Confidence Sets for the Break Date in Cointegrating Regressions [1]

Eiji Kurozumi

Department of Economics

Hitotsubashi University

Anton Skrobotov

Institute of Applied Economic Studies

Russian Presidential Academy of National Economy and Public Administration

This version: May 1, 2017

Abstract

In this paper, we propose constructing confidence sets for a break date in cointegrating regressions by inverting a test for the break location, which is obtained by maximizing the weighted average of power. It is found that the limiting distribution of the test depends on the number of I(1) regressors whose coefficients sustain structural change and the number of I(1) regressors whose coefficients are fixed throughout the sample. By Monte Carlo simulations, we then show that compared with a confidence interval developed by using the existing method based on the limiting distribution of the break point estimator under the assumption of the shrinking shift, the confidence set proposed in the present paper has a more accurate coverage rate, while the length of the confidence set is comparable. By using the method developed in this paper, we then investigate the cointegrating regressions of Russian macroeconomic variables with oil prices with a break.

JEL classification: C12, C21

Keywords: Confidence interval, structural change, cointegration, Russian economy, oil price

1. Introduction

This paper proposes constructing a confidence set for the change point in cointegrating regressions. Cointegration has long been an important concept for investigating the long-run relationships among macroeconomic variables.[2] To capture the long-run relationship, data over relatively long time-frames are often used in such an investigation. In this case, we should take into account that the economic structure may change during the sample period. For example, Campos, Ericsson, and Hendry (1996) investigate the effect of structural change on cointegration tests and Gregory and Hansen (1996a, b) propose tests for the null hypothesis of no cointegration that are robust to the existence of structural change, while tests for the null hypothesis of cointegration with a structural break are proposed by Carrion-i-Silvestre and Sansó (2006) and Arai and Kurozumi (2007). On the contrary, tests for structural change in the framework of cointegrating regressions have been proposed by Bai, Lumsdaine, and Stock (1998, BLS hereafter) and Kejriwal and Perron (2010). By using the tests presented in the literature in addition to the careful inspection of original data and economic events, we may find cointegrating relations with structural change. In this case, a statistical inference about the change point can be made by using the method proposed by BLS in the case of a single break, while multiple breaks were investigated by Kejriwal and Perron (2008a).

In the case of regressions using stationary variables, the break point is estimated by minimizing the sum of the squared residuals or by using the quasi-maximum likelihood method, while the confidence interval is constructed by using the limiting distribution of the break point estimator, as suggested by Bai (1997) and Bai and Perron (1998).[3] In this case, the crucial assumption made for the construction of the confidence interval is that the magnitude of the structural break shrinks to 0 at a rate slower than $1/\sqrt{T}$, as also assumed in BLS and Kejriwal and Perron (2008a). However, as demonstrated by Elliott and Müller (2007) and Chang and Perron (2015), a confidence interval based on the limiting distribution of the break point estimator tends to be too liberal when the magnitude of the break is not so large. Instead of using the limiting distribution of the change point estimator, Elliott and Müller

---

[2]Recently we can find a lot of work on large-dimensional VAR and/or Bayesian VAR methods but we focus on classical cointegration models using frequentist methods in this paper.

[3]Change point estimators have been investigated in the statistical and econometric literature. See, for example, Csörgő and Horváth (1997), Perron (2006) and Aue and Horváth (2013) among others.

(2007) propose constructing a confidence interval by inverting the locally best invariant test for the break location, which helps control the coverage rate.[4] However, the drawback of their method, as pointed out by Chang and Perron (2015), is that the confidence interval tends to be too wide. Indeed, it covers most of the sample period in some cases, thereby offering no useful information in practice. To overcome this drawback, Yamamoto (2016) pays attention to the estimation of the long-run variance for the construction of the test for the break location and proposes estimating it by taking the estimated break point into account, while Kurozumi and Yamamoto (2015) consider a similar method to Elliott and Müller (2007) but propose inverting the sup-type, average-type, and exponential-type tests for the break location, which can be obtained by maximizing the average power of a test. By Monte Carlo simulations, it is shown that these methods can better control the coverage rate and that the length of the confidence set becomes close to or smaller than that based on Bai (1997). On the contrary, Eo and Morley (2015) investigate a confidence set based on the likelihood ratio, while Harvey and Leybourne (2015) propose constructing a confidence set for the date of a break in level and trend without stochastic regressors using the LBI test. Their method is valid for both I(0) and I(1) processes. Further, Kurozumi (2016) extends the method of Kurozumi and Yamamoto (2015) to linear regression models with non-homogeneous regressors, particularly with a linear trend.

As in the above case of stationary regressions, controlling the coverage rate of the confidence interval of the break date in the case of cointegrating regressions may be difficult based on the methods of BLS and Kejriwal and Perron (2008a). Indeed, the simulation results reported in these papers are not necessarily satisfactory. Therefore, in this paper, we propose constructing a confidence set by inverting the test for the break location in cointegrating regressions. In this case, while the basic structure of the test is the same as in the case of regressions with stationary regressors, the limiting distribution becomes different. We derive the asymptotic distribution of the test depending on whether a linear trend is included in the regressions and/or the coefficient associated with the I(1) regressors sustain a structural change. As in the case of stationary regressions, the critical values depend on the location of the break fraction under the null hypothesis and it is inconvenient to tabulate them for all

---

[4]The duality between confidence regions and tests is a well-known property in the statistical literature.

the permissible break fractions in practical analysis. Instead, as in Kurozumi and Yamamoto (2015), we conduct response surface regressions and propose obtaining the critical values of the test for the break location in a simple formula. By Monte Carlo simulations, we show that our method can control the coverage rate of the confidence set better than BLS, while the size of the confidence set based on our method is comparable to that of BLS.

The rest of the paper is organized as follows. Section 2 introduces the model and assumptions. We separately consider models for whether a linear trend is included in the cointegrating regressions and/or the coefficients of the I(1) regressors change after the break point. In Section 3, we propose constructing the confidence set by inverting the test for the break location. The finite sample property of our method is investigated by Monte Carlo simulations in Section 4, and we apply the method developed in this paper to a Russian economic analysis in Section 5. Section 6 offers concluding remarks. The mathematical proofs are relegated to the supplementary on-line appendix A.

2. Model and Assumptions

Let us consider the following linear model:

$$y_t = w'_{b,t}\beta_b + w_{b,t}(\lambda_0)'\delta_b + w'_{f,t}\beta_f + e_t \tag{1}$$

for $t = 1, \cdots, T$, where $w_{b,t}$, $w_{b,t}(\lambda_0)$, and $w_{f,t}$ are p_b-, p_b-, and p_f-dimensional regressors, respectively, $w_{b,t}(\lambda_0) = 1(t > [\lambda_0 T])w_{b,t}$ with $1(\cdot)$ being an indicator function, λ_0 is a true break fraction, $[a]$ denotes the largest integer less than a, e_t is an error term, and β_b, δ_b, and β_f are p_b-, p_b-, and p_f-dimensional unknown coefficients, respectively. Since this paper aims to construct a confidence set for the break date, we assume that a one-time structural change occurred in the sample period and the true break date is denoted as $T_0 = [\lambda_0 T]$. Note that the coefficient associated with $w_{b,t}$ sustains a structural change from β_b to $\beta_b + \delta_b$, while β_f is fixed throughout the sample period.

Since we consider a cointegrating regression model, the regressors $w_{b,t}$ and/or $w_{f,t}$ include I(1) variables, which are given by

$$z_{b,t} = z_{b,t-1} + u^z_{b,t}$$
$$z_{f,t} = z_{f,t-1} + u^z_{f,t},$$

4

where $z_{b,t}$ and $z_{f,t}$ are p_b^z- and p_f^z-dimensional, respectively. In the following analysis, we mainly proceed without a drift in the I(1) variables even when $w_{b,t}$ includes a linear trend. As discussed in Remark 2, the main result in regressions with the I(1) variables with a drift remains the same by slightly modifying the test statistics in the next section.

In general, the error terms $u_{b,t}^z$ and $u_{f,t}^z$ are correlated with the regression error e_t in (1) and this correlation will become a nuisance parameter in our test. To eliminate the correlation, we include as regressors the leads-and-lags of the first differences of the I(1) variables in model (1), as suggested by Phillips and Loretan (1991), Saikkonen (1991), and Stock and Watson (1993). Such a model becomes

$$y_t = w_{b,t}'\beta_b + w_{b,t}(\lambda_0)'\delta_b + w_{f,t}'\beta_f + \sum_{j=-l}^{l} \pi_{b,j}'\Delta z_{b,t-j} + \sum_{j=-l}^{l} \pi_{f,j}'\Delta z_{f,t-j} + u_t, \qquad (2)$$

where Δ denotes the first differencing operator and both the leads-and-lags lengths are l for notational convenience, but they can be different as in, for example, Hayakawa and Kurozumi (2008) and Choi and Kurozumi (2012). In the following, we assume that u_t are uncorrelated with $u_{b,t-j}^z$ and $u_{f,t-j}^z$ for all j.

Remark 1 *In general, the leads-and-lags lengths are infinite and thus if we truncate them at some point l, then the error term u_t includes an additional noise caused by the truncation. However, this noise can be negligible asymptotically under additional assumptions, as proven by Saikkonen (1991), Arai and Kurozumi (2007), and Kejriwal and Perron (2008b) among others. We thus consider the finite leads-and-lags as in (2) only because we would like to avoid complicating the proof further; the results obtained in the paper also hold in the general infinite leads-and-lags case. Some examples of the case where the exact finite-order leads-and-lags are when the I(1) regressors are independent of the regression error e_t, or when the error terms, e_t, $u_{b,t}^z$, and $u_{f,t}^z$, are expressed as the finite order VAR model. We also note that in the general case of the infinite order leads-and-lags, the truncation order l should go to infinity as $T \to \infty$. For example, when the error term is a finite-order ARMA process, l is allowed to increase at a logarithmic rate and thus it can be chosen by usual information criteria such as the AIC and BIC. Details are discussed in the on-line appendix (see also Kejriwal and Perron, 2008b, in the case of usual cointegrating regressions).*

5

For notational convenience, we bundle up the additional stationary regressors $\Delta z_{b,t-j}$ and $\Delta z_{f,t-j}$ and denote them as a p_f^x-dimensional vector $x_{f,t}$. Since $x_{f,t}$ consists of the first differences of the I(1) regressors, we assume $E[x_{f,t}] = 0$ throughout the paper. In addition, because the coefficient associated with $x_{f,t}$ is fixed throughout the sample, we absorb $x_{f,t}$ into $w_{f,t}$ and then the model can be simply expressed as

$$
\begin{aligned}
y_t &= w_{b,t}'\beta_b + w_{b,t}(\lambda_0)'\delta_b + w_{f,t}'\beta_f + u_t \\
&= w_t(\lambda_0)'\beta + u_t,
\end{aligned}
\tag{3}
$$

where $w_{f,t}$ includes $x_{f,t}$ and $w_t(\lambda_0) = [w_{b,t}', w_{b,t}(\lambda_0)', w_{f,t}']'$. Model (3) can be expressed in vector form as

$$
\begin{aligned}
y &= W_b\beta_b + W_b(\lambda_0)\delta_b + W_f\beta_f + u \\
&= W(\lambda_0)\beta + u,
\end{aligned}
\tag{4}
$$

where $y = [y_1, \cdots, y_T]'$, $W_b = [w_{b,1}, \cdots, w_{b,T}]'$, $W_b(\lambda_0) = [0, \cdots, 0, w_{b,T_0+1}, \cdots, w_{b,T}]'$, $W_f = [w_{f,1}, \cdots, w_{f,T}]'$, $W(\lambda_0) = [W_b, W_b(\lambda_0), W_f]$, and $u = [u_1, \cdots, u_T]'$.

As a specification of the regressors $w_{b,t}$ and $w_{f,t}$, we consider the following commonly used models in practical analysis:

Model I-a : A constant and the I(1) regressors are included in the cointegrating relation and all the coefficients except for those associated with the first differences of the I(1) regressors sustain a break. That is, $w_{b,t} = [1, z_{b,t}']'$, $w_{f,t} = x_{f,t}$, $\beta_b = [\beta_{b,c}, \beta_{b,z}']'$, $\delta_b = [\delta_{b,c}, \delta_{b,z}']'$, and $\beta_f = \beta_{f,x}$:

$$
y_t = \beta_{b,c} + z_{b,t}'\beta_{b,z} + 1(t > [\lambda_0 T])\left(\delta_{b,c} + z_{b,t}'\delta_{b,z}\right) + x_{f,t}'\beta_{f,x} + u_t.
$$

Model I-b : A constant and the I(1) regressors are included in the cointegrating relation and only a constant term sustains a break. That is, $w_{b,t} = 1$, $w_{f,t} = [z_{f,t}', x_{f,t}']'$, $\beta_b = \beta_{b,c}$, $\delta_b = \delta_{b,c}$, and $\beta_f = [\beta_{f,z}', \beta_{f,x}']'$:

$$
y_t = \beta_{b,c} + 1(t > [\lambda_0 T])\delta_{b,c} + z_{f,t}'\beta_{f,z} + x_{f,t}'\beta_{f,x} + u_t.
$$

Model I-c : A constant and the I(1) regressors are included in the cointegrating relation and some of the coefficients associated with the I(1) regressors are fixed throughout the sample. That is, $w_{b,t} = [1, z'_{b,t}]'$, $w_{f,t} = [z'_{f,t}, x'_{f,t}]'$, $\beta_b = [\beta_{b,c}, \beta'_{b,z}]'$, $\delta_b = [\delta_{b,c}, \delta'_{b,z}]'$, and $\beta_f = [\beta'_{f,z}, \beta'_{f,x}]'$:

$$y_t = \beta_{b,c} + z'_{b,t}\beta_{b,z} + 1(t > [\lambda_0 T])\left(\delta_{b,c} + z'_{b,t}\delta_{b,z}\right) + z'_{f,t}\beta_{f,z} + x'_{f,t}\beta_{f,x} + u_t.$$

Model II-a : A constant, a linear trend, and the I(1) regressors are included in the cointegrating relation and all the coefficients except for those associated with the first differences of the I(1) regressors sustain a break. That is, $w_{b,t} = [1, t, z'_{b,t}]'$, $w_{f,t} = x_{f,t}$, $\beta_b = [\beta_{b,c}, \beta_{b,\tau}, \beta'_{b,z}]'$, $\delta_b = [\delta_{b,c}, \delta_{b,\tau}, \delta'_{b,z}]'$, and $\beta_f = \beta_{f,x}$:

$$y_t = \beta_{b,c} + \beta_{b,\tau}t + z'_{b,t}\beta_{b,z} + 1(t > [\lambda_0 T])\left(\delta_{b,c} + \delta_{b,\tau}t + z'_{b,t}\delta_{b,z}\right) + x'_{f,t}\beta_{f,x} + u_t.$$

Model II-b : A constant, a linear trend, and the I(1) regressors are included in the cointegrating relation and only a constant term and a linear trend sustain a break. That is, $w_{b,t} = [1, t]$, $w_{f,t} = [z'_{f,t}, x'_{f,t}]'$, $\beta_b = [\beta_{b,c}, \beta_{b,\tau}]'$, $\delta_b = [\delta_{b,c}, \delta_{b,\tau}]'$, and $\beta_f = [\beta'_{f,z}, \beta'_{f,x}]'$:

$$y_t = \beta_{b,c} + \beta_{b,\tau}t + 1(t > [\lambda_0 T])\left(\delta_{b,c} + \delta_{b,\tau}t\right) + z'_{f,t}\beta_{f,z} + x'_{f,t}\beta_{f,x} + u_t.$$

Model II-c : A constant, a linear trend, and the I(1) regressors are included in the cointegrating relation and some of the coefficients associated with the I(1) regressors are fixed throughout the sample. That is, $w_{b,t} = [1, t, z'_{b,t}]'$, $w_{f,t} = [z'_{f,t}, x'_{f,t}]'$, $\beta_b = [\beta_{b,c}, \beta_{b,\tau}, \beta'_{b,z}]'$, $\delta_b = [\delta_{b,c}, \delta_{b,\tau}, \delta'_{b,z}]'$, and $\beta_f = [\beta'_{f,z}, \beta'_{f,x}]'$:

$$y_t = \beta_{b,c} + \beta_{b,\tau}t + z'_{b,t}\beta_{b,z} + 1(t > [\lambda_0 T])\left(\delta_{b,c} + \delta_{b,\tau}t + z'_{b,t}\delta_{b,z}\right) + z'_{f,t}\beta_{f,z} + x'_{f,t}\beta_{f,x} + u_t.$$

Let $z_{u,t} = \sum_{j=1}^{t} u_j$ and $z_t = [z'_{b,t}, z'_{f,t}, z_{u,t}]' = [z'_{1,t}, z_{u,t}]'$, where $z_{1,t} = [z'_{b,t}, z'_{f,t}]'$. For model (3), we make the following assumption.

Assumption 1 *(a) The following weak convergences hold:*

$$\frac{1}{\sqrt{T}} z_{[rT]} \Rightarrow B^*(r) \quad uniformly\ over \quad 0 \leq r \leq 1, \tag{5}$$

7

$$\frac{1}{T}\sum_{t=1}^{[rT]} z_{1,t}u_t \Rightarrow \int_0^r B_1^*(r)dB_u^*(r)' \quad \textit{uniformly over} \quad 0 \le r \le 1, \tag{6}$$

where $B^*(r) = [B_1^*(r)', B_u^*(r)]' = [B_b^*(r)', B_f^*(r)', B_u^*(r)]'$ is a $(p_b^z + p_f^z + 1)$-dimensional Brownian motion with $E[B^*(s)B^*(t)'] = \min(s,t)\ \Omega = \min(s,t)\ diag\{\Omega_{11}, \omega_{uu}\}$ for $0 \le s, t \le 1$, $\Omega_{11} > 0$, $\omega_{uu} > 0$, and $\Rightarrow$ signifies the weak convergence under Skorohod topology.

(b) $x_{f,t}$ is a stationary process with $E[x_{f,t}] = 0$, $E[x_{f,t}u_{t-j}] = 0$ for all j, and the following relations hold:

$$\frac{1}{T}\sum_{t=1}^{T} x_{f,t}x_{f,t}' \xrightarrow{p} \Sigma_x > 0, \tag{7}$$

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_{f,t}u_t = O_p(1), \tag{8}$$

$$\frac{1}{T}\sum_{t=1}^{[rT]} z_t x_t' = O_p(1) \quad \textit{uniformly over} \quad 0 \le r \le 1, \tag{9}$$

where $\xrightarrow{p}$ signifies convergence in probability.

Assumption 1(a) implies that the functional central limit theorem holds for the I(1) regressors and the partial sum process based on u_t. Since the leads-and-lags of the I(1) regressors are included in model (3), u_t is supposed to be uncorrelated with all the leads-and-lags of $\Delta z_{b,t}$ and $\Delta z_{f,t}$, meaning that the so-called second-order bias does not appear in (6). The positive definiteness of Ω_{11} implies we do not allow for cointegration among the I(1) regressors. The theorem derived in this paper will not hold when z_{1t} are cointegrated. In practice, we may have several I(1) regressors and in that case, we first have to test for cointegration and confirm that there is no cointegrating relation among the I(1) regressors. Assumption 1(b) requires that the innovations driving the I(1) regressors are zero-mean stationary processes and that the well-known mild moment conditions also hold, as in (8) and (9).

3. Confidence Set for the Break Date

In this section, we consider constructing a confidence set for the break date by inverting the test for the location of the break point. For the unknown break point, we test for

$$H_N : T_0 = T_1 \quad \text{vs.} \quad H_A : T_0 = T_2 \tag{10}$$

8

with the significance level α, and if the null hypothesis is accepted, then we include T_1 in the confidence set; otherwise we exclude T_1 from the confidence set. By conducting this test for all the permissible break points, we obtain a confidence set for the break date with the confidence level $1 - \alpha$. Note that testing problem (10) can be expressed by using the break fraction as

$$H_N : \lambda_0 = \lambda_1 \quad \text{vs.} \quad H_A : \lambda_0 = \lambda_2 \tag{11}$$

where $\lambda_1 = T_1/T$ and $\lambda_2 = T_2/T$.

In this procedure, the confidence set becomes smaller as a test becomes more powerful and, therefore, we should construct a test as powerful as possible. However, it is not difficult to see that there exists no uniformly most powerful test for testing problem (10). Instead, following the literature, we consider constructing a test that maximizes the weighted average of power.

We first note that we cannot directly estimate (3) by using $w_{b,t}(\lambda_0)$ because $w_{b,t}(\lambda_0)$ depends on the unknown break fraction λ_0. Since the testing problem is given by (10), we consider estimating the model under the null hypothesis and construct the test statistic. Let $w_{b,t}(\lambda_1) = 1(t > [\lambda_1 T])w_{b,t}$ and $r_t(\lambda_0, \lambda_1) = w_{b,t}(\lambda_0) - w_{b,t}(\lambda_1)$. Then, model (3) can be expressed as

$$
\begin{aligned}
y_t &= w_{b,t}'\beta_b + w_{b,t}(\lambda_0)'\delta + w_{f,t}'\beta_f + u_t \\
&= w_{b,t}'\beta_b + w_{b,t}(\lambda_1)'\delta + w_{f,t}'\beta_f + u_t + r_t(\lambda_0, \lambda_1)'\delta \\
&= w_t(\lambda_1)'\beta + u_t(\lambda_0, \lambda_1),
\end{aligned}
\tag{12}
$$

where $w_t(\lambda_1) = [w_{b,t}', w_{b,t}(\lambda_1)', w_{f,t}']'$ and $u_t(\lambda_0, \lambda_1) = u_t + r_t(\lambda_0, \lambda_1)'\delta$. In vector form, this becomes

$$
\begin{aligned}
y &= W_b\beta_b + W_b(\lambda_0)\delta + W_f\beta_f + u \\
&= W_b\beta_b + W_b(\lambda_1)\delta + W_f\beta_f + u + R(\lambda_0, \lambda_1)\delta \\
&= W(\lambda_1)\beta + u(\lambda_0, \lambda_1),
\end{aligned}
\tag{13}
$$

where $W_b(\lambda_1)$ and $W(\lambda_1)$ are defined similarly to $W_b(\lambda_0)$ and $W(\lambda_0)$ in (4) with λ_0 replaced by λ_1, $R(\lambda_0, \lambda_1) = W_b(\lambda_0) - W_b(\lambda_1)$, and $u(\lambda_0, \lambda_1) = u + R(\lambda_0, \lambda_1)\delta$. From expression (13),

9

we restrict our attention to a class of tests invariant to the group of transformations given by $y \rightarrow y + W(\lambda_1)b$ and $\beta \rightarrow \beta + b$, where b is a $(2p_b + p_f)$-dimensional vector. In this case, it can be shown that there is no uniformly most powerful test even under restricted assumptions. Rather, the power of a test typically depends on the magnitude of the break, δ, and the location of the break under the alternative, λ_2, and thus it can be expressed as $P(\varphi$ rejects $H_N|\delta, \lambda_2)$ where φ is a test for (11) with the significance level α. Following Andrews and Ploberger (1994), Elliott and Müller (2007), and Kurozumi and Yamamoto (2015) among others, we consider maximizing the weighted average of $P(\varphi$ rejects $H_N|\delta, \lambda_2)$ over δ and λ_2 using some weighting functions, which is given by

$$\int \int P(\varphi \text{ rejects } H_N|\delta, \lambda_2)\, dQ_{\lambda_2}(\delta)dJ(\lambda_2) \tag{14}$$

where $Q_{\lambda_2}(\delta)$ and $J(\lambda_2)$ are non-negative measures on $\mathbb{R}^{p_b}$ and $(0,1)$, respectively. Typically, these weighting functions are chosen so that the asymptotic distribution of the test statistic becomes free of the nuisance parameters. In the case of cointegrating regressions, we set

$$Q_{\lambda_2}(\delta) \sim N\left(0, \frac{c\sigma^2}{T}\Sigma_\delta^{-1}\right) \quad \text{where} \quad \Sigma_\delta = \frac{1}{T}R(\lambda_2, \lambda_1)'M_{w_1}R(\lambda_2, \lambda_1)$$

with $M_{w_1} = I_T - W(\lambda_1)(W(\lambda_1)'W(\lambda_1))^{-1}W(\lambda_1)'$ and $R(\lambda_2, \lambda_1) = W_b(\lambda_2) - W_b(\lambda_1)$ (the choice of Q is discussed in the on-line appendix), while $J(\lambda_2)$ is a uniform measure on $\Lambda_\varepsilon = \{\lambda_2 : \varepsilon \leq \lambda_2 < \lambda_1 - \varepsilon,\ \lambda_1 + \varepsilon < \lambda_2 \leq 1 - \varepsilon\}$, in which both the end points and the neighborhood of λ_1 are excluded to avoid the explosive behavior of the test statistic. The value of ε must be determined by a researcher. Since small values of ε cover the wide range of the alternative, we will use $\varepsilon = 0.05$ in the rest of the paper. Then, in the same manner as Andrews and Ploberger (1994), Elliott and Müller (2007), and Kurozumi and Yamamoto (2015), under the additional assumption that $\{u_t\}$ is a sequence of i.i.d. normal random variables and $\{u_t\}$ is independent of $\{w_{b,t}\}$ and $\{w_{f,t}\}$, the test that maximizes the averaged power given by (14) rejects the null hypothesis when

$$
\begin{aligned}
\widetilde{LR}_T(\lambda_1) &= \int_{\lambda_2 \in \Lambda_\varepsilon} (1+c)^{-p_\beta/2} \exp\left\{ \frac{c}{2(1+c)\sigma^2} \right. \\
&\qquad \left. y'M_{w_1}R(\lambda_2, \lambda_1)\left(R(\lambda_2, \lambda_1)'M_{w_1}R(\lambda_2, \lambda_1)\right)^{-1} R(\lambda_2, \lambda_1)'M_{w_1}y \right\} d\lambda_2 > a
\end{aligned}
$$

for some value a. Since the localizing parameter c controls the weight of the magnitude of the break, the test can better detect a small break when c is close to zero whereas a test with

10

larger values of c is suitable for a large break. The choice of c will be discussed later. Note that the assumption of normality is used only for the derivation of the test statistic, and we do not assume normality in the remainder of the paper.

We first derive the limiting distribution of $\widetilde{LR}_T(\lambda_1)$ for each model. Let us partition the long-run variance Ω_{11} in Assumption 1(a) into Ω_{ij} for $i, j = b$ and f. Define

$$B_b(r) = \Omega_{bb}^{-1/2} B_b^*(r), \quad B_f(r) = \Omega_{ff}^{-1/2} B_f^*(r), \quad B_u(r) = \omega_{uu}^{-1/2} B_u^*(r),$$

$$B_{f \cdot b}(r) = \Omega_{f \cdot b}^{-1/2} \left(B_f^*(r) - \Omega_{fb} \Omega_{bb}^{-1} B_b^*(r) \right),$$

where $\Omega_{f \cdot b} = \Omega_{ff} - \Omega_{fb} \Omega_{bb}^{-1} \Omega_{bf}$. Note that they are all standard Brownian motions and that $B_b(r)$, $B_{f \cdot b}(r)$, and $B_u(r)$ are independent of each other.

Theorem 1 *Suppose Assumption 1 holds. Then, under the null hypothesis H_N, we have*

$$
\begin{aligned}
\widetilde{LR}_T &\Rightarrow \widetilde{LR}(\lambda_1) \\
&\equiv \int_{\lambda_2 \in \Lambda_\epsilon} (1+c)^{-p_\beta/2} \exp\left(\frac{c}{2(1+c)} G(\lambda_1, \lambda_2)' H(\lambda_1, \lambda_2)^{-1} G(\lambda_1, \lambda_2) \right) d\lambda_2,
\end{aligned}
$$

where (a) for Models I-a and II-a,

$G(\lambda_1, \lambda_2) =$
$$
\begin{cases}
\int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) dB_u(r) - \int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r)' dr \left(\int_0^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r)' dr \right)^{-1} \int_0^{\lambda_1} \tilde{B}_b(r) dB_u(r) & : \lambda_2 < \lambda_1 \\
-\int_{\lambda_1}^{\lambda_2} \tilde{B}_b(r) dB_u(r) - \int_{\lambda_1}^{\lambda_2} \tilde{B}_b(r) \tilde{B}_b(r)' dr \left(\int_{\lambda_1}^1 \tilde{B}_b(r) \tilde{B}_b(r)' dr \right)^{-1} \int_{\lambda_1}^1 \tilde{B}_b(r) dB_u(r) & : \lambda_2 > \lambda_1
\end{cases}
$$

$H(\lambda_1, \lambda_2) =$
$$
\begin{cases}
\int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r)' dr - \int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r)' dr \left(\int_0^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r)' dr \right)^{-1} \int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r) dr & : \lambda_2 < \lambda_1 \\
\int_{\lambda_1}^{\lambda_2} \tilde{B}_b(r) \tilde{B}_b(r)' dr - \int_{\lambda_1}^{\lambda_2} \tilde{B}_b(r) \tilde{B}_b(r)' dr \left(\int_{\lambda_1}^1 \tilde{B}_b(r) \tilde{B}_b(r)' dr \right)^{-1} \int_{\lambda_1}^{\lambda_2} \tilde{B}_b(r) \tilde{B}_b(r) dr & : \lambda_2 > \lambda_1
\end{cases}
$$

with $\tilde{B}_b(r) = [1, B_b(r)']'$ for Model I-a and $\tilde{B}_b(r) = [1, r, B_b(r)']'$ for Model II-a,
(b) for Models I-b and II-b,

$$
G(\lambda_1, \lambda_2) = \begin{cases}
\int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) dB_u(r) - J_1(\lambda_1, \lambda_2)' K(\lambda_1)^{-1} L(\lambda_1) & : \lambda_2 < \lambda_1 \\
-\int_{\lambda_1}^{\lambda_2} \tilde{B}_b(r) dB_u(r) - J_2(\lambda_1, \lambda_2)' K(\lambda_1)^{-1} L(\lambda_1) & : \lambda_2 > \lambda_1
\end{cases}
$$

$$
H(\lambda_1, \lambda_2) = \begin{cases}
\int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r)' dr - J_1(\lambda_1, \lambda_2)' K(\lambda_1)^{-1} J_1(\lambda_1, \lambda_2) & : \lambda_2 < \lambda_1 \\
\int_{\lambda_1}^{\lambda_2} \tilde{B}_b(r) \tilde{B}_b(r)' dr - J_2(\lambda_1, \lambda_2)' K(\lambda_1)^{-1} J_2(\lambda_1, \lambda_2) & : \lambda_2 > \lambda_1
\end{cases}
$$

11

with

$$J_1(\lambda_1, \lambda_2) = \left[\int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r)\tilde{B}_b(r)'dr, 0, \int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r)B_f(r)'dr \right]$$

$$J_2(\lambda_1, \lambda_2) = \left[0, -\int_{\lambda_1}^{\lambda_2} \tilde{B}_b(r)\tilde{B}_b(r)'dr, -\int_{\lambda_1}^{\lambda_2} \tilde{B}_b(r)B_f(r)'dr \right]$$

$$K(\lambda_1) = \begin{bmatrix} \int_0^{\lambda_1} \tilde{B}_b(r)\tilde{B}_b(r)'dr & 0 & \int_0^{\lambda_1} \tilde{B}_b(r)B_f(r)'dr \\ 0 & \int_{\lambda_1}^1 \tilde{B}_b(r)\tilde{B}_b(r)'dr & \int_{\lambda_1}^1 \tilde{B}_b(r)B_f(r)'dr \\ \int_0^{\lambda_1} B_f(r)\tilde{B}_b(r)'dr & \int_{\lambda_1}^1 B_f(r)\tilde{B}_b(r)'dr & \int_0^1 B_f(r)B_f(r)'dr \end{bmatrix}$$

$$L(\lambda_1) = \left[\left(\int_0^{\lambda_1} \tilde{B}_b(r)dB_u(r) \right)', \left(\int_{\lambda_1}^1 \tilde{B}_b(r)dB_u(r) \right)', \left(\int_0^1 B_f(r)dB_u(r) \right)' \right]'$$

with $\tilde{B}_b(r) = 1$ for Model I-b and $\tilde{B}_b(r) = [1, r]'$ for Model II-b,

(c) for Models I-c and II-c, $G(\lambda_1, \lambda_2)$, $H(\lambda_1, \lambda_2)$, $J_1(\lambda_1, \lambda_2)$, $J_2(\lambda_1, \lambda_2)$, $K(\lambda_1)$, and $L(\lambda_1)$ are defined as in Case (b) with $B_f(r)$ replaced by $B_{f \cdot b}(r)$ with $\tilde{B}(r) = [1, B_b(r)']'$ for Model I-c and $\tilde{B}(r) = [1, r, B_b(r)']'$ for Model II-c.

Remark 2 *When the I(1) regressors have a drift term, we typically consider including a linear trend in the regressions. Because a constant and a linear trend (with a break) are included as regressors in Models II-a to II-c, the test statistic $\widetilde{LR}_T(\lambda_1)$ is invariant to the drift term of the I(1) regressors if we replace $\Delta z_{b,t-j}$ and $\Delta z_{f,t-j}$ in $x_{f,t}$ with the demeaned versions, $\Delta z_{b,t-j} - \overline{\Delta z_b}$ and $z_{f,t-j} - \overline{\Delta z_f}$, respectively, to ensure that $E[x_{f,t}] = 0$. In this case, we obtain the same result as given in Theorem 1. Note that when the I(1) regressors have a drift term in Models I-a to I-c, the limiting distributions would be different from those given in Theorem 1, but we do not pursue these cases in this paper because it is natural and practically more relevant to include a linear trend as in Models II-a to II-c when the I(1) variables have a drift. We also note that the drift term of the I(1) regressors is allowed to sustain structural change, because it can be absorbed into the linear trend term (with a break). In this case, the demeaned versions of the I(1) regressors may be constructed using the estimated break date, as in the estimation of ω_{uu} explained at the end of this section.*

As we can see from Theorem 1, the limiting distributions become simpler for the models without the I(1) regressors with fixed coefficients because in those models, the regressor $w_{b,t}$,

whose coefficients sustain structural change, becomes asymptotically orthogonal to the other regressors $w_{f,t}$ with fixed coefficients. On the contrary, when the I(1) regressors are included in $w_{f,t}$, $w_{b,t}$ is correlated with $w_{f,t}$ even in the limit and thus the limiting distribution of the test statistic depends on p_f^z, the number of I(1) regressors with fixed coefficients.

Note that the limiting distribution depends on the localizing parameter c and we have to deal with c to implement the test. Following Andrews and Ploberger (1994), we consider the average-type test for which the weight is put on a small change ($c \to 0$) and the exponential-type test for which the large magnitude of a break is taken into account ($c \to \infty$). We can also consider the sup-type test following Andrews (1993), although it cannot be directly derived by the expression in Theorem 1. Then, their limiting expressions are given by

$$\text{sup-}LR(\lambda_1) = \sup_{\lambda_2 \in \Lambda_\epsilon} G(\lambda_1, \lambda_2)' H(\lambda_1, \lambda_2)^{-1} G(\lambda_1, \lambda_2), \tag{15}$$

$$\begin{aligned} \text{avg-}LR(\lambda_1) &= \lim_{c \to 0} \frac{2}{c} \left(\widetilde{LR}(\lambda_1) - 1 \right) \\ &= \int_{\lambda_2 \in \Lambda_\epsilon} G(\lambda_1, \lambda_2)' H(\lambda_1, \lambda_2)^{-1} G(\lambda_1, \lambda_2) d\lambda_2, \end{aligned} \tag{16}$$

$$\begin{aligned} \text{exp-}LR(\lambda) &= \lim_{c \to \infty} \log(1 + c)^{p_\beta/2} \widetilde{LR}(\lambda_1) \\ &= \log \int_{\lambda_2 \in \Lambda_\epsilon} \exp \left(\frac{1}{2} G(\lambda_1, \lambda_2)' H(\lambda_1, \lambda_2)^{-1} G(\lambda_1, \lambda_2) \right) d\lambda_2. \end{aligned} \tag{17}$$

The critical values of these distributions depend on λ_1 and it is inconvenient to tabulate them for all the permissible break fractions λ_1. Instead, we first calculate critical values for λ_1 ranging from 0.1 to 0.9 in increments of 0.01, for the given significance level, over $\Lambda_\varepsilon = \{\lambda_2 : \varepsilon \leq \lambda_2 < \lambda_1 - \varepsilon, \ \lambda_1 + \varepsilon < \lambda_2 \leq 1 - \varepsilon\}$ with $\varepsilon = 0.05$, where a standard Brownian motion is approximated by the normalized partial sums of standard normal random variables.[5] Next, we conduct the response surface regression by using the obtained critical values for each case as a function of λ_1. The regression form is given by

$$\text{critical value} = a_0 + a_{-1} \frac{1}{\ell + 1} + a_1 \ell + a_2 \ell^2 + a_3 \ell^3, \tag{18}$$

---

[5] As $\varepsilon = 0.05$, λ_1 may take values between 0.05 and 0.1 or 0.9 and 0.95. In these cases, the integral in the limiting expression is taken over the one-sided region ($\{\lambda_2 : \lambda_1 + \varepsilon < \lambda_2 \leq 1 - \varepsilon\}$ or $\{\lambda_2 : \varepsilon \leq \lambda_2 \leq \lambda_1 - \varepsilon\}$) when λ_1 is located near the end points, $\varepsilon \leq \lambda_1 \leq 2\varepsilon$ or $1 - 2\varepsilon \leq \lambda_1 \leq 1 - \varepsilon$, whereas the range of the integral is two-sided ($\{\lambda_2 : \varepsilon \leq \lambda_2 < \lambda_1 - \varepsilon, \ \lambda_1 + \varepsilon < \lambda_2 \leq \varepsilon\}$) when λ_1 is located far from the end points, $2\varepsilon < \lambda_1 < 1 - 2\varepsilon$. As a result, the response surface regression becomes complicated in this case, as shown by Kurozumi and Yamamoto (2015). To simplify our procedure as far as possible, we restrict the permissible values of the break fraction λ_1 to range from 0.1 to 0.9.

where $\ell = |\lambda_1 - 0.5|$. The obtained coefficients are summarized Tables B.1–B.6 in the on-line appendix.[6] Note that the regression coefficients a_0 to a_3 depend on the test statistics, significance level, number of $I(1)$ regressors whose coefficients sustain a break, and number of $I(1)$ regressors whose coefficients are fixed throughout the sample period.

From the limiting expressions (15)–(17), the corresponding finite sample versions, which can be implemented in practical analysis, are given by

$$\text{sup-}LR_T(T_1) \quad = \quad \max_{T_2 \in \mathcal{T}_\epsilon} F_{T_2}(T_1), \tag{19}$$

$$\text{avg-}LR_T(T_1) \quad = \quad \frac{1}{T^*} \sum_{T_2 \in \mathcal{T}_\epsilon} F_{T_2}(T_1), \tag{20}$$

$$\text{exp-}LR_T(T_1) \quad = \quad \log \left\{ \frac{1}{T^*} \sum_{T_2 \in \mathcal{T}_\epsilon} \exp \left(\frac{1}{2} F_{T_2}(T_1) \right) \right\}, \tag{21}$$

where $\mathcal{T}_\epsilon = \{T_2 : \epsilon T \leq T_2 < T_1 - \epsilon T, \ T_1 + \epsilon T < T_2 \leq (1 - \epsilon)T\}$, T^* is the number of T_2 included in $\mathcal{T}_\epsilon$, and $F_{T_2}(T_1)$ is the test statistic for the simple null hypothesis of $T_0 = T_1$ against the simple alternative of $T_0 = T_2$, which is given by

$$F_{T_2}(T_1) = \begin{cases} \left(\sum_{t=T_2+1}^{T_1} w_{b,t} \hat{u}_t \right)' \left(\tilde{\omega}_{uu} \hat{H} \right)^{-1} \left(\sum_{t=T_2+1}^{T_1} w_{b,t} \hat{u}_t \right) & : \quad T_2 < T_1 \\ \left(\sum_{t=T_1+1}^{T_2} w_{b,t} \hat{u}_t \right)' \left(\tilde{\omega}_{uu} \hat{H} \right)^{-1} \left(\sum_{t=T_1+1}^{T_2} w_{b,t} \hat{u}_t \right) & : \quad T_1 < T_T \end{cases} ,$$

where $\hat{u}_t$ are the regression residuals of y_t on $w_{b,t}$, $w_{b,t}(\lambda_1)$, and $w_{f,t}$,

$$\hat{H} = \sum_{t=1}^{T} \hat{r}_t(\lambda_2, \lambda) \hat{r}_t(\lambda_2, \lambda_1)'$$

with $\hat{r}_t$ being the regression residuals of $r_t(\lambda_2, \lambda_1)$ on $w_t(\lambda_1)$, and $\tilde{\omega}_{uu}$ is a consistent estimator of ω_{uu}. Note that we require $\tilde{\omega}_{uu}$ to be consistent under both the null and the alternative and thus, following Yamamoto (2016), we use the following nonparametric estimator:

$$\tilde{\omega}_{uu} = \tilde{\gamma}(0) + 2 \sum_{j=1}^{T} k(j, m) \tilde{\gamma}(j), \qquad \tilde{\gamma}(j) = \frac{1}{T} \sum_{t=j+1}^{T} \tilde{u}_t \tilde{u}_{t-j},$$

where $\tilde{u}_t$ are the regression residuals of y_t on $w_{b,t}$, $w_{b,t}(\lambda_1)$, $w_{f,t}$, and $1(|T_1 - \hat{T}_b| > p_b) w_{b,t}(\hat{\lambda}_b)$ with $\hat{T}_b$ $(\hat{\lambda}_b)$ being the break (fraction) estimator by minimizing the sum of the squared

---

[6]In the response surface regressions, the maximum relative error, (the fitted value − the obtained critical value)/the obtained critical value, among all the regressions was 0.028.

14

residuals, as considered by Kejriwal and Perron (2008a), and $k(j, m)$ is a kernel function such as the Bartlett and quadratic spectral (QS) kernel, as often used in practical analysis, with m being the bandwidth parameter. Note that $w_{b,t}(\hat{\lambda}_b)$ is multiplied by the indicator function $1(|T_1 - \hat{T}_b| > p_b)$ with p_b being the dimension of $w_b(t)$ (the regressors whose coefficients sustain structural change) to avoid the case where $\hat{T}_b$ is too close to T_1, so that the sample second moment of the regressors can be invertible in finite samples. See Yamamoto (2016) for details.

4. Finite Sample Property

In this section, we investigate the finite sample properties of the confidence sets constructed by inverting the three test statistics (19)–(21). We consider the case of only one I(1) regressor, where the data generating process is given by

$$
\begin{aligned}
\text{Model I-a:} \quad y_t &= 1 + z_t + 1(t > [\lambda_0 T])(\delta_{b,c} + \delta_{b,z} z_t) + u_t, \\
\text{Model I-b:} \quad y_t &= 1 + 1(t > [\lambda_0 T])\delta_{b,c} + z_t + u_t, \\
\text{Model II-a:} \quad y_t &= 1 + 3\frac{t}{T} + z_t + 1(t > [\lambda_0 T])\left(\delta_{b,c} + 3\delta_{b,\tau}\frac{t}{T} + \delta_{b,z} z_t\right) + u_t, \\
\text{Model II-b:} \quad y_t &= 1 + 3\frac{t}{T} + 1(t > [\lambda_0 T])\left(\delta_{b,c} + 3\delta_{b,\tau}\frac{t}{T}\right) + z_t + u_t.
\end{aligned}
$$

The I(1) regressor z_t and the error term u_t are generated by

$$
z_t = z_{t-1} + u_t^z, \quad u_t^z = \phi_z u_{t-1}^z + e_{z,t},
$$

$$
u_t = \phi_u u_{t-1} + e_{u,t},
$$

$$
\text{where} \quad \begin{bmatrix} e_{u,t} \\ e_{z,t} \end{bmatrix} \sim i.i.d.N(0, \Sigma) \quad \text{with} \quad \Sigma = \begin{bmatrix} 1 & \sigma_{zu} \\ \sigma_{zu} & 1 \end{bmatrix}.
$$

The sample size T is 100 and 300, and we consider the case where a one-time break occurs in the middle of the sample period, $\lambda_0 = 0.5$. We set the magnitude of the break to be $\delta_{b,c} = d/T^{1/4}$, $\delta_{b,\tau} = d/T^{1/4}$, and $\delta_{b,z} = d/T^{3/4}$ for $d = 4$, 8, 12, and 16, respectively. As a standard case, we consider the case with no serial correlation in u_t^z and u_t ($\phi_z = \phi_u = 0$), no endogeneity ($\sigma_{zu} = 0$), and the leads-and-lags length given by $l = 0$ (DGP1); we set $\phi_z = 0.5$, $\phi_u = 0$, and $\sigma_{zu} = 0$ with $l = 0$ to see the effect of the serial correlation in $u_{z,t}$ (DGP2), while

15

DGP3 allows for serial correlation in the regression errors u_t by setting $\phi_z = 0$, $\phi_u = 0.5$, $\sigma_{zu} = 0$, and $l = 0$. The effect of the endogeneity is investigated in DGP4 with $\phi_z = \phi_u = 0$, $\sigma_{zu} = 0.5$, where the leads-and-lags order l is selected by the BIC with the maximum lag order set to $[8(T/100)^{1/4}]$. Finally, DGP5 corresponds to the most general case in which serial correlation and endogeneity are allowed with $\phi_z = \phi_u = 0.5$, $\sigma_{zu} = 0.5$, and l selected by the BIC.[7] The long-run variance ω_{uu} is estimated using the QS kernel with the bandwidth selected by the automatic selection rule suggested by Andrews (1991). The confidence level is 0.95, the number of replications is 5000, and all the simulations are conducted by using the GAUSS matrix language.

Table 1 shows the coverage rates and lengths of the confidence sets for Model I-a. The column "BLS" is the result based on the confidence interval obtained by using the method proposed by BLS, while the columns "sup," "avg," and "exp" correspond to the confidence sets obtained by inverting the sup-type, average-type, and exponential-type tests, respectively. In DGP1, the coverage rate of BLS for $T = 100$ is much smaller than the nominal rate of 0.95 when the magnitude of the break is small, while it becomes close to 0.95 when a large structural change occurs. Similarly, the new methods do display some under-coverage for small d, albeit less severe than for BLS. Regarding the length of the confidence set, our three methods tend to construct larger confidence sets than BLS for small d, which may be due to more liberal coverage rate of BLS, while they tend be as small as BLS for large d. As a whole, the difference in the sizes of the confidence sets is relatively minor when the empirical coverage rates are close to 0.95.

From panel DGP2, we can see that the effects of serial correlation in the I(1) regressor on the coverage rates are very minor; they are close to those in DGP1. By contrast, positive serial correlation reduces the lengths of the coverage rates because the variation in the regressors becomes larger than the variance of u_t and the relative weight of the noise in the regression becomes smaller, thereby allowing us to estimate the model more efficiently.

On the contrary, serial correlation in the error term does affect the coverage rates, as shown in panel DGP3. For all the methods, the coverage rates become more liberal in this

---

[7]We also conducted simulations with a fixed leads-and-lags length using $l = [4(T/100)^{1/4}]$. The relative performance in this case is similar to the case with l selected by the BIC.

case. However, the coverage rate of the sup-type test recovers to the nominal rate as soon as the magnitude of the break becomes larger. In the case of positive serial correlation in u_t, the variation of the noise in the regression becomes large and the estimation becomes less efficient. As a result, the lengths of the confidence sets tend to be larger than those in the standard case.

Panel DGP4 suggests that the effect of the endogeneity is relatively minor, at least if we include the leads-and-lags by using the BIC. The coverage rates in DGP4 are close to those in DGP1, while the sizes of the confidence sets are slightly smaller.

DGP5 allows for serial correlation and endogeneity, and the results are a mixture of DGP2 to DGP4. Again, all the methods are liberal when $d = 4$; however, the coverage rate of the sup-type test becomes closest to 0.95 when $d \geq 8$, while the difference in the sizes of the confidence sets seems relatively minor.

When $T = 300$, all the methods perform better in terms of the coverage rates and the lengths of the confidence sets. In general, the relative performance is preserved in this case.

The results for Model I-b, II-a, aid II-b are summarized in Tables B.7–B.9 in the on-line appendix to save space. For Model I-b, the confidence set of BLS is liberal again when the magnitude of the break is small, whereas it becomes too conservative for a large break. For example, the coverage rate of BLS is 0.886 for $d = 4$, while it is 1.000 for $d = 16$ in DGP1. By contrast, our three methods are better able to control the coverage rate, although it is smaller than the nominal rate when $d = 4$ in DGP 3 and DGP5.

For Models II-a and II-b, the coverage rate of BLS becomes close to the nominal rate for the small break, whereas it is conservative for $d \geq 8$ for Model II-a and almost always 1.000 for Model II-b. On the contrary, the coverage rates of our three methods are relatively close to the nominal rate, although the sup-type test tends to be slightly conservative in some cases.

We also investigate the effect of the location of the true break point on the coverage rates. We conduct the simulations for $\lambda_0 = 0.3$, 0.5, and 0.7 using GDP5 and, as is seen in Tables B.10–B.13 in the on-line appendix, the location of the break point has only a minor effect for all the four method in Models I-a and I-b. whereas the performances in Models II-a and

II-b are more affected by λ_0 for small values of d. In these models, the coverage rates tends to be more liberal (conservative) when the location of the break is earlier (later).

Overall, the sup-type test is the best method for controlling the coverage rate, while the differences in the lengths of the confidence sets are minor as far as we can control the coverage rates of the tests, although these observations are limited to our simulation settings.

5. Empirical Application

In this section, we demonstrate the performance of our confidence sets by using Russian data. We consider the dependence of the Russian economy on oil prices. This positive long-run dependence originates from neoclassical theory of the capital accumulation channel, as discussed by Esfahani, Mohaddes, and Pesaran (2014), Idrisov, Kazakova and Polbin (2015) and Polbin and Skrobotov (2016). For example, in the Solow growth model, a part of oil revenue is saved in the form of the accumulation of physical capital. The following model is based on Polbin and Skrobotov (2016). This model distinguishes the dynamics of Russian GDP, consumption, and investment from the external economic environment by using a (univariate) cointegrating regression with structural breaks.[8]

First, we consider Russian quarterly real log GDP from 1999 to 2014 in constant 2003 prices (64 observations). Figure 1 illustrates two periods with different real GDP growth rates (refer to the left axis for GDP, consumption, and investment, while the right axis corresponds to oil prices). The first period shows recovering growth from approximately the beginning of the sample to mid-2008 (i.e., when the global financial crisis struck). During this period, the Russian economy grew approximately by 7% per year. This considerably high growth rate could have been induced by the country's economic recovery after the so-called transformation recession (accompanied by increasing production efficiency) as well as the considerable increase in oil prices (since 2003). After the structural break in 2008, a significant growth rate is observed only briefly after the crisis, 2008–2009 (recovering growth); when GDP reaches the pre-crisis level, the growth rate reduces significantly.

The long-run dependence of real GDP on oil prices can be expressed by the following

---

[8]Polbin and Skrobotov (2016) also consider the multivariate case based on the error correction model.

simple model:

$$\log y_t = \text{deterministic component} + \gamma \log p_t^{oil} + e_t \qquad (22)$$

where y_t is real output, p_t^{oil} is the real oil price, and γ is the long-run price elasticity of real output on the real oil price. If the process e_t is stationary, we interpret this relation as cointegration with the long-run dependence of output on oil prices. The structural break in mid-2008 can be induced by the reaction of the oil price dynamics as well as by the change in the long-run growth rate of the structural component of GDP. If only the first source matters, then there is no structural break in the deterministic component. If there is a change in the long-run growth rate of the structural component of GDP, then we can approximate this break as Model II-b:

$$\log y_t = \beta_{b,c} + \beta_{b,\tau} t + 1(t > [\lambda_0 T]) (\delta_{b,c} + \delta_{b,\tau} t) + \gamma \log p_t^{oil} + u_t. \qquad (23)$$

Throughout this section, we assume that the oil price is a random walk (see, inter alia, Alquist, Kilian, and Vigfusson, 2013). Then, as described by Polbin and Skrobotov (2016), the deterministic component in equation (22) is the unconditional expectation of the GDP growth rate in Russia. Therefore, if there is long-run (cointegrating) dependence of Russian real GDP on oil prices, the long-run GDP growth rate will coincide with the long-run growth rate of the structural component of GDP. A similar type of cointegrating relation for the Russian economy was considered in Rautava (2013) and Kuboniwa (2014). If there is no break, then the economy grows at a rate of $\beta_{b,\tau}$ under fixed oil prices. If there is a break, then the growth rate changes from $\beta_{b,\tau}$ to $\beta_{b,\tau} + \delta_{b,\tau}$. As in Polbin and Skrobotov (2016), according to neoclassical models of exogenous growth, the deterministic trends in (22) can be interpreted as long-run growth in labor efficiency in all sectors of the economy. This long-run growth ensures a balanced growth path of real GDP. In the Russian economy, the break in the deterministic component (in particular, the break in trend) in (23) originates from the end of the rapid growth period driven by the substantial reduction in the productivity gap between domestic and foreign economies and need to ensure growth at the expense of their own development.

We can also consider similar regressions for the real consumption and real investment series; real consumption and/or real investment may be cointegrated with oil prices with a

19

structural break in the trend function, as suggested by Figure 1 (see more discussion in Polbin and Skrobotov, 2016).

The data used in this section are logarithms of seasonally adjusted (by X-12-ARIMA) Russian quarterly real GDP, consumption, and investment from 1999 to 2014 in constant 2003 prices (source: Federal State Statistics Service, Russia) and real oil prices (Brent) from 1999 to 2014 (source: International Financial Statistics) deflated by seasonally adjusted CPI in the United States (source: Federal Reserve Economic Data). We investigate the data as follows:

1. Estimate the break date based on a regression without leads-and-lags.

2. Choose the number of leads-and-lags by the BIC (see Choi and Kurozumi, 2012).

3. Re-estimate the break date by using the selected leads-and-lags.

4. Test for the null hypothesis of cointegration with a break by using the test proposed by Arai and Kurozumi (2007).

5. If we do not reject the null hypothesis of cointegration, then construct the confidence set based on the leads-and-lags regression and estimated break date.

The first two steps imply that we choose the leads-and-lags under the alternative. By using the selected leads-and-lags, we re-estimate the break date in step 3. This estimated break date is then used for the LM-type test for the null hypothesis of cointegration with a break proposed by Arai and Kurozumi (2007) in step 4. We construct the confidence set in the final step if the null hypothesis of cointegration is not rejected. We note again that the real oil price is assumed to be a random walk as supposed in the literature.

Table 2 reports the estimated break dates. For GDP and consumption, they are in 2008, whereas a structural change may have occurred in 2006 in the relationship between investment and oil prices. By using the estimated break dates, we test for the null hypothesis of cointegration with a break for each series where the long-run variance is estimated based on the kernel method by using the QS kernel and data-dependent bandwidth rule proposed

20

by Andrews (1991), and we cannot reject the null hypothesis.[9] For completeness, we also test for the null hypothesis of cointegration without a break as proposed by Shin (1994), and we reject the null for GDP and investment at the 5% significance level and for consumption at the 10% significance level.[10] These results suggest a cointegrating relationship between Russian macroeconomic variables and oil prices because it is known that if series are cointegrated with a break, we do reject the null hypothesis of cointegration if we use the usual test without a break (Shin's test), whereas we fail to reject the null hypothesis if we appropriately introduce the break into the model, as in the case of unit root tests (Perron, 1989). We also checked the existence of a break by using the BIC and modified BIC proposed by Kurozumi and Tuvaandorj (2011) for discriminating between two models, with and without a break.[11] We find that the BIC and modified BIC provide evidence of the presence of break. From these results, we proceed with a cointegration model with a break as given by (23).

Once evidence of cointegration with a break is observed, we next construct confidence sets, as summarized in Figure 2. Note that our confidence sets provide additional evidence of the presence of a break because if the confidence set is empty, then we can conclude that there is no break. We select the different lengths of leads-and-lags by using the BIC, which are (leads, lags)=(0,0) for the regression with GDP, (1,5) for that with consumption, and (0,0) for that with investment. In the figure, the bold vertical line shows the estimated break date, which is also reported in Table 2. For the cointegrating regression with GDP reported in Figure 2(i), the sup-type and average-type tests return the same confidence sets ranging from 2008Q1 to 2009Q1, while those based on the exponential-type test and BLS are from 2008Q2 to 2008Q4. As observed in the simulation section, the coverage rate based on the sup-type test tends to be close to the nominal confidence level while it may be difficult to control the coverage rate for the other three methods; in particular, the coverage rate of BLS

---

[9]Model (23) is not included in Arai and Kurozumi (2007) and we obtained asymptotic critical values for the given break fractions from the simulations.

[10]We also employed the long-run variance estimators used in Carrion-i-Silvestre and Sansó (2006) and obtained the same results.

[11]The modified BIC proposed by Kurozumi and Tuvaandorj (2011) uses the additional penalty term related to uncertainty of the break point. See Kurozumi and Tuvaandorj (2011) for more details. We also note that Kejrival and Perron (2010) provided the sup-F test for only cointegration models without a linear trend and no limiting distribution and no corresponding set of critical values for the sup-F test are given in the case of models with a linear trend.

becomes either liberal or conservative depending on the case. Taking this result into account, the break point could be located outside the confidence interval estimated by BLS. Similarly, the confidence sets for the break dates in the cointegrating regressions with consumption and investment are reported in Figures 2(ii) and 2(iii). Similar to the regression with GDP, the confidence set based on the sup-type test is wider than the others; hence, considering the stable coverage rate of the sup-type test, the break point could again be located outside the periods covered by the other methods. We also observe that the confidence set based on our methods could return an asymmetric interval, while that based on BLS is always symmetric. To summarize, Russian quarterly macroeconomic variables are cointegrated with oil prices with a break and the possible break date could range more than one year.

We have so far applied our procedure for model (23) in which the coefficient of the real oil price, γ, is stable. For completeness, we also consider the model in which γ also sustains a structural change. Again, we find evidence of the presence of break by the information criteria, and the tests for cointegration show evidence of cointegration for all time series (see Table C.1 in Supplementary Online Appendix). The estimated break dates are the same except for investment, for which the break date moves from 2006Q4 to 2007Q1. As given in Figure C.1 in Supplementary Online Appendix, the confidence sets for GDP seem to be very close to that for the previous model except for the result by the exponential-type test (the confidence set is wider). For consumption, the confidence set based on the average-type test becomes narrower (only one point), while it becomes wider by using the exponential-type test. For investment, the confidence sets based on the sup-type and average-type tests remain the same while the confidence sets based on the exponential-test and BLS become wider.

6. Concluding Remarks

In this paper, we proposed constructing a confidence set for the break date in cointegrating regressions by inverting the test that maximizes the weighted average of power. We derived the limiting distribution of the test for the break location, which depends on whether a linear trend is included and/or the coefficients associated with the I(1) regressors sustain a structural change. By Monte Carlo simulations, we showed that the confidence set based on our test, particularly the sup-type test, can better control the coverage rate, while the size of

22

the confidence set is at least as small as the existing method. We then applied our method to analyze the Russian economy and showed that it is informative for the location of the break date.

References

[1] Alquist, R., L. Kilian and R. J. Vigfusson (2013). Forecasting the Price of Oil. *Handbook of Economic Forecasting* 2, 427-507.

[2] Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* 59, 817-858.

[3] Andrews, D. W. K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica* 61, 821-856 (Erratum, 71, 395-397).

[4] Andrews, D. W. K. and W. Ploberger (1994). Optimal Tests When a Nuisance Parameter Is Present Only Under the Alternative, *Econometrica* 62, 1383-1414.

[5] Arai, Y. and E. Kurozumi (2007). Testing for the Null Hypothesis of Cointegration with a Structural Break. *Econometric Reviews* 26, 705-7390.

[6] Aue, A. and L. Horváth (2013). Structural Breaks in Time Series. *Journal of Time Series Analysis* 34, 1-16.

[7] Bai, J. (1997). Estimation of a Change Point in Multiple Regression Models. *Review of Economics and Statistics* 79, 551-563.

[8] Bai, J., R. L. Lumsdaine and J. H. Stock (1998). Testing For and Dating Common Breaks in Multivariate Time Series. *Review of Economic Studies* 65, 395-432.

[9] Bai, J. and P. Perron (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica* 66, 47-78.

[10] Campos, J., N. R. Ericsson and D. F. Hendry (1996). Cointegration Tests in the Presence of Structural Breaks. *Journal of Econometrics* 70, 187-220.

[11] Carrion-i-Silvestre J. and A. J. Sansó (2006). Testing the Null of Cointegration with Structural Breaks. *Oxford Bulletin of Economics and Statistics* 68, 623-646.

[12] Chang, S. Y. and P. Perron (2015). A Comparison of Alternative Methods to Construct Confidence Intervals for the Estimate of a Break Date in Linear Regression Models. *Econometric Reviews*, DOI: 10.1080/07474938.2015.1122142.

[13] Choi, I. and E. Kurozumi (2012). Model Selection Criteria for the Leads-and-lags Cointegrating Regression. *Journal of Econometrics* 169, 224-238.

[14] Csörgő, M. and L. Horváth (1997). *Limit Theorems in Change-Point Analysis.* Wiley, New York.

[15] Elliott, G. and U. K. Müller (2007). Confidence Sets for the Date of a Single Break in Linear Time Series Regressions. *Journal of Econometrics* 141, 1196-1218.

[16] Eo, Y. and J. Morley (2015). Likelihood-ratio-based Confidence Sets for the Timing of Structural Breaks. *Quantitative Economics* 6, 463-497.

[17] Esfahani H. S., K. Mohaddes and M. H. Pesaran (2014). An Empirical Growth Model for Major Oil Exporters, *Journal of Applied Econometrics* 29, 1-21.

[18] Gregory, A. W. and B. E. Hansen (1996a). Residual-based Tests for Cointegration in Models with Regime Shifts. *Journal of Econometrics* 70, 99-126.

[19] Gregory, A. W. and B. E. Hansen (1996b). Tests for Cointegration in Models with Regime and Trend Shifts. *Oxford Bulletin of Economics and Statistics* 58, 555-560.

[20] Harvey, D. I. and S. J. Leybourne (2015). Confidence Sets for the Date of a Break in Level and Trend When the order of integration is unknown. *Journal of Econometrics* 184, 262-279.

[21] Hayakawa, K. and E. Kurozumi (2008). The Role of Leads in the Dynamic OLS Estimation of Cointegrating Regression Models. *Mathematics and Computers in Simulation* 79, 555-560.

[22] Idrisov, G., M. Kazakova, and A. Polbin (2015). A theoretical interpretation of the oil prices impact on economic growth in contemporary Russia. *Russian Journal of Economics* 1, 257-272.

[23] Kejriwal, M. and P. Perron (2008a). The Limiting Distribution of the Estimates in Cointegrated Regression Models with Multiple Structural Changes. *Journal of Econometrics* 146, 59-73.

[24] Kejriwal, M. and P. Perron (2008b). Data Dependent Rules for Selection of the Number of Leads and Lags in the Dynamic OLS Cointegrating Regression. *Econometric Theory* 24, 1425-1441.

[25] Kejriwal, M. and P. Perron (2010). Testing for Multiple Structural Changes in Cointegrated Regression Models. *Journal of Business & Economic Statistics* 28, 503-522.

[26] Kuboniwa, M. (2014). A Comparative Analysis of the Impact of Oil Prices on Oil-Rich Emerging Economies in the Pacific Rim. *Journal of Comparative Economics* 42, 328-339.

[27] Kurozumi, E. (2016). Construction of Confidence Sets for the Break Date in Regression Models with Non-homogeneous Regressors (in Japanese). *Journal of the Japan Statistical Society*, forthcoming.

[28] Kurozumi, E. and P. Tuvaandorj (2011). Model selection criteria in multivariate models with multiple structural changes. *Journal of Econometrics* 164, 218-238.

[29] Kurozumi, E. and Y. Yamamoto (2015). Confidence Sets for the Break Date Based on Optimal Tests. *Econometrics Journal* 18, 412-435.

[30] Perron, P. (1989). The Great Crash, the Oil Price shock, and the Unit Root Hypothesis. *Econometrica* 57, 1361-1401 (Erratum, 61, 249-249).

[31] Perron, P. (2006), Dealing with Structural Breaks, in T. C. Mills and K. Patterson, ed., *Palgrave Handbook of Econometrics, Vol.1: Econometric Theory*. Palgrave Macmillan, New York.

[32] Phillips, P. C. B. and M. Loretan (1991). Estimating Long-run Economic Equilibria. *Review of Economic Studies* 58, 407-436.

[33] Polbin A. and A. Skrobotov (2016). Testing for Structural Breaks in the Long-run Growth Rate of the Russian Economy (in Russian). *Mimeo.*

[34] Rautava, J. (2013). Oil Prices, Excess Uncertainty and Trend Growth. *Focus on European Economic Integration* 4, 77-87.

[35] Saikkonen, P. (1991). Asymptotically Efficient Estimation of Cointegration Regressions. *Econometric Theory* 7, 1-21.

[36] Shin, Y. (1994). A Residual- Based Test of the Null of Cointegration against the Alternative of No Cointegration. *Econometric Theory* 10, 91-115.

[37] Stock, J. H. and M. W. Watson (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica* 61, 783-820.

[38] Vogelsang, T. J. (1999). Sources of Nonmonotonic Power When Testing for a Shift in Mean of a Dynamic Time Series. *Journal of Econometrics* 88, 283-299.

[39] Yamamoto, Y. (2016). A Modified Confidence Set for the Structural Break Date in Linear Regression Models. *Econometric Reviews*, DOI: 10.1080/00927872.2016.1178892.

Table 1: Coverage rates and lengths of the confidence sets (Model I-a)

d	Coverage rates				Lengths of the confidence sets			
	BLS	sup	avg	exp	BLS	sup	avg	exp
DGP1				$T = 100$				
4	0.789	0.937	0.908	0.897	0.178	0.303	0.273	0.259
8	0.901	0.953	0.927	0.917	0.108	0.141	0.131	0.120
12	0.942	0.960	0.934	0.925	0.083	0.086	0.084	0.074
16	0.963	0.963	0.938	0.928	0.072	0.061	0.063	0.053
DGP2								
4	0.807	0.941	0.908	0.904	0.153	0.248	0.220	0.209
8	0.910	0.958	0.924	0.925	0.108	0.108	0.102	0.092
12	0.952	0.965	0.930	0.933	0.088	0.066	0.067	0.057
16	0.968	0.967	0.932	0.936	0.077	0.048	0.052	0.042
DGP3								
4	0.645	0.877	0.792	0.807	0.241	0.420	0.352	0.347
8	0.811	0.932	0.867	0.883	0.153	0.227	0.192	0.186
12	0.891	0.956	0.902	0.916	0.116	0.148	0.127	0.122
16	0.924	0.968	0.917	0.932	0.100	0.107	0.094	0.088
DGP4								
4	0.808	0.938	0.907	0.902	0.156	0.260	0.237	0.222
8	0.926	0.955	0.926	0.924	0.098	0.120	0.115	0.102
12	0.960	0.961	0.932	0.930	0.077	0.074	0.077	0.064
16	0.972	0.962	0.934	0.932	0.066	0.053	0.060	0.046
DGP5								
4	0.691	0.859	0.750	0.774	0.175	0.285	0.229	0.227
8	0.847	0.921	0.827	0.853	0.126	0.147	0.121	0.116
12	0.917	0.949	0.861	0.886	0.106	0.094	0.082	0.076
16	0.952	0.960	0.879	0.904	0.093	0.069	0.064	0.056
DGP1				$T = 300$				
4	0.833	0.953	0.940	0.940	0.126	0.234	0.225	0.213
8	0.918	0.959	0.944	0.944	0.072	0.100	0.103	0.092
12	0.954	0.960	0.945	0.946	0.051	0.060	0.064	0.055
16	0.963	0.962	0.946	0.948	0.043	0.041	0.046	0.039
DGP2								
4	0.852	0.952	0.936	0.939	0.117	0.181	0.174	0.163
8	0.925	0.957	0.940	0.943	0.069	0.071	0.076	0.066
12	0.954	0.958	0.942	0.945	0.051	0.043	0.050	0.041
16	0.971	0.959	0.942	0.945	0.043	0.031	0.037	0.030
DGP3								
4	0.686	0.928	0.885	0.892	0.204	0.404	0.361	0.354
8	0.833	0.950	0.911	0.918	0.120	0.198	0.180	0.171
12	0.900	0.959	0.922	0.929	0.088	0.122	0.115	0.107
16	0.932	0.965	0.929	0.935	0.071	0.086	0.083	0.076
DGP4								
4	0.848	0.953	0.938	0.937	0.112	0.196	0.188	0.177
8	0.935	0.958	0.942	0.943	0.066	0.082	0.086	0.075
12	0.956	0.960	0.943	0.945	0.047	0.049	0.054	0.046
16	0.968	0.961	0.944	0.947	0.038	0.034	0.040	0.033
DGP5								
4	0.756	0.926	0.874	0.884	0.164	0.276	0.240	0.235
8	0.889	0.949	0.901	0.916	0.106	0.123	0.112	0.105
12	0.931	0.958	0.909	0.925	0.082	0.074	0.071	0.065
16	0.949	0.962	0.915	0.929	0.066	0.052	0.052	0.046

Table 2: Empirical results

series: y_t, p_t^{oil}		
estimated break date	:	2008Q3
cointegration test with a break	:	0.057
cointegration test without a break	:	0.172**
series: c_t, p_t^{oil}		
estimated break date	:	2008Q4
cointegration test with a break	:	0.048
cointegration test without a break	:	0.117*
series: i_t, p_t^{oil}		
estimated break date	:	2006Q4
cointegration test with a break	:	0.049
cointegration test without a break	:	0.131**

Note: the superscripts *, **, and *** signify significance at 10%, 5% and 1%, respectively.
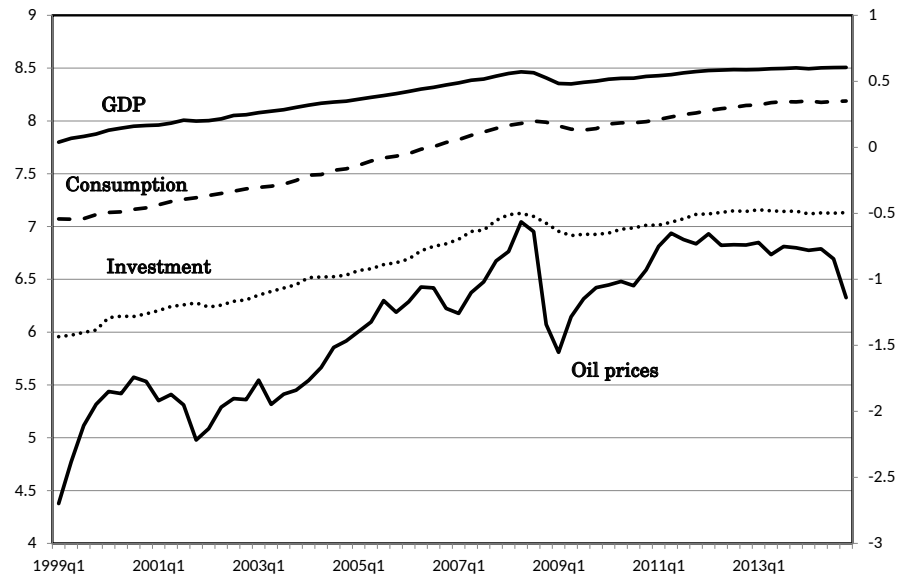
Figure 1: Russian macroeconomic series and oil prices

(i) y_t
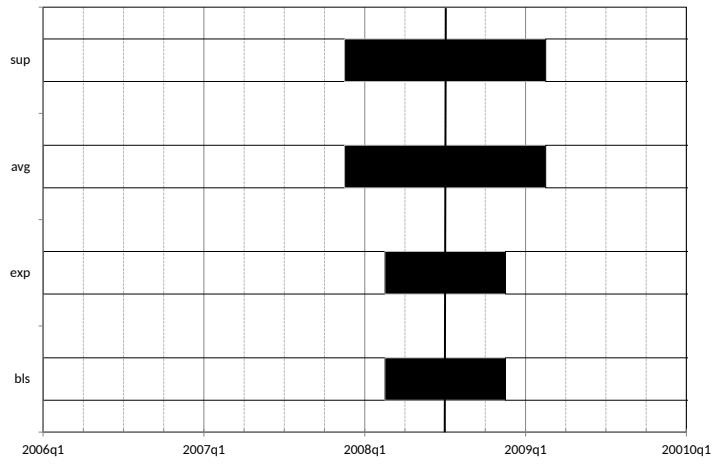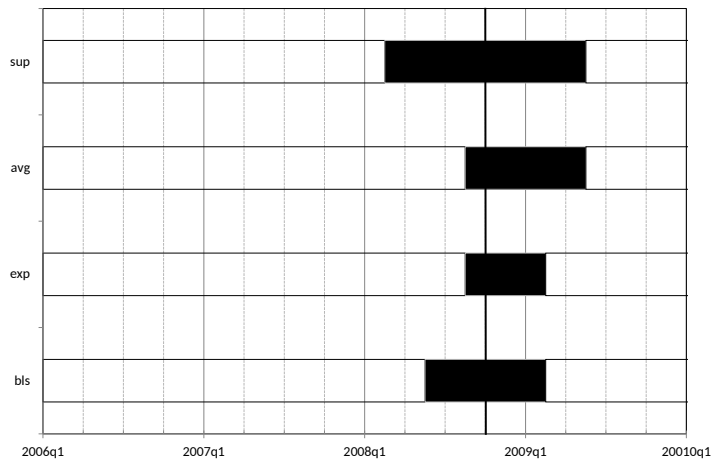
(ii) c_t
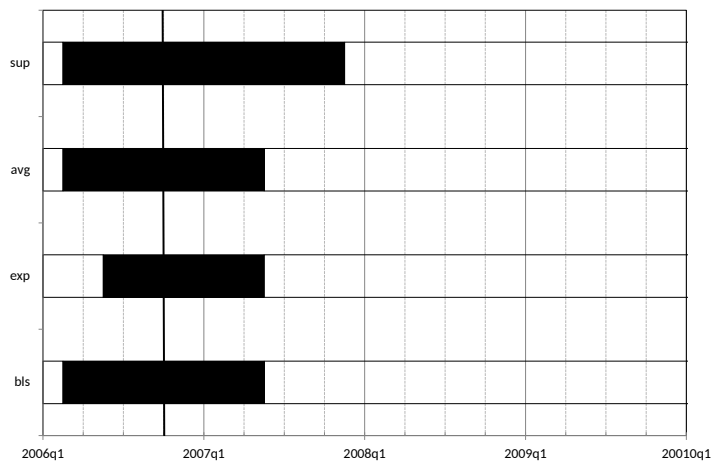
(iii) i_t

Figure 2: Confidence sets for a break

Supplementary Online Appendix to "Confidence Sets for the Break Date in Cointegrating Regressions"

Eiji Kurozumi   Anton Skrobotov

Department of Economics   Institute of Applied Economic Studies

Hitotsubashi University   Russian Presidential Academy of National
                                 Economy and Public Administration

A.1. Choice of the Variance of $Q_{\lambda_2}(\delta)$

Under the assumption that $\{u_t\}$ is a sequence of i.i.d. normal random variables and $\{u_t\}$ is independent of $\{w_{b,t}\}$ and $\{w_{f,t}\}$, the weighted average of power (conditional on $\{w_{b,t}\}$ and $\{w_{f,t}\}$) becomes

$$\int \int \exp\left(\frac{1}{\sigma^2}y'M_{w_1}R(\lambda_2,\lambda_1)\delta - \frac{1}{2\sigma^2}\delta'R(\lambda_2,\lambda_1)'M_{w_1}R(\lambda_2,\lambda_1)\delta\right)dQ_{\lambda_2}(\delta)dJ(\lambda_2). \quad (A.1)$$

Suppose that $Q_{\lambda_2}(\delta) \sim N(0,\Sigma)$. Then, the inner integral becomes, by integrating out δ,

$$\int \exp\left(\frac{1}{\sigma^2}y'M_{w_1}R(\lambda_2,\lambda_1)\delta - \frac{1}{2\sigma^2}\delta'R(\lambda_2,\lambda_1)'M_{w_1}R(\lambda_2,\lambda_1)\delta\right)dQ_{\lambda_2}(\delta)$$

$$= \left|I + \Sigma\frac{1}{\sigma^2}R(\lambda_2,\lambda_1)'M_{w_1}R(\lambda_2,\lambda_1)\right|^{-1/2}$$

$$\times \exp\left[\frac{1}{2\sigma^2}y'M_{w_1}R(\lambda_2,\lambda_1)\left(\sigma^2\Sigma^{-1} + R(\lambda_2,\lambda_1)'M_{w_1}R(\lambda_2,\lambda_1)\right)^{-1}R(\lambda_2,\lambda_1)'M_{w_1}y\right].$$

In order for the determinant to be tractable, we choose Σ such that it is proportional to $\sigma^2\left(R(\lambda_2,\lambda_1)'M_{w_1}R(\lambda_2,\lambda_1)\right)^{-1}$ and thus we set

$$Q_{\lambda_2}(\delta) \sim N\left(0,\frac{c\sigma^2}{T}\Sigma_\delta^{-1}\right) \quad \text{where} \quad \Sigma_\delta = \frac{1}{T}R(\lambda_2,\lambda_1)'M_{w_1}R(\lambda_2,\lambda_1).$$

In this case, it is easy to see that the determinat reduces to $|I + cI|^{-1/2} = (1+c)^{-p_\beta/2}$ while the inverse in the second term becomes

$$\left(\frac{1}{c}R(\lambda_2,\lambda_1)'M_{w_1}R(\lambda_2,\lambda_1) + R(\lambda_2,\lambda_1)'M_{w_1}R(\lambda_2,\lambda_1)\right)^{-1} = \frac{c}{1+c}\left(R(\lambda_2,\lambda_1)'M_{w_1}R(\lambda_2,\lambda_1)\right)^{-1}.$$

A.2. Proof of Theorem 1

Here, we prove the theorem for Model II-c. The other cases can be proven in the same manner. For the test statistic $\widetilde{LR}_T(\lambda_1)$, we note that

$$y'M_{w_1}R(\lambda_2,\lambda_1)\left(R(\lambda_2,\lambda_1)'M_{w_1}R(\lambda_2,\lambda_1)\right)^{-1}R(\lambda_2,\lambda_1)'M_{w_1}y$$
$$= \left(D_T'R(\lambda_2,\lambda_1)'M_{\tilde{w}_1}y\right)'\left(D_T'R(\lambda_2,\lambda_1)'M_{\tilde{w}_1}R(\lambda_2,\lambda_1)D_T\right)^{-1}\left(D_T'R(\lambda_2,\lambda_1)'M_{\tilde{w}_1}y\right),$$

where $D_T = diag\{T^{-1/2}, T^{-3/2}, T^{-1}\Omega_{bb}^{-1/2}\}$ and $M_{\tilde{w}_1} = \tilde{W}(\lambda_1)(\tilde{W}(\lambda_1)'\tilde{W}(\lambda_1))^{-1}\tilde{W}(\lambda_1)'$ with $\tilde{W}(\lambda_1) = [W_b - W_b(\lambda_1), W_b(\lambda_1), W_f]$. Note that D_T adjusts the convergence order as well as transforms the I(1) regressors $z_{b,t}$ in $R(\lambda_2,\lambda_1)$ to $\tilde{z}_{b,t} = \Omega_{bb}^{-1/2}z_{b,t}$ and

$$\frac{1}{\sqrt{T}}\tilde{z}_{b,[rT]} = \frac{1}{\sqrt{T}}\Omega_{bb}^{-1/2}z_{b,t} \Rightarrow B_b(r) \quad \text{uniformly over} \quad 0 \le r \le 1,$$

meaning that the resulting distribution becomes free of the nuisance parameters.

Let us consider the case where $\lambda_2 < \lambda_1$. We first note that

$$D_T'R(\lambda_2,\lambda_1)'M_{\tilde{w}_1}R(\lambda_2,\lambda_1)D_T$$
$$= D_T'R(\lambda_2,\lambda_1)'R(\lambda_2,\lambda_1)D_T$$
$$- D_T'R(\lambda_2,\lambda_1)'\tilde{W}(\lambda_1)\tilde{D}_T\left(\tilde{D}_T'\tilde{W}(\lambda_1)'\tilde{W}(\lambda_1)\tilde{D}_T\right)^{-1}\tilde{D}_T'\tilde{W}(\lambda_1)'R(\lambda_2,\lambda_1)D_T,$$

$$\text{where} \quad \tilde{D}_T = \begin{bmatrix} D_T & 0 & \begin{matrix} 0 \\ -\frac{1}{T}\Omega_{bb}^{-1}\Omega_{bf}\Omega_{f\cdot b}^{-1/2} \end{matrix} & 0 \\ 0 & D_T & \begin{matrix} 0 \\ -\frac{1}{T}\Omega_{bb}^{-1}\Omega_{bf}\Omega_{f\cdot b}^{-1/2} \end{matrix} & 0 \\ 0 & 0 & \frac{1}{T}\Omega_{f\cdot b}^{-1/2} & 0 \\ 0 & 0 & 0 & T^{-1/2}I_{p_f^x} \end{bmatrix}$$

with $\Omega_{f\cdot b} = \Omega_{ff} - \Omega_{fb}\Omega_{bb}^{-1}\Omega_{bf}$ as defined before Theorem 1. Note that the matrix $\tilde{D}_T$ transforms $z_{f,[rT]}$ in $\tilde{W}(\lambda_1)$ to $\tilde{z}_{f\cdot b,t} = \Omega_{f\cdot b}^{-1/2}(z_{f,t} - \Omega_{fb}\Omega_{bb}^{-1}z_{b,t})$ and

$$\frac{1}{\sqrt{T}}\tilde{z}_{f\cdot b,[rT]} = \frac{1}{\sqrt{T}}\Omega_{f\cdot b}^{-1/2}z_{f,[rT]} - \frac{1}{T}\Omega_{f\cdot b}^{-1/2}\Omega_{fb}\Omega_{bb}^{-1}z_{b,[rT]} \Rightarrow B_{f\cdot b}(r)$$

uniformly over $0 \le r \le 1$, which is independent of $B_b(r)$ and $B_u(r)$.

Then, under Assumption 1, we can see that

$$D_T' R(\lambda_2, \lambda_1)' R(\lambda_2, \lambda_1) D_T$$

$$= D_T' \sum_{t=T_2+1}^{T_1} w_{b,t} w_{b,t}' D_T$$

$$= \begin{bmatrix} \frac{T_1-T_2}{T} & \frac{1}{T^2}\sum_{t=T_2+1}^{T_1} t & \frac{1}{T\sqrt{T}}\sum_{t=T_2+1}^{T_1} \tilde{z}_{b,t}' \\ \frac{1}{T^2}\sum_{t=T_2+1}^{T_1} t & \frac{1}{T^3}\sum_{t=T_2+1}^{T_1} t^2 & \frac{1}{T^2\sqrt{T}}\sum_{t=T_2+1}^{T_1} t\tilde{z}_{b,t}' \\ \frac{1}{T\sqrt{T}}\sum_{t=T_2+1}^{T_1} \tilde{z}_{b,t} & \frac{1}{T^2\sqrt{T}}\sum_{t=T_2+1}^{T_1} t\tilde{z}_{b,t} & \frac{1}{T^2}\sum_{t=T_2+1}^{T_1} \tilde{z}_{b,t}\tilde{z}_{b,t}' \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \lambda_1 - \lambda_2 & \frac{\lambda_1^2-\lambda_2^2}{2} & \int_{\lambda_2}^{\lambda_1} B_b(r)' dr \\ \frac{\lambda_1^2-\lambda_2^2}{2} & \frac{\lambda_1^3-\lambda_2^3}{3} & \int_{\lambda_2}^{\lambda_1} r B_b(r)' dr \\ \int_{\lambda_2}^{\lambda_1} B_b(r) dr & \int_{\lambda_2}^{\lambda_1} r B_b(r) dr & \int_{\lambda_2}^{\lambda_1} B_b(r) B_b(r)' dr \end{bmatrix}$$

$$= \int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r)' dr, \qquad\qquad (A.2)$$

where $\tilde{B}_b(r) = [1, r, B_b(r)']'$. In the same manner, we have

$$D_T' R(\lambda_2, \lambda_1)' \tilde{W}(\lambda_1) \tilde{D}_T$$

$$= D_T' \sum_{t=T_2+1}^{T_1} w_{b,t} \tilde{w}_t(\lambda_1)' \tilde{D}_T$$

$$= \left[\begin{matrix} \frac{T_1-T_2}{T} & \frac{1}{T^2}\sum_{t=T_2+1}^{T_1} t & \frac{1}{T\sqrt{T}}\sum_{t=T_2+1}^{T_1} \tilde{z}_{b,t}' \\ \frac{1}{T^2}\sum_{t=T_2+1}^{T_1} t & \frac{1}{T^3}\sum_{t=T_2+1}^{T_1} t^2 & \frac{1}{T^2\sqrt{T}}\sum_{t=T_2+1}^{T_1} t\tilde{z}_{b,t}' \\ \frac{1}{T\sqrt{T}}\sum_{t=T_2+1}^{T_1} \tilde{z}_{b,t} & \frac{1}{T^2}\sum_{t=T_2+1}^{T_1} t\tilde{z}_{b,t} & \frac{1}{T^2}\sum_{t=T_2+1}^{T_1} \tilde{z}_{b,t}\tilde{z}_{b,t}' \end{matrix} \right.$$

$$\left. \begin{matrix} 0 & 0 & 0 & \frac{1}{T\sqrt{T}}\sum_{t=T_2+1}^{T_1} \tilde{z}_{f\cdot b,t}' & \frac{1}{T}\sum_{t=T_2+1}^{T_1} x_{f,t}' \\ 0 & 0 & 0 & \frac{1}{T^2\sqrt{T}}\sum_{t=T_2+1}^{T_1} t\tilde{z}_{f\cdot b,t}' & \frac{1}{T^2}\sum_{t=T_2+1}^{T_1} t x_{f,t}' \\ 0 & 0 & 0 & \frac{1}{T^2}\sum_{t=T_2+1}^{T_1} \tilde{z}_{b,f}\tilde{z}_{f\cdot b,t}' & \frac{1}{T\sqrt{T}}\sum_{t=T_2+1}^{T_1} \tilde{z}_{b,t} x_{f,t}' \end{matrix} \right]$$

$$\Rightarrow \left[\int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r)' dr, 0, \int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) B_{f\cdot b}(r)' dr, 0 \right] = [J_1(\lambda_1, \lambda_2), 0], \qquad (A.3)$$

and

$$\tilde{D}_T' \tilde{W}(\lambda_1)' \tilde{W}(\lambda_1) \tilde{D}_T$$

$$= \tilde{D}_T' \sum_{t=1}^{T} \tilde{w}_t(\lambda_1) \tilde{w}_t(\lambda_1)' \tilde{D}_T$$

$$\Rightarrow \begin{bmatrix} \int_0^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r)' dr & 0 & \int_0^{\lambda_1} \tilde{B}_b(r) B_{f\cdot b}(r)' dr & 0 \\ 0 & \int_{\lambda_1}^{1} \tilde{B}_b(r) \tilde{B}_b(r)' dr & \int_{\lambda_1}^{1} \tilde{B}_b(r) B_{f\cdot b}(r)' dr & 0 \\ \int_0^{\lambda_1} B_{f\cdot b}(r) \tilde{B}_b(r)' dr & \int_{\lambda_1}^{1} B_{f\cdot b}(r) \tilde{B}_b(r)' dr & \int_0^{1} B_{f\cdot b}(r) B_{f\cdot b}(r)' dr & 0 \\ 0 & 0 & 0 & \Sigma_x \end{bmatrix}$$

3

$$= \begin{bmatrix} K(\lambda_1) & 0 \\ 0 & \Sigma_x \end{bmatrix}. \tag{A.4}$$

Then, from (A.2)–(A.4), we have

$$D_T' R(\lambda_2, \lambda_1)' M_{\tilde{w}_1} R(\lambda_2, \lambda_1) D_T \Rightarrow \int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) \tilde{B}_b(r)' dr - J_1(\lambda_1, \lambda_2)' K(\lambda_1)^{-1} J_1(\lambda_1, \lambda_2). \tag{A.5}$$

Next, since $y = W(\lambda_1)\beta + u$ under the null hypothesis of $\lambda_0 = \lambda_1$, as can be seen in (13), we have

$$D_T' R(\lambda_2, \lambda_1)' M_{\tilde{w}_1} y$$
$$= D_T' R(\lambda_2, \lambda_1)' u - D_T' R(\lambda_2, \lambda_1)' \tilde{W}(\lambda_1) \tilde{D}_T \left(\tilde{D}_T' \tilde{W}(\lambda_1)' \tilde{W}(\lambda_1) \tilde{D}_T \right)^{-1} \tilde{D}_T' \tilde{W}(\lambda_1)' u.$$

Again, under Assumption 1, we have

$$D_T' R(\lambda_2, \lambda_1)' u = D_T' \sum_{t=T_2+1}^{T_1} w_{b,t} u_t$$
$$= \begin{bmatrix} \frac{1}{\sqrt{T}} \sum_{t=T_2+1}^{T_1} u_t \\ \frac{1}{T\sqrt{T}} \sum_{t=T_2+1}^{T_1} t u_t \\ \frac{1}{T} \sum_{t=T_2+1}^{T_1} \tilde{z}_{b,t} u_t \end{bmatrix}$$
$$\Rightarrow \omega_{uu}^{1/2} \int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) dB_u(r), \tag{A.6}$$

and

$$\tilde{D}_T' \tilde{W}(\lambda_1)' u = \tilde{D}_T' \sum_{t=1}^{T} \tilde{w}_t(\lambda_1) u_t$$
$$\Rightarrow \begin{bmatrix} \omega_{uu}^{1/2} \int_0^{\lambda_1} \tilde{B}_b(r) dB_u(r) \\ \omega_{uu}^{1/2} \int_{\lambda_1}^1 \tilde{B}_b(r) dB_u(r) \\ \omega_{uu}^{1/2} \int_0^1 B_{f \cdot b}(r) dB_u(r) \\ O_p(1) \end{bmatrix} = \begin{bmatrix} \omega_{uu}^{1/2} L(\lambda_1) \\ O_p(1) \end{bmatrix}. \tag{A.7}$$

From (A.3), (A.4), (A.6), and (A.7), we obtain

$$D_T' R(\lambda_2, \lambda_1)' M_{\tilde{w}_1} y \Rightarrow \omega_{uu}^{1/2} \left(\int_{\lambda_2}^{\lambda_1} \tilde{B}_b(r) dB_u(r) - J_1(\lambda_1, \lambda_2) K(\lambda_1)^{-1} L(\lambda_1) \right). \tag{A.8}$$

The limiting distribution of the test statistic is obtained from (A.5) and (A.8) for $\lambda_2 < \lambda_1$. The result for $\lambda_2 > \lambda_1$ is obtained similarly.

The theorem for the other models is proved in the same manner.∎

4

A.3. General Leads and Lags

In the main part of the article, we assumed that the leads-and-lags of the first differences of the I(1) variables are finite as given in (2). However, they are infinite in general, as pointed out in Remark 1, and thus in this appendix, we relax this assumption.

We start with the original regression model given by

$$y_t = w'_{b,t}\beta_b + w_{b,t}(\lambda_0)'\delta_b + w'_{f,t}\beta_f + e_t, \tag{A.9}$$

$$\Delta z_{b,t} = u^z_{b,t} \quad \text{and} \quad \Delta z_{f,t} = u^z_{f,t}.$$

Let $\xi_t = [e_t, u^{z\prime}_{b,t}, u^{z\prime}_{f,t}]'$ be a $(p^z_b + p^z_f + 1)$ vector. To allow for the infinite leads-and-lags expression, we make the following additional assumption as supposed in Saikkonen (1991) and Kejriwal and Perron (2008).

Assumption A.1 *(a) The process ξ_t is stationary with mean zero and the continuous spectral density matrix given by $f_{\xi\xi}(\lambda)$. (b) The spectral density matrix $f_{\xi\xi}(\lambda)$ is bounded away from zero so that $f_{\xi\xi}(\lambda) \geq \alpha I_{p^z_b + p^z_f + 1}$ for $\lambda \in [0, \pi]$ and some $\alpha > 0$. (c) The covariance function of ξ_t is absolutely summable; $\sum_{j=-\infty}^{\infty} \|E[\xi_t \xi_{t-j}]\| < \infty$ where $\|\cdot\|$ is the standard Euclidean norm. (d) The fourth order cumulants of ξ_t denoted by $\kappa_{ijk\ell}(m_1, m_2, m_3)$ satisfy*

$$\sum_{m_1, m_2, m_3 = -\infty}^{\infty} |\kappa_{ijkl}(m_1, m_2, m_3)| < \infty.$$

From Assumption A.1, we have (see also Brillinger, 1981),

$$e_t = \sum_{j=-\infty}^{\infty} \pi'_{b,j} u^z_{b,t-j} + \sum_{j=-\infty}^{\infty} \pi'_{f,j} u^z_{f,t-j} + u_t, \tag{A.10}$$

where u_t is a stationary process such that $E[u_t u^{z\prime}_{b,t-j}] = 0$ and $E[u_t u^{z\prime}_{f,t-j}] = 0$ for $j = 0, \pm 1, \pm 2, \cdots$. Inserting (A.10) into (A.9), we obtain

$$
\begin{aligned}
y_t &= w'_{b,t}\beta_b + w_{b,t}(\lambda_0)'\delta_b + w'_{f,t}\beta_f + \sum_{j=-\infty}^{\infty} \pi'_{b,j}\Delta z_{b,t-j} + \sum_{j=-\infty}^{\infty} \pi'_{f,j}\Delta z_{f,t-j} + u_t \\
&= w'_{b,t}\beta_b + w_{b,t}(\lambda_0)'\delta_b + w'_{f,t}\beta_f + \sum_{j=-l}^{l} \pi'_{b,j}\Delta z_{b,t-j} + \sum_{j=-l}^{l} \pi'_{f,j}\Delta z_{f,t-j} + u^*_t, \quad \text{(A.11)}
\end{aligned}
$$

where $u_t^* = u_t + v_t$ with $v_t = \sum_{|j|>l} \pi'_{b,j} \Delta z_{b,t-j} + \sum_{|j|>l} \pi'_{f,j} \Delta z_{f,t-j}$. In this case, the lag truncation parameter l is required to go to infinity as $T \to \infty$ and to satisfy the following condition (Kejriwal and Perron, 2008):

$$\frac{l^2}{T} \to 0, \tag{A.12}$$

$$l \sum_{|j|>l} (\|\pi_{b,j}\| + \|\pi_{f,j}\|) \to 0. \tag{A.13}$$

Note that (A.12) is the upper bound for l to go to infinity whereas (A.13) corresponds to the lower bound. As pointed out by Kejriwal and Perron (2008), when ξ_t is a finite order ARMA process, the leads-and-lags order l chosen by information criteria such as the AIC and BIC satisfy conditions (A.12) and (A.13).

We shall prove that both (A.5) and (A.8) holds under Assumption A.1, (A.12), and (A.13). Allowing for an abuse of notation, let $\tilde{w}_t^*(\lambda_1) = [(w_{b,t} - w_{b,t}(\lambda_1))', w_{b,t}(\lambda_1)', w'_{f,t}]'$ where $w_{f,t}$ does not include $x_{f,t}$, which consists of only the leads-and-lags of the first differences of the I(1) regressors. Accordingly, we partition $\tilde{D}_T$ into $diag\{\tilde{D}_T^*, T^{-1/2}I_{p_f^x}\}$. In this case, we have $\tilde{w}_t(\lambda_1) = [\tilde{w}_t^*(\lambda_1)', x'_{f,t}]'$. Then, we can write

$$D_T' R(\lambda_2, \lambda_1)' \tilde{W}(\lambda_1) \tilde{D}_T = \left[D_T' \sum_{t=T_2+1}^{T_1} w_{b,t} \tilde{w}_t^*(\lambda_1)' \tilde{D}_T^*, \ T^{-1/2} D_T' \sum_{t=T_2+1}^{T_1} w_{b,t} x'_{f,t} \right],$$

the first term of which converges to $J_1(\lambda_1, \lambda_2)$ as given in (A.3), while

$$\left\| T^{-1/2} D_T' \sum_{t=T_2+1}^{T_1} w_{b,t} x'_{f,t} \right\| = O_p\left(\frac{\sqrt{l}}{\sqrt{T}}\right) \tag{A.14}$$

using Lemma A.1(ii) of Kejriwal and Perron (2008). Similarly, we can write

$$\tilde{D}_T' \tilde{W}(\lambda_1)' \tilde{W}(\lambda_1) \tilde{D}_T = \left[\begin{array}{cc} \tilde{D}_T^{*\prime} \sum_{t=1}^T \tilde{w}_t^*(\lambda_1) \tilde{w}_t^*(\lambda_1)' \tilde{D}_T^* & T^{-1/2} \tilde{D}_T^{*\prime} \sum_{t=1}^T \tilde{w}_t^*(\lambda_1) x'_{f,t} \\ T^{-1/2} \sum_{t=1}^T x_{f,t} \tilde{w}_t^*(\lambda_1)' \tilde{D}_T^* & T^{-1} \sum_{t=1}^T x_{f,t} x'_{f,t} \end{array} \right],$$

the upper-left block converges to $K(\lambda_1)$ as given in (A.4), while

$$\left\| T^{-1/2} \tilde{D}_T^{*\prime} \sum_{t=1}^T \tilde{w}_t^*(\lambda_1) x'_{f,t} \right\| = O_p\left(\frac{\sqrt{l}}{\sqrt{T}}\right), \tag{A.15}$$

$$\left\| \frac{1}{T} \sum_{t=1}^T x_{f,t} x'_{f,t} \right\|_1 = O_p(1), \quad \left\| \left(\frac{1}{T} \sum_{t=1}^T x_{f,t} x'_{f,t} \right)^{-1} \right\|_1 = O_p(1), \tag{A.16}$$

6

using Lemma A.1(i) and (ii) of Kejriwal and Perron (2008). From (A.2)–(A.4) and (A.14)–(A.16), we have (A.5) because of (A.12).

Similarly, since $y = W(\lambda_1)\beta + u^*$ under the null hypothesis of $\lambda_0 = \lambda_1$, we have

$$
\begin{aligned}
& D_T' R(\lambda_2, \lambda_1)' M_{\tilde{w}_1} y \hspace{6cm} \text{(A.17)} \\
= \; & D_T' R(\lambda_2, \lambda_1)' u - D_T' R(\lambda_2, \lambda_1)' \tilde{W}(\lambda_1) \tilde{D}_T \left(\tilde{D}_T' \tilde{W}(\lambda_1)' \tilde{W}(\lambda_1) \tilde{D}_T \right)^{-1} \tilde{D}_T' \tilde{W}(\lambda_1)' u \\
& + D_T' R(\lambda_2, \lambda_1)' u - D_T' R(\lambda_2, \lambda_1)' \tilde{W}(\lambda_1) \tilde{D}_T \left(\tilde{D}_T' \tilde{W}(\lambda_1)' \tilde{W}(\lambda_1) \tilde{D}_T \right)^{-1} \tilde{D}_T' \tilde{W}(\lambda_1)' u.
\end{aligned}
$$

In the same way as the proof given in Appendix A.2, from (A.14)–(A.16) and using Lemma A.1(iii) of Kejriwal and Perron (2008), the first two terms of the right-hand side of (A.17) converge to (A.8). As a result, the rest we have to prove is that the last two terms of (A.17) converge to 0 in probability, which is proved using (A.14)–(A.16) and the following relations obtained using Lemma A.1(iv) and (v) of Kejriwal and Perron (2008),

$$
\left\| D_T' \sum_{t=T_2+1}^{T_1} w_{b,t} v_t \right\| = o_p \left(\frac{1}{l} \right)
$$

$$
\left\| \tilde{D}_T^{*\prime} \sum_{t=1}^{T} \tilde{w}_t^*(\lambda_1) v_t \right\| = o_p \left(\frac{1}{l} \right), \quad \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_{f,t} v_t \right\| = o_p \left(\frac{\sqrt{T}}{\sqrt{l}} \right). \quad \blacksquare
$$

References

[1] Brillinger, D. R. (1981). *Time Series Data Analysis and Theory*. Holden-Day, San Francisco.

[2] Kejriwal, M. and P. Perron (2008). Data Dependent Rules for Selection of the Number of Leads and Lags in the Dynamic OLS Cointegrating Regression. *Econometric Theory* 24, 1425-1441.

[3] Saikkonen, P. (1991). Asymptotically Efficient Estimation of Cointegration Regressions. *Econometric Theory* 7, 1-21.

Appendix B

Table B.1: Coefficients of the Response Surface Regressions (Model Ia)

	confidence level= 0.9				confidence level= 0.95			
	$p_b^z = 1$	2	3	4	$p_b^z = 1$	2	3	4
sup-LR_T								
a_0	695.025	230.321	611.599	446.265	527.107	194.143	614.973	295.263
a_{-1}	-682.721	-215.312	-594.541	-427.085	-513.026	-177.273	-596.134	-274.177
a_1	-677.973	-216.055	-591.452	-425.746	-511.464	-180.072	-593.118	-274.051
a_2	614.214	205.077	543.605	394.135	469.267	177.585	546.530	255.638
a_3	-359.150	-140.560	-332.187	-247.654	-280.410	-128.900	-335.239	-166.605
avg-LR_T								
a_0	-67.517	-4.742	38.664	-44.795	-25.396	14.564	-1.411	13.688
a_{-1}	70.655	9.170	-33.019	51.692	29.043	-9.570	7.700	-6.094
a_1	69.984	9.347	-32.545	50.884	28.021	-8.879	8.139	-6.433
a_2	-60.633	-7.170	32.583	-39.274	-16.550	10.989	-2.378	15.951
a_3	35.541	7.592	-16.432	20.054	5.967	-2.758	3.964	-11.637
exp-LR_T								
a_0	2.798	-12.572	53.978	36.773	97.372	102.603	1.783	129.459
a_{-1}	-0.123	16.283	-49.374	-31.246	-94.085	-98.245	3.528	-123.203
a_1	-0.384	15.638	-49.078	-31.513	-93.667	-97.790	3.760	-121.969
a_2	1.194	-13.246	44.108	30.038	84.168	88.665	-3.895	109.088
a_3	-2.076	5.542	-26.456	-21.076	-45.404	-50.997	0.252	-62.366

Table B.2: Coefficients of the Response Surface Regressions (Model Ib)

	confidence level= 0.9				confidence level= 0.95			
	$p_f^z = 1$	2	3	4	$p_f^z = 1$	2	3	4
sup-LR_T								
a_0	266.232	63.571	155.596	182.708	548.169	347.244	190.643	10.798
a_{-1}	-256.956	-54.017	-145.772	-172.655	-537.386	-336.214	-179.279	0.931
a_1	-256.999	-54.097	-144.358	-170.615	-534.243	-333.991	-177.364	0.950
a_2	242.167	53.478	131.757	154.259	486.989	306.444	163.143	1.629
a_3	-153.729	-41.001	-79.626	-91.230	-283.650	-182.508	-99.920	-7.094
avg-LR_T								
a_0	-28.100	-87.289	-44.889	-31.668	77.461	-53.186	-71.531	-89.269
a_{-1}	29.902	89.059	46.595	33.348	-75.256	55.330	73.620	91.285
a_1	29.896	88.601	46.847	33.434	-75.056	55.193	73.281	91.366
a_2	-29.054	-81.589	-45.638	-32.429	69.716	-51.809	-67.679	-87.243
a_3	21.657	48.817	30.398	21.538	-37.131	34.100	41.466	55.902
exp-LR_T								
a_0	14.082	5.506	-32.751	-46.655	146.433	33.369	-17.908	-102.067
a_{-1}	-12.478	-3.910	34.351	48.281	-144.355	-31.269	20.016	104.195
a_1	-12.745	-3.912	34.383	48.052	-143.521	-31.047	19.990	103.545
a_2	12.548	3.865	-31.986	-43.875	129.638	27.262	-17.976	-94.010
a_3	-7.824	-2.851	18.592	24.749	-71.139	-13.809	9.694	53.160

Table B.3: Coefficients of the Response Surface Regressions (Model Ic)

		confidence level= 0.9			confidence level= 0.95		
		$p_f^z = 1$	2	3	$p_f^z = 1$	2	3
$p_b^z = 1$	sup-LR_T						
	a_0	163.031	-70.596	66.972	37.957	-42.353	39.926
	a_{-1}	-150.319	83.405	-54.009	-23.612	56.782	-25.342
	a_1	-150.468	82.167	-54.751	-24.648	57.306	-23.862
	a_2	141.708	-67.818	56.836	30.489	-51.942	18.506
	a_3	-98.478	18.275	-51.103	-41.259	17.544	-18.796
	avg-LR_T						
	a_0	-26.579	-74.864	-36.055	-18.325	-155.061	-103.670
	a_{-1}	29.780	78.035	39.204	22.050	158.783	107.355
	a_1	29.494	77.806	39.254	21.893	157.819	107.142
	a_2	-24.988	-70.867	-36.294	-15.866	-140.760	-97.515
	a_3	15.644	43.709	24.280	9.116	82.143	60.743
	exp-LR_T						
	a_0	-23.472	-19.059	-61.708	-137.371	-136.612	-85.466
	a_{-1}	26.227	21.788	64.464	140.757	139.991	88.873
	a_1	25.873	22.148	64.148	139.533	139.353	88.336
	a_2	-23.011	-22.124	-58.419	-125.736	-128.447	-81.922
	a_3	11.256	13.606	31.669	69.001	74.790	48.487
$p_b^z = 2$	sup-LR_T						
	a_0	303.330	152.064		103.021	9.541	
	a_{-1}	-288.201	-136.822		-85.974	7.569	
	a_1	-285.274	-135.774		-86.839	6.744	
	a_2	257.835	125.711		87.238	-2.350	
	a_3	-161.720	-90.329		-73.026	-14.775	
	avg-LR_T						
	a_0	-29.874	-27.280		-43.158	-51.871	
	a_{-1}	34.407	31.809		48.310	57.023	
	a_1	33.970	31.594		48.051	56.689	
	a_2	-28.169	-28.281		-38.988	-48.233	
	a_3	17.654	20.746		23.029	30.478	
	exp-LR_T						
	a_0	-68.448	-46.246		-31.903	31.417	
	a_{-1}	72.220	50.010		36.376	-26.975	
	a_1	71.937	50.017		36.364	-26.210	
	a_2	-66.882	-47.770		-33.859	20.936	
	a_3	38.059	28.225		18.506	-10.155	
$p_b^z = 3$	sup-LR_T						
	a_0	394.751			313.344		
	a_{-1}	-377.388			-294.174		
	a_1	-375.966			-293.081		
	a_2	346.640			272.950		
	a_3	-217.638			-178.831		
	avg-LR_T						
	a_0	40.488			182.786		
	a_{-1}	-34.765			-176.403		
	a_1	-34.379			-174.802		
	a_2	33.773			164.069		
	a_3	-16.021			-90.843		
	exp-LR_T						
	a_0	21.183			-35.411		
	a_{-1}	-16.487			40.873		
	a_1	-16.666			39.931		
	a_2	15.812			-32.939		
	a_3	-11.680			12.955		

Table B.4: Coefficients of the Response Surface Regressions (Model IIa)

	confidence level= 0.9				confidence level= 0.95			
	$p_b^z = 1$	2	3	4	$p_b^z = 1$	2	3	4
sup-LR_T								
a_0	492.130	369.181	515.946	423.522	470.123	304.997	259.217	328.062
a_{-1}	-476.967	-351.765	-496.675	-402.336	-453.083	-285.782	-238.126	-304.848
a_1	-474.993	-350.083	-493.984	-400.520	-451.868	-284.049	-236.043	-303.513
a_2	439.066	317.486	451.864	368.371	414.827	258.147	214.645	279.579
a_3	-273.010	-192.033	-275.223	-230.669	-252.008	-161.675	-138.459	-178.937
avg-LR_T								
a_0	-46.922	-1.169	36.556	90.531	4.796	-13.991	105.999	108.377
a_{-1}	51.387	6.836	-29.658	-82.435	0.249	20.297	-98.391	-99.491
a_1	50.572	7.308	-29.493	-82.117	0.210	21.355	-97.361	-99.343
a_2	-40.291	-4.427	31.347	80.290	6.116	-17.640	93.101	99.859
a_3	21.002	5.001	-16.723	-44.961	-2.391	16.163	-47.740	-56.396
exp-LR_T								
a_0	40.067	103.672	156.828	186.745	98.406	99.830	72.113	11.858
a_{-1}	-36.364	-99.066	-151.349	-180.362	-94.030	-94.459	-65.837	-4.608
a_1	-36.349	-97.902	-149.611	-179.484	-93.976	-93.386	-64.653	-5.442
a_2	32.965	85.377	131.562	163.393	87.219	81.303	54.872	8.838
a_3	-19.352	-45.650	-71.917	-95.791	-52.039	-43.824	-28.958	-11.955

Table B.5: Coefficients of the Response Surface Regressions (Model IIb)

	confidence level= 0.9				confidence level= 0.95			
	$p_f^z = 1$	2	3	4	$p_f^z = 1$	2	3	4
sup-LR_T								
a_0	328.702	334.786	346.856	257.698	362.135	396.644	311.699	452.329
a_{-1}	−315.765	−321.771	−333.813	−244.557	−347.580	−381.994	−297.061	−437.608
a_1	−313.389	−318.486	−329.606	−241.258	−342.930	−377.776	−292.785	−431.854
a_2	287.581	288.122	292.281	212.770	302.797	336.200	256.617	381.415
a_3	−180.414	−175.235	−167.344	−121.852	−173.933	−192.287	−143.199	−212.898
avg-LR_T								
a_0	−47.527	26.397	16.797	15.562	−33.608	20.189	90.073	107.529
a_{-1}	50.677	−23.232	−13.645	−12.400	37.283	−16.484	−86.364	−103.796
a_1	50.875	−22.646	−12.646	−11.601	38.181	−15.249	−84.382	−102.383
a_2	−46.729	19.992	7.879	7.609	−34.755	12.295	72.716	91.846
a_3	30.778	−7.188	3.431	2.434	24.880	−0.170	−31.465	−45.685
exp-LR_T								
a_0	−42.071	23.915	76.243	132.601	−128.274	26.718	82.005	96.941
a_{-1}	44.739	−21.250	−73.595	−129.919	131.555	−23.446	−78.784	−93.699
a_1	44.922	−20.306	−71.722	−128.177	131.258	−22.277	−76.734	−91.967
a_2	−42.682	15.507	58.774	111.877	−121.773	15.717	63.289	78.577
a_3	25.522	−7.076	−27.017	−58.853	72.197	−3.896	−28.854	−38.897

Table B.6: Coefficients of the Response Surface Regressions (Model IIc)

		confidence level= 0.9			confidence level= 0.95		
		$p_f^z = 1$	2	3	$p_f^z = 1$	2	3
$p_b^z = 1$	sup-LR_T						
	a_0	335.274	348.389	451.206	128.014	524.278	144.918
	a_{-1}	-320.020	-333.177	-435.934	-111.019	-507.412	-127.889
	a_1	-318.772	-331.286	-433.120	-110.203	-502.047	-127.067
	a_2	297.613	305.824	397.958	103.123	449.871	119.633
	a_3	-192.373	-190.989	-243.560	-75.545	-259.791	-85.093
	avg-LR_T						
	a_0	48.859	77.250	5.872	6.978	-48.210	37.485
	a_{-1}	-44.417	-72.837	-1.431	-1.886	53.271	-32.442
	a_1	-43.858	-71.725	-1.230	-1.829	53.587	-31.341
	a_2	41.912	65.283	2.668	6.678	-46.686	29.331
	a_3	-21.831	-32.141	1.858	-1.975	32.086	-9.949
	exp-LR_T						
	a_0	35.927	32.331	11.703	110.311	126.707	126.436
	a_{-1}	-32.253	-28.693	-8.058	-105.973	-122.412	-122.153
	a_1	-31.709	-27.888	-7.561	-104.863	-120.880	-120.698
	a_2	27.280	22.968	5.399	94.251	106.734	108.314
	a_3	-14.855	-11.444	-2.367	-53.322	-56.984	-60.677
$p_b^z = 2$	sup-LR_T						
	a_0	468.094	205.131		211.642	45.048	
	a_{-1}	-450.681	-187.666		-192.428	-25.757	
	a_1	-447.824	-186.945		-191.186	-25.702	
	a_2	405.279	170.669		174.401	24.190	
	a_3	-240.676	-109.431		-112.630	-27.104	
	avg-LR_T						
	a_0	51.807	75.889		147.598	47.473	
	a_{-1}	-46.117	-70.191		-141.296	-41.090	
	a_1	-45.555	-69.196		-138.779	-40.375	
	a_2	44.600	63.949		125.323	41.374	
	a_3	-24.737	-32.509		-62.270	-21.834	
	exp-LR_T						
	a_0	73.477	44.888		119.248	121.963	
	a_{-1}	-68.866	-40.253		-113.902	-116.606	
	a_1	-68.134	-39.931		-112.207	-115.649	
	a_2	60.564	35.440		96.830	105.056	
	a_3	-34.944	-20.800		-50.816	-62.001	
$p_b^z = 3$	sup-LR_T						
	a_0	385.532			335.216		
	a_{-1}	-366.242			-314.067		
	a_1	-364.027			-311.014		
	a_2	333.115			279.456		
	a_3	-206.430			-169.314		
	avg-LR_T						
	a_0	64.721			-1.855		
	a_{-1}	-57.822			9.514		
	a_1	-57.221			9.304		
	a_2	55.161			-0.485		
	a_3	-28.746			0.244		
	exp-LR_T						
	a_0	132.675			68.744		
	a_{-1}	-127.162			-62.447		
	a_1	-126.196			-61.845		
	a_2	113.289			55.648		
	a_3	-64.996			-33.249		

Table B.7: Coverage rates and lengths of the confidence sets (Model I-b)

d	Coverage rates				Lengths of the confidence sets			
	BLS	sup	avg	exp	BLS	sup	avg	exp
DGP1				$T = 100$				
4	0.886	0.952	0.934	0.934	0.151	0.216	0.229	0.189
8	0.978	0.965	0.947	0.947	0.053	0.058	0.075	0.055
12	0.996	0.969	0.949	0.951	0.031	0.038	0.046	0.036
16	1.000	0.970	0.950	0.953	0.030	0.028	0.035	0.027
DGP2								
4	0.875	0.950	0.930	0.929	0.150	0.226	0.239	0.198
8	0.975	0.965	0.945	0.945	0.053	0.059	0.077	0.056
12	0.996	0.968	0.948	0.950	0.031	0.038	0.047	0.036
16	1.000	0.970	0.949	0.951	0.030	0.028	0.035	0.027
DGP3								
4	0.687	0.870	0.810	0.806	0.283	0.421	0.361	0.349
8	0.882	0.928	0.877	0.884	0.114	0.147	0.149	0.125
12	0.964	0.963	0.914	0.925	0.062	0.069	0.081	0.062
16	0.991	0.975	0.925	0.940	0.043	0.048	0.055	0.044
DGP4								
4	0.912	0.950	0.933	0.929	0.120	0.166	0.189	0.149
8	0.984	0.965	0.947	0.945	0.045	0.050	0.064	0.049
12	0.999	0.968	0.948	0.949	0.030	0.035	0.042	0.034
16	1.000	0.967	0.947	0.948	0.030	0.029	0.033	0.028
DGP5								
4	0.666	0.838	0.794	0.776	0.221	0.352	0.309	0.288
8	0.890	0.921	0.869	0.873	0.086	0.115	0.123	0.098
12	0.973	0.960	0.903	0.917	0.049	0.056	0.066	0.051
16	0.996	0.969	0.914	0.928	0.035	0.042	0.047	0.038
DGP1				$T = 300$				
4	0.925	0.954	0.954	0.949	0.087	0.118	0.158	0.119
8	0.971	0.958	0.957	0.952	0.026	0.043	0.057	0.044
12	0.990	0.959	0.958	0.954	0.016	0.029	0.038	0.029
16	0.995	0.959	0.958	0.954	0.010	0.021	0.028	0.022
DGP2								
4	0.925	0.953	0.953	0.949	0.087	0.120	0.160	0.121
8	0.969	0.958	0.957	0.952	0.026	0.043	0.057	0.044
12	0.990	0.959	0.957	0.953	0.016	0.029	0.038	0.029
16	0.995	0.960	0.957	0.954	0.010	0.022	0.028	0.022
DGP3								
4	0.776	0.924	0.901	0.898	0.236	0.380	0.351	0.332
8	0.918	0.948	0.923	0.924	0.074	0.100	0.125	0.095
12	0.961	0.958	0.933	0.937	0.036	0.053	0.068	0.052
16	0.985	0.962	0.937	0.941	0.023	0.039	0.049	0.039
DGP4								
4	0.936	0.953	0.949	0.946	0.066	0.084	0.119	0.086
8	0.978	0.957	0.952	0.950	0.021	0.037	0.049	0.038
12	0.991	0.957	0.952	0.951	0.013	0.025	0.032	0.025
16	0.999	0.958	0.952	0.951	0.010	0.019	0.024	0.019
DGP5								
4	0.788	0.914	0.894	0.885	0.185	0.297	0.286	0.259
8	0.914	0.944	0.918	0.920	0.055	0.074	0.096	0.072
12	0.965	0.955	0.927	0.932	0.028	0.045	0.056	0.044
16	0.989	0.958	0.930	0.937	0.019	0.034	0.041	0.033

Table B.8: Coverage rates and lengths of the confidence sets (Model II-a)

d		Coverage rates				Lengths of the confidence sets		
	BLS	sup	avg	exp	BLS	sup	avg	exp
DGP1				$T = 100$				
4	0.956	0.956	0.915	0.911	0.047	0.082	0.081	0.063
8	0.996	0.962	0.923	0.920	0.032	0.020	0.028	0.016
12	0.999	0.963	0.923	0.920	0.030	0.012	0.019	0.011
16	1.000	0.963	0.923	0.920	0.030	0.010	0.015	0.010
DGP2								
4	0.879	0.943	0.892	0.894	0.061	0.144	0.130	0.114
8	0.960	0.956	0.911	0.911	0.045	0.050	0.050	0.040
12	0.979	0.958	0.916	0.915	0.041	0.028	0.030	0.023
16	0.990	0.959	0.917	0.916	0.039	0.020	0.023	0.018
DGP3								
4	0.815	0.923	0.831	0.856	0.080	0.193	0.155	0.141
8	0.982	0.983	0.925	0.938	0.038	0.047	0.048	0.035
12	0.996	0.987	0.933	0.945	0.032	0.021	0.028	0.017
16	0.998	0.988	0.934	0.946	0.031	0.014	0.021	0.012
DGP4								
4	0.968	0.949	0.904	0.899	0.040	0.064	0.070	0.049
8	0.997	0.954	0.909	0.905	0.031	0.017	0.028	0.014
12	0.999	0.954	0.909	0.905	0.030	0.011	0.019	0.010
16	1.000	0.954	0.910	0.906	0.030	0.010	0.015	0.009
DGP5								
4	0.772	0.869	0.737	0.767	0.070	0.168	0.129	0.119
8	0.931	0.942	0.819	0.853	0.049	0.068	0.057	0.048
12	0.966	0.957	0.837	0.871	0.044	0.038	0.035	0.028
16	0.981	0.962	0.845	0.878	0.042	0.027	0.026	0.021
DGP1				$T = 300$				
4	0.967	0.962	0.939	0.948	0.024	0.044	0.051	0.039
8	0.995	0.963	0.941	0.950	0.012	0.014	0.018	0.013
12	0.998	0.963	0.941	0.950	0.011	0.008	0.012	0.008
16	1.000	0.963	0.941	0.950	0.010	0.006	0.010	0.006
DGP2								
4	0.910	0.959	0.936	0.941	0.038	0.098	0.096	0.086
8	0.968	0.962	0.939	0.945	0.024	0.033	0.035	0.029
12	0.983	0.963	0.939	0.945	0.020	0.018	0.021	0.017
16	0.988	0.963	0.939	0.946	0.017	0.012	0.016	0.012
DGP3								
4	0.855	0.950	0.899	0.915	0.056	0.168	0.151	0.133
8	0.975	0.974	0.927	0.944	0.020	0.034	0.035	0.029
12	0.995	0.977	0.931	0.947	0.013	0.018	0.020	0.016
16	0.997	0.977	0.931	0.947	0.011	0.012	0.015	0.011
DGP4								
4	0.979	0.962	0.939	0.943	0.020	0.035	0.040	0.031
8	0.997	0.962	0.939	0.944	0.011	0.012	0.017	0.011
12	0.999	0.962	0.939	0.944	0.010	0.007	0.012	0.006
16	0.999	0.962	0.940	0.944	0.010	0.005	0.009	0.005
DGP5								
4	0.815	0.934	0.865	0.885	0.055	0.166	0.142	0.136
8	0.934	0.959	0.898	0.916	0.032	0.060	0.054	0.049
12	0.971	0.966	0.908	0.926	0.024	0.032	0.031	0.027
16	0.980	0.967	0.910	0.929	0.022	0.021	0.022	0.018

Table B.9: Coverage rates and lengths of the confidence sets (Model II-b)

d	Coverage rates				Lengths of the confidence sets			
	BLS	sup	avg	exp	BLS	sup	avg	exp
DGP1				$T = 100$				
4	0.987	0.975	0.927	0.935	0.040	0.041	0.049	0.034
8	1.000	0.978	0.930	0.939	0.030	0.016	0.022	0.014
12	1.000	0.978	0.930	0.939	0.030	0.010	0.016	0.010
16	1.000	0.978	0.930	0.939	0.030	0.010	0.012	0.009
DGP2								
4	0.986	0.973	0.922	0.931	0.040	0.042	0.053	0.035
8	1.000	0.976	0.924	0.935	0.030	0.016	0.022	0.014
12	1.000	0.976	0.924	0.935	0.030	0.010	0.016	0.010
16	1.000	0.976	0.924	0.935	0.030	0.010	0.012	0.009
DGP3								
4	0.866	0.925	0.815	0.847	0.073	0.157	0.128	0.104
8	0.997	0.983	0.885	0.923	0.032	0.029	0.028	0.023
12	1.000	0.984	0.887	0.923	0.030	0.016	0.020	0.013
16	1.000	0.984	0.887	0.923	0.030	0.011	0.015	0.010
DGP4								
4	0.991	0.970	0.925	0.930	0.033	0.036	0.044	0.029
8	1.000	0.970	0.926	0.932	0.030	0.013	0.022	0.011
12	1.000	0.970	0.926	0.932	0.030	0.010	0.014	0.009
16	1.000	0.970	0.926	0.932	0.030	0.010	0.011	0.009
DGP5								
4	0.885	0.917	0.786	0.823	0.059	0.113	0.096	0.076
8	1.000	0.975	0.837	0.880	0.030	0.025	0.028	0.019
12	1.000	0.975	0.837	0.880	0.030	0.012	0.018	0.010
16	1.000	0.975	0.837	0.880	0.030	0.010	0.013	0.009
DGP1				$T = 300$				
4	0.979	0.960	0.940	0.945	0.019	0.024	0.030	0.022
8	0.999	0.962	0.942	0.947	0.010	0.011	0.013	0.010
12	1.000	0.962	0.942	0.947	0.010	0.008	0.009	0.007
16	1.000	0.962	0.942	0.947	0.010	0.005	0.008	0.005
DGP2								
4	0.978	0.960	0.939	0.944	0.019	0.024	0.031	0.022
8	0.999	0.961	0.941	0.945	0.010	0.011	0.013	0.010
12	1.000	0.961	0.941	0.946	0.010	0.008	0.009	0.007
16	1.000	0.961	0.941	0.946	0.010	0.005	0.008	0.005
DGP3								
4	0.901	0.943	0.889	0.902	0.049	0.108	0.111	0.081
8	0.994	0.965	0.911	0.927	0.017	0.021	0.023	0.018
12	1.000	0.965	0.913	0.928	0.010	0.013	0.014	0.012
16	1.000	0.965	0.913	0.928	0.010	0.010	0.011	0.009
DGP4								
4	0.989	0.960	0.939	0.942	0.016	0.020	0.025	0.019
8	1.000	0.960	0.940	0.942	0.010	0.010	0.013	0.010
12	1.000	0.960	0.940	0.942	0.010	0.007	0.010	0.006
16	1.000	0.960	0.940	0.942	0.010	0.004	0.008	0.004
DGP5								
4	0.919	0.946	0.886	0.901	0.037	0.067	0.076	0.052
8	0.996	0.965	0.905	0.922	0.014	0.018	0.020	0.016
12	1.000	0.965	0.904	0.922	0.010	0.012	0.014	0.011
16	1.000	0.965	0.905	0.922	0.010	0.009	0.012	0.008

Table B.10: Effect of the Location of the Break Point (Model I-a, DGP5)

	Coverage rates			Lengths		
d	$\lambda_0 = 0.3$	$\lambda_0 = 0.5$	$\lambda_0 = 0.7$	$\lambda_0 = 0.3$	$\lambda_0 = 0.5$	$\lambda_0 = 0.7$
BLS			$T = 100$			
4	0.648	0.691	0.705	0.184	0.175	0.165
8	0.844	0.847	0.866	0.130	0.126	0.111
12	0.911	0.917	0.921	0.106	0.106	0.088
16	0.948	0.952	0.954	0.093	0.093	0.080
			$T = 300$			
4	0.722	0.756	0.760	0.168	0.164	0.150
8	0.860	0.889	0.882	0.105	0.106	0.092
12	0.915	0.931	0.925	0.081	0.082	0.069
16	0.940	0.949	0.948	0.067	0.066	0.058
sup			$T = 100$			
4	0.839	0.859	0.862	0.323	0.285	0.276
8	0.915	0.921	0.925	0.161	0.147	0.142
12	0.942	0.949	0.951	0.099	0.094	0.089
16	0.956	0.960	0.961	0.071	0.069	0.063
			$T = 300$			
4	0.916	0.926	0.923	0.316	0.276	0.276
8	0.944	0.949	0.947	0.136	0.123	0.127
12	0.953	0.958	0.956	0.080	0.074	0.075
16	0.960	0.962	0.961	0.056	0.052	0.051
avg			$T = 100$			
4	0.754	0.750	0.781	0.265	0.229	0.230
8	0.840	0.827	0.851	0.136	0.121	0.124
12	0.867	0.861	0.876	0.092	0.082	0.084
16	0.883	0.879	0.889	0.072	0.064	0.065
			$T = 300$			
4	0.874	0.874	0.882	0.284	0.240	0.252
8	0.902	0.901	0.907	0.128	0.112	0.122
12	0.914	0.909	0.918	0.082	0.071	0.076
16	0.919	0.915	0.922	0.061	0.052	0.056
exp			$T = 100$			
4	0.752	0.774	0.786	0.257	0.227	0.221
8	0.848	0.853	0.861	0.127	0.116	0.113
12	0.885	0.886	0.890	0.081	0.076	0.072
16	0.903	0.904	0.906	0.059	0.056	0.052
			$T = 300$			
4	0.877	0.884	0.887	0.271	0.235	0.239
8	0.907	0.916	0.915	0.115	0.105	0.109
12	0.920	0.925	0.926	0.070	0.065	0.065
16	0.926	0.929	0.932	0.050	0.046	0.046

Table B.11: Effect of the Location of the Break Point (Model I-b, DGP5)

	Coverage rates			Lengths		
d	$\lambda_0 = 0.3$	$\lambda_0 = 0.5$	$\lambda_0 = 0.7$	$\lambda_0 = 0.3$	$\lambda_0 = 0.5$	$\lambda_0 = 0.7$
BLS			$T = 100$			
4	0.657	0.666	0.665	0.220	0.221	0.221
8	0.892	0.890	0.886	0.086	0.086	0.087
12	0.975	0.973	0.976	0.049	0.049	0.049
16	0.995	0.996	0.997	0.035	0.035	0.035
			$T = 300$			
4	0.778	0.788	0.779	0.185	0.185	0.187
8	0.919	0.914	0.923	0.056	0.055	0.056
12	0.971	0.965	0.968	0.028	0.028	0.028
16	0.990	0.989	0.989	0.019	0.019	0.019
sup			$T = 100$			
4	0.846	0.838	0.851	0.359	0.352	0.362
8	0.924	0.921	0.927	0.112	0.115	0.116
12	0.961	0.960	0.959	0.056	0.056	0.057
16	0.968	0.969	0.968	0.041	0.042	0.042
			$T = 300$			
4	0.908	0.914	0.909	0.312	0.297	0.309
8	0.942	0.944	0.938	0.077	0.074	0.075
12	0.954	0.955	0.950	0.045	0.045	0.045
16	0.957	0.958	0.955	0.034	0.034	0.034
avg			$T = 100$			
4	0.801	0.794	0.793	0.319	0.309	0.318
8	0.872	0.869	0.874	0.123	0.123	0.125
12	0.907	0.903	0.908	0.067	0.066	0.068
16	0.915	0.914	0.915	0.047	0.047	0.048
			$T = 300$			
4	0.892	0.894	0.890	0.302	0.286	0.299
8	0.920	0.918	0.915	0.102	0.096	0.101
12	0.929	0.927	0.926	0.058	0.056	0.058
16	0.931	0.930	0.929	0.043	0.041	0.043
exp			$T = 100$			
4	0.782	0.776	0.777	0.298	0.288	0.298
8	0.869	0.873	0.876	0.096	0.098	0.100
12	0.910	0.917	0.915	0.051	0.051	0.052
16	0.916	0.928	0.926	0.038	0.038	0.038
			$T = 300$			
4	0.883	0.885	0.880	0.273	0.259	0.271
8	0.918	0.920	0.913	0.074	0.072	0.073
12	0.931	0.932	0.925	0.044	0.044	0.044
16	0.933	0.937	0.930	0.033	0.033	0.033

Table B.12: Effect of the Location of the Break Point (Model II-a, DGP5)

	Coverage rates			Lengths		
d	$\lambda_0 = 0.3$	$\lambda_0 = 0.5$	$\lambda_0 = 0.7$	$\lambda_0 = 0.3$	$\lambda_0 = 0.5$	$\lambda_0 = 0.7$
BLS			$T = 100$			
4	0.659	0.772	0.817	0.088	0.070	0.062
8	0.878	0.931	0.941	0.057	0.049	0.043
12	0.940	0.966	0.974	0.048	0.044	0.040
16	0.967	0.981	0.985	0.046	0.042	0.037
			$T = 300$			
4	0.696	0.815	0.850	0.072	0.055	0.046
8	0.896	0.934	0.951	0.039	0.032	0.025
12	0.946	0.971	0.973	0.031	0.024	0.021
16	0.966	0.980	0.982	0.026	0.022	0.019
sup			$T = 100$			
4	0.811	0.869	0.884	0.242	0.168	0.137
8	0.912	0.942	0.940	0.102	0.068	0.054
12	0.939	0.957	0.953	0.055	0.038	0.031
16	0.953	0.962	0.957	0.038	0.027	0.023
			$T = 300$			
4	0.904	0.934	0.936	0.265	0.166	0.138
8	0.950	0.959	0.957	0.096	0.060	0.051
12	0.961	0.966	0.964	0.050	0.032	0.027
16	0.965	0.967	0.965	0.032	0.021	0.017
avg			$T = 100$			
4	0.695	0.737	0.770	0.188	0.129	0.113
8	0.797	0.819	0.827	0.088	0.057	0.055
12	0.821	0.837	0.842	0.056	0.035	0.037
16	0.834	0.845	0.849	0.042	0.026	0.029
			$T = 300$			
4	0.843	0.865	0.884	0.241	0.142	0.133
8	0.893	0.898	0.908	0.094	0.054	0.051
12	0.904	0.908	0.913	0.051	0.031	0.030
16	0.910	0.910	0.915	0.035	0.022	0.022
exp			$T = 100$			
4	0.701	0.767	0.790	0.175	0.119	0.098
8	0.818	0.853	0.852	0.073	0.048	0.039
12	0.846	0.871	0.868	0.040	0.028	0.024
16	0.862	0.878	0.874	0.028	0.021	0.018
			$T = 300$			
4	0.850	0.885	0.891	0.220	0.136	0.115
8	0.905	0.916	0.916	0.079	0.049	0.042
12	0.917	0.926	0.922	0.041	0.027	0.022
16	0.922	0.929	0.924	0.027	0.018	0.015

Table B.13: Effect of the Location of the Break Point (Model II-b, DGP5)

d	Coverage rates			Lengths		
	$\lambda_0 = 0.3$	$\lambda_0 = 0.5$	$\lambda_0 = 0.7$	$\lambda_0 = 0.3$	$\lambda_0 = 0.5$	$\lambda_0 = 0.7$
BLS			$T = 100$			
4	0.695	0.885	0.952	0.081	0.059	0.046
8	0.987	1.000	1.000	0.035	0.030	0.030
12	1.000	1.000	1.000	0.030	0.030	0.030
16	1.000	1.000	1.000	0.030	0.030	0.030
			$T = 300$			
4	0.803	0.919	0.961	0.061	0.037	0.027
8	0.985	0.996	0.999	0.020	0.014	0.011
12	0.999	1.000	1.000	0.011	0.010	0.010
16	1.000	1.000	1.000	0.010	0.010	0.010
sup			$T = 100$			
4	0.829	0.917	0.943	0.246	0.113	0.073
8	0.960	0.975	0.967	0.042	0.025	0.018
12	0.967	0.975	0.968	0.020	0.012	0.010
16	0.967	0.975	0.968	0.012	0.010	0.010
			$T = 300$			
4	0.923	0.946	0.960	0.219	0.067	0.044
8	0.965	0.965	0.969	0.026	0.018	0.015
12	0.966	0.965	0.969	0.016	0.012	0.010
16	0.966	0.965	0.969	0.012	0.009	0.007
avg			$T = 100$			
4	0.735	0.786	0.845	0.202	0.096	0.077
8	0.858	0.837	0.867	0.048	0.028	0.028
12	0.863	0.837	0.867	0.030	0.018	0.020
16	0.863	0.837	0.867	0.024	0.013	0.015
			$T = 300$			
4	0.882	0.886	0.917	0.229	0.076	0.060
8	0.919	0.905	0.924	0.035	0.020	0.018
12	0.921	0.904	0.923	0.019	0.014	0.013
16	0.921	0.905	0.923	0.015	0.012	0.012
exp			$T = 100$			
4	0.726	0.823	0.866	0.170	0.076	0.052
8	0.876	0.880	0.892	0.031	0.019	0.014
12	0.883	0.880	0.892	0.016	0.010	0.009
16	0.882	0.880	0.892	0.010	0.009	0.009
			$T = 300$			
4	0.878	0.901	0.925	0.171	0.052	0.036
8	0.928	0.922	0.933	0.023	0.016	0.014
12	0.931	0.922	0.933	0.015	0.011	0.009
16	0.931	0.922	0.933	0.011	0.008	0.007

Appendix C

Table C.14: Empirical results for model (23) in which γ sustains a structural change

| series: y_t, p_t^{oil}		
estimated break date	:	2008Q3
cointegration test with a break	:	0.048
cointegration test without a break	:	0.172**
series: c_t, p_t^{oil}		
estimated break date	:	2008Q4
cointegration test with a break	:	0.033
cointegration test without a break	:	0.117*
series: i_t, p_t^{oil}		
estimated break date	:	2007Q1
cointegration test with a break	:	0.046
cointegration test without a break	:	0.131**

Note: the superscripts *, **, and *** signify significance at 10%, 5% and 1%, respectively.

(i) y_t
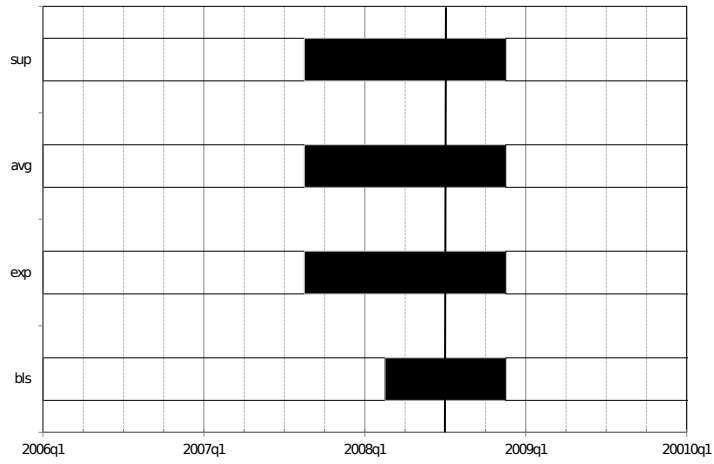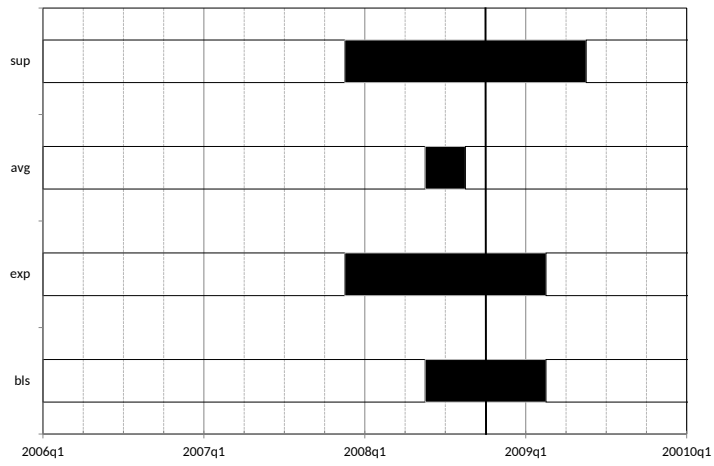
(ii) c_t
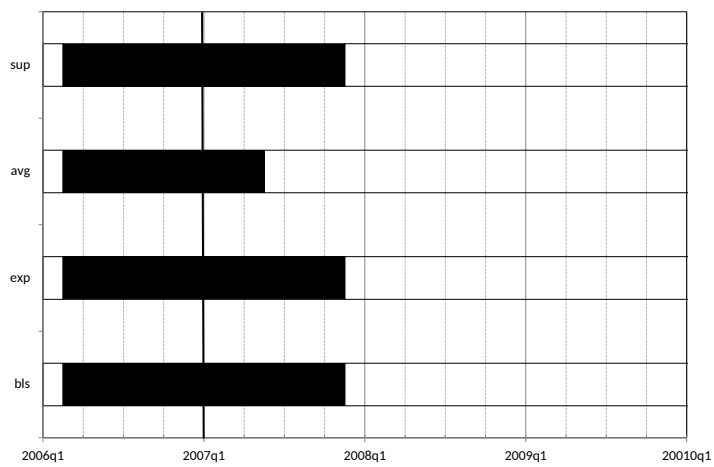
(iii) i_t

Figure C.1: Confidence sets for a break