Detecting Multiple Structural Breaks in Systems of Linear Regression Equations with Integrated and Stationary Regressors – Supplementary Material A

Karsten Schweikert*

University of Hohenheim

[Latest update: August 10, 2023]

1 Group LARS algorithm

We define some notation used in the exposition of the algorithm. Since our system is vectorized and the columns of $\boldsymbol{Z}$ have a specific structure in the change-point setting, we do not need to extend the correlation criterion as in Similä and Tikka (2006) to account for multiple responses. A simple re-partitioning before the most correlated set is computed allows us to use a modified version of the algorithm proposed by Chan et al. (2014) which itself is a specific adaptation of the group LARS algorithm outlined in Yuan and Lin (2006) to the univariate change-point setting.

We define the $Tq \times d$ matrix $\bar{Z} = I \otimes Z$, where the columns of Z contain the identical regressors for all responses. For $j = 1, \dots, Tq$, we define the d vector

$$\boldsymbol{B}_j(\nu) = \sum_{l=j}^{T} \bar{Z}_l' \nu_l.$$

Moreover, we define the $Tq \times d$ matrix $\boldsymbol{B}(\nu) = (\boldsymbol{B}_1'(\nu), \dots \boldsymbol{B}_{Tq}'(\nu))'$ which has q blocks of dimension $T \times d$. Now, we define the $T \times qd$ matrix $\boldsymbol{B}^*(\nu)$ re-partitioning $\boldsymbol{B}(\nu)$ so that the q blocks are concatenated horizontally. $\boldsymbol{B}_j^*(\nu)$ denotes the j-th row of $\boldsymbol{B}^*(\nu)$. The matrix $\boldsymbol{Z}_{\mathcal{A}}$ consists of all columns of $\boldsymbol{Z}$ that belong to the change-points contained in $\mathcal{A}$. The implementation of the modified group LARS algorithm on multiple change-points estimation is given below:

*Address: University of Hohenheim, Core Facility Hohenheim & Institute of Economics, Schloss Hohenheim 1 C, 70593 Stuttgart, Germany, e-mail: *karsten.schweikert@uni-hohenheim.de*

1. Initialization: specify K, the maximum number of change-points, and Δ, the minimum distance between change-points. Set $\mu^{[0]} = 0$, $k = 1$, $\nu^{[0]} = \boldsymbol{Y}$, $\mathcal{A}_0 = \{\emptyset\}$, and $\mathcal{T} = \{1, \ldots, T\}$.

2. Compute the current "most correlated set"

$$\mathcal{A}_k = \arg\max_{j \in \mathcal{T}} \|\boldsymbol{B}_j^*(\boldsymbol{\nu}^{[k-1]})\|_2.$$

3. Descent direction computation

$$\gamma_{\mathcal{A}_k} = (\boldsymbol{Z}'_{\mathcal{A}_k} \boldsymbol{Z}_{\mathcal{A}_k})^{-1} \boldsymbol{Z}'_{\mathcal{A}_k} \nu^{[k-1]}.$$

4. Descent step search: For $j \in \mathcal{T} \setminus \mathcal{A}_k$ define

$$\begin{aligned} a_j &= \|\boldsymbol{B}_j(\nu^{[k-1]})\|^2, & b_j &= \boldsymbol{B}'_j(\boldsymbol{Z}_{\mathcal{A}_k}\gamma_{\mathcal{A}_k})\boldsymbol{B}_j(\nu^{[k-1]}), \\ c_j &= \|\boldsymbol{B}_j(\boldsymbol{Z}_{\mathcal{A}_k}\gamma_{\mathcal{A}_k})\|^2, & d_j &= \max_{j \in \mathcal{T} \setminus \mathcal{A}_k} a_j. \end{aligned}$$

   Set $\alpha = \min_{j \in \mathcal{T} \setminus \mathcal{A}_k} a_j \equiv \alpha_{j^*}$, where

$$\begin{aligned} \alpha_j^+ &= \frac{(b_j - d_j) + \sqrt{(b_j - d_j)^2 - (a_j - d_j)(c_j - d_j)}}{c_j - d_j}, \\ \alpha_j^- &= \frac{(b_j - d_j) - \sqrt{(b_j - d_j)^2 - (a_j - d_j)(c_j - d_j)}}{c_j - d_j}, \end{aligned}$$

   and

$$\alpha_j = \begin{cases} \alpha_j^+ & \text{if } \alpha_j^+ \in [0, 1], \\ \alpha_j^- & \text{if } \alpha_j^- \in [0, 1]. \end{cases}$$

5. If $\alpha \neq 1$ or $k < K$, update $\mathcal{A}_{k+1} = \mathcal{A}_k \cup \{j^*\}$, $\mu^{[k]} = \mu^{[k-1]} + \alpha \boldsymbol{Z}_{\mathcal{A}_k}\gamma_{\mathcal{A}_k}$ and $\nu^{[k]} = Y - \mu^{[k]}$. Set $k = k + 1$ and go back to step 3. Otherwise, return $\mathcal{A}_k$ as the estimated change-points.

2 Backward elimination algorithm

The Backward elimination algorithm (BEA) successively eliminates breakpoints until no improvement in terms of the chosen criterion can be reached. For this purpose, we define

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

where $S_T(t_1, \ldots, t_m)$ is the least squares objective function for the pre-selected set of break-points and ω_T is the penalty function. The implementation of the BEA is given below:

1. Set $K = |\mathcal{A}_T|$, $\boldsymbol{t}_K = (t_{K,1}, \ldots, t_{K,K}) = \mathcal{A}_T$ and $V_K^* = IC(K, \mathcal{A}_T)$.
2. For $i = 1, \ldots, K$, compute $V_{K,i} = IC(K-1, \boldsymbol{t}_K \setminus \{t_{K,i}\})$. Set $V_{K-1}^* = \min_i V_{K,i}$.
3.
   - If $V_{K-1}^* > V_K^*$, then the estimated changepoints are $\mathcal{A}_T^* = \boldsymbol{t}_K$.
   - If $V_{K-1}^* \geq V_K^*$ and $K = 1$, then $\mathcal{A}_T^* = \emptyset$
   - If $V_{K-1}^* \geq V_K^*$ and $K > 1$, then set $j = \arg\min_i V_{K,i}$, $\boldsymbol{t}_{K-1} = \boldsymbol{t}_K \setminus \{t_{K-1,j}\})$ and $K = K - 1$. Go to step 2.

3 Additional simulation results

Table S1: Estimation of (multiple) structural breaks in the full model ($c = 0.5$)

	Panel A: Group LASSO with BEA				
	SB1: ($\tau = 0.5$)				
T	*pce*	τ			
100	67.9	0.502 (0.023)			
200	99.4	0.500 (0.012)			
400	99.9	0.500 (0.008)			
	SB2: ($\tau_1 = 0.33$, $\tau_2 = 0.67$)				
T	*pce*	τ_1	τ_2		
150	79.6	0.338 (0.034)	0.661 (0.026)		
300	97.2	0.335 (0.019)	0.666 (0.016)		
600	99.9	0.332 (0.010)	0.667 (0.008)		
	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	64.7	0.213 (0.034)	0.407 (0.031)	0.597 (0.030)	0.792 (0.028)
500	88.2	0.201 (0.016)	0.403 (0.012)	0.598 (0.010)	0.801 (0.014)
1000	99.7	0.200 (0.008)	0.401 (0.007)	0.598 (0.005)	0.800 (0.007)
	Panel B: Likelihood-based approach				
	SB1: ($\tau = 0.5$)				
T	*pce*	τ			
100	90.9	0.500 (0.030)			
200	93.2	0.500 (0.010)			
400	95.7	0.500 (0.005)			
	SB2: ($\tau_1 = 0.33$, $\tau_2 = 0.67$)				
T	*pce*	τ_1	τ_2		
150	94.1	0.326 (0.023)	0.667 (0.016)		
300	93.4	0.330 (0.009)	0.670 (0.007)		
600	95.8	0.330 (0.004)	0.670 (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	94.9	0.200 (0.012)	0.401 (0.012)	0.600 (0.011)	0.800 (0.009)
500	100	0.200 (0.006)	0.400 (0.005)	0.600 (0.004)	0.800 (0.004)
1000	96.7	0.200 (0.003)	0.400 (0.002)	0.600 (0.002)	0.800 (0.002)

Note: We use 1,000 replications of the data-generating process given in Equation (10) of the main text with $c = 0.5$. The variance of the error terms is $\sigma_\xi^2 = \sigma_e^2 = \sigma_u^2 = 1$. The first subpanel reports the results for one active breakpoint at $\tau = 0.5$, the second subpanel considers two active breakpoints at $\tau_1 = 0.33$ and $\tau_2 = 0.67$ and the third subpanel has four active breakpoints at $\tau_1 = 0.2$, $\tau_2 = 0.4$, $\tau_3 = 0.6$, and $\tau_4 = 0.8$. Standard deviations are given in parentheses. We conduct the $\sup(l + 1|l)$ test at the 5% level to determine the number of breaks.

Table S2: Estimation of (multiple) structural breaks in the full model ($c = 1.5$)

	Panel A: Group LASSO with BEA				
	SB1: ($\tau = 0.5$)				
T	*pce*	τ			
100	99.9	0.501 (0.010)			
200	99.9	0.500 (0.004)			
400	100	0.500 (0.002)			
	SB2: ($\tau_1 = 0.33$, $\tau_2 = 0.67$)				
T	*pce*	τ_1	τ_2		
150	93.7	0.338 (0.030)	0.660 (0.024)		
300	97.9	0.332 (0.016)	0.667 (0.014)		
600	99.9	0.332 (0.009)	0.668 (0.007)		
	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 (0.031)	0.404 (0.020)	0.597 (0.017)	0.788 (0.028)
500	98.1	0.203 (0.017)	0.402 (0.012)	0.598 (0.009)	0.803 (0.012)
1000	99.8	0.199 (0.008)	0.401 (0.005)	0.599 (0.005)	0.800 (0.008)
	Panel B: Likelihood-based approach				
	SB1: ($\tau = 0.5$)				
T	*pce*	τ			
100	90.0	0.500 (0.003)			
200	93.0	0.500 (0.002)			
400	95.7	0.500 (0.001)			
	SB2: ($\tau_1 = 0.33$, $\tau_2 = 0.67$)				
T	*pce*	τ_1	τ_2		
150	94.0	0.327 (0.003)	0.667 (0.002)		
300	92.9	0.331 (0.001)	0.670 (0.001)		
600	95.8	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	99.9	0.200 (0.008)	0.400 (0.013)	0.601 (0.032)	0.801 (0.038)
500	100	0.200 (0.001)	0.400 (0.001)	0.600 (0.001)	0.800 (0.001)
1000	97.8	0.200 (0.001)	0.400 (0.001)	0.600 (0.001)	0.800 (0.001)

Note: We use 1,000 replications of the data-generating process given in Equation (10) of the main text with $c = 1.5$. The variance of the error terms is $\sigma_\xi^2 = \sigma_e^2 = \sigma_u^2 = 1$. The first subpanel reports the results for one active breakpoint at $\tau = 0.5$, the second subpanel considers two active breakpoints at $\tau_1 = 0.33$ and $\tau_2 = 0.67$ and the third subpanel has four active breakpoints at $\tau_1 = 0.2$, $\tau_2 = 0.4$, $\tau_3 = 0.6$, and $\tau_4 = 0.8$. Standard deviations are given in parentheses. We conduct the $\sup(l+1|l)$ test at the 5% level to determine the number of breaks.

Table S3: Estimation of (multiple) structural breaks in the full model using the group LASSO with BEA ($c = 0.5$). Correlated errors.

T	pce	τ			
Panel A: Cross-correlated errors ($\rho = 0.95$)					
SB1: ($\tau = 0.5$)					
T	pce	τ			
100	90.6	0.500 (0.017)			
200	96.6	0.500 (0.009)			
400	98.9	0.500 (0.006)			
SB2: ($\tau_1 = 0.33$, $\tau_2 = 0.67$)					
T	pce	τ_1	τ_2		
150	91.2	0.336 (0.032)	0.661 (0.027)		
300	94.4	0.332 (0.018)	0.668 (0.014)		
600	98.5	0.331 (0.009)	0.669 (0.008)		
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	78.6	0.212 (0.030)	0.403 (0.024)	0.598 (0.020)	0.792 (0.026)
500	95.6	0.203 (0.017)	0.401 (0.013)	0.598 (0.010)	0.801 (0.014)
1000	98.9	0.200 (0.008)	0.400 (0.006)	0.599 (0.005)	0.800 (0.007)
Panel B: Cross-correlated ($\rho = 0.95$) and serially correlated ($\phi = 0.8$) errors					
SB1: ($\tau = 0.5$)					
T	pce	τ			
100	92.9	0.502 (0.067)			
200	99.5	0.501 (0.029)			
400	100	0.501 (0.014)			
SB2: ($\tau_1 = 0.33$, $\tau_2 = 0.67$)					
T	pce	τ_1	τ_2		
150	89.0	0.333 (0.042)	0.665 (0.037)		
300	98.7	0.330 (0.023)	0.670 (0.021)		
600	100	0.331 (0.011)	0.669 (0.010)		
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	73.4	0.212 (0.029)	0.402 (0.024)	0.599 (0.024)	0.795 (0.027)
500	95.1	0.202 (0.016)	0.400 (0.014)	0.599 (0.012)	0.800 (0.014)
1000	100	0.200 (0.008)	0.400 (0.007)	0.600 (0.006)	0.800 (0.007)

Note: We use 1,000 replications of the data-generating process given in Equation (10) of the main text with $c = 0.5$. The variance of the error terms is $\sigma_\xi^2 = \sigma_e^2 = \sigma_u^2 = 1$. In the first panel, we set the cross-correlation coeffcient to $\rho = 0.95$ and in the second panel, we additionally use AR(1) processes with autoregressive coefficient $\phi = 0.8$ to generate the error terms. The first subpanel reports the results for one active breakpoint at $\tau = 0.5$, the second subpanel considers two active breakpoints at $\tau_1 = 0.33$ and $\tau_2 = 0.67$ and the third subpanel has four active breakpoints at $\tau_1 = 0.2$, $\tau_2 = 0.4$, $\tau_3 = 0.6$, and $\tau_4 = 0.8$. Standard deviations are given in parentheses.

Table S4: Estimation of (multiple) partial structural breaks in the full model ($c = 0.5$)

	Group LASSO with BEA				
	SB1: ($\tau = 0.5$)				
T	*pce*	τ			
100	69.2	0.502 (0.022)			
200	99.0	0.500 (0.011)			
400	100	0.500 (0.005)			
	SB2: ($\tau_1 = 0.33$, $\tau_2 = 0.67$)				
T	*pce*	τ_1	τ_2		
150	79.0	0.339 (0.034)	0.661 (0.028)		
300	95.4	0.333 (0.018)	0.666 (0.017)		
600	99.5	0.332 (0.010)	0.668 (0.008)		
	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	77.6	0.218 (0.032)	0.406 (0.024)	0.597 (0.024)	0.791 (0.030)
500	94.4	0.205 (0.020)	0.403 (0.014)	0.598 (0.013)	0.801 (0.015)
1000	98.5	0.200 (0.008)	0.401 (0.006)	0.599 (0.006)	0.800 (0.008)

Note: We use 1,000 replications of the data-generating process given in Equation (10) of the main text with $c = 0.5$ but only the coefficients of the first equation change. Those changes are adjusted to ensure that the break magnitude is identical to the common break specification used to create Table Table S1. The variance of the error terms is $\sigma_\xi^2 = \sigma_e^2 = \sigma_u^2 = 1$. The first panel reports the results for one active breakpoint at $\tau = 0.5$, the second panel considers two active breakpoints at $\tau_1 = 0.33$ and $\tau_2 = 0.67$ and the third panel has four active breakpoints at $\tau_1 = 0.2$, $\tau_2 = 0.4$, $\tau_3 = 0.6$, and $\tau_4 = 0.8$. Standard deviations are given in parentheses.

Table S5: Estimation of (multiple) structural breaks in the full model ($c = 0.5$) with endogenous regressors

	Group LASSO with BEA				
	SB1: ($\tau = 0.5$)				
T	*pce*	τ			
100	92.7	0.501 (0.026)			
200	94.4	0.500 (0.012)			
400	96.8	0.500 (0.007)			
	SB2: ($\tau_1 = 0.33$, $\tau_2 = 0.67$)				
T	*pce*	τ_1	τ_2		
150	86.3	0.336 (0.036)	0.660 (0.029)		
300	98.7	0.335 (0.022)	0.665 (0.017)		
600	100	0.332 (0.011)	0.668 (0.008)		
	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	68.8	0.215 (0.038)	0.407 (0.034)	0.597 (0.033)	0.793 (0.032)
500	91.9	0.201 (0.017)	0.403 (0.013)	0.597 (0.011)	0.801 (0.014)
1000	99.9	0.200 (0.009)	0.401 (0.007)	0.598 (0.006)	0.799 (0.008)

Note: We use 1,000 replications of the data-generating process given in Equation (10) of the main text with $c = 0.5$. The variance of the error terms is $\sigma_\xi^2 = \sigma_e^2 = \sigma_u^2 = 1$. The error terms are correlated with the innovations of the first (second) integrated regressor with coefficient 0.5 (0.25). The first panel reports the results for one active breakpoint at $\tau = 0.5$, the second panel considers two active breakpoints at $\tau_1 = 0.33$ and $\tau_2 = 0.67$ and the third panel has four active breakpoints at $\tau_1 = 0.2$, $\tau_2 = 0.4$, $\tau_3 = 0.6$, and $\tau_4 = 0.8$. Standard deviations are given in parentheses.

Table S6: Estimation of (multiple) structural breaks in the full model with shrinking regimes

	Panel A: Group LASSO with BEA		
	SB1: ($\tau = 1 - 25/T$)		
T	*pce*	τ	
100	100	0.747 (0.019)	
200	99.5	0.873 (0.010)	
400	99.7	0.936 (0.006)	
800	99.8	0.967 (0.004)	
	SB2: ($\tau_1 = 0.5$, $\tau_2 = 0.5 + 25/T$)		
T	*pce*	τ_1	τ_2
150	84.2	0.505 (0.019)	0.661 (0.022)
300	91.9	0.500 (0.017)	0.584 (0.016)
600	94.1	0.499 (0.009)	0.542 (0.008)
	Panel B: Likelihood-based approach		
	SB1: ($\tau = 1 - 25/T$)		
T	*pce*	τ	
100	96.7	0.750 (0.006)	
200	97.3	0.854 (0.004)	
400	98.5	0.851 (0.005)	
800	98.7	0.850 (0.003)	
	SB2: ($\tau_1 = 0.5$, $\tau_2 = 0.5 + 25/T$)		
T	*pce*	τ_1	τ_2
150	94.3	0.505 (0.019)	0.661 (0.022)
300	95.5	0.463 (0.034)	0.615 (0.034)
600	96.3	0.443 (0.056)	0.596 (0.057)

Note: We use 1,000 replications of the data-generating process given in Equation (10). The variance of the error terms is $\sigma_\xi^2 = \sigma_e^2 = \sigma_u^2 = 1$. The first subpanel reports the results for the two step estimator with one active breakpoint at $\tau = 1 - 25/T$, the second subpanel reports the results for the likelihood based approach with trimming parameter 0.15. Standard deviations are given in parentheses. We conduct the $\sup(l + 1|l)$ test at the 5% level to determine the number of breaks. Critical values are provided in Qu and Perron (2007).

References

Chan, N.H., Yau, C.Y., Zhang, R.M., 2014. Group LASSO for Structural Break Time Series. Journal of the American Statistical Association 109, 590–599. doi:10.1080/01621459.2013.866566.

Qu, Z., Perron, P., 2007. Estimating and testing structural change in multivariate regressions. Econometrica 75, 459–502. doi:10.1111/j.1468-0262.2006.00754.x.

Similä, T., Tikka, J., 2006. Common subset selection of inputs in multiresponse regression. Proceedings of the 2006 International Joint Conference on Neural Networks , 1908–1915.

Yuan, M., Lin, Y., 2006. Model selection and estimation in regression with grouped variables. Journal of the Royal Statistical Society. Series B: Statistical Methodology 68, 49–67. doi:10.1111/j.1467-9868.2005.00532.x.