

**A Residual-Based ADF Test for  
Stationary Cointegration in  $I(2)$  Settings**

**Javier Gomez-Biscarri  
Javier Hualde**

**September 2014**

*Barcelona GSE Working Paper Series*

*Working Paper n° 779*

# A residual-based ADF test for stationary cointegration in $I(2)$ settings\*

Javier Gomez-Biscarri<sup>†</sup>

Javier Hualde<sup>‡</sup>

Universitat Pompeu Fabra  
and Barcelona GSE

Universidad Pública de Navarra

September 5, 2014

## Abstract

We propose a residual-based augmented Dickey-Fuller (ADF) test statistic that allows for detection of stationary cointegration within a system that may contain both  $I(2)$  and  $I(1)$  observables. The test is also consistent under the alternative of multicointegration, where first differences of the  $I(2)$  observables enter the cointegrating relationships. We find the null limiting distribution of this statistic and justify why our proposal improves over related approaches. Critical values are computed for a variety of situations. Additionally, building on this ADF test statistic, we propose a procedure to test the null of no stationary cointegration which overcomes the drawback, suffered by any residual-based method, of the lack of power with respect to some relevant alternatives. Finally, a Monte Carlo experiment is carried out and an empirical application is provided as an illustrative example.

*JEL Classification:* C12, C22, C32.

*Keywords:*  $I(2)$  systems; stationary cointegration; multicointegration; residual-based tests.

## 1 Introduction

The concept of cointegration has received much attention in the last two decades. Its importance stems from the fact that cointegration provides the link between the

---

\*We are grateful for the comments of three referees which have led to improvements in the paper.

<sup>†</sup>Javier Gomez-Biscarri acknowledges financial support from Universitat Pompeu Fabra, from the Barcelona GSE and from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Program for Centers of Excellence in R&D (SEV-2011-0075) and project ECO2011-25607.

<sup>‡</sup>Corresponding author: Javier Hualde, Universidad Pública de Navarra, Campus Arrosadía, 31006 Pamplona, Spain. Phone: +34948169674. Fax: +34948169721. email: javier.hualde@unavarra.es. Javier Hualde's research is supported by the Spanish Ministry of Economy and Competitiveness through project ECO2011-24304.

economic concept of (long-run) equilibrium and the statistical notions of nonstationarity and trending behavior: nonstationary variables may display relationships that are representative of long-run equilibria, in that deviations from the equilibrium are short-lasting. Since the seminal contributions of Engle and Granger (1987) and Johansen (1988), cointegration has been quite well studied both in uni-equation and system  $I(1)$  settings, where the observables may behave like  $I(1)$  ( $I(0)$  after differencing, where  $I(0)$  refers to covariance stationary after demeaning with nonzero and bounded spectral density) or stationary variables. However, many observables (especially nominal variables such as price indexes) are smoother than what  $I(1)$  behavior would suggest. For example, inflation rates have a behavior close to that of an  $I(1)$  variable which implies that (log)price indexes might be characterized as  $I(2)$ , for which twice differencing is necessary to achieve  $I(0)$ . Thus, structural models that involve aggregate prices could be combining variables with different integration orders (see Juselius, 1995, or Banerjee *et al.*, 2001, for two different illustrations of such settings). A similar rationale applies to nominal GDP and nominal wealth, which is the result of the time-accumulation of nominal income. Further evidence supporting the presence of  $I(2)$  trends in data sets can be found in Kongsted (2003, 2005), Kongsted and Nielsen (2004), Bacchiocchi and Fanelli (2005) and Johansen *et al.* (2010).

The analysis of cointegration involving  $I(2)$  variables is also relevant with respect to the issue of multicointegration (Granger and Lee, 1989, 1990, Lee, 1992), which is a particular case of the so-called polynomial cointegration. In this situation, two  $I(1)$  observables (flow variables) cointegrate and the cumulated cointegrating error ( $I(1)$  stock variable) cointegrates with at least one of the observables. Engsted and Johansen (1999) justified that such phenomenon should be modelled as an  $I(2)$  system, where multicointegration arises from cointegration between  $I(2)$  variables in levels and first differences (see also Engsted and Haldrup, 1999). Granger and Lee (1989) applied the concept of multicointegration to the relationship between production and sales (flow variables) in a given industry, exploring also the possibility of cointegration between the stock of inventories (accumulated change of inventory) and sales. This setting was also studied by Banerjee and Mizen (2006). Other works explore the existence of multicointegration between housing starts, completions (flow variables) and housing units under construction (stock) (Lee, 1992), government spending, revenues and debt (Leachman, 1996, Leachman and Francis, 2002, Leachman *et al.*, 2005), imports, exports and external debt (Leachman and Francis, 2000, 2002), or real per capita private consumption expenditure, real per capita disposable income and stock of consumer's wealth (Siliverstovs, 2006).

As in the  $I(1)$  case, two different approaches have been developed to examine

cointegration in  $I(2)$  systems. First, Johansen (1995a), Paruolo (1996), Rahbek *et al.* (1999), Nielsen and Rahbek (2007), among others, proposed cointegration tests within a vector autoregressive framework, which includes also the possibility of detecting multicointegration (see also, Gregoir and Laroque, 1994, Juselius, 1995, Engsted and Johansen, 1999). Alternatively, regression-based procedures have been also proposed. This methodology extends the Phillips and Ouliaris (1990) residual-based tests for cointegration to the  $I(2)$  setting; it has been pursued in uni-equation settings by Haldrup (1994) and, in the particular case of multicointegration, by Engsted *et al.* (1997).

Specifically, Haldrup (1994) developed a residual-based augmented Dickey-Fuller (ADF) test for the null of  $I(1)$  versus the alternative of stationary cointegration among a set of  $I(1)$  and  $I(2)$  observables. In Haldrup's model, the  $I(2)$  observables cointegrate (with rank one) to an  $I(1)$  cointegrating error, which under the null does not further cointegrate with the  $I(1)$  observables. The test is carried out by regressing an  $I(2)$  observable on the  $I(1)$  observables, the remaining  $I(2)$  series (which are assumed to be non-cointegrated) and deterministic terms. In view of the results of Haldrup's (1994) Theorem 4, the null limiting distribution of the test essentially depends on the number of  $I(1)$  and  $I(2)$  regressors. We find three empirically relevant limitations to this test. First, and more importantly, the result appears to be valid only if the coherence at frequency zero between the  $I(0)$  error input processes generating the  $I(1)$  and  $I(2)$  components of the system, respectively, is zero. This is a very stringent requirement, which is not in general satisfied if, e.g., this  $I(0)$  error is a vector autoregressive and moving average process. Therefore, in general, the null limiting distribution of Haldrup's ADF test statistic is not free of nuisance parameters. Second, the test assumes that the  $I(2)$  variables cointegrate with rank exactly equal to one, which in systems with several  $I(2)$  observables might not be the case. Finally, in Haldrup's setting multicointegration is not allowed (although in the discussion of his Lemma 2, he acknowledges that if polynomial multicointegration is considered, a slight modification of his theory is needed). However, Engsted *et al.* (1997) applied Haldrup's results to the multicointegration case in a simple bivariate setting, and suggested that Haldrup's (1994) critical values might be used. We, however, believe that this is not the case, given that, when multicointegration is present, one  $I(2)$  observable appears as regressor both in levels and in first differences, a circumstance which must affect the limiting distribution of the test statistic, and which is not captured by Haldrup's (1994) framework. Other works which, either explicitly applied Haldrup's test, or refer to results derived from it, include Haldrup (1998), Engsted and Haldrup (1999), Leachman and Francis (2000), Haldrup and Lildhold

(2002), Leachman *et al.* (2005), Kia (2008) and Kugler and Kaufmann (2008).

Our aim in this paper is to propose a regression-based procedure which could be generally applicable to detect stationary cointegration in  $I(2)$  settings. Specifically, this procedure is based on an appropriately modified version of Haldrup's (1994) residual ADF test statistic. Nicely, we find that allowing for nonzero coherence at frequency zero requires implementing a correction which is intimately related to the issue of multicointegration. In fact, the correction that leads to an ADF test statistic with a pivotal null limiting distribution makes the test consistent to the alternative of multicointegration, covering therefore the case of Engsted *et al.* (1997).

Additionally, building on this ADF test statistic, we propose a procedure to test the null of no stationary cointegration which overcomes the drawback, suffered by any residual-based method, of the lack of power with respect to some relevant alternatives. In any residual-based cointegration testing method (including those proposed for the  $I(1)$  setting) the choice of the left hand side variable is a critical issue. In particular, the tests proposed by Phillips and Ouliaris (1990), Haldrup (1994) or Engsted *et al.* (1997), do not have power if the chosen left hand side variable does not enter the stationary relation with nonzero coefficient. Our proposal sheds some light on how a consistent test for any stationary alternative can be constructed.

Given that likelihood-based procedures for analysis of  $I(2)$  systems have been developed (Johansen, 1995a, Paruolo, 1996, Rahbek, Kongsted and Jørgensen, 1999, Nielsen and Rahbek, 2007), it is warranted that we motivate the usefulness of our proposal. We find three main justifications. First, many economic models lead to equilibrium equations which might contain both  $I(1)$  or  $I(2)$  variables. In particular, some models deliver one single equilibrium condition or several, but one of them is of special interest to the researcher. Examples of these are the analyses of money demand equations (which involve possible  $I(2)$  variables such as nominal money and price indices, and variables with  $I(1)$  behavior, such as interest rates or real output: see, e.g., Stock and Watson, 1993, Haldrup, 1994, Bae and DeJong, 2007), structural models of the exchange rate (which lead to an expression of the exchange rate as a function of the differentials between domestic and foreign variables: some of these "exchange rate fundamentals" might be  $I(2)$ , such as money or prices, and some are possibly  $I(1)$ , such as real output or interest rates; see, e.g., Mark and Sul, 2001, Rapach and Wohar, 2002, Rossi, 2006) or purchasing power parity (PPP) models of the exchange rate (which postulate a relationship between domestic and foreign price indices, possibly  $I(2)$ , and the exchange rate, typically  $I(1)$ ; see, e.g., Rogoff, 1996, Caner and Kilian, 2001, Pedroni, 2004, Bacchiocchi and Fanelli, 2005). The researcher may be interested in testing these equilibrium relationships, without nec-

essarily attempting to give a full description of the cointegrating structure of the complete system.

Second, as pinpointed by Gomez-Biscarri and Hualde (2014) (Remark 6), even for  $I(1)$  systems, Johansen's methodology to infer the cointegrating rank is subject to sequentiality issues, which could arise in a more exacerbated way in  $I(2)$  systems. Note that this sequentiality is a very relevant issue in the determination of the cointegration indexes (see, e.g., Nielsen and Rahbek, 2007). Using regression-based techniques, Gomez-Biscarri and Hualde (2014) proposed an alternative way to deal with this sequentiality problem which appears to be a fair competitor to Johansen's methodology in  $I(1)$  settings. It is beyond the scope of the present paper to present a complete procedure to determine cointegration indexes in  $I(2)$  systems, but this can be achieved by extending the method in Gomez-Biscarri and Hualde (2014) appropriately. This extension would require the use of the test statistic proposed in the present paper.

Finally, as demonstrated by Gonzalo and Lee (1998) for the  $I(1)$  cointegrated setting, regression-based methods appear to be more robust than system methods under various circumstances. In an  $I(1)$  scenario, Gomez-Biscarri and Hualde (2014) showed situations where system approaches might show poor behaviour in small samples. In particular, regression-based methods to determine cointegration might be more parsimonious than full system maximum likelihood approaches, where estimation of a very large number of parameters might be required. We provide further evidence here along these lines.

The outline of the rest of the paper is as follows. In Section 2 we present an  $I(2)$  model and a residual-based ADF test statistic, and we derive its null asymptotic distribution. Section 3 comments on two important issues regarding the empirical implementation of this test statistic, namely, power properties and feasibility. Section 4 presents the results of a Monte Carlo experiment where we compare the performance of our procedure with that of a system method. An illustrative empirical example is discussed in Section 5, and we conclude in Section 6. Proofs are provided in the Appendix.

## 2 The ADF test: model, assumptions and properties

Before presenting our  $I(2)$  model we introduce some terminology. Formally, we say that a scalar or vector process  $\zeta_t$  is  $I(0)$  if  $\zeta_t - E(\zeta_t)$  is covariance stationary with

nonzero and bounded spectral density at all frequencies. Then a scalar or vector  $\xi_t$  is  $I(1)$  if  $\Delta\xi_t$  is  $I(0)$ , where  $\Delta = 1 - L$ ,  $L$  being the lag operator. Similarly,  $\xi_t$  is  $I(2)$  if  $\Delta^2\xi_t$  is  $I(0)$ . Note that if a vector is  $I(d)$ ,  $d = 1, 2$ , our definition implies that at least one of its individual components must be  $I(d)$ . The rest of the components might also be  $I(d)$  or, alternatively, they might have a smaller integration order. In this sense, our definition is similar to that of Johansen (1995b). Note also that this definition does not preclude the existence of components of an  $I(d)$  vector which are fractional processes ( $I(c)$ ,  $c$  being a real number smaller than  $d$ ), but the model proposed below will exclude this possibility. Next, we define cointegration for an  $I(d)$ ,  $d = 1, 2$ , process. Given a  $p \times 1$  process  $z_t \sim I(d)$ ,  $z_t$  is cointegrated if there exists a  $p \times 1$  vector  $\gamma \neq 0$  such that  $\gamma'z_t \sim I(c)$ , with  $c < d$ , prime denoting transposition. Again, this definition permits the existence of fractional linear combinations of the observables, but our model below excludes this possibility. Our definition of cointegration is similar to that of Johansen (1995b) and it is significantly more general than the standard notion of Engle and Granger (1987), where all observables are required to have identical integration orders. Note that according to our definition some of the cointegrating vectors might be unit vectors, just indicating that a particular observable has an integration order smaller than the order of the vector. As usual, the cointegrating rank among the elements of  $z_t$  is the number of linearly independent cointegrating vectors, and the space generated by these vectors will be denoted as cointegrating space.

As in Haldrup (1994), our purpose is to introduce an ADF statistic to test the null hypothesis of no stationary cointegration among the elements of a  $p$ -dimensional cointegrated  $I(2)$  vector of observables  $z_t$ . We assume that the cointegrating rank of  $z_t$  is  $r$ , where  $0 < r < p$ . Under the null,  $z_t$  is assumed to be generated by the model

$$\begin{pmatrix} I_r & B \\ 0 & I_{p-r} \end{pmatrix} (z_t - \mu_t) = \xi_t, \quad \begin{pmatrix} \Delta I_r & 0 \\ 0 & \Delta^2 I_{p-r} \end{pmatrix} \xi_t = \zeta_t, \quad (1)$$

where  $I_s$  is the  $s$ -rowed identity matrix,  $\zeta_t$  is a zero-mean  $I(0)$  vector process whose spectral density is finite and nonsingular at all frequencies,  $B$  is an  $r \times (p - r)$  matrix and  $\mu_t$  collects deterministic terms. Throughout, for any  $p$ -dimensional arbitrary vector  $\rho_t$ ,  $\rho_{(1)t}$  will be the vector collecting the first  $r$  components of  $\rho_t$ , while  $\rho_{(2)t}$  collects the rest, so that  $\rho_t = (\rho'_{(1)t}, \rho'_{(2)t})'$ . Thus, (1) could alternatively be written as

$$\begin{aligned} z_{(1)t} + Bz_{(2)t} &= \mu_{(1)t} + B\mu_{(2)t} + \xi_{(1)t}, \\ z_{(2)t} &= \mu_{(2)t} + \xi_{(2)t}, \end{aligned}$$

where the individual components of  $z_{(2)t}$  are  $I(2)$  and do not cointegrate. Model (1) captures a variety of situations where the cointegrating rank of  $z_t$  is  $r$ . If none of the rows of  $B$  is identically zero, all individual observables in  $z_t$  are  $I(2)$ . Alternatively, if  $B = 0$ , the  $r$  individual components in  $z_{(1)t}$  are  $I(1)$ . In this case, the  $r$  cointegrating relations are trivial. The situation where there are some  $I(1)$  components and the  $I(2)$  individual components cointegrate, is also covered by (1), whenever some (but not all) of the rows of  $B$  are identically zero. Note also that there is no loss of generality in the representation (1). If  $z_t$  cointegrates with rank  $r$ , a trivial extension of Theorem 1 of Gomez-Biscarri and Hualde (2014) (GBH hereinafter) ensures the existence of a  $(p - r)$ -dimensional subvector of  $z_t$  (say  $z_{(2)t}$ ) whose individual components are  $I(2)$  and do not cointegrate. These variables represent a set of common trends in the system. Also, collecting the rest of the observables in  $z_{(1)t}$ , by the same theorem, there exists an  $r \times (p - r)$  matrix  $B$  such that  $z_{(1)t} + Bz_{(2)t} \sim I(c)$ ,  $c < 2$ .

Haldrup (1994) considers the case where there are  $r - 1$   $I(1)$  observables, and the  $p - r + 1$   $I(2)$  observables cointegrate with rank one, so the cointegrating rank in  $z_t$  is  $r$ . Thus, in his setting,  $z_{(1)t}$  is composed of the  $r - 1$   $I(1)$  and one of the  $I(2)$  variables (the one which is not part of the common trends). His ADF test is based on the (ordinary least squares, OLS) regression of the  $I(2)$  variable in  $z_{(1)t}$  on the rest of the observables and deterministic terms. The null limiting distribution of this statistic is basically dependent on a vector of both nonintegrated and integrated Brownian motions, arising from the  $I(1)$  observables and the single cointegrating relation among the  $I(2)$  observables and from the  $I(2)$  common trends ( $z_{(2)t}$ ), respectively. However, unless these two types of Brownian motions are mutually independent (due for example to a zero coherence at frequency zero between the  $I(0)$  error input processes generating the  $I(1)$  and  $I(2)$  components, respectively), the typical decomposition (see, e.g., the proof of Lemma 2 in Haldrup, 1994) leading to standard (and mutually independent) nonintegrated and integrated Brownian motions is not valid. Therefore, in general, the limiting distribution of Haldrup's statistic is not free of nuisance parameters.

Fortunately, a simple correction can be carried out in the regression, so a proper orthogonalization can be achieved in general circumstances. Our proposed test statistic is based on residuals ( $\hat{u}_t$ ) arising from the regression of  $z_{1t}$  on  $z_{-1,t}$ ,  $\Delta z_{(2)t}$  and deterministic terms ( $c_t$ ), where  $z_{1t}$  is the first component of  $z_t$  (which obviously coincides with the first component of  $z_{(1)t}$ ) and  $z_{-1,t}$  collects the rest of elements of  $z_t$ . The inclusion of the extra regressors  $\Delta z_{(2)t}$  (first differences of the  $I(2)$  common trends) is what distinguishes our proposal from Haldrup's (1994). This simple modification leads to a pivotal null limiting distribution and, additionally, makes the test

consistent under the alternative of multicointegration (see Theorem 2 below).

We should comment on several crucial issues here. First, the test requires a choice for the variables in  $z_{(2)t}$ . In Section 3 we explain how this, along with  $r$ , can be inferred from the data. Second, as in Haldrup (1994), we account for deterministic terms by including a  $d \times 1$  vector of deterministic components  $c_t$  in the possible cointegrating regression. Although more general structures could have been allowed for, we focus on deterministic polynomial trends by letting  $c_t = (1, t, \dots, t^{d-1})'$ . Then we assume there exist  $r \times d$ ,  $(p-r) \times d$  matrices  $A_1$ ,  $A_2$ , respectively, such that

$$\mu_{(1)t} + B\mu_{(2)t} = -A_1c_t, \quad \mu_{(2)t} = -A_2c_t \quad (2)$$

(so it immediately follows that  $\Delta\mu_{(2)t} = -A_3c_t$ , for a corresponding  $(p-r) \times d$  matrix  $A_3$ ). Note that (2) just implies that the deterministic components characterizing the observables are general polynomial trends which are captured by  $c_t$ , the most relevant cases being  $c_t = 1, (1, t)', (1, t, t^2)'$ . Note also that defining  $D = \text{diag}\{1, n, \dots, n^{d-1}\}$ ,  $D^{-1}c_{[ns]} \rightarrow f(s)$ , as  $n \rightarrow \infty$ , where  $[\cdot]$  denotes integer part and  $f(s) = (1, s, \dots, s^{d-1})'$ .

Next, the null limiting distribution of our proposed test statistic is invariant to the choice of left hand side variable on the regression from which the residuals  $\hat{u}_t = (1, -\hat{\beta}')(z'_t, \Delta z'_{(2)t}, c'_t)'$  (where  $\hat{\beta}$  is the OLS estimator in this regression) are derived, as long as this choice is taken from  $z_{(1)t}$ . However, as in any residual-based test for cointegration, the choice of left hand side variable in the regression is important for power considerations and, again, we will address this issue in Section 3. Additionally, this null limiting distribution is also invariant to  $B$ ,  $A_i$ ,  $i = 1, 2, 3$ . The reason is that defining

$$T = \begin{pmatrix} I_r & B & 0 & A_1 \\ 0 & 0 & I_{p-r} & A_3 \\ 0 & I_{p-r} & 0 & A_2 \\ 0 & 0 & 0 & I_d \end{pmatrix},$$

then

$$\hat{u}_t = (1, -\hat{\beta}')T^{-1}T(z'_t, \Delta z'_{(2)t}, c'_t)' = (1, -\tilde{\theta}')v_t, \quad (3)$$

where  $v_t = (\xi'_{(1)t}, \Delta\xi'_{(2)t}, \xi'_{(2)t}, c'_t)'$  (see (1)), and  $\tilde{\theta}$  is the OLS estimator of  $v_{1t}$  on  $v_{-1,t}$  (where  $v_{1t}$  is the first component of  $v_t$ , and  $v_{-1,t}$  collects the rest of elements of  $v_t$ ). Noting (1),  $v_t$  is just a simple transformation of  $\xi_t$ ,  $c_t$ , which does not depend on  $B$  or  $A_i$ ,  $i = 1, 2, 3$ .

We introduce some assumptions which are similar to those in Chang and Park (2002). When applied to matrices, denote by  $\|\cdot\|$  the norm  $\|A\| = \sup_{\|x\| \leq 1} \|Ax\|$ , whereas  $\|\cdot\|$  applied to vectors is the usual Euclidean norm. Notice that if  $a_{ij}$  denotes

the  $(i, j)$ -th element of a  $p \times p$  matrix  $A$ ,  $\|A\|^2 \leq \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2$ .

**Assumption 1.** The process  $\zeta_t$  in (1) has representation

$$\zeta_t = A(L) \varepsilon_t, \text{ where } A(u) = I_p + \sum_{j=1}^{\infty} A_j u^j,$$

and the  $A_j$  are  $p \times p$  matrices such that:

- (i)  $\det(A(u)) \neq 0$ ,  $|u| = 1$ ;
- (ii)  $A(e^{i\lambda})$  is differentiable in  $\lambda$  with derivative in  $Lip(\eta)$ ,  $\eta > 1/2$ ;
- (iii)  $(\varepsilon_t, \mathcal{F}_t)$  is a martingale difference sequence with some filtration  $(\mathcal{F}_t)$  such that  $E(\varepsilon_t) = 0$ ,  $E(\varepsilon_t \varepsilon_t') = \Sigma$ ,  $\Sigma$  is positive definite,  $n^{-1} \sum_{t=1}^n \varepsilon_t \varepsilon_t' \rightarrow_p \Sigma$ ,  $E\|\varepsilon_t\|^u < K$  with  $u \geq 4$ , where  $K$  is some constant that depends only upon  $u$ .

Assumption 1 implies that  $\zeta_t$  is a fairly general linear process with martingale difference innovations. Note that (ii) implies the summation condition  $\sum_{j=1}^{\infty} j \|A_j\| < \infty$ , so (ii) and (iii) imply that  $\zeta_t$  is weakly stationary, whereas (iii) holds under suitable mixing conditions. In addition, Assumption 1 enables us to apply the multivariate invariance principle

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{[ns]} \zeta_t \Rightarrow B(s), \quad (4)$$

where  $B(s)$  is a  $p$ -vector Brownian motion with covariance matrix  $\Omega = A(1) \Sigma A(1)'$  and “ $\Rightarrow$ ” denotes weak convergence.

Given the previously defined residuals  $\hat{u}_t$ , the ADF test statistic is the  $t$ -ratio corresponding to the coefficient of  $\hat{u}_{t-1}$  in the regression of  $\Delta \hat{u}_t$  on  $\hat{u}_{t-1}, \Delta \hat{u}_{t-1}, \dots, \Delta \hat{u}_{t-q}$ . We will denote this  $t$ -ratio by  $t_n$ , and give its null limiting distribution in Theorem 1 below. As is well known (see, e.g., Phillips and Ouliaris, 1990), it is necessary in general to let  $q$  increase with  $n$ , for which we impose the following condition.

**Assumption 2.** Let  $q \rightarrow \infty$  and  $q = o(n^{1/3})$  as  $n \rightarrow \infty$ .

This condition guarantees the consistency of the estimators of autoregressive parameters in a particular autoregressive approximation (see, e.g., Berk, 1974, Chang and Park, 2002), which is a required step when calculating the null limiting distribution of our test statistic.

Before presenting the main result we introduce some additional notation. For a vector process  $G(s)$ ,  $G_1(s)$  denotes its first component and  $G_{-1}(s)$  the subvector resulting from omitting this first component. Also, given an arbitrary Brownian motion  $G(s)$ , define the integrated Brownian motion  $\bar{G}(s) = \int_0^s G(l) dl$ .

Let  $W(s)$  be a  $p$ -dimensional standard Brownian motion, let  $W_{(2)}(s)$  be the sub-vector made of the last  $p-r$  components of  $W(s)$  and let  $V(s) = (W'(s), \overline{W}'_{(2)}(s), f'(s))'$ . Finally, let  $Q(s) = \kappa'V(s)$ , where

$$\kappa = \left( 1, - \int_0^1 V_1(s) V'_{-1}(s) ds \left( \int_0^1 V_{-1}(s) V'_{-1}(s) ds \right)^{-1} \right)'$$

**Theorem 1.** Let  $z_t$  be generated by (1) and Assumptions 1 and 2 hold. Then, as  $n \rightarrow \infty$ ,

$$t_n \Rightarrow \Xi(p, r) \equiv \frac{\int_0^1 Q(s) dQ(s)}{\left( \int_0^1 Q^2(s) ds \right)^{\frac{1}{2}} \left( \kappa' \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix} \kappa \right)^{\frac{1}{2}}}. \quad (5)$$

The proof is provided in the Appendix. The distribution of the ADF test statistic  $t_n$  is free of nuisance parameters, but it depends on  $p$ ,  $r$  and the deterministic components  $c_t$  included in the regression that generated  $\hat{u}_t$ . Additionally, it can be easily shown that this test is consistent under particular alternatives of stationary cointegration, namely those where  $z_{1t}$  appears with nonzero coefficient in the stationary linear combination, including also any type of multicointegration (see Theorem 2 below) a possibility which is not contemplated by Haldrup's (1994) setting. Specifically, like other residual-based statistics,  $t_n$  diverges to  $-\infty$  as  $n \rightarrow \infty$ . We have calculated the simulated quantiles of  $\Xi(p, r)$  for series of length  $n = \{50, 100, 250, 500, 50,000\}$ , 200,000 replications and different  $(p, r)$  combinations. In particular, we generated the vector of observables  $z_t$  for cases  $p = 2, \dots, 6$ ,  $r = 1, \dots, p-1$ , choosing  $\zeta_t$  to be a  $p$ -dimensional zero mean normal innovation with covariance matrix  $I_p$  and independent over time. We computed the ADF statistic from the auxiliary regression

$$\Delta \hat{u}_t = \delta \hat{u}_{t-1} + \varphi_t,$$

where  $\hat{u}_t$  are residuals corresponding to three different characterizations of  $c_t$ , namely  $c_t = 1, (1, t)', (1, t, t^2)'$  (corresponding simulated quantiles are reported in Tables 1, 2, 3, respectively).

### 3 Implementing the test in practice: power considerations and feasibility

From a practical point of view our test requires knowledge of the cointegrating rank  $r$  (which characterizes the null limiting distribution (5)) and the identification of a valid set of  $p - r$   $I(2)$  common trends. Let us assume for the moment that this information is known to the researcher. Thus, given a set of observables  $z_t = (z'_{(1)t}, z'_{(2)t})'$ , where  $z_{(1)t}$  is  $r \times 1$  and  $z_{(2)t}$  collects a valid set of  $p - r$   $I(2)$  common trends, the question is how to design a test for the null of no stationary cointegration among levels and (possibly) first differences of the observables  $z_t$  (say  $H_0$ ) which is consistent under the alternative of stationary cointegration (say  $H_1$ ). As mentioned before, the problem arises because any of the  $r$  components in  $z_{(1)t}$  can be potentially chosen as the left hand side variable in the regression from which residuals  $\hat{u}_t$  are generated, but the corresponding test statistic only has power against alternatives where the chosen left hand side variable appears with nonzero coefficient in the stationary relation. This is in fact a problem suffered by any residual-based method for testing cointegration. However, there is a simple way to construct a test procedure for  $H_0$  which is consistent against any alternative of stationary cointegration (incidentally, the idea behind this testing procedure is also applicable in similar settings, such as the  $I(1)$  case). Denote by  $t_n^{(i)}$ ,  $i = 1, \dots, r$ , the test statistic derived from residuals arising from the OLS regression of  $z_{it}$  on the rest of the observables,  $\Delta z_{(2)t}$  and deterministic components, where we denote by  $\epsilon_{it}$  the  $i$ th component of an arbitrary vector  $\epsilon_t$ . Note that for any  $\alpha \in (0, 1)$ , (5) implies that there exist real numbers  $v_i(\alpha)$ ,  $i = 1, \dots, r$ , such that, under  $H_0$ ,

$$\Pr(t_n^{(i)} < v_i(\alpha)) \rightarrow \alpha, \text{ as } n \rightarrow \infty.$$

Note that we allow for different critical values  $v_i(\alpha)$ ,  $i = 1, \dots, r$ , corresponding to (possibly) different specifications of the deterministic components in each of the estimated regressions. Our proposed test is to reject  $H_0$  if  $t_n^{(i)} < v_i(\alpha/r)$  for at least one of the  $i$ 's, where  $i = 1, \dots, r$ . As shown below, this test has asymptotic level  $\alpha$  and is consistent under  $H_1$ . First, under  $H_0$ ,

$$\lim_{n \rightarrow \infty} \Pr(\text{Reject } H_0) = \lim_{n \rightarrow \infty} \Pr\left(\bigcup_{i=1}^r \{t_n^{(i)} < v_i(\alpha/r)\}\right) \leq \sum_{i=1}^r \lim_{n \rightarrow \infty} \Pr(t_n^{(i)} < v_i(\alpha/r)) \leq \alpha.$$

Under  $H_1$ , at least one of the components of  $z_{(1)t}$  must appear with nonzero coefficient in the stationary linear combination. Say, e.g., that such component is  $z_{jt}$  for  $j \in$

$\{1, \dots, r\}$ . Then,

$$\lim_{n \rightarrow \infty} \Pr(\text{Reject } H_0) \geq \lim_{n \rightarrow \infty} \Pr(t_n^{(j)} < v_j(\alpha/r)) = 1,$$

because, as mentioned before,  $t_n^{(j)}$  diverges (to  $-\infty$ ) as  $n \rightarrow \infty$ .

There is an additional point of concern regarding power: given that in the cointegrating regression we include first differences of a particular set of common trends ( $z_{(2)t}$ ) and noting that there might be alternative sets of valid common trends, we might wonder whether other multicointegrating relations (apart from those assessed by the test) are possible. Fortunately, Theorem 2 below rules out the existence of these alternative relations: if there are multicointegrating stationary relations, these must arise from combinations between  $z_t$  and  $\Delta z_{(2)t}$ .

**Theorem 2.** Let  $z_t$  be a  $p$ -dimensional cointegrated  $I(2)$  vector, with cointegrating rank  $r$ , where  $0 < r < p$ . Define two subspaces  $R, T$  of the cointegrating space ( $C$ ) in the following way:

- i. Given a  $p$ -dimensional vector  $\phi$ ,  $\phi \in R \subseteq C$  if there exists a  $p$ -dimensional vector  $\lambda(\phi)$  such that  $\phi' z_t + \lambda'(\phi) \Delta z_t$  is stationary;
- ii. Given a  $p$ -dimensional vector  $\phi$ ,  $\phi \in T \subseteq C$  if there exists a  $(p - r)$ -dimensional vector  $\rho(\phi)$  such that  $\phi' z_t + \rho'(\phi) \Delta \bar{z}_t$  is stationary, where  $\bar{z}_t$  is an arbitrary  $(p - r)$ -dimensional subvector of  $z_t$  with  $I(2)$  and not cointegrated individual components.

Then,  $R$  and  $T$  are identical subspaces ( $R = T$ ), that is, they contain the same elements.

The proof of Theorem 2 is given in the Appendix. Note that this result is parallel to that in Johansen (1995a), where multicointegration is tested with first differences of the common trends, which, in his setting, are particular linear combinations of the observables. Alternatively, we identify the common trends by the  $p - r$  dimensional vector of observables  $z_{(2)t}$ .

Next we address the issue of feasibility of the test. As anticipated, in practice,  $r$  and a correct choice for the  $I(2)$  common trends are unknown. In some cases, however, this might not be an issue of concern if the researcher were willing to make some assumptions based on generally accepted evidence. For example, in PPP analyses it is standard to look for a stationary relationship between a (log)exchange rate  $e_t$  and two (log)price levels  $p_t$  and  $p_t^*$ . As said before, these two prices might well be characterized as  $I(2)$  variables, whereas  $e_t$  is usually taken as  $I(1)$ . In addition,

in many cases it is well established that inflation rates of different countries are cointegrated, which implies that corresponding price levels also cointegrate. If this evidence is true, then  $r = 2$  and either of the two prices could act as the common trend. In this setting, Bacchiocchi and Fanelli (2005) derive an equilibrium PPP condition which, interestingly, includes, apart from the exchange rate and the two price levels, an inflation rate and the differential of the two inflation rates. Thus, due to the inclusion of the inflation rate, the equilibrium condition that Bacchiocchi and Fanelli (2005) use in their paper mimics the regression from which our test statistic would have been derived. In fact, our proposal allows for testing the stationarity of this equilibrium condition without resorting to a fully-fledged system method.

If the researcher does not want to work under such assumptions, or if there is no clear cut evidence from which such assumptions can be derived, we could infer  $r$  and the choice of common trends by a slight modification of the GBH procedure. In an  $I(1)$  setting, GBH proposed a sequential approach which relies on the residual-based ADF test of Phillips and Ouliaris (1990). This method leads to an estimator of  $r$  and to the identification of a set of common trends. The intuition behind their proposal is the following. First, if all pairs of observables are cointegrated, then necessarily  $r = p - 1$ . If not, there is at least a pair of non-cointegrated observables (common trends), and the next step is to test whether all trios containing this pair are cointegrated. If they are,  $r = p - 2$ , while if they are not, the next step is carried out. The procedure is finalized when all corresponding groups of observables are cointegrated, or, alternatively, when in the last possible step, cointegration among all observables is checked.

Nicely, the GBH method can be equally applied to infer the cointegrating rank in  $I(2)$  systems. There is however an important difference, because under the null of no cointegration, the residuals of the different cointegrating regressions are linear combinations of non-cointegrated  $I(2)$  variables. Hence, the critical values of Phillips and Ouliaris (1990) are not applicable. More importantly, the test based on these residuals is not consistent under the alternative of  $I(1)$  cointegration. However, performing the standard ADF test on first differences of these residuals sorts out this latter problem, although it is necessary to modify slightly the proof arguments of Phillips and Ouliaris (1990) in order to find the appropriate null limiting distribution of this ADF statistic (which differs from that in Phillips and Ouliaris, 1990; this is available from the authors upon request). In any case, this modified GBH procedure leads to an estimator of the rank  $r$  (say  $\hat{r}$ ) and, as a by-product, to the identification of a set of  $p - \hat{r}$  common trends. Thus a feasible version of the test to detect stationary cointegration consists on estimating  $r$ , making a data-based choice of a set of  $I(2)$

common trends and, based on that choice, performing the test as in the infeasible situation. We explore the behaviour of such feasible procedure in the next section.

## 4 Monte Carlo evidence

We investigate the finite sample performance of our test by means of a Monte Carlo experiment. We fix  $p = 4$ , and in all cases the analysis is based on 10,000 replications of series of lengths  $n = 100, 200, 500, 1,000$ . We generated  $\varepsilon_t$  as a Gaussian white noise with  $E(\varepsilon_t) = 0$ ,  $Var(\varepsilon_t) = I_4$ , and examine two different DGPs for the innovation vector  $\zeta_t$  (defined in Assumption 1):  $A(L) = I_4$  (WN),  $A(L) = (1 - 0.8L)^{-1} I_4$  (AR). We simulated 15 different models for  $z_t$  which can be classified according to their cointegrating rank and the type of cointegration. In all cases

$$\begin{pmatrix} I_r & B \\ 0 & I_{p-r} \end{pmatrix} z_t = w_t,$$

where:

$r = 3$ :  $B = (-1, -1, -1)'$ ,  $\Delta^2 w_{4t} = \zeta_{4t}$ , and

Model 1:  $w_{it} = \zeta_{it}$ ,  $i = 1, 2, 3$  (three  $I(0)$  relations, no multicointegration);

Model 2:  $w_{it} = \Delta z_{4t} + \zeta_{it}$ ,  $i = 1, 2, 3$  (three  $I(0)$  multicointegration relations);

Model 3:  $w_{it} = \zeta_{it}$ ,  $i = 1, 2$ ,  $\Delta w_{3t} = \zeta_{3t}$  (two  $I(0)$  relations, no multicointegration);

Model 4:  $w_{it} = \Delta z_{4t} + \zeta_{it}$ ,  $i = 1, 2$ ,  $\Delta w_{3t} = \zeta_{3t}$  (two  $I(0)$  multicointegration relations);

Model 5:  $w_{1t} = \zeta_{1t}$ ,  $\Delta w_{it} = \zeta_{it}$ ,  $i = 2, 3$  (one  $I(0)$  relation, no multicointegration);

Model 6:  $w_{1t} = \Delta z_{4t} + \zeta_{1t}$ ,  $\Delta w_{it} = \zeta_{it}$ ,  $i = 2, 3$  (one  $I(0)$  multicointegration relation);

Model 7:  $\Delta w_{it} = \zeta_{it}$ ,  $i = 1, 2, 3$  (no  $I(0)$  cointegration);

$r = 2$ :  $B = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$ ,  $\Delta^2 w_{it} = \zeta_{it}$ ,  $i = 3, 4$ , and

Model 8:  $w_{it} = \zeta_{it}$ ,  $i = 1, 2$  (two  $I(0)$  relations, no multicointegration);

Model 9:  $w_{it} = \Delta z_{3t} + \Delta z_{4t} + \zeta_{it}$ ,  $i = 1, 2$  (two  $I(0)$  multicointegration relations);

Model 10:  $w_{1t} = \zeta_{1t}$ ,  $\Delta w_{2t} = \zeta_{2t}$  (one  $I(0)$  relation, no multicointegration);

Model 11:  $w_{1t} = \Delta z_{3t} + \Delta z_{4t} + \zeta_{1t}$ ,  $\Delta w_{2t} = \zeta_{2t}$  (one  $I(0)$  multicointegration relation);

Model 12:  $\Delta w_{it} = \zeta_{it}$ ,  $i = 1, 2$  (no  $I(0)$  cointegration);

$r = 1$ :  $B = (-1, -1, -1)$ ,  $\Delta^2 w_{it} = \zeta_{it}$ ,  $i = 2, 3, 4$ , and

Model 13:  $w_{1t} = \zeta_{1t}$  (one  $I(0)$  relation, no multicointegration);

Model 14:  $w_{1t} = \Delta z_{2t} + \Delta z_{3t} + \Delta z_{4t} + \zeta_{1t}$  (one  $I(0)$  multicointegration relation);

Model 15:  $\Delta w_{1t} = \zeta_{1t}$  (no  $I(0)$  cointegration);

Models 7, 12, 15 are generated under the null  $H_0$  (there is no stationary cointegration), whereas the rest are generated under the alternative  $H_1$ , multicointegration being present in Models 2, 4, 6, 9, 11, 14. There are two parts to our experiment. First, we examine the power and size properties of our test statistic by presenting results for the infeasible version of the procedure presented in Section 3. By infeasible we mean that the cointegrating rank  $r$  and the correct choice of common trends are taken as known. Note that although the observables have been generated with zero mean, a constant is included in the different (possibly) cointegrating relations. Three different significance levels  $\alpha = \{.10, .05, .01\}$  and asymptotic critical values were used in the tests, and, in all cases, the number of lags in the ADF tests is chosen according to the BIC. Proportion of rejections are reported in Table 4. Overall, the performance is very satisfactory. Results for WN are in general superior to those for AR, especially when  $n = 100$ , where under AR the procedure exhibits oversizing and lack of power. However, as  $n$  increases, the rejection proportions behave as theory predicts, its performance being adequate for  $n = 200$  and excellent when  $n \geq 500$ .

In the second part of the experiment we compare the feasible procedure described in Section 3 with Johansen's  $I(2)$  method (as in Paruolo, 1996). Specifically we use the  ${}_0Q$  statistic introduced by Paruolo (1996), where in all cases the number of lags in the vector autoregression is set to the corresponding true value. This statistic is used in a sequential manner, and for each replication we record in Table 5 the proportion of cases where the sequential testing determines that there exists stationary cointegration. In our procedure the number of lags in the ADF tests is chosen according to the BIC. Also, when the estimated rank was  $\hat{r} = 0$ , the null is not rejected, whereas if  $\hat{r} = 4$ , then the GBH method for rank estimation in  $I(1)$  systems is applied: here, if there is evidence of cointegration, the null is rejected. Results are presented in Table 6. For the WN case, Paruolo's (1996) procedure is superior in terms of power (especially when the number of  $I(0)$  relations is small), although differences are small and our proposal displays comparable size behaviour (which improves as  $n$  increases). The picture changes substantially under the AR scenario, where our methods seems to be superior in most circumstances: here, Paruolo's (1996) approach displays bad size performance (although this is somewhat corrected as  $n$

increases), our method performing substantially better, with results comparable to those of the infeasible alternative. In any case, our results are quite remarkable, given that our design favours Johansen’s approach, for which we have used the true lag length (contrary to our procedure, where this choice is data-based).

Note finally that our Monte Carlo focuses on triangular representations of cointegrated systems. However, alternative specifications like those in Paruolo (1996) (vector autoregressive processes of order 2), present a similar picture.

## 5 An empirical application: markups and inflation

Banerjee *et al.* (2001) (BCR, hereafter) analyze a model of the markup of prices for a closed economy to show that there is a long run negative relationship between the markup of prices over cost and inflation. This implies that the real wage may respond positively to inflation. As a consequence, real activity (and unemployment) would be related in the long run to inflation, making the long run Phillips curve not vertical. Also, firm’s profitability (and stock returns) would be negatively correlated with inflation.

In order to justify the empirical analysis, BCR setup a model which delivers a solution for the long-run markup of the form

$$mu \equiv p - \delta ulc - (1 - \delta)pm = \omega_0 + \omega_1 x - \omega_2 \Delta p, \quad (6)$$

where  $\delta$ ,  $\omega_0$ ,  $\omega_1$ ,  $\omega_2$ , are parameters,  $mu$  denotes the markup,  $p$ ,  $ulc$ , and  $pm$  are prices, unit labor costs and import prices, respectively, and  $x$  captures shifts in the bargaining position of labor and firms. In particular,  $x$  includes variables that characterize the firm’s competitive environment. The relationship (6) expresses a long-run equilibrium among the variables involved. Under certain assumptions, BCR simplify the equation above by assuming that the competitive environment of the firm (variables in  $x$ ) is constant. Thus, they express the long-run markup as a function of the inflation rate exclusively. The long-run markup equation (6) is then estimated using quarterly Australian data that run from 1970:1 to 1995:2, proxying the core variables,  $p_t$ ,  $ulc_t$  and  $pm_t$ , with the private consumption deflator, the Australian Treasury’s measure of non-farm unit labor costs and the imports implicit price deflator, respectively.

BCR suggest that the three core variables are  $I(2)$ , so they consider scenarios where these variables cointegrate to  $I(1)$  or to stationarity or present multicointegration, as implied by the presence of  $\Delta p$  in (6). Thus, the setup of their long-run analysis is an adequate context for the test proposed in the present paper. BCR’s

setting assumes that the variables in  $x$  are all stationary, i.e., they are only present as determinants of short-run deviations from the long-run markup. Specifically, these variables include the unemployment rate, a measure of tax rates, oil prices and a measure of the number of labor strikes. There is evidence that the first three of these variables are  $I(1)$ , so BCR include them in first differences in the analysis. However, there seems to be no theoretical reason to omit the variables in  $x$  from the analysis of cointegration, which could, in principle, allow for a long run relation that involves the six nonstationary variables in the dataset.

We first carry out the BCR analysis by testing for an stationary relationship among the three core variables. First, our feasible procedure must infer from the data  $r$  and a set of common trends. As said before, a slight modification of the GBH procedure for  $I(1)$  systems can be applied here. This method, similarly to Johansen's analysis, starts by assuming a maximum possible integration order of the system, two in our case. Then we test for unit roots on first differences of the three observables: if all tests reject, then the system is not  $I(2)$  and we can apply an  $I(1)$  methodology. Our results, however, support the  $I(2)$  condition of the system, although, interestingly, we found no evidence of  $I(2)$  behavior in  $ulc_t$  or  $pm_t$  (BCR themselves acknowledge that this evidence is weak; see their footnote 19). In any case, we conclude that the vector of observables is  $I(2)$  and next we choose a common trend. GBH proposed a statistic which can be used for this choice, although, for space reasons, we omit this discussion. The chosen variable is  $p_t$ , which aligns with the output of previous unit root tests. Then, evidence of cointegration between  $ulc_t$  and  $p_t$  and also between  $pm_t$  and  $p_t$  is assessed. This is carried out by ADF tests applied to first differences of residuals obtained from the corresponding regressions (where an intercept was included). The ADF statistics took values -11.37 and -9.50, respectively, both significant at 1%, showing in both cases strong evidence of (possibly trivial) cointegration. Thus, the GBH procedure leads to  $\hat{r} = 2$  and to the choice of  $p_t$  as common trend. Given this, we carried out the test for stationary cointegration. In particular we estimated by OLS the regressions

$$\begin{aligned} ulc_t &= \alpha_0 + \alpha_1 p_t + \alpha_2 \Delta p_t + \alpha_3 pm_t + u_t, \\ pm_t &= \beta_0 + \beta_1 p_t + \beta_2 \Delta p_t + \beta_3 ulc_t + v_t. \end{aligned}$$

Following the procedure described in Section 3, the individual tests should be performed with size  $\alpha/2$ , and if either of them rejects, we would reject  $H_0$ , therefore finding evidence of a stationary markup. The corresponding ADF test statistics take values -2.98, -2.03 (the 5% critical value is -4.24), so  $H_0$  is not rejected (even at the

10% level). Hence the data do not support the existence of a stationary markup.

As done by BCR, an alternative approach would have been to analyze the possibility of cointegration in the (likely)  $I(1)$  system formed by the three observables  $p_t - ulc_t$ ,  $p_t - pm_t$ ,  $\Delta p_t$ . Again, our results (available upon request) do not support the existence of a stationary markup.

As hinted before, if we were willing to consider that the competitive environment may not be constant in the long-run, then some of the variables in  $x$  may enter the equilibrium relationship. Thus, we extend our exercise by including oil prices ( $pet_t$ ), unemployment rate ( $ue_t$ ) and a tax rate ( $tax_t$ ) in the cointegration analysis. If any of these three variables enters a stationary cointegrating relationship, the resulting cointegrating error may be interpreted as the markup net of persistent shocks and inflation. Following the same steps as before, we conclude that  $\hat{r} = 5$  and, again,  $p_t$  is the chosen common trend. Then our test is carried out by estimating five different relations where all observables but  $p_t$  are regressed on the rest of observables and  $\Delta p_t$  (we also include an intercept). The ADF test statistics derived from the regressions where  $ulc_t$ ,  $pm_t$ ,  $pet_t$ ,  $ue_t$ ,  $tax_t$  are the left hand side variables are -6.09, -3.11, -2.63, -4.08, -7.24, respectively. These numbers must be compared with appropriate critical values, which in this case are given by -5.56 (2%), -5.82 (1%), -6.37 (0.2%) (note that the individual tests have to be performed with  $\alpha/5$  significance level). Thus, even at 1%, we reject  $H_0$ , suggesting that there exists stationary cointegration among these six variables.

Combining the results from both analyses, we could say that there is evidence of a persistent markup, even accounting for the effect of inflation; however, the markup net of shocks to the competitive environment appears to be stationary. The two estimated equations for possible stationary relations among the variables are

$$\begin{aligned}\widehat{ulc}_t &= \frac{2.36}{(0.12)} + \frac{0.94}{(0.04)}p_t + \frac{3.00}{(0.37)}\Delta p_t + \frac{0.03}{(0.05)}pm_t - \frac{0.03}{(0.03)}pet_t + \frac{0.10}{(0.01)}ue_t - \frac{0.23}{(0.06)}tax_t, \\ \widehat{tax}_t &= \frac{2.86}{(0.31)} + \frac{0.51}{(0.16)}p_t + \frac{2.07}{(0.73)}\Delta p_t - \frac{0.57}{(0.15)}ulc_t + \frac{0.20}{(0.07)}pm_t + \frac{0.14}{(0.05)}pet_t + \frac{0.01}{(0.03)}ue_t,\end{aligned}$$

where standard errors are given in parenthesis. Signs and magnitudes are aligned with what theory would suggest.

## 6 Conclusions

Our main interest in the paper is the analysis of long run relationships that involve  $I(2)$  and, possibly,  $I(1)$  observables. The objective was to detect linear combinations of these observables that lead to stationary cointegrating errors. Cointegrating

regressions that combine the  $I(2)$  and  $I(1)$  observables can be used to test for this possibility, but care has to be exercised to make sure that these regressions are well specified. In particular, we show that a correction must be implemented in those regressions, which consists of including as additional regressors the first differences of non-cointegrated  $I(2)$  observables that characterize a set of common trends of the system.

Once this correction has been implemented, traditional ADF tests can be applied to the residuals from this cointegrating regression in order to test the null hypothesis of no stationary cointegration. We have derived the asymptotic distribution of the proposed ADF test and show that it depends on the number of observables  $p$  and on the number of  $I(2)$  common trends  $(p - r)$  (or, alternatively, on the cointegrating rank of the system,  $r$ ). We have tabulated the critical values of this distribution for a number of cases. Also, based on this ADF test statistic, we propose a procedure to test the null of no stationary cointegration which overcomes the drawback, suffered by any residual-based method, of the lack of power with respect to some relevant alternatives. We have compared the behavior of this test with that of a system alternative and show that it performs satisfactorily. Finally, we have illustrated the use of the test by means of an empirical analysis of markups and inflation.

## Appendix

**Proof of Theorem 1.** Using similar notation to that of Phillips and Ouliaris (1990), the ADF statistic is

$$t_n = \frac{\widehat{U}'_{-1} Q_{X_q} \Delta \widehat{U}}{\left(\widehat{U}'_{-1} Q_{X_q} \widehat{U}_{-1}\right)^{\frac{1}{2}} \widehat{\sigma}}, \quad (7)$$

where  $Q_{X_q} = I_{n-q-1} - X_q (X_q' X_q)^{-1} X_q'$ ,  $X_q = (x_{q,q+2}, \dots, x_{q,n})'$ ,  $x_{q,t} = (\Delta \widehat{u}_{t-1}, \dots, \Delta \widehat{u}_{t-q})'$ ,  $\widehat{U}_{-1} = (\widehat{u}_{q+1}, \dots, \widehat{u}_{n-1})'$ ,  $\Delta \widehat{U} = (\Delta \widehat{u}_{q+2}, \dots, \Delta \widehat{u}_n)'$ ,

$$\widehat{\sigma}^2 = \frac{1}{n - q - 1} \sum_t \left( \Delta \widehat{u}_t - \widehat{\alpha}_0 \widehat{u}_{t-1} - \sum_{j=1}^q \widehat{\alpha}_j \Delta \widehat{u}_{t-j} \right)^2,$$

where  $\Sigma_t = \Sigma_{t=q+2}^n$  and  $\widehat{\alpha}_j$ ,  $j = 0, \dots, q$ , are the ordinary least squares coefficients in the regression of  $\Delta \widehat{u}_t$  on  $\widehat{u}_{t-1}$ ,  $\Delta \widehat{u}_{t-1}$ ,  $\dots$ ,  $\Delta \widehat{u}_{t-q}$ . First, noting that  $\widehat{u}_t = \left(1, -\widehat{\theta}'\right) v_t$ , define

$$\widehat{\eta} = \begin{pmatrix} I_p & 0 & 0 \\ 0 & nI_{p-r} & 0 \\ 0 & 0 & n^{-1/2}D \end{pmatrix} \begin{pmatrix} 1 \\ -\widehat{\theta} \end{pmatrix}.$$

By (4) and the continuous mapping theorem

$$\widehat{\eta} \Rightarrow \eta \equiv \left( 1, - \int_0^1 X_1(s) X'_{-1}(s) ds \left( \int_0^1 X_{-1}(s) X'_{-1}(s) ds \right)^{-1} \right)', \quad (8)$$

where  $X(s) = \left( B'(s), \overline{B}'_{(2)}(s), f'(s) \right)'$ ,  $B_{(2)}(s)$  being the subvector made of the last  $p-r$  components of  $B(s)$ . We will stress the dependence of the ADF statistic on  $\widehat{\eta}$  by defining  $t_n(\widehat{\eta}) \equiv t_n$ . Theorem 1 follows on showing that, as  $n \rightarrow \infty$ ,  $q \rightarrow \infty$ ,

$$t_n(\widehat{\eta}) - t_n(\eta) = o_p(1), \quad (9)$$

$$t_n(\eta) \Rightarrow \Xi(p, r), \quad (10)$$

where  $t_n(\eta)$  is as  $t_n(\widehat{\eta})$ , just replacing  $\widehat{\eta}$  by  $\eta$ . We show (10) first. The proof will be based on the following result. Under our assumptions,  $n^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} \zeta_t$  is a mixing sequence (see, e.g., Rootzén, 1976, Phillips and Durlauf, 1986, Phillips and Ouliaris, 1990), so  $t_n(\widehat{\eta})$  is also mixing. Then, if (9) holds, by Lemma 2.6 of Rootzén (1976)  $t_n(\eta)$  is also a mixing sequence, so conditioning on  $\eta$  does not affect the analysis of the limiting distribution of  $t_n(\eta)$ . Thus, we would act as if  $\eta$  were fixed. Noting that by (3)

$$\Delta \widehat{u}_t = \left( 1, -\widehat{\theta}' \right) \begin{pmatrix} \zeta_t \\ \Delta \xi_{(2)t} \\ \Delta c_t \end{pmatrix} = \widehat{\eta}' \begin{pmatrix} \zeta_t \\ n^{-1} \Delta \xi_{(2)t} \\ n^{1/2} D^{-1} \Delta c_t \end{pmatrix},$$

define  $\bar{u}_t$  and  $\bar{x}_{q,t}$  as  $\widehat{u}_t$  and  $x_{q,t}$ , respectively, but replacing  $\widehat{\eta}$  by  $\eta$  in these latter expressions. There is a slight abuse of notation here because

$$\Delta \bar{u}_t = \eta' \begin{pmatrix} \zeta_t \\ n^{-1} \Delta \xi_{(2)t} \\ n^{1/2} D^{-1} \Delta c_t \end{pmatrix},$$

so, strictly speaking, a more appropriate (but more cumbersome) notation would be  $\bar{u}_{t,n}$ , given that this is a triangular array.

First, we show that as  $q \rightarrow \infty$  and  $n \rightarrow \infty$ ,

$$\left( \frac{1}{n} \sum_t \bar{x}_{q,t} \bar{x}'_{q,t} \right)^{-1} = O_p(1), \quad (11)$$

$$\frac{1}{n} \sum_t \bar{u}_{t-1} \bar{x}_{q,t} = O_p\left(q^{\frac{1}{2}}\right). \quad (12)$$

Let  $\eta = (\eta'_a, \eta'_b, \eta'_c)'$ , where  $\eta_a, \eta_b, \eta_c$ , are  $p \times 1, (p-r) \times 1, d \times 1$ , subvectors of  $\eta$ , respectively. Then  $\bar{x}_{q,t} = a_{q,t} + b_{q,t} + c_{q,t}$ , where

$$\begin{aligned} a_{q,t} &= (\eta'_a \zeta_{t-1}, \dots, \eta'_a \zeta_{t-q})', \quad b_{q,t} = n^{-1}(\eta'_b \Delta \xi_{(2)t-1}, \dots, \eta'_b \Delta \xi_{(2)t-q})', \\ c_{q,t} &= n^{\frac{1}{2}}(\eta'_c D^{-1} \Delta c_{t-1}, \dots, \eta'_c D^{-1} \Delta c_{t-q})'. \end{aligned}$$

In order to show (11), note that

$$\frac{1}{n} \sum_t \bar{x}_{q,t} \bar{x}'_{q,t} = C_n + R_n,$$

where

$$C_n = (I_q \otimes \eta'_a) E \left( \begin{pmatrix} \zeta_{t-1} \\ \vdots \\ \zeta_{t-q} \end{pmatrix} (\zeta'_{t-1}, \dots, \zeta'_{t-q}) \right) (I_q \otimes \eta_a) = (I_q \otimes \eta'_a) \Gamma_q (I_q \otimes \eta_a)$$

where  $\otimes$  denotes the Kronecker product,

$$\Gamma_q = \begin{pmatrix} \Gamma(0) & \Gamma(1) & \cdots & \Gamma(q-1) \\ \Gamma(-1) & \Gamma(0) & \cdots & \Gamma(q-2) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma(1-q) & \Gamma(2-q) & \cdots & \Gamma(0) \end{pmatrix},$$

with  $\Gamma(j) = E(\zeta_t \zeta'_{t-j})$ , and  $R_n$  collects remaining terms. First we show that  $\|R_n\| = O_p(qn^{-1/2})$ , which by Assumption 2 is  $o_p(1)$ . This result follows because it can be shown that under our conditions, uniformly in  $i, j$ ,

$$\begin{aligned} E \left\| \sum_t \zeta_{t-i} \Delta \xi'_{(2)t-j} \right\|^2 &= O(n^2), \quad E \left\| \sum_t \zeta_{t-i} \Delta c'_{t-j} D^{-1} \right\|^2 = O(1), \\ E \left\| \sum_t \Delta \xi_{(2)t-i} \Delta \xi'_{(2)t-j} \right\|^2 &= O(n^4), \quad E \left\| \sum_t \Delta \xi_{(2)t-i} \Delta c'_{t-j} D^{-1} \right\|^2 = O(n), \\ \left\| \sum_t D^{-1} \Delta c_{t-i} \Delta c'_{t-j} D^{-1} \right\|^2 &= O(n^{-2}), \quad E \left\| \sum_t (\zeta_{t-i} \zeta'_{t-j} - \Gamma(j-i)) \right\|^2 = O(n). \end{aligned}$$

Then by the properties of the norm

$$\begin{aligned} \left\| \frac{1}{n} \sum_t a_{q,t} b'_{q,t} \right\|^2 &= O_p \left( \frac{q^2}{n^2} \right), & \left\| \frac{1}{n} \sum_t a_{q,t} c'_{q,t} \right\|^2 &= O_p \left( \frac{q^2}{n} \right) \\ \left\| \frac{1}{n} \sum_t b_{q,t} b'_{q,t} \right\|^2 &= O_p \left( \frac{q^2}{n^2} \right), & \left\| \frac{1}{n} \sum_t b_{q,t} c'_{q,t} \right\|^2 &= O_p \left( \frac{q^2}{n^2} \right), \\ \left\| \frac{1}{n} \sum_t c_{q,t} c'_{q,t} \right\|^2 &= O_p \left( \frac{q^2}{n^2} \right), & \left\| \frac{1}{n} \sum_t a_{q,t} a'_{q,t} - C_n \right\|^2 &= O_p \left( \frac{q^2}{n} \right), \end{aligned}$$

so it immediately follows that  $\|R_n\| = O_p(qn^{-1/2})$ . Next  $\|C_n^{-1}\| = O_p(1)$  because  $\Gamma_q$  is positive definite and  $I_q \otimes \eta_a$  is a full rank  $qp \times q$  matrix. Then, given that

$$\left\| \left( \frac{1}{n} \sum_t \bar{x}_{q,t} \bar{x}'_{q,t} \right)^{-1} - C_n^{-1} \right\| \leq \left\| \left( \frac{1}{n} \sum_t \bar{x}_{q,t} \bar{x}'_{q,t} \right)^{-1} - C_n^{-1} \right\| \|R_n\| \|C_n^{-1}\| + \|R_n\| \|C_n^{-1}\|^2,$$

noting that  $\|C_n^{-1}\| = O_p(1)$ ,  $\|R_n\| = o_p(1)$ ,  $1 - \|R_n\| \|C_n^{-1}\| > 0$  with probability approaching one, so that

$$\left\| \left( \frac{1}{n} \sum_t \bar{x}_{q,t} \bar{x}'_{q,t} \right)^{-1} - C_n^{-1} \right\| \leq \frac{\|R_n\| \|C_n^{-1}\|^2}{1 - \|R_n\| \|C_n^{-1}\|} = O_p \left( \frac{q}{n^{\frac{1}{2}}} \right),$$

to conclude the proof of (11). Next, (12) follows by similar arguments noting that, uniformly in  $j$ ,

$$\begin{aligned} E \left\| \frac{1}{n} \sum_t \begin{pmatrix} \xi_{(1)t-1} \\ \Delta \xi_{(2)t-1} \end{pmatrix} \zeta'_{t-j} \right\|^2 &= O(1), & E \left\| \frac{1}{n^2} \sum_t \begin{pmatrix} \xi_{(1)t-1} \\ \Delta \xi_{(2)t-1} \end{pmatrix} \Delta \xi'_{(2)t-j} \right\|^2 &= O(1), \\ E \left\| \frac{1}{n^{\frac{1}{2}}} \sum_t \begin{pmatrix} \xi_{(1)t-1} \\ \Delta \xi_{(2)t-1} \end{pmatrix} \Delta c'_{t-j} D^{-1} \right\|^2 &= O(1), & E \left\| \frac{1}{n^2} \sum_t \xi_{(2)t-1} \zeta'_{t-j} \right\|^2 &= O(1) \\ E \left\| \frac{1}{n^3} \sum_t \xi_{(2)t-1} \Delta \xi'_{(2)t-j} \right\|^2 &= O(1), & E \left\| \frac{1}{n^{\frac{3}{2}}} \sum_t \xi_{(2)t-1} \Delta c'_{t-j} D^{-1} \right\|^2 &= O(1), \\ E \left\| \frac{1}{n^{\frac{1}{2}}} \sum_t D^{-1} c_{t-1} \zeta'_{t-j} \right\|^2 &= O(1), & E \left\| \frac{1}{n^{\frac{3}{2}}} \sum_t D^{-1} c_{t-1} \Delta \xi'_{(2)t-j} \right\|^2 &= O(1), \\ \left\| \sum_t D^{-1} c_{t-1} \Delta c'_{t-j} D^{-1} \right\|^2 &= O(1). \end{aligned}$$

Then we deal with  $\bar{U}'_{-1} Q_{\bar{X}_q} \bar{U}_{-1}$ , where  $\bar{U}_{-1}$ ,  $Q_{\bar{X}_q}$ , are defined as  $\hat{U}_{-1}$ ,  $Q_{X_q}$ , replacing

$\widehat{u}_t, x_{q,t}$ , by  $\bar{u}_t, \bar{x}_{q,t}$ , respectively. This is one of the components of the denominator of  $t_n(\eta)$  (see (7)), and by (11), (12),

$$\frac{1}{n^2} \bar{U}'_{-1} Q_{\bar{X}_q} \bar{U}_{-1} = \frac{1}{n^2} \sum_t \bar{u}_{t-1}^2 + O_p\left(\frac{q}{n}\right). \quad (13)$$

Partitioning

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}$$

according to  $B(s) = \left( B'_{(1)}(s), B'_{(2)}(s) \right)'$ , define

$$S = \begin{pmatrix} I_r & -\Omega_{12}\Omega_{22}^{-1} & 0 & 0 \\ 0 & I_{p-r} & 0 & 0 \\ 0 & 0 & I_{p-r} & 0 \\ 0 & 0 & 0 & I_d \end{pmatrix}.$$

Then

$$\begin{aligned} \frac{1}{n^2} \sum_t \bar{u}_{t-1}^2 &= \frac{1}{n^2} \eta' S^{-1} \sum_t S \begin{pmatrix} \xi_{(1)t-1} \\ \Delta \xi_{(2)t-1} \\ n^{-1} \xi_{(2)t-1} \\ n^{1/2} D^{-1} c_{t-1} \end{pmatrix} \\ &\times (\xi'_{(1)t-1}, \Delta \xi'_{(2)t-1}, n^{-1} \xi'_{(2)t-1}, n^{1/2} c'_{t-1} D^{-1}) S' (S')^{-1} \eta. \end{aligned}$$

First, note that

$$S (\xi'_{(1)t-1}, \Delta \xi'_{(2)t-1}, n^{-1} \xi'_{(2)t-1}, n^{1/2} c'_{t-1} D^{-1}) = (w'_t, \Delta \xi'_{(2)t}, n^{-1} \xi'_{(2)t}, n^{1/2} c'_{t-1} D^{-1})',$$

where here  $w_t = \xi_{(1)t} - \Omega_{12}\Omega_{22}^{-1}\Delta\xi_{(2)t}$  is an  $I(1)$  process such that the coherence at frequency zero between  $\Delta w_t$  and  $\Delta^2 \xi_{(2)t}$  is zero. Define  $Z(s) = \left( B'_{(1.2)}(s), B'_{(2)}(s), \bar{B}'_{(2)}(s), f'(s) \right)'$ ,  $B_{(1.2)}(s) = B_{(1)}(s) - \Omega_{12}\Omega_{22}^{-1}B_{(2)}(s)$ , noting that  $B_{(1.2)}(s)$  and  $B_{(2)}(s)$  are independent Brownian motions and  $B_{(1.2)}(s)$  has covariance matrix  $\Phi = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$ .

Then, by (4) and the continuous mapping theorem

$$\begin{aligned}
\frac{1}{n^2} \sum_t \bar{u}_{t-1}^2 &\Rightarrow \eta' S^{-1} \int_0^1 Z(s) Z'(s) ds (S')^{-1} \eta \\
&= \int_0^1 Z_1^2(s) ds \\
&\quad - \int_0^1 Z_1(s) Z'_{-1}(s) ds \left( \int_0^1 Z_{-1}(s) Z'_{-1}(s) ds \right)^{-1} \int_0^1 Z_1(s) Z_{-1}(s) ds,
\end{aligned} \tag{14}$$

because

$$\eta' S^{-1} = \left( 1, - \int_0^1 Z_1(s) Z'_{-1}(s) ds \left( \int_0^1 Z_{-1}(s) Z'_{-1}(s) ds \right)^{-1} \right). \tag{15}$$

As in Phillips and Ouliaris (1990), let

$$\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \Phi_{22} \end{pmatrix} = L' L, \quad \text{where } L = \begin{pmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{pmatrix},$$

where  $\phi_{11}$ ,  $\phi_{12}$ ,  $\Phi_{22}$ , are  $1 \times 1$ ,  $1 \times (r-1)$ ,  $(r-1) \times (r-1)$ , matrices, respectively,  $\phi_{21} = \phi'_{12}$ , and  $l_{11} = (\phi_{11} - \phi_{12} \Phi_{22}^{-1} \phi_{21})^{1/2}$ ,  $l_{21} = \Phi_{22}^{-1/2} \phi_{21}$ ,  $L_{22} = \Phi_{22}^{1/2}$ . Thus

$$Z(s) = \begin{pmatrix} B_{(1,2)}(s) \\ B_{(2)}(s) \\ \bar{B}_{(2)}(s) \\ f(s) \end{pmatrix} = \begin{pmatrix} L' & 0 & 0 & 0 \\ 0 & \Omega_{22}^{1/2} & 0 & 0 \\ 0 & 0 & \Omega_{22}^{1/2} & 0 \\ 0 & 0 & 0 & I_d \end{pmatrix} V(s).$$

Then, by (13), (14) and obvious manipulations

$$\frac{1}{n^2} \bar{U}'_{-1} Q_{\bar{X}_q} \bar{U}_{-1} \Rightarrow l_{11}^2 \int_0^1 Q^2(s) ds. \tag{16}$$

Next

$$\frac{1}{n} \bar{U}'_{-1} Q_{\bar{X}_q} \Delta \bar{U} = \frac{1}{n} \sum_t \bar{u}_{t-1} \Delta \bar{u}_t - \frac{1}{n} \sum_t \bar{u}_{t-1} \bar{x}'_{q,t} \left( \sum_t \bar{x}_{q,t} \bar{x}'_{q,t} \right)^{-1} \sum_t \Delta \bar{u}_t \bar{x}_{q,t}, \tag{17}$$

noting that  $\Delta \bar{u}_t = \zeta'_t \eta_a + n^{-1} \Delta \xi'_{(2)t} \eta_b + n^{1/2} \Delta c'_t D^{-1} \eta_c$ . First, by similar arguments to those in the proofs of (11), (12), it is simple to show that

$$\begin{aligned} \frac{1}{n^2} \sum_t \bar{x}_{q,t} \Delta \xi'_{(2)t} \eta_b &= O_p\left(\frac{q}{n}\right), \quad \frac{1}{n^{1/2}} \sum_t \bar{x}_{q,t} \Delta c'_t D^{-1} \eta_c = O_p\left(\frac{q}{n^{1/2}}\right) \\ \frac{1}{n} \sum_t b_{q,t} \zeta'_t \eta_a &= O_p\left(\frac{q}{n}\right), \quad \frac{1}{n} \sum_t c_{q,t} \zeta'_t \eta_a = O_p\left(\frac{q}{n^{1/2}}\right) \\ \left(\frac{1}{n} \sum_t \bar{x}_{q,t} \bar{x}'_{q,t}\right)^{-1} - \left(\frac{1}{n} \sum_t a_{q,t} a'_{q,t}\right)^{-1} &= o_p(1), \end{aligned}$$

which implies that (17) equals

$$\begin{aligned} &\frac{1}{n} \sum_t \bar{u}_{t-1} \left( \zeta'_t \eta_a - a'_{q,t} \left( \sum_t a_{q,t} a'_{q,t} \right)^{-1} \sum_t a_{q,t} \zeta'_t \eta_a \right) + \frac{1}{n^2} \sum_t \bar{u}_{t-1} \Delta \xi'_{(2)t} \eta_b \\ &+ \frac{1}{n^{1/2}} \sum_t \bar{u}_{t-1} \Delta c'_t D^{-1} \eta_c - \frac{1}{n} \sum_t \bar{u}_{t-1} b'_{q,t} \left( \sum_t a_{q,t} a'_{q,t} \right)^{-1} \sum_t a_{q,t} \zeta'_t \eta_a \\ &- \frac{1}{n} \sum_t \bar{u}_{t-1} c'_{q,t} \left( \sum_t a_{q,t} a'_{q,t} \right)^{-1} \sum_t a_{q,t} \zeta'_t \eta_a + o_p(1). \end{aligned} \quad (18)$$

We concentrate on the fourth term of (18). First, we show that

$$\frac{1}{n} \sum_t \bar{u}_{t-1} b'_{q,t} - \frac{1}{n^2} \sum_t \bar{u}_{t-1} \Delta \xi'_{(2)t} \eta_b e'_q = o_p(1), \quad \text{as } q \rightarrow \infty, n \rightarrow \infty, \quad (19)$$

where  $e_q$  is a  $q$ -dimensional vector of ones. The  $m$ th element of the row vector on the left of (19) equals

$$-\frac{1}{n^2} \sum_t \bar{u}_{t-1} \sum_{l=1}^m \Delta^2 \xi'_{(2)t-l+1} \eta_b, \quad \text{for } m = 1, \dots, q,$$

which can be easily shown to be  $O_p(qn^{-1})$  uniformly in  $m$ , so

$$\left\| \frac{1}{n} \sum_t \bar{u}_{t-1} b'_{q,t} - \frac{1}{n^2} \sum_t \bar{u}_{t-1} \Delta \xi'_{(2)t} \eta_b e'_q \right\|^2 = O_p\left(\frac{q^3}{n^2}\right),$$

to conclude (19). Next we show that

$$\frac{1}{n} \sum_t \bar{u}_{t-1} c'_{q,t} - \frac{1}{n^{1/2}} \sum_t \bar{u}_{t-1} \Delta c'_t D^{-1} \eta_c e'_q = o_p(1), \quad \text{as } q \rightarrow \infty, n \rightarrow \infty. \quad (20)$$

The  $m$ th element of the row vector on the left of (20) equals

$$-\frac{1}{n^{1/2}} \sum_t \bar{u}_{t-1} \sum_{l=1}^m \Delta^2 c'_{t-l+1} D^{-1} \eta_c, \quad \text{for } m = 1, \dots, q,$$

which can be easily shown to be  $O_p(qn^{-1})$  uniformly in  $m$ , so

$$\left\| \frac{1}{n} \sum_t \bar{u}_{t-1} b'_{q,t} - \frac{1}{n^{1/2}} \sum_t \bar{u}_{t-1} \Delta c'_t D^{-1} \eta_c e'_q \right\|^2 = O_p\left(\frac{q^3}{n^2}\right).$$

Then the sum of the second and fourth terms of (18) becomes

$$\frac{1}{n^2} \sum_t \bar{u}_{t-1} \Delta \xi'_{(2)t} \eta_b \left( 1 - e'_q \left( \sum_t a_{q,t} a'_{q,t} \right)^{-1} \sum_t a_{q,t} \zeta'_t \eta_a \right) + o_p(1), \quad (21)$$

whereas that of the third and fifth terms of (18) becomes

$$\frac{1}{n^{\frac{1}{2}}} \sum_t \bar{u}_{t-1} \Delta c'_t D^{-1} \eta_c \left( 1 - e'_q \left( \sum_t a_{q,t} a'_{q,t} \right)^{-1} \sum_t a_{q,t} \zeta'_t \eta_a \right) + o_p(1).$$

Next, as in Phillips and Ouliaris (1990), denote  $\rho_t = \eta'_a \zeta_t$ , which (conditional on  $\eta_a$ ) has an autoregressive representation

$$d(L) \rho_t = \gamma_t, \quad d(s) = \sum_{j=0}^{\infty} d_j s^j, \quad d_0 = 1,$$

where the sequence  $d_j$  is absolutely summable and  $\gamma_t$  is a zero-mean orthogonal sequence with variance  $d^2(1) \eta'_a \Omega \eta_a$ . Next, note that

$$1 - e'_q \left( \sum_t a_{q,t} a'_{q,t} \right)^{-1} \sum_t a_{q,t} \zeta'_t \eta_a = 1 + \sum_{j=1}^q \widehat{d}_j,$$

where  $-\widehat{d}_j$  is the estimated coefficient corresponding to  $\rho_{t-j}$ ,  $j = 1, \dots, q$ , in the regression of  $\rho_t$  on  $\rho_{t-1}, \dots, \rho_{t-q}$ . As in Lemma 3.4 of Chang and Park (2002),

$$1 - e'_q \left( \sum_t a_{q,t} a'_{q,t} \right)^{-1} \sum_t a_{q,t} \zeta'_t \eta_a \rightarrow_p d(1), \quad \text{as } q \rightarrow \infty, n \rightarrow \infty.$$

Then, by (18), (19), (21),

$$\begin{aligned}
\frac{1}{n}\bar{U}'_{-1}Q_{\bar{X}_q}\Delta\bar{U} &= \frac{d(1)}{n}\sum_t\bar{u}_{t-1}\left(\frac{\gamma_t}{d(1)}+\frac{1}{n}\Delta\xi'_{(2)t}\eta_b+n^{\frac{1}{2}}\Delta c'_tD^{-1}\eta_c\right)+\frac{1}{n}\sum_t\bar{u}_{t-1}(\hat{\gamma}_t-\gamma_t) \\
&+\frac{1}{n^2}\sum_t\bar{u}_{t-1}\Delta\xi'_{(2)t}\eta_b\left(\left(1-e'_q\left(\sum_t a_{q,t}a'_{q,t}\right)^{-1}\sum_t a_{q,t}\zeta'_t\eta_a\right)-d(1)\right), \\
&+\frac{1}{n^{\frac{1}{2}}}\sum_t\bar{u}_{t-1}\Delta c'_tD^{-1}\eta_c\left(\left(1-e'_q\left(\sum_t a_{q,t}a'_{q,t}\right)^{-1}\sum_t a_{q,t}\zeta'_t\eta_a\right)-d(1)\right) \\
&+o_p(1) \tag{22}
\end{aligned}$$

where  $\hat{\gamma}_t = \zeta'_t\eta_a - a'_{q,t}\left(\sum_t a_{q,t}a'_{q,t}\right)^{-1}\sum_t a_{q,t}\zeta'_t\eta_a$ . The second, third and fourth terms on the right side of (22) can be easily shown to be  $o_p(1)$ , whereas the first one can be analyzed by identical transformations to those employed in the proof of (16), so that

$$\frac{1}{n}\bar{U}'_{-1}Q_{\bar{X}_q}\Delta\bar{U} \Rightarrow d(1)l_{11}^2\int_0^1 Q(s)dQ(s).$$

Also, by previous arguments

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n}\Delta\bar{U}'Q_{\bar{X}_q}\Delta\bar{U}+o_p(1) \\
&= \frac{1}{n}\sum_t\eta'_a\zeta_t\zeta'_t\eta_a-\frac{1}{n}\sum_t\eta'_a\zeta_t a'_{q,t}\left(\sum_t a_{q,t}a'_{q,t}\right)^{-1}\sum_t\eta'_a\zeta_t a_{q,t}+o_p(1),
\end{aligned}$$

so that

$$\hat{\sigma}^2 \rightarrow_p d^2(1)\eta'_a\Omega\eta_a.$$

By (15),

$$\begin{aligned}
\eta'_a\Omega\eta_a &= \left(1,-\int_0^1 Z_1(s)Z'_{-1}(s)ds\left(\int_0^1 Z_{-1}(s)Z'_{-1}(s)ds\right)^{-1}\right)\begin{pmatrix}\Phi & 0 & 0 \\ 0 & \Omega_{22} & 0 \\ 0 & 0 & 0\end{pmatrix} \\
&\times\begin{pmatrix}1 \\ -\left(\int_0^1 Z_{-1}(s)Z'_{-1}(s)ds\right)^{-1}\int_0^1 Z_1(s)Z_{-1}(s)ds\end{pmatrix},
\end{aligned}$$

so

$$\eta_a' \Omega \eta_a = l_{11}^2 \kappa' \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix} \kappa,$$

by identical transformations to the ones employed before, to conclude the proof of (10).

Finally, we show (9). Clearly

$$t_n(\hat{\eta}) - t_n(\eta) = t_n(\hat{\eta}) - t(\hat{\eta}) - (t_n(\eta) - t(\eta)) + t(\hat{\eta}) - t(\eta),$$

where  $t(\cdot)$  is like  $t_n(\cdot)$ , but with the normalized summations replaced by the respective limits in distribution. First,  $t(\hat{\eta}) - t(\eta) = o_p(1)$ , by (8) and the continuous mapping theorem. Also, noting that  $\hat{\eta} = O_p(1)$ ,  $t_n(\hat{\eta}) - t(\hat{\eta}) = o_p(1)$  by tedious but simple calculations, showing that the difference between the individual components of  $t_n(\hat{\eta})$  with the corresponding ones in  $t(\hat{\eta})$  is  $o_p(1)$ . For identical reasons,  $t_n(\eta) - t(\eta) = o_p(1)$ , to conclude (9), and therefore complete the proof of Theorem 1.

**Proof of Theorem 2.** Let  $\bar{z}_t$  be an arbitrary  $(p-r)$ -dimensional subvector of  $z_t$  with  $I(2)$  and noncointegrated components (given that  $z_t$  is cointegrated with rank  $r$ , at least one of such subvectors exists). Collect the rest of components of  $z_t$  in the  $r \times 1$  vector  $\underline{z}_t$ . Then there exists a  $r \times (p-r)$  matrix  $A$  such that  $\underline{z}_t + A\bar{z}_t$  has a smaller integration order than 2. Without loss of generality set  $z_t = (\underline{z}_t', \bar{z}_t')'$ . If  $\underline{z}_t + A\bar{z}_t$  is stationary, the theorem holds trivially because  $R = T = C$ . The proof for the  $\underline{z}_t + A\bar{z}_t \sim I(1)$  situation is as follows. Let  $\phi \in T$ . Then  $\phi \in R$ , by setting  $\lambda(\phi) = (0_r', \rho'(\phi))'$ , where  $0_r$  denotes a  $r$ -dimensional vector of zeroes. Alternatively, if  $\phi \in R$ , there exists  $\lambda(\phi)$  such that  $\phi' z_t + \lambda'(\phi) \Delta z_t$  is stationary or, equivalently,

$$\phi' z_t + \underline{\lambda}'(\phi) \Delta \underline{z}_t + \bar{\lambda}'(\phi) \Delta \bar{z}_t \tag{23}$$

is stationary, where  $\lambda(\phi) = (\underline{\lambda}'(\phi), \bar{\lambda}'(\phi))'$  is partitioned according to  $z_t$ . From the cointegrating relations

$$\underline{\lambda}'(\phi) \Delta \underline{z}_t + \underline{\lambda}'(\phi) A \Delta \bar{z}_t$$

is stationary. Then adding and subtracting  $\underline{\lambda}'(\phi) A \Delta \bar{z}_t$  to (23), necessarily

$$\phi' z_t + \left( \bar{\lambda}'(\phi) - \underline{\lambda}'(\phi) A \right) \Delta \bar{z}_t$$

is stationary and, consequently,  $\phi \in T$ , to conclude the proof.

## REFERENCES

Bae, Y. and R. Dejong, 2007, Money demand function estimation by nonlinear

- cointegration. *Journal of Applied Econometrics* 22, 767-793.
- Bacchiocchi, E. and L. Fanelli, 2005, Testing the purchasing power parity through  $I(2)$  cointegration techniques. *Journal of Applied Econometrics* 20, 749-770.
- Banerjee, A., Cockerell, L. and B. Russell, 2001, An  $I(2)$  analysis of inflation and the markup. *Journal of Applied Econometrics* 16, 221-240.
- Banerjee, A. and P. Mizen, 2006, A re-interpretation of the linear quadratic model when inventories and sales are polynomially cointegrated. *Journal of Applied Econometrics* 21, 1249-1264.
- Berk, K.N., 1974, Consistent autoregressive spectral estimates. *Annals of Statistics* 2, 489-502.
- Caner, M. and L. Kilian, 2001, Size distortions of tests of the null hypothesis of stationarity: evidence and implications for the PPP debate. *Journal of International Money and Finance* 20, 639-657.
- Chang, Y. and J.Y. Park, 2002, On the asymptotics of ADF tests for unit roots. *Econometric Reviews* 21, 431-447.
- Engle, R.F. and C.W.J. Granger, 1987, Cointegration and error correction model. Representation, estimation and testing. *Econometrica* 55, 251-276.
- Engsted, T., Gonzalo, J. and N. Haldrup, 1997, Testing for multicointegration. *Economics Letters* 56, 259-266.
- Engsted, T. and N. Haldrup, 1999, Multicointegration in stock-flows models. *Oxford Bulletin of Economics and Statistics* 61, 237-254.
- Engsted, T. and S. Johansen, 1999, Granger's representation theorem and multicointegration, in Engle, R.F. and H. White (Eds.), *Cointegration, Causality, and Forecasting. A Festschrift in Honor of Clive W.J. Granger*. Oxford University Press, pp.200-211.
- Gomez-Biscarri, J. and J. Hualde, 2014, Regression-based analysis of cointegrated systems. *Barcelona GSE Working Paper No. 780*.
- Gonzalo, J. and T.H. Lee, 1998, Pitfalls in testing for long run relationships. *Journal of Econometrics* 86, 129-154.

- Granger, C.W.J. and T.H. Lee, 1989, Investigation of production, sales and inventory relations using multicointegration and non-symmetric error correction models. *Journal of Applied Econometrics* 4, S145-S159.
- Granger, C.W.J. and T.H. Lee, 1990, Multicointegration, in Rhodes, G.F., Fomby, T.B. (Eds.), *Advances in Econometrics*, vol. 8. JAI Press, Greenwich, CT, pp. 71-84.
- Gregoir, S. and G. Laroque, 1994, Polynomial cointegration estimation and test. *Journal of Econometrics* 63, 183-214.
- Haldrup, N., 1994, The asymptotics of single-equation cointegration regressions with  $I(1)$  and  $I(2)$  variables. *Journal of Econometrics* 63, 153-181.
- Haldrup, N., 1998, An econometric analysis of  $I(2)$  variables. *Journal of Economic Surveys* 12, 595-650.
- Haldrup, N. and P. Lildholdt, 2002, On the robustness of unit root tests in the presence of double unit roots. *Journal of Time Series Analysis* 23, 155-171.
- Johansen, S., 1988, Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12, 231-254.
- Johansen, S., 1995a, A statistical analysis of cointegration for  $I(2)$  variables. *Econometric Theory* 11, 25-29.
- Johansen, S., 1995b, Likelihood based inference in cointegrated vector autoregressive models. Oxford University Press.
- Johansen, S., Juselius, K., Frydman, R. and M. Goldberg, 2010, Testing hypotheses in an  $I(2)$  model with piecewise linear trends. An analysis of the persistent long swings in the Dmk/\$ rate. *Journal of Econometrics* 158, 117-129.
- Juselius, K., 1995, Do purchasing power parity and uncovered interest rate parity hold in the long-run? An example of likelihood inference in a multivariate time-series model. *Journal of Econometrics*, 69, 211-40.
- Kia, A., 2008, Fiscal sustainability in emerging countries: Evidence from Iran and Turkey. *Journal of Policy Modelling* 30, 957-972.
- Kongsted, H.C., 2003, An  $I(2)$  cointegration analysis of small-country import price determination. *Econometrics Journal* 6, 53-71.

- Kongsted, H.C., 2005, Testing the nominal-to-real transformation. *Journal of Econometrics* 124, 205-225.
- Kongsted, H.C. and H.B. Nielsen, 2004, Analysing  $I(2)$  systems by transformed vector autoregressions. *Oxford Bulletin of Economics and Statistics* 66, 379-397.
- Kugler, P. and S. Kaufmann, 2008, Does money matter for inflation in the Euro area? *Contemporary Economic Policy* 26, 590-606.
- Leachman, L.L., 1996, New evidence on the Ricardian equivalence theorem: a multicointegration approach. *Applied Economics* 28, 695-704.
- Leachman, L.L., Bester, A., Rosas, G. and P. Lange, 2005, Multicointegration and sustainability of fiscal practices. *Economic Inquiry* 43, 454-466.
- Leachman, L.L. and B.B. Francis, 2000, Multicointegration analysis of the sustainability of foreign debt. *Journal of Macroeconomics* 22, 207-227.
- Leachman, L.L. and B.B. Francis, 2002, Twin deficits: apparition or reality? *Applied Economics* 34, 1121-1132.
- Lee, T.H., 1992, Stock-flow relationships in housing construction. *Oxford Bulletin of Economics and Statistics* 54, 419-430.
- Mark, N.C. and D. Sul, 2001, Nominal exchange rates and monetary fundamentals: evidence from a small post-Bretton Woods panel. *Journal of International Economics* 53, 29-52.
- Nielsen, H.B. and A. Rahbek, 2007, The likelihood ratio test for cointegration ranks in the  $I(2)$  model. *Econometric Theory* 23, 615-637.
- Paruolo, P., 1996, On the determination of integration indices in  $I(2)$  systems. *Journal of Econometrics* 72, 313-356.
- Pedroni, P., 2004, Panel cointegration, asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory* 20, 597-625.
- Phillips P.C.B. and S.N. Durlauf, 1986, Multiple time series regression with integrated processes. *Review of Economic Studies* 53, 473-496.

- Phillips, P.C.B. and S. Ouliaris, 1990, Asymptotic properties of residual based tests for cointegration. *Econometrica* 58, 165-193.
- Rahbek, A., Kongsted, H.C. and C. Jørgensen, 1999, Trend stationarity in the  $I(2)$  cointegration model. *Journal of Econometrics* 90, 265-289.
- Rapach, D.E. and M.E. Wohar, 2002, Testing the monetary model of exchange rate determination: new evidence from a century of data. *Journal of International Economics* 58, 359-385.
- Rogoff, K., 1996, The purchasing power parity puzzle. *Journal of Economic Literature* 34, 647-668.
- Rootzén, H., 1976, Fluctuations of sequences which converge in distribution. *Annals of Probability* 4, 456-463.
- Rossi, B., 2006, Are exchange rates really random walks? Some evidence robust to parameter instability. *Macroeconomic Dynamics* 10, 20-38.
- Siliverstovs, B., 2006, Multicointegration in US consumption data. *Applied Economics* 38, 819-833.
- Stock J.H. and M.W. Watson, 1993, A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783-820.

Table 1. Critical values for the cointegration ADF test (intercept included in the cointegration regression)

$r$	$p$ $n \setminus \alpha$	2			3			4			5			6							
		.10	.05	.025	.01	.10	.05	.025	.01	.10	.05	.025	.01	.10	.05	.025	.01				
1	50	-3.64	-3.98	-4.29	-4.66	-4.44	-4.78	-5.10	-5.49	-5.14	-5.51	-5.84	-6.26	-5.79	-6.17	-6.53	-6.96	-6.40	-6.81	-7.18	-7.62
	100	-3.56	-3.88	-4.15	-4.49	-4.30	-4.62	-4.91	-5.24	-4.93	-5.25	-5.54	-5.88	-5.50	-5.83	-6.13	-6.49	-6.02	-6.35	-6.66	-7.04
	250	-3.52	-3.82	-4.08	-4.38	-4.21	-4.51	-4.78	-5.09	-4.81	-5.10	-5.36	-5.68	-5.33	-5.62	-5.89	-6.20	-5.80	-6.10	-6.36	-6.67
	500	-3.50	-3.79	-4.04	-4.34	-4.18	-4.48	-4.73	-5.03	-4.77	-5.06	-5.31	-5.61	-5.27	-5.56	-5.81	-6.11	-5.73	-6.02	-6.27	-6.57
	50,000	-3.49	-3.77	-4.02	-4.31	-4.16	-4.44	-4.68	-4.97	-4.73	-5.00	-5.25	-5.52	-5.21	-5.49	-5.75	-6.04	-5.66	-5.94	-6.18	-6.44
2	50					-4.03	-4.37	-4.68	-5.07	-4.77	-5.13	-5.46	-5.84	-5.43	-5.82	-6.17	-6.60	-6.06	-6.46	-6.82	-7.27
	100					-3.93	-4.24	-4.52	-4.86	-4.60	-4.92	-5.20	-5.54	-5.20	-5.53	-5.81	-6.16	-5.74	-6.07	-6.37	-6.74
	250					-3.87	-4.17	-4.44	-4.74	-4.51	-4.81	-5.07	-5.37	-5.06	-5.36	-5.62	-5.93	-5.56	-5.86	-6.12	-6.42
	500					-3.85	-4.14	-4.40	-4.70	-4.48	-4.76	-5.02	-5.32	-5.01	-5.30	-5.55	-5.85	-5.50	-5.79	-6.04	-6.33
	50,000					-3.83	-4.11	-4.36	-4.65	-4.45	-4.73	-4.97	-5.26	-4.98	-5.24	-5.48	-5.76	-5.44	-5.72	-5.97	-6.28
3	50									-4.38	-4.73	-5.06	-5.46	-5.08	-5.45	-5.79	-6.21	-5.72	-6.11	-6.46	-6.91
	100									-4.27	-4.58	-4.87	-5.20	-4.90	-5.22	-5.51	-5.85	-5.47	-5.80	-6.10	-6.45
	250									-4.19	-4.49	-4.76	-5.07	-4.79	-5.08	-5.35	-5.66	-5.30	-5.60	-5.87	-6.19
	500									-4.17	-4.46	-4.72	-5.01	-4.75	-5.04	-5.29	-5.60	-5.26	-5.54	-5.80	-6.10
	50,000									-4.16	-4.44	-4.68	-4.98	-4.71	-4.99	-5.24	-5.53	-5.21	-5.48	-5.73	-6.00
4	50													-4.71	-5.07	-5.40	-5.81	-5.38	-5.76	-6.10	-6.52
	100													-4.58	-4.90	-5.19	-5.53	-5.17	-5.49	-5.78	-6.12
	250													-4.49	-4.78	-5.05	-5.36	-5.04	-5.34	-5.59	-5.91
	500													-4.46	-4.75	-5.00	-5.31	-5.01	-5.30	-5.55	-5.86
	50,000													-4.43	-4.71	-4.96	-5.25	-4.96	-5.24	-5.49	-5.78
5	50																	-5.03	-5.40	-5.74	-6.15
	100																	-4.87	-5.19	-5.48	-5.82
	250																	-4.77	-5.07	-5.33	-5.64
	500																	-4.73	-5.02	-5.28	-5.57
	50,000																	-4.70	-4.99	-5.23	-5.52

The simulations were based upon 200,000 replications.  $p$  and  $r$  denote the total number of variables in the system and the cointegrating rank, respectively.

Table 2. Critical values for the cointegration ADF test (intercept and linear trend included in the cointegration regression)

$r$	$p$ $n \setminus \alpha$	2			3			4			5			6							
		.10	.05	.025	.01	.10	.05	.025	.01	.10	.05	.025	.01	.10	.05	.025	.01				
1	50	-4.08	-4.42	-4.74	-5.11	-4.82	-5.18	-5.51	-5.91	-5.49	-5.86	-6.21	-6.64	-6.11	-6.52	-6.88	-7.33	-6.72	-7.14	-7.52	-8.00
	100	-3.97	-4.28	-4.56	-4.89	-4.64	-4.96	-5.25	-5.60	-5.23	-5.56	-5.85	-6.20	-5.77	-6.11	-6.41	-6.78	-6.27	-6.61	-6.92	-7.29
	250	-3.90	-4.19	-4.45	-4.77	-4.53	-4.83	-5.09	-5.40	-5.09	-5.38	-5.64	-5.95	-5.57	-5.86	-6.12	-6.44	-6.02	-6.32	-6.60	-6.92
	500	-3.88	-4.16	-4.41	-4.71	-4.50	-4.79	-5.04	-5.34	-5.03	-5.32	-5.56	-5.87	-5.50	-5.80	-6.04	-6.34	-5.95	-6.24	-6.49	-6.80
	50,000	-3.86	-4.14	-4.39	-4.70	-4.46	-4.74	-5.00	-5.29	-4.99	-5.27	-5.51	-5.79	-5.44	-5.72	-5.99	-6.27	-5.88	-6.16	-6.39	-6.68
2	50					-4.42	-4.78	-5.10	-5.49	-5.13	-5.50	-5.85	-6.25	-5.76	-6.15	-6.51	-6.95	-6.37	-6.78	-7.16	-7.63
	100					-4.30	-4.61	-4.90	-5.24	-4.92	-5.24	-5.54	-5.88	-5.49	-5.82	-6.12	-6.47	-6.00	-6.34	-6.65	-7.03
	250					-4.21	-4.51	-4.77	-5.08	-4.80	-5.09	-5.36	-5.67	-5.62	-5.62	-5.88	-6.19	-5.79	-6.09	-6.09	-6.67
	500					-4.19	-4.47	-4.72	-5.02	-4.76	-5.05	-5.30	-5.60	-5.32	-5.56	-5.81	-6.11	-5.72	6.00	-6.26	-6.56
	50,000					-4.16	-4.45	-4.70	-4.99	-4.73	-5.01	-5.25	-5.54	-5.21	-5.49	-5.72	-6.01	-5.67	-5.93	-6.18	-6.46
3	50									-4.76	-5.11	-5.44	-5.85	-5.41	-5.79	-6.13	-6.56	-6.04	-6.44	-6.80	-7.26
	100									-4.60	-4.91	-5.20	-5.55	-5.19	-5.52	-5.81	-6.15	-5.73	-6.06	-6.35	-6.72
	250									-4.51	-4.81	-5.07	-5.38	-5.06	-5.36	-5.62	-5.92	-5.55	-5.84	-6.11	-6.43
	500									-4.47	-4.77	-5.02	-5.32	-5.02	-5.31	-5.57	-5.87	-5.49	-5.78	-6.04	-6.34
	50,000									-4.45	-4.74	-4.98	-5.27	-4.97	-5.26	-5.50	-5.78	-5.43	-5.71	-5.95	-6.22
4	50													-5.06	-5.43	-5.77	-6.19	-5.70	-6.09	-6.44	-6.88
	100													-4.88	-5.20	-5.49	-5.84	-5.45	-5.78	-6.08	-6.44
	250													-4.78	-5.08	-5.34	-5.65	-5.30	-5.60	-5.87	-6.18
	500													-4.75	-5.03	-5.29	-5.59	-5.25	-5.54	-5.79	-6.09
	50,000													-4.71	-4.99	-5.24	-5.52	-5.21	-5.49	-5.73	-6.02
5	50																	-5.37	-5.75	-6.09	-6.50
	100																	-5.17	-5.49	-5.79	-6.14
	250																	-5.04	-5.33	-5.60	-5.91
	500																	-5.00	-5.28	-5.54	-5.82
	50,000																	-4.97	-5.24	-5.48	-5.75

The simulations were based upon 200,000 replications.  $p$  and  $r$  denote the total number of variables in the system and the cointegrating rank, respectively.

Table 3. Critical values for the cointegration ADF test (intercept, linear and quadratic trend included in the cointegration regression)

$r$	$p$ $n \setminus \alpha$	2			3			4			5			6							
		.10	.05	.025	.01	.10	.05	.025	.01	.10	.05	.025	.01	.10	.05	.025	.01				
1	50	-4.50	-4.86	-5.18	-5.57	-5.20	-5.57	-6.33	-5.84	-6.24	-6.60	-7.05	-6.46	-6.87	-7.26	-7.73	-7.05	-7.49	-7.89	-8.37	
	100	-4.34	-4.66	-4.94	-5.29	-4.97	-5.29	-5.91	-5.53	-5.86	-6.15	-6.51	-6.04	-6.39	-6.70	-7.07	-6.53	-6.88	-7.19	-7.59	
	250	-4.25	-4.54	-4.81	-5.11	-4.83	-5.13	-5.70	-5.35	-5.64	-5.91	-6.22	-5.81	-6.11	-6.38	-6.70	-6.25	-6.55	-6.82	-7.14	
	500	-4.22	-4.51	-4.76	-5.06	-4.79	-5.07	-5.62	-5.29	-5.58	-5.83	-6.12	-5.74	-6.02	-6.27	-6.56	-6.16	-6.45	-6.70	-7.01	
	50,000	-4.20	-4.47	-4.70	-5.00	-4.74	-5.02	-5.57	-5.23	-5.52	-5.78	-6.08	-5.66	-5.96	-6.20	-6.47	-6.09	-6.36	-6.59	-6.86	
2	50					-4.82	-5.18	-5.50	-5.91	-5.49	-5.87	-6.22	-6.05	-6.10	-6.51	-6.87	-7.33	-6.71	-7.14	-7.53	-8.01
	100					-4.64	-4.96	-5.24	-5.59	-5.23	-5.56	-5.85	-6.21	-5.77	-6.10	-6.40	-6.76	-6.27	-6.61	-6.92	-7.28
	250					-4.54	-4.83	-5.09	-5.39	-5.08	-5.38	-5.64	-5.96	-5.57	-5.87	-6.13	-6.44	-6.02	-6.32	-6.59	-6.91
	500					-4.50	-4.78	-5.04	-5.34	-5.03	-5.32	-5.57	-5.87	-5.52	-5.81	-6.05	-6.35	-5.94	-6.23	-6.48	-6.77
	50,000					-4.47	-4.73	-4.99	-5.25	-4.99	-5.27	-5.51	-5.81	-5.45	-5.72	-5.95	-6.23	-5.87	-6.15	-6.39	-6.69
3	50									-5.12	-5.48	-5.82	-6.23	-5.76	-6.15	-6.50	-6.94	-6.38	-6.79	-7.16	-7.62
	100									-4.92	-5.24	-5.53	-5.89	-5.49	-5.82	-6.12	-6.47	-6.00	-6.34	-6.64	-7.00
	250									-4.81	-5.11	-5.36	-5.68	-5.33	-5.62	-5.89	-6.19	-5.79	-6.09	-6.35	-6.67
	500									-4.77	-5.05	-5.31	-5.60	-5.27	-5.56	-5.81	-6.11	-5.73	-6.02	-6.27	-6.57
	50,000									-4.74	-5.01	-5.23	-5.51	-5.23	-5.50	-5.75	-6.04	-5.66	-5.93	-6.16	-6.46
4	50													-5.42	-5.80	-6.16	-6.58	-6.03	-6.43	-6.79	-7.23
	100													-5.19	-5.51	-5.80	-6.15	-5.74	-6.07	-6.37	-6.73
	250													-5.06	-5.36	-5.62	-5.93	-5.55	-5.85	-6.12	-6.43
	500													-5.01	-5.30	-5.55	-5.86	-5.50	-5.79	-6.04	-6.34
	50,000													-4.98	-5.28	-5.51	-5.77	-5.44	-5.71	-5.94	-6.19
5	50																	-5.69	-6.09	-6.44	-6.87
	100																	-5.45	-5.78	-6.07	-6.44
	250																	-5.30	-5.60	-5.86	-6.17
	500																	-5.25	-5.54	-5.80	-6.09
	50,000																	-5.20	-5.47	-5.72	-5.97

The simulations were based upon 200,000 replications.  $p$  and  $r$  denote the total number of variables in the system and the cointegrating rank, respectively.

Table 4. Proportion of rejection of the null of no stationary cointegration for the different models (infeasible tests)

$\zeta_t$	Model \ $\alpha$	$n = 100$			$n = 200$			$n = 500$			$n = 1000$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
WN	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	.993	.990	.983	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	6	.993	.988	.981	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	7	.173	.113	.037	.116	.075	.021	.094	.052	.012	.086	.044	.009
	8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	10	.992	.988	.982	1.00	1.00	.999	1.00	1.00	1.00	1.00	1.00	1.00
	11	.992	.987	.981	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	12	.199	.125	.046	.136	.082	.026	.109	.057	.013	.100	.051	.010
	13	.992	.987	.979	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	14	.992	.989	.978	1.00	1.00	.999	1.00	1.00	1.00	1.00	1.00	1.00
	15	.209	.127	.045	.141	.082	.023	.111	.055	.014	.103	.054	.011
AR	1	.711	.559	.274	.999	.991	.898	1.00	1.00	1.00	1.00	1.00	1.00
	2	.711	.561	.267	.998	.991	.896	1.00	1.00	1.00	1.00	1.00	1.00
	3	.617	.479	.221	.990	.967	.799	1.00	1.00	1.00	1.00	1.00	1.00
	4	.603	.467	.221	.989	.962	.792	1.00	1.00	1.00	1.00	1.00	1.00
	5	.436	.329	.147	.916	.831	.577	1.00	1.00	1.00	1.00	1.00	1.00
	6	.444	.324	.145	.912	.829	.564	1.00	1.00	1.00	1.00	1.00	1.00
	7	.180	.113	.039	.123	.067	.019	.096	.052	.011	.085	.047	.010
	8	.678	.538	.273	.989	.961	.792	1.00	1.00	1.00	1.00	1.00	1.00
	9	.673	.524	.262	.990	.959	.778	1.00	1.00	1.00	1.00	1.00	1.00
	10	.503	.372	.181	.910	.820	.557	1.00	1.00	1.00	1.00	1.00	1.00
	11	.495	.376	.182	.908	.821	.550	1.00	1.00	1.00	1.00	1.00	1.00
	12	.216	.142	.048	.141	.082	.023	.111	.061	.016	.098	.054	.011
	13	.600	.465	.238	.942	.865	.606	1.00	1.00	1.00	1.00	1.00	1.00
	14	.595	.463	.237	.937	.865	.599	1.00	1.00	1.00	1.00	1.00	1.00
	15	.241	.156	.056	.156	.098	.025	.118	.062	.017	.113	.053	.010

10,000 replications were carried out for each sample size  $n$ . Three different significance levels  $\alpha = \{.10, .05, .01\}$  were used in the tests. The number of lags in the ADF tests is chosen according to the BIC. The innovation vector  $\zeta_t$  is generated as  $\zeta_{jt} = \varepsilon_{jt}$ ,  $j = 1, \dots, 4$  (WN) and  $\zeta_{jt} = 0.8\zeta_{j,t-1} + \varepsilon_{jt}$ ,  $j = 1, \dots, 4$  (AR) with Gaussian  $\varepsilon_t$  such that  $E(\varepsilon_t) = 0$ ,  $Var(\varepsilon_t) = I_4$ . Critical values from Table 1,  $p = 4$ ,  $r = 3$  (Models 1-7),  $p = 4$ ,  $r = 2$  (models 8-12),  $p = 4$ ,  $r = 1$  (models 13-15). Models 7, 12 and 15 are the models with no stationary cointegration.

Table 5. Proportion of rejection of the null of no stationary cointegration for the different models (tests of Paruolo, 1996)

$\zeta_t$	Model \ $\alpha$	$n = 100$			$n = 200$			$n = 500$			$n = 1000$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
WN	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	.990	.973	.889	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	6	.999	.998	.983	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	7	.182	.108	.033	.149	.078	.020	.124	.064	.019	.114	.066	.017
	8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	10	.982	.947	.809	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	.997	.977	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	12	.183	.109	.028	.155	.083	.021	.135	.073	.016	.136	.069	.016
	13	.958	.910	.717	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	14	.999	.997	.971	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	15	.187	.099	.025	.161	.081	.020	.144	.076	.020	.130	.071	.019
AR	1	.679	.531	.279	.993	.982	.915	1.00	1.00	1.00	1.00	1.00	1.00
	2	.704	.558	.312	.997	.987	.932	1.00	1.00	1.00	1.00	1.00	1.00
	3	.595	.446	.200	.931	.872	.660	1.00	1.00	1.00	1.00	1.00	1.00
	4	.619	.472	.226	.951	.885	.698	1.00	1.00	1.00	1.00	1.00	1.00
	5	.513	.369	.148	.714	.579	.323	.997	.991	.946	1.00	1.00	1.00
	6	.551	.390	.168	.748	.616	.359	.998	.995	.963	1.00	1.00	1.00
	7	.445	.305	.108	.409	.289	.115	.219	.131	.043	.160	.090	.027
	8	.674	.521	.279	.903	.828	.574	1.00	1.00	1.00	1.00	1.00	1.00
	9	.698	.560	.307	.920	.848	.621	1.00	1.00	1.00	1.00	1.00	1.00
	10	.625	.490	.224	.709	.570	.299	.994	.979	.897	1.00	1.00	1.00
	11	.669	.520	.258	.753	.617	.329	.997	.991	.938	1.00	1.00	1.00
	12	.605	.449	.202	.452	.314	.115	.241	.147	.047	.179	.104	.033
	13	.773	.645	.377	.723	.587	.316	.987	.966	.848	1.00	1.00	1.00
	14	.795	.673	.408	.759	.634	.356	.995	.982	.909	1.00	1.00	1.00
	15	.772	.644	.358	.513	.367	.159	.259	.166	.050	.188	.115	.026

10,000 replications were carried out for each sample size  $n$ . Three different significance levels  $\alpha = \{.10, .05, .01\}$  were used in the tests. The number of lags in the VAR is set to the true value of zero (WN) and one (AR). The innovation vector  $\zeta_t$  is generated as  $\zeta_{jt} = \varepsilon_{jt}$ ,  $j = 1, \dots, 4$  (WN) and  $\zeta_{jt} = 0.8\zeta_{j,t-1} + \varepsilon_{jt}$ ,  $j = 1, \dots, 4$  (AR) with Gaussian  $\varepsilon_t$  such that  $E(\varepsilon_t) = 0$ ,  $Var(\varepsilon_t) = I_4$ . Models 7, 12 and 15 are the models with no stationary cointegration.

Table 6. Proportion of rejection of the null of no stationary cointegration for the different models (feasible tests)

$\zeta_t$	$Model \setminus \alpha$	$n = 100$			$n = 200$			$n = 500$			$n = 1000$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
WN	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	3	1.00	.999	.997	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	.994	.993	.990	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	.957	.942	.920	.984	.980	.970	1.00	.999	.997	1.00	1.00	1.00
	6	.890	.872	.827	.904	.893	.895	.937	.953	.958	.935	.968	.992
	7	.190	.120	.037	.158	.092	.028	.122	.064	.017	.103	.056	.012
	8	.999	.999	.997	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	9	.995	.993	.987	1.00	1.00	.999	1.00	1.00	1.00	1.00	1.00	1.00
	10	.934	.954	.959	.990	.988	.979	1.00	1.00	.999	1.00	1.00	1.00
	11	.876	.859	.785	.866	.867	.842	.934	.940	.911	.957	.978	.985
	12	.184	.117	.043	.144	.092	.028	.116	.067	.018	.104	.057	.014
	13	.981	.971	.947	.999	.998	.995	1.00	1.00	1.00	1.00	1.00	1.00
	14	.899	.902	.868	.898	.906	.889	.939	.949	.945	.944	.975	.989
	15	.195	.127	.039	.153	.091	.029	.128	.074	.020	.115	.063	.017
AR	1	.688	.534	.257	.997	.984	.864	1.00	1.00	1.00	1.00	1.00	1.00
	2	.823	.696	.401	.999	.995	.950	1.00	1.00	1.00	1.00	1.00	1.00
	3	.597	.461	.212	.986	.951	.762	1.00	1.00	1.00	1.00	1.00	1.00
	4	.679	.550	.268	.988	.966	.852	1.00	1.00	1.00	1.00	1.00	1.00
	5	.436	.318	.139	.908	.821	.553	1.00	.999	.997	1.00	1.00	1.00
	6	.369	.255	.097	.836	.759	.506	.957	.980	.996	.950	.973	.996
	7	.161	.085	.015	.118	.062	.018	.092	.046	.009	.080	.040	.009
	8	.616	.485	.247	.976	.939	.729	1.00	1.00	1.00	1.00	1.00	1.00
	9	.694	.559	.271	.991	.974	.875	1.00	1.00	1.00	1.00	1.00	1.00
	10	.479	.352	.167	.899	.801	.533	1.00	1.00	.999	1.00	1.00	1.00
	11	.351	.235	.087	.822	.716	.451	.974	.988	.997	.970	.986	.998
	12	.167	.086	.017	.127	.075	.020	.094	.051	.014	.081	.043	.010
	13	.582	.448	.229	.930	.847	.593	1.00	1.00	1.00	1.00	1.00	1.00
	14	.326	.211	.054	.840	.728	.416	.964	.985	.997	.964	.983	.998
	15	.174	.094	.023	.159	.101	.028	.109	.060	.018	.102	.048	.012

10,000 replications were carried out for each sample size  $n$ . Three different significance levels  $\alpha = \{.10, .05, .01\}$  were used in the tests. The number of lags in the ADF tests is chosen according to the BIC. The innovation vector  $\zeta_t$  is generated as  $\zeta_{jt} = \varepsilon_{jt}$ ,  $j = 1, \dots, 4$  (WN) and  $\zeta_{jt} = 0.8\zeta_{j,t-1} + \varepsilon_{jt}$ ,  $j = 1, \dots, 4$  (AR) with Gaussian  $\varepsilon_t$  such that  $E(\varepsilon_t) = 0$ ,  $Var(\varepsilon_t) = I_4$ . Critical values from Table 1,  $p = 4$ ,  $r = 3$  (Models 1-7),  $p = 4$ ,  $r = 2$  (models 8-12),  $p = 4$ ,  $r = 1$  (models 13-15). Models 7, 12 and 15 are the models with no stationary cointegration.