ are independent and x_t are iid with finite fourth moments):

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[\delta T]} x_{t-1}\epsilon_t = \xi_T(\delta) \Rightarrow \sigma\Sigma_{xx}^{1/2}W_k(\cdot)$$

   where $W_k(\cdot)$ is k-dimensional Brownian motion, and σ^2 is the variance of ϵ_t.

To derive the limiting distribution of F statistics, we must look at the behavior of SSR. Let $\hat\beta$ be the OLS estimate from the sample from $t = 1, .., r$.

$$\begin{aligned} SSR_{1,r} =& \sum_{t=1}^{r}(y_t - \hat\beta' x_{t-1})^2 \\ =& \sum_{t=1}^{r}\epsilon_t^2 - 2(\hat\beta-\beta)'\sum^{r} x_{t-1}\epsilon_t + (\hat\beta-\beta)'\sum^{r} x_{t-1}x'_{t-1}(\hat\beta-\beta) \end{aligned}$$

Then, since $(\hat\beta - \beta)' = (\sum_{t=1}^{r} x_{t-1}x'_{t-1})^{-1}(\sum_{t=1}^{r} x_{t-1}\epsilon_t)$, we have

$$SSR_{1,r} = \sum_{t=1}^{r}\epsilon_t^2 - (\sum_{t=1}^{r} x_{t-1}\epsilon_t)'(\sum_{t=1}^{r} x_{t-1}x'_{t-1})^{-1}(\sum_{t=1}^{r} x_{t-1}\epsilon_t)$$

Combining this, we have, for $[\tau T] = r$:

$$\begin{aligned} SSR_{1,r} - \sum_{t=1}^{r}\epsilon_t^2 =& (\frac{1}{\sqrt{T}}\sum_{t=1}^{r} x_{t-1}\epsilon_t)'(\frac{1}{T}\sum_{t=1}^{r} x_{t-1}x'_{t-1})^{-1}(\frac{1}{\sqrt{T}}\sum_{t=1}^{r} x_{t-1}\epsilon_t) \\ =& \xi_T(\tau)'\Psi_T(\tau)^{-1}\xi_T(\tau) \\ \Rightarrow& \frac{\sigma^2}{\tau}W_k(\tau)'W_k(\tau) \end{aligned}$$

Now, looking at the numerator of F, we have

$$\begin{aligned} & (SSR_{1,T} - (SSR_{1,[T\tau]} + SSR_{[T\tau],T}))/k \\ \Rightarrow& \frac{\sigma^2}{k}\left(-W_k(1)'W_k(1) + \frac{W_k(\tau)'W_k(\tau)}{\tau} + \frac{(W_k(1) - W_K(\tau))'(W_k(1) - W_k(\tau))}{1-\tau}\right) \\ \Rightarrow& \frac{\sigma^2}{k}\frac{(W_k(\tau) - \tau W_k(1))'(W_k(\tau) - \tau W_k(1))}{\tau(1-\tau)} \end{aligned}$$

Under the null (of no break)

$$\frac{SSR_{1,[T\tau]} + SSR_{[T\tau],T)}}{T-2k} \to^p \sigma^2$$

Let us introduce a process $B_k(\tau) = W_k(\tau) - \tau W_k(1)$ called a Brownian bridge. It is a linear transformation of a Brownian motion that is required to be 0 at $t = 0$ and $t = 1$. Thus we have convergence of processes (as a weak convergence of processes)

$$F_T(\tau) \Rightarrow \frac{B_k(\tau)' B_k(\tau)}{k\tau(1-\tau)}$$

Now we can get asymptotic distributions of test statistics as they are continuous functionals of $F_T()$:

$$Q \Rightarrow \sup_{\delta \le \tau \le 1-\delta} \frac{B_k(\tau)' B_k(\tau)}{\tau(1-\tau)}$$

$$MW \Rightarrow \frac{1}{1-2\delta} \int_\delta^{1-\delta} \frac{B_k(\tau)' B_k(\tau)}{k\tau(1-\tau)} d\tau$$

$$AP \Rightarrow \log\left(\frac{1}{1-2\delta} \int_\delta^{1-\delta} \exp\left\{ \frac{B_k(\tau)' B_k(\tau)}{k\tau(1-\tau)} \right\} d\tau \right)$$

Conclusions and Remarks:

- A nice feature of the last statements is that the limiting distributions do not depend on nuisance parameters. We can simulate these distributions to find critical values. The test for H_0 : no break vs. H_a : there is one break can be performed by calculating the Q statistic in sample and comparing it to the simulated critical values. We will reject large values of Q. An alternative approach is bootstrapping the Q statistic.
- Critical values for Q test are bigger than for Chow test(known break date).
- If errors are auto-correlated one needs to correct for that (use HAC version of Chow test).
- Uniform Law of Large Numbers Assumption implicitly assumes stationarity of x_t.

More than one break

- Theoretically it is possible to get asymptotic distributions for statistics testing for two breaks.
- Practically: you have to simulate critical values, which in a case of even just 2 breaks is heavy computational burden. And complexity increases with dimensionality rapidly.
- Ideologically: If you allow for multiple breaks in a relatively small sample probably you should model the situation differently. Alternatives: describe a process for breaks to occur (deterministic or stochastic), or introduce continuously changing coefficients.

Recursive Estimation

Another idea for how to test for breaks is to look whether the estimation of β on different horizons is stable (so called recursive estimation). Let:

$$\tilde{\beta}(r/T) = (\sum_{t=1}^{r} x_{t-1} x_{t-1}')^{-1} (\sum_{t=1}^{r} x_{t-1} y_t)$$

One can calculate estimates of $\tilde{\beta}$ recursively and look at their stability. If $\tilde{\beta}$ changes a lot, it is a sign of a break. Formally, we need to formalize what does it mean "changes a lot", that is , we have to find the asymptotic distribution. Suppose H_0 : no breaks is true and β is the true coefficient. Then,

$$\sqrt{T}(\tilde{\beta}(\tau) - \beta) = \Psi_T(\tau)^{-1}\xi_T(\tau) \Rightarrow \frac{\sigma\Sigma_{xx}^{-1/2}W_k(\tau)}{\tau}$$

The problem is the nuisance parameters, β and Σ_{xx}. We can eliminate them from the asymptotic distribution by estimating β by $\tilde{\beta}(1)$ and looking at the associated t-statistic:

$$t_t(\tau) = \hat{\sigma}_\epsilon^{-1}(\frac{1}{T}\sum x_{t-1}x'_{t-1})^{1/2}\sqrt{T}(\tilde{\beta}(\tau) - \tilde{\beta}(1)) \Rightarrow \frac{W_k(\tau)}{\tau} - W_k(1)$$

and as above, you can use test statistic $\sup_{\delta\leq\tau}|t_T(\tau)|$ and critical values simulated from $\sup_\tau\{\frac{W_k(\tau)}{\tau} - W_k(1)\}$. There are many ways to test for breaks, and the way to derive the limiting distribution of test statistics is by using the FCLT.

Unit Root tests with a Break

Consider a unit root process with possible break in trend. The model is $y_t^* = y_t + d_t$ where y_t is a random walk and $d_t = c + \gamma\mathbf{1}_{t>t_0}$ is a trend. How do we then test for a unit root? Two observations:

1. The distribution of Dickey-Fuller statistics is very sensitive to the trend specification. If the trend in misspecified, for example if y_t^* regressed on const, y_{t-1}^* (the break is ignored), then you are likely to accept a unit root, even if there is no one. The intuition is that both a break and a unit root lead to permanent changes in y_t.

2. If the trend is correctly specified say $D_t^\tau = (1, \mathbf{1}_{t>[\tau T]})$ is for a known break date at time $t_0 = [\tau T]$, then we can get the asymptotics of Dickey-Fuller statistics under unit root assumption. For example,

$$T(\hat{\rho} - 1) \Rightarrow \frac{\int_0^1 P_{d^\tau}W(s)dW(s)}{\int_0^1 (P_{d^\tau}W(s))^2 ds},$$

   where $P_N = I - N(N'N)^{-1}N$ is a linear projection to a space perpendicular to N, and $d^\tau = (1, \mathbf{1}_{t>\tau})'$. In practice, you would simulate critical values imposing the correct break date.

Perron (1989) found that if you allow for breaks during the Great Depression (1929) and during the oil shocks (1973), then you reject the null of unit root in most macro series. Christiano(1992) and Zivot and Andrews(1992) objected to this by arguing that it is not fair to treat the break dates as known. If you test for unit roots without assuming the break dates are known, then you have to change a statistic. One suggestion would be

$$t_{DF}^{min} = \min_{\tau\in[\delta_0,\delta_1]} t(\tau),$$

where $t(\tau)$ is Dickey-Fuller t-statistic for a known break date at $t_0 = [\tau T]$. That surely, does not change the value of statistic from what Perron obtained, but suggest different critical values. As a result, one cannot reject the null of unit root in most series if assume "unknown break date".

Cointegration

Spurious Regression

Let x_t and y_t be two independent random walks,

$$y_t = y_{t-1} + z_t$$
$$x_t = x_{t-1} + u_t$$

where z_t and u_t are iid and independent of each other. Suppose we run OLS of y_t on x_t.

$$y_t = \beta x_t + e_t$$

The true value of β is 0. However, if you simulate this model and test the hypothesis $H_0 : \beta = 0$ using OLS t-statistic you will likely faulty reject it in a large portion of cases. That can be seen from the following lines.

$$\hat{\beta} = \frac{\frac{1}{T^2}\sum y_t x_t}{\frac{1}{T^2}\sum x_t^2}$$

Let assume that the functional central limit theorem works:

$$\xi_T(\tau) = \frac{1}{\sqrt{T}}\begin{pmatrix} x_{[\tau T]} \\ y_{[\tau T]} \end{pmatrix} \Rightarrow W(\tau) = \begin{pmatrix} W_1(\tau) \\ W_2(\tau) \end{pmatrix}$$

where $W(\tau)$ is a 2-dimensional Brownian motion. Since integrals are continuous functionals, we have:

$$\frac{1}{T}\sum \xi_T(t/T)\xi_T(t/T)' \Rightarrow \int W(\tau)W(\tau)'d\tau$$

Thus,

$$\hat{\beta} = \frac{\frac{1}{T^2}\sum y_t x_t}{\frac{1}{T^2}\sum x_t^2} \Rightarrow \frac{\int W_1 W_2 dt}{\int W_2^2 dt}$$

That is, $\hat{\beta}$ is not consistent. You won't receive zero even in very large samples. Next, we can notice that the estimate of variance of error term diverges:

$$\widehat{\sigma}^2 = \frac{1}{T}SSR = \frac{1}{T}\sum_{t=1}^{T}(y_t - \widehat{\beta}x_t)^2 = T(\frac{1}{T^2}\sum_{t=1}^{T} y_t^2 - 2\beta\frac{1}{T^2}\sum_{t=1}^{T} y_t x_t + \widehat{\beta}^2\frac{1}{T^2}\sum_{t=1}^{T} x_t^2)$$

$$\frac{1}{T}\widehat{\sigma}^2 \Rightarrow \left(\int W_1^2 dt - \frac{(\int W_1 W_2 dt)^2}{\int W_2^2 dt}\right)$$

As a result,

$$t = \frac{\widehat{\beta}}{\widehat{\sigma}}\sqrt{\sum x_t^2} = \sqrt{T}\frac{\widehat{\beta}}{\sqrt{\frac{1}{T}\widehat{\sigma}^2}}\sqrt{\frac{1}{T^2}\sum x_t^2}$$

We can see that $|t| \to^p \infty$, that is, in large samples you will faulty reject the null in almost all samples. The important point here is that with non-stationary regressors, we get behavior of OLS estimates.

14.384 Time Series Analysis
Fall 2013