Inference Related to Locally Ordered and Common Breaks in a Multivariate System with Joined Segmented Trends

Ye Li∗ Pierre Perron†
Boston University Boston University
April 3, 2013

Abstract

The issues addressed in this paper are related to testing for common breaks and constructing confidence intervals for locally ordered breaks in bivariate linear time trend regressions with changes in the slopes such that series are joined at the break dates. The common break test considered is a likelihood ratio type test. The null hypothesis is that one of the break dates from one series is common with one of the break dates from the other one, while the alternative hypothesis is that the breaks dates are not the same and need not be separated by a positive fraction of the sample size. In both cases, the estimation method is quasi-maximum likelihood. We provide results about the consistency, rate of convergence and asymptotic distribution of the test statistic and about the limit distribution of the estimates of the locally ordered break dates. Simulation results show that both the test and the coverage rate provided by the limit distribution have good finite sample properties.

JEL Classification Number: C22.
Keywords: Locally Ordered Breaks, Common Breaks, Multiple Regressions, Joined Segmented Trend.

∗Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (yechina@bu.edu).
†Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (perron@bu.edu).
1 Introduction

Issues related to structure breaks have received a lot of attention in the statistics and econometrics literature (see Perron, 2006, for a survey). Substantial advances have been made to cover general models in the context of estimating and testing structural breaks in both single and multiple equations systems. In the single equation case, Bai (1997) studies the least squares estimation of a single change point in regressions involving stationary and/or trending regressors. He derives the consistency, rate of convergence and the limiting distributions of change point estimates under general conditions on the regressors and the errors. Bai and Perron (1998) extend the testing and estimation analysis to the case of multiple structure changes, while Bai and Perron (2003) present an efficient algorithm to obtain the break date estimates as the global minimizers of the overall sum of squared residuals.

Much of the work in the literature concentrated on the case where the regressors and errors are stationary. Nevertheless, issues related to structure changes are also important in the context of trending regressors and non-stationary time series. Perron and Zhu (2005) considered a linear trend function subject to a one-time change in the parameters. They analyzed the consistency, rate of convergence and limiting distributions of the parameters with errors that can be stationary or have an autoregressive unit root. They considered three different models: a “joint broken trend”, a “local disjoint broken trend” and a “global disjoint broken trend”. It was that each case involves different asymptotic results, in particular pertaining to the rate of convergence and the asymptotic distribution of the estimates of break dates. The model we consider in this paper is the “joint broken trend” model, whereby the slope of the trend changes and the series is joined at the time of the break.

Advances have also been made for structural change problems in the context of a multiple equations system. Bai, Lumsdaine and Stock (1998) develop the methods to construct a confidence interval for the estimate of a single break date in multivariate system, assuming a priori the break date to be common across all equations. Bai (2000) analyses multiple structure changes in both the regression coefficients and the variance-covariance matrix in vector autoregressive models. Qu and Perron (2007) provide a comprehensive treatment of issues related to estimation, inference and computation with multiple structural changes that occur at unknown dates in linear multivariate regression models, including VAR, certain linear panel data models, and seemingly unrelated regression. They also introduce a novel structure labelled as “locally ordered breaks”. These occur when one has prior knowledge of which coefficient is subject to the first break and when the subsequent break date is local
in the sense that the distance is not a positive fraction of the sample size (and can even be a fixed small number), as usually assumed in the structural change literature. They provide appropriate methods for estimation, inference and testing procedure for such locally ordered breaks in the context of a bivariate system with stationary regressors. As for testing in a multivariate system, Oka and Perron (2011) provides a general framework to testing common breaks across or within equations in the multivariate system with stationary, trending and/or unit root regressors. The null hypothesis is that some subsets of regressors share some common breaks, with the break date separated by a positive fraction of the sample size, while the alternative hypothesis is that the break dates are not the same and need not to be asymptotically distinct. Li and Perron (2012), building on the work of Oka and Perron (2001), extend the analysis of locally ordered breaks to cover systems with stationary, trending and/or unit root regressors.

A problem with the analysis of Oka and Perron (2011) and Li and Perron (2012) is that, while trends are permitted, the trend function cannot be restricted to be joined at the time of the break. As evidenced by many series in macroeconomics, finance and even climate change (e.g., Estrada, Perron and Martinez, 2013). The latter case is indeed the motivation behind this paper as global and hemispheric temperatures as well as radiative forcings (e.g., greenhouse gases) are well approximated by a linear trend with a one-time change in slope near 1955 with the noise component being stationary. As shown in Perron and Zhu (2005), the limit results with joined segmented trends are very different from locally or globally disjoined trends. Hence, the need for a separate treatment. The aim of this paper is to provide the relevant results to test for common breaks in a bivariate system of trending series, as in Oka and Perron (2011), as well as to extend the results of Qu and Perron (2007) and Li and Perron (2012) for the analysis of locally ordered breaks in such systems.

The structure of this paper is as follows. Section 2 presents the model and results about the test for common breaks. Section 3 discusses inference about locally ordered break dates. Section 4 provides some simulation results to assess the adequacy of the limit distribution in providing useful approximations in finite sample. Section 5 offers brief concluding remarks and technical derivations are contained in an appendix.

2 Testing Common Breaks

We adopt a framework similar to that of Perron and Zhu (2005), extended to have multiple breaks and considering a bivariate system. Each variable is represented by a linear trend with multiple changes in slope such that the trend function is joined at each break date.
There are $m_1$ breaks in the trend of the first variable and $m_2$ breaks for the second variable. More specifically, the bivariate system is, for $t = 1, \ldots, T$, with $T$ the sample size:

$$
\begin{align*}
    y_{1t} &= \mu_1^0 + \beta_1^0 t + \sum_{j_1=1}^{m_1} \delta_{1j_1}^0 \mathbf{1}(t \geq K_{1j_1}^0)(t - K_{1j_1}^0) + u_{1t} \\
    y_{2t} &= \mu_2^0 + \beta_2^0 t + \sum_{j_2=1}^{m_2} \delta_{2j_2}^0 \mathbf{1}(t \geq K_{2j_2}^0)(t - K_{2j_2}^0) + u_{2t}
\end{align*}
$$

where $K_{1j_1}^0$ ($j_1 = 1, \ldots, m_1$) are the break dates for the changes (with magnitudes $\delta_{1j_1}^0$) in the trend of the first variable and $K_{2j_2}^0$ ($j_2 = 1, \ldots, m_2$) are break dates for the changes (with magnitudes $\delta_{2j_2}^0$) in the trend of the second variable. In matrix form, we have, with $y_t = (y_{1t}, y_{2t})'$ and $U_t = (u_{1t}, u_{2t})'$,

$$
y_t = X_t^0 \theta^0 + U_t
$$

with

$$
X_t^0 = \\begin{bmatrix}
1 & t & \mathbf{1}(t \geq K_{11}^0)(t - K_{11}^0) & \cdots & \mathbf{1}(t \geq K_{1m_1}^0)(t - K_{1m_1}^0) \\
0 & \cdots & & & 0 \\
0 & \cdots \\
1 & t & \mathbf{1}(t \geq K_{11}^0)(t - K_{11}^0) & \cdots & \mathbf{1}(t \geq K_{2m_2}^0)(t - K_{2m_2}^0)
\end{bmatrix}
$$

and

$$
\theta^0 = [\mu_1^0 \beta_1^0 \delta_{11}^0 \cdots \delta_{1m_1}^0 \mu_2^0 \beta_2^0 \delta_{21}^0 \cdots \delta_{2m_2}^0].
$$

It is assumed that $U_t$ has mean 0 and covariance matrix $\Sigma^0$.

For the bivariate system, it is assumed that there are $m$ breaks, denoted by $T_{1}^0, \ldots, T_{m}^0$. We are interested in testing whether a break date for one variable is common to one for the other variable. Hence, the null and alternative hypotheses are:

$$
\begin{align*}
    H_0 : \quad & K_{1l_1}^0 = K_{2l_2}^0 = T_{l}^0 \\
    H_1 : \quad & K_{1l_1}^0 \neq K_{2l_2}^0
\end{align*}
$$

for some $l_1$ and $l_2$. Note that under the null hypothesis the total number of breaks is $m = m_1 + m_2 - 1$ while under the alternative hypothesis it is $m = m_1 + m_2$. We consider testing only for a single common break dates. The analysis could be extended to cover the more general case of testing for multiple common breaks. We do not pursue this extension here as the case of more practical interest is that for which there is a single common break.

To make the notation clear consider the following simple example. There are 4 breaks in the first equation and 2 breaks in the second equation. The order of the break dates
are as follows: the first equation has the first break at $K_{11}^0$, the second break is common to both equations, so that $K_{12}^0 = K_{21}^0$. The next break is in the first equation at $K_{13}^0$ followed by a break in the second equation at $K_{22}^0$. The last break is in the first equation at $K_{14}^0$. Hence, there are 5 breaks in total given by $T_1^0 = K_{11}^0, T_2^0 = K_{12}^0 = K_{21}^0, T_3^0 = K_{13}^0, T_4^0 = K_{22}^0, T_5^0 = K_{14}^0$.

The estimation methods we considered is Quasi-Maximum Likelihood assuming serially uncorrelated Gaussian errors. We define the sets $K_1^0 = \{K_{11}^0, \ldots, K_{1m_1}^0\}$ and $K_2^0 = \{K_{21}^0, \ldots, K_{2m_2}^0\}$. Also $\tau^0 = \{T_1^0, \ldots, T_m^0\}$ is the set of true break dates for the bivariate system. The candidate break dates for the two equations are $K_1 = \{K_{11}, \ldots, K_{1m_1}\}$ and $K_2 = \{K_{21}, \ldots, K_{2m_2}\}$. Under $H_0$, the set of candidate break dates are $\tau = \{T_1, \ldots, T_{t_1}, \ldots, T_m\}$, while under $H_1$ it is $\tau^1 = \{T_1, \ldots, T_{t_1}, T_{2l_1}, \ldots, T_{m_1}\}$ (??).

Let $LR_T$ be the likelihood ratio under $H_1$, $LR_T$ be the likelihood ratio under $H_0$, $lr_T$ be the log-likelihood ratio under $H_1$ and $lr_T$ be the log-likelihood ratio under $H_0$. Also, $\hat{X}_t$ denotes the regressor matrix and $(\hat{\theta}, \hat{\Sigma})$ the estimates of the coefficients and the variance-covariance matrix under $H_0$. Under $H_1$ we use the notation $\hat{X}_t$ and $(\hat{\theta}, \hat{\Sigma})$. Then,

$$\hat{LR_T} = \frac{\prod_{t=1}^{T} f(Y_t|\hat{X}_t, \hat{\theta}, \hat{\Sigma})}{\prod_{t=1}^{T} f(Y_t|X_t^0, \theta^0, \Sigma^0)}$$

$$\tilde{LR_T} = \frac{\prod_{t=1}^{T} f(Y_t|\tilde{X}_t, \tilde{\theta}, \tilde{\Sigma})}{\prod_{t=1}^{T} f(Y_t|X_t^0, \theta^0, \Sigma^0)}$$

and

$$\hat{lr_T} = \sum_{t=1}^{T} \log f(Y_t|\hat{X}_t, \hat{\theta}, \hat{\Sigma}) - \sum_{t=1}^{T} \log f(Y_t|X_t^0, \theta^0, \Sigma^0)$$

$$\tilde{lr_T} = \sum_{t=1}^{T} \log f(Y_t|\tilde{X}_t, \tilde{\theta}, \tilde{\Sigma}) - \sum_{t=1}^{T} \log f(Y_t|X_t^0, \theta^0, \Sigma^0)$$

The test statistic is the likelihood ratio test that compares the values of the likelihood function under the null hypothesis of a common break date and the alternative hypothesis of distinct break dates. It is defined as:

$$CB_T = \max_{T_1, \ldots, T_{2l_1}, \ldots, T_m \in \tau^1} \hat{lr_T}(\hat{X}_t, \hat{\theta}, \hat{\Sigma}) - \max_{T_1, \ldots, T_{t_1}, \ldots, T_m \in \tau} \tilde{lr_T}(\tilde{X}_t, \tilde{\theta}, \tilde{\Sigma})$$

2.1 Theoretical Results

We now consider the limit distribution of the test statistic for common breaks. We first state the assumptions needed to obtain the required results.

**Assumption 1** $0 < \lambda_1^0 < \ldots < \lambda_m^0 < 1$, with $T_j^0 = [T\lambda_j^0]$. 
Assumption 2 $\delta^0_{1j_1} \neq 0$ and $\delta^0_{2j_2} \neq 0$ for $j_1 = 1, ..., m_1$ and $j_2 = 1, ..., m_2$.

Assumption 3 Let $\mathcal{F}_t = \sigma$-field $\{\cdots, u_{t-2}, u_{t-1}\}$. If $u_t$ is weakly stationary within each segment, then (a) $\{u_t, \mathcal{F}_t\}$ forms a strongly mixing (a-mixing) sequence with size $-4r/(r-2)$ for some $2 < r < 8$. (b) $E(u_t) = 0$ and $\sup_t ||u_t||_{2r+4} < M < \infty$ for some $\delta > 0$ and $M > 0$, (c) let $S_{k,j}(l) = \sum_{k=1}^{k+4} u_t, j = 1, ..., m + 1$, for each $e \in R^n$ of length 1, $\text{var}(<e, S_{k,j}(0)>) \geq v(k)$ for some function $v(k) \to \infty$ as $k \to \infty$ (with $<\cdot>$, the usual inner product). If $u_t$ is not weakly stationary within each segment, we assume that (a)-(c) holds, and in addition, that there exists a positive definite matrix $\Omega = [\omega_{i,s}]$ such that for any $i, s = 1, ..., p$, we have, uniformly in $l$, $|k^{-1}E((S_{k,j}(l)), (S_{k,j}(l))_s)| \leq C_2k^{-\psi}$, for some $C_2, \psi > 0$. It is also assumed that $\{u_t' - \Sigma^0_{(j)}\}$ satisfies the conditions stated in this assumption.

Assumptions 1 and 2 are standard and simply state that the break dates are asymptotically distinct (i.e., each regime increases proportionally with the sample size $T$) and the changes in the parameters are non-zero at the break dates. Assumption A3 determines the dependence structure of the processes $u_t$. In particular, they imply that $u_t$ are short memory processes having bounded fourth moments. The assumptions are imposed to obtain a functional central limit theorem, a generalized Hajek and Renyi (1955) type inequality and a strong law of large numbers that allow us to show the estimates of break dates are consistent and to derive the rate of convergence. The conditions are mild in the sense that they allow for substantial conditional heteroskedasticity and autocorrelation. They are the same as those in Oka and Perron (2011) and Li and Perron (2012), so that we can use some of their results for trending series. Note, in particular, that under assumption A3, we have for $\eta_t = (\Sigma^0)^{-1/2}u_t$, $T^{-1/2} \sum_{t=1}^{T_0} \eta_t \Rightarrow W(\lambda^0_j) = (W_1(\lambda^0_j), W_2(\lambda^0_j))'$ where $W_1(\cdot)$ and $W_2(\cdot)$ are independent Wiener processes and “$\Rightarrow$” denotes weak convergence under the Skorohod topology.

We start with some preliminary results about the rate of convergence of the various estimates, whose proofs follow the developments in Perron and Zhu (2005) and are, hence, omitted.

Proposition 1 a) (Rate of Convergence Under $H_1$) Under Assumptions 1-3 and assuming the break dates not to be common, the following quantities are all $O_p(1): \sqrt{T}(\hat{T}_i - T^0_i) (i = 1, ..., m), \sqrt{T}(\hat{\mu}_i - \mu^0_i) (i = 1, 2), T^{3/2}(\hat{\beta}_i - \beta^0_i) (i = 1, 2), T^{3/2}(\hat{\delta}_{1j_1} - \delta^0_{1j_1}) (j_1 = 1, ..., m_1)$,
Theorem 1

The limit distribution of the test statistic is stated in the following Theorem proved in the appendix.

Then, within the sets $C^{(1)}_M$ and $C^{(2)}_M$, the common break test statistic is:

$$CM_T = \max_{\tilde{T}_1, \ldots, \tilde{T}_m; \tilde{T}_{m-1}} \tilde{tr}_T^2(\tilde{X}_t, \theta^0, \Sigma^0) - \max_{\tilde{T}_1, \ldots, \tilde{T}_m; \tilde{T}_{m-1}} \tilde{tr}_T^2(\tilde{X}_t, \theta^0, \Sigma^0)$$

with

$$\tilde{tr}_T^2(\tilde{X}_t, \theta^0, \Sigma^0) = \frac{1}{2} \left[ \sum_{t=1}^T \theta^0(X_t^0 - \hat{X}_t)(\Sigma^0)^{-1}(X_t^0 - \hat{X}_t)' \theta^0 + 2 \sum_{t=1}^T \theta^0(X_t^0 - \hat{X}_t)(\Sigma^0)^{-1}U_t \right]$$

and

$$\tilde{tr}_T^2(\tilde{X}_t, \theta^0, \Sigma^0) = \frac{1}{2} \left[ \sum_{t=1}^T \theta^0(X_t^0 - \hat{X}_t)(\Sigma^0)^{-1}(X_t^0 - \hat{X}_t)' \theta^0 + 2 \sum_{t=1}^T \theta^0(X_t^0 - \hat{X}_t)(\Sigma^0)^{-1}U_t \right]$$

The limit distribution of the test statistic is stated in the following Theorem proved in the appendix.

**Theorem 1** Under the Assumptions 1-3, we have:

$$CB_T \Rightarrow \max_{s_1, s_2} \hat{H}(s_1^*, s_2^*) - \max_{s_1, s_2} \tilde{H}(s_1, s_2)$$

with

$$\hat{H}(s_1^*, s_2^*) = \max_{s_1, s_2} -\frac{1}{2} \left[ 2(\sum_{j=1}^{l} tr((\Sigma^0)^{-1/2} \hat{A}_1) + \sum_{j=1}^{m+1} tr((\Sigma^0)^{-1/2} \hat{A}_2)) + \sum_{j=1}^{l} tr((\Sigma^0)^{-1} \hat{B}_1) + \sum_{j=1}^{m+1} tr((\Sigma^0)^{-1} \hat{B}_2) \right]$$
where \( s_1^* = \{ T_1^{1/2}(\hat{K}_{11} - K_{11}^0), ..., T_1^{1/2}(\hat{K}_{1m_1} - K_{1m_1}^0) \}, s_2^* = \{ T_1^{1/2}(\hat{K}_{21} - K_{21}^0), ..., T_1^{1/2}(\hat{K}_{2m_2} - K_{2m_2}^0) \} \). Also,

\[
\bar{H}(s_1, s_2) = -\frac{1}{2} \left[ 2 \sum_{j=1}^{m+1} tr((\Sigma^0)^{-1/2} \bar{A} + \sum_{j=1}^{m+1} (\lambda_j^0 - \lambda_{j-1}^0) tr((\Sigma^0)^{-1} \bar{B})) \right]
\]

with \( s_1 = \{ T_1^{1/2}(\hat{K}_{11} - K_{11}^0), ..., T_1^{1/2}(\hat{K}_{1m_1} - K_{1m_1}^0) \}, s_2 = \{ T_1^{1/2}(\hat{K}_{21} - K_{21}^0), ..., T_1^{1/2}(\hat{K}_{2m_2} - K_{2m_2}^0) \}; K_{11} = K_{2l_1} \} \), and the elements of the 2 \times 2 matrices \( \bar{A}, \bar{B}, \bar{A}_1, \bar{A}_2, \bar{B}_1 \) and \( \bar{B}_2 \) are defined by:

\[
\bar{A}(1, 1) = (\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1})^2
\]
\[
\bar{A}(1, 2) = \bar{A}(2, 1) = (\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1}^0)(\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2}^0)
\]
\[
\bar{A}(2, 2) = (\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2})^2
\]

\[
\bar{B}(1, 1) = (\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1}^0)(W_1(\lambda_j^0) - W_1(\lambda_{j-1}^0))
\]
\[
\bar{B}(1, 2) = (\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1}^0)(W_1(\lambda_j^0) - W_1(\lambda_{j-1}^0))
\]
\[
\bar{B}(2, 1) = (\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2}^0)(W_2(\lambda_j^0) - W_2(\lambda_{j-1}^0))
\]
\[
\bar{B}(2, 2) = (\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2}^0)(W_2(\lambda_j^0) - W_2(\lambda_{j-1}^0))
\]

\[
\bar{A}_1(1, 1) = (\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1}^0)(W_1(\lambda_j^0) - W_1(\lambda_{j-1}^0))
\]
\[
\bar{A}_1(1, 2) = (\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2}^0)(W_1(\lambda_j^0) - W_1(\lambda_{j-1}^0))
\]
\[
\bar{A}_1(2, 1) = (\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1}^0)(W_2(\lambda_j^0) - W_2(\lambda_{j-1}^0))
\]
\[
\bar{A}_1(2, 2) = (\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2}^0)(W_2(\lambda_j^0) - W_2(\lambda_{j-1}^0))
\]

\[
\bar{A}_2(1, 1) = (\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1} + \delta_{si_1}^0 s_{i_1} + \sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1}^0)(W_1(\lambda_j^0) - W_1(\lambda_{j-1}^0))
\]
\[
\bar{A}_2(1, 2) = (\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2}^0)(W_1(\lambda_j^0) - W_1(\lambda_{j-1}^0))
\]
\[
\bar{A}_2(2, 1) = (\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1} + \delta_{si_1}^0 s_{i_1} + \sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1}^0)(W_2(\lambda_j^0) - W_2(\lambda_{j-1}^0))
\]
\[
\bar{A}_2(2, 2) = (\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2}^0)(W_2(\lambda_j^0) - W_2(\lambda_{j-1}^0))
\]

\[
\hat{B}_1(1, 1) = (\lambda_j^0 - \lambda_{j-1}^0)(\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1})^2
\]
\[
\hat{B}_1(1, 2) = (\lambda_j^0 - \lambda_{j-1}^0)(\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1}^0)(\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2}^0)
\]
\[
\hat{B}_1(2, 1) = (\lambda_j^0 - \lambda_{j-1}^0)(\sum_{i=1}^{j_1-1} \delta_{si_1}^0 s_{i_1}^0)(\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2}^0)
\]
\[
\hat{B}_1(2, 2) = (\lambda_j^0 - \lambda_{j-1}^0)(\sum_{i=1}^{j_2-1} \delta_{si_2}^0 s_{i_2}^0)^2
\]
\[
\hat{B}_2(1,1) = (\lambda_j^0 - \lambda_{j-1}^0)(\sum_{i_1=1}^{l_i-1} \delta_{i_1}^0 s_{i_1} + \delta_{i_1}^0 s_{i_1} + \sum_{i_1=l_1+1}^{j_i-1} \delta_{i_1}^0 s_{i_1}) \\
\hat{B}_2(1,2) = \hat{B}_2(2,1) \\
= (\lambda_j^0 - \lambda_{j-1}^0)(\sum_{i_1=1}^{l_i-1} \delta_{i_1}^0 s_{i_1} + \delta_{i_1}^0 s_{i_1} + \sum_{i_1=l_1+1}^{j_i-1} \delta_{i_1}^0 s_{i_1}) \\
\times (\sum_{i_2=1}^{l_2-1} \delta_{i_2}^0 s_{i_2} + \delta_{i_2}^0 s_{i_2} + \sum_{i_2=l_2+1}^{j_2-1} \delta_{i_2}^0 s_{i_2}) \\
\hat{B}_2(2,2) = (\lambda_j^0 - \lambda_{j-1}^0)(\sum_{i_2=1}^{l_2-1} \delta_{i_2}^0 s_{i_2} + \delta_{i_2}^0 s_{i_2} + \sum_{i_2=l_2+1}^{j_2-1} \delta_{i_2}^0 s_{i_2})^2
\]

Remark 1 Ye, give the results for the special case with only one break in each equation.

3 Inference about Locally Ordered Breaks

We now consider the problem of forming a confidence interval for a pair of locally ordered breaks for changes in the slope of a trend function joined at the time of the break. Again, we consider a bivariate system with a single break in each equation. We also omit other possible breaks that are not locally ordered since confidence intervals for these can be constructed in the usual manner (see, Qu and Perron, 2007, and Li and Perron, 2011). Accordingly, the model is:

\[
y_{1t} = \mu_1^0 + \beta_1^0 t + \delta_1^0 1(t \geq K_1^0) (t - K_1^0) + u_{1t} \\
y_{2t} = \mu_2^0 + \beta_2^0 t + \delta_2^0 1(t \geq K_2^0) (t - K_2^0) + u_{2t}
\]

Again, \(U_t = (u_{1t}, u_{2t})\) has mean 0, variance \(\Sigma^0\) and satisfies Assumption A3.

Definition 1 Locally Ordered Breaks (LOB): Let \(v_T\) be a sequence of positive numbers that satisfies \(v_T \rightarrow 0\) and \(T^{1/2} v_T / (\log^2 T) \rightarrow \infty\). \(K_1^0\) and \(K_2^0\) are said to be locally ordered if \(K_1^0 \leq K_2^0\), and \(v_T^2 (K_2^0 - K_1^0) \leq M_T\) with \(M_T \rightarrow 0\) as \(T \rightarrow \infty\).

Remark 2 The condition \(v_T^2 (K_2^0 - K_1^0) \leq M_T\) with \(M_T \rightarrow 0\) implies that \((K_2^0 - K_1^0) / T \rightarrow 0\). Hence, asymptotically the distance between the break dates becomes a negligible portion of the sample size. If \(\lim_{T \rightarrow \infty} (K_2^0 - K_1^0) / T > 0\), then the two break dates are asymptotically distinct and the usual asymptotic distribution theory applies.

The method of estimation considered is restricted quasi-maximum likelihood that assumes serially uncorrelated Gaussian errors. Conditional on a given partition of the sample \(K = (K_1, K_2)\), the Gaussian quasi-likelihood function is

\[
L_T(K, \beta, \Sigma) = \prod_{t=1}^{T} f(y_t | X_t; \theta, \Sigma)
\]
where
\[ f(y_t|X_t; \theta, \Sigma) = \frac{1}{(2\pi)^{1/2}|\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2}(y_t - X'_t\theta)\Sigma^{-1}(y_t - X'_t\theta) \right\} \]
and the quasi-likelihood ratio is
\[ LR_T(K, \beta, \Sigma) = \frac{\prod_{t=1}^{T} f(y_t|X_t; \theta, \Sigma)}{\prod_{t=1}^{T} f(y_t|X_t; \theta^0, \Sigma^0)}. \]

We wish to obtain the values of \((K_1, K_2, \theta, \Sigma)\), which maximizes \(LR_T\) subject to restrictions \(g(\theta^0, vec(\Sigma^0)) = 0\), if applicable. Let \(lr_T(\cdot)\) denote the log-likelihood ratio and \(rlr_T(\cdot)\) the restricted log-likelihood ratio, the objective function is then
\[ rlr_T(K, \beta, \Sigma) = lr_T(K, \theta, \Sigma) + \lambda g(\theta^0, vec(\Sigma^0)) \]
and the estimates are
\[ (\hat{K}, \hat{\theta}, \hat{\Sigma}) = \arg \max_{(K_1, K_2, \theta, \Sigma)} rlr_T(K, \theta, \Sigma). \]

The supremum with respect to \((K_1, K_2)\) is taken over a restricted set of partitions \(\mathbb{K}_\varepsilon\). For a small number \(\varepsilon > 0\), \(\mathbb{K}_\varepsilon = \{(K_1, K_2) : K_1 \geq [T\varepsilon], K_2 - K_1 \geq h, T - K_2 \geq [T\varepsilon]\}\), where \(h\) is at least as large as the number of parameters to be estimated in each regime. We do not impose the restriction that the number of observations in single regime increases as the sample size increases, as is common when dealing with non-local break dates.

3.1 Theoretical Results

We again start with some preliminary results about the rate of convergence of the estimates. Throughout, we assume that \(\delta^0_1\) and \(\delta^0_2\) are non-zero and that Assumption 3 holds.

**Proposition 2** Under the stated conditions, the following quantities are \(O_p(1)\):
\[ \sqrt{T}(\hat{K}_1 - K^0_1), \sqrt{T}(\hat{K}_2 - K^0_2), \sqrt{T}(\hat{\mu}_1 - \mu^0_1), \sqrt{T}(\hat{\mu}_2 - \mu^0_2), T^{3/2}(\hat{\beta}_1 - \beta^0_1), T^{3/2}(\hat{\beta}_2 - \beta^0_2), T^{3/2}(\hat{\delta}_1 - \delta^0_1), T^{3/2}(\hat{\delta}_2 - \delta^0_2). \]

The results follow directly from ???. Given the rates of convergence, we can analyze the likelihood function under the compact set
\[ C_M = \{\mu_i, \beta_i, \delta_i, K_i : \sqrt{T}(\mu_i - \mu^0_i) \leq M, T^{3/2}(\beta_i - \beta^0_i) \leq M, T^{3/2}(\delta_i - \delta^0_i) \leq M, \sqrt{T}(K_i - K^0_i) \leq M\} \]

The log-likelihood function is given by:
so that to derive the asymptotic distribution of the estimates of the break dates, we need to the true values; Case 1: we have the following six cases for the position of the estimates of the break dates relative to the set $K$

$$lr_T = -\frac{T}{2}[\log |\Sigma| - \log |\Sigma^0|]$$

$$-\frac{1}{2}[\sum_{t=1}^{T}[(Y_t - X_{t}^0\theta)^{\Sigma^{-1}}(Y_t - X_{t}^0\theta) - \sum_{t=1}^{T}(Y_t - X_{t}^0\theta)^{\Sigma^{-1}}(Y_t - X_{t}^0\theta)]]$$

$$-\frac{1}{2}[\sum_{t=1}^{T}\theta^{0\prime}(X_{t}^0 - X_t)(\Sigma^0)^{-1}(X_{t}^0 - X_t)^{\prime\theta^0} + 2\sum_{t=1}^{T}\theta^{0\prime}(X_{t}^0 - X_t)(\Sigma^0)^{-1}U_t] + o_p(1)$$

Then, within the set $C_M$, this log-likelihood function can be separated in two parts (except for terms that converge in probability to 0), namely,

$$lr_T^1 = -\frac{T}{2}[\log |\Sigma| - \log |\Sigma^0|]$$

$$-\frac{1}{2}[\sum_{t=1}^{T}[(Y_t - X_{t}^0\theta)^{\Sigma^{-1}}(Y_t - X_{t}^0\theta) - \sum_{t=1}^{T}(Y_t - X_{t}^0\theta)^{\Sigma^{-1}}(Y_t - X_{t}^0\theta)]]$$

$$lr_T^2 = -\frac{1}{2}\sum_{t=1}^{T}\theta^{0\prime}(X_{t}^0 - X_t)(\Sigma^0)^{-1}(X_{t}^0 - X_t)^{\prime\theta^0} + 2\sum_{t=1}^{T}\theta^{0\prime}(X_{t}^0 - X_t)(\Sigma^0)^{-1}U_t]$$

so that to derive the asymptotic distribution of the estimates of the break dates, we need only focus on $lr_T^2$. Given the rates of convergence and the definition of locally ordered breaks, we have the following six cases for the position of the estimates of the break dates relative to the true values; Case 1: $\hat{K}_1 \leq K_1^0 \leq \hat{K}_2 \leq K_2^0$, Case 2: $\hat{K}_1 \leq K_1^0 \leq K_2^0 \leq \hat{K}_2$, Case 3: $\hat{K}_1 \leq K_1^0 \leq \hat{K}_2 \leq K_2^0$, Case 4: $K_1^0 \leq \hat{K}_1 \leq K_2^0 \leq \hat{K}_2$, Case 5: $K_1^0 \leq \hat{K}_1 \leq \hat{K}_2 \leq K_2^0$, Case 6: $K_1^0 \leq \hat{K}_2 \leq \hat{K}_1 \leq \hat{K}_2$. It turns out, however, that the limit distribution of the estimates of the break dates is the same in all cases, as stated in the following theorem.

**Theorem 2** Under the stated conditions,

$$\left[\begin{array}{c} T^{1/2}(\hat{K}_1 - K_1^0) \\
T^{1/2}(\hat{K}_2 - K_2^0) \end{array}\right] \Rightarrow \arg \max_{s_1, s_2} H(s_1, s_2)$$

$$= -\frac{1}{2}\{tr((\Sigma^0)^{-1}\left[\begin{array}{cc} (1 - \lambda_2^0)\delta_1^0 s_1^2 & (1 - \lambda_2^0)\delta_1^0 s_1 s_2 \\
(1 - \lambda_2^0)\delta_1^0 s_1 s_2 & (1 - \lambda_2^0)\delta_2^0 s_2^2 \end{array}\right])

+2tr((\Sigma^0)^{-1/2}\left[\begin{array}{cc} \delta_1^0 s_1(W_1(1) - W_1(\lambda_2^0)) & \delta_2^0 s_2(W_1(1) - W_1(\lambda_2^0)) \\
\delta_1^0 s_1(W_2(1) - W_2(\lambda_2^0)) & \delta_2^0 s_2(W_2(1) - W_2(\lambda_2^0)) \end{array}\right])\}$$

10
Remark 3 In the case with $\Sigma^0 = I$, we have
\[
\begin{bmatrix}
T^{1/2}(K_1 - K^0_1) \\
T^{1/2}(K_2 - K^0_2)
\end{bmatrix} \Rightarrow \arg \max_{s_1, s_2} H(s_1, s_2)
\]
\[
= -(1/2)(1 - \lambda^0_2)(\delta^{02}_1 s_1^2 + \delta^{02}_2 s_2^2)
- (\delta^0_1 s_1 (W_1(1) - W_1(\lambda^0_2)) + \delta^0_2 s_2 (W_2(1) - W_2(\lambda^0_2)))
\]

The first order conditions yield,
\[
-(1 - \lambda^0_2)\delta^{02}_1 s_1 = \delta^0_1 (W_1(1) - W_1(\lambda^0_2))
\]
\[
-(1 - \lambda^0_2)\delta^{02}_2 s_2 = \delta^0_2 (W_2(1) - W_2(\lambda^0_2))
\]

So that,
\[
s_1 = -\frac{W_1(1) - W_1(\lambda^0_2)}{(1 - \lambda^0_2)\delta^0_1}
\]
\[
s_2 = -\frac{W_2(1) - W_2(\lambda^0_2)}{(1 - \lambda^0_2)\delta^0_2}
\]

These imply that
\[
\begin{bmatrix}
T^{1/2}(\hat{K}_1 - K^0_1) \\
T^{1/2}(\hat{K}_2 - K^0_2)
\end{bmatrix} \Rightarrow N\left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{(1 - \lambda^0_2)\delta^0_1} & 0 \\ 0 & \frac{1}{(1 - \lambda^0_2)\delta^0_2} \end{bmatrix}\right)
\]

which corresponds to the result obtained by Perron and Zhu (2005).

4 Simulation Result

In this section, we provide Monte Carlo simulation results to assess the adequacy of the asymptotic distribution in providing useful approximations to the finite sample distribution of the locally ordered break dates. We also we provide simulation results concerning the size and power of the common breaks test statistic. All results are based on 1,000 replications and the Wiener process are approximate by the partial sums of a sequence of i.i.d. $N(0, 1)$ random variable of length 1,000.

4.1 Simulation Result for Testing Common Breaks

We start by considering the exact size and power of the common breaks test. The Data Generating Process (DGP) is specified by:
\[
y_{1t} = \mu^0_1 + \beta^0_1 t + \delta^0_1 1(t \geq K^0_1)(t - K^0_1) + u_{1t}
\]
\[
y_{2t} = \mu^0_2 + \beta^0_2 t + \delta^0_2 1(t \geq K^0_2)(t - K^0_2) + u_{2t}
\]
The sample size is set to $T = 100$, and we use the following parameter values

$$
\begin{align*}
\mu_1^0 &= 1, \beta_1^0 = 0.5, \delta_1^0 = 0.2 \\
\mu_2^0 &= 0.8, \beta_2^0 = 0.5, \delta_2^0 = 0.3
\end{align*}
$$

with

$$U_t = (u_1, u_2)’ \sim i.i.d. N(0, I_2)$$

Under the null hypothesis, there is one common break at mid-sample, i.e., $K_1^0 = K_2^0 = 50$. The nominal and exact quantiles of the asymptotic distribution of the test are presented below. The results show a close correspondence so that the test has no size distortions.

<table>
<thead>
<tr>
<th></th>
<th>Asymptotic</th>
<th>Exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>95%</td>
<td>96%</td>
<td></td>
</tr>
<tr>
<td>90%</td>
<td>89.9%</td>
<td></td>
</tr>
<tr>
<td>85%</td>
<td>83%</td>
<td></td>
</tr>
<tr>
<td>80%</td>
<td>77.4%</td>
<td></td>
</tr>
<tr>
<td>75%</td>
<td>73.7%</td>
<td></td>
</tr>
<tr>
<td>70%</td>
<td>68.9%</td>
<td></td>
</tr>
</tbody>
</table>

To assess the power of the test, we set $K_1^0 = 50$ and vary the date of the break in the second equation with the following values: $K_2^0 = 40, 45, 60, 65, 70$. The power of the test is shown in the following figure:
This is not a power function. What is it? Power should equal to size at 50 and increase as we move away either side. What is the size of the test used??

4.2 Simulation Result for Locally Ordered Breaks

We now consider the adequacy of the limit distribution for the estimates of locally ordered break dates. The DGP used is:

\[ y_{1t} = \mu_1^0 + \beta_1^0 t + \delta_1^0 1(t \geq K_1^0)(t - K_1^0) + u_{1t} \]

\[ y_{2t} = \mu_2^0 + \beta_2^0 t + \delta_2^0 1(t \geq K_2^0)(t - K_2^0) + u_{2t} \]

The sample size is set to \( T = 100 \), and we use the following parameter values:

\[
\begin{align*}
\mu_1^0 &= 1, \beta_1^0 = 0.5, \delta_1^0 = 0.2 \\
\mu_2^0 &= 1, \beta_2^0 = 0.5, \delta_2^0 = 0.3 \\
K_1^0 &= 45, K_2^0 = 50
\end{align*}
\]

with

\[ U_t = (u_1, u_{2t})' \sim i.i.d. N(0, I_2) \]???

[Is that the case??? if so the two estimates are simply normally distributed and uncorrelated. Not very useful???]

The histograms of the bivariate asymptotic and finite sample distributions are presented below. One can see that the correspondence is quite close.
Though, the experiments are quite limited they indicate that the limit distributions derived can provide tests with reliable size in finite samples and decent power. Also, the limit distribution of the estimates of the locally ordered breaks provide a good approximation...
to the finite sample distribution. This shows the usefulness of the results derived.

5 Conclusions

In this paper, we considered the issues of testing for common breaks in a bivariate system described by a linear trend function with changes in slope such that the series are joined at the time of the breaks. We also considered the problem of performing joint inference about the estimates of the break dates in the same setup when the break dates are locally ordered. We provided the rate of convergence, the asymptotic distribution for the estimates of the locally ordered breaks and the asymptotic distribution of the common break test statistic. Limited simulation results showed that the theoretical results derived deliver good approximations in finite samples.
References


Appendix

Proof of Theorem 1 (asymptotic distribution of the common break test statistic). For \( t \in [T^0_{j-1} + 1, T^0_j] \), the regressor matrix in regime \( j \) is:

\[
X^0_{tj} = \begin{bmatrix}
1 & t - K^0_{11} & \cdots & t - K^0_{1j-1} & 0 & \cdots & 0 & 0 \\
& 1 & t - K^0_{21} & \cdots & t - K^0_{2j-1} & 0 & \cdots & 0
\end{bmatrix}
\]

so that the data for the first equation is in its \( j_1 \)-th regime and that for the second equation is in its \( j_2 \)-th regime. Under \( H_0 \), for the \( j \)-th regime with \( t \in [T_{j-1} + 1, T_j] \):

\[
\tilde{X}_{tj} = \begin{bmatrix}
1 & t & \cdots & t - \tilde{K}_{11} & \cdots & t - \tilde{K}_{1j-1} & 0 & \cdots & 0 & 0 \\
& 1 & t & \cdots & t - \tilde{K}_{21} & \cdots & t - \tilde{K}_{2j-1} & 0 & \cdots & 0
\end{bmatrix}
\]

so that

\[
X^0_{tj} - \tilde{X}_{tj} =
\begin{bmatrix}
0 & \tilde{K}^0_{11} - K^0_{11} & \cdots & \tilde{K}_{1j-1} - K^0_{1j-1} & 0 & \cdots & 0 & 0 \\
& 0 & \tilde{K}_{21} - K^0_{21} & \cdots & \tilde{K}_{2j-1} - K^0_{2j-1} & 0 & \cdots & 0
\end{bmatrix}
\]

For simplicity, assume that \( T_j < T^0_j \), then for \( t \in [T_j + 1, T^0_j] \), depending on which equation has a change at \( T^0_j \), the form of \( \tilde{X}_{tj+1} \) could be either: a) with the first equation having a break at \( T^0_j \), so that:

\[
\tilde{X}_{tj+1} = \begin{bmatrix}
1 & t - \tilde{K}_{11} & \cdots & t - \tilde{K}_{1j-1} & t - \tilde{K}_{1j} & \cdots & 0 & 0 \\
& 1 & t - \tilde{K}_{21} & \cdots & t - \tilde{K}_{2j-1} & 0 & \cdots & 0
\end{bmatrix}
\]

b) with the second equation having a break at \( T^0_j \), in which case

\[
\tilde{X}_{tj+1} = \begin{bmatrix}
1 & t - \tilde{K}_{11} & \cdots & t - \tilde{K}_{1j-1} & 0 & \cdots & 0 & 0 \\
& 1 & t - \tilde{K}_{21} & \cdots & t - \tilde{K}_{2j-1} & t - \tilde{K}_{2j} & \cdots & 0
\end{bmatrix}
\]

c) with both equations having a break at \( T^0_j \), in which case

\[
\tilde{X}_{tj+1} = \begin{bmatrix}
1 & t - \tilde{K}_{11} & \cdots & t - \tilde{K}_{1j} & \cdots & 0 & 0 & 0 \\
& 1 & t - \tilde{K}_{21} & \cdots & t - \tilde{K}_{2j} & \cdots & 0
\end{bmatrix}
\]
We first have

\[
\sum_{t=1}^{T} \theta'(X_t^0 - \tilde{X}_t) (\Sigma^0)^{-1} U_t
\]

\[
= \sum_{j=1}^{m+1} \sum_{t=T_{j-1}^0+1}^{T_j^0} \theta'(X_{tj}^0 - \tilde{X}_{tj}) (\Sigma^0)^{-1} U_t
\]

\[
+ \sum_{j=1}^{m} \sum_{t=T_{j-1}^0+1}^{T_j^0} \theta'(X_{tj}^0 - \tilde{X}_{t+1}) (\Sigma^0)^{-1} U_t
\]

\[
- \sum_{j=1}^{m} \sum_{t=T_{j-1}^0+1}^{T_j^0} \theta'(X_{tj}^0 - \tilde{X}_{tj}) (\Sigma^0)^{-1} U_t
\]

\[
= \sum_{j=1}^{m+1} \sum_{t=T_{j-1}^0+1}^{T_j^0} \theta'(X_{tj}^0 - \tilde{X}_{tj}) (\Sigma^0)^{-1} U_t + o_p(1)
\]

\[
= \sum_{j=1}^{m+1} \sum_{t=T_{j-1}^0+1}^{T_j^0} \left[ \sum_{i=1}^{j-1} \delta_{11i}^0 (K_{1i}^0 - K_{1i}^0) \sum_{i_2=1}^{j-1} \delta_{1i_2}^0 (K_{1i_2}^0 - K_{1i_2}^0) \right] (\Sigma^0)^{-1} U_t
\]

\[
= \sum_{j=1}^{m+1} tr((\Sigma^0)^{-1/2} \sum_{t=T_{j-1}^0+1}^{T_j^0} \eta_t \sum_{i=1}^{j-1} \delta_{11i}^0 (K_{1i}^0 - K_{1i}^0) \sum_{i_2=1}^{j-1} \delta_{1i_2}^0 (K_{1i_2}^0 - K_{1i_2}^0))
\]

\[
= \sum_{j=1}^{m+1} tr((\Sigma^0)^{-1/2} \sum_{t=T_{j-1}^0+1}^{T_j^0} \eta_t \sum_{i=1}^{j-1} \delta_{11i}^0 (K_{1i}^0 - K_{1i}^0) \sum_{i_2=1}^{j-1} \delta_{1i_2}^0 (K_{1i_2}^0 - K_{1i_2}^0))
\]

\[
= \sum_{j=1}^{m+1} tr((\Sigma^0)^{-1/2} \sum_{t=T_{j-1}^0+1}^{T_j^0} \eta_t \sum_{i=1}^{j-1} \delta_{11i}^0 (K_{1i}^0 - K_{1i}^0) \sum_{i_2=1}^{j-1} \delta_{1i_2}^0 (K_{1i_2}^0 - K_{1i_2}^0))
\]

\[
\Rightarrow \sum_{j=1}^{m+1} tr((\Sigma^0)^{-1/2} \sum_{t=T_{j-1}^0+1}^{T_j^0} \eta_t \sum_{i=1}^{j-1} \delta_{11i}^0 (W_1(\lambda_{j}^0) - W_1(\lambda_{j-1}^0)) \sum_{i_2=1}^{j-1} \delta_{1i_2}^0 (W_1(\lambda_{j}^0) - W_1(\lambda_{j-1}^0)))
\]

\[
= \sum_{j=1}^{m+1} tr((\Sigma^0)^{-1/2} B)
\]

with

\[
\widetilde{B} = \left[ \begin{array}{cc}
(\sum_{i=1}^{j-1} \delta_{11i}^0 s_{1i_1})(W_1(\lambda_{j}^0) - W_1(\lambda_{j-1}^0)) & \left( \sum_{i_2=1}^{j-1} \delta_{1i_2}^0 s_{1i_2})(W_1(\lambda_{j}^0) - W_1(\lambda_{j-1}^0)) \right)
\end{array} \right]
\]

In what follows, we provide details for case (a) for which the first equation has a break at \( T_{j}^0 \) and indicate the changes needed for the two other cases. Now,
\[
\begin{align*}
\sum_{t=1}^{T} \theta_0^0 (X_t^0 - \tilde{X}_t)(\Sigma^0)^{-1}(X_t^0 - \tilde{X}_t)'\theta_0^0 \\
= \sum_{j=1}^{m+1} \sum_{t=T_{j-1}+1}^{T_j} \theta_0^0 (X_{tj}^0 - \tilde{X}_{tj})(\Sigma^0)^{-1}(X_{tj}^0 - \tilde{X}_{tj})'\theta_0^0 \\
- \sum_{j=1}^{m} \sum_{t=T_{j}+1}^{T_{j+1}} \theta_0^0 (X_{tj}^0 - \tilde{X}_{tj})(\Sigma^0)^{-1}(X_{tj}^0 - \tilde{X}_{tj})'\theta_0^0 \\
+ \sum_{j=1}^{m} \sum_{t=T_{j}+1}^{T_j} \theta_0^0 (X_{tj}^0 - \tilde{X}_{tj+1})(\Sigma^0)^{-1}(X_{tj}^0 - \tilde{X}_{tj+1})'\theta_0^0 \\
= \sum_{j=1}^{m+1} \sum_{t=T_{j-1}+1}^{T_j} \left[ \sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{1i_1}^0) + \sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{1i_2}^0) \right] \\
\times (\Sigma^0)^{-1} \begin{bmatrix}
\sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{1i_1}^0) \\
\sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{1i_2}^0)
\end{bmatrix}

\times (\Sigma^0)^{-1} \begin{bmatrix}
\sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{1i_1}^0) + \delta_{1j_1}^0 (t - K_{1j_1}) + \sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{1i_2}^0) \\
\sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{1i_2}^0)
\end{bmatrix}
\times (\Sigma^0)^{-1} \begin{bmatrix}
\sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{1i_1}^0) \\
\sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{1i_2}^0)
\end{bmatrix}
\Rightarrow o_p(1)

= \sum_{j=1}^{m+1} \sum_{t=T_{j-1}+1}^{T_j} \left[ \sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{1i_1}^0) + \sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{1i_2}^0) \right] \\
\times (\Sigma^0)^{-1} \begin{bmatrix}
\sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{1i_1}^0) \\
\sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{1i_2}^0)
\end{bmatrix}

= \sum_{j=1}^{m+1} \sum_{t=T_{j-1}+1}^{T_j} \left[ \sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{1i_1}^0) + \sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{1i_2}^0) \right] \\
\times (\Sigma^0)^{-1} \begin{bmatrix}
\sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{1i_1}^0) \\
\sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{1i_2}^0)
\end{bmatrix}

= \sum_{j=1}^{m+1} (\lambda_j^0 - \lambda_{j-1}^0) tr((\Sigma^0)^{-1})

= \sum_{j=1}^{m+1} (\lambda_j^0 - \lambda_{j-1}^0) tr((\Sigma^0)^{-1} \tilde{A})

\text{with}

\tilde{A} = \\
\begin{bmatrix}
\left( \sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 s_{1i_1} \right)^2 & \left( \sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 s_{1i_1} \right) \left( \sum_{i_2=1}^{j_2-1} \delta_{2i_2}^0 s_{2i_2} \right) \\
\left( \sum_{i_2=1}^{j_2-1} \delta_{2i_2}^0 s_{2i_2} \right)^2 & \left( \sum_{i_2=1}^{j_2-1} \delta_{2i_2}^0 s_{2i_2} \right) \left( \sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 s_{1i_1} \right)
\end{bmatrix}

A.3
Remark 4 For case (b), with the second equation having a break at \( T_j^0 \), the term (A.1) is replaced by

\[
\sum_{j=1}^{m} \sum_{t=T_{j}+1}^{T_j^0} \left[ \sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{11}^0) + \delta_{1j_1}^0 (t - K_{1j_1}) \sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{11}^0) + \delta_{2j_1}^0 (t - K_{2j_1}) \right] \times (\Sigma^0)^{-1} \left[ \sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{11}^0) + \delta_{1j_1}^0 (t - K_{1j_1}) \sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{11}^0) + \delta_{2j_1}^0 (t - K_{2j_1}) \right]
\]

and for case (c) for which both equations have a break at \( T_j^0 \), it is replaced by

\[
\sum_{j=1}^{m} \sum_{t=T_j+1}^{T_j^0} \left[ \sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{11}^0) \sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{11}^0) + \delta_{2j_1}^0 (t - K_{2j_1}) \right] \times (\Sigma^0)^{-1} \left[ \sum_{i_1=1}^{j_1-1} \delta_{1i_1}^0 (K_{1i_1} - K_{11}^0) \sum_{i_2=1}^{j_2-1} \delta_{1i_2}^0 (K_{1i_2} - K_{11}^0) \right]
\]

Hence,

\[
\tilde{r}_T^1 \Rightarrow \max_{s_1,s_2} - (1/2)[2 \sum_{j=1}^{m+1} \text{tr}(\Sigma^0)^{-1/2} \tilde{B}) + \sum_{j=1}^{m+1} (\lambda_j^0 - \lambda_{j-1}^0) \text{tr}(\Sigma^0)^{-1/2} \tilde{A}]
\]

with

\[
s_1 = \{ T^{1/2} (\bar{K}_{11} - K_{11}^0), \ldots, T^{1/2} (\bar{K}_{1m_1} - K_{1m_1}^0) \}
\]

\[
s_2 = \{ T^{1/2} (\bar{K}_{21} - K_{21}^0), \ldots, T^{1/2} (\bar{K}_{2m_2} - K_{2m_2}^0) \text{ with } K_{1i} = K_{2i} \}
\]

Under \( H_1 \), we have similar results as under \( H_0 \) for regimes other than the \( l \)th one. However, for the \( l \)th regime, things are different given that we estimate \( K_{1l}^0 \) and \( K_{2l}^0 \). Now suppose we have the case \( T_{l-1}^0 < T_{l1}^0 < T_{l2}^0 < T_{l}^0 \) (note the subscripts 1 and 2 for \( T_{l1}^0 \) and \( T_{l2}^0 \) are not related to the number of equations, but rather to their order of occurrence). Then a) for \( t \in [T_{l-1}^0 + 1, T_{l}^0] \):

\[
\begin{bmatrix}
1 & t & t - K_{11}^0 & \cdots & t - K_{11,l1-1}^0 & 0 & \cdots & 0 \\
0 & 1 & t & t - K_{21}^0 & \cdots & t - K_{21,l2-1}^0 & 0 & \cdots & 0 \\
\end{bmatrix}
\]

b) for \( t \in [T_{l-1}^0 + 1, T_{l1}^0] \):

\[
\begin{bmatrix}
1 & t & t - K_{11}^0 & \cdots & t - K_{11,l1-1}^0 & 0 & \cdots & 0 \\
0 & 1 & t & t - K_{21}^0 & \cdots & t - K_{21,l2-1}^0 & 0 & \cdots & 0 \\
\end{bmatrix}
\]

A-4
c) for $t \in [T_{1t} + 1, T_{2t}]$, if the first break is in the first equation

$$
\hat{X}_{t_{1t}} = \begin{bmatrix}
1 & t & t - K_{11}^0 & \cdots & t - \hat{K}_{11t} - 1 & t - \hat{K}_{11t} & \cdots & 0 \\
& 1 & t & t - K_{21}^0 & \cdots & t - \hat{K}_{21t} - 1 & 0 & \cdots & 0
\end{bmatrix}
$$

while if the first break is in the second equation

$$
\hat{X}_{t_{2t}} = \begin{bmatrix}
1 & t & t - K_{11}^0 & \cdots & t - \hat{K}_{11t} - 1 & \cdots & 0 \\
& 1 & t & t - K_{21}^0 & \cdots & t - \hat{K}_{21t} - 1 & t - \hat{K}_{21t} & \cdots & 0
\end{bmatrix}
$$

d) for $t \in [T_{2t} + 1, T_{t}^0]$:

$$
\hat{X}_{t_{t+1}} = \begin{bmatrix}
1 & t & t - K_{11}^0 & \cdots & t - \hat{K}_{11t} - 1 & t - \hat{K}_{11t} & \cdots & 0 \\
& 1 & t & t - K_{21}^0 & \cdots & t - \hat{K}_{21t} - 1 & t - \hat{K}_{21t} & \cdots & 0
\end{bmatrix}
$$

Using these results, we have

$$
\sum_{t=1}^{T} \theta^0 (X_t^0 - \hat{X}_t) (\Sigma^0)^{-1} U_t = \sum_{t=1}^{T} \sum_{j=t}^{T} \left[ \sum_{i_1=1}^{j-1} \delta_{1i_1} (K_{1i_1} - K_{1i_1}^0) \sum_{i_2=1}^{j-2} \delta_{2i_2} (K_{2i_2} - K_{2i_2}^0) \right] (\Sigma^0)^{-1} U_t
$$

$$
+ \sum_{j=t}^{T} \sum_{i_1=1}^{m+1} \delta_{1i_1} (K_{1i_1} - K_{1i_1}^0) + \delta_{1i_1} (K_{1i_1} - K_{1i_1}^0)
$$

$$
+ \sum_{j_1=1}^{j-1} \sum_{j_2=1}^{j_1-1} \delta_{2j_2} (K_{2j_2} - K_{2j_2}^0) + \delta_{2j_2} (K_{2j_2} - K_{2j_2}^0)
$$

$$
+ \sum_{j_2=1}^{j-1} \sum_{j_1=1}^{j_2-1} \delta_{2j_2} (K_{2j_2} - K_{2j_2}^0) (\Sigma^0)^{-1} U_t
$$

which is given by:

$$
\sim \sum_{j=1}^{T} tr((\Sigma^0)^{-1/2} \hat{A}_1) + \sum_{j=1}^{m+1} tr((\Sigma^0)^{-1/2} \hat{A}_2)
$$

with

$$
\hat{A}_1 = \begin{bmatrix}
(\sum_{i_1=1}^{j-1} \delta_{1i_1} s_{1i_1} (W_1(\lambda_j^0) - W_1(\lambda_{j-1}^0))) & (\sum_{i_2=1}^{j-1} \delta_{2i_2} s_{2i_2} (W_1(\lambda_j^0) - W_1(\lambda_{j-1}^0))) \\
(\sum_{i_1=1}^{j-1} \delta_{1i_1} s_{1i_1} (W_2(\lambda_j^0) - W_2(\lambda_{j-1}^0))) & (\sum_{i_2=1}^{j-1} \delta_{2i_2} s_{2i_2} (W_2(\lambda_j^0) - W_2(\lambda_{j-1}^0)))
\end{bmatrix}
$$

and the elements of $\hat{A}_2$ are given by:

$$
\hat{A}_2(1, 1) = (\sum_{i_1=1}^{j-1} \delta_{1i_1} s_{1i_1} + \delta_{1i_1} s_{1i_1}) (W_1(\lambda_j^0) - W_1(\lambda_{j-1}^0))
$$

$$
\hat{A}_2(1, 2) = (\sum_{i_2=1}^{j-1} \delta_{2i_2} s_{2i_2} + \delta_{2i_2} s_{2i_2}) (W_1(\lambda_j^0) - W_1(\lambda_{j-1}^0))
$$

$$
\hat{A}_2(2, 1) = (\sum_{i_1=1}^{j-1} \delta_{1i_1} s_{1i_1} + \delta_{1i_1} s_{1i_1}) (W_2(\lambda_j^0) - W_2(\lambda_{j-1}^0))
$$

$$
\hat{A}_2(2, 2) = (\sum_{i_2=1}^{j-1} \delta_{2i_2} s_{2i_2} + \delta_{2i_2} s_{2i_2}) (W_2(\lambda_j^0) - W_2(\lambda_{j-1}^0))
$$
Using similar arguments, we have:

\[
\sum_{t=1}^{T} (X_t - \hat{X}_t)(\Sigma^0)^{-1}(X_t - \hat{X}_t)'\theta^0
\]

\[
= \sum_{j=1}^{T} \sum_{t=j+1}^{T} \left[ \sum_{i=1}^{t-1} \delta_{01i}(K_{1i} - K_{01i}) \sum_{j=1}^{t-2} \delta_{2ij}(K_{2ij} - K_{2ij}) \right]
\times (\Sigma^0)^{-1} \left[ \begin{array}{c} \sum_{i=1}^{t-1} \delta_{01i}(K_{1i} - K_{01i}) \\ \sum_{j=1}^{t-2} \delta_{2ij}(K_{2ij} - K_{2ij}) \end{array} \right] \\
+ \sum_{j=1}^{T} \sum_{t=j+1}^{T} \left[ \sum_{i=1}^{t-1} \delta_{01i}(K_{1i} - K_{01i}) + \delta_{11i}(K_{1i} - K_{11i}) \right] \\
+ \sum_{j=1}^{T} \sum_{t=j+1}^{T} \left[ \sum_{i=1}^{t-1} \delta_{01i}(K_{1i} - K_{01i}) + \sum_{j=1}^{t-2} \delta_{2ij}(K_{2ij} - K_{2ij}) + \delta_{21i}(K_{21i} - K_{21i}) \right] \\
\times (\Sigma^0)^{-1} \left[ \begin{array}{c} \sum_{i=1}^{t-1} \delta_{01i}(K_{1i} - K_{11i}) + \delta_{11i}(K_{1i} - K_{11i}) \\ \sum_{j=1}^{t-2} \delta_{2ij}(K_{2ij} - K_{2ij}) + \delta_{21i}(K_{21i} - K_{21i}) \end{array} \right]
\]

\[
\Rightarrow \sum_{j=1}^{T} \sum_{t=j+1}^{T} \left( (\Sigma^0)^{-1/2} \hat{A}_1 \right) \right] + \sum_{j=1}^{T} \sum_{t=j+1}^{T} \left( (\Sigma^0)^{-1/2} \hat{A}_2 \right)
\]

where

\[
\hat{B}_1 = \left[ \begin{array}{cc} (\lambda_j - \lambda_{j-1}) (\sum_{i=1}^{t-1} \delta_{01i}s_{1i})^2 & (\lambda_j - \lambda_{j-1}) (\sum_{i=1}^{t-1} \delta_{01i}s_{1i})(\sum_{j=1}^{t-2} \delta_{2ij}s_{2ij}) \\ (\lambda_j - \lambda_{j-1}) (\sum_{i=1}^{t-1} \delta_{01i}s_{1i})(\sum_{j=1}^{t-2} \delta_{2ij}s_{2ij}) & (\lambda_j - \lambda_{j-1}) (\sum_{j=1}^{t-2} \delta_{2ij}s_{2ij})^2 \end{array} \right]
\]

and the elements of \(\hat{B}_2\) are:

\[
\hat{B}_2(1,1) = (\lambda_j - \lambda_{j-1}) (\sum_{i=1}^{t-1} \delta_{01i}s_{1i})^2 \\
\hat{B}_2(2,2) = (\lambda_j - \lambda_{j-1}) (\sum_{j=1}^{t-2} \delta_{2ij}s_{2ij})^2 \\
\hat{B}_2(1,2) = \hat{B}_2(2,1)
\]

Collecting results:

\[
\hat{r}_T^1 = \max_{s_1, s_2} \left( 2 \sum_{j=1}^{T} \sum_{t=j}^{T} tr(\Sigma^0)^{-1/2} \hat{A}_1) + \sum_{j=1}^{T} \sum_{t=j+1}^{T} tr(\Sigma^0)^{-1/2} \hat{A}_2) \right]
\]

Proof of Theorem 2 (asymptotic distribution of the estimates of locally ordered breaks): The limit distribution of the estimates can be obtained from the maximization of

\[
l_T^2 = -(1/2) \sum_{t=1}^{T} \theta^0(X_t - X_t)(\Sigma^0)^{-1}(X_t - X_t)\theta^0 + 2 \sum_{t=1}^{T} \theta^0(X_t - X_t)(\Sigma^0)^{-1}U_t
\]

A-6
Given the rates of convergence and the definition of locally ordered breaks, we have 6 possible cases for the position of the candidate values for the estimates of the break dates relative to the true values; Case 1: $K_1 \leq K_0^1 \leq K_2 \leq K_0^2$, Case 2: $K_1 \leq K_0^1 \leq K_2 \leq K_0^2$, Case 3: $K_1 \leq K_2 \leq K_0^1 \leq K_0^2$, Case 4: $K_1^0 \leq K_1 \leq K_2 \leq K_0^2$, Case 5: $K_1^0 \leq K_1 \leq K_2 \leq K_0^2$, Case 6: $K_1^0 \leq K_0^2 \leq K_1 \leq K_2$. We derive the results only for Case 1, the arguments being similar for the others. We first have,

$$
\sum_{t=1}^{T} \theta^0(X_t - X_t)(\Sigma^0)^{-1}U_t
= \sum_{t=K_1^0+1}^{K_0^2} [\delta_1^0(K_1 - K_0^0) 0](\Sigma^0)^{-1}U_t
+ \sum_{t=K_0^1+1}^{T} [\delta_1^0(K_1 - K_0^0) \delta_2^0(K_2 - K_0^0)](\Sigma^0)^{-1}U_t
$$

For the other term,

$$
\sum_{t=1}^{T} \theta^0(X_t - X_t)(\Sigma^0)^{-1}(X_t^0 - X_t)\theta^0
= \sum_{t=K_1^0+1}^{K_0^2} [\delta_1^0(K_1 - K_0^0) 0](\Sigma^0)^{-1} \begin{bmatrix}
\delta_1^0(K_1 - K_0^0) \\
0
\end{bmatrix}
+ \sum_{t=K_0^1+1}^{T} [\delta_1^0(K_1 - K_0^0) \delta_2^0(K_2 - K_0^0)](\Sigma^0)^{-1} \begin{bmatrix}
\delta_1^0(K_1 - K_0^0) \\
\delta_2^0(K_2 - K_0^0)
\end{bmatrix}
$$

Collecting terms,

$$
lr_T^2 = -\frac{1}{2} \left\{ \sum_{t=K_1^0+1}^{K_0^2} [\delta_1^0(K_1 - K_0^0) 0](\Sigma^0)^{-1} \begin{bmatrix}
\sum_{t=K_0^1+1}^{K_0^2} [\delta_1^0(K_1 - K_0^0)] \\
0
\end{bmatrix}
+ \sum_{t=K_0^1+1}^{T} [\delta_1^0(K_1 - K_0^0) \delta_2^0(K_2 - K_0^0)](\Sigma^0)^{-1} \begin{bmatrix}
\sum_{t=K_0^1+1}^{K_0^2} [\delta_1^0(K_1 - K_0^0)] \\
\delta_2^0(K_2 - K_0^0)
\end{bmatrix}
\right\}
- \left\{ \sum_{t=K_1^0+1}^{K_0^2} [\delta_1^0(K_1 - K_0^0) 0](\Sigma^0)^{-1}U_t
+ \sum_{t=K_0^1+1}^{T} [\delta_1^0(K_1 - K_0^0) \delta_2^0(K_2 - K_0^0)](\Sigma^0)^{-1}U_t \right\}
\Rightarrow -\frac{1}{2} \left\{ tr((\Sigma^0)^{-1} \left\{ (1 - \lambda_0^2)\delta_1^0 s_1^2 (1 - \lambda_0^0)\delta_1^0 s_1 s_2 \\
(1 - \lambda_0^0)\delta_1^0 s_1 s_2 (1 - \lambda_0^0)\delta_2^0 s_2^2 \\
+ 2tr((\Sigma^0)^{-1/2} \left\{ [1^0 s_1(W_1 - W_1(\lambda_0^2)) \delta_2^0 s_2(W_1(1) - W_1(\lambda_0^2))] \\
[1^0 s_1(W_2 - W_2(\lambda_0^2)) \delta_2^0 s_2(W_2(1) - W_2(\lambda_0^2))] \right\} \right\}
\right\}
$$
This implies that

\[
\begin{bmatrix}
T^{1/2}(\tilde{K}_1 - K_1^0) \\
T^{1/2}(\tilde{K}_2 - K_2^0)
\end{bmatrix} \Rightarrow \arg\max_{s_1, s_2} H(s_1, s_2)
\]

\[
= -\frac{1}{2} tr((\Sigma^0)^{-1}) 
\begin{bmatrix}
(1 - \lambda_2^0)\delta_1^{02}s_1^2 & (1 - \lambda_2^0)\delta_1^{0}\delta_2^{0}s_1s_2 \\
(1 - \lambda_2^0)\delta_1^{0}\delta_2^{0}s_1s_2 & (1 - \lambda_2^0)\delta_2^{02}s_2^2
\end{bmatrix}
\]

\[
+ 2 tr((\Sigma^0)^{-1/2}) 
\begin{bmatrix}
\delta_1^{0}s_1(W_1(1) - W_1(\lambda_2^0)) & \delta_2^{0}s_2(W_1(1) - W_1(\lambda_2^0)) \\
\delta_1^{0}s_1(W_2(1) - W_2(\lambda_2^0)) & \delta_2^{0}s_2(W_2(1) - W_2(\lambda_2^0))
\end{bmatrix}
\]