Quasi-likelihood ratio tests for cointegration, cobreaking, and cotrending

Josep Lluís Carrion-i-Silvestre and Dukpa Kim

AQR-IREA research group, Department of Econometrics, Statistics and Applied Economics, Faculty of Economics and Business, University of Barcelona, Barcelona, Spain; Department of Economics, Korea University, Seoul, Korea

ABSTRACT

We consider a set of variables with two types of nonstationary features, stochastic trends and broken linear trends. We develop tests that can determine whether there is a linear combination of these variables under which the nonstationary features can be canceled out. The first test can determine whether stochastic trends can be eliminated and thus whether cointegration holds, regardless of whether structural breaks in linear trends are eliminated. The second test can determine whether both stochastic trends and breaks in linear trends are simultaneously removed and thus whether cointegration and cobreaking simultaneously hold. The third test can determine whether not only breaks in linear trends but also linear trends themselves are eliminated along with stochastic trends and thus whether both cointegration and cotrending hold.

KEYWORDS

Cobreaking; cointegration; cotrending; multiple structural breaks

JEL CLASSIFICATION

C22

1. Introduction

Early developments in the time series literature focus on distinguishing between a difference stationary process and a trend stationary process. In the recent unit root (UR) literature, one line of research focuses on the case in which a stochastic trend and a structural break (SB) in the deterministic function coexist. For example, see Perron (1989); Carrion-i-Silvestre et al. (2009); Harris et al. (2009); Harvey et al. (2012, 2013); and Kim and Perron (2009). One motivation for this type of research is that a stochastic trend and a SB in the deterministic function are separate objects, and thus, the existence of one implies nothing about the existence of the other. However, popular UR tests are developed assuming either the presence or absence of a SB. These UR tests do not lead to correct inference when the assumption regarding SBs fails. Thus, the goal in the recent UR literature is to devise a UR test robust to the presence or absence of SBs.

The issues analyzed in this paper are closely related to this recent UR literature. Suppose that a researcher finds, for example via the UR tests proposed by the aforementioned authors, that the variables under investigation have both a stochastic trend and a broken linear trend. Her goal is to test whether there is a stable relationship among those variables. Can she use any of the tests proposed in the literature for this purpose, e.g., cointegration (CI) tests allowing for a SB or SB tests for CI models? If not, what is the alternative? This paper intends to answer these questions.

Most methods proposed in the econometrics literature assume that variables have either a stochastic trend (CI models) or a SB (cobreaking or cotrending models), but not both. Even CI models with SBs mostly assume that SBs are exogenously given only to the CI equation and do not affect individual variables. However, we are interested in the case in which a SB in the CI equation exists because
SBs in individual variables do not cancel one another out. In Section 2, we discuss in greater detail what statistical issues arise when existing tests are applied to our problem. Additionally, see Section 5 for the Monte Carlo simulation results supporting our argument.

The traditional definition of CI says that a set of trending variables is stochastically cointegrated if a certain linear combination of the variables yields a stationary process around a linear trend and that it is deterministically cointegrated if that linear combination also eliminates the linear trend. Given variables with both a stochastic trend and a broken linear trend, a linear combination eliminates not only all SBs but also the linear trend itself. The traditional deterministic CI holds when our CI and CT jointly hold, but not when our CI and cobreaking jointly hold.

When CI does not hold, there is no stable relationship regardless of whether CB or CT holds. Hence, a test that distinguishes between CI and no CI regardless of CB and CT is of interest. We call this type of test a robust CI test because it is robust to the existence or absence of breaks. The robust CI test we develop uses the null hypothesis of CI and is designed to control size regardless of CB and CT. Rejection of the robust CI test implies no stable relationship, while nonrejection implies CI with CB and CT still being unfounded. To test for stronger relationships, we also develop a joint test for CI and CB and another for CI and CT.

All three tests are obtained in a unified framework, which extends Jansson’s (2005) approach. We consider two models, one with shifts in the mean and the other with intercept shifts in the linear trend. We analyze both the known and unknown break date cases.

Section 2 presents the models and then compares the new tests with other tests in the literature. Section 3 offers test statistics and their asymptotic distributions. Section 4 discusses endogenous regressors. Section 5 contains some Monte Carlo experiment results. Section 6 has an empirical illustration of U.S. budget sustainability. Section 7 concludes. The critical values of our tests are reported in Supplementary Appendix I, and all mathematical derivations are collected in Supplementary Appendix II. Additional Monte Carlo simulation results and auxiliary results for the empirical illustration are provided in Supplementary Appendix III.

2. Model

We assume that the observation \((y_t, x'_t)\) is generated by

\[
\begin{bmatrix}
  y_t \\
  x_t
\end{bmatrix} = \begin{bmatrix}
  x' \\
  x^\tau
\end{bmatrix} d_t + \begin{bmatrix}
  y_0 \\
  x_0^\tau
\end{bmatrix},
\]

where \(y_t\) is a scalar random variable, \(x_t\) is a \(p_x \times 1\) random vector, \(d_t\) is a \(p_d \times 1\) vector of deterministic functions in time, \(x'\) and \(x^\tau\) are coefficient matrices, and \(y_0\) and \(x_0^\tau\) are stochastic components. We consider two specifications for \(d_t\). That is, \(d'_t = [DU_t(T_0), \ldots, DU_t(T_m)]\) in Model I and \([DU_t(T_0), \ldots, DU_t(T_m), t]\), in Model II, where \(DU_t(T_i) = 1\) for \(t > T_i\) and 0 elsewhere, and \(T_0 = 0\). The deterministic function has \(m\) breaks in each model. Model I has mean changes, and Model II has a linear trend with shifts in the intercept. The coefficient matrices \(x'\) and \(x^\tau\) are defined as

\[\begin{equation}
\end{equation}\]

1. See Campbell and Perron (1991) and Ogaki and Park (1997), for example.

2. Our cobreaking and cotrending are special cases of Chapman and Ogaki’s (1993) cotrending or Hendry and Mizon’s (1998) cobreaking. Hatanaka and Yamada’s (2003) cotrending refers to the case in which there is no slope change, while intercept changes are allowed. Bierens (1997, 2000) considers cotrending of nonlinear time functions that are more general than broken linear trends.

3. A model with a slope change is of interest. A slope change can be located consistently even in the presence of a stochastic trend, which leads to a different statistical procedure. We explore this model in a separate paper.
\[ \chi' = \begin{cases} \{ \mu^y_0, \ldots, \mu^y_m \}, & \text{in Model I} \\ \{ \mu^x_0, \ldots, \mu^x_m, \psi^x_0 \}, & \text{in Model II} \end{cases} \]

Hence, \( \mu^y_i \) and \( \mu^x_i \) denote the intercepts in the \((i+1)^{th}\) regime, and \( \psi^y_0 \) and \( \psi^x_0 \) are the slopes.

Let \( \pi = (\pi_1, \ldots, \pi_m) \) be a vector of break fractions where each \( \pi_i \) satisfies \( T_i = [\pi_i T] \) with \( [c] \) being the integer part of \( c \). In what follows, it will be important to distinguish between the true break dates and break dates used in the regression equation. We will use \((T^0_1, \ldots, T^0_m)\) and \( \pi^0 = (\pi^0_1, \ldots, \pi^0_m) \) where \( T_i^0 = [\pi^0_i T] \) to denote the true break dates and true break fractions. We use the following assumption regarding the break dates.

**Assumption 1.** \( T_i^0 = [\pi^0_i T] \), and \( \pi^0_i - \pi^0_j \geq a > 0 \) for all \( j < i \) where \( \pi^0_0 = 0 \) and \( \pi^0_{m+1} = 1 \).

The stochastic components \( y^0_i \) and \( x^0_i \) are assumed to be integrated of order 1, I(1). Hence, \( y_i \) and \( x_i \) have both stochastic and broken deterministic trends. Our main focus is to test whether a certain linear combination of \((y_i, x_i')\) becomes more stable in the sense that stochastic trends, broken deterministic trends, or both cancel one another out. Such a linear combination, if it exists, is assumed to be unique up to scalar multiplication and to always include \( y_i \). We write this linear relationship in a regression form as

\[ y_t = \beta' x_t + \alpha' d_t + v_t, \quad \text{for } t = 1, \ldots, T, \]  

with \( v_t = y^0_t - \beta' x^0_t \) and \( \alpha' = (\chi' - \chi' \beta)' \) in Model I and \((\mu_0, \ldots, \mu_m, \psi_0)\) in Model II. We make the following assumptions regarding the random components.

**Assumption 2.** (i) \( v_t = v_{1t} + v_{2t} \); (ii) \( \Delta v_{1t} = \epsilon_{t} - \theta \epsilon_{t-1} \) with \( \epsilon_{10} = \epsilon_{0}, \epsilon_{t} = \sum_{i=0}^{\infty} c_i \eta_{t-i} \) with \( \sum_{i=0}^{\infty} ||c|| < C, \omega^2_i = (\sum_{i=0}^{\infty} c_i)^2 > 0 \) and \( \eta_t \sim i.i.d.(0,1) \); (iii) \( u^c_t = \Delta x^c_t = \sum_{i=0}^{\infty} G_i \eta^x_{t-i} \) with \( \sum_{i=0}^{\infty} ||G_i|| < C, \gamma^x_t = \sum_{i=0}^{\infty} G_i \eta^x_{t-i} \) is of full column rank, \( \eta_t \sim i.i.d.(0,G_p) \), and \( \eta_t \) and \( \eta^x_t \) are independent; (iv) \( v_{2t} = \sum_{j=-\infty}^{\infty} u_{t-j} \gamma^y_{-j} \) with \( \sum_{j=-\infty}^{\infty} ||\gamma|| < C \).

\( \Delta \) denotes the first difference operator. The disturbance \( \epsilon_t \) is stationary with its long-run variance being \( \omega^2 > 0 \). Hence, \( v_{1t} \) can be either I(0) or I(1), depending on the value of \( \theta \). The stochastic part of the regressors, \( x^c_t \), is I(1) and the full column rank condition for \( G_p \) excludes the case where CI holds among the elements in \( x^0_t \) without \( y^0_t \). \( v_{2t} \) is always stationary and captures the correlation between the regressors and the regression error.

**Assumption 3.** \( ||\psi^y_0|| > 0 \).

Assumption 3 implies that the linear trend exists in at least one element of \( x_t \). Since the existence of a linear trend in \( x_t \) affects the limiting distribution, this assumption will be used whenever needed. We also use the following definitions.

1. \((y_t, x'_t)\) is CI if and only if \( \theta = 1 \).
2. \((y_t, x'_t)\) is CB if and only if \( \mu_1 = \ldots = \mu_m = 0 \).
3. \((y_t, x'_t)\) is CT if and only if it is CB and \( \psi_0 = 0 \) in Model II.

The next diagram summarizes the data generating processes (DGP) that we consider.\(^4\)

| \(\text{CB (} \mu_i = 0, \forall i \geq 1 \) | \(\text{CT (} \psi_0 = 0 \) | \(\text{No CT (} \psi_0 \neq 0 \) | \(\text{No CB (} \mu_i \neq 0, \forall i \geq 1 \) |
|---|---|---|---|
| CI (\(\theta = 1\)) | 1 | 2 | 3 |
| No CI (\(\theta \neq 1\)) | 5 | 6 |  | 4 |

\(^4\) Model I assumes that there is no linear trend and thus CB implies CT.
As explained, if there is no CI, the variables under consideration do not possess any stable relationship, irrespective of CB and CT. Hence, a test that can distinguish DGPs 1, 2, and 3 from DGPs 4, 5, and 6 will be useful. This motivates us to develop a robust CI test. To show a stronger relationship than CI, it is natural to test for the joint null hypothesis of CI and CB (DGPs 1 and 2) or the joint null of CI and CT (DGP 1). Hence, we also develop a joint test for CI and CB and a joint test for CI and CT. In the following, we discuss what issues one can encounter when tests that are available in the literature are used.

First, the CI tests developed in line with UR tests such as Phillips and Ouliaris (1990) take the null hypothesis of no CI (DGP 4 and/or DGP5) and are designed to have power against the alternative of CI (DGPs 1 and 2). To parallel the UR tests, these tests should be called no CI tests. It is well known that these tests do not have much power against the alternative of CI with no CB (DGP 3), which is certainly a stronger relationship than no CI. Hence, modified tests such as the Gregory and Hansen (1996) test are proposed. The Gregory and Hansen test is specifically designed to have proper power against all of DGPs 1, 2, and 3, and thus, rejections from this test are often regarded as supporting CI. Furthermore, because of the ability to detect all of DGPs 1, 2, and 3, the Gregory and Hansen test is even regarded as a robust CI test. However, the null hypothesis employed is not just no CI (DGPs 4, 5, and 6). It should be described as no CI but CB (DGPs 4 and 5 only). It is not well known how this test would behave when neither CI nor CB is true (DGP 6). If it does not control size under DGP 6, it cannot be used as a robust CI test.

The Gregory and Hansen test is similar to the Zivot and Andrews (1992), Perron and Vogelsang (1992), and Perron (1997) UR tests. Kim and Perron (2009) show that the probability the Zivot and Andrews test rejects the null of a UR varies dramatically across the kinds and magnitudes of breaks if both a UR and a SB in the deterministic function exist. The asymptotic null distribution of the Zivot and Andrews test is derived with assuming that there is no SB. Hence, if there is a SB in the data, it affects the finite sample distribution. Similar properties are expected for the Gregory and Hansen test, which we demonstrate via Monte Carlo simulation. Rejecting the Gregory and Hansen test should not be regarded as supporting CI. The rejection may have been possible only due to a SB.

Inoue (1999) considers a CI rank test that allows for the breaking of deterministic trends. His treatment of SBs is similar to Gregory and Hansen’s. The null assumes CI rank at most $r$ with CB, while the alternative assumes CI rank greater than $r$ with no CB. This approach is not consistent in treating SBs, since SBs appear or disappear depending on whether we are under the null or the alternative hypothesis but also depending on the step of the sequential testing procedure. In the first step, the null does not consider a SB, but the alternative does; if we reject the null hypothesis, the alternative hypothesis of the first step becomes the null hypothesis of the second step with imposing that there is no break.

While our tests are based on the single equation CI model, some robust CI rank tests are developed in the context of the vector error correction model. Johansen et al. (2000), Saikkonen and Lütkepohl (2000), and Trenkler et al. (2008) consider cointegrating rank tests that allow for breaks in deterministic trends. Their tests are robust to the existence of breaks in the deterministic components, but they assume known break dates. Lütkepohl et al. (2004) extend Saikkonen et al. (2000) to allow for a break at an unknown date. The CI rank is assumed to be at most $r$ under the null and greater than $r$ under the alternative. When testing for the null that $r = 0$, the test controls size irrespective of CB and the test can be regarded as a robust CI test. Harris et al. (2016) also develop a robust CI rank test. Their model allows for a change in the slope of the trend at an unknown date.

The tests proposed by Carrion-i-Silvestre and Sansó (2006) and Arai and Kurozumi (2007) can be understood in a similar manner, although the testing strategy for these tests is the opposite of the Gregory and Hansen type tests. These tests are extensions of Shin’s (1994) CI test to allow for
breaks in the deterministic time function. They take the null of CI and are designed to have proper power against no CI regardless of CB/CT (DGPs 4, 5, and 6). However, they require breaks to exist under the null. Hence, these tests control size only when there is no CB (DGP 3). These tests are expected, though not shown, to have a well-defined limiting distribution in the simultaneous presence of CI and CB (DGPs 1 and 2). Thus, they can be made robust to CB/CT simply by using bigger percentiles of the two limiting distributions obtained under ‘DGPs 1 and 2’ and ‘DGP 3’.5

Many CI tests with the null of CI can be used as a joint test because they are general specification tests with nontrivial power not only under DGPs 4 and 5 but also under DGPs 3 and 6. This group includes the ones proposed by Shin (1994), Leybourne and McCabe (1993), Jansson (2005), and Kurozumi and Arai (2008). Since these tests are developed without considering the possibility of no CB, they leave some room for power improvement under DGPs 3 and 6, while they tend to have good power under DGPs 4 and 5. One good example is the nonmonotonic power function of the Shin test. The Shin test is an extension of the KPSS test proposed by Kwiatkowski et al. (1992), which is the same as Perron’s (1991) Q test. As is well documented in Vogelsang (1999) and Perron (2005), these stationarity tests have a nonmonotonic power function, which means that the power can drop even to zero as the magnitude of the break gets large. This occurs because the long-run variance estimate is contaminated by the breaks that are not accounted for in the construction of those tests. The bigger the unaccounted breaks are, the more severely the test loses power. In Section 5, we demonstrate, via Monte Carlo experiments, that the Shin test exhibits a nonmonotonic power function when CB is violated severely (DGPs 3 and 6).

SB tests for CI models can be understood in a similar way. This group includes Hansen (1992), Kuo (1998), Hao (1996), Seo (1998), Hansen (2003), Qu (2007), and Kejriwal and Perron (2010). The null hypothesis of these tests is the joint of CI and CB (DGPs 1 and 2), which is the same as the CI tests in the previous group. The difference is that these tests are designed to have power against no CB (DGP 3) instead of no CI (DGPs 4 and 5). However, these tests are also general specification tests and have power under DGPs 4 and 5.

With the joint null of CI and CB/CT, it is not obvious why tests should be designed to have power against only one type of violation. We show that our joint tests, developed considering both types of violation, have better power properties than the aforementioned CI tests or SB tests in CI models.

As mentioned in the introduction, many of the aforementioned tests assume that there is no break in each variable. Only the relationship among the variables is allowed to change. We deal with a different situation in which breaks in the relationship equation are caused by breaks in each variable and aim at giving a comprehensive answer regarding what to do when practitioners need to test for a potential long-run relationship with variables that are I(1) nonstationary with SBs. In principle, we cannot simply ignore breaks detected in the univariate analysis. These breaks may or may not be present in the relationship equation due to possible CB, and this must be assessed in a rigorous way. Furthermore, the case in which only the relationship equation has breaks is a special case of our model that sets \( \mu_i^x = 0 \) for all \( i \) in (1), and thus, our results hold regardless of the origin of breaks.

3. Test statistics and asymptotic distributions

3.1. Known break dates

The first testing problem concerns the null hypothesis of CI (\( \theta = 1 \)) against the alternative of no CI (\( \theta = \hat{\theta} < 1 \)). We devise our test statistic as a quasi-likelihood ratio by assuming that

---

5 In fact, this is how we develop our robust CI test. We derive two limiting distributions of our test statistic under the case of CB and no CB and suggest use of the bigger percentiles of the two distributions.
\( \epsilon_t \sim \text{i.i.d. } N(0, \sigma^2_\epsilon) \). In addition, we assume that the regressor vector \( x_t \) is exogenous (i.e., \( \nu_{2t} = 0 \) in Assumption 2). The treatment of endogenous regressors is postponed to the next section for the sake of expositional simplicity. The model in (2) in matrix notation becomes

\[
y = D\alpha + X\beta + \Psi_{1/2}^b, \quad y = (y_1, \ldots, y_T)', D = (d_1, \ldots, d_T)', X = (x_1, \ldots, x_T)', \epsilon = (\epsilon_1, \ldots, \epsilon_T)', \text{ and } \Psi_{1/2}^b \text{ is a lower triangular matrix with 1 on the diagonal and 1 off the diagonal.}
\]

Since the goal is to develop a robust CI test, we focus on tests that are invariant to location shifts in the conditional mean of \( y \). Following Jansson (2005), the maximal invariant we consider is

\[
P'y \sim N(0, \sigma^2_\epsilon P' \Psi_0 P),
\]

where \( P \) is defined such that \( PP' = M = I_T - Z(Z'Z)^{-1}Z' \) with \( Z = [D, X] \) and \( PP' = I_T - \rho_{t-1} \). The maximal invariant depends on \( \theta \) and the break dates used in \( D \). The log-likelihood of the maximal invariant is given as

\[
\ell_T(\theta, \pi, \sigma^2_\epsilon) = -\frac{1}{2} \log |P' \Psi_0 P| - \frac{1}{2\sigma^2_\epsilon} y'y(P' \Psi_0 P)^{-1} P'y
\]

\[
= -\frac{1}{2} \log |Z' \Psi_0^{-1} Z| - \frac{1}{2\sigma^2_\epsilon} y'(\Psi_0^{-1} - \Psi_0^{-1} Z(Z' \Psi_0^{-1} Z)^{-1} Z' \Psi_0^{-1}) y.
\]

The point optimal invariant (POI) test rejects for large values of

\[
-2(\ell_T(1, \pi_0^c, \sigma^2_\epsilon) - \ell_T(\bar{\theta}, \pi_0^c, \sigma^2_\epsilon)),
\]

which defines our infeasible robust CI test. Let the log-likelihood of \( y \) be \( \ell_T(\alpha, \beta|\theta, \pi_0^c, \sigma^2_\epsilon) \). Then, it follows that \( \ell_T(\theta, \pi_0^c, \sigma^2_\epsilon) = \max_{\alpha, \beta} \ell_T(\alpha, \beta|\theta, \pi_0^c, \sigma^2_\epsilon) \), and the POI test in (5) is the same as

\[
-2(\max_{\alpha, \beta} \ell_T(\alpha, \beta|\theta, \pi_0^c, \sigma^2_\epsilon) - \max_{\alpha, \beta} \ell_T(\alpha, \beta|\bar{\theta}, \pi_0^c, \sigma^2_\epsilon)).
\]

The test statistic for the joint null of CI and CB is given by

\[
-2 \max_{\alpha, \beta, \ell_0, \alpha=0} \ell_T(\alpha, \beta|1, \pi_0^c, \sigma^2_\epsilon) - \max_{\alpha, \beta} \ell_T(\alpha, \beta|\bar{\theta}, \pi_0^c, \sigma^2_\epsilon)
\]

where \( R_{cb} \alpha = (\mu_1, \ldots, \mu_m)' \) and \( c_{cb} = m \log(T) \).

The test statistic for the joint null of CI and CT in Model II is given by

\[
-2 \max_{\alpha, \beta, \ell_0, \alpha=0} \ell_T(\alpha, \beta|1, \pi_0^c, \sigma^2_\epsilon) - \max_{\alpha, \beta} \ell_T(\alpha, \beta|\bar{\theta}, \pi_0^c, \sigma^2_\epsilon)
\]

where \( R_{ct} \alpha = (\mu_1, \ldots, \mu_m, \psi_0)' \) and \( c_{ct} = \log \left( \psi_0^T \psi_0^2 / (\psi_0^T G_2 \psi_0^2)^2 \right) + (m + 2) \log(T) \). The constants \( c_{cb} \) and \( c_{ct} \) are included to warrant proper convergence of the test statistics.\(^6\) The first term in \( c_{ct} \) can be estimated consistently from the first differences of \( x_t \).

Let \( \hat{\omega}_\epsilon^2 \) be a consistent estimator for \( \omega^2_\epsilon \), the long-run variance of \( \epsilon_t \), and let \( \hat{\epsilon}_{ct} \) be such that \( \hat{\epsilon}_{ct} - c_{ct} = \sigma_p(1) \). Then, using the elements defined above, the feasible robust CI test can be computed as

\[
Q_{ct} = -2(\ell_T(1, \pi_0^c, \hat{\omega}_\epsilon^2) - \ell_T(\bar{\theta}, \pi_0^c, \hat{\omega}_\epsilon^2)),
\]

the feasible joint test for CI and CB is given by

\[
Q_{cb} = -2 \max_{\alpha, \beta, \ell_0, \alpha=0} \ell_T(\alpha, \beta|1, \pi_0^c, \hat{\omega}_\epsilon^2) - \max_{\alpha, \beta} \ell_T(\alpha, \beta|\bar{\theta}, \pi_0^c, \hat{\omega}_\epsilon^2) + c_{cb},
\]

\(^6\) These constants originate from the log term of the Gaussian log-likelihood function, which involves trending regressors. See the proofs in Supplementary Appendix II to see how these constants are calculated.
and the feasible joint test for CI and CT in Model II is given by
\[
\mathcal{Q}_{ct} = -2 \left( \max_{\alpha, \beta} \ell_T(\alpha, \beta|1, \pi^0, \hat{\omega}_n^2) - \max_{\alpha, \beta} \ell_T(\alpha, \beta|\hat{\theta}, \pi^0, \hat{\omega}_n^2) \right) + \tilde{c}_{ct}.
\]

To describe the limiting distribution of the test statistics, we introduce the following notations. Let \( du(r, \pi_i) \) and \( b(r) \) be the limit counterpart of \( DU_i(T_i) \) and \( B_i \) so that \( du(r, \pi_i) = 1(r > \pi_i) \) and \( b(r) = r \). We also rewrite \( \theta \) and \( \tilde{\theta} \) using local-to-unity parameters, so that \( \theta = 1 - \lambda/T \) and \( \tilde{\theta} = 1 - \tilde{\lambda}/T \). \( V(r) \) and \( W(r) \) are independent standard Wiener processes of dimensions one and \( p_x \), respectively. We denote the first \( n \) elements of \( W(r) \) by \( W_n(r) \). We define \( V^k(s) = V(s) + \lambda \int_0^s V(r)dr, \) \( V^k_1(s) = V^k(s) - \tilde{\lambda} \int_0^s e^{\tilde{\lambda}(s-r)}V^k(r)dr, \) and \( Q^k_i(s) = Q^k_i(s)p, n) - \tilde{\lambda} \int_0^s e^{\tilde{\lambda}(s-r)}Q^k_i(r)p, n)dr, \) where \( Q^k_i(s|p, n) = (du(s, \pi_0), ..., du(s, \pi_{p-1}), W_n(s))' \) in Model I and \( (du(s, \pi_0), ..., du(s, \pi_{p-1}), b(s), W_n(s))' \) in Model II. Finally, \( \Phi_1(\tilde{\lambda}, \tilde{\lambda}) = 2\tilde{\lambda} \int_0^1 V^k_1(s) dV^k(s) - \tilde{\lambda}^2 \int_0^1 V^k_1(s)^2 ds \) and \( \Phi_2(\tilde{\lambda}, \tilde{\lambda}; p, n) = [\int_0^1 Q^k_i(s) dV^k_1(s)]\int_0^1 Q^k_i(s) Q^k_i(s)\int_0^1 Q^k_i(s) dV^k_1(s) - \log [\int_0^1 Q^k_i(s) Q^k_i(s)\int_0^1 Q^k_i(s) dV^k_1(s)]ds]. \)

The next theorem states the asymptotic distribution of the \( \mathcal{Q}_r \) statistic.

**Theorem 1.** Let \( \theta = 1 - \lambda/T, \tilde{\theta} = 1 - \tilde{\lambda}/T. \) Then, under Assumptions 1~2 with \( v_{2t} = 0 \), we have
\[
\mathcal{Q}_r \Rightarrow \phi_r(\tilde{\lambda}, \tilde{\lambda}, \pi^0) = \Phi_1(\tilde{\lambda}, \tilde{\lambda}) - \Phi_2(\tilde{\lambda}, 0, \pi^0; m + 1, p_x) + \Phi_2(\tilde{\lambda}, \tilde{\lambda}, \pi^0; m + 1, p_x).
\]

With \( \pi^0 \) known, the test statistic is exactly invariant to \( \alpha \), the CB and CT parameters. This implies that CB and CT may or may not hold under both the null and alternative hypotheses. Hence, the test is robust to the presence or absence of breaks. Note also that Assumption 3 is not needed here. Theorem 1 states the asymptotic distribution of \( \mathcal{Q}_r \) not only under the null (\( \theta = 1 \)) but also under the local alternative hypothesis (\( \tilde{\theta} < 1 \) or \( \tilde{\lambda} > 0 \)). Further, the limiting distribution depends on the composition of the deterministic trend (\( d_i \)), the number of stochastic regressors (\( p_x \)), the break fractions (\( \pi^0 \)), and the local-to-unity parameter (\( \tilde{\lambda} \)). The usual recommendation in the econometrics literature is to choose \( \tilde{\lambda} \) so that the local asymptotic power curve is tangent to the theoretical power envelope with 50% power, which means that the choice of \( \tilde{\lambda} \) will depend on the break dates. This unnecessarily complicates the testing procedure with unknown break dates. Instead, we propose using \( \tilde{\lambda} \) that makes the average power over possible combinations of break fractions achieve 50% when \( \lambda = \tilde{\lambda} \) so that \( \tilde{\lambda} \) does not depend on break positions.

The suggested values for \( \tilde{\lambda} \) are reported in Supplementary Appendix 1. The average power is obtained from all combinations of break fractions in one decimal place that satisfy Assumption 1 with \( a = 0.2 \). With this choice of \( \tilde{\lambda}, \mathcal{Q}_r \) is still a POI test with the local power curve tangent to the envelope around 50% rather than exactly at 50%. Using the suggested values of \( \tilde{\lambda} \), we report selected percentiles of \( \phi(0, \tilde{\lambda}, \pi^0) \) in Supplementary Appendix 1 for one and two SBs cases.\(^7\) The next two theorems state the limiting distribution of the \( \mathcal{Q}_{cb} \) and \( \mathcal{Q}_{ct} \) statistics under their respective null hypotheses.

**Theorem 2.** Let \( \theta = 1 \) and \( \tilde{\theta} = 1 - \tilde{\lambda}/T. \)

i. If CB holds in Models I and II, we have under Assumptions 1~2 with \( v_{2t} = 0 \),
\[
\mathcal{Q}_{cb} \Rightarrow \phi_{cb}(0, \tilde{\lambda}, \pi^0) \equiv \Phi_1(0, \tilde{\lambda}) - \Phi_2(0, 0, \pi^0; 1, p_x) + \Phi_2(0, \tilde{\lambda}, \pi^0; m + 1, p_x).
\]

ii. If CT holds in Model II, we have under Assumptions 1~3 with \( v_{2t} = 0 \),
\[
\mathcal{Q}_{ct} \Rightarrow \phi_{ct}(0, \tilde{\lambda}, \pi^0) \equiv \Phi_1(0, \tilde{\lambda}) - \Phi_2(0, 0, \pi^0; 1, p_x-1) + \Phi_2(0, \tilde{\lambda}, \pi^0; m + 1, p_x).
\]

---

\(^7\) In creating all asymptotic critical values, a Wiener process is approximated with 2,000 steps, and the number of replications is 20,000.
As can be seen, the limiting distributions depend on \(\tilde{\lambda}\), on the number and position of the SBs \((\pi^0)\) and on the number of stochastic regressors \((p_x)\). Selected percentiles of \(\varphi_{cb}(0, \tilde{\lambda}, \pi^0)\) and \(\varphi_{ct}(0, \tilde{\lambda}, \pi^0)\) are approximated by simulation, as reported in Supplementary Appendix I. Note that Assumption 3 is needed for Theorem 2 (ii). When Assumption 3 does not hold, there is no linear trend in \(x_t\) and the limiting distribution of \(Q_{ct}\) changes.

### 3.2. Unknown break dates

We now address the case of unknown break dates, in which analysts decide to use estimated break dates to compute the proposed test statistics. Unfortunately, when there are only intercept shifts, the break dates can be estimated from neither the levels nor the first differences of \((y_t, x_t')\). They can be estimated from the relationship equation in (2) only if CB does not hold while CI holds. This is a general feature shared by all proposals in the literature, including that of Lütkepohl et al. (2004). The only way to overcome this issue is to assume that the break magnitude is large so that it is increasing with \(T\), making the level shift non-negligible in the limit. While we do not pursue this approach in this paper, we suggest below a conservative approach that avoids assuming a particular magnitude of the SBs, covering the possible situation of uncertainty about whether the SBs have occurred.

We treat the break dates as additional parameters to be estimated. The break dates will be searched over the set \(\Pi(m) = \{(\pi_1, ..., \pi_m) | \pi_i - \pi_j \geq a > 0 \text{ for all } 0 \leq j < i \leq m + 1 \text{ where } \pi_0 = 0 \text{ and } \pi_{m+1} = 1\}. Then, the feasible test statistics that can be computed for the unknown break dates case are

\[
\hat{Q}_r = -2 \left( \max_{x, \beta, \pi} \ell_T(x, \beta, \pi | x' \hat{\beta}, \hat{\omega}_x^2) - \max_{x, \beta, \pi} \ell_T(x, \beta, \pi | \hat{\beta}, \hat{\omega}_x^2) \right)
\]

\[
\hat{Q}_{cb} = -2 \left( \max_{x, \beta, \pi} \ell_T(x, \beta, \pi | x' \hat{\beta}, \hat{\omega}_x^2) - \max_{x, \beta, \pi} \ell_T(x, \beta, \pi | \hat{\beta}, \hat{\omega}_x^2) \right) + \hat{c}_{cb},
\]

\[
\hat{Q}_{ct} = -2 \left( \max_{x, \beta, \pi} \ell_T(x, \beta, \pi | x' \hat{\beta}, \hat{\omega}_x^2) - \max_{x, \beta, \pi} \ell_T(x, \beta, \pi | \hat{\beta}, \hat{\omega}_x^2) \right) + \hat{c}_{ct},
\]

where maximization with respect to \(\pi\) is taken over \(\Pi(m)\).

It is worth mentioning that, in principle, it could be possible to define two different estimators for the break fractions, depending on whether the estimates are obtained under the null or alternative hypotheses, and using the notation in (4), \(\hat{\pi} = \arg\max_{\pi} \ell_T(1, \pi, \hat{\omega}_x^2)\) and \(\tilde{\pi} = \arg\max_{\pi} \ell_T(\hat{\beta}, \pi, \hat{\omega}_x^2)\). The consistency of break fraction estimator \(\hat{\pi}\) at rate \(T\) is well established in various SB models when CI holds; for instance, see Carrion-i-Silvestre and Sansó (2006). The next lemma shows the asymptotic equivalence of \(\hat{\pi}\) and \(\tilde{\pi}\).

**Lemma 1.** Let \(\theta = 1\) and \(\bar{\theta} = 1 - \frac{\bar{\lambda}}{T}\). Assume that CB does not hold so that \(\mu_i \neq 0\) for all \(i\). Then, under Assumptions 1 2. with \(v_{2t} = 0\), we have \(\|\hat{\pi} - \pi\| = o_p(T^{-1})\).

Taking into account these elements, let us first state the following theorem the asymptotic null distributions of the robust CI test statistics for the unknown break date case.

**Theorem 3.** Let \(\theta = 1\) and \(\bar{\theta} = 1 - \frac{\bar{\lambda}}{T}\). Suppose that Assumptions 1 2. with \(v_{2t} = 0\) hold.

i. If \(\|\hat{\pi} - \pi^0\| = O_p(T^{-1})\) and \(\|\tilde{\pi} - \pi\| = o_p(T^{-1})\), we have in Models I and II

\[
\hat{Q}_r \Rightarrow \varphi_r(0, \bar{\lambda}, \pi^0),
\]

where \(\varphi_r(0, \bar{\lambda}, \pi^0)\) is as defined in Theorem 1.
If CB holds, we have in Models I and II

\[ \hat{Q}_r \Rightarrow \varphi_r^{\text{max}}(0, \hat{\lambda}, m) \equiv \Phi_1(0, \hat{\lambda}) - \max_{\pi \in \Pi(m)} \Phi_2(0, 0, \pi; m + 1, p_x) + \max_{\pi \in \Pi(m)} \Phi_2(0, \hat{\lambda}, \pi; m + 1, p_x). \]

The result in (i) pertains to the case of no CB (DGP 3) and the result in (ii) pertains to the case of CB (DGP 1 and 2). As can be seen, the asymptotic distribution of \( \hat{Q}_r \) is the same as that of \( Q_r \), with no CB, while it takes a different form with CB. This makes the empirical use of these statistics difficult, since a practitioner might not be sure about the presence of SBs. To control size over all of DGP 1, 2, and 3, we propose using a conservative approach by taking the larger of the two critical values coming from these distributions.

When CB is known to hold, Jansson’s (2005) test is the POI test. Naturally one might consider a test procedure that pretests for intercept shifts and chooses the optimal CI test depending on the pretest result. The only available candidate for the pretest is a SB test for the intercept of a CI equation, because intercept shifts are not detectable without CI. As explained earlier, however, such a SB test rejects with a probability approaching one the null of no break as long as there is no CI, irrespective of CB. The intended power gain then fails to materialize because the \( \hat{Q}_r \) test is selected over Jansson’s test even when there is CB with no CI. Thus, we do not pursue such a sequential method.

Selected percentiles of \( \varphi_r^{\text{max}}(0, \hat{\lambda}, m) \) are reported in Supplementary Appendix I. When multiple breaks are involved, intermediate cases not covered by the above theorem exist. They occur when some breaks cancel out but others do not, and they correspond to no CB in our definition since there is at least one break appearing in the relationship. Hence, some break date estimators are consistent for existing breaks, while the other break date estimators remain random. The limiting distribution of the test statistic will then be intermediate of the two distributions given by the above theorem. We do not explicitly derive these distributions, although they can be obtained from the results of this paper. Because our goal is to control size irrespective of CB, we only need to analyze the extreme cases that can yield the greatest critical value. See also the Monte Carlo simulation results in Section 5.

One practical issue is that the two break date estimates \( \hat{\pi} \) and \( \hat{\lambda} \) can be different in practice. In such a case, one suggestion is to use only \( \hat{\pi} \). Let \( \hat{Q}_r = -2(\ell_T(1, \hat{\pi}, \hat{\sigma}_T^2) - \ell_T(\hat{\theta}, \hat{\pi}, \hat{\sigma}_T^2)) \). Then, \( \hat{Q}_r \geq \hat{Q}_r \), and the test becomes only conservative. Next, we provide the asymptotic null distributions for the joint tests.

**Theorem 4.** Let \( \theta = 1 \) and \( \tilde{\theta} = 1 - \frac{\hat{\lambda}}{T} \).

i. If CB holds in Models I and II, we have under Assumptions 1 ~ 2 with \( v_{2t} = 0 \),

\[ \hat{Q}_{cb} \Rightarrow \varphi_{cb}^{\text{max}}(0, \hat{\lambda}, m) \equiv \Phi_1(0, \hat{\lambda}) - \max_{\pi \in \Pi(m)} \Phi_2(0, 0, \pi; m + 1, p_x) + \max_{\pi \in \Pi(m)} \Phi_2(0, \hat{\lambda}, \pi; m + 1, p_x). \]

ii. If CT holds in Model II, we have under Assumptions 1 ~ 3 with \( v_{2t} = 0 \),

\[ \hat{Q}_{ci} \Rightarrow \varphi_{ci}^{\text{max}}(0, \hat{\lambda}, m) \equiv \Phi_1(0, \hat{\lambda}) - \Phi_2(0, 0, \pi; m + 1, p_x - 1) + \max_{\pi \in \Pi(m)} \Phi_2(0, \hat{\lambda}, \pi; m + 1, p_x). \]

Selected percentiles of \( \varphi_{cb}^{\text{max}}(0, \hat{\lambda}, m) \) and \( \varphi_{ci}^{\text{max}}(0, \hat{\lambda}, m) \) are reported in Supplementary Appendix I. While an optimal property is difficult to establish for the joint test, we note that \( \hat{Q}_{cb} \) consists of two terms, \( \hat{Q}_{cb} = Q_1 + \hat{Q}_r \), where

---

8 The tests proposed by Carrion-i-Silvestre and Sansó (2006) and by Arai and Kurozumi (2007) are comparable with our tests. The main difference is that they require breaks to exist when the break dates are unknown. Hence, no asymptotic distribution comparable with our \( \varphi_r^{\text{max}}(0, \hat{\lambda}, m) \) is derived for these tests.

9 Supplementary Appendix I also reports the percentiles of \( \varphi_{cb}^{\text{max}}(0, \hat{\lambda}, m) \) and \( \varphi_{ci}^{\text{max}}(0, \hat{\lambda}, m) \) for \( m = 0 \), which correspond to regular cointegration tests that do not consider broken deterministic trends.
which is similar to the sup-type SB test in the cointegrated equation explored by Hansen (1992) and Kejriwal and Perron (2010). Andrews and Ploberger (1994) show that the sup-type SB test in general is not optimal although it is asymptotically admissible. Similarly, $\hat{Q}_{ct} = Q_2 + Q_{cb}$ where

$$Q_2 = -2 \left( \max_{z, \beta \text{ s.t. } R_{eb} = 0} \ell_T(z, \beta | 1, \pi, \hat{\omega}_e^2) - \max_{z, \beta, \pi} \ell_T(z, \beta | 1, \pi, \hat{\omega}_e^2) \right) + c_{cb},$$

which is a test for the significance of the linear trend.

One issue in using the $\hat{Q}_{cb}$ and $\hat{Q}_{ct}$ tests is the choice of $m$. We now write $\hat{Q}_{cb}(m)$ and $\hat{Q}_{ct}(m)$ for $\hat{Q}_{cb}$ and $\hat{Q}_{ct}$ to signify the assumed number of breaks. When the break dates are unknown, it is unlikely that the number of breaks is known. A natural way to proceed is to extend the principle of likelihood ratio by treating $m$ as another parameter. This would simply require computing $\hat{Q}_{cb}(m)$ or $\hat{Q}_{ct}(m)$ from a range of $m$ values and taking the maximum. Since the percentiles of $\hat{Q}_{cb}(m)$ or $\hat{Q}_{ct}(m)$ are increasing in $m$, the maximum would be mostly determined by the $\hat{Q}_{cb}(m)$ or $\hat{Q}_{ct}(m)$ statistic with the largest $m$ value, failing to deliver the intended power property. Hence, we consider the following modified statistic given by

$$\hat{Q}^{D\max}_{cb} \equiv \max_{0 \leq m \leq M} \beta_m^{-1} \left[ \hat{Q}_{cb}(m) - \alpha_m \right]$$

and

$$\hat{Q}^{D\max}_{ct} \equiv \max_{0 \leq m \leq M} \beta^{-1}_m \left[ \hat{Q}_{ct}(m) - \alpha_m \right],$$

with $\alpha_m$ and $\beta_m$ to be defined below. These tests are related to the WDmax test proposed by Bai and Perron (1998). The following theorem states the asymptotic null distribution of the $\hat{Q}^{D\max}_{cb}$ and $\hat{Q}^{D\max}_{ct}$ statistics.

**Theorem 5.** Let $\theta = 1$ and $\bar{\theta} = 1 - \lambda / T$.

i. If CB holds in Models I and II, we have under Assumptions 1 ~ 2 with $\nu_{2t} = 0$,

$\hat{Q}^{D\max}_{cb} \Rightarrow Q^{D\max}_{cb} \equiv \max_{0 \leq m \leq M} \beta^{-1}_m \left[ \phi^{max}_{cb}(0, \bar{\lambda}_m, m) - \alpha_m \right],$

where $\phi^{max}_{cb}(0, \bar{\lambda}_m, m)$ is as defined in Theorem 4.

ii. If CT holds in Model II, then we have under Assumptions 1 ~ 3 with $\nu_{2t} = 0$,

$\hat{Q}^{D\max}_{ct} \Rightarrow Q^{D\max}_{ct} \equiv \max_{0 \leq m \leq M} \beta^{-1}_m \left[ \phi^{max}_{ct}(0, \bar{\lambda}_m, m) - \alpha_m \right],$

where $\phi^{max}_{ct}(\lambda, \bar{\lambda}_m, m)$ is as defined in Theorem 4.

Depending on $\alpha_m$ and $\beta_m$, $\hat{Q}^{D\max}_{cb}$ and $\hat{Q}^{D\max}_{ct}$ will show different power characteristics. We set $\alpha_m$ to be the asymptotic 95th percentile of $\phi^{max}_{cb}(0, \bar{\lambda}_m, m)$ or $\phi^{max}_{ct}(0, \bar{\lambda}_m, m)$ and $\beta_m$ to be the difference between the asymptotic 99th and 95th percentiles. Our choice of $\alpha_m$ and $\beta_m$ is made to make the critical values at 5% and 1% size become zero and one, respectively, for each $m$. 10 Using this choice of $\alpha_m$ and $\beta_m$, selected percentiles of $\phi^{D\max}_{cb}$ and $\phi^{D\max}_{ct}$ are reported in Supplementary Appendix I.

10 To see this, let $cv_{95}(m)$ and $cv_{99}(m)$ be the 95th and 99th percentiles. Then,

$$P \left( \beta^{-1}_m (\hat{Q}_{cb}(m) - \alpha_m) > 0 \right) = P \left( (\hat{Q}_{cb}(m) - cv_{95}(m)) > 0 \right) = 0.05,$$

$$P \left( \beta^{-1}_m (\hat{Q}_{cb}(m) - \alpha_m) > 1 \right) = P \left( (\hat{Q}_{cb}(m) - cv_{99}(m)) > (cv_{99}(m) - cv_{95}(m)) \right) = 0.01.$$
4. Endogenous regressors

We now discuss endogenous regressors. While the approach taken by Jansson (2005) is one possibility, we use the dynamic ordinary least squares (DOLS) by Saikkonen (1991). From Assumption 2, let \( v_t = \sum_{j=-k}^{k} u_{t-j}^\epsilon + \omega_t \) with \( \omega_t = \sum_{|j| > k} u_{t-j}^\epsilon \). The model in matrix notation is given by \( y = DX + X\beta + UX\gamma + \omega + \Psi_{1/2}^e \), where \( \gamma = (\gamma'_1, \ldots, \gamma'_p)' \). When we ignore \( \phi \) and \( \omega \), the maximal invariant of interest is \( P'_W y \) where \( P'_W = I_T - W(W'W)^{-1}W' = M_W \) with \( W = [Z, DX] \) and \( Z = [D, X] \). This is the same as the maximal invariant in (3) except for \( DX \). The log-likelihood of \( P'_W y \) is given as

\[
\ell_T(\bar{\theta}, \pi, \sigma^2_t) = -\frac{1}{2} \log |W'\Psi^{-1}W| - \frac{1}{2\sigma^2_t} y'(\Psi^{-1} - \Psi^{-1}W(W'\Psi^{-1}W)^{-1}W'\Psi^{-1})y.
\]

We make the following assumption, which is a standard in the literature.

**Assumption 4.** As \( T \to \infty, k \to \infty \) such that \( k^2 / T \to 0 \) and \( T^{1/2}\sum_{|j| \geq k} ||\gamma|| \to 0 \).

A finite number of leads and lags of \( DX \) augmented in the regression will fix endogeneity only partially due to \( \omega \). However, we increase the number of leads and lags as the sample size grows making \( \omega \) negligible. Assumption 4 provides the relevant conditions on the number of leads and lags, \( k \). The condition that \( \sum_{|j| \geq k} ||\gamma|| = o(T^{-1/2}) \) regulates the lower bound for \( k \). Note that the lower bound depends on the decay rate of \( ||\gamma|| \). For example, this condition permits \( k \) to increase at a logarithmic rate, if \( ||\gamma|| \) decays exponentially. The \( \phi \) term also has no effect asymptotically. For the robust CI test, we show in Supplementary Appendix II that under Assumptions 1 \( \sim \) 4,

\[
-2(\ell_T(1, \pi^0, \sigma^2_t) - \ell_T(\bar{\theta}, \pi^0, \sigma^2_t)) = \log |Z'Z| + \frac{1}{\sigma^2_t} e_0'(I - Z(Z'Z)^{-1}Z')e_0 \\
+ \log |Z'\Psi^{-1}Z| - \frac{1}{\sigma^2_t} e_0'(\Psi^{-1} - \Psi^{-1}Z(Z'\Psi^{-1}Z)^{-1}Z'\Psi^{-1})e_0 + o_p(1),
\]

which implies that the effects of additional regressors \( DX \) are asymptotically negligible and the asymptotic distributions obtained in the previous section continue to hold. Extensions for joint tests are straightforward. One only needs to add \( DX \) to each regression in the construction of the test statistics.

5. Monte Carlo simulation

The DGP is the same as (1). Recall that \( \alpha = \alpha' - \alpha^\beta \beta \). We set \( \alpha' \) to be a matrix of 1/2, and \( \beta \) to be a \( p_x \times 1 \) vector of \( 1/p_x; \beta = (1/p_x, \ldots, 1/p_x)' \). \( \alpha' \) is determined in such a way that \( \alpha \) can have a desired value. Recall that \( \alpha' = (\mu_0, \ldots, \mu_m, \psi_0) \). We always set \( \mu_0 = 0 \) as the simulation results are invariant. Hence, \( \mu_i = 0 \) \( (i \neq 0) \) corresponds to CB and \( \psi_0 = 0 \) to CT. \( x_t^0 \) is generated by accumulating independent standard normal variates. \( y_t^0 \) is created as a sum of \( \beta'x_t^0 \) and \( v_t = \sum_{j=1}^T (\epsilon_j - \theta \epsilon_{j-1}) \) with \( \epsilon_i \) being an independent standard normal variate and \( v_{10} = \epsilon_0 = 0 \). Hence,

\[\text{In fact, a weaker condition } k^{1/2} \sum_{|j| \geq k} ||\gamma|| \to 0 \text{ can be used, which allows the use of information criteria; see Kejriwal and Perron (2008).}\]
Table 1. Finite sample sizes.

(i) DGP 1 (CI + CT)

| Model | $T$  | $\hat{Q}_{r}(1)$ | $\hat{Q}_{ct}(1)$ | $\hat{Q}_{ct}(2)$ | $\hat{Q}_{ct}(2)$ | $\hat{Q}_c^{D_{max}}$ | $\hat{Q}_c^{D_{max}}$ |
|-------|-----|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| I     | 120 | 0.044           | 0.032           | n.a.            | 0.027           | n.a.            | 0.039           | n.a.            |
|       | 360 | 0.042           | 0.034           | n.a.            | 0.027           | n.a.            | 0.040           | n.a.            |
| II    | 120 | 0.035           | 0.027           | 0.028           | 0.029           | 0.015           | 0.013           | 0.045           | 0.040           |
|       | 360 | 0.035           | 0.038           | 0.036           | 0.027           | 0.032           | 0.024           | 0.041           | 0.037           |

(ii) DGP 3 (CI + No CB, $\mu_i = 2.0$)

| $\sigma^2$ | $T$  | $\hat{Q}_{r}(1)$ | $\hat{Q}_{r}(2)$ | $\hat{Q}_r(1)$ | $\hat{Q}_r(2)$ |
|-------------|-----|-----------------|-----------------|-----------------|-----------------|
| (0.5)       | 120 | 0.062           | 0.033           | 0.060           | 0.044           |
|             | 360 | 0.057           | 0.036           | 0.061           | 0.055           |
| (0.3, 0.7)  | 120 | 0.968           | 0.058           | 0.847           | 0.071           |
|             | 360 | 0.999           | 0.058           | 0.997           | 0.064           |

$\theta = 1$ corresponds to CI. We also consider only exogenous regressors by setting $\nu_{2t} = 0$. The number of stochastic regressors is set at $p_x = 1$ and the number of repetitions is 5,000 for the size and 1,000 for the power. To focus on the case of unknown break dates, we compute $\hat{Q}_r(m)$, $\hat{Q}_c(m)$, and $\hat{Q}_ct(m)$ for $m = 1$ and 2, $\varphi_c^{D_{max}}$, and $\varphi_{ct}^{D_{max}}$ with $M = 2$. We simulate the following six cases.

DGP 1 (CI + CT) $\theta = 1$, $\mu_i = 0$, $\psi_0 = 0$.
DGP 2 (CI + CB) $\theta = 1$, $\mu_i = 0$, $\psi_0 = 0.3$.
DGP 3 (CI + No CB) $\theta = 1$, $\mu_i = 1$, 2, 3, 6, and 9, $\psi_0 = 0.3$.
DGP 4 (CI + CT) $\theta = 0$, 0.25, 0.5 and 1.05 ~ 1.25 in steps of 0.1, $\mu_i = 0$, $\psi_0 = 0$.
DGP 5 (CI + CB) $\theta = 0$, 0.25, 0.5 and 1.05 ~ 1.25 in steps of 0.1, $\mu_i = 0$, $\psi_0 = 0.3$.
DGP 6 (No CI + No CB) $\theta = 1.05$ ~ 1.25 in steps of 0.1 with $\mu_i = 10(\theta - 1)$, and $\theta = 0$ with $\mu_i = 3$ ~ 9 in steps of 3, $\psi_0 = 0.3$.

Table 1 has the finite sample sizes of our tests obtained using the asymptotic critical values for the nominal size of 5%. In (i), the results are obtained under DGP 1 (CI + CT) with one intercept change for each variable but no break in the relationship equation due to CB. For the robust CI tests, $\hat{Q}_r(m)$, the bigger percentiles of $\varphi_r^{max}(0, \hat{\lambda}, m)$ and $\varphi_r(0, \hat{\lambda}, \hat{\pi})$ should be used. However, we use the bigger percentiles of $\varphi_r^{max}(0, \hat{\lambda}, m)$ and $\varphi_r(0, \hat{\lambda}, \pi^0)$ for simplicity. As can be checked in Supplementary Appendix I, the percentiles of $\varphi_r(0, \hat{\lambda}, \pi)$ do not vary much with $\pi$ and the simplification rarely affects the empirical sizes. Overall, our tests show rejection rates close to or slightly smaller than the nominal 5% due to the use of asymptotic critical values, so that better results would be expected if finite sample critical values were used instead.

In (ii), the results for the $\hat{Q}_r(m)$ test are obtained under DGP 3 (CI + No CB) with $\mu_i = 2.0$. Since there is no CB, we consider two different combinations of break fractions, 0.5 and (0.3, 0.7). When the number of breaks specified by $\hat{Q}_r(m)$ is greater than or equal to the true number of breaks, the robust CI test controls size near the nominal 5%, as desired. For example, $\hat{Q}_r(1)$ controls size properly only for the case of one break and $\hat{Q}_r(2)$ does so for up to two breaks.

---

12 See Supplementary Appendix III for the case of endogenous regressors both with and without the DOLS corrections discussed in Section 4.

13 For all tests, the long-run variance is computed by a heteroskedasticity and autocorrelation consistent covariance estimator with the quadratic spectral kernel, for which the bandwidth parameter is selected using the Andrews (1991) data-dependent method with an AR(1) approximation. In doing so, the residuals obtained under the null are used.
When there are more breaks than specified by $\hat{Q}_r(m)$, the robust CI test rejects with a large probability. This is a desirable feature, because the existence of any additional break beyond those specified under the null hypothesis implies an unstable relationship.

Table 2 reports size corrected null rejection probabilities under DGPs 1 ~ 6 with $m=1$ and $T=120$ for Model II. We also simulate the Gregory and Hansen test, the Shin test, the Lütkepohl, Saikkonen, and Trenkler test and the Kejriwal and Perron test. The Gregory and Hansen test (GH test) is selected to represent the group of no CI tests allowing for breaks; the Shin test to represent the group of CI tests using the joint null of CI and CB/CT; the Lütkepohl, Saikkonen, and Trenkler test (LST test) to represent a robust CI rank test; and the Kejriwal and Perron test (KP test) to represent the group of SB tests for CI models. The KP test we simulate is the one with one break and is thus directly comparable with the $\hat{Q}_{ct}(1)$ test.

Under DGP 1 (CI + CT), all tests except for the GH test and the LST test show 5% rejection rates because the finite sample critical values are obtained from DGP1. The finite sample critical values for the GH test and the LST tests are obtained under DGP 4 with $\theta=0$. Since there is CI under DGP 1, the GH test and the LST test reject with a large probability.

Under DGP 2, CT does not hold while the CI and CB still hold. As such, the $\hat{Q}_{ct}(1)$ and $\hat{Q}_{ct}^\text{max}$ tests in Model II correctly detect the breakdown of CT under DGP 2. The other tests exhibit rejection probabilities similar to the ones under DGP 1.

Under DGP 3 (CI + No CB), the rejection rate of the $\hat{Q}_r(1)$ test is slightly over 5% but stable across the values of $\mu_1$. This is again due to the fact that we are not using the conservative approach. The $\hat{Q}_{ct}(1), \hat{Q}_{ct}^\text{max}, \text{ and KP tests show similar powers, while the Shin test falls short. This is very encouraging for the } \hat{Q}_{ct}(1) \text{ test because the KP test is specifically developed to detect DGP 3. Note that the rejection rate of the Shin test drops as } \mu_1 \text{ increases in both models, which}

\begin{table}[h]
\centering
\begin{tabular}{cccccccccc}
\hline
$\Theta$ & $\mu_1$ & $\psi_1$ & $\hat{Q}_r(1)$ & $\hat{Q}_{ct}(1)$ & $\hat{Q}_{cb}(1)$ & $\hat{Q}_{ct}^\text{max}$ & $\hat{Q}_{cb}^\text{max}$ & Shin & KP & GH & LST \\
\hline
DGP 1 & 1.0 & 0.0 & 0.0 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.99 & 0.99 \\
DGP 2 & 1.0 & 0.0 & 0.3 & 0.05 & 0.06 & 1.0 & 0.06 & 1.0 & 0.03 & 0.05 & 0.99 \\
DGP 3 & 1.0 & 1.0 & 0.3 & 0.08 & 0.38 & 1.0 & 0.34 & 1.0 & 0.22 & 0.38 & 0.98 \\
DGP 4 & 0.0 & 0.0 & 0.0 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 \\
DGP 5 & 0.0 & 0.0 & 0.3 & 0.05 & 0.06 & 1.0 & 0.06 & 1.0 & 0.03 & 0.05 & 0.99 \\
DGP 6 & 0.0 & 0.0 & 0.3 & 0.05 & 0.06 & 1.0 & 0.06 & 1.0 & 0.03 & 0.05 & 0.99 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{cccccccccc}
\hline
$\Theta$ & $\mu_1$ & $\psi_1$ & $\hat{Q}_r(1)$ & $\hat{Q}_{ct}(1)$ & $\hat{Q}_{ct}(1)$ & $\hat{Q}_{ct}^\text{max}$ & $\hat{Q}_{ct}^\text{max}$ & Shin & KP & GH & LST \\
\hline
DGP 1 & 1.0 & 0.0 & 0.0 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.99 & 0.99 \\
DGP 2 & 1.0 & 0.0 & 0.3 & 0.05 & 0.06 & 1.0 & 0.06 & 1.0 & 0.03 & 0.05 & 0.99 \\
DGP 3 & 1.0 & 1.0 & 0.3 & 0.08 & 0.38 & 1.0 & 0.34 & 1.0 & 0.22 & 0.38 & 0.98 \\
DGP 4 & 0.0 & 0.0 & 0.0 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 \\
DGP 5 & 0.0 & 0.0 & 0.3 & 0.05 & 0.06 & 1.0 & 0.06 & 1.0 & 0.03 & 0.05 & 0.99 \\
DGP 6 & 0.0 & 0.0 & 0.3 & 0.05 & 0.06 & 1.0 & 0.06 & 1.0 & 0.03 & 0.05 & 0.99 \\
\hline
\end{tabular}
\end{table}

14 The results for Model I are available from the authors upon request.

15 To compute the long-run variance, residuals obtained under the null hypothesis are used for the Shin test, while those under the alternative hypothesis are used for the Gregory and Hansen and the Kejriwal and Perron tests. The VAR of order one is used for the Lütkepohl, Saikkonen, and Trenkler test.
is the well-known nonmonotonic power phenomenon. The GH test always rejects under DGP 3 and shows the intended ability to detect CI regardless of CB/CT. The LST test also rejects with a large probability but is less powerful than the GH test.

Under DGP 4 (No CI + CT), the picture is quite different. As expected, the rejection rate of the $\hat{Q}_c(1)$ test increases as $\theta$ deviates from one, which shows that the robust CI test has the intended power. All of the $\hat{Q}_{cb}(1)$, $\hat{Q}_{ct}(1)$, $\hat{Q}_{cb}^{D_{max}}$, and $\hat{Q}_{ct}^{D_{max}}$ tests also show the same pattern. Under DGP 4, CB/CT holds but these joint tests correctly reject the null hypothesis due to the violation of CI. Among these tests, the joint tests for CI and CT generally have the greatest power, the joint tests for CI and CB the second greatest, and the robust CI tests the least. The power difference is natural given the different null hypotheses of these tests. The $\hat{Q}_{cb}^{D_{max}}$ test has greater power than the $\hat{Q}_{cb}(1)$ test, and the $\hat{Q}_{ct}^{D_{max}}$ test has greater power than the $\hat{Q}_{ct}(1)$ test. This seems to be due to the fact that the $\hat{Q}_{cb}^{D_{max}}$ and $\hat{Q}_{ct}^{D_{max}}$ tests are also based on $\hat{Q}_{cb}(0)$ and $\hat{Q}_{ct}(0)$, which direct power only toward the violation of CI.\footnote{In general, the $\hat{Q}_c(m)$, $\hat{Q}_{cb}(m)$, and $\hat{Q}_{cb}(m)$ tests exhibit lower power as $m$ gets bigger, when there is no break in the data. Some relevant simulation results are reported in Supplementary Appendix III.}

The Shin test does not show a nonmonotonic power function. However, it is much less powerful than the other tests unless the DGP is very close to the null ($\theta$ close to one). It is interesting to see that the KP test is often more powerful than the Shin test even though it is designed as a SB test. However, the KP test is still far less powerful than our joint tests. The rejection rate for the GH test and for the LST test is 5% when $\theta = 0$ since the critical value is obtained here. Their rejection rates are supposed to be about 5% under the entire DGP 4 but they increase to 100% when $\theta$ is close to one. We can write from the Beveridge–Nelson decomposition that $v_t = (1-\theta) \sum_{j=1}^{l} e_j + \theta e_t$, and thus $(1-\theta)^2$ stands for the variance of the innovation to the stochastic trend of $v_t$. When $\theta$ is close to one, the random walk portion of $v_t$ can be masked by the stationary component $\theta e_t$ and the data might look as if there were CI, which makes the GH test falsely reject the null of no CI. The size of the LST test is also close to 5% only when $\theta = 0$.\footnote{Another issue for the LST test is that it assumes the DGP for $(\Delta y_t^x, \Delta x_t^y)$ to be a VAR of finite order. However, under DGP 4–6, $(\Delta y_t^x, \Delta x_t^y)$ is a moving average of order one with no equivalent finite-order VAR representation unless $\theta = 0$.}

Under DGP 5 (No CI + CB), the rejection rates for the $\hat{Q}_{ct}(1)$ and $\hat{Q}_{ct}^{D_{max}}$ tests jump to one, because these tests are rejecting the null for the violated CT as well. The rejection rates for the other tests are similar to those under DGP 4.

Lastly, under DGP 6 (No CI + No CB), neither CI nor CB holds. The rejection rates for the $\hat{Q}_{ct}(1)$ and $\hat{Q}_{ct}^{D_{max}}$ tests are still one due to their ability to detect CT breakdown. The $\hat{Q}_{cb}(1)$ and $\hat{Q}_{cb}^{D_{max}}$ perform better than the Shin and KP tests with a good margin. The rejection rate for the GH test is increasing in $\mu_1$ even when $\theta = 0$. As mentioned in Section 2, this is because the assumption of CB made under the null of the GH test breaks down. With this result, rejecting the GH test should not be interpreted as supporting CI. The rejection might have resulted from failed CB, while CI does not hold. Then, not rejecting the null of the GH test supports no CI with CB while rejection does not support any particular DGP. This is a very serious drawback because neither rejection nor nonrejection implies any meaningful relation among the variables under the test. The LST test is relatively free of this problem. When $\theta = 0$, the rejection rate is stable below 5% for all values of $\mu_1$.

6. Empirical illustration

The global economic crisis has created tensions in macroeconomic aggregates that must be monitored to drive economic recovery. Fiscal deficit sustainability has been an area in which
Table 3. U.S. Government Budget Sustainability.

| Leads & lags | $\hat{Q}_r(1), \hat{\pi}$ | $\hat{Q}_r(2), \hat{\pi}$ | Joint CI/CB | Joint CI/CT |
|--------------|----------------|----------------|-------------|-------------|
| 1            | 15.08^a (78)  | 16.24^a (32,79) | 18.17^a     | 23.13^b     | 4.55^a      | 33.59^a | 38.87^a | 6.45^a |
| 3            | 14.45^a (78)  | 14.71^a (32,79) | 18.57^b     | 23.58^b     | 4.48^a      | 29.83^b | 34.98^b | 4.98^b |
| 5            | 13.15^a (78)  | 13.64^a (32,79) | 18.45^a     | 23.59^b     | 4.30^a      | 26.76^a | 32.16^a | 4.17^a |
| 7            | 12.55^a (78)  | 14.08^a (33,78) | 18.33^a     | 23.62^b     | 4.18^a      | 25.61^a | 31.11^a | 3.90^a |

Note: The numbers in parentheses are the break fractions estimated under the null of CI. Superscripts a, b, and c refer to rejection at the 1%, 5%, and 10% levels of significance, respectively.

Economists have focused their attention, since other macroeconomic imbalances (e.g., current account imbalances or the debt of the economy) depend on government fiscal policy. Directly connected to the budget sustainability is the issue of whether the deficit series itself – defined as the difference between government revenues and expenditures – or any other linear combination of revenues and expenditures is I(0) stationary. Hence, one strategy that has been followed in the literature to assess fiscal deficit sustainability relies on discovering the strongest possible relationship between revenues and expenditures using CI analysis.

Earlier studies in the literature have already applied various statistical tests to analyze the existence of an equilibrium relationship between revenues and expenditures, and the results are quite mixed – see, for example, Haug (1991), Liu and Tanner (1994), Quintos (1995), Martin (2000), Bajo-Rubio et al. (2008), and Holmes (2010), among others. For instance, Holmes (2010) applies Gregory and Hansen’s (1996) test statistics and, although the overall results are mixed, some evidence of CI is claimed upon rejecting the null hypothesis of no CI. However, this does not imply that there is CI, because the rejection might result from the presence of a SB under the null hypothesis, which is ruled out in the framework used by Gregory and Hansen (1996). In this regard, the robust CI test proposed in this paper should be of interest, since it is designed to control size regardless of the presence of SBs.

The study conducted in Bajo-Rubio et al. (2008) is based on Shin’s (1994) CI test, which does not reject the null of deterministic CI, the joint of CI and CT in terms of our definition. With this nonrejection, they consider the system to be cointegrated and continue to apply Bai and Perron’s (1998) multiple SB test and find four regimes. Instead of Bai and Perron’s (1998) test, a better alternative which has become available would be Kejriwal and Perron’s (2010) multiple SBs test; note that this proposal has been specifically developed for cointegrating equations. As explained earlier, nonrejection of the Shin test combined with rejection of the SB test does not imply piecewise CI. Both the Shin test and the SB test should be viewed as a joint test of CI/CB or CI/CT, and nonrejection of one test and rejection of the other test should be interpreted as only conflicting evidence due to the different power characteristics. As above, the joint tests that are proposed in this paper should be of interest, since they offer more reliable power over various unstable relationships.

In this section, we conduct an empirical illustration focusing on US fiscal deficit sustainability using CI analysis that considers the possible existence of SBs. We apply our tests to quarterly US government revenues and expenditures as percentages of GDP from 1947.I to 2010.II.18 We use only Model II, since an upward trend in each variable is clearly visible from the time plot. In constructing the test statistics, we correct for the endogeneity bias via the DOLS device suggested by Saikkonen (1991), which augments the regression with leads and lags of the differentiated stochastic regressors. Instead of selecting a particular number of leads and lags, we report our results for a range of values, from one to seven in steps of two. Finally, it should be mentioned that the

18 In Supplementary Appendix III, we provide a figure of the time plots, the order of integration analysis for the revenues and expenditures series, and test results for the presence of SBs.
main conclusion does not change at all with the choice of the left-hand side variable. Consequently, we only present the results obtained with the expenditures being the dependent variable and the revenues the regressor.

Table 3 (i) reports the results of the robust CI test for Model II. For each number of allowed breaks, we reject the null of CI at the 1% level of significance in all cases, which strongly suggests that there is no CI. Table 3 (ii) reports the test results of the joint tests in Model II. Using the relevant asymptotic critical values, we again reject the joint null of CI/CB or CI/CT, at least at the 5% level of significance, in all cases; note that rejection of the null hypothesis is found at the 1% level of significance in most cases. Taking all these elements into consideration, we can conclude that the test statistics that have been computed do not support any stable relationship between the revenues and the expenditures (i.e., CI is not found in this case).

7. Conclusion

We consider a set of variables that have both stochastic and broken linear trends. Of main interest is whether there is a stable relationship among the variables. Whenever a CI model with SBs is considered, one should carefully consider where the breaks originate. Our model deals with the case in which the relationship equation is subject to breaks because of breaks in each variable. When it is certain that each variable has no break, but exogenous breaks are given to the relationship equation only, the traditional methods proposed in the literature might suffice. We extensively discuss what problems might emerge when the existing methods are applied to a set of variables with both stochastic and broken linear trends, and develop three new tests, a robust CI test, a joint test for CI and CB, and a joint test for CI and CT.

Funding

Josep Lluís Carrion-i-Silvestre acknowledges financial support from the Spanish Ministerio de Ciencia y Tecnología, Agencia Española de Investigación (AEI) and European Regional Development Fund (ERDF, EU) under grant ECO2017-83255-C3-1-P. Dukpa Kim acknowledges financial support from the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2014S1A5A8018632). Dukpa Kim also acknowledges that this research is supported by the Korea University Grant (K1613441).

References

Andrews, D. W. K., Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative hypothesis. *Econometrica* 62(6):1383–1414.
Arai, Y., Kurozumi, E. (2007). Testing for the null hypothesis of cointegration with a structural break. *Econometric Reviews* 26(6):705–739.
Bai, J., Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* 66(1):47–78.
Bajo-Rubio, O., Díaz-Roldán, C., Esteve, V. (2008). US deficit sustainability revisited: A multiple structural change approach. *Applied Economics* 40(12):1609–1613.
Bierens, H. J. (1997). Nonparametric cointegration analysis. *Journal of Econometrics* 77(2):379–404.
Bierens, H. J. (2000). Nonparametric nonlinear co-trending analysis, with an application to inflation and interest in the U.S. *Journal of Business and Economic Statistics* 18(3):323–337.
Campbell, J. Y., Perron, P. (1991). Pitfalls and opportunities: What macroeconomists should know about URs. *NBER Macroeconomics Annual* 6:141–201.
Carrion-I-Silvestre, J. L., Kim, D., Perron, P. (2009). GLS-based unit root tests with multiple structural breaks both under the null and the alternative hypotheses. *Econometric Theory* 25(6):1754–1792.
Carrion-I-Silvestre, J. L., Sansó, A. (2006). Testing the null of cointegration with structural breaks. *Oxford Bulletin of Economics and Statistics* 68(5):623–646.
Chapman, D. A., Ogaki, M. (1993). Cotrending and the stationarity of the real interest rate. Economics Letters 42(2–3):133–138.

Gregory, A. W., Hansen, B. E. (1996). Residual-based tests for cointegration in models with regime shifts. Journal of Econometrics 70(1):99–126.

Harris, D., Harvey, D., Leybourne, S. J., Robert Taylor, A. M. (2009). Testing for a unit root in the presence of a possible break in trend. Econometric Theory 25(6):1545–1588.

Hansen, D., Leybourne, S. J., Robert Taylor, A. M. (2016). Tests of the co-integration rank in VAR models in the presence of a possible break in trend at an unknown point. Journal of Econometrics 192(2):451–467.

Hansen, B. E. (1992). Tests for parameter instability in regressions with I(1) processes. Journal of Business and Economic Statistics 10(3):321–335.

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

Hao, K. (1996). Testing for structural change in cointegrated regression models: Some comparisons and generalizations. Econometric Reviews 15(4):401–429.

Harvey, D., Leybourne, S. J., Robert Taylor, A. M. (2012). Unit root testing under a local break in trend. Journal of Econometrics 167(1):140–167.

Harvey, D., Leybourne, S. J., Robert Taylor, A. M. (2013). Testing for unit root in the possible presence of multiple trend breaks using minimum Dickey–Fuller statistics. Journal of Econometrics 177(2):265–284.

Hatanaka, M., Yamada, H. (2003). Co-Trending: A Statistical System Analysis of Economic Trends. Springer.

Haug, A. A. (1991). Cointegration and government borrowing constraints: Evidence for the United States. Journal of Business and Economic Statistics 9(1):97–101.

Hendry, D. F., Mizon, G. E. (1998). Exogeneity, causality and co-breaking in economic policy analysis of a small econometric model of money in the UK. Empirical Economics 23(3):267–294.

Holmes, M. J. (2010). Nonlinearities, co-trending and budget balance sustainability. Economic Analysis and Policy 40(3):369–376.

Inoue, A. (1999). Tests of cointegrating rank with a trend-break. Journal of Econometrics 90(2):215–237.

Jansson, M. (2005). Point optimal tests of the null hypothesis of cointegration. Journal of Econometrics 124(1):187–201.

Johansen, S., Mosconi, R., Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. The Econometrics Journal 3(2):216–249.

Kejriwal, M., Perron, P. (2008). Data dependent rules for selection of the number of leads and lags in the dynamic OLS cointegrating regression. Econometric Theory 24:1425–1441.

Kejriwal, M., Perron, P. (2010). Testing for multiple structural changes in cointegrated regression models. Journal of Business and Economic Statistics 28(4):503–522.

Kim, D., Perron, P. (2009). Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses. Journal of Econometrics 148(1):1–13.

Kuo, B. (1998). Test for partial parameters stability in regressions with I(1) processes. Journal of Econometrics 86:337–368.

Kurozumi, E., Arai, Y. (2008). Test for the null hypothesis of cointegration with reduced size distortion. Journal of Time Series Analysis 29(3):476–500.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics 54(1–3):159–178.

Leybourne, S. J., McCabe, B. P. M. (1993). A simple test for cointegration. Oxford Bulletin of Economics and Statistics 56(1):97–103.

Liu, P., Tanner, E. (1994). Is the budget deficit “too large”?: Some further evidence. Economic Inquiry 32(3):511–518.

Lütkepohl, H., Saikkonen, P., Trenkler, C. (2004). Testing for the cointegrating rank of a VAR process with level shift at unknown time. Econometrica 72(2):647–662.

Martin, G. (2000). US deficit sustainability: A new approach based on multiple endogenous breaks. Journal of Applied Econometrics 15(1):83–105.

Ogaki, M., Park, J. Y. (1997). A cointegration approach to estimating preference parameters. Journal of Econometrics 82(1):107–134.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57(6):1361–1401.

Perron, P. (1991). A Test for Changes in a Polynomial Trend Function for a Dynamic Time Series. Mimeograph: Princeton University Department of Economics.

Perron, P. (1997). Further evidence on breaking trend functions in macroeconomic variables. Journal of Econometrics 80(2):355–385.

Perron, P. (2005). Dealing with structural breaks. In: Patterson, K., Mills, T.C., eds., Palgrave Handbook of Econometrics 1: Econometric Theory, pp. 278–352.
Perron, P., Vogelsang, T. J. (1992). Nonstationarity and level shifts with an application to purchasing power parity. *Journal of Business & Economic Statistics* 10(3):301–320.

Phillips, P. C. B., Ouliaris, S. (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica* 58(1):165–193.

Saikkonen, P. (1991). Asymptotically efficient estimation of cointegration regressions. *Econometric Theory* 7(01):1–21.

Seo, B. (1998). Tests for structural change in cointegrated systems. *Econometric Theory* 14(2):222–259.

Qu, Z. (2007). Searching for cointegration in a dynamic system. *The Econometrics Journal* 10(3):580–604.

Quintos, C. E. (1995). Sustainability of the deficit process with structural shifts. *Journal of Business and Economic Statistics* 13(4):409–417.

Saikkonen, P., Lütkepohl, H., Lütkepohl, H. (2000). Testing for the cointegrating rank of a VAR process with structural shifts. *Journal of Business and Economic Statistics* 18(4):451–464.

Shin, Y. (1994). A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory* 10(01):91–115.

Trenkle, C., Saikkonen, P., Lütkepohl, H. (2008). Testing for the cointegrating rank of a VAR process with level shift and trend break. *Journal of Time Series Analysis* 29(2):331–358.

Vogelsang, T. (1999). Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series. *Journal of Econometrics* 88(2):283–299.

Zivot, E., Andrews, D. W. K. (1992). Further evidence on the great crash, the oil price shock and the unit root hypothesis. *Journal of Business and Economic Statistics* 10(3):251–270.