Inference on the time of break

Štěpána Lazarová (Queen Mary)

Javier Hidalgo (LSE)

Cass Business School, London

11 December 2006

Linear regression with break

$$
\begin{aligned}
y_t &= \alpha + \beta_1' x_t + u_t \qquad & t = 1, \ldots, k_0 \\
y_t &= \alpha + \beta_2' x_t + u_t \qquad & t = k_0 + 1, \ldots, T
\end{aligned}
$$

$\delta = \delta_T = \beta_2 - \beta_1 \neq 0 \ldots$ size of break

$k_0 = [\tau_0 T] \qquad \tau_0 \in (0, 1) \qquad \tau_0$ unknown

Goals

- estimation of the date of break: $\hat{k}$
- inference about the date of break

Standard assumption:

Size of break is diminishing as the sample size increases

Main questions:

motivation?

justification?

can it be relaxed?

Cost of assuming shrinking break:

Rate of convergence

- Shrinking break
  $T \to \infty$: $\qquad \delta = \delta_T \to 0, \qquad T\,||\delta_T||^2 \to \infty$
  $$\hat{k} - k_0 = O_p\left(||\delta_T||^{-2}\right)$$

- Fixed break
  $\delta_T = \delta$
  $$\hat{k} - k_0 = O_p\left(1\right) = o_p\left(||\delta_T||^{-2}\right)$$

Fixed break: tighter confidence intervals for $\hat{k}$

Asymptotic distribution

- Shrinking break

$$v_T^2 \left(\hat{k} - k_0\right) \xrightarrow{d} \arg\min_{\tau} \bar{W}(\tau)$$

where $v_T \sim \|\delta_T\|$ and

$$\bar{W}(\tau) = \begin{cases} W_1(\tau) + \frac{|\tau|}{2} & \tau > 0, \\ W_2(-\tau) + \frac{|\tau|}{2} & \tau \leq 0, \end{cases}$$

W_1, W_2... independent standard Brownian motion processes on $[0, \infty)$

- distribution free
- nuisance parameter free
- distribution function known explicitly

Asymptotic distribution

- Fixed break

$$\hat{k} - k_0 \xrightarrow{d} \arg\min_k W(k)$$

where

$$W(k) = \begin{cases} \delta' \sum_{t=1}^{k} x_t x_t' \delta - 2\delta' \sum_{t=1}^{k} x_t u_t & k \geq 1, \\ 0 & k = 0 \\ \delta' \sum_{t=k}^{0} x_t x_t' \delta + 2\delta \sum_{t=k}^{0} x_t u_t & k \leq -1. \end{cases}$$

  - δ unknown
  - x_t, u_t ...unknown distribution

Inference under fixed break

- Hinkley (1970)
  - assumption: distribution of x_t, u_t is known
  - simple case of $x_t = 1$ and $u_t \sim$ iid $N(0,1)$
  - analytical solution for density function

- Antoch, Hušková (1999)

$$\begin{aligned} y_t^* &= \hat{\alpha} + \hat{\beta}_1' x_t + u_t^* \qquad & t = 1, \ldots, \hat{k} \\ y_t^* &= \hat{\alpha} + \hat{\beta}_2' x_t + u_t^* \qquad & t = \hat{k}+1, \ldots, T \end{aligned}$$

  iid data: Efron's (1979) bootstrap

Our paper

- Assume the size of break is fixed

- Allow for strongly dependent data

- $\hat{k}\ldots$ least squares estimator

- Estimate distribution of $\hat{k}$ by a bootstrap procedure

Results

- Assumption of fixed break technically possible

- Distribution of $\hat{k}$ can be approximated by bootstrap

Assumptions

- x_t, u_t strictly stationary, mutually independent

- $Ex_t = 0$, $Eu_t = 0$, finite fourth moments

- $0 \leq d_x + d_u < \frac{1}{2}$

Bootstrap procedure

(1) obtain bootstrap sample of u_t^*

(2) construct bootstrap sample

$$\begin{aligned} y_t^* &= \hat{\alpha} + \hat{\beta}_1' x_t + u_t^* \qquad t = 1, \ldots, \hat{k} \\ y_t^* &= \hat{\alpha} + \hat{\beta}_2' x_t + u_t^* \qquad t = \hat{k}+1, \ldots, T \end{aligned}$$

(3) estimate the time of break $\hat{k}^*$

(4) repeat many times and approximate distribution of $\hat{k} - k_0$ by the bootstrap distribution of $\hat{k}^* - \hat{k}$

Bootstrapping dependent data

- parametric bootstrap
  - AR(p) parametric process - Franke, Kreiss (1992)
- nonparametric bootstrap
  - blocks - Carlstein (1986), Künsch (1989),
  - subsampling - Politis, Romano (1992)
  - sieve bootstrap - Bühlmann (1997, 1998)
- frequency domain bootstrap
  - periodogram (Franke, Härdle (1992))
  - discrete Fourier transform (Hidalgo (2003), Lazarová (2005, 2006))

Bootstrapping dependent u_t

no model assumed - nonparametric procedure needed

Wold decomposition:

$$
\begin{aligned}
u_t &= \sum_{j=0}^{\infty} b_j \varepsilon_{t-j} \qquad b_0 = 1 \\
&= B(L)\,\varepsilon_t \\
&= (1-L)^{-d}\,\Psi(L)\,\varepsilon_t \qquad\qquad 0 < d < \frac{1}{2}
\end{aligned}
$$

$B(z) = \sum_{j=0}^{\infty} b_j z^j$ $\qquad$ $\Psi(z) = \sum_{j=0}^{\infty} \psi_j z^j$

$\varepsilon_t \ldots$ white noise $\qquad$ $\sum_{j=0}^{\infty} \psi_j^2 < \infty$

$$
f_u(\lambda) = \left|1 - e^{-i\lambda}\right|^{-2d} \left|\Psi\left(e^{-i\lambda}\right)\right|^2 f_\varepsilon(\lambda)
$$

$u_t = (1 - L)^{-d} \Psi(L) \varepsilon_t$

if we knew u_t, d, Ψ:

1. deconvolve

$$\varepsilon_t = \Psi(L)^{-1} (1 - L)^d u_t$$

2. resample $\varepsilon_t \longrightarrow$ obtain ε_t^*

3. construct

$$u_t^* = (1 - L)^{-d} \Psi(L) \varepsilon_t^*$$

but u_t, d, Ψ not known

Estimators of u_t, d, $\mathbf{\Psi}$

u_t... estimated by OLS residuals $\hat{u}_t$

d... estimated from residuals by e.g. local Whittle estimator of Robinson (1995)

$\mathbf{\Psi}$... obtain estimate $\hat{f}_u$ of spectral density f_u

put $\left|\hat{\mathbf{\Psi}}\left(e^{-i\lambda_j}\right)\right|^2 = \hat{f}_u(\lambda)\left(1 - e^{-i\lambda_j}\right)^{2\hat{d}}$

obtain $\hat{\mathbf{\Psi}}\left(e^{-i\lambda_j}\right)$ by the canonical decomposition

With estimated $\hat{u}_t$, $\hat{d}$, $\hat{\mathbf{\Psi}}$:

1. $\hat{\varepsilon}_t = \hat{\mathbf{\Psi}}(L)^{-1}(1-L)^{\hat{d}}\,\hat{u}_t$

2. resample $\hat{\varepsilon}_t \longrightarrow$ obtain $\hat{\varepsilon}_t^*$

3. construct

$$u_t^* = (1-L)^{-\hat{d}}\,\hat{\mathbf{\Psi}}(L)\,\hat{\varepsilon}_t^*$$

with appropriate truncation

Bootstrap procedure

(1) obtain bootstrap sample of u_t^* by the above method

(2) construct bootstrap sample

$$\begin{aligned} y_t^* &= \hat{\alpha} + \hat{\beta}_1' x_t + u_t^* \qquad t = 1, \ldots, \hat{k} \\ y_t^* &= \hat{\alpha} + \hat{\beta}_2' x_t + u_t^* \qquad t = \hat{k}+1, \ldots, T \end{aligned}$$

(3) estimate the time of break $\hat{k}^*$

(4) repeat many times and approximate distribution of $\hat{k} - k_0$ by the bootstrap distribution of $\hat{k}^* - \hat{k}$

Results

Under regularity conditions,

(1) for any finite K,

$$\left(u^*_{k_0-K},\ldots.u^*_{k_0+K}\right)\xrightarrow{d^*}\left(u_{k_0-K},\ldots.u_{k_0+K}\right)$$

(2)

$$\hat{k}^*-\hat{k}=O_{p^*}(1)$$

(3)

$$\hat{k}^*-\hat{k}\xrightarrow{d^*}\arg\min_k W(k)$$

(4)

$$T^{1/2}\begin{pmatrix}\hat{\beta}^*_1-\hat{\beta}_1\\ \hat{\beta}^*_2-\hat{\beta}_2\end{pmatrix}\xrightarrow{d^*}N(0,V)$$

Conclusions

- Breaks should be assumed fixed

- Breaks can be assumed fixed

- Distribution of $\hat{k}$ can be estimated by a bootstrap procedure

- Bootstrap based on two-step prewhitening (fractional differencing and sieve)

Further steps and possible extensions

- more than one break

- deterministic components

- bootstrap in time domain
  Kapetanios and Psaradakis (2006), Poskitt (2006)
  - comparison of performance

Lag order

- Bühlmann (1997): AR(p_T) with

$$p_n = o\left((n/\log n)^{1/4}\right)$$

- our procedure:

$$\begin{aligned} M &= o\left((n/\log n)^{1/3}\right) \\ M^{-1} &= o\left(n^{-1/2}\right) \end{aligned}$$

  - our procedure: greater number of lags
  - this is to be expected as our procedure allows for strong dependence of data

Bootstrap procedure

1. Obtain centered least squares residuals

   $$\hat{u}_t = (y_t - \bar{y}) - \hat{\beta}' (x_t - \bar{x}) - \hat{\delta}' (\hat{z}_t - \bar{z}), \quad t = 1, ..., n.$$

2. Estimate d by the local Whittle estimator $\hat{d}$ proposed by Robinson (1995b),

   $$\hat{d} = \arg \min_{a \in [0,\Delta]} H(a),$$

   where $0 < \Delta < 1/2$,

   $$H(a) = \log\left(\frac{1}{m}\sum_{j=1}^{m} \lambda_j^{2a} I_{\hat{u}\hat{u}}\left(\lambda_j\right)\right) - \frac{2a}{m}\sum_{j=1}^{m} \log \lambda_j$$

   for integer $m \in [1, [n/2])$, where $I_{\hat{u}\hat{u}}(\lambda) = |w_{\hat{u}}(\lambda)|^2 / (2\pi)$ is the periodogram of $\hat{u}_t$, with the bandwidth m satisfying condition $1/m + m/n \to 0$ as $n \to \infty$.

3. Let

   $$\hat{h}(\lambda) = \frac{1}{2m+1}\sum_{j=-m}^{m} \left|1 - e^{-i(\lambda + \lambda_j)}\right|^{2\hat{d}} I_{\hat{u}\hat{u}}\left(\lambda + \lambda_j\right).$$

4. Let $M = \left[\frac{n}{4m}\right]$ and compute

$$\hat{c}_r = \frac{1}{n} \sum_{l=m+1}^{[n/2]} \log \hat{h}_l e_l^{ir\lambda_j}, \quad r = 1, ..., M.$$

5. Let

$$\begin{aligned} \hat{\Psi}\left(e^{-i\lambda_j}\right) &= \exp\left\{-\sum_{r=1}^{M} \hat{c}_r e^{-ir\lambda_j}\right\}, \\ \hat{g}(\lambda) &= \left(1 - e^{-i\lambda}\right)^{-\hat{d}}. \end{aligned}$$

6. Estimate innovations $\hat{\varepsilon}_t$ as

$$\hat{\varepsilon}_t = \frac{1}{n^{1/2}} \sum_{j=1}^{n} e^{it\lambda_j} \widehat{\hat{R}}_j w_{\hat{u}}\left(\lambda_j\right), \quad t = 1, ..., n,$$

where

$$\begin{aligned} \widehat{\hat{R}}_j &= \sum_{l=0}^{M} \hat{\rho}_l e^{-il\lambda_j}, \quad j = 1, ..., n, \\ \hat{\rho}_l &= \frac{1}{n} \sum_{j=1}^{n-1} \hat{R}_j e^{il\lambda_j}, \quad l = 0, ..., M, \\ \hat{R}_j &= \hat{g}^{-1}\left(\lambda_j\right) \hat{\Psi}^{-1}\left(e^{-i\lambda_j}\right), \quad j = 1, ..., n-1. \end{aligned}$$

7. Draw a random sample $\varepsilon^* = \left(\varepsilon_1^*, \varepsilon_2^*, ..., \varepsilon_n^*\right)'$ with replacement from the empirical distribution of the residuals $\hat{\varepsilon}_t$, $P^*\left(\varepsilon_i^* = \hat{\varepsilon}_j\right) = n^{-1}$ for $i, j = 1, \ldots, n$, and compute the discrete Fourier transform w_{ε^*} of ε^*, that is

$$w_{\varepsilon^*}\left(\lambda_j\right) = \frac{1}{n^{1/2}} \sum_{t=1}^{n} \varepsilon_t^* e^{-it\lambda_j}, \quad j = 1, \ldots, n.$$

8. Compute

$$u_t^* = \frac{1}{n^{1/2}} \sum_{j=1}^{n} e^{it\lambda_j} \widehat{\hat{B}}_j w_{\varepsilon^*, j}\left(\lambda_j\right), \quad t = 1, ..., n,$$

where

$$\begin{aligned} \widehat{\hat{B}}_j &= \sum_{l=0}^{M} \hat{b}_l e^{-il\lambda_j}, \quad j = 1, \ldots, n, \\ \hat{b}_l &= \frac{1}{n} \sum_{j=1}^{n-1} \hat{B}_j e^{il\lambda_j}, \quad l = 0, \ldots, M, \\ \hat{B}_j &= \hat{g}\left(\lambda_j\right) \hat{\boldsymbol{\Psi}}\left(e^{-i\lambda_j}\right), \quad j = 1, \ldots, n-1. \end{aligned}$$

9. Construct bootstrap sample y_t^*,

$$y_t^* = \hat{\beta}' x_t + \hat{\delta}' \hat{z}_t + u_t^*, \quad t = 1, ..., n.$$

Estimation of spectral density

prewhitening, smoothing

$$\hat{f}_v(\lambda) = \frac{1}{2m+1} \sum_{j=-m}^{m} \left|\lambda + \lambda_j\right|^{2\hat{d}} I_{\hat{u}\hat{u}}\left(\lambda + \lambda_j\right)$$

where

$$m = m(T) \to 0 \qquad \text{as } T \to \infty$$

Examining parameter instability

- Model diagnostics

  Model adequate?

  More structure needed?

- Instability itself of interest

  Alternative: One-time structural break

  Was there a break? When?

Linear regression with break

$$\begin{aligned} y_t &= \alpha + \beta_1' x_t + u_t \qquad t = 1, \ldots, k_0 \\ y_t &= \alpha + \beta_2' x_t + u_t \qquad t = k_0 + 1, \ldots, T \end{aligned}$$

Least squares estimator $\hat{k}$ of k_0

$\hat{u}_k \ldots$vector of residuals of regression with break at k

$S(k) = ||\hat{u}_k||^2 \ldots$ sum of squares

Estimator of the breakpoint:

$$\hat{k} = \arg\min_k S(k)$$

Assumptions

- $\{x_t\}$ and $\{u_t\}$ are strictly stationary linear processes

  f_{xx}, f_{uu}... spectral density functions of $\{x_t\}$, $\{u_t\}$

-

$$x_t = \sum_{j=0}^{\infty} a_j \xi_{t-j}, \quad \sum_{j=0}^{\infty} a_j^2 < \infty, \quad a_0 = I$$

$$u_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} b_j^2 < \infty, \quad b_0 = 1$$

- $\{\xi_t\}$ is a stochastic process that satisfies $E\xi_t = 0$, $E\xi_t\xi_t' = \Xi$, ξ_t has finite fourth moments

- $\{\varepsilon_t\}$ is a stochastic process that is independent of $\{\xi_t\}$ and that satisfies $E\varepsilon_t = 0$, $E\varepsilon_t^2 = \sigma_\varepsilon^2$ and $E\varepsilon_t^4 < \infty$

- The matrix functions

$$A(\lambda) = \sum_{j=0}^{\infty} a_j e^{ij\lambda} \text{ and } B(\lambda) = \sum_{j=0}^{\infty} b_j e^{ij\lambda}$$

satisfy the following:

1. $\left|A_{jj}(\lambda)\right| \sim C_{x,j}\lambda^{-d_{x,j}}$, $|B(\lambda)| \sim C_u\lambda^{-d_u}$ as $\lambda \to 0+$, $d_{x,j}, d_u \in \left(-\frac{1}{2}, \frac{1}{2}\right)$

2. $A(\lambda)$ and $B(\lambda)$ are twice continuously differentiable on $(0, \pi]$ and $\left\|\frac{dA(\lambda)}{d\lambda}\right\| = O\left(\frac{||A(\lambda)||}{\lambda}\right)$, $\left|\frac{dB(\lambda)}{d\lambda}\right| = O\left(\frac{|B(\lambda)|}{\lambda}\right)$,

3. $||A(\lambda)|| > 0$ and $|B(\lambda)| > 0$ for all $\lambda \in (0, \pi]$.

- $\Omega = 2\pi \int_{-\pi}^{\pi} f_{xx}(\lambda) f_{uu}(\lambda) d\lambda < \infty$,

  $\Sigma = E(x_t^2) > 0$

Consistent estimators of Σ, Ω

$$\begin{aligned}\Sigma &= E x_t x_t' \\ \hat{\Sigma} &= \frac{1}{T}\sum_{t=1}^{T} x_t x_t'\end{aligned}$$

and

$$\begin{aligned}\Omega &= 2\pi \int_{-\pi}^{\pi} f_{xx}(\lambda)\, f_{uu}(\lambda) d\lambda \\ \hat{\Omega} &= \frac{4\pi^2}{T}\sum_{j=1}^{T-1} I_{xx}\left(\lambda_j\right) I_{\hat{u}\hat{u},j}\left(\lambda_j\right)\end{aligned}$$

as proposed by Robinson (1998)

Fourier transform

Discrete Fourier transform of v_t for $t = 1, \ldots, T$:

$$w_v(\lambda) = (2\pi T)^{-\frac{1}{2}} \sum_{t=1}^{T} v_t e^{it\lambda} \qquad 0 \leq \lambda < 2\pi$$

Fourier frequencies: $\lambda_j = 2\pi j / T \qquad j = 1, \ldots, T$

Inverse discrete Fourier transform:

$$v_t = \left(\frac{2\pi}{T}\right)^{-\frac{1}{2}} \sum_{j=1}^{T} w_v\left(\lambda_j\right)$$

Periodogram:

$$I_v(\lambda) = |w_v(\lambda)|^2$$

Time domain

Berk (1974)

$$\sum_{j=0}^{\infty} \phi_j v_{t-j} = \varepsilon_t \qquad \phi_0 = 1$$

AR(∞) representation of the process

Truncation

$$v_t = -\sum_{j=0}^{k(T)} \phi_j v_{t-j} + \varepsilon_t$$

Some necessary conditions:

$$\begin{aligned} \frac{k^3}{T} &\rightarrow 0 \ldots \text{consistency} \\ \frac{k}{T} &\rightarrow 0 \ldots \text{asymptotic normality} \end{aligned}$$

Bühlmann (1997, 1998): sieve bootstrap

$$v_t^* = -\sum_{j=0}^{k(T)} \hat{\phi}_j v_{t-j}^* + \varepsilon_t^*$$

Robinson (1987)

$$\sum_{j=0}^{\infty} \hat{\phi}_j \left(\theta\right) v_{t-j} = \hat{\varepsilon}_t \qquad t = 1, \ldots, T$$

recovering innovations

Possible solutions

(1) assume distribution of x_t, u_t is known

(2) assume size of break is shrinking with sample size

(3) estimate distribution of x_t, u_t

Cumulative distribution function of $\arg\min_{\tau} W(\tau)$ known

Picard (1985)

Bhattacharya (1987)

Yao (1987)

Bai (1994, 1997a,b)

Bai and Perron (1998)

Fiteni (2002, 2004)

Lazarova (2005, 2006)