A Single-Equation Cointegration Estimator Robust to Variance Breaks

Nikolaos Kourogenis* Ekaterini Panopoulou†‡ Nikitas Pittis§

February 2007

Abstract

This paper investigates the performance of standard estimators for estimating a single cointegrating vector in the context of a triangular system that exhibits a single break in the elements of the covariance matrix of the errors. It is shown that both the limiting distribution of OLS and that of the associated t-statistic are significantly affected by the presence of such shifts. Based on this result, we propose a new OLS-based estimator which corrects the ordinary one, not only for endogeneity but also for possible variance shifts. A small Monte Carlo study suggests that the new procedure performs significantly better than the standard cointegration methods, especially when the variances of the errors jump upwards towards the end of the sample.

JEL Classification: C22

Keywords: cointegration, variance breaks, limiting distributions, t-statistics.

Acknowledgments: We are grateful to K. Abadir, A. Antzoulatos, G. Hardouvelis, D. Malliaropulos and seminar participants at various seminars and conferences for helpful suggestions and comments. The usual disclaimer applies.

*Department of Banking and Financial Management, University of Piraeus.

†Department of Statistics and Insurance Science, University of Piraeus & IIIS, Trinity College Dublin.

‡*Correspondence to:* Ekaterini Panopoulou Department of Statistics and Insurance Science, University of Piraeus, Greece. Email: apano@unipi.gr. Tel: 0030 210 4142728. Fax: 0030 210 4142340.

§Department of Banking and Financial Management, University of Piraeus.

1 Introduction

Some recent empirical studies have shown that the volatility of the innovations that drive many macroeconomic and financial series has experienced a sharp decline since the mid 1980s (see, for example, Kim and Nelson 1999, McConnell and Perez-Quiros 2000, Stock and Watson 2002, Sensier and van Dijk 2004). This phenomenon is particularly profound for the case of the U.S. economy and it is usually referred to as the "Great Moderation". This empirical evidence has motivated extensive theoretical research that investigates whether variance breaks affect standard methods for analyzing univariate time series. More specifically, a great deal of research efforts have been devoted to study the effects of such breaks on the usual unit root and stationarity tests (see, for example, Hamori and Tokihisa 1997, Kim, Leybourne and Newbold 2002, Burridge and Taylor 2001, Busetti and Taylor 2003, Cavaliere 2004a, 2004b, Cavaliere and Taylor 2005). The results of these studies have demonstrated that the presence of a permanent shift in the error variance, not accounted for by the testing procedure, seriously affects inferences on the time-heterogeneity properties of the relevant series.

More recently, Cavaliere and Taylor (2006) analyze the issue of shifts in the error variance in a multivariate framework. In particular, they assume a multivariate cointegrated time series model and analyze the effects of several types of variance breaks that occur either in the common stochastic trends or in the cointegrating relations on the tests for the null hypothesis of cointegration. They show that such breaks affect the limiting distributions of the relevant test statistics under both the null hypothesis of cointegration and the alternative of no cointegration, thus affecting both the size and the power properties of these tests.

Despite the problems mentioned above, an applied researcher is still likely to make the correct decision concerning the cointegration properties of his model. For example, if the null hypothesis of cointegration is true, the applied researcher faces a significant probability of not rejecting the null, even by means of a relatively over-sized test. In such a case, his next step is to employ an asymptotically efficient estimator in order to estimate the cointegrating vector(s). The question that naturally arises at this point is the following: "what are the asymptotic properties of such an estimator and those of the associated test statistic in the presence of a variance break?" In other words, even if the applied researcher is fortunate enough to make the correct inferences about the cointegrability of his model, the presence of a variance break is likely to distort his inferences on the cointegrating vector itself.

This paper aims at analyzing the performance of various estimation methods for cointegration in the case of abrupt changes in the variances of the errors of a cointegrated model. Concerning the latter, we adopt the triangular system, proposed by Phillips (1988, 1991), and assume that the variance of the cointegration error, the variance of the error that drives the regressor, along with the covariance of these errors are subject to abrupt upward and/or downward shifts at a single point in the sample. For simplicity, we assume that the errors are temporally independent. Under this set of assumptions, we first derive the limiting distribution of the OLS estimator. It is shown that the presence of a variance break seriously affects the limiting distribution of OLS and that of the corresponding t-statistic. The usual corrections on OLS, which in the absence of any breaks result in the usual cointegration estimators, such as the Fully Modified Least Squares (FMLS) estimator of Phillips and Hansen (1990), or the Dynamic OLS estimator, (DOLS) (see Saikonnen 1991, Phillips and Loretan 1991 and Stock and Watson 1993) are not sufficient to restore asymptotic optimality in the presence of a variance break. Indeed, it is shown that the FMLS and DOLS t-statistics suffer from severe size distortions if the elements of the covariance matrix of the errors experience an upward shift at a specific point in the sample. Similar behavior is exhibited by the system-based Johansen (JOH) t-statistic (see Johansen, 1991). For these cases, we propose a new OLS-based estimator which corrects the ordinary OLS estimator, not only for endogeneity but also for possible variance shifts. It is shown that the new estimator performs very well even when the sample is small and the variance breaks are quite large. It is also shown that the performance of the new estimator is comparable to that of FMLS or DOLS in the case of no variance break.

The remainder of the paper is organized as follows. Section 2 introduces the model and derives the limiting distribution of OLS in the case of variance breaks. Based on these results, it suggests several modifications on OLS that result in a feasible t-statistic which is distributed as $N(0,1)$ under the null hypothesis. Section 3 presents the results from a Monte Carlo study that compare the size performance of the t-statistics which are based on standard cointegration estimators with that of the new t-statistic, under various combinations of shifts in the covariance matrix of the errors. Section 4 concludes the paper. All the proofs can be found in the Appendix.

2 The Model and Theoretical Results

Let $\mathbf{z}_t$ and $\mathbf{u}_t$ be two bivariate processes, with $\mathbf{z}_t = [y_t, x_t]^\top$ and $\mathbf{u}_t = [u_{1t}, u_{2t}]^\top$. We further assume that the generating mechanism for y_t is given by the system:

$$\begin{aligned} y_t &= \theta x_t + u_{1t} \\ x_t &= x_{t-1} + u_{2t} \end{aligned} \tag{1}$$

Concerning the error process, $\mathbf{u}_t$, we assume that it experiences a unique structural break at time $[sT]$, $0 \leq s < 1$, as described by the next assumption.

Assumption A:

$$\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \sim IID\left(0, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}\right), \ 1 \leq t < [sT] \ ,$$

$$\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \sim IID\left(0, \begin{bmatrix} \sigma_1'^2 & \sigma_{12}' \\ \sigma_{12}' & \sigma_2'^2 \end{bmatrix}\right), \ [sT] \leq t \leq T, \text{ and}$$

where, in general, $\sigma_1^2 \neq \sigma_1'^2$, $\sigma_2^2 \neq \sigma_2'^2$ and $\sigma_{12} \neq \sigma_{12}'$, and there exists $B > 0$ such that

$$E[u_{2t}^4] < B$$

for every $t = 1, 2, ..., T$.

Let us define the new error, v_t, as follows

$$v_t = u_{1t} - \frac{\sigma_{12}}{\sigma_2^2} u_{2t}, t < [sT],$$

and

$$v_t = u_{1t} - \frac{\sigma_{12}'}{\sigma_2'^2} u_{2t}, t \geq [sT].$$

Then

$$Var(v_t) = \sigma_v^2 = \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}, t < [sT],$$

and

$$Var(v_t) = \sigma_v'^2 = \sigma_1'^2 - \frac{\sigma_{12}'^2}{\sigma_2'^2}, t \geq [sT].$$

Obviously, $E[u_{2t} v_t] = 0$, for all t.

Next, we define the functions, $D(\cdot)$, $Q(\cdot)$ which will be used in the derivation of the relevant limiting distributions:

Definition: $D(\cdot)$, $Q(\cdot)$ are two functions defined on the interval $[0,1]$, such that

$$\begin{array}{ll} D(r) = \sigma_2 \text{ and } Q(r) = \sigma_v & \text{for } r < s \text{ and} \\ D(r) = \sqrt{\sigma_2'^2 + (\sigma_2^2 - \sigma_2'^2)\frac{s}{r}} \text{ and } Q(r) = \sigma_v' & \text{for } s \leq r. \end{array}$$

Then, the following theorem holds

Theorem 1: *If equations (1) and assumption A hold then*

$$T(\hat{\theta} - \theta) \xrightarrow{L} \frac{\frac{1}{2}\left(s\sigma_{12} + (1-s)\sigma_{12}' + \frac{\sigma_{12}'}{\sigma_2'^2} D^2(1) W_2^2(1) + \left(\frac{\sigma_{12}}{\sigma_2^2} - \frac{\sigma_{12}'}{\sigma_2'^2}\right) D^2(s) W_2^2(s)\right) + \int_0^1 Q(r) D\left(r\right) W_2(r) dW_1(r)}{\int_0^1 D^2(r) W_2^2(r) dr} \tag{2}$$

where $W_1(\cdot)$, $W_2(\cdot)$ *are two independent Wiener processes.*

Remark 1: Note that the superconsistency of the OLS estimator of θ under a variance break is maintained.

Remark 2: The numerator of equation (2) can be separated in two terms. The first term is described by the factors before the stochastic integral and is a result of a possible nonzero covariance between u_{1t} and u_{2t}. The second term is the stochastic integral that also arises in the zero covariance case. Considering the definitions of $Q(\cdot)$ and $D\left(\cdot\right)$, we see that both terms of the numerator as well as the denominator are perturbed in a complicated way due to s and the variance change.

Remark 3: Observe that in (2) the trivial case of $s = 0$, that is when there is no break in covariance matrix of the errors, produces an expression similar to that of the relevant literature (see inter alia, Park and Phillips 1988, Phillips 1991 and White 2001).

Theorem 1 makes clear that the OLS estimator suffers asymptotically from non-constant endogeneity effects. Our next task is to modify OLS in a way such that these effects will be eliminated. To this end, we first define the statistic G_T :

$$\begin{aligned} G_T &= \hat{\theta} - \frac{\frac{\sigma_{12}}{\sigma_2^2}\sum_{t=1}^{[sT]-1} x_t u_{2t} + \frac{\sigma'_{12}}{\sigma'^2_2}\sum_{t=[sT]}^{T} x_t u_{2t}}{\sum_{t=1}^{T} x_t^2} & (3) \\ &= \frac{\sum_{t=1}^{T} x_t y_t - \frac{\sigma_{12}}{\sigma_2^2}\sum_{t=1}^{[sT]-1} x_t u_{2t} - \frac{\sigma'_{12}}{\sigma'^2_2}\sum_{t=[sT]}^{T} x_t u_{2t}}{\sum_{t=1}^{T} x_t^2}. \\ &= \frac{\theta\sum_{t=1}^{T} x_t^2 + \sum_{t=1}^{[sT]-1} x_t u_{1t} - \frac{\sigma_{12}}{\sigma_2^2}\sum_{t=1}^{[sT]-1} x_t u_{2t} + \sum_{t=[sT]}^{T} x_t u_{1t} - \frac{\sigma'_{12}}{\sigma'^2_2}\sum_{t=[sT]}^{T} x_t u_{2t}}{\sum_{t=1}^{T} x_t^2} \\ &= \theta + \frac{\sum_{t=1}^{T} x_{t-1} v_t}{\sum_{t=1}^{T} x_t^2} & (4) \end{aligned}$$

Note that the error term u_{2t} is obtained directly from the equation $u_{2t} = x_t - x_{t-1}$. Then, using the asymptotic result of Theorem 1, we prove that the G_T-based t-statistic is asymptotically distributed as $N(0,1)$.

Theorem 2: *If equations (1) and assumption A hold, then*

$$\begin{aligned} \tilde{t}_b\Big|_{x_1,x_2,\ldots} &\equiv \left.\frac{G_T - \theta}{s.e.\,(G_T)}\right|_{x_1,x_2,\ldots} = \left.\frac{(G_T-\theta)\sum_{t=1}^{T} x_t^2}{\sqrt{\sum_{t=1}^{T} Q^2\left(\frac{t}{T}\right) x_t^2}}\right|_{x_1,x_2,\ldots} = & (5) \\ &= \left.\frac{\sum_{t=1}^{T} x_t y_t - \frac{\sigma_{12}}{\sigma_2^2}\sum_{t=1}^{[sT]-1} x_t u_{2t} - \frac{\sigma'_{12}}{\sigma'^2_2}\sum_{t=[sT]}^{T} x_t u_{2t} - \theta\sum_{t=1}^{T} x_t^2}{\sqrt{\sum_{t=1}^{[sT]-1}\sigma_v^2 x_t^2 + \sum_{t=[sT]}^{T}\sigma'^2_v x_t^2}}\right|_{x_1,x_2,\ldots} \xrightarrow{L} N(0,1) & (6) \end{aligned}$$

as $T \to \infty$.

The problem of finding consistent estimators for the variance parameters would be solved if we could estimate the parameter s, that is the timing of the break. In order deal with this problem we follow the method of maximizing a statistic with respect to a supposed break parameter r. This statistic is defined as follows:

$$A_T(r) = \left|\sum_{t=1}^{[rT]-1}\frac{(\Delta x_t)^2}{[rT]-1} - \sum_{t=[rT]}^{T}\frac{(\Delta x_t)^2}{T-[rT]+1}\right|, \ \frac{2}{T} < r \le 1 \ .$$

Then, the following result is proved:

Theorem 3: *If assumption A holds and* $0 < \alpha < \beta < 1$, *then* $A_T(r) \xrightarrow{p} \frac{1-s}{1-r}\left|\sigma_2^2 - \sigma'^2_2\right|$ *if* $a \le r \le s$ *and* $A_T(r) \xrightarrow{p} \frac{s}{r}\left|\sigma_2^2 - \sigma'^2_2\right|$ if $s < r \le \beta$, *uniformly for every* $r \in [\alpha, \beta]$.

Note that the maximum value of the $p\lim$ of $A_T(r)$ is $\left|\sigma_2^2 - \sigma'^2_2\right|$ and is obtained when $r = s$.

If we set

$$\hat{s}_T = \frac{1}{T} \arg \max_{k=[\alpha T],[\alpha T]+1,\ldots,[\beta T]} \left\{ \left| \sum_{t=1}^{k} \frac{(\Delta x_t)^2}{k} - \sum_{t=T-k+1}^{T} \frac{(\Delta x_t)^2}{T-k} \right| \right\} \tag{7}$$

then, by definition, $A_T(\hat{s}_T) \geq A_T(s)$ for every T with probability 1. Using this and Theorem 3 we can prove the following:

Theorem 4: *If assumption A holds and we can find* $\alpha, \beta \in (0,1)$ *such that* $\alpha < s < \beta$ *then* $A_T(\hat{s}_T) \xrightarrow{p} \left|\sigma_2^2 - \sigma_2'^2\right|$ *and* $\hat{s}_T \xrightarrow{p} s$.

Note that α and β are important for the definition of $\hat{s}_T$. Moreover, $\hat{s}_T T$ is an integer, therefore $[\hat{s}_T T] = \hat{s}_T T$. From the proofs of Theorems 3 and 4 (with minor modifications) we obtain consistent estimators for the covariance matrices of u_{1t} and u_{2t} before and after the variance break:

Corollary 1: *If assumption A holds and we can find* $\alpha, \beta \in (0,1)$ *such that* $\alpha < s < \beta$ *then*

a) $\hat{\sigma}_i^2 = \frac{1}{\hat{s}_T T-1} \sum_{t=1}^{\hat{s}_T T-1} u_{it}^2 \xrightarrow{p} \sigma_i^2$, $i = 1, 2, v$,

b) $\hat{\sigma}_i'^2 = \frac{1}{T-\hat{s}_T T+1} \sum_{t=\hat{s}_T T}^{T} u_{it}'^2 \xrightarrow{p} \sigma_i'^2$, $i = 1, 2, v$,

c) $\hat{\sigma}_{12} = \frac{1}{\hat{s}_T T-1} \sum_{t=1}^{\hat{s}_T T-1} u_{1t} u_{2t} \xrightarrow{p} \sigma_{12}$, and

d) $\hat{\sigma}_{12}' = \frac{1}{T-\hat{s}_T T+1} \sum_{t=\hat{s}_T T}^{T} u_{1t}' u_{2t}' \xrightarrow{p} \sigma_{12}'$.

For our final theorem we also need the following lemma:

Lemma 1: *If assumption A holds and we can find* $\alpha, \beta \in (0,1)$ *such that* $\alpha < s < \beta$ *then*

a) $\frac{1}{T}\left(\sum_{t=1}^{\hat{s}_T T-1} x_t u_{2t} - \sum_{t=1}^{[sT]-1} x_t u_{2t}\right) \xrightarrow{p} 0$,

b) $\frac{1}{T}\left(\sum_{t=\hat{s}_T T}^{T} x_t u_{2t} - \sum_{t=[sT]}^{T} x_t u_{2t}\right) \xrightarrow{p} 0$,

c) $\frac{1}{T}\left(\sum_{t=1}^{\hat{s}_T T-1} x_t^2 - \sum_{t=1}^{[sT]-1} x_t^2\right) \xrightarrow{p} 0$, and

d) $\frac{1}{T}\left(\sum_{t=\hat{s}_T T}^{T} x_t^2 - \sum_{t=[st]}^{T} x_t^2\right) \xrightarrow{p} 0$.

Applying now the results of Corollary 1 and Lemma 1 to Theorem 2 we obtain the following theorem that provides us with a feasible t-statistic:

Theorem 5: *If equations (1) and assumption A hold and we can find* $\alpha, \beta \in (0,1)$ *such that* $\alpha < s < \beta$ *then*

$$\hat{t}_b\Big|_{x_1,x_2,\ldots} \equiv \left. \frac{\sum_{t=1}^{T} x_t y_t - \frac{\hat{\sigma}_{12}}{\hat{\sigma}_2^2} \sum_{t=1}^{\hat{s}_T T-1} x_t u_{2t} - \frac{\hat{\sigma}_{12}'}{\hat{\sigma}_2'^2} \sum_{t=\hat{s}_T T}^{T} x_t u_{2t} - \theta \sum_{t=1}^{T} x_t^2}{\sqrt{\hat{\sigma}_v^2 \sum_{t=1}^{\hat{s}_T T-1} x_t^2 + \hat{\sigma}_v'^2 \sum_{t=\hat{s}_T T}^{T} x_t^2}} \right|_{x_1,x_2,\ldots} \xrightarrow{L} N(0,1) \;, \tag{8}$$

as $T \to \infty$.

3 Monte Carlo results

In this section we conduct a Monte Carlo study that aims at comparing the small sample properties of the new t-statistic, t_b, with those of the t-statistics that correspond to standard cointegration methods, such as OLS, DOLS, FMLS and JOH.[1] The new t-statistic, t_b, is calculated from equation (8) with the time of the break estimated from equation (7). To evaluate the performance of the aforementioned estimators, we compute the empirical sizes of t_{OLS}, t_{DOLS}, t_{JOH}, t_{FMLS} and t_b, for testing the hypothesis $\theta = 1$ for nominal sizes of 5%. We generate 3000 series of length 200 starting with $u_{10} = u_{20} = 0$, and then discard the initial 100 observations, thus generating a sample size, T, of 100 observations. Alternative sample sizes are also considered, ranging from $T = 50$ to $T = 5000$. Below, we report the results only for $T = 100$ and $T = 300$.[2]

We assume that $\mathbf{u}_t$ follows a bivariate, zero-mean, Normal independent process. The covariance matrix of $\mathbf{u}_t$ experiences a break of size a at period sT, $0 < s < 1$. The time of the break, s, takes values in the interval [0.1, 0.9], by steps of 0.1. We also include the case $s = 0$, which corresponds to the case of no variance breaks. Various combinations of changes in the covariance matrix of $\mathbf{u}_t$, at a single point of the sample, are analyzed. Below, we discuss the cases of "variance increase" and "variance decline". In particular, the elements of the covariance matrix change from $\sigma_1^2 = \sigma_2^2 = 1$ to $\sigma_1^2 = \sigma_2^2 = a^2$, at s, where $a = 2$, 10, 100 for the case of "variance increase" and $a = 0.5$, 0.1, and 0.01 for the case of "variance decline". In both cases the covariance σ_{12} changes in a way that keeps the correlation coefficient, ρ_{12}, between u_{1t} and u_{2t} constant between the pre- and post-break periods. We consider various values of ρ_{12}, corresponding to positive, zero and negative correlation between the errors.

The results for the case of "variance increase" are reported in Tables 1a, 1b and 1c for the cases of $\rho_{12} = 0.6$, $\rho_{12} = 0$ and $\rho_{12} = -0.6$, respectively. This set of results may be summarized as follows:

(i) The standard t-statistics, namely t_{OLS}, t_{DOLS}, t_{JOH} and t_{FMLS} suffer from size distortions which are maximized when the increases in σ_1^2 and σ_2^2 occur towards the end of the sample. These distortions are particularly severe when the variance break is large. For example, for $\rho_{12} = 0.6$, $a = 10$, $T = 100$, and $s = 0.9$, the empirical sizes of t_{OLS}, t_{DOLS}, t_{JOH} and t_{FMLS}

[1] Since the errors in the model are martingale differences, it is sufficient to augment (1) only by Δx_t, thus we employ DOLS(1). With respect to the FMLS estimator, we employ the Bartlett kernel and set the bandwidth equal to 3. Finally, the lag-order of the Vector Autoregressive Model, on which JOH is based, is set equal to one.

[2] The full set of results for alternative sample sizes are available upon request from the authors.

are 59.03%, 44.07%, 41.83% and 41.93% respectively. This pattern remains largely the same for alternative values of ρ_{12}. As expected, t_{OLS} exhibits the worst performance, since it does not attempt any correction whatsoever for the correlation between u_{1t} and u_{2t}. It is also worth noting that t_{OLS} is significantly oversized, even in the case of $\rho_{12} = 0$, in which OLS is the optimal estimator under no variance breaks.

(ii) The new t-statistic, t_b, works very well for small values of s. Nevertheless, some size distortions are still present in the cases of large values of s and a. For example, for $a = 10$, $T = 100$, and $s = 0.9$, the empirical size of t_b is equal to 14.03%. When the sample size increases, the size distortions of t_b decline albeit at a very slow rate. For example, when T increases from $T = 100$ to $T = 300$, the size of t_b decreases from 14.03% to 13.70%. This size reduction reflects the fact that the empirical distribution of t_b does approach the standard normal distribution as the sample size increases. On the contrary, the divergencies between the distributions of t_{DOLS}, t_{JOH}, t_{FMLS} and $N(0,1)$ do not tend to disappear. Specifically, the empirical size of t_b hovers around the nominal size of 5% for $T = 1000$ irrespective of the magnitude of the break, while the respective figures for t_{JOH} and t_{FMLS} are over 35%.[3] To illustrate these issues, we report Figures 1A, 1B, 1C and 1D, which plot the kernel estimates of the empirical distributions of t_{DOLS}, t_{JOH}, t_{FMLS} and t_b respectively against the standard normal for $T = 1000$, $a = 10$, $\rho_{12} = 0.6$ and $s = 0.9$. While the distribution of t_b is almost identical to $N(0,1)$, the other three t-statistics exhibit large distributional divergencies from $N(0,1)$. Specifically, the distributions of t_{DOLS}, t_{JOH} and t_{FMLS} appear more platykurtic and dispersed than the standard normal and consequently the empirical 2.5% and 97.5% critical values are much larger (in absolute value) than 1.96.

(iii) In the case of zero or negative correlation between the errors (Tables 1b and 1c), the empirical distribution of t_b approaches the standard normal distribution at a faster rate compared to the case of positive correlation, while the behavior of the other estimators remains broadly unchanged.

(iv) More importantly, the price that we have to pay for employing t_b in the absence of a variance break is negligible. Specifically, when $s = 0$, that is when there is no variance breaks, the empirical size of t_b is close to 5% for all the parameter settings and the sample sizes considered.

Next we consider the case of "variance decline". The results, reported in Tables 2a, 2b and

[3]Similar size distortions are present even for the case of a particularly large size of $T = 5000$.

2c for the cases of $\rho_{12} = 0.6$, $\rho_{12} = 0$ and $\rho_{12} = -0.6$, respectively, may be summarized as follows:

(i) In the case of a downward shift in σ_1^2 and σ_2^2, the effect of the time of the shift on the performance of t-statistics is the opposite to that of the case of an upward shift; the earlier the downward shift has occurred, the greater the distributional divergencies of t_{OLS}, t_{DOLS}, t_{JOH}, t_{FMLS} from the standard normal are.

(ii) Although the distributional divergencies mentioned above are clearly translated to size distortions for the t_{DOLS} statistic, they have a milder effect on the empirical sizes of t_{JOH}, t_{FMLS}. For example, for $a = 0.01$, $\rho_{12} = 0.6$ and $T = 100$, the empirical size of t_{DOLS} reaches its maximum value of 63.10% for $s = 0.1$, whereas the corresponding sizes of t_{FMLS} and t_{JOH} do not exceed 16%. These effects mainly stem from the fact that the empirical distributions of t_{JOH}, t_{FMLS} are much more leptokurtic and concentrated around the mean than the distribution of t_{DOLS}. Table 3 reports the mean, standard deviation, skewness and kurtosis coefficients, as well as the 2.5% and 97.5% points of the empirical distributions of the t-statistics under consideration, for the case of $\rho_{12} = 0.6$, $T = 300$, $a = 10$ and $a = 0.1$. It can be seen that in the case of "variance decline" and $s = 0.1$ the kurtosis coefficients of t_{JOH}, t_{FMLS} are 8.31 and 7.55 respectively, whereas that of t_{DOLS} is just 3.15. Moreover, the respective standard deviations are reasonably close to one for t_{JOH} and t_{FMLS}, while the associated value of t_{DOLS} is more than double reaching 2.61. As a result, the 2.5% and 97.5% points of the empirical distribution of t_{FMLS} are -2.54 and 2.18 respectively, whereas the corresponding points for t_{DOLS} are -5.16 and 5.32.

(iii) The performance of the new t-statistic, t_b, for the case of "variance decline" is the mirror image of that for the case of "variance increase". In particular, the new method works very well for large values of s, whereas some size distortions are present in the cases of small sample sizes and small values of s and a. Again, when the sample size increases, these size distortions decline at a rate which seems to be faster than that of the case of "variance increase". For example, when T increases from $T = 100$ to $T = 300$, the size of t_b decreases from 13.83% to 6.93% for the case $\rho_{12} = 0.6$, and $a = 0.01$.

(iv) Figures 2A, 2B, 2C and 2D report the kernel estimates of the empirical distributions of t_{DOLS}, t_{JOH}, t_{FMLS} and t_b respectively for $T = 1000$, $a = 0.1$, $\rho = 0.6$ and $s = 0.1$. A picture similar to that observed for the case of "variance increase" is observed; the distribution of t_b is very close to $N(0, 1)$, whereas the one of t_{DOLS} is quite different from the standard normal.

Interestingly, the empirical distributions of t_{JOH} and t_{FMLS} exhibit apparent differences between the cases of "variance increase" and "variance decline": the tails of these distributions are much heavier in the former than in the latter case. This fact explains why t_{JOH} and t_{FMLS} are significantly more oversized in the case of "variance increase" than in the case of "variance decline".

4 Conclusions

In this paper we examined the effects of a single shift in the elements of the covariance matrix of the cointegration error and the error that drives the regressor on the performance of several cointegration estimators. We first showed analytically that the limit distribution of the simple OLS estimator is significantly affected by the presence of a variance break. Based on this, we derived a modified t-statistic which is based on the estimation of the break point and it is asymptotically distributed as $N(0,1)$. The new t-statistic accounts for any possible configuration of changes in the covariance matrix of the errors, provided that these changes occur at a single common point in the sample. Simulation evidence suggests that the gains in statistical inference from employing the new t-statistic, instead of the t-statistics that are based on standard cointegration methods, may be quite significant especially when a "variance increase" occurs towards the end of the sample or a "variance decline" occurs early in the sample.

The analysis of this paper may be extended in at least three directions: First, the assumption of independence in the errors of the model may be replaced by that of asymptotic independence (mixing). Second, the number of regressors can increase from 1 to k. In this framework, however, it would be less realistic to claim that the time of the break is common for all the regressors. Third, additional deterministic components, such as constants and/or time trends may be included in the cointegrating regression. All these issues are currently under research by the authors.

References

[1]Burridge P. and A.M.R. Taylor, 2001, On regression-based tests for seasonal unit roots in the presence of periodic heteroscedasticity, *Journal of Econometrics*, 104, 91-117.

[2]Busetti, F. and A.M.R. Taylor, 2003, Variance shifts, structural breaks, and stationarity tests, *Journal of Business & Economic Statistics*, 21(4), 510-531.

[3]Cavaliere, G., 2004a, Unit root tests under time-varying variances, *Econometric Reviews*, 23, 259-292.

[4]Cavaliere, G., 2004b, Testing stationarity under a permanent variance shift, *Economics Letters*, 82(3), 403-408.

[5]Cavaliere, G. and A.M.R. Taylor, 2005, Stationarity tests under time-varying second moments, *Econometric Theory*, 21(6), 1112-1129.

[6]Cavaliere, G. and A.M.R. Taylor, 2006, Testing the null of co-integration in the presence of variance breaks, *Journal of Time Series Analysis*, 27, 613-636.

[7]Hamori, S. and A. Tokihisa, 1997, Testing for a unit root in the presence of a variance shift, *Economics Letters,* 57, 245–253.

[8]Johansen, S., 1991, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica*, 59, 1551-1580.

[9]Kim, C-J and C. R. Nelson, 1999, Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle, *The Review of Economics and Statistics*, 81(4), 608-616.

[10]Kim, T. H., S. J. Leybourne and P. Newbold, 2002, Unit root tests with a break in innovation variance, *Journal of Econometrics,* 109, 365–387.

[11]McConnell, M. and G. Perez Quiros, 2000, Output fluctuations in the United States: What has changed since the early 1980s?, *American Economic Review*, 1464-1476.

[12]Park, J. Y. and P.C.B. Phillips, 1988, Statistical inference in regressions with integrated processes: Part 1, *Econometric Theory*, 4, 468-497.

[13]Phillips, P.C.B., 1988, Reflections on Econometric Methodology. *Economic Record,* 64, 334-359.

[14]Phillips, P.C.B., 1991, Optimal inference in cointegrated systems, *Econometrica*, 59, 283-306.

[15]Phillips, P.C.B. and B.E. Hansen, 1990, Statistical inference in instrumental variables regression with I(1) processes, *Review of Economic Studies,* 57, 99-125.

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

[17]Saikkonen, P., 1991, Asymptotically Efficient Estimation of the Cointegration Regressions, *Econometric Theory*, 7,1, 1-27.

[18]Sensier, M. and D. van Dijk, 2004, Testing for volatility changes in U.S. macroeconomic time series, *The Review of Economics and Statistics*, 86(3), 833-839.

[19]Stock, J.H. and M.W. Watson, 1993, A simple estimator of cointegrating vectors in higher order integrated systems, *Econometrica,* 61, 783-820.

[20]Stock, J.H. and M.W. Watson, 2002, Has the business cycle changed and why? *NBER Macroeconomics Annual.*

[21]White, H., 2001, *Asymptotic Theory for Econometricians*, Academic Press.

APPENDIX

Proof of Theorem 1: We start from the following expression

$$\hat{\theta} - \theta = \frac{\sum_{t=1}^{T} x_t u_{1t}}{\sum_{t=1}^{T} x_t^2}$$

Let us set $X_T(r) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} u_{2t}$ for $\frac{1}{T} \leq r \leq 1$ and $X_T(r) = 0$ for $0 \leq r \leq \frac{1}{T}$. Then, for $0 < r < s$ we have

$$X_T(r) = \frac{\sqrt{[rT]}}{\sqrt{T}} \frac{1}{\sqrt{[rT]}} \sum_{t=1}^{[rT]} u_{2t} \xrightarrow{L} \sigma_2 \sqrt{r} N(0,1) \equiv \sigma_2 N(0,r)$$

For $s \leq r \leq 1$ we also have

$$\begin{aligned}
X_T(r) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]-1} u_{2t} + \frac{1}{\sqrt{T}} \sum_{t=[sT]}^{[rT]} u_{2t} = \frac{\sqrt{[sT]-1}}{\sqrt{T}} \frac{1}{\sqrt{[sT]-1}} \sum_{t=1}^{[sT]-1} u_{2t} \\
&\quad + \frac{\sqrt{[rT]-[sT]+1}}{\sqrt{T}} \frac{1}{\sqrt{[rT]-[sT]+1}} \sum_{t=[sT]}^{[rT]} u_{2t} \\
&\quad \xrightarrow{L} \sigma_2 \sqrt{s} N_1(0,1) + \sigma_2' \sqrt{r-s} N_2(0,1) \\
&\equiv \sigma_2 \left(N_1(0,s) + N_2(0,r-s) \right) \\
&= \sqrt{\sigma_2^2 \frac{s}{r} + \sigma_2'^2 \frac{(r-s)}{r}} N(0,r) = \sqrt{\sigma_2'^2 + (\sigma_2^2 - \sigma_2'^2) \frac{s}{r}} N(0,r) \qquad (9)
\end{aligned}$$

where for the special case of $s = r = 0$ the definition of $X_T(r)$ yields $X_T(r) = 0$ with probability 1. From equation (9) we observe that the process $X_T(\cdot)$ divided only by the standard deviation σ_2 does not converge to a standard Wiener process. If, however, we set

$$Z_T(r) = \frac{1}{D(r)} X_T(r) \ ,$$

then for every $r \in [0,1]$ we have

$$Z_T(r) \xrightarrow{L} N(0,r) \ .$$

Therefore, since u_{2t} are independent we obtain

$$Z_T(r) \xrightarrow{L} W_2(r) \ ,$$

where $W_2(\cdot)$ is a standard Wiener process. The last result allows us to obtain the asymptotic distribution of the denominator of the OLS bias as follows

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T} x_t^2 &= \int_0^1 X_T^2(r)dr = \int_0^1 D^2(r)Z_T^2(r)dr \\ &\xrightarrow{L} \int_0^1 D^2(r)W_2^2(r)dr\ . \end{aligned} \tag{10}$$

With respect to the numerator of the OLS bias, first observe that we can write $u_{1t} = \frac{\sigma_{12}}{\sigma_2^2}u_{2t} + v_t, t < [sT]$, and $u_{1t} = \frac{\sigma'_{12}}{\sigma_2'^2}u_{2t} + v_t, t \geq [sT]$ where $E[u_{2t}v_t] = 0$. Then,

$$\begin{aligned}\sum_{t=1}^{T} x_t u_{1t} &= \frac{\sigma_{12}}{\sigma_2^2}\sum_{t=1}^{[sT]-1} x_t u_{2t} + \frac{\sigma'_{12}}{\sigma_2'^2}\sum_{t=[sT]}^{T} x_t u_{2t} + \sum_{t=1}^{T} x_t v_t \\ &= \frac{\sigma_{12}}{\sigma_2^2}\sum_{t=1}^{[sT]-1} \frac{1}{2}\left(x_t^2 - x_{t-1}^2 + u_{2t}^2\right) + \frac{\sigma'_{12}}{\sigma_2'^2}\sum_{t=[sT]}^{T} \frac{1}{2}\left(x_t^2 - x_{t-1}^2 + u_{2t}^2\right) \\ &\quad + \sum_{t=1}^{T} x_{t-1}v_t + \sum_{t=1}^{T} u_{2t}v_t \\ &= \frac{1}{2}\frac{\sigma_{12}}{\sigma_2^2}\left(\sum_{t=1}^{[sT]-1} u_{2t}^2 + x_{[sT]-1}^2 - x_0^2\right) + \frac{1}{2}\frac{\sigma'_{12}}{\sigma_2'^2}\left(\sum_{t=[sT]}^{T} u_{2t}^2 + x_T^2 - x_{[sT]}^2\right) \\ &\quad + \sum_{t=1}^{T} x_{t-1}v_t + \sum_{t=1}^{T} u_{2t}v_t \\ &= \frac{1}{2}\left(\frac{\sigma_{12}}{\sigma_2^2}\sum_{t=1}^{[sT]-1} u_{2t}^2 + \frac{\sigma'_{12}}{\sigma_2'^2}\sum_{t=[sT]}^{T} u_{2t}^2 + \frac{\sigma'_{12}}{\sigma_2'^2}x_T^2 + \frac{\sigma_{12}}{\sigma_2^2}x_{[sT]-1}^2 - \frac{\sigma'_{12}}{\sigma_2'^2}x_{[sT]}^2\right. \\ &\quad \left. - \frac{\sigma_{12}}{\sigma_2^2}x_0^2\right) + \sum_{t=1}^{T} x_{t-1}v_t + \sum_{t=1}^{T} u_{2t}v_t \end{aligned} \tag{11}$$

For the first two terms of (11) we can write

$$\frac{\sum_{t=1}^{[sT]-1} u_{2t}^2}{T} = \frac{[sT]-1}{T}\frac{1}{[sT]-1}\sum_{t=1}^{[sT]-1} u_{2t}^2 \xrightarrow{p} s\sigma_2^2 \tag{12}$$

and

$$\frac{\sum_{t=[sT]}^{T} u_{2t}^2}{T} = \frac{T-[sT]+1}{T}\frac{1}{T-[sT]+1}\sum_{t=[sT]}^{T} u_{2t}^2 \xrightarrow{p} (1-s)\sigma_2'^2 \tag{13}$$

Since $\frac{x_T^2}{T} = X_T^2(1)$ we have that

$$\frac{x_T^2}{T} \xrightarrow{L} D^2(1)W_2^2(1)\ . \tag{14}$$

Moreover,

$$\frac{x_0^2}{T} \xrightarrow{p} 0 \ . \tag{15}$$

For the remaining terms of (11) we have

$$\begin{aligned}\frac{\sigma_{12}}{\sigma_2^2}x_{[sT]-1}^2 - \frac{\sigma_{12}'}{\sigma_2'^2}x_{[sT]}^2 &= \frac{\sigma_{12}}{\sigma_2^2}x_{[sT]-1}^2 - \frac{\sigma_{12}'}{\sigma_2'^2}\left(x_{[sT]-1}^2 + u_{2[sT]}^2 + 2x_{[sT]-1}u_{2[sT]}\right) \\ &= \left(\frac{\sigma_{12}}{\sigma_2^2} - \frac{\sigma_{12}'}{\sigma_2'^2}\right)x_{[sT]-1}^2 - \frac{\sigma_{12}'}{\sigma_2'^2}\left(u_{2[sT]}^2 + 2x_{[sT]-1}u_{2[sT]}\right)\end{aligned}$$

Now

$$\frac{x_{[sT]-1}^2}{T} \xrightarrow{L} D^2(s)W_2^2(s) \tag{16}$$

$$\frac{u_{2[sT]}^2}{T} \xrightarrow{p} 0 \tag{17}$$

and since $x_{[sT]-1}$ and $u_{2[sT]}$ are independent

$$\frac{x_{[sT]-1}u_{2[sT]}}{T} \xrightarrow{p} 0 \tag{18}$$

Let us now try to treat the next term of equation (11). In order to obtain an asymptotic result and following the usual treatment for the constant variance case, we must first construct a new Wiener process. To simplify things we denote by σ_3^2 the variance of v_t. It is easy to see that $Var(v_t) = \sigma_3^2 = \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}$ for $0 \leq t < [sT]$ and $Var(v_t) = \sigma_3'^2 = \sigma_1'^2 - \frac{\sigma_{12}'^2}{\sigma_2'^2}$ for $[sT] \leq t$. If we set

$$w_t = \begin{array}{ll} \frac{1}{\sigma_3}v_t & \text{for } t < [sT] \text{ and} \\ \frac{1}{\sigma_3'}v_t & \text{for } t \geq [sT] \end{array}$$

we can define by $W_1(\cdot)$ the corresponding to w_t Wiener process, i.e. if we set

$$W_T(r) = \frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]} w_t,\ 0 \leq r \leq 1$$

then we have

$$W_T(r) \xrightarrow{L} W_1(r),\ 0 \leq r \leq 1$$

where $W_1(\cdot)$ is a standard Wiener process. Since u_{2t} and v_t are not correlated, $W_1(\cdot)$ and $W_2(\cdot)$

are independent. Considering this and the definition of the stochastic integral we obtain

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} x_{t-1}v_t &= \sum_{t=1}^{T} Q\left(\frac{t}{T}\right) D\left(\frac{t-1}{T}\right) Z_T\left(\frac{t-1}{T}\right)\left(W_T\left(\frac{t}{T}\right)-W_T\left(\frac{t-1}{T}\right)\right) \\
&= \sum_{t=1}^{T} Q\left(\frac{t-1}{T}\right) D\left(\frac{t-1}{T}\right) Z_T\left(\frac{t-1}{T}\right)\left(W_T\left(\frac{t}{T}\right)-W_T\left(\frac{t-1}{T}\right)\right) \\
& \quad +\sigma_2 Z_T\left(\frac{[sT]-1}{T}\right)\left(\frac{w_{[sT]}}{\sqrt{T}}\right) \\
& \quad \xrightarrow{L} \int_0^1 Q(r)D(r)W_2(r)dW_1(r). \qquad (19)
\end{aligned}
$$

Finally, for the last term of equation (11) we have that the stochastic sequence $u_{2t}v_t$ has zero mean and uniformly bounded variance. Therefore,

$$
\frac{\sum_{t=1}^{T} u_{2t}v_t}{T} \xrightarrow{p} 0. \qquad (20)
$$

Applying (12) - (20) to (11) we obtain

$$
T(\hat{\theta}-\theta) \xrightarrow{L} \frac{\frac{1}{2}\left(s\sigma_{12}+(1-s)\sigma_{12}'+\frac{\sigma_{12}'}{\sigma_2'^2}D^2(1)W_2^2(1)+\left(\frac{\sigma_{12}}{\sigma_2^2}-\frac{\sigma_{12}'}{\sigma_2'^2}\right)D^2(s)W_2^2(s)\right)+\int_0^1 Q(r)D(r)W_2(r)dW_1(r)}{\int_0^1 D^2(r)W_2^2(r)dr} \qquad (21)
$$

Proof of Theorem 2: From (4) we have

$$
G_T-\theta = \frac{\sum_{t=1}^{T} x_{t-1}v_t}{\sum_{t=1}^{T} x_t^2}
$$

Consider the standard deviation of G_T conditional on $\{x_0, x_1, \ldots, x_T\}$,

$$
\begin{aligned}
s.e(G_T|_{x_1,x_2,\ldots}) &= \left(E\left[\left(\sum_{t=1}^{T}\frac{x_t}{\sum_{t=1}^{T} x_t^2}v_t\right)^2\Bigg|_{x_1,x_2,\ldots}\right]\right)^{1/2} \\
&= \frac{\sqrt{\sigma_3^2\sum_{t=1}^{[sT]-1} x_t^2+\sigma_3'^2\sum_{t=[sT]}^{T} x_t^2}}{\sum_{t=1}^{T} x_t^2}
\end{aligned}
$$

Thus, from (11), we obtain a t-statistic for G_T

$$
\tilde{t}_b = \frac{(G_T-\theta)\sum_{t=1}^{T} x_t^2}{\sqrt{\sum_{t=1}^{T} Q^2\left(\frac{t}{T}\right)x_t^2}} = \frac{\sum_{t=1}^{T} x_{t-1}v_t}{\sqrt{\sum_{t=1}^{T} Q^2\left(\frac{t}{T}\right)x_t^2}} \qquad (22)
$$

From (19) we have that conditional on the limiting process $W_2(\cdot)$, which is produced by the x_t's, the limit of the nominator of (22) is equal to $\int_0^1 Q(r)D\left(r\right)W_2(r)dW_1(r)\mid_{W_2(\cdot)}$, which has a zero mean normal distribution with variance equal to

$$E\left[\left(\sigma_2\sigma_3\int_0^1 Q(r)D\left(r\right)W_2(r)dW_1(r)\mid_{W_2(\cdot)}\right)^2\right]=\int_0^1 Q^2(r)D^2\left(r\right)W_2^2(r)dr \tag{23}$$

Moreover, using the definitions of $Z_T(r)$ and $X_T(r)$ we have

$$\frac{1}{T^2}\sum_{t=1}^{T}Q^2\left(\frac{t}{T}\right)x_t^2=\int_0^1 Q^2\left(r\right)D^2\left(\frac{[rT]}{T}\right)Z_T^2(r)dr\xrightarrow{L}\int_0^1 Q^2(r)D^2\left(r\right)W_2^2(r)dr \tag{24}$$

This suggests that

$$\frac{1}{T^2}\sum_{t=1}^{[qT]}Q^2\left(\frac{t}{T}\right)x_t^2=\int_0^q Q^2\left(r\right)D^2\left(\frac{[rT]}{T}\right)Z_T^2(r)dr\rightarrow\int_0^q Q^2(r)D^2\left(r\right)W_2^2(r)dr\ ,\ 0\le q\le 1, \tag{25}$$

where, particularly in (25), $W_2(r)$ represents a realization of $W_2(\cdot)$ and "$\rightarrow$" denotes the point-wise convergence. Combining (23) and (24) in (22) we can finally conclude that for known $x_1, x_2, \ldots$, the statistic

$$\left.\tilde{t}_b\right|_{x_1,x_2,\ldots}=\left.\frac{(G_T-\theta)\sum_{t=1}^T x_t^2}{\sqrt{\sum_{t=1}^T Q^2\left(\frac{t}{T}\right)x_t^2}}\right|_{x_1,x_2,\ldots}\xrightarrow{L}N(0,1)\ . \tag{26}$$

Proof of Theorem 3: Let $\alpha\le r<s$. We set $Q_T(r)=\frac{1}{[rT]-1}\sum_{t=1}^{[rT]-1}u_{2t}^2$. Then, for $\varepsilon>0$,

$$\begin{aligned}P\left(|Q_T(r)-\sigma_2^2|>\varepsilon\right)&\le\frac{1}{\varepsilon^2}E\left[\left(Q_T(r)-\sigma_2^2\right)^2\right]=\frac{1}{\varepsilon^2}\left(E\left[Q_T^2(r)\right]+\sigma_2^4-2\sigma_2^2E\left[Q_T(r)\right]\right)\\&\le\frac{1}{\varepsilon^2}\left(\frac{[rT]-1}{\left([rT]-1\right)^2}B+\frac{\left([rT]-2\right)\left([rT]-1\right)}{\left([rT]-1\right)^2}\sigma_2^4-\sigma_2^4\right)\\&=\frac{1}{\varepsilon^2}\left(\frac{1}{[rT]-1}B-\frac{1}{[rT]-1}\sigma_2^4\right)\le\frac{1}{\varepsilon^2\left([aT]-1\right)}\left(B-\sigma_2^4\right),\end{aligned}$$

where B is defined in Assumption A.

Therefore

$$Q_T(r)\xrightarrow{p}\sigma_2^2 \tag{27}$$

uniformly for every $r\in[\alpha,s)$. Assuming that for T large enough, $[rT]<=[sT]-1$, we set

$R_T(r) = \frac{1}{T-[rT]+1} \sum\limits_{t=[rT]}^{[sT]-1} u_{2t}^2$ and $S_T(r) = \frac{1}{T-[rT]+1} \sum\limits_{t=[sT]}^{T} u_{2t}^2$. Then

$$R_T(r) + S_T(r) = \frac{1}{T-[rT]+1} \sum_{t=[rT]}^{T} u_{2t}^2$$

Moreover

$$P\left(\left|R_T(r) - \frac{[sT]-[rT]}{T-[rT]+1}\sigma_2^2\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2} E\left[\left(R_T(r) - \frac{[sT]-[rT]}{T-[rT]+1}\sigma_2^2\right)^2\right]$$
$$\leq \frac{1}{\varepsilon^2 \left(T-[rT]+1\right)^2} \left(([sT]-[rT])B + ([sT]-[rT]-1)([sT]-[rT])\sigma_2^4 - ([sT]-[rT])^2\sigma_2^4\right)$$
$$\leq \frac{1}{\varepsilon^2 \left(T-[sT]+1\right)^2} \left(([sT]-[aT])B - ([sT]-[rT])\sigma_2^4\right) \leq \frac{([sT]-[aT])}{\varepsilon^2 \left(T-[sT]+1\right)^2} B \tag{28}$$

Therefore $R_T(r) - \frac{[sT]-[rT]}{T-[rT]+1}\sigma_2^2 \xrightarrow{p} 0$ uniformly for every $r \in [\alpha, s)$. On the other hand, considering that if $e \leq d$ are two positive real numbers then

$$d - e - 1 < [d] - [e] < d - e + 1$$

and

$$d - e + 1 < d - [e] + 1 < d - e + 2$$

we can easily prove that

$$\frac{[sT]-[rT]}{T-[rT]+1}\sigma_2^2 \rightarrow \frac{s-r}{1-r}\sigma_2^2$$

uniformly for every $r \in [\alpha, s)$, and conclude that

$$R_T(r) \xrightarrow{p} \frac{s-r}{1-r}\sigma_2^2 \tag{29}$$

uniformly for every $r \in [\alpha, s)$. Finally, using similar arguments we obtain that

$$S_T(r) \xrightarrow{p} \frac{1-s}{1-r}\sigma_2'^2 \tag{30}$$

uniformly for every $r \in [\alpha, s)$. As now $A_T(r) = |Q_T(r) - (R_T(r) + S_T(r))|$, we have shown that

$$A_T(r) \xrightarrow{p} \frac{1-s}{1-r}\left|\sigma_2^2 - \sigma_2'^2\right| \tag{31}$$

uniformly for every $r \in [\alpha, s)$. In a similar way, it is easy to see that when $s < r \leq \beta$,

$$A_T(r) \xrightarrow{p} \frac{s}{r} \left|\sigma_2^2 - \sigma_2'^2\right| \tag{32}$$

uniformly for every $r \in (s, b]$ and

$$A_T(s) \xrightarrow{p} \left|\sigma_2^2 - \sigma_2'^2\right| \tag{33}$$

Proof of Thorem 4: Since $A_T(\hat{s}_T) \geq A_T(s)$ for every T with probability 1 and the limit in probability of $A_T(r)$ is maximized when $r = s$ (see Theorem 3), from the uniform convergence proved in Theorem 3 we immediately have that

$$A_T(\hat{s}_T) \xrightarrow{p} \left|\sigma_2^2 - \sigma_2'^2\right| \tag{34}$$

Let us now suppose that $\hat{s}_T$ does not converge in probability to s. This means that there exist $\varepsilon, \delta > 0$ and a subsequence $\{T_n\}_{n \geq 1}$ such that for every $n \geq 1$ $P(|\hat{s}_{T_n} - s| > \varepsilon) > \delta$. If we set $d = \max\left\{\frac{1-s}{1-s+\varepsilon}, \frac{s}{s+\varepsilon}\right\}$ then $d < 1$, and $d\left|\sigma_2^2 - \sigma_2'^2\right|$ is the maximum value of the limit in probability of $A_T(r)$ for $r \in [\alpha, s-\varepsilon] \cup [s+\varepsilon, \beta]$. We also can find d' such that $d < d' < 1$. Then, from Theorem 3 we have that there exists $\delta' > 0$ such that

$$\liminf P(A_{T_n}(\hat{s}_{T_n}) \leq d' \left|\sigma_2^2 - \sigma_2'^2\right|) \geq \delta'$$

The last relationship clearly contradicts (34) and the proof is complete.

Proof of Lemma 1: We prove only (a), since a similar approach yields (b), (c) and (d). Set $m_T = \min\{\hat{s}_T T, [sT]\}$ and $M_T = \max\{\hat{s}_T T, [sT]\}$. Then, from Theorem 4 we have that $\frac{m_T}{T}, \frac{M_T}{T} \xrightarrow{p} s$. Then,

$$\begin{aligned}
\left|\frac{1}{T}\left(\sum_{t=1}^{\hat{s}_T T-1} x_t u_{2t} - \sum_{t}^{[sT]-1} x_t u_{2t}\right)\right| &= \left\{\begin{array}{cc} 0 & \text{if } \hat{s}_T T = [sT] \\ \frac{1}{T}\left|\sum_{t=m_T}^{M_T-1} x_t u_{2t}\right| & \text{if } \hat{s}_T T \neq [sT] \end{array}\right\} \\
&= \left\{\begin{array}{cc} 0 & \text{if } \hat{s}_T T = [sT] \\ \left|\frac{1}{T}\sum_{t=m_T}^{M_T-1} u_{2t}^2 + \frac{1}{T}\sum_{t=m_T}^{M_T-1} x_{t-1}u_{2t}\right| & \text{if } \hat{s}_T T \neq [sT] \end{array}\right\}
\end{aligned} \tag{35}$$

We also set $\sigma^2_{2m} = \max\{\sigma^2_2, \sigma'^2_2\}$. For every $\varepsilon > 0$, $\varepsilon' > 0$ and for T large enough,

$$
\begin{aligned}
&P\left(\left|\frac{1}{T}\left(\sum_{t=1}^{\hat{s}_T T-1} x_{t-1}u_{2t} - \sum_{t}^{[sT]-1} x_{t-1}u_{2t}\right)\right| > \varepsilon\right) = P\left(\left|\frac{1}{T}\sum_{t=m_T}^{M_T-1} x_{t-1}u_{2t}\right| > \varepsilon\ ;\ \hat{s}_T T \neq [sT]\right) \\
&\leq P\left(\left|\frac{1}{T}\sum_{t=m_T}^{M_T-1} x_{t-1}u_{2t}\right| > \varepsilon\ ;\ \hat{s}_T T \neq [sT]; \frac{M_T - 1 - m_T}{T} < \varepsilon'\right) + P\left(\left|\frac{M_T - m_T}{T}\right| > \varepsilon'\right) \\
&\leq P\left(\left|\frac{1}{T}\sum_{t=[sT]}^{[sT]+l} x_{t-1}u_{2t}\right| > \varepsilon\right) + P\left(\left|\frac{1}{T}\sum_{t=[sT]-l-1}^{[sT]-1} x_{t-1}u_{2t}\right| > \varepsilon\right) + P\left(\left|\frac{M_T - m_T}{T}\right| > \varepsilon'\right) \\
&\qquad \text{for some } l \in \{0, 1, 2, ...[\varepsilon' T]\} \text{ (see also the definition of } m_T \text{ and } M_T\text{)}, \\
&\leq \frac{1}{\varepsilon T^2}\sum_{t=[sT]}^{[sT]+l} E\left[x^2_{t-1}u^2_{2t}\right] + \frac{1}{\varepsilon T^2}\sum_{t=[sT]-l-1}^{[sT]-1} E\left[x^2_{t-1}u^2_{2t}\right] + P\left(\left|\frac{M_T - m_T}{T}\right| > \varepsilon'\right) \\
&\leq \frac{1}{\varepsilon T^2}\left(l([sT] + l - 1)\sigma^4_{2m} + l([sT] - 2)\sigma^4_{2m}\right) + P\left(\left|\frac{M_T - m_T}{T}\right| > \varepsilon'\right) \\
&\leq \frac{1}{\varepsilon T^2}[\varepsilon' T](2[sT] + [\varepsilon' T] - 3)\sigma^4_{2m} + P\left(\left|\frac{M_T - m_T}{T}\right| > \varepsilon'\right) \\
&\qquad\qquad \to \frac{\varepsilon'}{\varepsilon}(2s + \varepsilon')\sigma^4_{2m} \text{ as } T \to \infty.
\end{aligned}
$$

Since ε' was chosen arbitrarily, we conclude that

$$
\frac{1}{T}\left(\sum_{t=1}^{\hat{s}_T T-1} x_{t-1}u_{2t} - \sum_{t}^{[sT]-1} x_{t-1}u_{2t}\right) \xrightarrow{p} 0
$$

Therefore, from (35) and the proof of Theorems 3 and 4 (see also Corollary 1) we conclude that (a) holds.

Table 1α. Empirical 5% size of cointegration estimators as a function of the breakpoint (ρ_{12}=0.6)

Panel A

T=100						T=300				
α=2	σ_{11}=1	σ_{12}=0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	14.63	5.13	5.70	5.83	5.63	13.73	5.00	5.20	5.03	4.90
0.1	14.03	5.23	7.00	7.97	6.63	14.30	4.67	5.97	5.83	5.30
0.2	15.23	5.43	8.67	9.23	6.73	16.10	5.07	6.53	6.97	5.43
0.3	16.30	5.57	8.73	9.47	6.80	16.30	5.77	8.53	8.40	6.53
0.4	19.90	7.13	9.20	9.83	7.80	18.17	7.00	9.27	9.60	6.77
0.5	20.93	7.33	11.30	11.50	8.60	20.00	7.97	9.60	9.63	7.30
0.6	22.53	8.93	11.97	12.47	8.83	22.20	8.33	10.90	11.07	7.57
0.7	24.00	9.33	11.50	12.07	9.03	23.03	9.43	9.30	9.27	6.67
0.8	24.07	9.70	11.60	11.80	8.77	24.70	9.53	9.73	9.57	7.70
0.9	22.83	8.97	9.47	9.90	9.00	21.97	9.20	7.73	7.77	6.00

Panel B

T=100						T=300				
α=10	σ_{11}=1	σ_{12}=0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	14.63	5.13	5.70	5.83	5.63	13.73	5.00	5.20	5.03	4.90
0.1	14.83	5.20	7.30	8.10	6.63	14.40	4.60	6.43	6.47	4.93
0.2	15.80	5.63	9.43	10.00	6.63	16.87	5.13	7.70	8.10	5.23
0.3	17.73	7.33	11.57	12.23	7.53	18.50	6.83	11.20	10.73	6.30
0.4	23.37	9.37	14.23	15.23	7.93	23.00	9.27	13.33	13.77	7.00
0.5	26.60	12.07	18.40	18.60	9.20	28.57	11.97	17.07	16.93	7.70
0.6	33.50	17.03	23.27	24.47	11.40	35.27	16.07	21.20	20.83	9.67
0.7	41.53	24.13	28.87	29.73	12.70	41.33	22.70	25.93	25.63	12.13
0.8	50.50	34.40	37.97	39.00	13.50	52.37	32.03	32.80	32.23	13.60
0.9	59.03	44.07	41.83	41.93	14.03	60.77	42.13	39.50	38.40	13.70

Panel C

T=100						T=300				
α=100	σ_{11}=1	σ_{12}=0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	14.63	5.13	5.70	5.83	5.63	13.73	5.00	5.20	5.03	4.90
0.1	14.97	5.10	7.20	8.07	6.60	14.30	4.70	6.33	6.57	4.83
0.2	16.07	5.63	9.23	9.87	6.73	17.13	5.17	7.77	8.03	5.33
0.3	17.87	7.47	11.83	12.37	7.33	18.77	7.03	11.30	11.20	6.43
0.4	23.40	9.47	14.80	16.30	8.53	23.33	9.57	13.73	14.47	7.23
0.5	26.83	12.03	19.60	19.53	9.20	29.13	12.33	17.70	17.90	8.37
0.6	34.43	18.03	25.63	25.70	11.33	35.90	16.90	22.83	23.07	9.97
0.7	42.70	26.17	33.57	34.10	13.23	43.30	24.03	29.27	29.37	13.10
0.8	53.23	38.83	47.53	47.63	15.17	55.80	35.73	40.10	40.67	15.03
0.9	66.40	54.83	70.23	64.27	16.93	69.17	54.53	59.67	58.17	17.40

Table 1b. Empirical 5% size of cointegration estimators as a function of the breakpoint (ρ_{12}=0)

Panel A

T=100						T=300				
α=2	σ_{11}=1	σ_{12}=0	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	5.30	5.33	6.17	6.73	6.60	5.20	5.17	5.27	5.27	5.07
0.1	4.70	4.87	6.63	7.90	6.33	4.70	4.90	5.73	6.27	5.07
0.2	4.87	4.90	7.93	8.87	6.80	4.87	4.63	7.53	7.77	5.63
0.3	5.63	4.93	7.73	8.77	6.47	5.83	5.70	8.07	8.40	5.37
0.4	6.20	5.87	9.83	11.07	7.37	6.67	7.00	8.83	9.40	5.63
0.5	7.70	7.67	10.63	11.23	8.00	7.53	7.07	9.20	9.93	6.33
0.6	8.60	8.70	10.87	12.13	8.67	9.03	9.13	11.40	11.73	6.77
0.7	9.20	9.00	11.70	12.60	9.20	10.03	9.93	10.10	10.53	6.60
0.8	9.87	9.77	11.13	11.50	8.63	9.57	9.77	10.60	10.73	6.50
0.9	9.10	9.03	10.10	11.23	9.37	8.13	8.23	7.63	8.00	5.80

Panel B

T=100						T=300				
α=10	σ_{11}=1	σ_{12}=0	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	5.30	5.33	6.17	6.73	6.60	5.20	5.17	5.27	5.27	5.07
0.1	4.57	4.63	7.03	8.33	5.93	4.83	4.97	5.70	6.20	4.97
0.2	5.40	5.43	9.50	10.73	6.60	5.13	4.97	8.53	9.17	5.27
0.3	6.70	6.77	11.43	12.17	6.93	6.67	6.37	9.97	10.73	5.10
0.4	8.10	8.17	13.90	15.70	7.73	9.73	9.83	14.33	14.77	6.23
0.5	11.50	12.17	18.47	19.23	9.17	11.57	11.37	16.43	16.77	6.57
0.6	16.27	16.73	23.60	25.30	11.57	16.37	16.50	21.93	23.23	7.30
0.7	23.70	24.53	29.40	30.87	12.20	22.83	22.87	26.23	26.93	8.17
0.8	34.27	35.07	36.63	38.93	14.33	32.53	32.53	32.70	32.83	8.33
0.9	41.97	43.00	41.83	40.17	13.27	42.60	43.30	38.87	39.40	8.07

Panel C

T=100						T=300				
α=100	σ_{11}=1	σ_{12}=0	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	5.30	5.33	6.17	6.73	6.60	5.20	5.17	5.27	5.27	5.07
0.1	4.57	4.53	7.00	8.37	6.03	4.70	4.87	5.57	6.03	4.77
0.2	5.53	5.60	9.43	10.73	6.67	5.07	5.00	8.57	9.20	5.43
0.3	6.67	6.77	11.57	12.23	7.03	6.80	6.60	10.23	11.17	5.10
0.4	8.27	8.27	14.40	15.90	7.80	10.03	10.03	14.90	15.07	6.07
0.5	11.83	12.17	19.20	20.10	9.13	12.00	11.63	17.13	17.60	6.30
0.6	17.20	17.53	25.17	26.93	11.43	16.83	17.27	23.70	24.17	7.03
0.7	25.30	26.23	33.10	34.30	12.73	24.23	24.20	29.03	29.87	8.43
0.8	38.10	39.57	47.70	48.47	16.23	36.27	35.97	40.20	40.57	8.70
0.9	51.80	54.63	69.97	62.17	15.43	54.97	55.03	59.93	60.47	8.80

Table 1c. Empirical 5% size of cointegration estimators as a function of the breakpoint (ρ_{12}=-0.6)

Panel A

T=100						T=300				
α=2	σ_{11}=1	σ_{12}=-0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	13.10	5.50	6.70	6.73	6.53	14.03	4.43	4.37	4.43	4.33
0.1	13.07	5.47	6.93	7.37	6.50	14.37	4.30	4.93	5.30	4.23
0.2	15.27	5.03	7.17	8.10	6.20	14.40	4.63	6.77	7.20	5.00
0.3	16.43	5.23	8.00	8.83	6.67	16.33	4.87	7.40	7.63	4.53
0.4	17.40	6.00	9.67	10.30	7.00	17.40	6.27	8.33	8.33	5.47
0.5	19.27	7.80	10.30	10.90	7.67	19.87	6.73	9.40	9.43	6.00
0.6	20.43	7.57	10.43	11.10	7.27	22.83	8.27	10.60	11.10	6.43
0.7	23.67	9.10	10.47	11.70	7.80	23.00	9.57	10.53	10.87	6.93
0.8	24.10	9.90	11.70	11.93	9.00	23.37	10.37	11.07	10.90	6.27
0.9	22.77	9.47	9.80	11.13	8.77	23.00	8.90	8.33	8.13	6.00

Panel B

T=100						T=300				
α=10	σ_{11}=1	σ_{12}=-0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	13.10	5.50	6.70	6.73	6.53	14.03	4.43	4.37	4.43	4.33
0.1	12.87	5.27	6.93	7.70	6.17	14.47	4.13	4.93	5.30	3.97
0.2	16.37	5.07	8.57	9.67	6.57	15.47	4.90	8.17	7.97	4.70
0.3	18.83	6.27	10.73	11.63	7.10	18.67	5.60	9.47	9.80	4.67
0.4	21.00	8.63	14.13	14.63	7.67	22.23	9.60	12.47	12.47	5.73
0.5	26.37	11.57	18.23	18.80	8.70	28.63	11.80	14.87	15.87	5.87
0.6	33.37	16.20	22.63	24.13	10.47	35.13	16.37	21.10	21.47	6.83
0.7	43.27	23.50	29.30	30.93	12.27	41.53	23.53	27.53	27.43	8.10
0.8	50.73	34.50	38.30	38.67	13.53	51.50	32.10	33.53	32.73	8.00
0.9	60.60	42.93	42.57	42.90	14.13	61.37	42.90	38.50	37.57	7.43

Panel C

T=100						T=300				
α=100	σ_{11}=1	σ_{12}=-0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	13.10	5.50	6.70	6.73	6.53	14.03	4.43	4.37	4.43	4.33
0.1	12.90	5.33	7.03	7.70	6.07	14.40	3.97	5.03	5.40	4.10
0.2	16.53	5.07	8.80	9.80	6.50	15.60	5.03	8.10	8.03	4.47
0.3	19.03	6.37	10.67	11.70	7.00	18.60	5.63	9.50	10.27	4.73
0.4	21.13	8.97	14.00	14.93	7.70	22.13	9.67	13.17	13.27	5.63
0.5	27.30	12.13	18.90	19.10	8.83	29.03	12.10	16.23	17.00	5.73
0.6	33.93	17.27	25.00	26.17	10.93	35.70	16.93	22.37	22.93	7.23
0.7	44.90	25.43	34.20	35.17	12.47	43.23	24.60	29.47	29.67	8.23
0.8	53.20	37.93	47.53	46.13	14.70	54.80	35.67	40.57	40.00	8.70
0.9	68.67	54.37	69.47	65.27	15.83	68.63	54.63	58.17	58.53	7.83

Table 2a. Empirical 5% size of cointegration estimators as a function of the breakpoint (ρ_{12}=0.6)

Panel A

	T=100					T=300				
α=0.5	σ_{11}=1	σ_{12}=0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	14.63	5.13	5.70	5.83	5.63	13.73	5.00	5.20	5.03	4.90
0.1	14.87	8.60	5.93	5.60	6.43	14.97	8.30	4.27	4.03	5.13
0.2	16.20	9.40	5.07	4.37	5.67	15.33	8.70	4.50	4.33	5.17
0.3	14.80	8.90	5.27	4.60	5.23	14.90	8.77	4.53	4.23	5.20
0.4	13.40	8.77	4.37	4.03	4.87	14.57	9.03	4.37	4.37	5.13
0.5	14.03	8.50	5.03	4.93	6.00	12.47	8.00	4.80	4.53	5.90
0.6	11.83	6.87	4.97	5.13	6.00	11.70	6.03	4.10	4.03	5.07
0.7	12.03	6.00	4.90	5.13	6.00	11.10	5.90	4.27	4.13	4.83
0.8	10.90	4.93	5.77	4.97	5.80	10.97	4.97	4.30	4.30	5.13
0.9	11.60	4.63	5.90	6.13	5.83	10.63	4.67	3.93	4.03	4.63

Panel B

	T=100					T=300				
α=0.1	σ_{11}=1	σ_{12}=0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	14.63	5.13	5.70	5.83	5.63	13.73	5.00	5.20	5.03	4.90
0.1	46.87	45.60	10.67	9.17	11.47	45.17	42.87	8.57	7.43	6.83
0.2	38.90	37.10	9.70	7.47	7.70	34.23	33.50	8.67	7.90	5.80
0.3	27.47	25.97	8.43	6.97	7.30	26.27	23.23	7.30	6.97	4.77
0.4	20.87	17.90	6.87	5.80	6.23	19.40	18.23	6.43	5.77	5.73
0.5	16.77	13.30	6.13	5.50	6.70	15.77	13.33	5.27	5.03	5.43
0.6	13.03	10.13	5.47	5.07	5.73	12.50	8.47	3.73	3.90	5.13
0.7	12.17	7.67	5.23	5.17	6.33	10.53	6.90	4.70	4.67	5.33
0.8	10.47	5.03	5.03	4.73	6.03	9.70	5.40	4.67	4.47	5.60
0.9	10.60	4.63	5.50	5.37	6.03	9.73	4.63	3.93	4.00	4.57

Panel C

	T=100					T=300				
α=0.01	σ_{11}=1	σ_{12}=0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	14.63	5.13	5.70	5.83	5.63	13.73	5.00	5.20	5.03	4.90
0.1	62.00	63.10	15.67	12.97	13.83	56.60	55.93	11.47	9.93	6.93
0.2	43.07	41.87	11.33	8.17	8.17	38.43	37.17	9.90	9.00	6.43
0.3	29.40	28.10	8.43	7.30	7.77	27.43	25.47	7.43	7.20	5.07
0.4	21.63	19.03	7.47	6.43	6.33	20.37	18.33	6.67	6.03	5.77
0.5	16.97	13.60	6.23	5.97	6.47	16.13	13.57	5.13	5.00	5.40
0.6	13.33	10.07	5.30	5.33	5.83	12.40	8.63	4.00	4.00	4.77
0.7	12.20	7.90	5.47	5.40	5.93	10.53	7.07	4.80	4.57	5.17
0.8	10.17	5.13	5.03	4.50	6.00	9.87	5.40	4.63	4.47	5.53
0.9	10.57	4.50	5.33	5.30	5.93	9.77	4.60	3.97	4.07	4.57

Table 2b. Empirical 5% size of cointegration estimators as a function of the breakpoint (ρ_{12}=0)

Panel A

	T=100					T=300				
α=0.5	σ_{11}=1	σ_{12}=0	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	5.30	5.33	6.17	6.73	6.60	5.20	5.17	5.27	5.27	5.07
0.1	7.90	8.07	5.60	5.67	6.43	8.63	8.47	3.90	4.07	5.20
0.2	9.33	9.37	5.00	5.13	6.17	9.17	9.27	3.90	3.93	4.87
0.3	9.03	8.87	5.47	5.17	5.93	9.60	9.57	4.13	4.13	4.87
0.4	8.30	8.30	5.67	5.63	6.90	8.30	8.37	3.70	3.57	5.27
0.5	8.13	8.07	4.83	5.03	5.57	8.60	8.37	4.33	4.23	5.70
0.6	6.40	6.37	5.33	5.30	5.83	6.27	5.97	3.90	4.00	5.40
0.7	5.47	5.47	5.37	4.83	5.73	5.83	5.83	4.00	4.10	5.00
0.8	5.07	4.87	5.13	5.20	5.80	5.17	4.90	4.87	5.17	5.83
0.9	4.90	4.87	6.13	6.07	6.07	4.43	4.47	3.97	4.23	4.90

Panel B

	T=100					T=300				
α=0.1	σ_{11}=1	σ_{12}=0	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	5.30	5.33	6.17	6.73	6.60	5.20	5.17	5.27	5.27	5.07
0.1	44.60	45.20	10.93	8.63	8.97	42.07	42.70	8.33	7.60	5.90
0.2	35.30	36.53	9.67	7.97	6.93	32.83	33.43	8.30	7.87	5.73
0.3	24.70	24.80	8.47	7.33	7.50	24.53	24.63	6.87	6.77	5.50
0.4	17.57	17.87	7.77	7.03	7.20	17.07	17.07	5.23	5.13	5.23
0.5	12.93	13.07	5.80	5.77	6.10	13.17	13.13	5.23	5.07	5.73
0.6	8.70	9.20	5.27	4.97	5.87	8.47	8.03	4.17	4.37	4.73
0.7	6.90	6.70	5.07	5.10	5.83	6.70	6.80	3.77	3.63	5.03
0.8	5.37	5.33	4.83	5.00	5.53	5.77	5.80	4.43	4.60	5.67
0.9	4.83	4.97	5.73	5.80	5.60	4.73	4.73	4.03	4.27	5.03

Panel C

	T=100					T=300				
α=0.01	σ_{11}=1	σ_{12}=0	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	5.30	5.33	6.17	6.73	6.60	5.20	5.17	5.27	5.27	5.07
0.1	60.87	62.73	15.43	12.00	12.60	53.70	55.20	11.10	9.93	6.37
0.2	39.93	41.57	10.63	8.40	7.93	37.10	37.50	9.33	8.77	6.17
0.3	27.20	27.70	8.63	8.13	7.57	25.70	25.67	7.17	7.13	5.60
0.4	18.57	19.07	8.40	7.53	7.17	17.73	17.87	5.53	5.23	5.37
0.5	13.97	13.53	6.20	5.87	6.17	13.40	13.60	5.60	5.40	5.63
0.6	8.93	9.27	5.33	5.20	6.03	8.60	8.47	4.17	4.23	4.60
0.7	7.27	6.80	4.93	4.90	5.80	6.80	6.60	3.90	3.83	5.10
0.8	5.23	5.27	5.03	5.17	5.33	5.83	6.03	4.43	4.50	5.77
0.9	4.90	4.80	5.63	5.90	5.80	4.87	4.63	4.07	4.33	5.00

Table 2c. Empirical 5% size of cointegration estimators as a function of the breakpoint (ρ_{12}=-0.6)

Panel A

	T=100					T=300				
α=0.5	σ_{11}=1	σ_{12}=-0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	13.10	5.50	6.70	6.73	6.53	14.03	4.43	4.37	4.43	4.33
0.1	14.70	8.03	5.87	5.67	6.27	16.10	8.27	3.83	3.57	4.87
0.2	15.97	9.47	5.30	5.13	5.80	15.87	9.00	3.93	3.87	4.57
0.3	14.70	9.60	5.13	4.70	5.67	15.43	9.50	4.87	4.87	5.87
0.4	14.10	8.40	5.17	5.23	5.83	13.93	8.17	4.70	4.37	5.10
0.5	13.27	7.37	4.40	4.77	5.30	12.70	7.67	4.07	3.83	4.43
0.6	11.60	6.63	4.67	4.60	5.27	13.00	6.00	4.70	4.63	5.40
0.7	10.43	5.30	5.07	5.00	6.07	11.10	5.13	4.30	4.47	5.70
0.8	10.20	5.67	5.50	5.27	6.23	10.57	4.47	5.03	4.63	5.40
0.9	11.17	4.87	5.27	5.47	5.13	11.33	4.80	4.50	4.43	5.20

Panel B

	T=100					T=300				
α=0.1	σ_{11}=1	σ_{12}=-0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	13.10	5.50	6.70	6.73	6.53	14.03	4.43	4.37	4.43	4.33
0.1	47.07	44.80	10.73	9.70	10.60	45.13	43.23	9.83	8.53	6.20
0.2	36.53	35.60	9.80	8.00	7.80	35.37	33.30	8.23	7.53	6.10
0.3	27.50	25.83	8.37	7.83	8.30	28.10	24.57	7.33	6.67	6.70
0.4	21.23	17.87	7.20	7.30	6.93	20.97	17.70	6.43	5.97	6.23
0.5	17.67	13.17	6.00	5.80	6.47	15.43	11.73	4.93	4.97	4.83
0.6	12.80	8.87	5.57	5.33	5.80	14.37	8.20	4.80	4.97	5.07
0.7	11.10	7.03	5.47	5.20	6.10	11.03	5.73	4.87	4.70	5.33
0.8	9.63	5.90	5.20	5.13	5.83	9.60	5.27	4.70	4.30	5.30
0.9	10.80	5.13	5.13	5.30	5.70	10.63	5.00	4.50	4.53	5.27

Panel C

	T=100					T=300				
α=0.01	σ_{11}=1	σ_{12}=-0.6	σ_{22}=1							
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	13.10	5.50	6.70	6.73	6.53	14.03	4.43	4.37	4.43	4.33
0.1	61.63	63.00	16.33	11.93	14.47	56.80	55.90	12.60	10.53	7.23
0.2	41.27	40.87	11.17	9.17	8.93	39.17	38.10	9.03	8.47	6.37
0.3	29.37	27.40	9.00	8.27	7.97	30.10	26.70	7.87	7.47	6.83
0.4	22.03	19.30	7.57	7.57	7.10	21.70	18.40	6.27	5.97	5.97
0.5	17.63	13.77	6.20	5.73	6.50	15.50	12.27	4.87	4.87	4.37
0.6	13.30	9.27	5.70	5.27	5.77	14.57	8.50	4.73	4.90	4.97
0.7	11.10	7.10	5.57	5.40	6.03	10.93	5.87	4.77	4.93	5.23
0.8	9.27	5.90	4.97	5.00	5.97	9.70	5.30	4.57	4.17	5.20
0.9	10.87	4.93	5.17	5.27	5.67	10.63	4.97	4.53	4.53	5.23

Table 3. Empirical Distribution of t-statistics (ρ_{12}=0.6, T=300)

Panel A

Mean t						Mean t				
α=10	σ_{11}=1	σ_{12}=0.6	σ_{22}=1			α=0.1				
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	0.899	-0.018	-0.018	0.003	-0.014					
0.1	0.938	-0.022	-0.019	0.018	-0.002	0.702	0.024	-0.015	-0.065	-0.066
0.2	1.063	0.039	0.046	0.089	0.055	0.621	-0.074	-0.004	-0.023	-0.029
0.3	1.088	-0.013	-0.007	0.047	0.004	0.630	-0.029	0.000	0.004	-0.013
0.4	1.219	0.030	0.028	0.097	0.041	0.626	-0.061	-0.006	-0.001	-0.020
0.5	1.315	-0.022	-0.017	0.071	0.005	0.656	-0.019	-0.002	0.010	0.006
0.6	1.489	0.000	0.012	0.132	0.030	0.695	-0.014	-0.009	0.000	-0.012
0.7	1.677	-0.019	-0.005	0.141	0.026	0.725	0.025	0.002	0.015	-0.001
0.8	2.006	0.005	0.054	0.286	0.071	0.746	-0.001	0.003	0.018	0.007
0.9	2.468	0.009	-0.016	0.382	0.076	0.800	-0.026	-0.024	-0.007	-0.026

Panel B

Standard deviation						Standard deviation				
α=10	σ_{11}=1	σ_{12}=0.6	σ_{22}=1			α=0.1				
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	0.937	1.007	1.018	1.016	1.015	0.937	1.007	1.018	1.016	1.015
0.1	0.926	0.989	1.066	1.070	1.012	2.479	2.606	1.221	1.143	1.063
0.2	0.940	1.011	1.126	1.132	1.025	1.956	2.028	1.130	1.092	1.035
0.3	0.986	1.067	1.225	1.231	1.059	1.626	1.648	1.052	1.033	1.003
0.4	1.042	1.173	1.336	1.339	1.098	1.398	1.444	1.025	1.014	1.021
0.5	1.150	1.258	1.408	1.413	1.119	1.240	1.299	0.983	0.976	1.019
0.6	1.273	1.421	1.563	1.576	1.185	1.094	1.146	0.950	0.945	0.999
0.7	1.455	1.661	1.786	1.774	1.285	1.009	1.064	0.943	0.939	0.995
0.8	1.718	1.987	2.079	2.051	1.383	0.948	1.006	0.965	0.964	1.020
0.9	2.317	2.591	2.512	2.448	1.394	0.910	0.975	0.958	0.958	0.988

Panel C

Skewness						Skewness				
α=10	σ_{11}=1	σ_{12}=0.6	σ_{22}=1			α=0.1				
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	-0.011	-0.018	-0.004	-0.022	-0.015	-0.011	-0.018	-0.004	-0.022	-0.015
0.1	0.103	-0.014	-0.012	-0.026	-0.023	-0.351	0.009	-0.305	-0.389	-0.063
0.2	0.069	-0.090	-0.133	-0.130	-0.144	-0.125	-0.003	-0.154	-0.159	-0.054
0.3	0.280	0.044	0.076	0.053	0.056	-0.025	-0.040	0.087	0.068	0.024
0.4	0.187	-0.040	-0.018	-0.040	-0.049	-0.034	-0.062	-0.010	0.027	0.006
0.5	0.156	0.001	0.030	-0.004	0.011	-0.009	-0.128	-0.194	-0.174	-0.103
0.6	0.116	-0.050	-0.027	-0.040	-0.067	0.046	-0.035	-0.036	-0.052	-0.045
0.7	0.169	-0.018	0.068	0.047	0.011	-0.031	0.022	-0.029	-0.024	0.007
0.8	0.079	0.037	0.104	0.078	0.212	-0.030	0.065	0.038	0.030	0.067
0.9	0.214	0.017	0.072	0.240	0.095	-0.093	-0.024	-0.037	-0.025	-0.046

Table 3 (continued)

Panel D

Kurtosis						Kurtosis				
α=10	σ_{11}=1	σ_{12}=0.6	σ_{22}=1			α=0.1				
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	2.942	2.987	3.070	3.063	3.086	2.942	2.987	3.070	3.063	3.086
0.1	2.875	3.029	3.113	3.121	3.081	3.421	3.154	8.308	7.554	3.045
0.2	3.174	3.163	3.075	3.104	3.147	3.082	3.166	5.891	5.690	3.184
0.3	3.107	3.115	3.063	3.112	3.083	3.108	3.055	5.257	5.238	3.350
0.4	3.163	3.232	3.098	3.068	3.213	3.141	3.035	4.519	4.421	3.161
0.5	2.956	3.072	3.052	2.997	3.264	3.027	3.263	3.912	3.891	3.135
0.6	3.255	3.104	3.075	3.125	3.406	3.231	3.065	3.764	3.667	2.940
0.7	3.306	3.256	3.307	3.261	3.708	3.455	3.186	3.438	3.406	3.203
0.8	3.182	3.180	3.402	3.456	4.752	3.317	3.065	3.101	3.068	2.943
0.9	3.486	3.397	3.906	3.809	4.287	3.234	3.154	3.191	3.160	3.027

Panel E

$t_{0.025}$						$t_{0.025}$				
α=10	σ_{11}=1	σ_{12}=0.6	σ_{22}=1			α=0.1				
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	-0.953	-2.003	-2.009	-1.985	-1.963	-0.953	-2.003	-2.009	-1.985	-1.963
0.1	-0.810	-1.943	-2.123	-2.077	-2.007	-4.584	-5.161	-2.638	-2.542	-2.214
0.2	-0.785	-1.933	-2.240	-2.240	-2.066	-3.245	-4.154	-2.443	-2.331	-2.076
0.3	-0.686	-2.092	-2.364	-2.359	-2.047	-2.500	-3.283	-2.187	-2.150	-1.953
0.4	-0.727	-2.306	-2.597	-2.505	-2.170	-2.170	-2.965	-2.095	-2.017	-2.005
0.5	-0.848	-2.471	-2.731	-2.701	-2.153	-1.761	-2.657	-2.029	-2.015	-2.053
0.6	-0.936	-2.910	-3.057	-3.024	-2.309	-1.447	-2.343	-1.897	-1.907	-1.981
0.7	-1.100	-3.309	-3.530	-3.341	-2.562	-1.332	-2.080	-1.999	-1.937	-1.979
0.8	-1.375	-3.875	-3.970	-3.799	-2.673	-1.129	-1.931	-1.896	-1.881	-1.963
0.9	-1.878	-5.139	-5.003	-4.421	-2.624	-1.023	-1.907	-1.910	-1.888	-1.956

Panel F

$t_{0.975}$						$t_{0.975}$				
α=10	σ_{11}=1	σ_{12}=0.6	σ_{22}=1			α=0.1				
s	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b	t_{OLS}	t_{DOLS}	t_{JOH}	t_{FMLS}	t_b
0.0	2.752	1.912	1.973	1.923	1.935	2.752	1.912	1.973	1.923	1.935
0.1	2.802	1.911	2.042	2.059	1.918	5.226	5.315	2.441	2.176	1.955
0.2	2.926	2.006	2.121	2.191	1.930	4.304	3.935	2.397	2.268	1.948
0.3	3.168	2.100	2.395	2.475	2.101	3.828	3.132	2.225	2.222	1.937
0.4	3.388	2.343	2.613	2.652	2.183	3.414	2.776	2.156	2.135	2.020
0.5	3.659	2.453	2.763	2.829	2.232	3.131	2.492	1.934	1.917	1.942
0.6	4.044	2.728	3.049	3.147	2.359	2.844	2.217	1.828	1.839	1.956
0.7	4.693	3.315	3.590	3.756	2.694	2.759	2.132	1.899	1.917	2.001
0.8	5.492	3.905	4.383	4.449	2.932	2.597	2.025	1.956	1.926	2.047
0.9	7.495	5.223	5.041	5.484	2.934	2.604	1.925	1.873	1.867	1.866

Figure 1. Distribution of t-statistic of cointegration estimators

(α=10, ρ_{12}=0.6, s=0.9, T=1000)

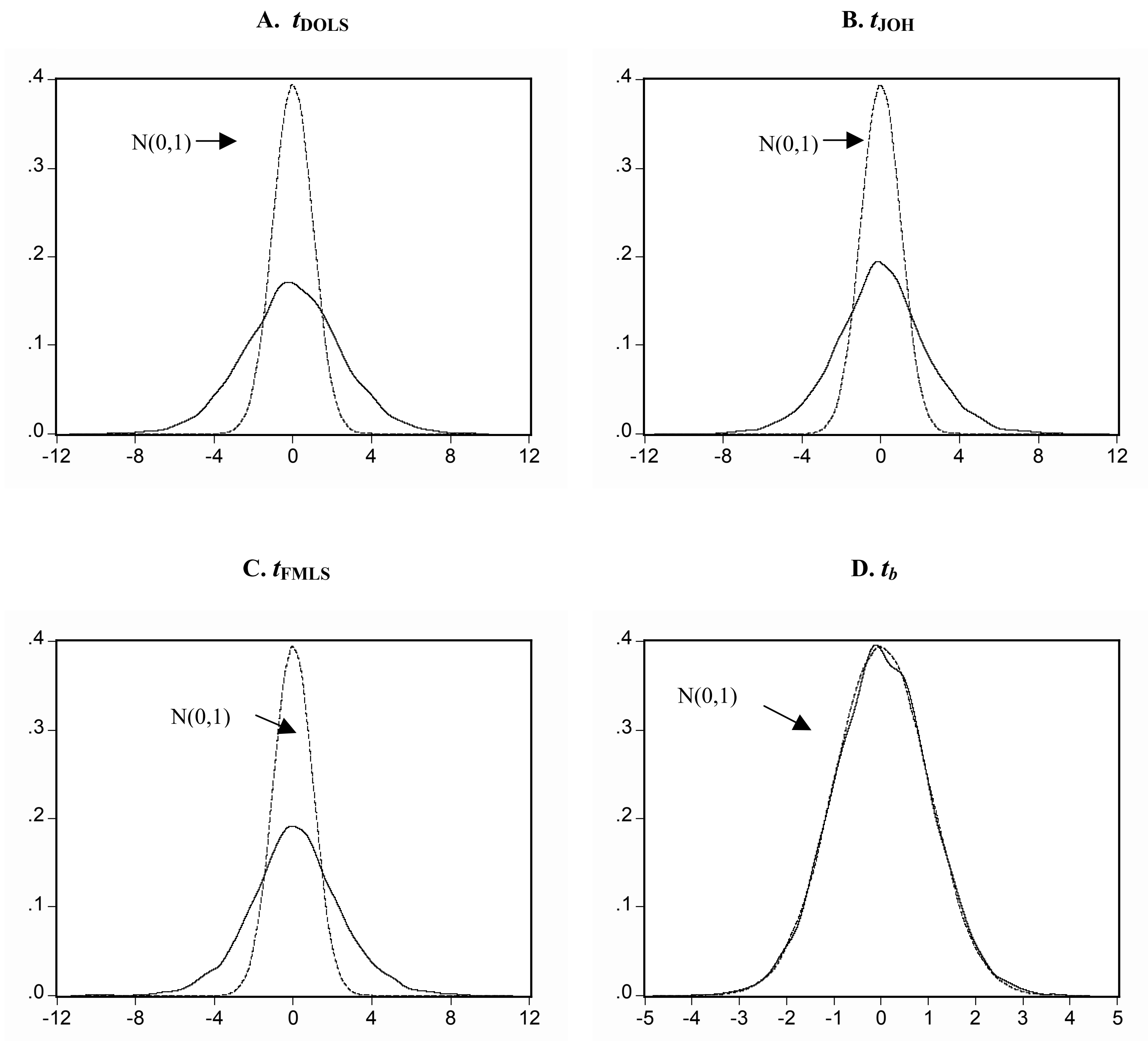

Figure 2. Distribution of t-statistic of cointegration estimators

(α=0.1, ρ_{12}=0.6, s=0.1, T=1000)