

Tests of the Co-integration Rank in VAR Models in the Presence of a Possible Break in Trend at an Unknown Point

David Harris, Steve Leybourne, Robert Taylor

U. of Melbourne, U. of Nottingham, U. of Essex

Conference on "New Trends and Developments in Econometrics", Bank of Portugal

June 3rd 2016

This research is funded by the ESRC under award number ES/M01147X/1.

Motivation – 1

- ▶ Macroeconomic series typically characterized by piecewise linear (or broken) trend functions; see, inter alia, Stock and Watson (1996, 1999, 2005) and Perron and Zhu (2005).

Motivation – 1

- ▶ Macroeconomic series typically characterized by piecewise linear (or broken) trend functions; see, inter alia, Stock and Watson (1996, 1999, 2005) and Perron and Zhu (2005).
- ▶ In the univariate setting this has spurred a large literature on testing for an autoregressive unit root when a trend break may be present in the data.

Motivation – 1

- ▶ Macroeconomic series typically characterized by piecewise linear (or broken) trend functions; see, inter alia, Stock and Watson (1996, 1999, 2005) and Perron and Zhu (2005).
- ▶ In the univariate setting this has spurred a large literature on testing for an autoregressive unit root when a trend break may be present in the data.
 - ▶ Perron (1989) - unit root tests which omit regressors to model a trend break present in the data have non-pivotal limiting null distributions and are inconsistent under stable root alternatives.

Motivation – 1

- ▶ Macroeconomic series typically characterized by piecewise linear (or broken) trend functions; see, inter alia, Stock and Watson (1996, 1999, 2005) and Perron and Zhu (2005).
- ▶ In the univariate setting this has spurred a large literature on testing for an autoregressive unit root when a trend break may be present in the data.
 - ▶ Perron (1989) - unit root tests which omit regressors to model a trend break present in the data have non-pivotal limiting null distributions and are inconsistent under stable root alternatives.
 - ▶ However, if there's no break his known breakdate suggested approach loses considerable power by including unnecessary trend break regressors.

Motivation – 1

- ▶ Macroeconomic series typically characterized by piecewise linear (or broken) trend functions; see, inter alia, Stock and Watson (1996, 1999, 2005) and Perron and Zhu (2005).
- ▶ In the univariate setting this has spurred a large literature on testing for an autoregressive unit root when a trend break may be present in the data.
 - ▶ Perron (1989) - unit root tests which omit regressors to model a trend break present in the data have non-pivotal limiting null distributions and are inconsistent under stable root alternatives.
 - ▶ However, if there's no break his known breakdate suggested approach loses considerable power by including unnecessary trend break regressors.
 - ▶ Subsequent approaches focus on the unknown breakdate case - see Zivot and Andrews (1992) [minimum ADF] and Perron (1997) [estimate breakdate].

Motivation – 2

- ▶
 - ▶ ZA procedure non-pivotal unless no break is assumed to have occurred.

Motivation – 2



- ▶ ZA procedure non-pivotal unless no break is assumed to have occurred.
- ▶ Perron (1997) approach is problematic as the break point estimator has a non-degenerate limit distribution when no break is present which changes the limit null distrn. Size-controlled inference achieved by using so-called conservative critical values corresponding to the case where no break is present, with an associated loss of efficiency where a break is present.

Motivation – 2



- ▶ ZA procedure non-pivotal unless no break is assumed to have occurred.
- ▶ Perron (1997) approach is problematic as the break point estimator has a non-degenerate limit distribution when no break is present which changes the limit null distrn. Size-controlled inference achieved by using so-called conservative critical values corresponding to the case where no break is present, with an associated loss of efficiency where a break is present.
- ▶ As a result, Carrion-i-Silvestre *et al.* (2009), Harris *et al.* (2009) and Kim and Perron (2009) advocate approaches based on the use of pre-tests for the presence of a trend break.

Motivation – 3

- ▶ In the vector case, un-modelled trend breaks cause similar problems in the rank tests of Johansen (1995). Inoue (1999) documents large losses in finite sample power with Johansen's standard trace and maximum eigenvalue tests when an un-modelled trend break is present in the data. Our simulations also show this causes substantial over-sizing in the tests.

Motivation – 3

- ▶ In the vector case, un-modelled trend breaks cause similar problems in the rank tests of Johansen (1995). Inoue (1999) documents large losses in finite sample power with Johansen's standard trace and maximum eigenvalue tests when an un-modelled trend break is present in the data. Our simulations also show this causes substantial over-sizing in the tests.
- ▶ Johansen *et al.* (2000) develop LR tests, analogous to those in Perron (1989), for the case where the break in the trend function occurs at a *known* point. They consider single and multiple level break and trend breaks.

Motivation – 4

- ▶ Saikkonen and Lütkepohl (2000) allow for a level break (but no trend break) at a known date within a components DGP; Lütkepohl *et al.* (2003) for a level break (no trend break) at an *unknown* point, and Trenkler *et al.* (2007) for a trend break at a known date, propose further co-integration rank tests, in each case using the pseudo-GLS de-trending method outlined in Saikkonen and Lütkepohl (2000).

Motivation – 4

- ▶ Saikkonen and Lütkepohl (2000) allow for a level break (but no trend break) at a known date within a components DGP; Lütkepohl *et al.* (2003) for a level break (no trend break) at an *unknown* point, and Trenkler *et al.* (2007) for a trend break at a known date, propose further co-integration rank tests, in each case using the pseudo-GLS de-trending method outlined in Saikkonen and Lütkepohl (2000).
- ▶ Inoue (1999), develops Zivot and Andrews (1992) type co-integration rank tests by calculating with-break implementations of the Johansen (1995) statistics over all possible break dates and basing a test on the most positive of these.

Motivation – 4

- ▶ Saikkonen and Lütkepohl (2000) allow for a level break (but no trend break) at a known date within a components DGP; Lütkepohl *et al.* (2003) for a level break (no trend break) at an *unknown* point, and Trenkler *et al.* (2007) for a trend break at a known date, propose further co-integration rank tests, in each case using the pseudo-GLS de-trending method outlined in Saikkonen and Lütkepohl (2000).
- ▶ Inoue (1999), develops Zivot and Andrews (1992) type co-integration rank tests by calculating with-break implementations of the Johansen (1995) statistics over all possible break dates and basing a test on the most positive of these.
- ▶ All of these procedures assume that the VAR lag length is known.

Motivation – 5

- ▶ Significant lacunae therefore exist in the vector case:

Motivation – 5

- ▶ Significant lacunae therefore exist in the vector case:
 - ▶ Firstly, the approach in Inoue (1999) is infeasible in practice because, like Zivot and Andrews (1992), it cannot allow a trend break to occur under the null.

Motivation – 5

- ▶ Significant lacunae therefore exist in the vector case:
 - ▶ Firstly, the approach in Inoue (1999) is infeasible in practice because, like Zivot and Andrews (1992), it cannot allow a trend break to occur under the null.
 - ▶ Second, the tests considered in Johansen *et al.* (2000) and Trenkler *et al.* (2007) both assume that the trend break date is known to the practitioner. Moreover, these tests essentially assume that a trend break does indeed occur and, hence, would be expected to unnecessarily sacrifice a considerable degree of finite sample power when no break occurs. Indeed it should be noted that Lütkepohl *et al.* (2003) need to impose that a break does occur otherwise they run into the same problems outlined above for the Perron (1997) procedure where no break is present.

Motivation – 5

- ▶ Significant lacunae therefore exist in the vector case:
 - ▶ Firstly, the approach in Inoue (1999) is infeasible in practice because, like Zivot and Andrews (1992), it cannot allow a trend break to occur under the null.
 - ▶ Second, the tests considered in Johansen *et al.* (2000) and Trenkler *et al.* (2007) both assume that the trend break date is known to the practitioner. Moreover, these tests essentially assume that a trend break does indeed occur and, hence, would be expected to unnecessarily sacrifice a considerable degree of finite sample power when no break occurs. Indeed it should be noted that Lütkepohl *et al.* (2003) need to impose that a break does occur otherwise they run into the same problems outlined above for the Perron (1997) procedure where no break is present.
 - ▶ Third, known VAR lag length assumption.

Contributions – 1

- ▶ We focus on the trend break case, but allowing for a simultaneous level break.

Contributions – 1

- ▶ We focus on the trend break case, but allowing for a simultaneous level break.
- ▶ Propose new procedures similar in spirit to the approaches of Carrion-i-Silvestre *et al.* (2009), Harris *et al.* (2009) and Kim and Perron (2009).

Contributions – 1

- ▶ We focus on the trend break case, but allowing for a simultaneous level break.
- ▶ Propose new procedures similar in spirit to the approaches of Carrion-i-Silvestre *et al.* (2009), Harris *et al.* (2009) and Kim and Perron (2009).
- ▶ The first step in the procedure is to estimate the putative break date. Three such estimators are considered: first difference, ML and unrestricted.

Contributions – 1

- ▶ We focus on the trend break case, but allowing for a simultaneous level break.
- ▶ Propose new procedures similar in spirit to the approaches of Carrion-i-Silvestre *et al.* (2009), Harris *et al.* (2009) and Kim and Perron (2009).
- ▶ The first step in the procedure is to estimate the putative break date. Three such estimators are considered: first difference, ML and unrestricted.
- ▶ Based on these estimators, an information-based method using a Schwarz-type criterion is then employed to select between the version of the model which includes a trend break (for the estimated break date) and that which does not. A Schwarz-type criterion is also used to select the autoregressive lag length.

Contributions – 1

- ▶ We focus on the trend break case, but allowing for a simultaneous level break.
- ▶ Propose new procedures similar in spirit to the approaches of Carrion-i-Silvestre *et al.* (2009), Harris *et al.* (2009) and Kim and Perron (2009).
- ▶ The first step in the procedure is to estimate the putative break date. Three such estimators are considered: first difference, ML and unrestricted.
- ▶ Based on these estimators, an information-based method using a Schwarz-type criterion is then employed to select between the version of the model which includes a trend break (for the estimated break date) and that which does not. A Schwarz-type criterion is also used to select the autoregressive lag length.
- ▶ Rank tests are then computed appropriate to the model selected by these criteria.

Contributions – 2

- ▶ For each of the proposed procedures we establish that:

Contributions – 2

- ▶ For each of the proposed procedures we establish that:
 - ▶ (i) The estimator of the break fraction is consistent for the true break fraction;

Contributions – 2

- ▶ For each of the proposed procedures we establish that:
 - ▶ (i) The estimator of the break fraction is consistent for the true break fraction;
 - ▶ (ii) The information-based methods based on this estimator consistently select between the with-break and without-break variants of the model,

Contributions – 2

- ▶ For each of the proposed procedures we establish that:
 - ▶ (i) The estimator of the break fraction is consistent for the true break fraction;
 - ▶ (ii) The information-based methods based on this estimator consistently select between the with-break and without-break variants of the model,
 - ▶ (iii) The resulting trace statistics can be validly compared to known break date critical values in trend break case and to the without break critical values in the no break case.

Contributions – 3

- ▶ The information-based methods therefore allow us to correctly identify (in the limit) if we need to allow for a trend break in the model. This implies that where a break is not present we will not see the loss in efficiency incurred by including a redundant trend break regressor in the model, and at the same time where a trend break is present we will not see the potentially large impact on the size and power properties of the rank tests that result from omitting the trend break.

Contributions – 3

- ▶ The information-based methods therefore allow us to correctly identify (in the limit) if we need to allow for a trend break in the model. This implies that where a break is not present we will not see the loss in efficiency incurred by including a redundant trend break regressor in the model, and at the same time where a trend break is present we will not see the potentially large impact on the size and power properties of the rank tests that result from omitting the trend break.
- ▶ We present Monte Carlo simulation evidence which suggests that a full ML-based procedure, based on the Johansen *et al.* (2000) set-up, is preferred and generally works very well even for a relatively small sample size with finite sample performance quite close to that seen for the benchmark rank tests which would obtain with knowledge of whether a trend break is present.

The Trend Break VAR Model – 1

- ▶ Following Trenkler *et al.* (2007), we consider the n -dimensional time series process $y_t := (y_{1t}, \dots, y_{nt})'$, $t = 1, \dots, T$, generated according to the following DGP

$$y_t = \mu_{0,0}d_{0,t}(0) + \mu_{1,0}d_{1,t}(0) + \mu_{0,1}d_{0,t}(b) + \mu_{1,1}d_{1,t}(b) + u_t, \quad (1)$$

where $d_{0,t} := 1$, $d_{1,t}(0) := t$, $d_{0,t}(b)$ is a level shift dummy and $d_{1,t}(b) := 0 \vee (t - b)$ is a trend break dummy.

The Trend Break VAR Model – 1

- ▶ Following Trenkler *et al.* (2007), we consider the n -dimensional time series process $y_t := (y_{1t}, \dots, y_{nt})'$, $t = 1, \dots, T$, generated according to the following DGP

$$y_t = \mu_{0,0}d_{0,t}(0) + \mu_{1,0}d_{1,t}(0) + \mu_{0,1}d_{0,t}(b) + \mu_{1,1}d_{1,t}(b) + u_t, \quad (1)$$

where $d_{0,t} := 1$, $d_{1,t}(0) := t$, $d_{0,t}(b)$ is a level shift dummy and $d_{1,t}(b) := 0 \vee (t - b)$ is a trend break dummy.

- ▶ We parameterise the breakpoint in terms of the break fraction λ (assumed unknown) where $0 < \lambda_L \leq \lambda \leq \lambda_U < 1$, by $b = \lfloor \lambda T \rfloor$. Notice therefore that b is constrained to lie in the set $B := [\lfloor T\lambda_L \rfloor, \lfloor T\lambda_U \rfloor]$.

The Trend Break VAR Model – 1

- ▶ Following Trenkler *et al.* (2007), we consider the n -dimensional time series process $y_t := (y_{1t}, \dots, y_{nt})'$, $t = 1, \dots, T$, generated according to the following DGP

$$y_t = \mu_{0,0}d_{0,t}(0) + \mu_{1,0}d_{1,t}(0) + \mu_{0,1}d_{0,t}(b) + \mu_{1,1}d_{1,t}(b) + u_t, \quad (1)$$

where $d_{0,t} := 1$, $d_{1,t}(0) := t$, $d_{0,t}(b)$ is a level shift dummy and $d_{1,t}(b) := 0 \vee (t - b)$ is a trend break dummy.

- ▶ We parameterise the breakpoint in terms of the break fraction λ (assumed unknown) where $0 < \lambda_L \leq \lambda \leq \lambda_U < 1$, by $b = \lfloor \lambda T \rfloor$. Notice therefore that b is constrained to lie in the set $B := [\lfloor T\lambda_L \rfloor, \lfloor T\lambda_U \rfloor]$.
- ▶ A trend break exists in y_t only if $\mu_{1,1} \neq 0$ in (1) (ie. where at least one element of the vector $\mu_{1,1}$ is non-zero); unlike previous contributions to this literature, we won't assume that $\mu_{1,1} \neq 0$ or that the true break fraction λ^* is known.

The Trend Break VAR Model – 2

- ▶ The model in (1) is completed by specifying the usual p th order reduced rank VAR (CVAR) indeterministic version of the model of Johansen (1995) for u_t ; that is,

$$\Delta u_t = \alpha \beta' u_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta u_{t-j} + e_t, \quad t = 1, \dots, T \quad (2)$$

where $u_t := (u_{1t}, \dots, u_{nt})'$, $e_t := (e_{1t}, \dots, e_{nt})'$, with u_{1-p}, \dots, u_0 , taken to be fixed in the statistical analysis.

The Trend Break VAR Model – 2

- ▶ The model in (1) is completed by specifying the usual p th order reduced rank VAR (CVAR) indeterministic version of the model of Johansen (1995) for u_t ; that is,

$$\Delta u_t = \alpha \beta' u_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta u_{t-j} + e_t, \quad t = 1, \dots, T \quad (2)$$

where $u_t := (u_{1t}, \dots, u_{nt})'$, $e_t := (e_{1t}, \dots, e_{nt})'$, with u_{1-p}, \dots, u_0 , taken to be fixed in the statistical analysis.

- ▶ $\{e_t\}$ are taken to satisfy a globally stationary martingale difference assumption (Davidson, 1994, pp.454-455):

Assumption 1: *The innovations $\{e_t\}$ form a martingale difference sequence with respect to the filtration \mathcal{F}_t , where $\mathcal{F}_{t-1} \subseteq \mathcal{F}_t$ for $t = \dots, -1, 0, 1, 2, \dots$, satisfying: (i) the global homoskedasticity condition: $\frac{1}{T} \sum_{t=1}^T \mathbb{E}(e_t e_t' | \mathcal{F}_{t-1}) \xrightarrow{p} \Sigma$, where Σ is full-rank, and (ii) $\mathbb{E} \|e_t\|^4 \leq K < \infty$.*

The Trend Break VAR Model – 3

- ▶ As is routine, we also impose the standard so-called ‘ $I(1, r)$ conditions’ of Johansen (1995) on the parameters of (2) in order to rule out things like explosive, $I(2)$, and seasonal unit root processes.

Assumption 2: *The following conditions hold on the parameters of (2): (i) The autoregressive lag order p satisfies $1 \leq p < \infty$; (ii) $|(I_n - \sum_{j=1}^{p-1} \Gamma_j z^j)(1 - z) - \alpha\beta'z| = 0$ implies $|z| > 1$ or $z = 1$, and (iii) $|\alpha'_\perp \Gamma \beta_\perp| \neq 0$, where $\Gamma := (I_n - \sum_{j=1}^{p-1} \Gamma_j)$.*

The Trend Break VAR Model – 3

- ▶ As is routine, we also impose the standard so-called ‘ $I(1, r)$ conditions’ of Johansen (1995) on the parameters of (2) in order to rule out things like explosive, $I(2)$, and seasonal unit root processes.

Assumption 2: *The following conditions hold on the parameters of (2): (i) The autoregressive lag order p satisfies $1 \leq p < \infty$; (ii) $|(I_n - \sum_{j=1}^{p-1} \Gamma_j z^j)(1 - z) - \alpha \beta' z| = 0$ implies $|z| > 1$ or $z = 1$, and (iii) $|\alpha'_\perp \Gamma \beta_\perp| \neq 0$, where $\Gamma := (I_n - \sum_{j=1}^{p-1} \Gamma_j)$.*

- ▶ Under Assumption 2, u_t is integrated of order one ($I(1)$) with co-integration rank r , and the co-integrating relations $\beta' u_t - E(\beta' u_t)$ are stationary. Part (i) of Assumption 2 assumes that the lag length parameter p is finite, but crucially does not assume that it is known to the practitioner.

The Trend Break VAR Model – 4

- ▶ An alternative formulation of (1)-(2) is considered in Johansen *et al.* (2000). Multiplying (1) through by the lag polynomial $(I_n - \sum_{j=1}^{p-1} \Gamma_j L^j) \Delta - \alpha \beta' L$ and re-arranging yields the VECM form

$$\begin{aligned} \Delta y_t = & \delta_{0,0} d_{0,t}(0) + \delta_{0,1} d_{0,t}(b) + \sum_{j=0}^{p-1} \delta_{-1,j} d_{-1,t}(b+j) \\ & + \alpha (\beta' y_{t-1} + \delta'_{1,0} d_{1,t-1}(0) + \delta'_{1,1} d_{1,t-1}(b)) + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + e_t \end{aligned} \quad (3)$$

where $d_{-1,t}(b) := 1_{(t=b+1)}$ is an impulse dummy; cf. Equation (5) of Trenkler *et al.* (2007), where explicit formulae for the $\delta_{i,j}$, $i, j = 0, 1$ and $\delta_{-1,j}$, $j = 0, \dots, p-1$, coefficient vectors are provided.

The Trend Break VAR Model – 5

- ▶ The VECM form in (3) includes a (broken) linear trend but does so in such a way that its coefficients are restricted to exclude the possibility of a quadratic (broken) trend in y_t .

The Trend Break VAR Model – 5

- ▶ The VECM form in (3) includes a (broken) linear trend but does so in such a way that its coefficients are restricted to exclude the possibility of a quadratic (broken) trend in y_t .
- ▶ Regardless of whether we work with the components form in (1)-(2), as in Trenkler *et al.* (2007), or the VECM form in (3) as in Johansen *et al.* (2000), our interest is focussed on testing the usual null that the co-integration rank is (less than or equal to) r , $H(r)$, against $H(n)$, but without assuming any prior knowledge of whether $\mu_{1,1} = 0$ or $\mu_{1,1} \neq 0$ in (1), and in the case where $\mu_{1,1} \neq 0$ without prior knowledge of the trend break location, λ .

Co-integration Rank Test Procedures – 1

- ▶ The procedures we develop are all done within the usual reduced rank regression framework of Johansen (1995) and Johansen *et al.* (2000).

Co-integration Rank Test Procedures – 1

- ▶ The procedures we develop are all done within the usual reduced rank regression framework of Johansen (1995) and Johansen *et al.* (2000).
- ▶ In our preferred method, labelled SC-VECM, break date estimation is done using (quasi) MLE on (3) under $H(r)$. Given this breakpoint, an adaptation of the usual Schwarz information criterion [SC] is then used to select between the model with a break and the model with break excluded (i.e. (3) with $\delta_{0,1} = 0, \delta_{1,1} = 0$), both estimated under $H(r)$. The usual trace test for $H(r)$ is then performed on the selected model, using critical values appropriate to the selected model.

Co-integration Rank Test Procedures – 2

- ▶ It is also possible to estimate the breakpoint applying a standard least squares estimator to the first differences of the data, and correspondingly use SC to choose between the with-break and without-break models in the context of a simple random walk model. We call this the SC-DIFF procedure.

Co-integration Rank Test Procedures – 2

- ▶ It is also possible to estimate the breakpoint applying a standard least squares estimator to the first differences of the data, and correspondingly use SC to choose between the with-break and without-break models in the context of a simple random walk model. We call this the SC-DIFF procedure.
- ▶ Finally, we also consider SC-VAR, which carries out the breakpoint estimation and SC selection between the with-break and without-break models in an unrestricted VAR, i.e. with $r = n$.

Co-integration Rank Test Procedures – 2

- ▶ It is also possible to estimate the breakpoint applying a standard least squares estimator to the first differences of the data, and correspondingly use SC to choose between the with-break and without-break models in the context of a simple random walk model. We call this the SC-DIFF procedure.
- ▶ Finally, we also consider SC-VAR, which carries out the breakpoint estimation and SC selection between the with-break and without-break models in an unrestricted VAR, i.e. with $r = n$.
- ▶ Let us now detail these three approaches.

SC-VECM – 1

The SC-VECM procedure can then be described as follows.

- ▶ Step 1. For each of $p = 1, \dots, \bar{p}$, define the MLE of the breakpoint under $H(r)$,

$$\hat{b}_{r,p} := \arg \max_{b \in B} \hat{\ell}_T(r; D_{0,b/T}, D_{1,b/T}, p). \quad (4)$$

where $\hat{\ell}_T$ is the maximised quasi log-likelihood associated with (3) - where $D_{0,b/T}$ contains the constant and level break variables and $D_{1,b/T}$ contains the trend and trend break dummies. Correspondingly, $\hat{\lambda}_{r,p} := \hat{b}_{r,p}/T$.

SC-VECM – 1

The SC-VECM procedure can then be described as follows.

- ▶ Step 1. For each of $p = 1, \dots, \bar{p}$, define the MLE of the breakpoint under $H(r)$,

$$\hat{b}_{r,p} := \arg \max_{b \in B} \hat{\ell}_T(r; D_{0,b/T}, D_{1,b/T}, p). \quad (4)$$

where $\hat{\ell}_T$ is the maximised quasi log-likelihood associated with (3) - where $D_{0,b/T}$ contains the constant and level break variables and $D_{1,b/T}$ contains the trend and trend break dummies. Correspondingly, $\hat{\lambda}_{r,p} := \hat{b}_{r,p}/T$.

- ▶ Step 2. Define the SC for the trend break model to be

$$SC_1(p; r, \lambda) := -2\hat{\ell}_T(r; D_{0,\lambda}, D_{1,\lambda}, p) + (n + r + 2 + n^2 p) \log T,$$

set lag length $\hat{p}_{1,r} := \arg \min_{p \in \{1, \dots, \bar{p}\}} SC_1(p; n, \hat{\lambda}_{r,p})$, where $\hat{\lambda}_{r,p}$ is the estimate of λ^* obtained in Step 1. Notice that $\hat{p}_{1,r}$ is selected under $H(n)$.

SC-VECM – 2

- ▶ Step 3. Define the SC for the model excluding the trend break to be $(\iota_0$ and τ_0 the constant and linear trend)

$$SC_0(p; r) := -2\hat{\ell}_T(r; \iota_0, \tau_0; p) + (n^2 p) \log T,$$

with selected lag length $\hat{p}_0 := \arg \min_{p \in \{1, \dots, \bar{p}\}} SC_0(p; n)$.
Again notice that \hat{p}_0 is selected under $H(n)$.

SC-VECM – 2

- ▶ Step 3. Define the SC for the model excluding the trend break to be $(\iota_0$ and τ_0 the constant and linear trend)

$$SC_0(p; r) := -2\hat{\ell}_T(r; \iota_0, \tau_0; p) + (n^2 p) \log T,$$

with selected lag length $\hat{p}_0 := \arg \min_{p \in \{1, \dots, \bar{p}\}} SC_0(p; n)$.
Again notice that \hat{p}_0 is selected under $H(n)$.

- ▶ Step 4. Choose the model with trend break by setting:
 $\hat{p} = \hat{p}_{1,r}$ and $(X_0, X_1) = (D_{0, \hat{\lambda}_{r, \hat{p}}}, D_{1, \hat{\lambda}_{r, \hat{p}}})$ if

$$SC-VECM : SC_1(\hat{p}_{1,r}; r, \hat{\lambda}_{r, \hat{p}_{1,r}}) \leq SC_0(\hat{p}_0; r);$$

and setting $\hat{p} = \hat{p}_0$ and $(X_0, X_1) = (\iota_0, \tau_0)$ otherwise.

SC-VECM – 2

- ▶ Step 3. Define the SC for the model excluding the trend break to be $(\iota_0$ and τ_0 the constant and linear trend)

$$SC_0(p; r) := -2\hat{\ell}_T(r; \iota_0, \tau_0; p) + (n^2 p) \log T,$$

with selected lag length $\hat{p}_0 := \arg \min_{p \in \{1, \dots, \bar{p}\}} SC_0(p; n)$.
Again notice that \hat{p}_0 is selected under $H(n)$.

- ▶ Step 4. Choose the model with trend break by setting:
 $\hat{p} = \hat{p}_{1,r}$ and $(X_0, X_1) = (D_{0, \hat{\lambda}_{r, \hat{p}}}, D_{1, \hat{\lambda}_{r, \hat{p}}})$ if

$$SC-VECM : SC_1(\hat{p}_{1,r}; r, \hat{\lambda}_{r, \hat{p}_{1,r}}) \leq SC_0(\hat{p}_0; r);$$

and setting $\hat{p} = \hat{p}_0$ and $(X_0, X_1) = (\iota_0, \tau_0)$ otherwise.

- ▶ Step 5. The trace test statistic of $H(r)$ against $H(n)$ is then

$$q_T(X_0, X_1; \hat{p}) := 2 \left(\hat{\ell}_T(n; X_0, X_1, \hat{p}) - \hat{\ell}_T(r; X_0, X_1, \hat{p}) \right).$$



SC-VECM – 3

- ▶ Step 4 constitutes a pre-test for the presence of a break which, by design, has size which shrinks to zero as T diverges. The same requirement is needed on the trend break pre-tests used in the univariate case by Harris *et al.* (2009) and Carrion-i-Silvestre *et al.* (2009).

SC-VECM – 3

- ▶ Step 4 constitutes a pre-test for the presence of a break which, by design, has size which shrinks to zero as T diverges. The same requirement is needed on the trend break pre-tests used in the univariate case by Harris *et al.* (2009) and Carrion-i-Silvestre *et al.* (2009).
- ▶ The part of the SC-type penalty corresponding to the trend break is $(n + r + 2) \log T$. There are n parameters in $\delta_{0,1}$, r parameters in $\delta_{1,1}$ and the unknown breakpoint parameter is given a penalty of 2, the latter following from the theoretical results provided in Zhang and Siegmund (2007), Kurozumi and Tuvaandorj (2011) and Kim (2012). We found that 2 indeed gave better finite sample results than the usual penalty of 1, in that the latter did not appear to penalise the inclusion of the break sufficiently, with the trend break retained too often when no break occurred, resulting in correspondingly lower power.

SC-VECM – 4

- ▶ The lag length is selected for both the model including a break and the model excluding a break. Although the breakpoint estimation and break selection is done under $H(r)$, it is necessary to select p under $H(n)$ (i.e. from the VAR in levels). Failure to do so leads to power losses for the trace test; see Lütkepohl (2005) and Lütkepohl and Saikkonen (1999).

SC-VECM – 4

- ▶ The lag length is selected for both the model including a break and the model excluding a break. Although the breakpoint estimation and break selection is done under $H(r)$, it is necessary to select p under $H(n)$ (i.e. from the VAR in levels). Failure to do so leads to power losses for the trace test; see Lütkepohl (2005) and Lütkepohl and Saikkonen (1999).
- ▶ When a trend break occurs, \hat{p}_0 may be inconsistent because it is based on a misspecified deterministic. Nevertheless, we show that the selection of the trend break in step 4 is still consistent, implying that the resulting lag length estimator \hat{p} is consistent whether or not a trend break is present in the DGP. An alternative is change step 4 to $SC_1(\hat{p}_{1,r}; r, \hat{\lambda}_{r, \hat{p}_{1,r}}) \leq SC_0(\hat{p}_{1,r}; r)$, so that only the lag length estimator $\hat{p}_{1,r}$ is used. No impact in the limit but in simulations we found this to give inferior performance.

SC-DIFF – 1

The SC-DIFF procedure can then be described as follows.

- ▶ Step 1. Use $\hat{b}_{0,1}$ defined in (4) and the resulting $\hat{\lambda}_{0,1} := \hat{b}_{0,1}/T$; that is the breakpoint and break fraction estimates are obtained setting $r = 0$ and $p = 1$.

SC-DIFF – 1

The SC-DIFF procedure can then be described as follows.

- ▶ Step 1. Use $\hat{b}_{0,1}$ defined in (4) and the resulting $\hat{\lambda}_{0,1} := \hat{b}_{0,1}/T$; that is the breakpoint and break fraction estimates are obtained setting $r = 0$ and $p = 1$.
- ▶ Step 2. Choose the model with trend break by setting:
 $(X_0, X_1) = (D_{0, \hat{\lambda}_{0,1}}, D_{1, \hat{\lambda}_{0,1}})$ if

$$\text{SC-DIFF} : SC_1(1; 0, \hat{\lambda}_{0,1}) \leq SC_0(1; 0),$$

where $SC_1(.; ., .)$ and $SC_0(.; .)$ are as defined in Steps 2 and 3, respectively, of SC-VECM; and setting $(X_0, X_1) = (\iota_0, \tau_0)$ otherwise.

SC-DIFF – 2

- ▶ Step 3. If the break is selected in Step 2, set

$$\hat{p} = \hat{p}_{1,0} := \arg \min_{p \in \{1, \dots, \bar{p}\}} SC_1(p; n, \hat{\lambda}_{0,1}).$$

If the break is not selected in Step 2, set

$$\hat{p} := \arg \min_{p \in \{1, \dots, \bar{p}\}} SC_0(p; n).$$

SC-DIFF – 2

- ▶ Step 3. If the break is selected in Step 2, set

$$\hat{p} = \hat{p}_{1,0} := \arg \min_{p \in \{1, \dots, \bar{p}\}} SC_1(p; n, \hat{\lambda}_{0,1}).$$

If the break is not selected in Step 2, set

$$\hat{p} := \arg \min_{p \in \{1, \dots, \bar{p}\}} SC_0(p; n).$$

- ▶ Step 4. The trace test statistic of $H(r)$ against $H(n)$ is then given by

$$q_T(X_0, X_1; \hat{p}) := 2 \left(\hat{\ell}_T(n; X_0, X_1, \hat{p}) - \hat{\ell}_T(r; X_0, X_1, \hat{p}) \right).$$



SC-DIFF – 3

- ▶ $\hat{b}_{0,1}$ can be viewed as the multivariate extension of the trend break estimator discussed in Harris *et al.* (2009) which is based on applying the univariate level break estimator proposed in Bai (1994) to the first differences of the data.

SC-DIFF – 3

- ▶ $\hat{b}_{0,1}$ can be viewed as the multivariate extension of the trend break estimator discussed in Harris *et al.* (2009) which is based on applying the univariate level break estimator proposed in Bai (1994) to the first differences of the data.
- ▶ The SC-DIFF approach imposes $r = 0$ and $p = 1$ for the breakpoint estimator and break selection steps. Although based on a misspecified model when either $r > 0$ or $p > 1$, SC-DIFF is still able to consistently discriminate between the trend break and no trend break models in such cases.

SC-DIFF – 3

- ▶ $\hat{b}_{0,1}$ can be viewed as the multivariate extension of the trend break estimator discussed in Harris *et al.* (2009) which is based on applying the univariate level break estimator proposed in Bai (1994) to the first differences of the data.
- ▶ The SC-DIFF approach imposes $r = 0$ and $p = 1$ for the breakpoint estimator and break selection steps. Although based on a misspecified model when either $r > 0$ or $p > 1$, SC-DIFF is still able to consistently discriminate between the trend break and no trend break models in such cases.
- ▶ Unlike SC-VECM, SC-DIFF uses only a single breakpoint estimator, $\hat{b}_{0,1}$, across all r and p . As a result, the lag length selection is the same for every r if a sequence of $H(r)$, $r = 0, 1, \dots$, hypotheses are being tested in a sequential procedure.

SC-VAR – 1

- ▶ The SC-VAR procedure is similar to SC-VECM except that the breakpoint estimator in Step 1 is calculated under $H(n)$, and carries out the SC-based choice between the trend break and no trend break model also under $H(n)$.

SC-VAR – 1

- ▶ The SC-VAR procedure is similar to SC-VECM except that the breakpoint estimator in Step 1 is calculated under $H(n)$, and carries out the SC-based choice between the trend break and no trend break model also under $H(n)$.
- ▶ Although we show that the large sample properties of SC-VECM and SC-VAR are the same, in our Monte Carlo simulations we found SC-VAR to be always (and often substantially) inferior to SC-VECM. We will therefore not discuss SC-VAR further here.

Asymptotics – 1

In the paper we establish the following results:

- ▶ First that the break fraction estimators $\hat{\lambda}_{r,p}$ are consistent for the true break fraction, λ^* at rate $O_p(T^{-1})$, where a break occurs. This rate holds regardless of the true co-integrating rank, r^* . Moreover, it also holds regardless of the true autoregressive lag length, p^* , since correct specification of the lag length is not necessary for the consistent estimation of the break fraction.

Asymptotics – 1

In the paper we establish the following results:

- ▶ First that the break fraction estimators $\hat{\lambda}_{r,p}$ are consistent for the true break fraction, λ^* at rate $O_p(T^{-1})$, where a break occurs. This rate holds regardless of the true co-integrating rank, r^* . Moreover, it also holds regardless of the true autoregressive lag length, p^* , since correct specification of the lag length is not necessary for the consistent estimation of the break fraction.
- ▶ Second, we show that each of the three SC criteria are consistent in that, with probability converging to one as T diverges, they correctly discriminate between the trend break and no trend break models.

Asymptotics – 1

In the paper we establish the following results:

- ▶ First that the break fraction estimators $\hat{\lambda}_{r,p}$ are consistent for the true break fraction, λ^* at rate $O_p(T^{-1})$, where a break occurs. This rate holds regardless of the true co-integrating rank, r^* . Moreover, it also holds regardless of the true autoregressive lag length, p^* , since correct specification of the lag length is not necessary for the consistent estimation of the break fraction.
- ▶ Second, we show that each of the three SC criteria are consistent in that, with probability converging to one as T diverges, they correctly discriminate between the trend break and no trend break models.
- ▶ Third, the trace statistics constructed using the estimated break fraction and lag length are asymptotically equivalent to those based on the true break fraction and true lag length where a break occurs.

Asymptotics – 2

- ▶ The second and third results crucially depend on the rate of consistency established for the break fraction estimator in the first result, where a break occurs.

Asymptotics – 2

- ▶ The second and third results crucially depend on the rate of consistency established for the break fraction estimator in the first result, where a break occurs.
- ▶ The results allow us to conclude that where a trend break is present the trace statistics based on the estimated trend break points are asymptotically equivalent under the null hypothesis to the corresponding trace statistics based on the true (unknown) break point for each of the procedures. Appropriate asymptotic critical values are given in the paper or can be obtained from the response surface given in Table 4 of Johansen *et al.* (2000).

Asymptotics – 2

- ▶ The second and third results crucially depend on the rate of consistency established for the break fraction estimator in the first result, where a break occurs.
- ▶ The results allow us to conclude that where a trend break is present the trace statistics based on the estimated trend break points are asymptotically equivalent under the null hypothesis to the corresponding trace statistics based on the true (unknown) break point for each of the procedures. Appropriate asymptotic critical values are given in the paper or can be obtained from the response surface given in Table 4 of Johansen *et al.* (2000).
- ▶ Where no trend break is present, our results show that all three procedures correctly select the no break model for sufficiently large samples. The resulting no break trace statistic has the usual restricted linear trend limiting distribution tabulated in Table 15.4 of Johansen (1995).

Asymptotics – 3

- ▶ The trace statistics which result from the three procedures are consistent at rate $O_p(T)$ when the true co-integration rank is such that $r^* > r$. This result holds regardless of whether a trend break is present in the data or not. This implies, therefore, that the usual sequential approach to determining the co-integration rank - this procedure starts with $r = 0$ and sequentially raises r by one until for $r = \hat{r}$ the trace test statistic does not exceed the ξ level critical value for the test - outlined in Johansen (1995) can still be employed using the trace statistics which obtain from any of the three procedures. In particular, these sequential approaches will lead to the selection of the correct co-integrating rank with probability $(1 - \xi)$ in large samples, again regardless of whether a trend break occurs or not.

Simulations – 1

- ▶ We report a Monte Carlo study based on the VAR(2) simulation DGP:

$$y_t = \begin{pmatrix} y_t^{(1)} \\ (n-r) \times 1 \\ y_t^{(0)} \\ r \times 1 \end{pmatrix} = \begin{pmatrix} \mu_{0,1}^{(1)} & \mu_{1,1}^{(1)} \\ (n-r) \times 1 & (n-r) \times 1 \\ \mu_{0,1}^{(0)} & \mu_{1,1}^{(0)} \\ r \times 1 & r \times 1 \end{pmatrix} \begin{pmatrix} d_{0,t}(b^*) \\ d_{1,t}(b^*) \end{pmatrix} + \begin{pmatrix} u_t^{(1)} \\ u_t^{(0)} \end{pmatrix}$$

where

$$\left(I_n - \begin{pmatrix} a_{1,1} I_{n-r} & 0 \\ 0 & a_{0,1} I_r \end{pmatrix} L \right) \left(I_n - \begin{pmatrix} a_2 I_{n-r} & 0 \\ 0 & a_2 I_r \end{pmatrix} L \right) \begin{pmatrix} u_t^{(1)} \\ u_t^{(0)} \end{pmatrix} = \begin{pmatrix} e_t^{(1)} \\ e_t^{(0)} \end{pmatrix}$$

where the superscript (1) denotes the $I(1)$ component under $H(r)$ and superscript (0) the $I(0)$ component. Here $|a_{0,1}| < 1$, $|a_2| < 1$, while $a_{1,1} = 1$ for $H(r)$ and $|a_{1,1}| < 1$ for $H(n)$.

Simulations – 2

- ▶ The disturbances are generated by $e_t^{(1)} \sim \text{i.i.d.}N(0, I_{n-r})$ and to allow for cross-correlation, we specify

$$e_t^{(0)} = \rho \kappa e_t^{(1)} + \sqrt{1 - \rho^2} \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}N(0, I_r)$$

where κ is an $r \times (n - r)$ matrix of ones. Here ρ controls the degree of cross-correlation (where relevant) between the $I(0)$ and $I(1)$ parts of the system.

Simulations – 2

- ▶ The disturbances are generated by $e_t^{(1)} \sim \text{i.i.d.}N(0, I_{n-r})$ and to allow for cross-correlation, we specify

$$e_t^{(0)} = \rho \kappa e_t^{(1)} + \sqrt{1 - \rho^2} \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}N(0, I_r)$$

where κ is an $r \times (n - r)$ matrix of ones. Here ρ controls the degree of cross-correlation (where relevant) between the $I(0)$ and $I(1)$ parts of the system.

- ▶ We set $b^* = \lfloor \lambda^* T \rfloor$, for the set of trend break fractions $\lambda^* = 0.25, 0.50, 0.75$, and $\mu_{i,1}^{(j)} = c \iota$, $i, j = 0, 1$, where ι is a vector of ones and c is a scalar constant controlling the break magnitude. For simplicity, this imposes the same magnitudes for level and trend breaks, and in the $I(1)$ and $I(0)$ directions, with all breaks occurring at date b^* . The values $c = 0.8, 0.4, 0.2$ are used, along with $c = 0$, representing the case when no breaks occur.

Simulations – 3

- ▶ In this simulation DGP, when $a_{1,1} = 1$ the first $n - r$ components of u_t are $I(1)$ and the remaining r components are $I(0)$, implying r co-integrating vectors of the form $\beta = (0_{r \times (n-r)} : I_r)'$. When $|a_{1,1}| < 1$ the process is stationary in all directions (has rank n).

Simulations – 3

- ▶ In this simulation DGP, when $a_{1,1} = 1$ the first $n - r$ components of u_t are $I(1)$ and the remaining r components are $I(0)$, implying r co-integrating vectors of the form $\beta = (0_{r \times (n-r)} : I_r)'$. When $|a_{1,1}| < 1$ the process is stationary in all directions (has rank n).
- ▶ The diagonal structure of the simulation DGP may appear restrictive but in fact is quite general because the DGP (and the statistical methods used) is invariant to taking orthogonal linear combinations of the columns of α and β .

Simulations – 4

- ▶ We compare the finite sample size and power properties of the SC-VECM and SC-DIFF based LR tests with the VECM trace tests which: (i) always include the trend break with break fraction estimated under $H(r)$ (i.e. using $\hat{\lambda}_{r,\hat{p}_{1,r}}$ defined in SC-VECM), which we denote Break-VECM; (ii) never includes a trend break (appropriate for $c = 0$), denoted as VECM. Since the SC-VECM procedure selects between these two individual tests, they provide an informal benchmark for the performance of the SC-procedures.

Simulations – 4

- ▶ We compare the finite sample size and power properties of the SC-VECM and SC-DIFF based LR tests with the VECM trace tests which: (i) always include the trend break with break fraction estimated under $H(r)$ (i.e. using $\hat{\lambda}_{r, \hat{p}_{1,r}}$ defined in SC-VECM), which we denote Break-VECM; (ii) never includes a trend break (appropriate for $c = 0$), denoted as VECM. Since the SC-VECM procedure selects between these two individual tests, they provide an informal benchmark for the performance of the SC-procedures.
- ▶ None of the tests assume a priori knowledge of p , but determine its value in the manner outlined before, assuming a maximum possible value of $\bar{p} = 4$. Results based on 10,000 Monte Carlo replications for tests at the nominal (asymptotic) 0.05 level, for sample sizes of $T = 100$ and 200 .

Table 2. Finite sample size and power; estimated lag length; $n = 2$, $r = 0$, $p = 1$

λ^*	c	SC-VECM				SC-DIFF				Break-VECM				VECM				
$T = 100$																		
		$a_{1,1}:$	1.00	0.90	0.80	0.70	1.00	0.90	0.80	0.70	1.00	0.90	0.80	0.70	1.00	0.90	0.80	0.70
0.25	0.8		0.079	0.159	0.502	0.894	0.066	0.151	0.495	0.891	0.065	0.139	0.484	0.887	0.676	0.796	0.951	0.996
0.50	0.8		0.073	0.122	0.410	0.839	0.063	0.112	0.397	0.832	0.072	0.122	0.410	0.839	0.795	0.875	0.968	0.997
0.75	0.8		0.071	0.136	0.481	0.891	0.064	0.133	0.476	0.889	0.076	0.144	0.488	0.896	0.459	0.417	0.654	0.892
0.25	0.4		0.093	0.193	0.506	0.868	0.090	0.195	0.509	0.870	0.080	0.136	0.401	0.759	0.176	0.205	0.517	0.874
0.50	0.4		0.066	0.089	0.274	0.634	0.064	0.092	0.276	0.637	0.075	0.121	0.362	0.739	0.173	0.102	0.275	0.634
0.75	0.4		0.046	0.045	0.148	0.443	0.046	0.046	0.148	0.444	0.075	0.124	0.395	0.764	0.074	0.039	0.133	0.427
0.25	0.2		0.059	0.151	0.537	0.920	0.059	0.151	0.540	0.923	0.079	0.137	0.397	0.774	0.073	0.152	0.541	0.925
0.50	0.2		0.056	0.074	0.242	0.628	0.055	0.074	0.243	0.631	0.080	0.135	0.377	0.764	0.068	0.073	0.242	0.629
0.75	0.2		0.044	0.071	0.275	0.693	0.045	0.071	0.276	0.696	0.080	0.127	0.377	0.747	0.049	0.070	0.275	0.695
0.00	0.0		0.049	0.165	0.708	0.987	0.050	0.165	0.711	0.989	0.083	0.148	0.463	0.877	0.051	0.165	0.713	0.991
$T = 200$																		
		$a_{1,1}:$	1.00	0.94	0.88	0.82	1.00	0.94	0.88	0.82	1.00	0.94	0.88	0.82	1.00	0.94	0.88	0.82
0.25	0.8		0.049	0.169	0.671	0.977	0.048	0.167	0.669	0.976	0.048	0.169	0.671	0.977	0.974	1.000	1.000	1.000
0.50	0.8		0.050	0.136	0.571	0.951	0.047	0.133	0.568	0.950	0.050	0.136	0.571	0.951	0.997	1.000	1.000	1.000
0.75	0.8		0.055	0.165	0.658	0.976	0.052	0.161	0.656	0.976	0.056	0.165	0.658	0.976	0.922	0.993	1.000	1.000
0.25	0.4		0.114	0.343	0.784	0.977	0.110	0.343	0.784	0.977	0.056	0.164	0.596	0.920	0.367	0.457	0.839	0.988
0.50	0.4		0.066	0.168	0.580	0.918	0.063	0.165	0.577	0.914	0.055	0.132	0.523	0.892	0.420	0.396	0.754	0.967
0.75	0.4		0.049	0.101	0.388	0.791	0.048	0.101	0.389	0.790	0.056	0.154	0.593	0.927	0.172	0.104	0.346	0.768
0.25	0.2		0.076	0.186	0.617	0.952	0.076	0.186	0.617	0.952	0.060	0.148	0.490	0.840	0.099	0.187	0.618	0.953
0.50	0.2		0.052	0.070	0.289	0.732	0.052	0.070	0.289	0.733	0.057	0.132	0.465	0.835	0.083	0.069	0.288	0.732
0.75	0.2		0.040	0.051	0.238	0.679	0.040	0.051	0.238	0.679	0.062	0.132	0.472	0.818	0.051	0.050	0.236	0.679
0.00	0.0		0.046	0.235	0.875	0.999	0.046	0.235	0.876	0.999	0.061	0.167	0.617	0.969	0.047	0.235	0.877	0.999

Table 3. Finite sample size and power; estimated lag length; $n = 2$, $r = 0$, $p = 2$, $a_2 = 0.5$

λ^*	c	SC-VECM				SC-DIFF				Break-VECM				VECM				
$T = 100$																		
		$a_{1,1}:$	1.00	0.80	0.60	0.40	1.00	0.80	0.60	0.40	1.00	0.80	0.60	0.40	1.00	0.80	0.60	0.40
0.25	0.8		0.098	0.332	0.652	0.750	0.090	0.311	0.662	0.730	0.117	0.288	0.626	0.733	0.186	0.472	0.860	0.981
0.50	0.8		0.091	0.226	0.560	0.660	0.094	0.254	0.591	0.633	0.115	0.254	0.571	0.662	0.164	0.320	0.808	0.984
0.75	0.8		0.076	0.206	0.510	0.665	0.090	0.284	0.637	0.715	0.114	0.281	0.625	0.740	0.083	0.127	0.436	0.728
0.25	0.4		0.083	0.324	0.539	0.672	0.083	0.340	0.564	0.678	0.121	0.274	0.543	0.624	0.104	0.340	0.562	0.682
0.50	0.4		0.082	0.155	0.309	0.362	0.088	0.162	0.329	0.378	0.121	0.254	0.518	0.595	0.093	0.140	0.301	0.363
0.75	0.4		0.073	0.141	0.268	0.222	0.081	0.141	0.272	0.237	0.118	0.254	0.525	0.636	0.074	0.119	0.247	0.216
0.25	0.2		0.075	0.373	0.626	0.748	0.081	0.387	0.645	0.751	0.121	0.294	0.597	0.652	0.082	0.388	0.645	0.753
0.50	0.2		0.077	0.248	0.412	0.366	0.087	0.249	0.417	0.371	0.124	0.287	0.590	0.634	0.080	0.246	0.415	0.367
0.75	0.2		0.077	0.281	0.487	0.431	0.084	0.285	0.497	0.437	0.121	0.282	0.572	0.627	0.078	0.284	0.496	0.435
0.00	0.0		0.078	0.414	0.733	0.916	0.080	0.432	0.752	0.918	0.120	0.308	0.669	0.747	0.078	0.433	0.753	0.921
$T = 200$																		
		$a_{1,1}:$	1.00	0.90	0.80	0.70	1.00	0.90	0.80	0.70	1.00	0.90	0.80	0.70	1.00	0.90	0.80	0.70
0.25	0.8		0.089	0.402	0.875	0.984	0.061	0.361	0.896	0.993	0.073	0.340	0.860	0.983	0.311	0.690	0.978	0.999
0.50	0.8		0.070	0.276	0.797	0.970	0.065	0.283	0.831	0.984	0.073	0.277	0.797	0.970	0.320	0.592	0.959	0.999
0.75	0.8		0.058	0.287	0.828	0.983	0.065	0.348	0.885	0.993	0.071	0.327	0.850	0.985	0.137	0.253	0.766	0.974
0.25	0.4		0.075	0.365	0.827	0.967	0.062	0.371	0.842	0.976	0.073	0.278	0.706	0.902	0.111	0.377	0.848	0.979
0.50	0.4		0.062	0.172	0.614	0.893	0.059	0.217	0.677	0.918	0.072	0.250	0.692	0.896	0.097	0.162	0.606	0.902
0.75	0.4		0.049	0.123	0.468	0.800	0.057	0.167	0.516	0.817	0.071	0.259	0.697	0.893	0.059	0.102	0.431	0.787
0.25	0.2		0.065	0.424	0.894	0.985	0.063	0.430	0.903	0.989	0.078	0.295	0.754	0.940	0.073	0.431	0.903	0.989
0.50	0.2		0.058	0.224	0.645	0.892	0.059	0.224	0.639	0.891	0.076	0.281	0.746	0.935	0.067	0.220	0.637	0.891
0.75	0.2		0.059	0.275	0.714	0.923	0.060	0.277	0.717	0.924	0.077	0.282	0.729	0.920	0.062	0.274	0.715	0.925
0.00	0.0		0.057	0.523	0.980	0.999	0.061	0.530	0.986	0.999	0.078	0.333	0.862	0.990	0.057	0.531	0.986	0.999

Table 4. Finite sample size and power; estimated lag length; $n = 2$, $r = 1$, $p = 1$, $a_{0,1} = 0.0$, $\rho = 0.0$

λ^*	c	SC-VECM				SC-DIFF				Break-VECM				VECM				
$T = 100$																		
		$a_{1,1}$:	1.00	0.85	0.70	0.65	1.00	0.85	0.70	0.65	1.00	0.85	0.70	0.65	1.00	0.85	0.70	0.65
0.25	0.8		0.051	0.212	0.746	0.881	0.083	0.300	0.743	0.846	0.051	0.212	0.746	0.881	0.386	0.749	0.945	0.955
0.50	0.8		0.062	0.182	0.689	0.844	0.064	0.173	0.637	0.775	0.062	0.182	0.689	0.844	0.396	0.710	0.891	0.898
0.75	0.8		0.060	0.204	0.731	0.867	0.057	0.194	0.635	0.739	0.060	0.204	0.731	0.868	0.198	0.259	0.351	0.361
0.25	0.4		0.052	0.201	0.716	0.844	0.096	0.211	0.367	0.402	0.052	0.203	0.729	0.868	0.141	0.214	0.364	0.399
0.50	0.4		0.061	0.177	0.666	0.814	0.038	0.048	0.090	0.101	0.061	0.178	0.672	0.827	0.079	0.042	0.068	0.073
0.75	0.4		0.060	0.194	0.699	0.832	0.011	0.007	0.020	0.024	0.060	0.195	0.711	0.854	0.019	0.003	0.006	0.008
0.25	0.2		0.050	0.186	0.659	0.801	0.058	0.215	0.606	0.710	0.053	0.199	0.694	0.837	0.066	0.216	0.606	0.711
0.50	0.2		0.061	0.170	0.626	0.756	0.020	0.018	0.021	0.022	0.063	0.174	0.655	0.806	0.026	0.017	0.015	0.014
0.75	0.2		0.052	0.180	0.648	0.766	0.026	0.065	0.152	0.186	0.057	0.191	0.695	0.836	0.028	0.064	0.148	0.182
0.00	0.0		0.046	0.311	0.925	0.979	0.047	0.313	0.929	0.983	0.073	0.199	0.699	0.843	0.048	0.314	0.931	0.985
$T = 200$																		
		$a_{1,1}$:	1.00	0.94	0.88	0.82	1.00	0.94	0.88	0.82	1.00	0.94	0.88	0.82	1.00	0.94	0.88	0.82
0.25	0.8		0.050	0.140	0.493	0.886	0.046	0.133	0.471	0.857	0.050	0.140	0.493	0.886	0.302	0.641	0.945	0.997
0.50	0.8		0.060	0.124	0.437	0.847	0.060	0.122	0.425	0.817	0.060	0.124	0.437	0.847	0.350	0.647	0.943	0.995
0.75	0.8		0.049	0.132	0.471	0.876	0.052	0.136	0.470	0.863	0.049	0.132	0.471	0.876	0.278	0.567	0.914	0.991
0.25	0.4		0.048	0.141	0.493	0.886	0.166	0.463	0.805	0.924	0.048	0.141	0.493	0.886	0.261	0.499	0.837	0.938
0.50	0.4		0.055	0.121	0.430	0.839	0.124	0.327	0.636	0.811	0.055	0.121	0.430	0.839	0.258	0.451	0.746	0.860
0.75	0.4		0.047	0.133	0.465	0.870	0.053	0.108	0.220	0.296	0.047	0.133	0.465	0.870	0.090	0.112	0.208	0.271
0.25	0.2		0.048	0.142	0.482	0.881	0.069	0.112	0.239	0.314	0.048	0.142	0.482	0.881	0.080	0.113	0.239	0.313
0.50	0.2		0.055	0.123	0.426	0.830	0.021	0.007	0.014	0.023	0.055	0.123	0.427	0.830	0.031	0.006	0.012	0.018
0.75	0.2		0.050	0.133	0.463	0.867	0.005	0.001	0.002	0.003	0.050	0.133	0.463	0.867	0.007	0.000	0.000	0.000
0.00	0.0		0.049	0.200	0.720	0.984	0.049	0.200	0.720	0.985	0.063	0.130	0.446	0.850	0.049	0.200	0.721	0.985

Table 5. Finite sample size and power; estimated lag length; $n = 2$, $r = 1$, $p = 1$, $a_{0,1} = 0.5$, $\rho = 0.0$

λ^*	c	SC-VECM				SC-DIFF				Break-VECM				VECM				
$T = 100$																		
		$a_{1,1}$:	1.00	0.80	0.60	0.40	1.00	0.80	0.60	0.40	1.00	0.80	0.60	0.40	1.00	0.80	0.60	0.40
0.25	0.8		0.049	0.338	0.897	0.964	0.054	0.351	0.876	0.932	0.048	0.330	0.892	0.970	0.386	0.860	0.961	0.961
0.50	0.8		0.062	0.297	0.877	0.962	0.053	0.274	0.850	0.929	0.062	0.297	0.879	0.967	0.389	0.817	0.915	0.902
0.75	0.8		0.059	0.324	0.847	0.902	0.051	0.309	0.838	0.868	0.060	0.330	0.902	0.976	0.172	0.178	0.157	0.137
0.25	0.4		0.047	0.207	0.419	0.500	0.072	0.195	0.293	0.351	0.047	0.278	0.730	0.849	0.122	0.195	0.286	0.343
0.50	0.4		0.048	0.160	0.344	0.395	0.019	0.024	0.055	0.066	0.057	0.252	0.742	0.874	0.067	0.005	0.008	0.020
0.75	0.4		0.041	0.155	0.293	0.334	0.010	0.011	0.026	0.027	0.052	0.282	0.779	0.887	0.015	0.000	0.001	0.001
0.25	0.2		0.050	0.294	0.749	0.894	0.055	0.330	0.755	0.895	0.055	0.272	0.750	0.886	0.061	0.332	0.757	0.896
0.50	0.2		0.035	0.107	0.200	0.255	0.018	0.013	0.015	0.020	0.057	0.248	0.741	0.875	0.024	0.010	0.007	0.010
0.75	0.2		0.031	0.130	0.268	0.365	0.024	0.074	0.131	0.213	0.055	0.270	0.747	0.877	0.025	0.072	0.125	0.208
0.00	0.0		0.047	0.544	0.989	0.996	0.046	0.551	0.995	1.000	0.071	0.327	0.908	0.987	0.047	0.553	0.996	1.000
$T = 200$																		
		$a_{1,1}$:	1.00	0.92	0.84	0.76	1.00	0.92	0.84	0.76	1.00	0.92	0.84	0.76	1.00	0.92	0.84	0.76
0.25	0.8		0.051	0.239	0.790	0.993	0.045	0.225	0.768	0.980	0.051	0.239	0.790	0.993	0.511	0.945	1.000	1.000
0.50	0.8		0.057	0.201	0.739	0.986	0.056	0.193	0.718	0.971	0.057	0.201	0.739	0.986	0.550	0.939	0.999	1.000
0.75	0.8		0.050	0.221	0.774	0.993	0.050	0.223	0.763	0.983	0.050	0.221	0.774	0.993	0.446	0.886	0.995	1.000
0.25	0.4		0.048	0.229	0.770	0.946	0.124	0.389	0.577	0.634	0.048	0.230	0.777	0.982	0.289	0.436	0.581	0.620
0.50	0.4		0.057	0.197	0.711	0.949	0.043	0.130	0.295	0.363	0.057	0.198	0.714	0.978	0.231	0.170	0.145	0.126
0.75	0.4		0.050	0.216	0.748	0.909	0.022	0.035	0.089	0.101	0.050	0.216	0.759	0.985	0.062	0.004	0.004	0.006
0.25	0.2		0.045	0.207	0.679	0.803	0.069	0.167	0.331	0.466	0.048	0.221	0.748	0.962	0.082	0.167	0.328	0.464
0.50	0.2		0.056	0.184	0.642	0.743	0.019	0.001	0.006	0.007	0.057	0.190	0.694	0.957	0.032	0.000	0.000	0.000
0.75	0.2		0.049	0.197	0.671	0.712	0.011	0.005	0.003	0.005	0.051	0.209	0.741	0.965	0.015	0.003	0.000	0.000
0.00	0.0		0.049	0.360	0.951	1.000	0.049	0.360	0.951	1.000	0.062	0.211	0.743	0.987	0.050	0.361	0.952	1.000

Table 6. Break inclusion frequency for SC-VECM, $n = 2$

λ^*	c	$r = 0, p = 1$				$r = 0, p = 2, a_2 = .5$				$r = 1, p = 1, a_{0,1} = 0, \rho = 0$				$r = 1, p = 1, a_{0,1} = .5, \rho = 0$				
$T = 100$																		
		$a_{1,1} :$	1.00	0.90	0.80	0.70	1.00	0.80	0.60	0.40	1.00	0.85	0.70	0.65	1.00	0.80	0.60	0.40
0.25	0.8		0.897	0.942	0.943	0.922	0.557	0.548	0.648	0.716	1.000	1.000	1.000	0.999	0.990	0.971	0.889	0.862
0.50	0.8		0.980	0.993	0.993	0.992	0.710	0.791	0.895	0.881	1.000	1.000	1.000	1.000	0.999	0.998	0.989	0.983
0.75	0.8		0.926	0.961	0.965	0.962	0.635	0.666	0.747	0.765	1.000	1.000	1.000	1.000	0.998	0.990	0.940	0.923
0.25	0.4		0.266	0.090	0.058	0.049	0.346	0.172	0.125	0.060	0.999	0.998	0.982	0.967	0.832	0.589	0.310	0.323
0.50	0.4		0.426	0.220	0.144	0.116	0.410	0.250	0.190	0.106	1.000	0.999	0.994	0.987	0.950	0.761	0.454	0.438
0.75	0.4		0.297	0.128	0.076	0.063	0.368	0.198	0.142	0.066	1.000	0.999	0.988	0.977	0.929	0.686	0.369	0.365
0.25	0.2		0.095	0.020	0.017	0.019	0.300	0.120	0.084	0.032	0.960	0.956	0.903	0.870	0.414	0.312	0.191	0.217
0.50	0.2		0.130	0.027	0.022	0.023	0.316	0.136	0.087	0.030	0.994	0.993	0.968	0.946	0.771	0.548	0.267	0.278
0.75	0.2		0.102	0.023	0.019	0.019	0.305	0.125	0.085	0.027	0.982	0.980	0.944	0.919	0.635	0.457	0.241	0.255
0.00	0.0		0.055	0.014	0.014	0.015	0.287	0.108	0.080	0.035	0.035	0.018	0.020	0.022	0.051	0.034	0.045	0.044
$T = 200$																		
		$a_{1,1} :$	1.00	0.94	0.88	0.82	1.00	0.90	0.80	0.70	1.00	0.94	0.88	0.82	1.00	0.92	0.84	0.76
0.25	0.8		0.998	1.000	1.000	1.000	0.673	0.737	0.867	0.949	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.50	0.8		1.000	1.000	1.000	1.000	0.845	0.946	0.982	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.75	0.8		0.999	1.000	1.000	1.000	0.731	0.814	0.921	0.971	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.25	0.4		0.469	0.282	0.194	0.153	0.274	0.117	0.127	0.152	1.000	1.000	1.000	1.000	1.000	0.999	0.985	0.908
0.50	0.4		0.706	0.689	0.621	0.554	0.369	0.211	0.248	0.314	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.967
0.75	0.4		0.502	0.351	0.261	0.202	0.292	0.138	0.150	0.184	1.000	1.000	1.000	1.000	1.000	0.999	0.988	0.921
0.25	0.2		0.087	0.011	0.009	0.010	0.179	0.056	0.054	0.055	1.000	1.000	1.000	1.000	0.988	0.973	0.892	0.695
0.50	0.2		0.142	0.023	0.014	0.014	0.206	0.067	0.067	0.073	1.000	1.000	1.000	1.000	0.998	0.992	0.942	0.773
0.75	0.2		0.096	0.014	0.009	0.008	0.186	0.059	0.059	0.061	1.000	1.000	1.000	1.000	0.995	0.984	0.922	0.730
0.00	0.0		0.027	0.005	0.004	0.005	0.156	0.046	0.043	0.043	0.010	0.005	0.004	0.004	0.013	0.006	0.005	0.006

Simulations – 5

Here we summarise the main findings of the Monte Carlo experiments:

- ▶ The SC-VECM and SC-DIFF tests behave similarly for $r = 0$, but behave very differently for $r = 1$ (and $r > 1$). Here, SC-DIFF can be prone to low size and very low power when $c > 0$. In contrast, SC-VECM is well size-controlled everywhere and frequently has the ability to secure close to the better levels of power available from the VECM and Break-VECM tests in the environments for which they are intended to operate. On this basis we recommend the SC-VECM procedure for practical use.

Simulations – 5

Here we summarise the main findings of the Monte Carlo experiments:

- ▶ The SC-VECM and SC-DIFF tests behave similarly for $r = 0$, but behave very differently for $r = 1$ (and $r > 1$). Here, SC-DIFF can be prone to low size and very low power when $c > 0$. In contrast, SC-VECM is well size-controlled everywhere and frequently has the ability to secure close to the better levels of power available from the VECM and Break-VECM tests in the environments for which they are intended to operate. On this basis we recommend the SC-VECM procedure for practical use.
- ▶ The presence of stationary autocorrelation introduces some size distortions and power losses into all the SC-based procedures co-integration tests (commensurate with the effects on the Break-VECM and VECM tests), and can also make the SC selection of the break more difficult.

Simulations – 6

- ▶ The size of the trend break generally affects the SC break selection in predictable ways, with larger breaks easier to detect. Variations in the break fraction λ^* , on the other hand, can produce unexpected effects on SC-VECM through its differing effects on the SC step and the benchmark Break-VECM and VECM tests - generally, as might be expected, a break in the middle of the sample is easiest for the SC to detect, while the rejection frequencies of the misspecified VECM test in particular can be considerably greater for earlier breaks than later ones, and the interactions of these effects produce variations in the performance of the SC-VECM procedure that may appear unexpected but turn out to be somewhat explicable in these terms.

Conclusions – 1

- ▶ Considered the problem of testing for the co-integration rank in VAR processes of unknown lag order when a break in the deterministic trend component may be present at an unknown point in the sample.

Conclusions – 1

- ▶ Considered the problem of testing for the co-integration rank in VAR processes of unknown lag order when a break in the deterministic trend component may be present at an unknown point in the sample.
- ▶ We outlined an approach based on the use of information criteria designed to choose the autoregressive lag length and to select between the trend break and no trend break models, using a consistent estimate of the break fraction in the former case.

Conclusions – 1

- ▶ Considered the problem of testing for the co-integration rank in VAR processes of unknown lag order when a break in the deterministic trend component may be present at an unknown point in the sample.
- ▶ We outlined an approach based on the use of information criteria designed to choose the autoregressive lag length and to select between the trend break and no trend break models, using a consistent estimate of the break fraction in the former case.
- ▶ These procedures were shown to deliver asymptotically correctly sized and consistent co-integration rank tests regardless of whether a trend break occurs or not. By selecting the no break model when no trend break is present, these procedures avoid the potentially large power losses associated with tests which assume that a trend break date has occurred, when in fact it has not.

Conclusions – 2

- ▶ Interesting possible extensions of this work could allow for: the possibility of multiple trend breaks; wild bootstrap implementations of these procedures to allow for variance non-stationarity in the innovations; and, pseudo-GLS de-trending of the form considered in Saikkonen and Lütkepohl (2000), Lütkepohl *et al.* (2003) and Trenkler *et al.* (2007).

Conclusions – 2

- ▶ Interesting possible extensions of this work could allow for: the possibility of multiple trend breaks; wild bootstrap implementations of these procedures to allow for variance non-stationarity in the innovations; and, pseudo-GLS de-trending of the form considered in Saikkonen and Lütkepohl (2000), Lütkepohl *et al.* (2003) and Trenkler *et al.* (2007).
- ▶ We have focussed on *stochastic* rather than *deterministic* co-integration. In the latter the co-integrating vector also eliminates deterministic non-stationarity in the data, corresponding to the restrictions that $\delta_{1,0} = 0$ and $\delta_{1,1} = 0$ in (3). Versions of the rank tests proposed here which impose these restrictions could be used instead. This could result in more powerful tests where those restrictions hold on (3), but would come at the expense of uncontrolled size where they did not.