Karaman Örsal, Deniz Dilan; Arsova, Antonia

Conference Paper

A panel cointegration rank test with structural breaks and cross-sectional dependence

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Time Series Econometrics, No. D01-V3

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Karaman Örsal, Deniz Dilan; Arsova, Antonia (2016) : A panel cointegration rank test with structural breaks and cross-sectional dependence, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Time Series Econometrics, No. D01-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:
https://hdl.handle.net/10419/145822

A panel cointegration rank test with structural breaks and cross-sectional dependence

Antonia Arsova[1†], Deniz Dilan Karaman Örsal[1,2‡]

[1] *Center for Methods*
[2] *Institute of Economics*

Leuphana Universität Lüneburg
Scharnhorststr. 1, 21335 Lüneburg, Germany

August 26, 2016

Abstract

This paper proposes a new panel cointegration rank test which allows for a linear time trend with breaks and cross-sectional dependence. The new correlation-augmented inverse normal (CAIN) test is based on a novel modification of the inverse normal method and combines the p-values of individual likelihood-ratio trace statistics. A Monte Carlo study demonstrates its robustness to cross-sectional dependence and its superior size and power properties compared to other meta-analytic tests used in practice. The test is applied to investigate the exchange rate pass-through into import prices across different industries in seven Eurozone countries.

Keywords: panel cointegration rank test, structural breaks, cross-sectional dependence, common factors, likelihood-ratio, time trend
JEL classification: C12, C15, C33

*Financial support by the German Research Foundation (DFG) through the project KA-3145/1 − 2 is gratefully acknowledged.

†E-mail: antonia.arsova@leuphana.de
‡E-mail: deniz.oersal@leuphana.de

1 Introduction

Panel unit root and cointegration tests have been developed since the early 2000s with the aim to increase the power of single-unit tests. The so-called "first generation" tests rely on the assumption of independence between the panel units. In macroeconomic panel data, however, cross-sectional dependence arises naturally due to common shocks or spillover effects. If not accounted for, it may bias the outcome of the tests by inflating the type-I error rate above the nominal significance level. Another issue often observed in longer time series is that of structural breaks; it may also invalidate the test results if left unattended.

Focusing on cointegration rank testing, we address both issues simultaneously by extending the cointegration rank test of Trenkler et al. (2007) (henceforth the TSL test) to panel data with cross-sectional dependence. Under the assumptions of the test structural breaks are allowed in the deterministic parts of the data generating process (DGP), i.e. level and trend slope, but not in the cointegrating vector. This is in line with its interpretation as a long-run equilibrium relationship between the variables in the system. Furthermore, this framework allows for structural breaks both under the null and the alternative hypothesis, as the breaks do not affect the stochastic properties of the DGP. Our preference for the TSL test over another likelihood-based alternative, the cointegration rank test of Johansen et al. (2000) (henceforth JMN test), is motivated by its superior finite-sample properties demonstrated by Trenkler et al. (2007) and our own simulations.

Panel cointegration testing in the presence of structural breaks and cross-sectional dependence has only recently gained attention from researchers, leading to the development of the so called "third-generation tests". Westerlund and Edgerton (2008), for example, propose a simple panel test for no cointegration allowing for a level shift and a break in the cointegrating relation, but not for a break in the deterministic trend. They assume that the cross-sectional dependence is driven by stationary unobserved common factors, which might be seen as restrictive in practice. Banerjee and Carrion-i Silvestre (2015) relax these restrictions in their no-cointegration test and propose test with level shifts, level shifts and break in the cointegrating relation, or level and trend shifts. However, in the latter case they allow only for homogeneous number of breaks and break dates across the panel units. They point out that "the difficulty [in allowing for heterogeneous break dates] essentially lies in the dependence of the critical values of the tests on the location of the break dates when trend breaks are present". While the same holds true for the TSL test, our approach to combining information from individual cross-sections into a panel test accommodates heterogeneous number of breaks and break dates across units.

We extend the TSL test to the panel setting resorting to a new p-value combination method which allows the p-values to be correlated. p-value combination approaches offer much more flexibility than traditional pooling of individual test statistics, as they allow the specification of the deterministic terms, the lag order, the number and the location of the breaks, and even the time span of the data to vary over cross-sections.

Recent research on panel unit root and cointegration testing has benefited significantly from the "reinvention" of already existing methods for combining possibly dependent p-values. One example is a modification of the Bonferroni procedure proposed by Simes (1986) and employed by Hanck (2013) in his panel unit root test. In a Monte Carlo study he demonstrates its robustness to cross-sectional dependence induced by common factors. This avenue is further explored in the direction of cointegration rank testing in dependent panels by Arsova and Örsal (2016a), who also show empirically that a sufficient condition for the validity of Simes' procedure is not violated in a common-factor-driven

panel framework.[1]

In the present work we adopt a new, augmented version of yet another p-value combination method initially proposed by Stouffer et al. (1949) — the inverse normal test. The novelty of our approach lies in that we explicitly model the degree of cross-sectional correlation between the probits of the individual statistics and use it as a variance-inflation factor in the panel test statistic. In this regard the proposed test is similar to the modified inverse normal method of Hartung (1999), which also uses an estimate of the cross-sectional correlation of the probits, however in combination with a somewhat arbitrary correction factor κ.[2] Our estimator for the unobserved correlation of the probits is modelled as a function of a quantity easily measurable in practice, namely the average absolute cross-sectional correlation of the residuals of the individual VAR models. In a Monte Carlo study we demonstrate that our correlation-augmented inverse normal (CAIN) method for combining p-values of individual TSL tests has good size and power properties in finite samples. Its performance is preferable to that of Hartung's method over the whole range of possible values for the cross-sectional correlation. The test can be used to determine the cointegrating rank of the observed time series even when the cross-sectional dependence is driven by unobserved common factors, without decomposing the data into idiosyncratic and common components and testing these separately.

The remainder of the paper is organized as follows. The next section briefly reviews the test of Trenkler et al. (2007), while its extension to panel data by a modification of the normal inverse method is given in Section 3. Section 4 presents the response surface approach to modelling the cross-sectional correlation between the probits of the individual test statistics. Section 5 discusses the results of a Monte Carlo study comparing the CAIN-TSL test with other meta-analytic tests for DGPs with various degrees of heterogeneous cross-sectional dependence. Section 6 illustrates the use of the CAIN-TSL test by investigating the exchange rate pass-through (ERPT) in the Euro-area and the last section concludes. Supplementary material is provided in the Appendix.

2 The TSL test for the cointegrating rank

Our aim is to develop a panel test for the cointegrating rank which allows for structural breaks in the deterministic parts of the DGP. A natural first step is to take an existing single-unit test and to extend it to the panel setting. Two well-known alternatives are the likelihood-based test of Johansen et al. (2000) and its GLS detrended counterpart proposed by Trenkler et al. (2007). Both tests allow for up to two breaks in the level and/or trend slope, whereas the break dates are assumed to be known.[3] The problem of estimating the cointegrating rank of a VAR process with level shifts whose dates are unknown has been addressed by Lütkepohl et al. (2004) and Saikkonen et al. (2006); unfortunately no such test has been developed yet for the case of breaks in the trend slope. Therefore, another strategy for the break dates determination has to be adopted if trend breaks need to be considered.

[1] Sarkar (1998) shows that for the Simes' procedure to hold for p-values of dependent statistics, their multivariate distribution has to be *multivariate totally positive* (MTP_2). We refer to the latter article for a definition of this property.

[2] Hartung's approach has been used to account for cross-sectional dependence in panel unit root tests by Demetrescu et al. (2006) and Costantini and Lupi (2013).

[3] Theoretically more than two breaks could be specified, however response surface regressions approximating the limiting distributions of the test statistics are available for a maximum of two breaks for both tests. Including more than two breaks may also be seen as problematic in time series of a limited time span.

The first question that arises in practice is whether breaks in the trend have occurred, or allowing only for level shifts might be enough.[4] It can be answered by the newly proposed tests of Kejriwal and Perron (2010) and Sobreira and Nunes (2015), which are applicable to both $I(0)$ and $I(1)$ processes. The authors propose also sequential testing procedures to find out the number of trend breaks. Next, the break dates and the presence of unit roots have to be determined. This can be done simultaneously by using, for example, the univariate unit root tests of Lee and Strazicich (2003) or Lee et al. (2012), or by their panel extensions by Im et al. (2010) and Payne et al. (2015). All these tests allow for up to two breaks in the trend of the series and determine the break locations as a by-product. Alternatively, the test of Carrion-i Silvestre et al. (2009) can be used. Then the cointegration analysis may proceed by employing a suitable cointegration rank test. Our preference lies with the TSL test for the reasons outlined next.

Trenkler et al. (2007) find that the JMN test with a single trend break displays excessive size distortions in systems of larger dimensions while their test is correctly sized. Our own Monte Carlo simulations confirm these findings also when two breaks in the trend slope are present; the results are briefly summarized in Table 1. The JMN test is severely oversized for small T, with the size distortions persisting even for $T = 500$. The TSL test, in contrast, has approximately correct size at the 5% level for all values of T. Trenkler et al. (2007) further write that in the case of larger systems the power curves are very flat, making it "difficult to draw proper conclusions regarding the cointegrating rank". This fact, along with its preferable finite-sample properties, is our main motivation for extending the TSL test to the panel setting. Next we briefly describe its assumptions and the computation of the test statistic.

Table 1: Empirical size at 5% nominal level of the JMN and TSL tests with two structural breaks, $H_0 : r = 0$

Test	$T = 50$	$T = 100$	$T = 200$	$T = 500$
$\text{JMN}_{0.25,0.5}$	0.545	0.209	0.116	0.070
$\text{JMN}_{0.25,0.75}$	0.531	0.206	0.111	0.075
$\text{TSL}_{0.25,0.5}$	0.047	0.046	0.042	0.044
$\text{TSL}_{0.25,0.75}$	0.052	0.051	0.043	0.045

Subscripts denote break locations at relative sample lengths. Simulations based on a 3-variate VAR(2) process as in Wagner and Hlouskova (2010) with 5000 replications.

Let the observed data $Y_{it} = (Y_{1,it}, \ldots, Y_{m,it})'$ for cross-sectional unit i $(i = 1, \ldots, N)$ be generated by a stochastic VAR(s_i) process X_{it} added to a deterministic process. The latter consists of a constant, linear time trend and structural breaks in both the level and the trend slope at known individual-specific time(s) τ_i:

$$Y_{it} = \mu_{0i} + \mu_{1i}t + \delta_{0i}d_{it} + \delta_{1i}b_{it} + X_{it}, \quad t = 1, \ldots, T. \tag{1}$$

Here μ_{ji} and δ_{ji} $(j = 0, 1)$ are unknown $(m \times 1)$ parameter vectors, while d_{it} and b_{it} are dummy variables defined by $d_{it} = b_{it} = 0$ for $t < \tau_i$, and $d_{it} = 1$ and $b_{it} = t - \tau_i + 1$ for $t \geq \tau_i$. The break dates are assumed to occur at individual-specific fixed fractions of the sample size: $\tau_i = [T\lambda_i]$ with $0 < \underline{\lambda_i} < \lambda_i < \overline{\lambda_i}$, where $\underline{\lambda_i}$ and $\overline{\lambda_i}$ are specified real numbers

[4] The power loss from using a trend-breaks test when only level shifts are present can be substantial, see e.g. Trenkler et al. (2007). This issue, however, would be less of a concern in the panel setting, where the power increases as the number of cross-sections grows (see Table 3).

and $[\cdot]$ denotes the integer part of the argument. In other words, the breaks are assumed not to occur in the very beginning or in the very end of the sample, while $\underline{\lambda_i}$ and $\overline{\lambda_i}$ are allowed to be arbitrarily close to 0 and 1, respectively. The stochastic processes X_{it} are assumed to be at most $I(1)$ and cointegrated with cointegrating rank at most r:

$$X_{it} = A_{1i}X_{i,t-1} + \ldots A_{s_i i}X_{i,t-s_i} + \varepsilon_{it}, \quad t = 1, \ldots, T_i. \tag{2}$$

It is assumed that the $(m \times 1)$ vector $\varepsilon_{it} \sim iid(0, \Omega_i)$, where Ω_i is a positive definite matrix for each i. Further it is assumed that ε_{it} have finite moments of order $(4 + \nu)$ for some $\nu > 0$, $\forall i$.

Denoting the pairwise cross-sectional correlations of the elements of ε_{it} by $\rho_{il,jk} := corr(\varepsilon_{it,l}, \varepsilon_{jt,k})$ for $i, j = 1, \ldots, N$ and $l, k = 1, \ldots, m$, we make the following assumptions.

Assumption 1 The average absolute pairwise cross-sectional correlation between the innovations to the same variable converges to some fixed value $\rho_\varepsilon \geq 0$ as $N \to \infty$:

$$\lim_{N\to\infty} \frac{1}{mN(N-1)} \sum_{i\neq j}^{N} \sum_{l=1}^{m} |\rho_{il,jl}| = \rho_\varepsilon \,. \tag{3}$$

Assumption 2 The average absolute pairwise cross-sectional correlation between the innovations to different variables converges to zero:

$$\lim_{N\to\infty} \frac{1}{mN(N-1)} \sum_{i\neq j}^{N} \sum_{l\neq k}^{m} |\rho_{il,jk}| = 0. \tag{4}$$

We note that Assumption 1 is not strictly necessary, but rather eases the interpretation of the estimated average absolute correlation coefficient $\hat{\rho}_\varepsilon$. Albeit seemingly restrictive, Assumption 2 is rather a technical one, necessary for the computation of a suitable estimator of ρ_ε. It is motivated by the fact that strong correlations are less likely to occur between the innovations to different variables across units compared to those between the same variables. Furthermore, our Monte Carlo simulations demonstrate that the proposed panel test is robust to a certain degree of deviation from this assumption, hence it should not hinder the applicability of the test in practice. Both assumptions hold when, for example, a spatial type of dependence is assumed, or when the dependence is driven by variable-specific common factors.

The computation of the individual TSL test statistics proceeds by estimating the deterministic terms by reduced rank regression taking into account the structural breaks, and then computing a likelihood-ratio (LR) trace statistic from the trend-adjusted observations. For details on the procedure we refer to Trenkler et al. (2007).

3 The correlation-augmented inverse normal test

Let p_i denote the p-values of the individual TSL statistics for units $i = 1, \ldots, N$. Let t_i denote the corresponding probits, i.e. $t_i = \Phi^{-1}(p_i)$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. Assuming independence of the individual test statistics (and hence of their p-values and the corresponding probits t_i),

the inverse normal test has a standard $N(0,1)$ limiting distribution:

$$t = \frac{\sum_{i=1}^{N} \Phi^{-1}(p_i)}{\sqrt{N}} = \frac{\sum_{i=1}^{N} t_i}{\sqrt{N}} \Rightarrow N(0,1). \tag{5}$$

The inverse normal method was first introduced to the panel unit root testing literature by Choi (2001), who demonstrates by simulations that under cross-sectional independence it outperforms other p-value combination alternatives, in particular Fisher's inverse Chi-square method employed for panel unit root testing by Maddala and Wu (1999). As a further advantage of the inverse normal test Choi (2001) points out its applicability to panels with both finite and infinite cross-sectional dimension N.

Assuming multivariate normal distribution of the probits, Hartung (1999) proposes a modification of the weighted inverse normal method to accommodate dependence between the original test statistics. The real-valued weights of the individual probits are denoted by λ_i and are such that $\sum_{i=1}^{N} \lambda_i \neq 0$. In practice it is often assumed that $\lambda_i = 1, \forall i$. The dependence is captured by a single correlation coefficient ρ_t, which can be interpreted as a "mean correlation approximating the case of possibly different correlations between the transformed statistics" (Hartung, 1999). The variance of the denominator in (5) is then augmented with an estimator of the correlation between the individual probits $\hat{\rho}_{\mathrm{t}}^{*}$:

$$t\left(\hat{\rho}_{\mathrm{t}}^{*}, \kappa\right) = \frac{\sum_{i=1}^{N} \lambda_i t_i}{\sqrt{\sum_{i=1}^{N} \lambda_i^2 + \left[\left(\sum_{i=1}^{N} \lambda_i\right)^2 - \sum_{i=1}^{N} \lambda_i^2\right] \left[\hat{\rho}_{\mathrm{t}}^{*} + \kappa \cdot \sqrt{\frac{2}{(N+1)}}(1 - \hat{\rho}_{\mathrm{t}}^{*})\right]}}. \tag{6}$$

The estimator $\hat{\rho}_{\mathrm{t}}^{*}$ is computed as

$$\hat{\rho}_{\mathrm{t}}^{*} = \max\{-\frac{1}{N-1}, \hat{\rho}_{\mathrm{t}}\}, \text{ where} \tag{7}$$

$$\hat{\rho}_{\mathrm{t}} = 1 - \frac{1}{N-1} \sum_{i=1}^{N} \left(t_i - \frac{1}{N} \sum_{i=1}^{N} t_i\right)^2. \tag{8}$$

The correction term $\kappa\sqrt{\frac{2}{(N+1)}}(1 - \hat{\rho}_{\mathrm{t}}^{*})$, which simply scales the standard deviation of $\hat{\rho}_{\mathrm{t}}$ by a factor κ, aims to avoid a systematic underestimation of the denominator in eq. (6). For the κ parameter Hartung suggests two alternative values: $\kappa_1 = 0.2$ and $\kappa_2 = 0.1 \cdot \left(1 + \frac{1}{N-1} - \hat{\rho}_{\mathrm{t}}^{*}\right)$, where κ_2 is suitable mainly for smaller $\hat{\rho}_{\mathrm{t}}^{*}$. However, he provides no guidance as to where the threshold between "small" and "large" $\hat{\rho}_{\mathrm{t}}^{*}$ should be.

Demetrescu et al. (2006) are the first to employ Hartung's modified inverse normal method to develop a panel unit root test allowing for cross-sectional dependence by generalizing Hartung's approach in two directions. First, they prove that the correlation between the individual test statistics needs not be constant for the limiting $N(0,1)$ distribution to hold. Second, they show that a necessary and sufficient condition for the limiting normality of the panel test statistic is the multivariate distribution of the individual test statistics to have a Gaussian copula. This theoretical result is, however, difficult to verify in practice; hence the authors proceed to demonstrate by simulation the applicability of Hartung's method to dependent individual ADF unit root tests. They employ the version of the test with $\kappa_2 = 0.2$ and unit weights $\lambda_i = 1, \forall i$, and find that for medium and strong cross-sectional correlation it generally observes the nominal significance level at 5% and 10%. For weak correlation the test is rather undersized and therefore not recommended

for use with more than 5 cross-sectional units.

Following the approach of Demetrescu et al. (2006), Costantini and Lupi (2013) use Hartung's modification with the value $\kappa_1 = 0.2$ in their simple panel-CADF test for unit roots. In order to mitigate the size distortions for weak cross-sectional dependence, they propose a type of "switching algorithm" between the regular and the modified inverse normal test ((5) and (6), respectively) based on the outcome of the CD test for cross-sectional correlation of Pesaran (2015). When the CD test rejects the null of weak correlation, (6) is used, otherwise Choi's test (5) is utilized. We argue that this strategy is sub-optimal because of two reasons. Firstly, the implicit null hypothesis and the finite-sample properties of the CD test depend on the relative expansion rates of T and N. It tends to over-reject if T is large relative to N and the exponent of cross-sectional dependence $\alpha \in (1/4, 1/2]$, a situation which in our view might well occur in macroeconometric panels.[5] Hence Hartung's modification might get preferred over Choi's test too often, leading to a loss of power. Secondly, the CD statistic accounts for the *average* correlation coefficient so that in its calculation large correlations of the opposite sign will cancel out. We argue that even negative correlations between the innovations to the DGPs lead to positive correlation between the individual test statistics. As our Monte Carlo study demonstrates, a better way to quantify the degree of cross-sectional dependence is to look at the mean *absolute* residual correlation.

We therefore propose a new, improved version of the inverse normal test for combination of correlated individual TSL test statistics for cointegration with structural breaks. It does not require neither switching between Choi's test and its Hartung's modification, nor arbitrary choice between different variance-inflating correction factors. This new test, which we name CAIN-TSL test, is based on a novel estimate $\tilde{\rho}_t$ of the correlation between the probits. The panel test statistic for the composite null hypothesis

$$H_0 : r_i = r, \, \forall i = 1, \ldots, N \tag{9}$$

against the alternative

$$H_1 : r_i > r \;\text{ for at least one } i \tag{10}$$

is given by

$$t(\tilde{\rho}) = \frac{\sum_{i=1}^{N} t_i}{\sqrt{N + (N^2 - N) \cdot \tilde{\rho}_{\mathrm{t}}}} \underset{H_0}{\sim} N(0, 1). \tag{11}$$

The CAIN-TSL test can be applied for testing $H_0 : r_i = r$, $\forall i$ at each step $r = 0, \ldots, m-1$ of the sequential rank testing procedure. We estimate $\tilde{\rho}_{\mathrm{t}}$ as a function of the dimension of the system m, the hypothesized cointegrating rank r and the estimated mean absolute correlation ρ_ε between the innovations of the individual DGPs. This function results from a response surface regression of the values of $\tilde{\rho}_{\mathrm{t}}$, estimated by simulation, on m, r and ρ_ε; the procedure is described in detail in the next section.

[5] We refer to Pesaran (2015) for a definition of the exponent of cross-sectional dependence.

4 Response surface regressions for the correlation of the individual probits

Common shocks to the innovations of the DGPs for the individual units, or generally cross-sectional correlation of the innovations leads to correlation between the individual TSL test statistics. This correlation transfers also to the p-values and their probits, and therefore has to be taken into account in the construction of the inverse normal test statistic in order to achieve a correctly sized test.

We argue that ρ_{t} can be inferred from its origin ρ_ε, which, in turn, can be consistently estimated in practice as N and T grow. It is, however, difficult to derive analytically how correlation between the innovations translates into correlation between the LR statistics, as the latter are complex non-linear functions of the observed data. We therefore resort to simulation methods to estimate this link. In our large-scale simulation study we estimate the average correlation of the probits $\tilde{\rho}_{\mathrm{t}}$ for different values of the absolute cross-sectional correlation between the innovations ρ_ε, controlling for the system dimension m, the hypothesized cointegrating rank r, and the time and cross-sectional dimensions T and N, respectively. We then estimate the relationship $\rho_t = \rho_t(\rho_\varepsilon, m, r)$ by response surface regression.

Estimation of the average absolute cross-sectional correlation between the process innovations The cornerstone of our proposed solution $\rho_{\mathrm{t}} = \rho_{\mathrm{t}}(\rho_\varepsilon, m, r)$ is an estimator $\hat{\rho}_\varepsilon$ of the cross-sectional correlation of the process innovations. In this regard we follow Pesaran (2015) and estimate $\hat{\rho}_\varepsilon$ from the residuals of the individual VAR(s_i) models under the null hypothesis $H_0 : r = 0$. Our estimation methodology, however, differs from his one in two aspects. First, as we are dealing with a panel of multivariate systems, the average cross-sectional correlation coefficient needs to be redefined for this richer data structure. We assume that strong cross-correlations are more likely to appear between the shocks to the same variables than between those to different variables. Hence averaging has to be performed only across the pairwise correlations between the same-variable residuals in order to avoid underestimation of ρ_ε and subsequently of ρ_{t}. Such underestimation would lead to an oversized panel test. That is, for $\hat{\rho}_{il,jl}$ denoting the estimated correlation between the residuals for variable l in units i and j, $\hat{\rho}_\varepsilon$ is to be estimated from the average of $|\hat{\rho}_{il,jl}|$ over $l = 1, \ldots, m$ and $i \neq j$ where $i, j = 1, \ldots, N$. Second, we take the absolute value of the estimated correlations $\hat{\rho}_{il,jl}$. In simple averaging positive and negative correlations will cancel out, thus leading to underestimation of the true degree of cross-sectional correlation.

Therefore, the estimated absolute residual correlations $|\hat{\rho}_{il,jl}|$ are averaged over l and $i \neq j$, resulting in $\hat{\rho}_\varepsilon$:

$$\hat{\rho}_\varepsilon = \frac{2}{mN(N-1)} \sum_{i=2}^{N} \sum_{j=i+1}^{N} \sum_{l=1}^{m} |\hat{\rho}_{il,jl}|. \tag{12}$$

Estimating $\hat{\rho}_\varepsilon$ only under $H_0 : r_i = 0, \forall i$ in the beginning of the sequential testing procedure for $r = 0, \ldots, m-1$ requires further justification. Ideally, the correlation between the probits would be inferred from the correlation between the residuals under each null hypothesis and then used in the correlation-augmented inverse normal panel test. In order to keep the procedure feasible for practical work, however, we decide to do this inference only once by determining $\hat{\rho}_\varepsilon$ only under the null of no cointegration, similarly to the lag orders s_i of the VAR processes. We note that the estimation error in $\hat{\rho}_\varepsilon$ is

negligible under the true $H_0 : r_i = 0, \forall i, m$, while $\hat{\rho}_\varepsilon$ gets only slightly underestimated for higher true cointegrating ranks (see Table 9 in the Appendix). The resulting slight underestimation of $\tilde{\rho}_\mathrm{t}$, however, would be negligible in practice for moderate ρ_ε; for higher ρ_ε it could even be beneficial, given that the TSL test is severely undersized when testing for higher ranks under the null.

Simulations design We estimate the relationship between ρ_ε and ρ_t in a large-scale simulation study. The data are generated according to (1) with the stochastic processes X_{it} following an m-variate VAR(1) Toda processes (see e.g. Toda (1995)) for $m = 2, 3, 4, 5$:

$$X_{it} = \begin{pmatrix} I_d & 0_{(d\times r)} \\ 0_{(r\times d)} & \Psi_r \end{pmatrix} X_{i,t-1} + \varepsilon_{it}, \tag{13}$$

where I_d denotes the identity matrix of dimension $d = m - r$, $0_{(d\times r)}$ is a zero matrix of the corresponding dimension and r is the cointegrating rank of the process. Ψ_r is a diagonal matrix of dimension $(r \times r)$ with diagonal elements $\psi_i \sim i.i.d.\, U(0, 1)$, $i = 1, \ldots, r$ which govern the dynamics of the stationary elements of the process for $r > 0$; the uniform $(0, 1)$ distribution of ψ_i provides for generality. As the individual LR trace statistics of the TSL test are invariant to the values of the deterministic terms μ_{0i}, μ_{1i}, δ_{0i} and δ_{1i} for given VAR order s_i and break date(s) τ_i (Trenkler et al., 2007, p. 340), we have set $\mu_{ji} = \delta_{ji} = 0$, $j = 0, 1$.

For the estimation of the individual LR trace statistics for each $i = 1, \ldots, N$ the number of breaks is randomly chosen between 1 and 2 with 50% chance each, and the break locations are chosen at relative sample length(s) $\lambda_i \sim i.i.d.\, U(0.15, 0.85)$. In the case of two breaks the minimal distance between them is set to $0.2T$. These values are in line with what is usually assumed in the literature on cointegration with structural breaks – the breaks are not allowed to be too close to the beginning or to the end of the sample, neither are they allowed to be too close to each other.

The innovations ε_{it} are drawn from a multivariate normal distribution with variance-covariance matrix Σ, where $\Sigma = R \otimes \Omega$ is generated as in Wagner and Hlouskova (2010) with

$$R = \begin{pmatrix} 1 & \rho_\varepsilon & \cdots & \rho_\varepsilon \\ \rho_\varepsilon & 1 & \cdots & \ddots \\ \vdots & \ddots & \ddots & \rho_\varepsilon \\ \rho_\varepsilon & \cdots & \rho_\varepsilon & 1 \end{pmatrix}_{(N\times N)}, \tag{14}$$

and $\Omega_{(m\times m)}$ being a random correlation matrix generated independently for each replication as described in Costantini and Lupi (2013, p. 283).

For each value of m, testing $H_0 : r_i = r_0$ for $r_0 = 0, \ldots, m - 1$ is considered in a separate experiment.

We let ρ_ε vary over a grid of 24 equally spaced values in the range $[0.04, 0.96]$, in order to be able to adequately fit an interpolating curve to the estimated average values $\tilde{\rho}^*_\mathrm{t}$ of ρ_t over $\hat{\rho}_\varepsilon$.

Each simulation experiment is repeated 100000 times for three combinations of the panel dimensions T and N: $(T, N) \in \{(500, 5), (500, 10), (1000, 5)\}$. For each combination (T, N) an average $\tilde{\rho}^*_\mathrm{t}$ is computed by means of Fisher's Z-transformation from the $(N \times N)$ correlation matrix of the probits based on the $(100000 \times N)$ matrix of independent observations on $(t_1, \ldots, t_N)$. For given ρ_ε and T, the estimated $\tilde{\rho}^*_\mathrm{t}$ is practically invariant

to the number of cross-sections, which is not surprising, as it is logical to expect that it would depend on the degree of dependence between the processes of any two cross-sections, but not on the number of units. Preliminary simulations corroborate this conjecture – see the right panel of Figure 1 in the Appendix. For robustness the estimated $\tilde{\rho}_{\mathrm{t}}^{*}$'s for all three combinations of (T, N) are subsequently modelled in the response surface regressions for $\tilde{\rho}_{\mathrm{t}}^{*} = \hat{\tilde{\rho}}_{\mathrm{t}}(\rho_{\varepsilon}, m, r)$.

In practice, by the Law of Large Numbers increasing N would lead to more precise estimation of $\hat{\rho}_{\varepsilon}$ and subsequently also of $\tilde{\rho}_{\mathrm{t}}^{*}$. Because of the large number of replications in the current simulation exercise, considering a wider range of values for N is not necessary and we concentrate only on values which are typical for macroeconometric studies. With regard to the choice of values for the parameter T, we motivate it by the fact that the estimated $\tilde{\rho}_{\mathrm{t}}^{*}$ converges to ρ_{t} from below as $T \to \infty$ (see the left panel of Figure 1), with the differences between $T = 500$ and $T = 1000$ being virtually negligible for all practical purposes. Indeed, for small T the estimated $\tilde{\rho}_{\mathrm{t}}^{*}$ might be lower than its asymptotic large-T value and could thus potentially get overestimated. We argue, however, that the use of the large-T $\tilde{\rho}_{\mathrm{t}}^{*}$ even for small T's in the panel setting would have beneficial rather than detrimental effects. It is well known that individual likelihood-based cointegration tests tend to be oversized for $H_0 : r = 0$ when T is small, and these size distortions can get magnified in the panel setting as the cross-sectional dimension increases (see, e.g., Demetrescu and Hanck (2012) and Arsova and Örsal (2016b)). Using the asymptotic $\tilde{\rho}_{\mathrm{t}}^{*}$ might help mitigate this issue, as slight overestimation of the cross-sectional correlation of the probits might inflate the variance of the panel statistic thus offsetting the inherent size distortions of the individual tests.

Response surface regressions We now turn our attention to modelling the estimator $\hat{\tilde{\rho}}_{\mathrm{t}}$ of $\tilde{\rho}_{\mathrm{t}}^{*}$. To obtain the response surface, we regress $\tilde{\rho}_{\mathrm{t}}^{*}$ on polynomials of the system dimension m, the cointegrating rank under the null hypothesis r, and the mean absolute correlation between the innovations ρ_{ε}. No constant is included in the regression and all regressors are multiples of ρ_{ε}, so that all estimates $\hat{\tilde{\rho}}_{\mathrm{t}}$ of $\tilde{\rho}_{\mathrm{t}}^{*}$ are equal to 0 when $\rho_{\varepsilon} = 0$.

The goodness of fit measure of the estimated regression is $R^2 = 0.9993$, which renders the approximation very good for practical purposes. The estimated coefficients are given in Table 2. They can be used to compute the estimate $\hat{\tilde{\rho}}_{\mathrm{t}}$ of the unknown correlation of the probits given m, r and the estimated $\hat{\rho}_{\varepsilon}$ for any system dimension $m \leq 5$. Then the CAIN-TSL test statistic is computed as in (11) with $\hat{\tilde{\rho}}_{\mathrm{t}}$ in the place of the unknown ρ_t.

Table 2: Response surface coefficients for correlation between the probits of the individual TSL LR trace statistics

Term	Estimated coefficient
${\rho_\varepsilon}^2$	0.6319575
$\sqrt{m} \cdot {\rho_\varepsilon}^2$	−0.5193669
$\sqrt{m} \cdot {\rho_\varepsilon}^4$	0.2721753
$\frac{r}{m} \cdot {\rho_\varepsilon}^2$	0.1821374
$\frac{r}{m} \cdot {\rho_\varepsilon}^4$	−0.0856903
$(r \cdot \rho_\varepsilon)^2$	0.0041125
$r \cdot {\rho_\varepsilon}^2$	0.0766267
$r \cdot {\rho_\varepsilon}^4$	−0.1008678
$\sqrt{m-r} \cdot {\rho_\varepsilon}^2$	0.1874919
$\frac{1}{m-r} \cdot {\rho_\varepsilon}^2$	0.1410229
$\frac{1}{m-r} \cdot {\rho_\varepsilon}^4$	−0.2029126
$(m-r)^2 \cdot {\rho_\varepsilon}^2$	0.0052557
$(m-r)^4 \cdot {\rho_\varepsilon}^4$	−0.0000327

5 Monte Carlo study

5.1 Simulation study design

The finite sample properties of the CAIN-TSL are examined by simulations in an empirically relevant case using a three-variate VAR(2) DGP as in the study of Wagner and Hlouskova (2010).

The general form the DGP is:

$$Y_{it} = \mu_{0i} + \mu_{1i} t + \delta_{0i} d_{it} + \delta_{1i} b_{it} + X_{it}, \tag{15}$$

$$X_{it} = \begin{pmatrix} a_{11}^i & 0 & 0 \\ 0 & a_{12}^i & 0 \\ 0 & 0 & a_{13}^i \end{pmatrix} X_{i,t-1} + \begin{pmatrix} a_{21}^i & 0 & 0 \\ 0 & a_{22}^i & 0 \\ 0 & 0 & a_{23}^i \end{pmatrix} X_{i,t-2} + u_{it}, \tag{16}$$

$$u_{it} = \gamma_i' f_t + \varepsilon_{it}, \tag{17}$$

$$\varepsilon_{it} \sim i.i.d.\, N(0, \Omega_i). \tag{18}$$

The cointegrating properties of the process are determined by the roots q_{1j}^i, q_{2j}^i of the autoregressive polynomial, which are linked to the coefficients of the autoregressive matrices by $a_{1j}^i = \frac{1}{q_{1j}^i} + \frac{1}{q_{2j}^i}$ and $a_{2j}^i = -\frac{1}{q_{1j}^i q_{2j}^i}, j = 1, 2, 3$.

Following Wagner and Hlouskova (2010), for a system with cointegrating rank zero we set $q_{1j}^i = 1$ and $q_{2j}^i \sim U(1.8, 3)$, $j = 1, 2, 3$. Power is investigated in a setting when all roots are sufficiently away from unity (case A) and also in a near-unit root setting (case B). For cointegrating rank one in case A we let $q_{11}^i \sim U(1.3, 1.7)$ or, in case B, $q_{11}^i \sim U(1, 1.3)$, while $q_{21}^i \sim U(1.5, 2.5)$, $q_{1j}^i = 1$ and $q_{2j}^i \sim U(1.8, 3)$ for $j = 2, 3$ in both cases. Finally, for cointegrating rank two we again let $q_{11}^i \sim U(1.3, 1.7)$ for case A or $q_{11}^i \sim U(1, 1.3)$ for case B. The remaining roots for both cases are $q_{12}^i, q_{2j}^i \sim U(1.5, 2.5)$ for $j = 1, 2$, while $q_{13}^i = 1$ and $q_{23}^i \sim U(1.8, 3)$. All roots are drawn separately for each unit.

We let a 3-dimensional vector of variable-specific common factors $f_t \sim i.i.d.\, N(0, I_3)$ drive the cross-sectional dependence through heterogeneous loadings. The factor loadings

γ_i are simulated as diagonal $(k \times m)$ matrices with (a) $i.i.d.\,U\,(-0.4, 0.4)$, (b) $i.i.d.\,U\,(0, 1)$ or (c) $i.i.d.\,U\,(-1, 3)$ entries, drawn separately for each unit. The robustness of the CAIN-TSL test against violations of Assumption 2 is investigated by letting γ_i be an unrestricted matrix with $i.i.d.\,U\,(0, 1)$ or $i.i.d.\,U\,(-1, 3)$ elements. This corresponds to all factors affecting all variables simultaneously.

Throughout we set $\mu_{ji} = \delta_{ji} = 0$, $j = 0, 1$, as in Trenkler et al. (2007). We again allow for a random number of breaks (1 or 2, with 50% chance each) and the break locations are chosen at relative fraction(s) of the sample size $\lambda_i \sim i.i.d.\,U(0.15, 0.85)$. In the case of two breaks the minimal distance between them is set to $0.2T$. The processes X_{it} are initialised with 0 and the first 50 observations are discarded to mitigate the effect of initial values. The individual-specific random correlation matrices Ω_i of the idiosyncratic errors ε_{it} are simulated, as before, as described in Costantini and Lupi (2013).

We consider all combinations of $T \in \{100, 200\}$ and $N \in \{5, 15, 25\}$. The lag order is assumed to be known and is set to its true value. The simulations are carried out in GAUSS and the number of replications is 5000. Nominal significance level $\alpha = 0.05$ applies in all cases.

The performance of the CAIN-TSL test is compared to that of the standard inverse normal test of Choi without correction for the cross-sectional dependence, both variants of Hartung's modified inverse normal test with κ_1 and κ_2, respectively, and the multiple testing procedure of Simes (1986). For the latter test the individual p-values of the test statistics are ordered in ascending order as $p_{(1)} \leq \ldots \leq p_{(N)}$, and the joint null hypothesis $H_0 : r_i = r, \forall i$, is rejected at significance level α if $p_{(i)} \leq \frac{i\alpha}{N}$ for any $i = 1, \ldots, N$.

5.2 Simulation results

The size and power results under $H_0 : r_i = 0, \forall i$ when the true rank is zero and one, respectively, are presented in Table 3.

When the factor loading matrix γ_i is diagonal with entries $U(-0.4, 0.4)$, the cross-sectional dependence is very weak. We note that due to the factor loadings being centered around zero, the mean (over all replications) estimated cross-sectional correlation between the residuals of the individual VAR models in first differences is zero, while the estimated mean absolute cross-sectional correlation is $\bar{\hat{\rho}}_\varepsilon = 0.089$ for $T = 100$ and $\bar{\hat{\rho}}_\varepsilon = 0.068$ for $T = 200$. The mean (again over all simulations) estimated correlation between the probits for the CAIN-TSL test is $\hat{\bar{\rho}}_{\mathrm{t}} = 0.001$. In this case all variants of the inverse normal test, including CAIN-TSL, become undersized as N grows. This results from the individual TSL tests begin slightly undersized; the size distortions get magnified in the panel setting as N increases. Hartung's test with κ_1 has the most severe size distortion, inline with the findings of Demetrescu et al. (2006) that it is not suitable for more than $N = 5$ units when the cross-sectional correlation is low. Simes' test gets slightly oversized with increasing N, and the size of CAIN-TSL is comparable with that of Choi's test and Hartung's test with κ_2. In terms of power, the CAIN-TSL test performs best together with Choi's in both case A and case B, as it correctly captures the low cross-sectional correlation of the probits.

Diagonal factor loading matrix with $U(0, 1)$ entries generates moderate cross-sectional correlation between the innovations – the mean estimated absolute cross-sectional correlation of the innovations is $\bar{\hat{\rho}}_\varepsilon = 0.186$ for $T = 100$ and $\bar{\hat{\rho}}_\varepsilon = 0.180$ for $T = 200$. As the estimated average correlations between the probits are already around $\hat{\bar{\rho}}_{\mathrm{t}} = 0.007$ and $\hat{\bar{\rho}}_{\mathrm{t}} = 0.005$, respectively, Choi's test has size close to the nominal one, along with the CAIN-TSL test. Both Hartung's tests again become undersized as N grows. Considering power, Choi's test performs best because the nominator of the panel test statistic is

inflated by the unattended cross-sectional dependence. The CAIN-TSL test has slightly lower power than Choi's for $T = 100$ in case B, but in case A and for $T = 200$ in case B it performs equally well.

Relatively strong cross-sectional correlation is introduced by a diagonal factor loading matrix with $U(-1, 3)$ entries. The mean estimated absolute cross-sectional correlation of the innovations $\bar{\hat{\rho}}_\varepsilon$ is already 0.413 for both $T = 100$ and $T = 200$, with estimated average correlation between the probits varying between 0.020 and 0.046. In this case Choi's test and Hartung's test with κ_2 become oversized for large N, while the size of Hartung's test with κ_1 and of CAIN-TSL fluctuates around the desired 5% level. In terms of power, in both cases A and B the CAIN-TSL outperforms Hartung's tests and the test of Simes', being second only compared to the oversized standard inverse normal. We note that due to the high correlation between the cross-sections, the power increase over N for all tests is less significant compared to the case of low cross-sectional correlation.

Lastly, with unrestricted factor loadings we investigate the performance of the tests when all variables are assumed to be correlated to the same extent, on average, over the cross-sections. When $\gamma_i \sim U(0, 1)$, the mean estimated absolute correlation between the innovations over all replications is 0.355. Choi's test becomes oversized for high N, while both Hartung's κ_1 and the CAIN-TSL tests tend to become undersized. The CAIN-TSL test performs best in terms of power without becoming oversized for both cases A and B. When $\gamma_i \sim U(-1, 3)$, the mean estimated absolute correlation between the innovations over all replications is already 0.44. The CAIN-TSL test tends to become oversized as N grows with the empirical size reaching 10% for $T = 200$ and $N = 25$ at the 5% level, while Hartung's κ_1 test observes the nominal size. As these experimental settings violate Assumption 2 for the CAIN-TSL test, we conclude that it is robust to such violations when the estimated mean absolute cross-sectional correlation between the innovations $\hat{\rho}_\varepsilon$ is low to moderate, or when the number of cross-sections is small. For higher estimated $\hat{\rho}_\varepsilon$ or when $N > 10$ we'd rather recommend to use Hartung's modification with $\kappa_1 = 0.2$.

Size and power under $H_0 : r_i = 1, \forall i$ for cases A and B are presented in Tables 4 and 5, respectively. The undersized individual TSL test leads to all five panel tests being severely undersized. As a result, the power in case A with $T = 100$ is significantly lower when compared to the corresponding power in Table 3; for $T = 200$ all tests perform well despite being undersized. In the presence of near-unit roots (case B, see Table 5), however, the power against hypothesized rank one when the true cointegrating rank is two is almost non-existent for short T. It nevertheless increases significantly over N when $T = 200$. For all correlation settings in case A Choi's test exhibits the highest power, closely followed by the CAIN-TSL test. The same holds for case B when the cross-sectional correlation is low to moderate (e.g. the factor loadings γ_i are diagonal matrices with $U(-0.4, 0.4)$ or $U(0, 1)$ entries). When the cross-sectional correlation is larger (e.g. diagonal factor loading matrices with $U(-1, 3)$ entries or unrestricted matrices with $U(0, 1)$ entries), the CAIN-TSL test has already slightly lower power than Hartung's tests. This is because it precisely corrects for the high cross-sectional correlation between the probits, which turns out to be unnecessary given the severely undersized results of the individual TSL tests. In these cases Choi's test is again the most powerful test (without becoming oversized under the true null $H_0 : r_i = 1, \forall i$), while Simes' test is again the least powerful one.

These results can be summarized as follows. When there is only low mean absolute cross-sectional correlation $\hat{\rho}_\varepsilon$, in terms of size the CAIN-TSL performs comparably to Choi's and Hartung's κ_2 test, being as powerful as Choi's and more powerful than both Hartung's tests. This conclusion holds regardless of the null hypothesis under consideration. When the cross-sectional correlation is high, the CAIN-TSL test offers best

size-power trade-off under the null hypothesis of no cointegration. This is particularly important in view of the sequential rank testing procedure, which begins with $H_0 : r_i = 0, \forall i$. Also, even when Assumption 2 is violated, the CAIN-TSL test may be used to test for no cointegration when the estimated $\hat{\rho}_\varepsilon$ is low to average, or when the number of cross-sections is not too large. Otherwise Hartung's test with κ_1 is recommended, as it controls the size better. When testing for higher cointegrating ranks under high cross-sectional correlation, then it might be preferable to employ Choi's test without any correction for the cross-sectional dependence, or Hartung's κ_2 test. They would yield high power without running the risk of an inflated Type I error probability, due to the lower than nominal size of the individual TSL test under such null hypotheses. Simes' simple intersection test, although almost correctly sized in all settings, is, in general, least powerful. It outperforms the other tests in terms of power only under null hypotheses $H_0 : r_i = r, \forall r$ for $r > 0$ when T is small; it these cases, however, its power is trivial.

6 Empirical illustration

We illustrate the use of the CAIN-TSL test by applying it to investigate the exchange rate pass-through in the Euro area. We employ the same model and data used by Banerjee and Carrion-i Silvestre (2015). The panel consists of monthly observations over the period 1995–2005 ($T = 123$) for nine different industries across seven Euro-Area countries: France, Netherlands, Germany, Italy, Ireland, Greece and Spain; for more details on the data we refer to the latter paper. The long-run relationship between (the logarithms of) the import price mp_t, the exchange rate er_t and the foreign price fp_t is given by

$$mp_t = \alpha + \beta t + b \cdot er_t + c \cdot fp_t + u_t,$$

where a linear trend term is included to reflect the trending behaviour of the individual variables.

Banerjee and Carrion-i Silvestre (2015) apply their proposed residual-based tests for panel cointegration with level shifts, and with level shifts and breaks in the cointegrating relationships. They find evidence of cointegration with one or two heterogeneous breaks in the levels and in the cointegrating relation, although not in the observable variables alone, but only after allowing up to twelve unobserved nonstationary common components into the model. They do not test for a stable cointegrating relation with trend breaks, as their proposed panel test for this setting does not accommodate heterogeneous break dates for the different units. Our aim, however, is to investigate exactly whether a stable long-term relationship exists, possibly around a broken linear trend. We further seek for evidence of cointegration between the observed variables, without decomposing them into unobservable common and idiosyncratic components. In this regard our analysis may be viewed as complementary to that of Banerjee and Carrion-i Silvestre (2015).

For determining the break dates we follow a mixed approach. We first run the univariate unit root tests with trend breaks of Narayan and Popp (2010) and Im et al. (2010) and let them determine the slope break dates endogenously for each series. After analysing clusters of break dates for the exchange rate and foreign price series, we identify the major economic and financial events that likely have triggered the change in trend slope.[6]

In this way we identify two major break points. The first one is in May 2000 and marks the end of a period of almost steady depreciation of the Euro against the US dollar since its introduction in 1999. On March 10, 2000 the NASDAQ Composite Index closed at (by

[6]The unit root test results indicate that the most series are integrated; they are available upon request.

then) an all-time high of over 5000, and the decline from this peak signalled the beginning of the dot-com bubble burst. In an effort to protect the economy, the Federal Reserve raised the federal funds rate by 25 basis points on March 21 and subsequently by another 50 basis points on May 16, effectively pinning the federal funds rate to a historic high of 6.5% until the beginning of 2001. From May 2000 until around May 2002 the exchange rate fluctuated around a level of 0.9 USD/EUR.

From May 2002 the Euro began appreciating against the dollar, and the foreign prices (fp) in our system began rising as well. Several factors can be pointed out as contributing to this change. The first one are the terrorist attacks on the World Trade Center in New York on September 11, 2001. In the weeks following the attacks investors became unsure about how terrorism would affect the US economy. Investors' confidence in the securities markets was next shaken by the accounting fraud scandals which preceded the bankruptcy of several major US corporations, e.g. Enron, Tyco and WorldCom, in the period December 2001 – May 2002. The decline of the dollar relative to the Euro can also be related to the US trade deficit – the share of the US in total world exports has been declining since 2001. At the same time, the Euro was physically introduced in January 2002 and became a full fledged currency in March 2002, after the coins and notes of the national currencies of the Euro-area countries were officially withdrawn from circulation.

In determining the break dates for the subsequent CAIN-TSL cointegration tests we only consider potential breaks in the exchange rate and foreign price series. The reasoning is that as the import prices are assumed to be determined by the exchange rate and the foreign prices, it is the structural breaks in these two variables which are carried over. We note that the exchange rate series have very similar dynamics across countries during the period 1996–1999, and that there is a single exchange rate for all countries following the introduction of the Euro. Furthermore, the foreign price series are also identical for all countries and only differ across industries. Therefore assuming homogeneous break dates across countries is not too restrictive.

For the cointegration analysis we consider nine industry panels, each with $N = 7$ country cross-sections. In this way we are able to test the ERPT hypothesis separately in each industry, as the degree of exchange rate pass-through is likely to differ across industries.

We first test for cointegration using the test of Saikkonen and Lutkepohl (2000) (henceforth SL) allowing for an intercept and trend, but no structural breaks. In this way we are able to identify the industries in which "classical" cointegration around a stable trend holds despite the individual series exhibiting trend breaks. In such cases we would say that the series are rather "co-breaking", with the breaks in trend slope appearing around the same time and being similar over the individual series, and hence not influencing the cointegrating relation. For combining the individual-unit statistics into a panel statistic which is robust to cross-sectional dependence, we employ the modified Inverse Normal method of Hartung. A small-scale simulation study revealed that the CAIN test proposed here is not suitable for combining the individual SL test statistics without structural breaks; the panel test is rather oversized.[7] However, Hartung's approach with $\kappa_1 = 0.2$ does control the type I error probability well in such setting. Hence, we use it to look for evidence of cointegration at the panel level without considering structural breaks. Table 6 reveals that the industries for which this holds at the usual 5% level are 3: Mineral fuels, 4: Animal

[7]These results are omitted for brevity, but are available upon request. They further indicate that the CAIN test could be employed, however, to combine a mixture of individual SL and TSL test statistics, if the number of SL statistics without breaks is low compared to that of TSL statistics, and also when N is relatively small.

and vegetable oils and 8: Manufactured goods. For the latter, however, we remark that the panel test rejects the null of no cointegration only because of the extremely small p-value for Germany; none of the p-values for the other countries are close to 0, hence we include this industry into the subsequent analysis with structural breaks as well. Further we note that the p-values of Hartung's panel tests for industries 2: Crude materials (inedible, except fuels), 5: Chemicals and related products and 6: Manufactured goods classified chiefly by materials indicate rejection of the null hypothesis of no cointegration only at the 10% level.

When testing for cointegration around a broken trend with the CAIN-TSL test, we first consider each of the two identified break dates separately. Introducing a break in May 2000 does not lead to rejecting the null hypothesis of no cointegration for any industry, except for Industry 4: Animal and vegetable oils, a result which was found also without considering any breaks.[8] Hence despite being selected by the univariate unit root tests, May 2000 turns out not to be a relevant break date in the cointegration analysis.

We next consider a level shift and a trend break in May 2002. The results of the individual TSL tests and the panel CAIN-TSL tests are presented in Tables 7 and 8, respectively. The high degree of cross-sectional dependence, resulting primarily from the similar exchange rate dynamics and the identical foreign prices series, is reflected in the high value of the estimated $\hat{\rho}_{\varepsilon}$ (see Table 8). We find that including this trend break has a significant impact on the results from the cointegration analysis for at least some industries. For example, in Industry 1: Beverages and tobacco we now find evidence of one stable long-run relation, as well as for Industry 8: Manufactured goods. For industry 5: Chemicals and related products including the break results in finding two cointegrating relations. On the contrary, we find that such a trend break in the presumed cointegrating relation is not relevant for industries 2: Crude materials and related products, 6: Manufactured goods classified chiefly by materials and 7: Machines, transport equipment. These findings remain robust when a break in any month in the period February – June 2002 is considered. Furthermore, the results are not fundamentally altered when considering two breaks – in May 2000 and May 2002, respectively, although the second cointegrating relationship for Industry 5: Chemicals and related products is not captured in this case. This might be due to the loss of power which naturally occurs if considering non-existent breaks.

We conclude that considering the possibility of a trend break plays an important role in establishing a cointegrating relationship between import prices, exchange rate and foreign prices for at least three different industries in the data set under investigation.

7 Conclusion

In this paper we propose a new meta-analytic approach (CAIN) to test for the cointegrating rank in panels where structural breaks and cross-sectional dependence are allowed for. It is an extension of the likelihood-based rank test of Trenkler et al. (2007) (TSL), and requires only the p-values of the individual LR trace statistics of the TSL test. The CAIN-TSL test is based on a modification of the popular inverse normal method for p-values combination, employing a novel estimator for the unknown correlation between the probits. We propose a way to estimate this correlation as a function of the system dimension, the cointegrating rank under the null hypothesis and the average absolute cross-sectional correlation between the residuals of the individual VAR models in first differences. The latter is easily estimable in practice and provides an easy-to-interpret measure of the degree of cross-sectional dependence. In a Monte Carlo study we demonstrate the superior

[8] For brevity we do not report these results, but they are available upon request.

properties of the CAIN-TSL test in comparison with other meta-analytic approaches recently proposed in the panel unit-root literature. An application of the test to investigate the exchange rate pass-through in a panel of Euro-area countries provides an illustration of its usefulness in practice.

Table 3: Monte Carlo study results, not near-unit root processes (case A) and near-unit root processes (case B).
Empirical size and power under $H_0 : r_i = 0, \forall i = 1, \ldots, N$.

	True cointegrating rank 0: Size						True cointegrating rank 1, case A: Power						True cointegrating rank 1, case B: Power					
	T=100			T=200			T=100			T=200			T=100			T=200		
	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$
Test	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$					
Choi	0.05	0.04	0.03	0.04	0.04	0.03	0.56	0.92	0.99	0.99	1.00	1.00	0.16	0.27	0.36	0.58	0.92	0.99
Hartung κ_1	0.04	0.02	0.01	0.04	0.01	0.01	0.43	0.68	0.79	0.96	1.00	1.00	0.13	0.15	0.17	0.51	0.80	0.93
Hartung κ_2	0.05	0.04	0.03	0.05	0.03	0.02	0.46	0.73	0.82	0.97	1.00	1.00	0.16	0.23	0.28	0.54	0.85	0.95
Simes	0.06	0.06	0.07	0.05	0.05	0.06	0.36	0.50	0.57	0.94	1.00	1.00	0.14	0.17	0.19	0.48	0.68	0.78
CAIN-TSL	0.05	0.04	0.03	0.04	0.04	0.03	0.56	0.92	0.99	0.99	1.00	1.00	0.16	0.27	0.35	0.58	0.92	0.99
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Choi	0.05	0.05	0.05	0.04	0.05	0.05	0.49	0.84	0.95	0.98	1.00	1.00	0.13	0.23	0.30	0.49	0.84	0.95
Hartung κ_1	0.04	0.02	0.01	0.04	0.02	0.01	0.35	0.56	0.66	0.93	1.00	1.00	0.10	0.12	0.12	0.42	0.65	0.79
Hartung κ_2	0.05	0.04	0.03	0.05	0.04	0.04	0.38	0.62	0.71	0.93	1.00	1.00	0.13	0.18	0.20	0.44	0.71	0.84
Simes	0.05	0.06	0.06	0.05	0.05	0.06	0.29	0.39	0.45	0.89	0.99	1.00	0.11	0.13	0.14	0.37	0.51	0.60
CAIN-TSL	0.04	0.04	0.04	0.04	0.04	0.04	0.48	0.83	0.93	0.98	1.00	1.00	0.13	0.21	0.26	0.49	0.82	0.94
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Choi	0.06	0.08	0.11	0.06	0.09	0.12	0.42	0.71	0.82	0.95	1.00	1.00	0.13	0.22	0.29	0.42	0.69	0.82
Hartung κ_1	0.04	0.03	0.04	0.05	0.04	0.04	0.28	0.42	0.47	0.85	0.97	0.99	0.10	0.11	0.12	0.32	0.47	0.57
Hartung κ_2	0.06	0.06	0.06	0.06	0.06	0.07	0.31	0.47	0.53	0.86	0.97	0.99	0.12	0.16	0.17	0.35	0.54	0.64
Simes	0.06	0.06	0.06	0.05	0.05	0.06	0.24	0.31	0.33	0.79	0.93	0.96	0.10	0.11	0.11	0.28	0.37	0.42
CAIN-TSL	0.04	0.05	0.05	0.05	0.05	0.06	0.38	0.61	0.70	0.94	1.00	1.00	0.12	0.15	0.17	0.38	0.59	0.70
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Choi	0.05	0.05	0.06	0.05	0.06	0.07	0.66	0.95	0.99	1.00	1.00	1.00	0.20	0.37	0.49	0.70	0.96	0.99
Hartung κ_1	0.04	0.02	0.02	0.04	0.02	0.02	0.52	0.77	0.87	0.98	1.00	1.00	0.17	0.23	0.27	0.62	0.90	0.97
Hartung κ_2	0.05	0.04	0.04	0.05	0.05	0.04	0.55	0.80	0.89	0.99	1.00	1.00	0.20	0.31	0.37	0.65	0.93	0.98
Simes	0.06	0.06	0.06	0.06	0.05	0.06	0.45	0.60	0.68	0.97	1.00	1.00	0.17	0.22	0.26	0.58	0.81	0.89
CAIN-TSL	0.04	0.03	0.03	0.04	0.04	0.04	0.63	0.92	0.98	1.00	1.00	1.00	0.18	0.29	0.35	0.67	0.93	0.99

Notes: Rejection frequencies at 5% significance level, 5000 replications. Power is not size-adjusted, as size-adjustment would not be available in practice.

Table 4: Monte Carlo study results, case A: not near-unit root processes.
Empirical size and power under $H_0 : r_i = 1, \forall i = 1, \ldots, N$.

	True cointegrating rank 1, hypothesized rank 1: Size						True cointegrating rank 2, hypothesized rank 1: Power					
	T=100			T=200			T=100			T=200		
	$N = 5$	$N = 15$	$N = 25$	$N = 5$	$N = 15$	$N = 25$	$N = 5$	$N = 15$	$N = 25$	$N = 5$	$N = 15$	$N = 25$
Test	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$					
Choi	0.001	0.000	0.000	0.003	0.000	0.000	0.41	0.78	0.92	0.99	1.00	1.00
Hartung κ_1	0.001	0.000	0.000	0.005	0.000	0.000	0.28	0.46	0.56	0.97	1.00	1.00
Hartung κ_2	0.002	0.000	0.000	0.007	0.001	0.000	0.32	0.54	0.63	0.98	1.00	1.00
Simes	0.010	0.006	0.009	0.016	0.017	0.016	0.23	0.27	0.27	0.95	1.00	1.00
CAIN-TSL	0.001	0.000	0.000	0.003	0.000	0.000	0.41	0.77	0.91	0.99	1.00	1.00
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Choi	0.002	0.000	0.000	0.006	0.001	0.000	0.40	0.75	0.88	0.99	1.00	1.00
Hartung κ_1	0.004	0.000	0.000	0.006	0.001	0.000	0.26	0.42	0.51	0.97	1.00	1.00
Hartung κ_2	0.005	0.001	0.000	0.008	0.001	0.000	0.28	0.49	0.58	0.97	1.00	1.00
Simes	0.009	0.007	0.004	0.020	0.019	0.024	0.20	0.24	0.25	0.95	1.00	1.00
CAIN-TSL	0.002	0.000	0.000	0.005	0.001	0.000	0.39	0.72	0.85	0.99	1.00	1.00
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Choi	0.004	0.003	0.002	0.015	0.017	0.017	0.40	0.67	0.79	0.99	1.00	1.00
Hartung κ_1	0.003	0.001	0.001	0.014	0.007	0.005	0.24	0.34	0.41	0.95	1.00	1.00
Hartung κ_2	0.004	0.002	0.001	0.018	0.012	0.011	0.27	0.40	0.47	0.96	1.00	1.00
Simes	0.008	0.007	0.005	0.024	0.018	0.019	0.18	0.21	0.21	0.92	0.99	1.00
CAIN-TSL	0.002	0.001	0.000	0.010	0.008	0.005	0.34	0.55	0.63	0.98	1.00	1.00
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Choi	0.003	0.002	0.002	0.009	0.007	0.006	0.43	0.78	0.90	0.99	1.00	1.00
Hartung κ_1	0.003	0.001	0.001	0.008	0.004	0.003	0.30	0.49	0.60	0.98	1.00	1.00
Hartung κ_2	0.005	0.002	0.002	0.011	0.007	0.005	0.33	0.56	0.67	0.98	1.00	1.00
Simes	0.011	0.010	0.007	0.019	0.014	0.018	0.24	0.29	0.31	0.96	1.00	1.00
CAIN-TSL	0.002	0.001	0.000	0.006	0.002	0.002	0.40	0.69	0.81	0.99	1.00	1.00

Notes: Rejection frequencies at 5% significance level, 5000 replications. Power is not size-adjusted, as size-adjustment would not be available in practice.

Table 5: Monte Carlo study results, case B: near-unit root processes.
Empirical size and power under $H_0 : r_i = 1, \forall i = 1, \ldots, N$.

	True cointegrating rank 1, hypothesized rank 1: Size						True cointegrating rank 2, hypothesized rank 1: Power					
	T=100			T=200			T=100			T=200		
	$N = 5$	$N = 15$	$N = 25$	$N = 5$	$N = 15$	$N = 25$	$N = 5$	$N = 15$	$N = 25$	$N = 5$	$N = 15$	$N = 25$
Test	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$					
Choi	0.000	0.000	0.000	0.001	0.000	0.000	0.02	0.01	0.01	0.37	0.66	0.83
Hartung κ_1	0.000	0.000	0.000	0.002	0.000	0.000	0.03	0.01	0.00	0.34	0.55	0.70
Hartung κ_2	0.000	0.000	0.000	0.003	0.000	0.000	0.03	0.02	0.01	0.38	0.65	0.81
Simes	0.005	0.003	0.002	0.008	0.005	0.006	0.04	0.03	0.04	0.33	0.46	0.53
CAIN-TSL	0.000	0.000	0.000	0.001	0.000	0.000	0.02	0.01	0.01	0.37	0.66	0.83
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Choi	0.000	0.000	0.000	0.002	0.000	0.000	0.03	0.02	0.01	0.38	0.67	0.82
Hartung κ_1	0.001	0.000	0.000	0.001	0.000	0.000	0.02	0.01	0.00	0.32	0.53	0.66
Hartung κ_2	0.001	0.000	0.000	0.002	0.000	0.000	0.03	0.02	0.01	0.36	0.63	0.76
Simes	0.004	0.002	0.001	0.008	0.005	0.006	0.04	0.03	0.03	0.30	0.41	0.47
CAIN-TSL	0.000	0.000	0.000	0.002	0.000	0.000	0.03	0.01	0.01	0.37	0.64	0.79
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Choi	0.000	0.000	0.000	0.003	0.001	0.000	0.04	0.05	0.05	0.39	0.64	0.76
Hartung κ_1	0.000	0.000	0.000	0.002	0.000	0.000	0.02	0.02	0.02	0.31	0.47	0.58
Hartung κ_2	0.001	0.000	0.000	0.003	0.001	0.001	0.03	0.03	0.04	0.34	0.55	0.65
Simes	0.003	0.002	0.001	0.008	0.004	0.005	0.03	0.03	0.03	0.27	0.35	0.40
CAIN-TSL	0.000	0.000	0.000	0.001	0.000	0.000	0.03	0.02	0.02	0.34	0.52	0.61
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Choi	0.000	0.000	0.000	0.002	0.000	0.000	0.03	0.02	0.02	0.39	0.66	0.80
Hartung κ_1	0.000	0.000	0.000	0.001	0.000	0.000	0.03	0.01	0.01	0.37	0.57	0.71
Hartung κ_2	0.001	0.000	0.000	0.002	0.000	0.000	0.04	0.02	0.01	0.41	0.67	0.81
Simes	0.004	0.003	0.002	0.008	0.006	0.006	0.05	0.05	0.04	0.37	0.52	0.62
CAIN-TSL	0.000	0.000	0.000	0.001	0.000	0.000	0.02	0.01	0.01	0.35	0.57	0.68

Notes: Rejection frequencies at 5% significance level, 5000 replications. Power is not size-adjusted, as size-adjustment would not be available in practice.

Table 6: Individual SL cointegrating rank tests without structural breaks and industry-specific panel tests employing Hartung's combination of p-values

	0: Food and live animals chiefly for food					1: Beverages and tobacco					2: Crude materials (inedible, except fuels)				
		$H_0: r=0$		$H_0: r=1$			$H_0: r=0$		$H_0: r=1$			$H_0: r=0$		$H_0: r=1$	
Country	lag $\hat{s}$	SL LR	p-value	SL LR	p-value	lag $\hat{s}$	SL LR	p-value	SL LR	p-value	lag $\hat{s}$	SL LR	p-value	SL LR	p-value
France	2	19.55	0.42	7.44	0.61	4	19.08	0.46	8.53	0.49	2	19.74	0.41	4.91	0.88
Netherlands	2	13.83	0.84	6.67	0.70	4	19.40	0.43	7.34	0.62	3	20.83	0.34	8.23	0.52
Germany	3	16.78	0.63	5.69	0.81	4	26.65	0.09	6.43	0.73	4	24.35	0.16	8.68	0.47
Italy	3	21.53	0.29	9.69	0.37	4	21.07	0.32	9.63	0.37	4	26.36	0.09	4.00	0.94
Ireland	4	17.62	0.57	7.74	0.58	4	22.70	0.23	15.00	0.07	2	31.43	0.02	7.96	0.55
Greece	3	24.49	0.15	6.54	0.72	4	19.78	0.41	8.77	0.46	2	29.87	0.03	12.16	0.18
Spain	2	21.77	0.28	7.79	0.57	4	16.80	0.63	8.07	0.54	4	22.97	0.22	7.65	0.59
Hartung κ_1		−0.15	0.44	0.35	0.64		−0.45	0.33	−0.17	0.43		−1.36	0.09	0.44	0.67
Hartung κ_2		−0.16	0.44	0.35	0.64		−0.45	0.33	−0.17	0.43		−1.39	0.08	0.46	0.68
	3: Mineral fuelds, lubricants and related materials					4: Animal and vegetable oils, fats and waxes					5: Chemicals and related products				
		$H_0: r=0$		$H_0: r=1$			$H_0: r=0$		$H_0: r=1$			$H_0: r=0$		$H_0: r=1$	
Country	lag $\hat{s}$	SL LR	p-value	SL LR	p-value	lag $\hat{s}$	SL LR	p-value	SL LR	p-value	lag $\hat{s}$	SL LR	p-value	SL LR	p-value
France	3	21.52	0.29	3.39	0.97	4	22.58	0.23	6.56	0.71	3	30.18	0.03	6.04	0.77
Netherlands	2	19.89	0.40	4.88	0.88	3	32.92	0.01	4.25	0.93	3	23.23	0.20	6.44	0.73
Germany	4	29.48	0.04	5.73	0.80	4	29.19	0.04	7.27	0.63	3	28.71	0.05	5.38	0.84
Italy	2	42.03	0.00	5.39	0.84	4	17.09	0.61	7.69	0.58	4	26.08	0.10	7.18	0.64
Ireland	3	23.82	0.18	4.19	0.93	4	18.68	0.49	9.79	0.36	3	18.69	0.49	6.05	0.77
Greece	4	27.00	0.08	5.53	0.82	4	38.25	0.00	8.27	0.52	3	30.97	0.02	10.02	0.34
Spain	3	23.04	0.21	4.86	0.88	4	20.43	0.36	5.72	0.80	3	22.38	0.24	4.62	0.90
Hartung κ_1		−3.15	0.00	1.27	0.90		−3.48	0.00	0.51	0.70		−1.53	0.06	0.70	0.76
Hartung κ_2		−3.53	0.00	1.28	0.90		−4.27	0.00	0.52	0.70		−1.57	0.06	0.71	0.76
	6: Manufactured goods classified chiefly by materials					7: Machines, transport equipment					8: Manufactured goods				
		$H_0: r=0$		$H_0: r=1$			$H_0: r=0$		$H_0: r=1$			$H_0: r=0$		$H_0: r=1$	
Country	lag $\hat{s}$	SL LR	p-value	SL LR	p-value	lag $\hat{s}$	SL LR	p-value	SL LR	p-value	lag $\hat{s}$	SL LR	p-value	SL LR	p-value
France	4	17.92	0.55	7.08	0.65	3	19.53	0.42	7.05	0.66	4	18.17	0.53	4.98	0.87
Netherlands	4	25.17	0.13	6.52	0.72	3	16.98	0.62	4.56	0.91	4	23.54	0.19	12.00	0.19
Germany	3	23.65	0.18	7.30	0.63	4	17.78	0.56	5.91	0.78	2	47.87	0.00	19.30	0.01
Italy	3	31.34	0.02	7.26	0.63	3	19.44	0.43	5.43	0.83	4	13.77	0.84	4.73	0.89
Ireland	3	26.43	0.09	4.58	0.90	4	20.26	0.37	5.44	0.83	4	23.26	0.20	6.17	0.76
Greece	3	25.38	0.12	13.78	0.10	2	22.72	0.23	4.68	0.90	4	21.66	0.29	11.25	0.24
Spain	4	29.49	0.04	3.71	0.96	3	16.40	0.66	5.82	0.79	4	19.25	0.44	9.35	0.40
Hartung κ_1		−1.48	0.07	0.85	0.80		−0.09	0.47	0.96	0.83		−2.41	0.01	−0.47	0.32
Hartung κ_2		−1.51	0.07	0.90	0.82		−0.09	0.47	0.97	0.83		−2.95	0.00	−0.58	0.28

Notes: Hartung's approach to combination of p-values is used to combine the individual SL LR trace statistics into a panel test statistic with a $N(0,1)$ distribution. The lag order $\hat{s}$ has been selected by the MAIC criterion of Qu and Perron (2007) under the null hypothesis of no cointegration.

Table 7: Industry-specific CAIN-TSL panel cointegrating rank tests with a level shift and trend break in May 2002

	0: Food and live animals chiefly for food					1: Beverages and tobacco					2: Crude materials (inedible, except fuels)				
		$H_0: r = 0$		$H_0: r = 1$			$H_0: r = 0$		$H_0: r = 1$			$H_0: r = 0$		$H_0: r = 1$	
Country	lag $\hat{s}$	TSL LR	p-value	TSL LR	p-value	lag $\hat{s}$	TSL LR	p-value	TSL LR	p-value	lag $\hat{s}$	TSL LR	p-value	TSL LR	p-value
France	3	26.22	0.23	8.81	0.69	4	26.59	0.21	11.45	0.43	3	17.61	0.78	8.87	0.69
Netherlands	3	23.76	0.36	5.78	0.93	4	30.27	0.09	17.18	0.09	3	21.83	0.49	11.60	0.42
Germany	3	26.35	0.22	5.04	0.96	4	41.80	0.00	15.16	0.16	4	18.89	0.69	13.77	0.24
Italy	4	21.92	0.48	9.39	0.63	4	28.14	0.15	11.81	0.40	4	27.02	0.19	11.35	0.44
Ireland	4	34.83	0.02	9.53	0.62	4	30.09	0.09	20.09	0.03	4	32.47	0.05	12.21	0.36
Greece	3	21.37	0.52	7.45	0.82	4	24.02	0.34	11.76	0.40	3	26.28	0.22	21.84	0.02
Spain	4	17.93	0.76	6.88	0.86	4	27.27	0.18	6.16	0.91	3	22.28	0.46	14.91	0.18
	6: Manufactured goods classified chiefly by materials					Industry 7: Machines, transport equipment					8: Manufactured goods				
		$H_0: r = 0$		$H_0: r = 1$			$H_0: r = 0$		$H_0: r = 1$			$H_0: r = 0$		$H_0: r = 1$	
Country	lag $\hat{s}$	TSL LR	p-value	TSL LR	p-value	lag $\hat{s}$	TSL LR	p-value	TSL LR	p-value	lag $\hat{s}$	TSL LR	p-value	TSL LR	p-value
France	4	16.44	0.84	8.91	0.68	4	22.21	0.46	4.51	0.98	4	27.51	0.17	8.41	0.73
Netherlands	4	25.70	0.25	7.83	0.78	3	22.00	0.48	11.75	0.40	4	38.84	0.01	8.09	0.76
Germany	3	29.25	0.11	6.49	0.89	4	17.47	0.78	7.75	0.79	3	52.65	0.00	13.05	0.29
Italy	4	30.67	0.08	8.99	0.67	4	22.17	0.46	10.50	0.52	4	30.41	0.09	9.91	0.58
Ireland	4	15.28	0.89	4.90	0.97	4	22.82	0.42	9.99	0.57	4	23.00	0.41	11.65	0.41
Greece	3	26.65	0.21	8.50	0.72	3	30.40	0.09	10.82	0.49	4	49.57	0.00	16.60	0.11
Spain	4	15.70	0.88	8.98	0.68	4	12.43	0.97	3.89	0.99	4	31.56	0.06	10.43	0.53
	5: Chemicals and related products														
		$H_0: r = 0$		$H_0: r = 1$		$H_0: r = 2$									
Country	lag $\hat{s}$	TSL LR	p-value	TSL LR	p-value	TSL LR	p-value								
France	3	35.01	0.02	13.71	0.25	3.81	0.43								
Netherlands	3	32.85	0.04	11.75	0.40	5.51	0.22								
Germany	3	36.45	0.02	20.36	0.03	2.44	0.67								
Italy	4	39.37	0.01	18.31	0.06	1.25	0.89								
Ireland	4	32.83	0.04	19.62	0.04	1.81	0.79								
Greece	3	34.56	0.03	16.71	0.10	3.90	0.41								
Spain	4	28.23	0.15	9.74	0.60	2.11	0.73								

Notes: The lag order $\hat{s}$ has been selected by the MAIC criterion of Qu and Perron (2007) under the null hypothesis of no cointegration.

Table 8: Individual TSL cointegrating rank tests with a level shift and trend break in May 2002

Industry	$H_0 : r_i \leq r, \forall i$	CAIN-TSL				Hartung κ_1		Hartung κ_2	
		$\hat{\rho}_\varepsilon$	$\hat{\bar{\rho}}_t$	t-stat	p-value	t-stat	p-value	t-stat	p-value
0: Food and live animals cheifly for food	$r = 0$	0.69	0.16	−0.85	0.20	−0.66	0.25	−0.69	0.25
1: Beverages and tobacco	$r = 0$	0.66	0.14	−2.40	0.01	−1.66	0.05	−1.71	0.04
	$r = 1$	0.66	0.15	−0.98	0.16	−1.20	0.11	−1.34	0.09
2: Crude materials (inedible, except fuels)	$r = 0$	0.70	0.17	−0.57	0.28	−0.46	0.32	−0.48	0.32
5: Chemicals and related products	$r = 0$	0.63	0.13	−3.73	0.00	−2.04	0.02	−2.05	0.02
	$r = 1$	0.63	0.13	−2.03	0.02	−1.48	0.07	−1.53	0.06
	$r = 2$	0.63	0.14	0.52	0.70	0.34	0.63	0.34	0.63
6: Manufactured goods classified chiefly by materials	$r = 0$	0.71	0.18	−0.18	0.43	−0.31	0.38	−0.38	0.35
7: Machines, transport equipment	$r = 0$	0.63	0.12	0.26	0.60	0.28	0.61	0.31	0.62
8: Manufactured goods	$r = 0$	0.70	0.17	−3.75	0.00	−6.38	0.00	−7.82	0.00
	$r = 1$	0.70	0.17	−0.11	0.46	−0.07	0.47	−0.08	0.47

Notes: $H_0 : r_i \leq 1$ is tested only if $H_0 : r_i = 0$ is rejected; similarly for higher ranks under the null hypothesis.

Appendix

Figure 1: Cross-sectional correlation of the probits against cross-sectional correlation of the innovations, bi-variate system with true cointegrating rank zero, testing for $H_0 : r_i = 0, \forall i$

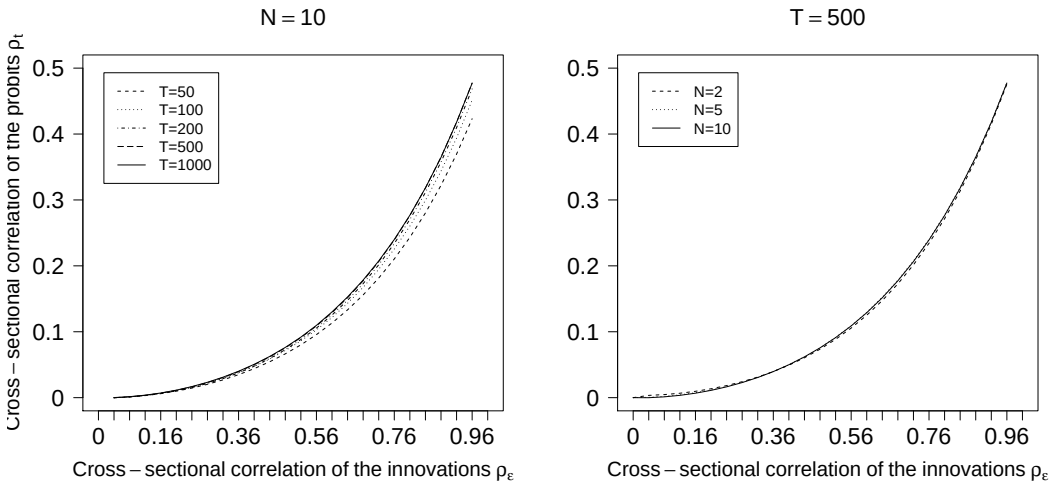

Table 9: True and estimated ρ_ε under $H_0 : r = 0$

	Estimated $\hat{\rho}_\varepsilon$		
True ρ_ε	True rank 0	True rank 1	True rank 2
0.2	0.200	0.196	0.192
0.4	0.400	0.392	0.384
0.6	0.600	0.589	0.577
0.8	0.800	0.786	0.771

Notes: Results from large-scale simulation study with $m = 3$, $T = 500$ and $N = 5$.

References

Arsova A, Örsal DDK. 2016a. An intersection test for the cointegrating rank in dependent panel data. Working Paper 357, Leuphana University Lüneburg.

Arsova A, Örsal DDK. 2016b. Likelihood-Based Panel Cointegration Test in the Presence of a Linear Time Trend and Cross-Sectional Dependence. *Econometric Reviews* ISSN 0747-4938. URL `http://dx.doi.org/10.1080/07474938.2016.1183070`

Banerjee A, Carrion-i Silvestre JL. 2015. Cointegration in Panel Data with Structural Breaks and Cross-Section Dependence. *Journal of Applied Econometrics* **30**: 1–23. ISSN 08837252.

Carrion-i Silvestre JL, Kim D, Perron P. 2009. GLS-Based Unit Root Tests With Multiple Structural Breaks Under Both the Null and the Alternative Hypotheses. *Econometric Theory* **25**: 1754. ISSN 0266-4666.

Choi I. 2001. Unit root tests for panel data. *Journal of International Money and Finance* **20**: 249–272. ISSN 02615606.

Costantini M, Lupi C. 2013. A Simple Panel-CADF Test for Unit Roots. *Oxford Bulletin of Economics and Statistics* **75**: 276–296. ISSN 03059049.

Demetrescu M, Hanck C. 2012. Unit Root Testing in Heteroscedastic Panels Using the Cauchy Estimator. *Journal of Business & Economic Statistics* **30**: 256–264. ISSN 0735-0015.

Demetrescu M, Hassler U, Tarcolea AI. 2006. Combining Significance of Correlated Statistics with Application to Panel Data. *Oxford Bulletin of Economics and Statistics* **68**: 647–663. ISSN 03059049.

Hanck C. 2013. An intersection test for panel unit roots. *Econometric Reviews* **32**: 183–203. ISSN 0747-4938.

Hartung J. 1999. A Note on Combining Dependent Tests of Significance. *Biometrical Journal* **41**: 849–855.

Im KS, Lee J, Tieslau M. 2010. Panel LM Unit Root Tests with Trend Shifts. Working Paper 2010-1, Federal Deposit Insurance Corporation, Center for Financial Research.

Johansen S, Mosconi R, Nielsen B. 2000. Cointegration analysis in the presence of structural breaks in the deterministic trend. *Econometrics Journal* **3**: 216–249. ISSN 13684221.

Kejriwal M, Perron P. 2010. A sequential procedure to determine the number of breaks in trend with an integrated or stationary noise component. *Journal of Time Series Analysis* **31**: 305–328. ISSN 01439782.

Lee J, Strazicich MC. 2003. Minimum Lagrange multiplier unit root test with two structural breaks. *Review of Economics and Statistics* **85**: 1082–1089. ISSN 0034-6535.

Lee J, Strazicich MC, Meng M. 2012. Two-Step LM Unit Root Tests with Trend-Breaks. *Journal of Statistical and Econometric Methods* **1**: 81–107.

Lütkepohl H, Saikkonen P, Trenkler C. 2004. Testing for the Cointegrating Rank of a VAR Process with Level Shift at Unknown Time. *Econometrica* **72**: 647–662.

Maddala GS, Wu S. 1999. A comparative study of panel unit root tests. *Oxford Bulletin of Economics and Statistics* **Special Issue**: 631–652. ISSN 03059049.

Narayan PK, Popp S. 2010. A new unit root test with two structural breaks in level and slope at unknown time. *Journal of Applied Statistics* **37**: 1425–1438. ISSN 0266-4763.

Payne JE, Anderson S, Lee J, Cho MH. 2015. Do per capita health care expenditures converge among OECD countries? Evidence from unit root tests with level and trend-shifts. *Applied Economics* **47**: 5600–5613. ISSN 0003-6846.

Pesaran MH. 2015. Testing Weak Cross-Sectional Dependence in Large Panels. *Econometric Reviews* **34**: 1089–1117. ISSN 0747-4938.

Qu Z, Perron P. 2007. A modified information criterion for cointegration tests based on a VAR approximation. *Econometric Theory* **23**: 638–685.

Saikkonen P, Lutkepohl H. 2000. Trend Adjustment Prior to Testing for the Cointegrating Rank of a Vector Autoregressive Process. *Journal of Time Series Analysis* **21**: 435–456. ISSN 0143-9782.

Saikkonen P, Lütkepohl H, Trenkler C. 2006. Break date estimated for VAR processes with level shift with an application to cointegration testing. *Econometric Theory* **22**: 15–68.

Sarkar SK. 1998. Some Probability Inequalities for Ordered MTP2 Random Variables: A Proof of the Simes Conjecture. *The Annals of Statistics* **26**: 494–504.

Simes RJ. 1986. An improved Bonferroni procedure for multiple tests of significance. *Biometrika* **73**: 751–754.

Sobreira N, Nunes LC. 2015. Tests for Multiple Breaks in the Trend with Stationary or Integrated Shocks. *Oxford Bulletin of Economics and Statistics* : n/aISSN 03059049.

Stouffer S, DeVinney L, Suchmen E. 1949. *The American Soldier: Adjustment During Army Life.* Princeton, NJ: Princeton University Press.

Toda HY. 1995. Finite sample performance of likelihood ratio tests for cointegrating ranks in vector autoregressions. *Econometric Theory* **11**: 1015–1032.

Trenkler C, Saikkonen P, Lütkepohl H. 2007. Testing for the Cointegrating Rank of a VAR Process with Level Shift and Trend Break. *Journal of Time Series Analysis* **29**: 331–358. ISSN 0143-9782.

Wagner M, Hlouskova J. 2010. The Performance of Panel Cointegration Methods: Results from a Large Scale Simulation Study. *Econometric Reviews* **29**: 182–223. ISSN 0747-4938.

Westerlund J, Edgerton DL. 2008. A Simple Test for Cointegration in Dependent Panels with Structural Breaks. *Oxford Bulletin of Economics and Statistics* **70**: 665–704. ISSN 03059049.