Efficiently Detecting Multiple Structural Breaks in Systems of Linear Regression Equations with Integrated and Stationary Regressors

Karsten Schweikert

University of Hohenheim

August 24, 2022

1. Introduction

Motivation

- Ignoring structural breaks in time series regression models leads to inconsistently estimated coefficients.
- Particularly important for regression models involving long spans of data, which are more likely to be affected by structural breaks.
- Focus on linear regressions with multivariate responses and a mix of integrated and stationary regressors. Background
- Precisely determine the number of structural breaks, their timing, and simultaneously estimate the model's coefficients for each regime,
- ... but do so with an algorithm that keeps computational costs as low as possible.

1. Introduction

Literature

Available approaches:

- Grid search over all possible combinations of breakpoints and optimize a suitable information criterion (Yao, 1988; Liu et al., 1997).
- Almost exact segmentation algorithms like the likelihood-based approach of Qu and Perron (2007), Eo and Morley (2015), Li and Perron (2017), and Oka and Perron (2018). Uses a dynamic programming algorithm to reduce the number of estimations. Details
- Penalized regression similar to the two-step approach proposed by Chan et al. (2014). First, uses the group LARS algorithm to find an initial set of breakpoint candidates, then uses backward elimination to eliminate irrelevant breaks.

Approach	Comp. costs	Precision
Grid search	$O(T^M)$	$+++$
Likelihood-based	$O(T^2)$	$+++$
Penalized regression	$O(M^3 + MT)$	$++$

2. Methodology

We consider the following potentially cointegrated system in triangular form

$$\begin{aligned} Y_t &= AX_t + \delta t + \mu + Bw_t + u_t, \qquad t = 1, 2, \ldots, \\ X_t &= X_{t-1} + \xi_t, \end{aligned} \tag{1}$$

where

- Y_t is a $q \times 1$ vector of dependent variables,
- X_t is a $r \times 1$ vector of integrated regressors,
- w_t is a $s \times 1$ vector of stationary regressors,
- u_t and ξ_t are stationary error processes.

2. Methodology

Our model (1)

$$\begin{aligned} Y_t &= AX_t + \delta t + \mu + Bw_t + u_t, \qquad t = 1, 2, \ldots, \\ X_t &= X_{t-1} + \xi_t, \end{aligned}$$

allows for very flexible specifications and covers several special cases:

- SUR models for $A = 0$ and $\delta = 0$,
- VAR(s) models for $A = 0$ and $\delta = 0$ and $w_t = (Y_{t-1}, \ldots, Y_{t-s})'$, see also Gao et al. (2020) and Safikhani and Shojaie (2022),
- or dynamically augmented cointegrating regressions if w_t contains the leads and lags of changes in X_t.

2. Methodology

Assumptions

Assumption (1)

(i) $u_t = \sum_{j=0}^{\infty} C_j \epsilon_{t-j} = C(L)\epsilon_t$, $\xi_t = \sum_{j=0}^{\infty} D_j e_{t-j} = D(L)e_t$, $C(1)$ *and* $D(1)$ *are full rank,* $\sum_{j=0}^{\infty} j\|C_j\| < \infty$ *and* $\sum_{j=0}^{\infty} j\|D_j\| < \infty$, (ϵ_t, e_t) *are i.i.d. with finite* $4 + a$ $(a > 0)$ *moment.* w_t *is a mean-zero second order stationary process with uniformly bounded* $4 + a$ *moment.*
(ii) Further, we require that

$$\sup_T E\left|\frac{1}{T}\sum_{i=1}^{t} X_{l,i}u_i\right|^{4+\epsilon} < \infty, \qquad \text{for } 1 \le l \le r,\ 1 \le t \le T \text{ and some } \epsilon > 0,$$

and

$$\sup_T E\left|\frac{1}{T}\sum_{i=1}^{t} w_{l,i}u_i\right|^{4+\epsilon} < \infty, \qquad \text{for } 1 \le l \le s,\ 1 \le t \le T \text{ and some } \epsilon > 0.$$

2. Methodology

Assumptions

Assumption (2)

The error process u_t is independent of the regressors for all leads and lags.

Assumption (3)

$$E\left(\frac{tX_{l,t}^2}{X_{l,1}^2 + X_{l,2}^2 + \cdots + X_{l,t}^2}\right) \leq M_X, \qquad \forall t \geq 1 \text{ and } l = 1, \ldots, r.$$

Assumption (4)

$$E\left(\frac{tw_{l,t}^2}{w_{l,1}^2 + w_{l,2}^2 + \cdots + w_{l,t}^2}\right) \leq M_w, \qquad \forall t \geq 1 \text{ and } l = 1, \ldots, s.$$

2. Methodology

We rewrite Equation (1) in its stacked form and apply scaling factors:

$$Y_t = (Z_t' \otimes I)\theta + u_t, \tag{2}$$

where

- $Z_t = (T^{-1/2}X_t', T^{-1}t, 1, w_t')'$ and
- $\theta = \mathrm{Vec}(A, \delta, \mu, B)$ is a $d = q(r + 2 + s)$ column vector, concatenating the coefficients for each regressor over all equations.

We apply scaling factors so that the order of all regressors is the same (Li and Perron, 2017; Oka and Perron, 2018).

2. Methodology

We assume that the true data generating process includes m^0 true, but unknown, (partial) structural breaks,

$$Y_t = (Z_t' \otimes I)\theta_{t_0} + \sum_{k=1}^{m} d(t_k)(Z_t' \otimes I)S'S\theta_{t_k} + u_t, \qquad (3)$$

where

- $d(t_k) = 0$ for $t \leq t_k$, $d(t_k) = 1$ for $t \geq t_k$,
- S is a selection matrix, and
- m denotes the total number of structural break candidates.

If we do not have any prior knowledge about the number and timing of structural breaks, each point in time has to be considered as a potential breakpoint.

Estimate the model in Equation (3) with $m = T$ under the condition that the set $\theta(T) = \{\theta_1, \theta_2, \ldots, \theta_T\}$ exhibits a certain sparse nature (m^0 nonzero groups).

2. Methodology

To use a convenient matrix notation, we define

$$\mathcal{Z} = \begin{pmatrix} Z_1' & 0 & 0 & \dots & 0 \\ Z_2' & Z_2' & 0 & \dots & 0 \\ Z_3' & Z_3' & Z_3' & \dots & 0 \\ \vdots & & & & \\ Z_T' & Z_T' & Z_T' & \dots & Z_T' \end{pmatrix}, \tag{4}$$

$\mathcal{Y} = (Y_1', \dots, Y_T')'$, $\mathcal{U} = (u_1', \dots, u_T')'$ and $\boldsymbol{\theta}(T) = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_T)'$.
Furthermore, we define $\boldsymbol{Y} = \text{Vec}(\mathcal{Y})$, $\boldsymbol{Z} = I \otimes \mathcal{Z}$, and $\boldsymbol{U} = \text{Vec}(\mathcal{U})$.

Now, the system for T breakpoint candidates can be rewritten as

$$\boldsymbol{Y} = \boldsymbol{Z}\boldsymbol{\theta}(T) + \boldsymbol{U}, \tag{5}$$

where $\boldsymbol{Y} \in \mathbb{R}^{Tq\times 1}$, $\boldsymbol{Z} \in \mathbb{R}^{Tq\times Td}$, $\boldsymbol{\theta}(T) \in \mathbb{R}^{Td\times 1}$, and $\boldsymbol{U} \in \mathbb{R}^{Tq\times 1}$.

2. Methodology

Estimate the set of coefficient changes $\boldsymbol{\theta}(T)$ by minimizing the following penalized least squares objective function (Yuan and Lin, 2006):

$$Q^*(\boldsymbol{\theta}(T)) = \frac{1}{T}\left\|\boldsymbol{Y} - \boldsymbol{Z}\boldsymbol{\theta}(T)\right\|^2 + \lambda_T \sum_{i=1}^{T} \|\boldsymbol{\theta}_i\|, \qquad (6)$$

where λ_T is a tuning parameter and $\|\cdot\|$ denotes the L_2-norm.

Minimizing the objective function in (6) yields the group LASSO estimator $\tilde{\boldsymbol{\theta}}(T)$.

2. Methodology

Asymptotic theory – first step estimator

Assumption (5)

(i) The break magnitudes are bounded to satisfy
$m_\theta = \min_{1\leq j\leq m_0+1} \|\boldsymbol{\theta}^0_{t^0_{j-1}}\| \geq \nu > 0$ *and*
$M_\theta = \max_{1\leq j\leq m_0+1} \|\boldsymbol{\theta}^0_{t^0_{j-1}}\| \leq \mathcal{V} < \infty$.
(ii) $\min_{1\leq j\leq m_0+1} |t^0_j - t^0_{j-1}|/T\gamma_T \to \infty$ *for some* $\gamma_T \to 0$ *and* $\gamma_T/\lambda_T \to \infty$ *as* $T \to \infty$.

Theorem (1)

Under Assumption 1 and Assumption 5, if $\lambda_T = 2dc_0(\log T/T)^{1/4}$ *for some* $c_0 > 0$*, then there exists some* $C > 0$ *such that with probability greater than* $1 - \frac{C}{c_0^4 \log T}$,

$$\frac{1}{T}\|\boldsymbol{Z}\left(\tilde{\boldsymbol{\theta}}(T) - \boldsymbol{\theta}^0(T)\right)\|^2 \leq 4dc_0\left(\frac{\log T}{T}\right)^{\frac{1}{4}}(m_0+1)M_\theta.$$

2. Methodology

Asymptotic theory – first step estimator

We define the Hausdorff distance $d_H(A, B) = \max_{b \in B} \min_{a \in A} |b - a|$ with $d_H(A, \emptyset) = d_H(\emptyset, B) = 1$, where $\emptyset$ is the empty set.

Theorem (2)

If Assumption 1 - 5 hold, then as $T \to \infty$

$$P\left(|\mathcal{A}_T| \geq m_0\right) \to 1,$$

and

$$P\left(d_H(\mathcal{A}_T, \mathcal{A}) \leq T\gamma_T\right) \to 1.$$

The group LASSO estimator **overestimates the number of breaks** which necessitates a second step refinement to solve the change-point problem!

2. Methodology

Second step estimator

To distinguish between active and non-active breakpoints in the set $\mathcal{A}_T$, we employ an information criterion for the second step.

We define $\widehat{\widehat{\boldsymbol{\theta}}}_j$, $1 \leq j \leq m_0$ as the least squares estimator of $\boldsymbol{\theta}_j^0$, based on breakpoints estimated in the first step. Further, we define the sum of squared residuals over all q equations as

$$S_T(t_1, \ldots, t_m) = \sum_{j=1}^{m+1} \sum_{t=t_{j-1}}^{t_j - 1} \| Y_t - \bar{Z}_t \sum_{s=1}^{j} \boldsymbol{K}' \widehat{\widehat{\boldsymbol{\theta}}}_s \|^2, \quad (7)$$

where $\bar{Z}_t = (Z_t' \otimes I)$. For m and the breakpoints $\boldsymbol{t} = (t_1, \ldots, t_m)$, we can define the information criterion (IC)

$$IC(m, \boldsymbol{t}) = S_T(t_1, \ldots, t_m) + m\omega_T, \quad (8)$$

where ω_T is the penalty term.

2. Methodology

Second step estimator

We determine the number of breaks and their timing by solving

$$(\widehat{\widehat{m}}, \widehat{\widehat{\boldsymbol{t}}}) = \arg \min_{\substack{m \in \{1,\ldots,|\mathcal{A}_T|\} \\ \boldsymbol{t}=(t_1,\ldots,t_m)\subset \mathcal{A}_T}} IC(m, \boldsymbol{t}). \tag{9}$$

Theorem (3)

If Assumption 1 - 5 hold and ω_T satisfies the conditions $\lim_{T\to\infty} T\gamma_T/\omega_T = 0$ and $\lim_{T\to\infty} \omega_T / \min_{1\leq i\leq m_0} |t_i^0 - t_{i-1}^0| = 0$, then, as $T \to \infty$, $(\widehat{\widehat{m}}, \widehat{\widehat{\boldsymbol{t}}})$ satisfies

$$P\left(\widehat{\widehat{m}} = m_0\right) \to 1,$$

and it exists a constant $B > 0$ such that

$$P\left(\max_{1\leq i\leq m_0} |\widehat{\widehat{t}}_i - t_i^0| \leq BT\gamma_T\right) \to 1.$$

2. Methodology

Second step estimator

- If the number of breaks in $\mathcal{A}_T$ is large, use the backward elimination algorithm (BEA) to determine m and $\boldsymbol{t}$.
- The BEA successively eliminates breakpoints until no improvement in terms of the IC can be reached.
- We can show the same consistency results for the BEA.
- The conditions for ω_T given in Theorem 3 are satisfied, e.g., for $\omega_T = CT^{3/4} \log T$, $C > 0$ (BIC: $\omega_T = C \log T$).
- C can be chosen analogously to the BIC, penalizing the total number of coefficients.

3. Simulation

In our simulation experiment, we consider model specifications with one, two and four breakpoints, respectively. The following DGP is employed

$$\begin{array}{rclcrcl} Y_t &=& A_t X_t + \delta_t t + \mu + B_t w_t + u_t, & & u_t &\sim& N(0, \Sigma_u), \\ X_t &=& X_{t-1} + \xi_t, & & \xi_t &\sim& N(0, \Sigma_\xi), \\ w_t &=& \Phi w_{t-1} + e_t, & & e_t &\sim& N(0, \Sigma_e), \end{array} \tag{10}$$

where $X_t = (X_{1t}, X_{2t}, \ldots, X_{Nt})'$, $\Sigma_u = diag(\sigma_u^2)$, $\Sigma_\xi = diag(\sigma_\xi^2)$ and $\Sigma_e = diag(\sigma_e^2)$.
For the main results, setting $q = 2$ and $c = 1$, we use the following coefficient matrices:

$$A_0 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad A_i = A_{i-1} + c \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad i = 1, \ldots, m,$$

$$B_0 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad B_i = B_{i-1} + c \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad i = 1, \ldots, m.$$

3. Simulation

Group LARS + BEA

T	pce	τ			
	SB1: ($\tau = 0.5$)				
100	98.9	0.501 (0.014)			
200	100	0.500 (0.007)			
400	100	0.500 (0.003)			
800	100	0.500 (0.001)			
	SB2: ($\tau_1 = 0.33$, $\tau_2 = 0.67$)				
T	pce	τ_1	τ_2		
150	91.7	0.337 (0.030)	0.660 (0.024)		
300	98.4	0.334 (0.017)	0.667 (0.014)		
600	99.8	0.332 (0.009)	0.668 (0.007)		
1200	100	0.331 (0.004)	0.669 (0.003)		
	SB4: ($\tau_1 = 0.2$, $\tau_2 = 0.4$, $\tau_3 = 0.6$, $\tau_4 = 0.8$)				
T	pce	τ_1	τ_2	τ_3	τ_4
250	89.0	0.217 (0.030)	0.404 (0.022)	0.596 (0.019)	0.788 (0.028)
500	98.2	0.203 (0.017)	0.402 (0.012)	0.598 (0.009)	0.803 (0.012)
1000	99.9	0.199 (0.008)	0.401 (0.006)	0.599 (0.005)	0.800 (0.008)
2000	100	0.200 (0.003)	0.401 (0.003)	0.599 (0.002)	0.800 (0.003)

3. Simulation

Likelihood-based approach

SB1: ($\tau = 0.5$)

T	pce	τ
100	91.3	0.499 (0.041)
200	93.0	0.500 (0.010)
400	94.5	0.500 (0.005)
800	94.7	0.500 (0.003)

SB2: ($\tau_1 = 0.33$, $\tau_2 = 0.67$)

T	pce	τ_1	τ_2
150	94.0	0.327 (0.005)	0.667 (0.004)
300	95.0	0.330 (0.002)	0.670 (0.002)
600	96.1	0.330 (0.001)	0.670 (0.001)
1200	95.5	0.330 (0.001)	0.670 (0.001)

SB4: ($\tau_1 = 0.2$, $\tau_2 = 0.4$, $\tau_3 = 0.6$, $\tau_4 = 0.8$)

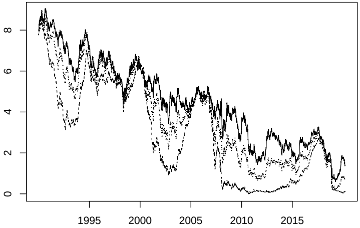
T	pce	τ_1	τ_2	τ_3	τ_4
250	100	0.200 (0.004)	0.400 (0.004)	0.600 (0.004)	0.800 (0.004)
500	96.7	0.200 (0.002)	0.400 (0.001)	0.600 (0.001)	0.800 (0.001)
1000	95.6	0.200 (0.001)	0.400 (0.001)	0.600 (0.001)	0.800 (0.001)
2000	95.1	0.200 (0.001)	0.400 (0.000)	0.600 (0.000)	0.800 (0.000)

4. Empirical application

- Detecting structural breaks in a simple US term structure model.
  EHT Literature
- Using daily data from January 1990 to July 2021 on fitted yields on zero coupon US bonds with 10-year ($r_{10y,t}$), 5-year ($r_{5y,t}$), and 1-year ($r_{1y,t}$) maturity.

Estimating the term structure model,

$$\begin{aligned} r_{10y,t} &= \mu_1 + \beta_1 r_{1y,t} + u_{1,t} \\ r_{5y,t} &= \mu_2 + \beta_2 r_{1y,t} + u_{2,t}. \end{aligned} \tag{11}$$



4. Empirical application

- We estimate the model with a dynamic OLS specification adding two leads and lags of $\Delta r_{1y,t}$.
- The coefficient estimates for the full sample without accounting for any structural breaks are $\hat{\beta}_1 = 0.764(0.116)$ and $\hat{\beta}_2 = 0.894(0.078)$.
- The expectations hypothesis (EHT) is rejected at the 5% significance level for the 10-year maturity but not for the 5-year maturity.
- We pre-specify a large maximum number of breaks, $M = 40$, and maintain a minimum break distance of two month (50 daily observations).
- The two-step estimator detects four structural breaks (November 1994, March 2003, August 2010, and March 2015).

Federal funds rate

4. Empirical application

Regimes:	$r_{10y,t}$: $\hat{\mu}_1$	$\hat{\beta}_1$	$r_{5y,t}$: $\hat{\mu}_2$	$\hat{\beta}_2$
1990 m01 - 1994 m11	5.013 (0.261)	0.464 (0.082)	3.141 (0.154)	0.678 (0.055)
1994 m11 - 2003 m04	4.000 (0.211)	0.392 (0.125)	2.297 (0.163)	0.653 (0.093)
2003 m04 - 2010 m08	3.569 (0.329)	0.267 (0.109)	2.020 (0.220)	0.556 (0.077)
2010 m08 - 2015 m03	2.311 (0.259)	-0.158 (0.812)	1.103 (0.162)	0.736 (0.532)
2015 m03 - 2021 m07	1.220 (0.235)	0.669 (0.168)	0.525 (0.191)	0.842 (0.131)

- The pairwise cointegrating vectors for each regime are substantially different from $(1, -1)$. Null hypothesis is not rejected for the two most recent regimes.
- The fourth regime from August 2010 to March 2015 is characterized by very unusual coefficients. It is also associated with an unusually steep yield curve (zero target rate).
- In summary, accounting for multiple structural breaks in the term structure model reveals some important differences for subsamples of the data and leads to a rejection of the EHT for most of the sampling period.

5. Conclusion

- Proposed a computationally efficient alternative to the existing likelihood-based approach solving the change-point problem in multivariate systems with a mix of integrated and stationary regressors.
- The two-step estimator (group LARS + BEA) is much faster but less precise than the likelihood-based approach.
- Future research: structural breaks in VECMs (involves a mix of integrated and stationary variables).

List of References I

Bulkley, G. and Giordani, P.: 2011, Structural breaks, parameter uncertainty, and term structure puzzles, *Journal of Financial Economics* **102**(1), 222–232.

Campbell, J. Y. and Shiller, R. J.: 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* **58**(3), 495–514.

Chan, N. H., Yau, C. Y. and Zhang, R.-M.: 2014, Group LASSO for Structural Break Time Series, *Journal of the American Statistical Association* **109**(506), 590–599.

Eo, Y. and Morley, J.: 2015, Likelihood-ratio-based confidence sets for the timing of structural breaks, *Quantitative Economics* **6**(2), 463–497.

Froot, K. A.: 1989, New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates, *The Journal of Finance* **44**(2), 283–305.

Gao, W., Yang, H. and Yang, L.: 2020, Change points detection and parameter estimation for multivariate time series, *Soft Computing* **24**, 6395–6407.

Hansen, P. R.: 2003, Structural changes in the cointegrated vector autoregressive model, *Journal of Econometrics* **114**(2), 261–295.

Lanne, M.: 1999, Near Unit Roots and the Predictive Power of Yield Spreads for Changes in Long-Term Interest Rates, *The Review of Economics and Statistics* **81**(3), 393–398.

Li, Y. and Perron, P.: 2017, Inference on locally ordered breaks in multiple regressions, *Econometric Reviews* **36**(1-3), 289–353.

List of References II

Liu, J., Wu, S. and Zidek, J. V.: 1997, On Segmented Multivariate Regression, *Statistica Sinica* **7**, 497–525.

Oka, T. and Perron, P.: 2018, Testing for common breaks in a multiple equations system, *Journal of Econometrics* **204**(1), 66–85.

Qu, Z. and Perron, P.: 2007, Estimating and testing structural change in multivariate regressions, *Econometrica* **75**(2), 459–502.

Safikhani, A. and Shojaie, A.: 2022, Joint Structural Break Detection and Parameter Estimation in High-Dimensional Nonstationary VAR Models, *Journal of the American Statistical Association* **117**(537), 251–264.

Saikkonen, P.: 1991, Asymptotically efficient estimation of cointegration regressions, *Econometric Theory* **7**(1), 1–21.

Sarno, L., Thornton, D. L. and Valente, G.: 2007, The empirical failure of the expectations hypothesis of the term structure of bond yields, *Journal of Financial and Quantitative Analysis* **42**(1), 81–100.

Shiller, R. J.: 1979, The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure, *Journal of Political Economy* **87**(6), 1190–1219.

Shiller, R. J., Campbell, J. Y. and Schoenholtz, K. L.: 1983, Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates, *Brookings Papers on Economic Activity* (1), 173–222.

List of References III

Thornton, D. L.: 2018, Greenspan's Conundrum and the Fed's Ability to Affect Long-Term Yields, *Journal of Money, Credit and Banking* **50**(2-3), 513–543.

Tillmann, P.: 2007, Inflation regimes in the US term structure of interest rates, *Economic Modelling* **24**(2), 203–223.

Yao, Y. C.: 1988, Estimating the number of change-points via Schwarz' criterion, *Statistics and Probability Letters* **6**(3), 181–189.

Yuan, M. and Lin, Y.: 2006, Model selection and estimation in regression with grouped variables, *Journal of the Royal Statistical Society. Series B: Statistical Methodology* **68**(1), 49–67.

Example: bivariate linear regression

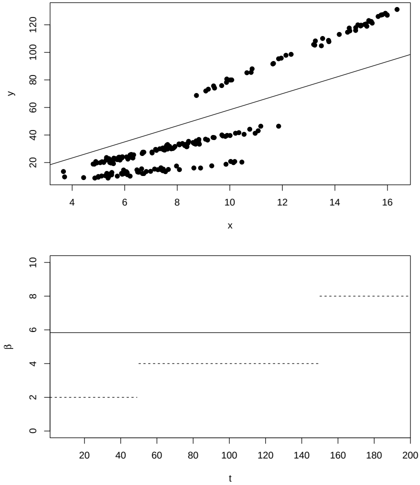
Linear regression:

$$y_t = \beta x_t + u_t, \qquad t = 1, \ldots, 200.$$

True breaks located at $t = (50, 150)$ with $\beta_1 = 2$, $\beta_2 = 4$, $\beta_3 = 8$.

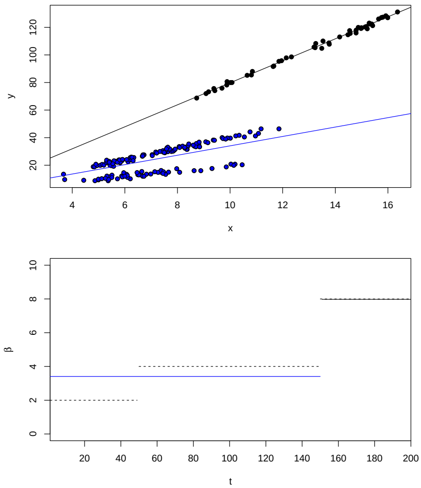
1. No breaks: OLS regression
2. One break: OLS + grid search optimization
3. Two breaks: OLS + grid search optimization

No breaks:



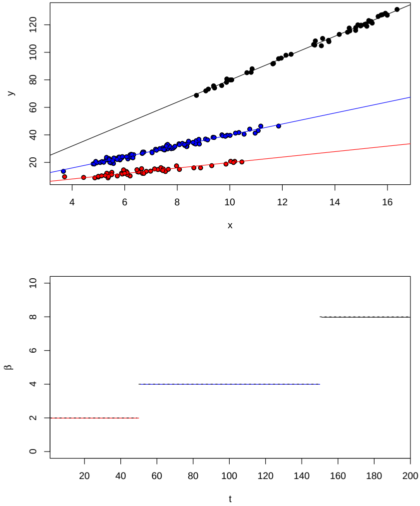
y
20
40
60
80
100
120
x
4
6
8
10
12
14
16
β
0
2
4
6
8
10
t
20
40
60
80
100
120
140
160
180
200

One break:



y
120
100
80
60
40
20
4
6
8
10
12
14
16
x
β
10
8
6
4
2
0
20
40
60
80
100
120
140
160
180
200
t

Two breaks:



back

Background

- A cointegrating regression is defined by

$$y_t = \beta x_t + \mu + u_t,$$

where x_t is a scalar, nonstationary, unit root process ($\Delta x_t = \xi_t$) and ξ_t, u_t are stationary innovation terms.

- Implies a long-run relationship between x_t and y_t (e.g. demand for money, term structure of interest rates, law of one price).
- The OLS estimator $\hat{\beta}$ is superconsistent (rate T instead of $\sqrt{T}$).
- $\hat{\beta}$ has a mixed normal limiting distribution. The serial correlation between ξ_t and u_t creates a second order bias $\rightarrow$ add leads and lags of Δx_t (Saikkonen, 1991),

$$y_t = \beta x_t + \mu + \sum_{j=-K}^{K} \gamma_j \Delta x_{t-j} + \tilde{u}_t.$$

back

Likelihood-based approach

- QMLE: specify Gaussian error terms although it is known that this assumption is (almost certainly) violated.
- The likelihood function under Gaussianity is largely determined by the fit.
- Reduces dimensionality by mapping many unkown coefficients to one value of the likelihood function.
- Uses likelihood ratio tests to determine the number of breaks.
- Dynamic programming to reduce the number of possible breakpoint candidates.

back

Expectations hypothesis of the term structure

Present value model in the notation of Campbell and Shiller (1991):

$$Y_t = \theta(1-\delta)\sum_{i=0}^{\infty}\delta^i E_t y_{t+1} + c,$$

- Y_t is the long-term yield,
- y_t is the one-period yield,
- θ is the coefficient of proportionality, $\theta = 1$ in case of bonds (see Shiller, 1979; Shiller et al., 1983),
- δ is the discount factor ($0 < \delta < 1$) and the constant c is the liquidity premium.

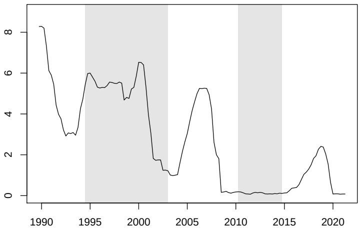
back

EHT literature

- Early studies report that the expectations hypothesis fails in empirical practice (Froot, 1989; Campbell and Shiller, 1991).
- Structural breaks are named as one of the important reasons for this failure (Lanne, 1999; Hansen, 2003; Sarno et al., 2007; Bulkley and Giordani, 2011).
- Several studies investigate whether regime shifts in the term structure of interest rates are related to changes in monetary policy (Tillmann, 2007; Thornton, 2018).

back

Estimated breakpoints and the EFFR



- BP1 (1994m11): increasing effective federal funds rate (EFFR).
- BP2 (2003m04): recession and decreasing EFFR.
- BP3 (2010m08): zero target rate as a reaction to the GFC.
- BP4 (2015m03): increasing EFFR.

back