Multicointegration, polynomial cointegration and I(2) cointegration with structural breaks. An application to the sustainability of the US external deficit[1]

By Vanessa Berenguer-Rico and Josep Lluís Carrion-i-Silvestre

AQR-IREA
Departament d'Econometria, Estadística i Economia Espanyola
Universitat de Barcelona

Abstract
In this paper we model the multicointegration relation, allowing for one structural break. Since multicointegration is a particular case of polynomial or I(2) cointegration, our proposal can also be applied in these cases. The paper proposes the use of a residual-based Dickey-Fuller class of statistic that accounts for one known or unknown structural break. Finite sample performance of the proposed statistic is investigated by using Monte Carlo simulations, which reveals that the statistic shows good properties in terms of empirical size and power. We complete the study with an empirical application of the sustainability of the US external deficit. Contrary to existing evidence, the consideration of one structural break leads to conclude in favour of the sustainability of the US external deficit.

Keywords: I(2) processes, multicointegration, polynomial cointegration, structural break, sustainability of external deficit

JEL Codes: C12, C22

1 The authors gratefully acknowledge the financial support of the Ministerio de Ciencia y Tecnología from the grant SEJ2005-08646/ECON. AQR-IREA Research Group. Department of Econometrics, Statistics and Spanish Economy. University of Barcelona. Av. Diagonal, 690. 08034 Barcelona. Tel: +34 934021826, fax: +34 934021821, and e-mails: vberenguer@ub.edu and carrion@ub.edu.

1 Introduction

It is generally true that any linear combination of I(1) series is also I(1) by the dominant property of the stochastic trends. However, it is possible that linear combinations of I(1) stochastic processes generate an I(0) time series, in which case the time series are said to be cointegrated in the sense of Engle and Granger (1987) – hereafter, the so-called first level of cointegration. Moreover, it could also be possible that the cumulated cointegrated residuals, I(1) by definition, cointegrate with the I(1) original variables. In this case, a deeper level of cointegration – henceforth, second level of cointegration – occurs between the original time series. Granger and Lee (1989) denote this sort of cointegration as "multicointegration" and show that it is a long-run relationship that might be expected to occur in economics. One appealing example is given by the consumption life-cycle hypothesis. If income and consumption are I(1) variables the economic theory establishes that saving must be stationary, so that income and consumption have to be cointegrated. Furthermore, if we cumulate saving, which can be seen as the stock of consumers' wealth, the life-cycle hypothesis predicts that wealth must cointegrate with consumption – see Siliverstovs (2003) for an application of multicointegration to the analysis of the life-cycle hypothesis of consumption for the US. Other relevant applications can be found in the literature including the analysis of production and sales – Granger and Lee (1989) – housing units started and new housing units completed – see Lee (1992) and Engsted and Haldrup (1999) – the Ricardian Equivalence – see Leachman (1996) – the twin deficits hypothesis – see Leachman and Francis (2002) – and the sustainability of fiscal practices – see Leachman, Bester, Rosas and Lange (2005). The common feature that all these empirical applications share is that they involve stock-flow relationships among economic variables.

The current definition of multicointegration assumes invariant parameters. However, the longer the time period the higher the probability that structural breaks appear affecting the behaviour of time series. Nowadays, it is well known that this feature may cause bias in the statistical inference that is obtained when dealing with time series – Perron (1989). To this end, unit root and stationarity tests have been designed to accommodate the presence of structural breaks. Besides, cointegration testing procedures have also been devised to account for the presence of instability either in the deterministic and/or in the cointegrating vector of the long-run relationship. Perron (2006) offers a comprehensive overview of the field.

As described above, the concept of multicointegration is naturally derived from the definition of cointegration, which in turn might be subject to the effects of structural breaks – see Gregory and Hansen (1996). Therefore, the standard framework of analysis defined in Granger and Lee (1989), and Engsted, Gonzalo and Haldrup (1997) should be extended to accommodate the presence of structural breaks. In this paper we have addressed this issue dealing with the different specifications that may arise depending on whether the structural break affects only the first and/or the second level of cointegration. Once the models are defined, we propose to test the null hypothesis of non-multicointegration without structural break against the alternative hypothesis of multicointegration with a structural break using a residual-based Dickey-Fuller class statistic. It is worth mentioning that though the set-up that is used in this paper is bases on the multicointegration framework, one appealing feature of our proposal is that it can be used to analyze long-run relationships among I(2) stochastic processes, provided that multicointegration can be understood as I(2) cointegration or polynomial cointegration. To the best of our knowledge, this has not been previously addressed in the literature. The paper provides different sets of critical values depending on the number and nature of the stochastic regressors involved in the model and on the deterministic specification that is used.

Our proposal is illustrated through the investigation of the sustainability of the external deficit in the US. This topic has attracted the attention of economists both from a theoretical and empirical point of view, especially in recent times where the external deficit of the US economy has been increasing. Previous empirical approaches in Leachman and Francis (2000, 2002) have used the multicointegration framework to analyze the relationship between flow variables (exports and imports) and the stock variable (foreign debt). However, none of them has considered the presence of structural breaks in the analysis, which is shown to lead to misleading conclusions. The statistics that have been applied in this paper allow us to conclude that the US time series of exports and imports can be characterized as I(1) stochastic processes, which have been affected by one structural break. When we account for the presence of one unknown structural break in the analysis, we find evidence that supports the existence of multicointegration relationships among the US exports and imports, which contradicts previous conclusions in the literature.

The paper is organized as follows. Section 2 describes the models that are specified depending on the effect that the structural breaks have on the parameters of the different components of the models. Section 3 describes the statistical procedure to test the null hypothesis of non-multicointegration against the alternative hypothesis of multicointegration with one structural break. The limiting distribution of the statistic is derived and critical values are provided. Section 4 focuses on the finite sample performance of the proposal. To be specific, we analyze the empirical size and power of the test statistic, as well as the use of information criteria to select among the different specifications that have been proposed in the paper. In Section 5 we illustrate the application of the proposal focusing on the sustainability of the external deficit in the US. Finally, Section 6 presents some concluding remarks. All derivations are collected in the Appendix.

2 Multicointegration with structural break

Cointegration is a necessary condition for the presence of multicointegration as defined in Granger and Lee (1989). Thus, if we consider one dimensional time series $\{y_t\}_0^n$ and m_1-dimensional time series $\{x_t\}_0^n$ all being I(1) non-stationary stochastic processes, these variables are assumed to satisfy the following standard cointegration model:

$$y_t = c_t'\alpha + x_t'\beta + \vartheta_t, \tag{1}$$

$t = 1, \ldots, n$, where c_t is an s_0-dimensional deterministic sequence of general form – typically, $c_t = 0$, $c_t = 1$ and $c_t = (1, t)'$ – and where $\{\vartheta_t\}_0^n$ is an I(0) series. Suppose that the cumulated cointegration residuals, $S_t = \sum_{j=1}^t \vartheta_j$, cointegrate with either $\{y_t\}_0^n$

and/or $\{x_t\}_0^n$, in which case we obtain the standard multicointegration model, that is

$$S_t = m_t'\zeta + x_t'\gamma + u_t, \tag{2}$$

where m_t is the s_1-dimensional deterministic sequence and where $\{u_t\}_0^n$ is an I(0) series.

Following Engsted, Gonzalo and Haldrup (1997), we can write (2) as:

$$Y_t = Cm_t'\mu + x_t'\gamma + X_t'\beta + u_t, \tag{3}$$

where $Y_t = \sum_{j=1}^t y_j$ and $X_t = \sum_{j=1}^t x_j$ are I(2) variables, and $Cm_t = \left(\sum_{j=1}^t c_j', m_t'\right)'$ is the new m_0-deterministic component associated to multicointegration relation (3), with $m_0 = s_0 + s_1$ and $\mu = (\alpha', \zeta')'$. The specification given by (3) can be written using the Phillips' triangular representation:

$$\begin{aligned} Y_t &= Cm_t'\mu_0 + Y_t^0; \qquad \Delta^2 Y_t^0 = v_t && (4)\\ x_t &= Cm_t'\mu_1 + x_t^0; \qquad \Delta x_t^0 = \varepsilon_{1t} && (5)\\ X_t &= Cm_t'\mu_2 + X_t^0; \qquad \Delta^2 X_t^0 = \varepsilon_{2t}, && (6) \end{aligned}$$

where x_t^0, X_t^0 are m_1 and m_2-dimensional I(1) and I(2) processes, respectively, and where the $w_t = (v_t, \varepsilon_{1t}', \varepsilon_{2t}')'$ stochastic processes involved in the definition of the data-generating process (DGP) are assumed to be a strong-mixing sequence satisfying the multivariate invariance principle in Phillips and Durlauf (1986). We partition Ω conformably with w_t, so that

$$\Omega = \begin{bmatrix} \omega_{00} & \omega_{01} & \omega_{02} \\ \omega_{10} & \Omega_{11} & \Omega_{12} \\ \omega_{20} & \Omega_{21} & \Omega_{22} \end{bmatrix} = \Sigma + \Lambda + \Lambda', \tag{7}$$

where $\Sigma = E(w_1 w_1')$ and $\Lambda = \sum_{k=2}^{\infty} E(w_1 w_k')$. For subsequent use we also define

$$\Delta = \Sigma + \Lambda, \tag{8}$$

which can be conveniently partitioned as Ω. In (7) the diagonal submatrices Ω_{11} and Ω_{22} are assumed to be positive definite such that x_t^0 and X_t^0 are not permitted to be individually cointegrated.

The component Cm_t in (3) to (6) denotes the deterministic part of the model. Haldrup (1994) and Engsted, Gonzalo and Haldrup (1997) define the standard multicointegration case without structural breaks through three different specifications: (i) $Cm_t = 1$ – hereafter, Model 1 – for the constant case, (ii) $Cm_t = (1, t)'$ – denoted as Model 2 – for the time trend case, and (iii) $Cm_t = (1, t, t^2)'$ – henceforth, Model 3 – for the quadratic time trend case. The definition of this deterministic component constitutes the first way in which we can introduce the presence of the structural break in the model. Thus, the most general situation that is considered in this paper defines the vector of deterministic regressors as $Cm_t = (1, t, t^2, DU_t, DT_t^*)'$ where $DU_t = 1(t > T_b)$ and $DT_t^* = (t - T_b) 1(t > T_b)$, with $1(t > T_b)$ the indicator function, $T_b = [\lambda n]$ denoting the break point, λ the break fraction parameter, $\lambda \in \Lambda$, with Λ a closed subset of $(0, 1)$, and $[\cdot]$ being the integer part. Note that the inclusion of the DU_t dummy variable captures the presence of a level shift, while DT_t^* aims to capture the slope shift.

The second way in which the structural break can enter in the model is through the stochastic part. Thus, the Y_t^0 in (4) is one-dimensional I(2) stochastic process linked to x_t^0 and X_t^0 through

$$Y_t^0 - x_t^{0\prime}\gamma_1 - x_t^{0\prime}\gamma_2 1(t > T_b) - X_t^{0\prime}\beta_1 - X_t^{0\prime}\beta_2 1(t > T_b) = u_t, \tag{9}$$

where the processes x_t^0, X_t^0, Y_t^0 are initialized at $t = 1, 0, 0$, respectively, without loss of generality. Note that we can only allow the cointegrating vector at the first level to change setting $\beta_2 \neq 0$ and $\gamma_2 = 0$. Similarly, we can allow the multicointegrating vector to change, but not the first level cointegrating vector, through the specification of $\gamma_2 \neq 0$ and $\beta_2 = 0$. Finally, the vectors of the two levels of cointegration are allowed to change if $\gamma_2 \neq 0$ and $\beta_2 \neq 0$.

Table 1: Model specification

	Deterministic part	Stochastic part
Model 1	$Cm_t = 1$	$\gamma_2 = \beta_2 = 0$
Model 2	$Cm_t = (1, t)'$	$\gamma_2 = \beta_2 = 0$
Model 3	$Cm_t = (1, t, t^2)'$	$\gamma_2 = \beta_2 = 0$
Model 4	$Cm_t = (1, t, DU_t, DT_t^*)'$	$\gamma_2 = \beta_2 = 0$
Model 5	$Cm_t = (1, t, t^2, DU_t, DT_t^*)'$	$\gamma_2 = \beta_2 = 0$
Model 6	$Cm_t = (1, t, DU_t, DT_t^*)'$	$\gamma_2 = 0$ and $\beta_2 \neq 0$
Model 7	$Cm_t = (1, t, DU_t, DT_t^*)'$	$\gamma_2 \neq 0$ and $\beta_2 = 0$
Model 8	$Cm_t = (1, t, t^2, DU_t, DT_t^*)'$	$\gamma_2 \neq \beta_2 \neq 0$

Taken together, we can define the general conditional model:

$$\begin{aligned} Y_t &= Cm_t'\mu + x_t'\gamma_1 + x_t'\gamma_2 1\,(t > T_b) + X_t'\beta_1 + X_t'\beta_2 1\,(t > T_b) + u_t \qquad (10) \\ &= X_t^m\,(\lambda)'\,\theta + u_t \qquad (11) \\ \Delta^d u_t &= v_t, \qquad (12) \end{aligned}$$

which allows the specification of several models depending on the effect of the structural break. Table 1 summarizes the specifications for the standard multicointegration framework of analysis without structural breaks, as well as those considered in this paper that account for one structural break. As mentioned above, the specifications given by Models 1 to 3 are those proposed in Engsted, Gonzalo and Haldrup (1997). Models 4 and 5 introduce the presence of the structural break only through the deterministic component, so that the cointegration vectors of the two levels remain unchanged. Model 6 deals with changes both in the deterministic component and in the cointegrating vector of the first level, but not in the cointegrating vector of the second level. Model 7 controls for those breaks affecting both the deterministic component and the cointegrating vector of the second level, but not the one of the first level of cointegration. Finally, Model 8 is the most general specification, which allows for effects on all the parameters of the model.

There are in this I(2) system several cointegration possibilities depending on the order of integration of u_t in (10), i.e. $\Delta^d u_t = v_t$ with $d = 0, 1, 2$. When $d = 2$ neither cointegration nor multicointegration exist since there are not any common stochastic trends – i.e. u_t process is integrated of order two. When $d = 1$ there is cointegration only at the first level. Finally, when $d = 0$ we conclude that the variables y_t and x_t are multicointegrated in such a way that all stochastic trends are cancelled in the multicointegration relationship. The next section proposes the statistic to test the presence of multicointegration accounting for the presence of one structural break. As mentioned above, the proposal generalizes the analysis of Haldrup (1994) and Engsted, Gonzalo and Haldrup (1997) to test for either polynomial cointegration or cointegration among I(2) variables where the long-run elements are permitted to change during the sample period.

3 Testing for multicointegration with structural break

We can test the null hypothesis of non-multicointegration against the alternative hypothesis of multicointegration with structural breaks using a residual-based augmented Dickey-Fuller (ADF) class of test statistics. It is worth noticing that, as in Gregory and Hansen (1996), our set-up permits the presence of one structural break only under the alternative hypothesis. As pointed out in Haldrup (1994), in many situations it is likely that cointegration to at least I(1) level will occur, which leads us to test the null hypothesis of cointegration at the first level, i.e. $u_t \sim I(1)$, against the alternative hypothesis of multicointegration with structural breaks, $u_t \sim I(0)$. In order to assess the integration order for $\hat{u}_t$, we estimate the ADF-type regression,

$$\Delta \hat{u}_t = \delta \hat{u}_{t-1} + \sum_{j=1}^{p} \varphi_j \Delta \hat{u}_{t-j} + \eta_t \tag{13}$$

and consider the t-ratio ADF statistic $t_\delta(\lambda)$ for testing the null hypothesis that $\delta = 0$ computed from the OLS estimation of (13). Theorem 1 presents the limiting distribution of the ADF statistic for the known break case.

Theorem 1 *Let Y_t be generated according to (10) and (12) with $d = 1$. Then for $n \to \infty$, $T_b \to \infty$ in a way that $\lambda = T_b/n$ remains constant, and with $p = O_p(n^{1/3})$, the $t_\delta(\lambda)$ test statistic computed from the estimated regression (13) converges to*

$$t_\delta(\lambda) \Rightarrow \left(\int_0^1 W^{*2}(r, \lambda) \right)^{-1/2} \left(\int_0^1 W^*(\lambda, r)\, dW^*(r, \lambda) \right)$$

where $\Rightarrow$ denotes weak convergence of the associated probability measure on the unit interval [0,1], and $W^*(\lambda, r) = W_0(r) - W(\lambda, r)' \left(\int_0^1 W(\lambda, r)' W(\lambda, r)\, dr\right)^{-1} \left(\int_0^1 W(\lambda, r)' W_0(r)\, dr\right)$, *with $W_0(r)$ a standard Brownian motion and $W(\lambda, r)$ a vector of the limit of the elements that define the deterministic component, Brownian motions and integrated Brownian motions defined in the Appendix.*

See the Appendix for the outline of the proof. As can be seen from Theorem 1 the limiting distribution of the ADF statistic depends on the deterministic component, on the number of I(1) and I(2) stochastic regressors involved in the model – m_1 and m_2 respectively – and on the break fraction (λ) nuisance parameter. Asymptotic and finite sample critical values for different combinations of m_1, m_2 and λ values have been computed, although they are not reported to save space – they are available upon request.

From an empirical point of view, it would be more interesting to consider the situation in which the structural break is unknown. In order to deal with this case we follow Gregory and Hansen's (1996) methodology. In brief, the approach proceeds in two stages. First, the multicointegration ADF test is computed for each possible break point, $\lambda \in \Lambda$, which defines a sequence of statistics. Second, the infimum of the sequence of ADF statistics is taken, which defines the statistic:

$$t_\delta^*(\lambda) = \inf_{\lambda \in \Lambda} t_\delta(\lambda),$$

where $\hat{T}_b = \arg\min_{\lambda \in \Lambda} t_\delta^*(\lambda)$. The following Theorem provides the limiting distribution of the $t_\delta^*(\lambda)$ statistic.

Theorem 2 *Let Y_t be generated according to (10) and (12) with $d = 1$. Under the null hypothesis of non-multicointegration, the $t_\delta^*(\lambda)$ statistic converges to*

$$t_\delta^*(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left[\left(\int_0^1 W^{*2}(r, \lambda) \right)^{-1/2} \left(\int_0^1 W^*(\lambda, r)\, dW^*(r, \lambda) \right) \right].$$

The proof is given in the Appendix. As above, the distribution depends upon the deterministic component, m_1 and m_2, but not on the break fraction parameter. Tables

2 to 4 report critical values for Models 4 to 8 when $m_1 = \{1,2,3,4\}$ and $m_2 = \{1,2\}$, for four sample sizes $n = \{50, 100, 250, 500\}$. To be specific, we have followed Haldrup (1994) and Engsted, Gonzalo and Haldrup (1997) and define $(m_2 + 1)$ $I(2)$ stochastic processes – one for the endogenous variable and m_2 for the regressors – using partial sum of partial sum of iid $N(0,1)$, whereas for the m_1 $I(1)$ stochastic processes we have used using partial sum of iid $N(0,1)$. The order (p) of the parametric correction in (13) is selected using the t-sig criterion in Ng and Perron (1995) with $p_{\max} = 6$ as the maximum number of lags. The simulations were based upon 5,000 replications and, following Zivot and Andrews (1992), and Gregory and Hansen (1996), we define $\Lambda = [[0.15n], [0.85n]]$.

4 Finite sample performance

This section offers the Monte Carlo results concerning the performance of the statistic proposed in this paper. The results are organized in two subsections. First of all, we analyze the empirical size and power for the standard multicointegration case that does not consider the presence of structural breaks. The study of the standard multicointegration framework is interesting since, to the best of our knowledge, it has not been previously investigated in the literature. Furthermore, it constitutes the benchmark for the following analysis. Second, we focus on the statistic that has been designed to accommodate the presence of one unknown structural break. After the empirical size and power have been analyzed, we essay the use of the Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) as instruments to select the type of structural break, i.e. the model specification that better captures the effects of the structural break.

4.1 Standard multicointegration without structural break

The design of the experiments bases on previous contributions in the literature. Thus, following Kremers, Ericsson and Dolado (1992), we use the common factor representation for multicointegrated processes to obtain the empirical size and power properties of the statistics designed in Granger and Lee (1989), and Engsted, Gonzalo and Haldrup (1997).

Suppose that x_t, y_t are I(1) cointegrated processes so that $z_t = x_t - Ay_t \sim I(0)$. The standard common factor representation for cointegrated processes is given by:

$$x_t = AW_t + \eta_{1t} \qquad y_t = W_t + \eta_{2t}, \tag{14}$$

where W_t is an I(1) process, and η_{1t}, η_{2t} are both I(0) processes. Using this kind of representation for multicointegrated processes we have

$$x_t = AW_t + \alpha_1 \Delta W_t + \xi_{1t} \qquad y_t = W_t + \alpha_2 \Delta W_t + \xi_{2t}, \tag{15}$$

where ξ_{1t}, ξ_{2t} are both I(-1), with A being a constant. Note that we can go from the null hypothesis of non-multicointegration to the alternative hypothesis of multicointegration depending on the values of α_1 and α_2, and/or the definition of the ξ_{1t} and ξ_{2t} processes. Thus, note that we are under the null hypothesis of non-multicointegration when $\alpha_1 = \alpha_2 = 0$, while the alternative hypothesis of multicointegration is obtained when $\alpha_1 \neq \alpha_2 \neq 0$. Furthermore, if we define

$$\xi_{1t} = \Delta v_t = \Delta(\rho_1 v_{t-1} + \varepsilon_{1t}); \qquad \xi_{2t} = \Delta w_t = \Delta(\rho_2 w_{t-1} + \varepsilon_{2t}),$$

where $\varepsilon_{1t}, \varepsilon_{2t}$ are both I(0), then, we are under the null hypothesis of non-multicointegration when $\rho_1 = \rho_2 = 1$ – so that $\xi_{1t} \sim I(0)$ and $\xi_{2t} \sim I(0)$ – while the alternative hypothesis of multicointegration is achieved when $\rho_1, \rho_2 < 1$ – so that $\xi_{1t} \sim I(-1)$ and $\xi_{2t} \sim I(-1)$.

Let us first focus on the two-step procedure proposed by Granger and Lee (1989) using the following model:

$$x_t = DS_t + u_t, \tag{16}$$

which does not include any deterministic component and where D denotes a constant. Table 5 reports rejection frequencies for $n = \{50, 100\}$ using the critical values corresponding to the 5% level of significance drawn from MacKinnon (1991). Throughout the paper, the number of replications is set at 10,000. As can be seen from Table 5, the empirical size, i.e. $\alpha_1 = \alpha_2 = 0$ and $\rho_1 = \rho_2 = 1$, is close to the nominal one. As expected, the empirical power of the statistic increases with the sample size regardless of the combination of the values of α_i and ρ_i, $i = 1, 2$. Furthermore, note that when ρ_i remains fixed, the empirical power of the statistic increases with α_i, $i = 1, 2$. Finally, for fixed α_i the power increases when the ρ_i parameters move away from the null hypothesis, $i = 1, 2$.

Table 5 also reports the results corresponding to the one-step procedure in Engsted, Gonzalo and Haldrup (1997) using the following model:

$$\textstyle\sum_{j=1}^{t} x_j = Cm_t'\mu + A\sum_{j=1}^{t} y_j + Dy_t + u_t, \tag{17}$$

with two deterministic component specifications, i.e. $Cm_t = (1, t)'$ and $Cm_t = (1, t, t^2)'$, where x_t and y_t are generated according to (15) using the same set of parameters as above. As can be seen, the empirical size of the statistic is close to the nominal one, regardless of the deterministic specification that is used. Compared to the two step-procedure of Granger and Lee (1989), the empirical power of the one-step based statistic in Engsted, Gonzalo and Haldrup (1997) for fixed ρ_i is invariant to α_i, $i = 1, 2$, due to the implicit common factor restriction that is imposed by the ADF test. In order to understand the reason for this different behaviour we can think of the common factor representation for multicointegrated processes.

Note that Granger and Lee (1989) impose the condition $(\alpha_1 - A\ \alpha_2) \neq 0$ to ensure the presence of multicointegration. However, when $\alpha_i = 0$, $i = 1, 2$, $X_t \left(= \sum_{j=1}^{t} x_j\right)$ and $Y_t \left(= \sum_{j=1}^{t} y_j\right)$ polynomially cointegrate by construction, although x_t and y_t do not multicointegrate. That is the reason why the one-step procedure shows higher power than the two-step procedure for any value of α_i, $i = 1, 2$. In fact, when $\alpha_i = 0$, $i = 1, 2$, the common factor restriction is satisfied by (16), but not by (17). To see this, we can plug the common factor representation in (16) and (17) when $\alpha_i = 0$, $i = 1, 2$,

$$(AW_t + \xi_{1t}) = D\sum_{j=1}^{t} \xi_{1j} - DA\sum_{j=1}^{t} \xi_{2j} + u_t, \tag{18}$$

$$\sum_{j=1}^{t} AW_j + \sum_{j=1}^{t} \xi_{1j} = A\sum_{j=1}^{t} W_j + \sum_{j=1}^{t} \xi_{2j} + D\left(W_t + \xi_{2t}\right) + u_t. \tag{19}$$

Since the estimator of D is super-consistent and converges to zero – note that there is not multicointegration in this case – the common factor restriction holds in (18), but not in (19). Note that we always run the risk of finding evidence in favour of multicointegration when in fact there might be polynomial cointegration by construction, but not multicointegration. In order to detect this kind of *spurious* multicointegration associated with the one-step procedure we suggest practitioners to test the significance of D in (17).

4.2 Multicointegration with structural break

The DGP that is used is given by:

$$Y_t = X_t^m\left(\lambda\right)' \theta + u_t,$$

where $X_t^m\left(\lambda\right) = \left(1, t, t^2, DU_t, DT_t^*, x_t', x_t'1\left(t > T_b\right), X_t', X_t'1\left(t > T_b\right)\right)'$, with x_t and y_t generated according to (15). As above, the empirical size is investigated using $(\alpha_i, \rho_i) = (0, 1)$, $i = 1, 2$. As for the empirical power analysis, we only report results for $(\alpha_i, \rho_i) = (0.5, 0.9)$, $i = 1, 2$, since conclusions do not change for the other configurations that have been essayed. To analyze the effect of the magnitude of the structural break on the power of the test we have considered two sets of parameters. The first set of parameters accounts for small effects of the structural break, i.e. $\theta = \theta_1 = (1,\ 0.02,\ 0.01,\ 3,\ 0.06,\ 1.5,\ 1,\ 2,\ 0.1)'$, while the second one is for large effects of the structural breaks, i.e. $\theta = \theta_2 = (1,\ 0.02,\ 0.01,\ 9,\ 0.18,\ 1.5,\ 3,\ 2,\ 0.3)'$. Note that this specification of the DGP corresponds to the most general model considered in the paper – Model 8 – so that the other specifications can be obtained as particular cases. Finally, the design of the experiment considers $\lambda = \{0.25, 0.50, 0.75\}$ to investigate the effect of the break point position on the empirical power.

As can be seen from Table 6, the $t_\delta^*\left(\lambda\right)$ statistic has the correct size in all cases, since the empirical size is close to the nominal one. Concerning the empirical power of the

$t_{\delta}^{*}(\lambda)$ statistic, we can see that the power increases with sample size, the magnitude of the structural break, the break point position, and the type of the break. The increase of the power when the magnitude of the break point increases is to be expected, since in this case it is easier to detect. The fact that the empirical power increases as the break point moves away from the beginning of the time period has also been pointed out in earlier literature – see Gregory and Hansen (1996). For $n = 50$ the statistic shows low power, specially for those cases where the magnitude of the break is small and the break point is located at the beginning of the period. Nevertheless, these values are similar to others obtained in the literature – see, Engle and Granger (1987), among others. Results not reported here, indicate that when ρ_i moves away from one the power of the statistic increases – for instance, for $(\alpha_i, \rho_i) = (0.5, 0)$, $i = 1, 2$, the power is around 0.95.

To sum up, simulation experiments indicate that the test statistics proposed in this paper show good properties in terms of empirical size and power, with values that resemble those for the standard multicointegration framework.

4.2.1 Model selection

Throughout the paper we have proposed several models to capture the different way in which the structural break can affect the components of the model under the alternative hypothesis of multicointegration. Unless there are strong analyst's priors, practitioners may doubt about which of these models or types of break to use. Consequently and from an empirical point of view, it is interesting to think of the way to determine the most appropriate specification of the type of break. To this end, and following Calvo, Montañés and Olloqui (2005) for the unit root tests with structural breaks, we analyze in this section the performance of the AIC and BIC information criteria when used as instruments to select from the different specifications that have been proposed. For each of the five DGP's that we have considered in the paper, we have estimated the five models and recorded the relative frequencies of the selected model. The design of the Monte Carlo experiment is the same as the one described above for the power analysis with one structural break.

The results of the experiment are presented in Tables 7 and 8 for both the small (θ_1) and large (θ_2) break cases respectively. The column labelled as DGP indicates the true specification, while the columns labelled as Models 4 to 8 indicate the model specification that is estimated. As expected, the larger the magnitude of the structural break, the better the classification that is obtained using the AIC and BIC information criteria. Furthermore, when the magnitude of the structural break is given by θ_2 we only find one error in the classification of the models using the AIC information criterion, and four using the BIC information criterion – the errors are remarked in bold in Table 8. When both the magnitude of the break and the sample size are small, the number of incorrect classifications increases regardless of the information criteria that is used – see bold numbers in Table 7. However, we can see that in these cases the BIC information criterion always tends to select more parsimonious models than the true ones, whereas the AIC information criterion under-specifies the true model in six out of the ten situations of incorrect classifications.

All in all, these results show that AIC and BIC can be used to select from the models that have been proposed in this paper, since they provide good guidance, specially in those cases where both the sample size and the magnitude of the structural break are large. However, in small samples the AIC information criterion has been shown to be more conservative than the BIC one, since the misspecification errors, when they exist, do not necessarily point to parsimonious specifications.

5 The unstable nature of the sustainability of external deficit in the US

In this section we analyze the existence of changes in the determination of intertemporal budgeting of the external sector and sustainability of external deficit using the procedures developed above. From a theoretical point of view, intertemporal external budget constraints have been introduced in models with open economies. Macroeconomic accounting identities establish that the current account (CA) is equal to exports (E) less imports (M) plus net remittances ($rm^f - rm$)

$$M_t - E_t + (rm_t - rm_t^f) = CA_t. \tag{20}$$

It is also equal to negative one times the capital account (KA) due to the balance of payments identity. Thus,

$$-CA_t = KA_t, \tag{21}$$

that is,

$$M_t - E_t + (rm_t - rm_t^f) = KA_t. \tag{22}$$

Following Leachman and Francis (2000), assuming that the net flow of labor income is zero, equation (22) can be written as

$$M_t - E_t + iB_{t-1}^n = B_t - B_t^f = \Delta B_t^n, \tag{23}$$

where i is the real rate of interest on foreign debt, B^n net borrowing from foreigners or net capital inflows, and B and B^f represent domestic borrowing form foreign agents and foreign borrowing from domestic agents, respectively. Previous equations are characterizations of the period-by-period balance of payments and equate the current account deficit (surplus) to capital inflow (outflow) or net borrowing (lending) from (to) abroad.

Forward substitution of Equation (23) gives

$$\begin{aligned} &(1 + i_{t-1})B_{t-1}^n + (M_t - E_t) + \\ &\lim_{T\to\infty} \sum_{\gamma=1}^{T} \left[\prod_{s=1}^{\gamma} [1/(1 + r_{t+s})(M_{t+s} - E_{t+s})] \right] = \\ &\lim_{T\to\infty} \left[\prod_{s}^{T} [1/(1 + r_{t+s})B_{t+s}^n \right], \end{aligned} \tag{24}$$

where r is a time varying discount rate. Equation (24) is the intertemporal external budget constraint, which equilibrates the present discounted value of current account deficits plus the beginning period foreign debt to the present discounted value of foreign borrowing or capital inflow. Sustainable policies with regards to the external sector must satisfy the intertemporal budget constraint given in (24) along with the transversality condition that requires a zero limit on future external debt discounted at a rate r – see Leachman and Francis (2002).

From an empirical point of view, several methodologies have been applied to study sustaintability of foreign US debt. We concentrate on Leachman and Francis (2000) since they use multicointegration techniques to analyze intertemporal external budget balance in the US economy. If exports and imports are to be multicointegrated, these variables have to be bound together by two equilibrating forces. The first cointegrating relationship within a multicointegrated system reflects the flow equilibrium force, while the second relationship reflects the deeper stock-flow relation. These two equilibrating forces should reflect then the sources of the sustainability of foreign debt via intertemporal external budget balance.

Leachman and Francis (2000) split the sample period that is analyzed in two subsamples: 1947-1973 and 1974-1994. Based on standard multicointegration techniques that do not allow for the presence of structural breaks, they find evidence about the presence of multicointegration in the current account system in the first subperiod. Such a relationship ensured that the external sector intertemporally moved towards balance even in "bad" states of nature. However, this kind of long-run relationship was not found in the second subsample. Hence, they conclude that in the second subperiod the US may no longer adhere to its intertemporal budget constraint and may be engaged in a Ponzi gamble that could lead to diminished welfare – see Leachman and Francis (2000) for further details.

Implicitly, the analysis of Leachman and Francis (2000) assumed the presence of instability since they decided to split the period of analysis in two. Similarly, Leachman and Francis (2002) test the Twin Deficits Hypothesis for the US economy splitting their period sample (1948-1992) in 1973. The more recent literature on the external sector has emphasized the importance of structural breaks in the determination of the current account. Edwards (2001) discusses several events, including the oil price shocks in the

1970s and the debt crisis in 1982. Hatemi and Shukur (2002) find a structural break at the beginning of the 1990s, which they attribute to the New Economy, globalization and the integration of the former socialist economies into the world economy. Although all these proposals evidence the existence of structural breaks in the current account behavior, none of them have tested the sustainability of the foreign debt in a multicointegration framework that considers the presence of structural breaks. Furthermore, note that splitting the sample size in two different subperiods as in Leachman and Francis (2000, 2002) implies reducing the number of observations that is used when testing for the presence of multicointegration relationships, which in turn causes loss of power of the test statistic. Instead, it is possible to apply the methodology that has been proposed in this paper, which allows us to consider the whole sample period taking into account the presence of one unknown structural break. Following Leachman and Francis (2000), the data set on exports and imports for the US economy has been taken from the Federal Reserve Bank of St. Louis. They are quarterly observation in billions of real 2000 dollars from 1947:Q1 through 2005:Q4.

We first proceed to investigate the stochastic properties of the univariate series describing foreign deficit. Since structural breaks might be affecting the time series, we first compute the $Exp - W_{FS}$ break test of Perron and Yabu (2005) to test whether there is a structural break affecting the time series regardless of the order of integration. The specification that is chosen is given by Model III in Perron and Yabu (2005), which considers that the structural break may affect both the level and the slope of the time trend. Results in Table (9) show that the null hypothesis of absence of structural breaks is rejected at the 5% level of statistical significance. Consequently, the analysis of the order of integration has to consider the presence of structural breaks. To this end, we have computed the M-class tests proposed in Carrion-i-Silvestre, Kim and Perron (2006) that accommodates the presence of one structural break. As can be seen, the null hypothesis of unit root with a structural break that affects the level and the slope of the time series cannot be rejected at the 5% level of significance by any of the M-class tests that have been computed – the break point for the exports is estimated at 1985:Q3, while the one for the imports is 1990:Q3.

This situation indicates that structural breaks should be taken into account when testing for the presence of cointegration and multicointegration relationships among the US imports and exports – following Leachman and Francis (2000), we specify imports as the dependent variable and exports as the explanatory one. Table 10 presents the results that are obtained when we compute the ADF statistic proposed in Gregory and Hansen (1996) as well as the cointegration test in Carrion-i-Silvestre and Sansó (2006). As mentioned above, the test in Gregory and Hansen (1996) specifies the null hypothesis of no cointegration, while the one in Carrion-i-Silvestre and Sansó (2006) is cointegration with one structural break. Concerning the test of Gregory and Hansen (1996), we can see that the null hypothesis of no cointegration between imports and exports is only rejected at the 5% level for the specification given by Model C/S – Model C accounts for a change in the level with no time trend and no change in the cointegrating vector, model C/T for a model including a time trend with a change in the level, but no change either in the slope of the time trend or in the cointegrating vector, and, finally, Model C/S considers a change in the level with no time trend and change in the cointegrating vector. The evidence is more striking when using the cointegration test in Carrion-i-Silvestre and Sansó (2006), since the null hypothesis of cointegration with structural break between imports and exports cannot be rejected at the 5% level of significance for any of the six different specifications that they propose. Therefore, we have found evidence of a long-run relationship between imports and exports that have been affected by the presence of one structural break.

This fact allows us to investigate in a multivariate framework the long-run relationships between exports and imports by using the multicointegration procedure. Results reported in Table (11) indicate that the null hypothesis of non-multicointegration is rejected at the 5% level of significance when the specification is given by Models 6 and 8, where the break point is estimated at 1992:Q4 and 1992:Q2 respectively. Note that Model 6 is preferred to Model 8 according to the BIC information criterion. It has to be stressed that, regardless of the specification that is selected, the estimated break point is located at the beginning of the 1990s, which is in accordance with the discussion in Hatemi and Shukur (2002) that attribute this change to the New Economy, globalization and the integration of the former socialist economies into the world economy. Therefore, our results support the intertemporal external budget balance for the complete period 1947-2005 in the US economy, with a changing way of balancing. Note that the break has been found in the first level of cointegration, that is the flow relation between exports and imports is changing along time but, according to the BIC information criterion, not the stock-flow relationship among debt, exports and imports.

6 Conclusions

In this paper we generalize the concept of multicointegration considering the presence of one structural break, which can affect either the deterministic component and/or the cointegrating vectors of the two levels of cointegration. The contribution in this paper goes beyond the multicointegration framework since our approach has wider application when the interest is testing for polynomial cointegration and, in general, cointegration among I(2) variables in a single-equation framework accounting for one structural break.

The paper has devised a residual based ADF statistic to test the null hypothesis of non-multicointegration against the alternative hypothesis of multicointegration with one structural break. Simulation experiments show that the finite sample performance of the statistic shows good properties in terms of empirical size and power. Moreover, the use of the AIC and BIC information criteria has been investigated when these statistics are used to select among the different specifications that have been devised. Based on simulation evidence we conclude that these information criteria can be used to select from different types of break, specially when both the magnitude of the structural break and the sample size are large.

We illustrate the use of the statistic to analyze the sustainability of the external deficit for the US. Previous evidence in the literature indicates that some doubt can be cast on the sustainability of US external deficit, specially in the last quarter of the twentieth century. We have shown that this conclusion was obtained because previous analyses did not consider the presence of structural breaks, which causes misleading conclusions. Thus, the study that has been conducted reveals that imports and exports can be considered as I(1) processes with one structural break, which implies that cointegration and multicointegration analyses have to account for this feature. Thus, when the presence of one structural break is considered, the evidence points to the sustainability of the US external deficit with a structural break that has been estimated at the beginning of the 1990s.

References

Akaike, H., 1973. Information theory and the extension of the maximum likelihood principle, in: Petrow, B.N., Csaki, F., (Eds.), Second International Symposium of Information Theory, Akailesoniai-Kiudo.

Billingsley, P., 1968. Convergence of probability measures. Wiley, New York, NY.

Carrion-i-Silvestre, J. Ll., Kim, D., Perron, P., 2006. GLS-based unit root tests with multiple structural breaks both under the null and the alternative hypotheses. Working paper. Department of Economics. Boston University.

Carrion-i-Silvestre, J. Ll., Sansó, A., 2006. Testing the null of cointegration with structural breaks. Oxford Bulletin of Economics and Statistics 68, 623-646.

Chang, Y., Park, J. Y., 2002. On the asymptotics of ADF tests for unit roots. Econometric Reviews, 21, 431-447.

Edwards, S., 2001. Does the Current Account Matter?. Mimeo. Cambridge: NBER.

Engle, R. F., Granger, C. W. J., 1987. Co-integration and error correction: representation, estimation and testing. Econometrica 55, 251-276.

Engsted, T., Gonzalo, J., Haldrup, N., 1997. Testing for multicointegration. Economic Letters 56, 259-266.

Engsted, T. and Haldrup, N. 1999. Multicointegration in Stock-Flow Models. Oxford Bulletin of Economics and Statistics 61, 237-254.

Granger, C. W. J., Lee, T., 1989. Multicointegration. Advances in Econometrics 8, 71-84.

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

Haldrup, N., 1994. The asymptotics of single-equation cointegration regressions with I(1) and I(2) variables. Journal of Econometrics 63, 151-181.

Hansen, B. E., 1992. Convergence to stochastic integrals for dependent heterogeneous processes. Econometric Theory 8, 489-500.

Hatemi, A. J., Shukur, G., 2002. Multivariate-Based Causality Tests of Twin Deficits in the US. Journal of Applied Statistics 29, 812-824.

Leachman, L. L., 1996. New evidence on the Ricardian Equivalence Theorem: A Multicointegration Approach. Applied Economics, 28, 695-704.

Leachman, L. L., Francis, B., 2002. Twin Deficits: Apparition or Reality?. Applied Economics, 34, 1121-1132.

Leachman, L. L., Bester, A., Rosas, G., Lange, P., 2005. Multicointegration and Sustainability of Fiscal Practices. Economic Inquiry, 43, 454-466.

Lee, T., 1992. Stock flow relationships in US housing construction. Oxford Bulletin of Economics and Statistics, 54, 419-430.

Kremers, J.M., Ericsson, N.R., Dolado, J.J., 1992. The Power of Cointegration Tests. Oxford Bulletin of Economics and Statistics 54, 325-348.

MacKinnon, J., 1991. Critical values for cointegration tests, in: R. F. Engle and C. W. J. Granger, (Eds.), Long Run Economic Relationships, Oxford University Press, Oxford.

Perron, P., 1989. The Great Crash, the Oil Price Shock and the Unit Root Hypothesis. Econometrica 57, 1361-1401.

Perron, P., 2006. Dealing with Structural Breaks, in Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory, K. Patterson and T.C. Mills (eds.), Palgrave Macmillan, 278-352.

Perron, P., Yabu, T., 2005. Testing for Shifts in Trend with an Integrated or Stationary Noise Component. Working Paper. Department of Economics. Boston University.

Phillips, P.C.B., Durlauf, S. N., 1986. Multiple Time Series Regression with Integrated Processes. Review of Economic Studies LIII, 473-495.

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics 6, 461-464.

Siliverstovs, B., 2003. Multicointegration in US consumption data. Discussion Paper, num. 382. DIW Berlin, German Institute for Economic Research.

A Mathematical appendix

The following Lemma is used throughout the paper to derive the statements in the Theorems.

Lemma 1 *Let* $z_t(\lambda) = (Cm_t', x_t^{0\prime}, x_t^{0\prime} 1(t > T_b), X_t^{0\prime}, X_t^{0\prime} 1(t > T_b))'$ *be the vector that contains the deterministic terms, I(1) and I(2) stochastic trends of the model defined in (10). Thus, as* $n \to \infty$*, the following moments converge jointly to*

$$
\begin{aligned}
&(a)\ n^{-1} D_n z'(\lambda) z(\lambda) D_n \Rightarrow \textstyle\int_0^1 B(\lambda, r)' B(\lambda, r)\, dr \\
&(b)\ n^{-1/2} D_n z'(\lambda) v \Rightarrow \textstyle\int_0^1 B(\lambda, r)'\, dB_0(r) \\
&\qquad\qquad +(0, \Delta_{10}, (1-\lambda)\Delta_{10}, 0, 0)' \\
&(c)\ n^{-3/2} D_n z'(\lambda) \Delta^{-1} v \Rightarrow \textstyle\int_0^1 B(\lambda, r)' B_0(r) dr \\
&(d)\ n^{-5/2} D_n z'(\lambda) \Delta^{-2} v \Rightarrow \textstyle\int_0^1 B(\lambda, r)' \overline{B_0(r)} dr,
\end{aligned}
$$

where D_n *denotes the scaling matrix.*

Proof. Let us focus on the specification given by Model 8, which is the most general one considered in this paper. This model uses the deterministic component given by $Cm_t = (1, t, t^2, DU_t, DT_t^*)'$, for which we can define the associated scaling matrix $D_0 = diag\{1, n^{-1/2}, n^{-1}, 1, n^{-1/2}\}$. Thus, note that $D_0 Cm_{[nr]} \to f(\lambda, r)$, with $f(r, \lambda) = (1, r, r^2, du(\lambda, r), dt^*(\lambda, r))'$, $r = t/n$, $du(\lambda, r) = 1(r > \lambda)$, $dt^*(\lambda, r) = (r - \lambda) 1(r > \lambda)$, the limiting functions of the deterministic components. For the I(1) stochastic regressors we have $n^{-1/2} x_t^0 \Rightarrow B_1(r)$ by the Donsker's Theorem, where $B_1(r)$ denotes a vector of m_1 Brownian processes defined on $[0, 1]$. For the I(2) stochastic regressors we can establish weak convergence towards $n^{-3/2} X_t^0 \Rightarrow \int_0^r B_2(s) ds \equiv \overline{B_2(r)}$ by the Continuous Mapping Theorem (CMT) – see Billingsley (1968) – with $B_2(s)$ being a vector of m_2 Brownian processes defined on $[0, 1]$. Similarly, $n^{-1/2} x_t^0 1(t > T_b) \Rightarrow B_1(r) 1(r > \lambda)$ and $n^{-3/2} X_t^0 1(t > T_b) \Rightarrow \overline{B_2(r)} 1(r > \lambda)$. Using all these elements it is straightforward to see that

$$
\begin{aligned}
n^{-1/2} D_n z_t(\lambda) &\Rightarrow \left(f(\lambda, r)', B_1'(r), B_1'(r) 1(r > \lambda), \overline{B_2'(r)}, \overline{B_2'(r)} 1(r > \lambda) \right)' \\
&\equiv B(\lambda, r),
\end{aligned}
$$

where $D_n = diag\{D_0, D_1, D_2\}$, with $D_1 = diag\{n^{-1/2}, ..., n^{-1/2}\}$ and $D_2 = diag\{n^{-3/2}, ...n^{-3/2}\}$ the scaling matrices involving the I(1) and I(2) stochastic regressors. Therefore, statements (a) to (c) easily follow from the application of the CMT. ■

A.1 Proof of Theorem 1

Consider the OLS estimation of the specification given in (11). By using the fact that for the general case $u_t \sim I(d)$

$$
n^{1/2-d} \hat{u}_t = n^{1/2-d} u_t - n^{1/2-d} X_t^m(\lambda)' (\hat{\theta} - \theta),
$$

that is,

$$
\begin{aligned}
n^{1/2-d} \hat{u}_t &= n^{1/2-d} u_t - n^{1/2-d} X_t^m(\lambda)' \left(X^m(\lambda)' X^m(\lambda) \right)^{-1} X^m(\lambda)' u \\
&= n^{1/2-d} u_t - n^{1/2-d} z_t(\lambda)' D_n \left(D_n z(\lambda)' z(\lambda) D_n \right)^{-1} D_n z(\lambda)' u \\
&= n^{1/2-d} u_t - z_t(\lambda)' D_n \left(n^{-1} D_n z(\lambda)' z(\lambda) D_n \right)^{-1} \left(n^{-1/2-d} D_n z(\lambda)' u \right).
\end{aligned}
$$

Under the null hypothesis of non-multicointegration $d = 1$, so that using the statements in Lemma 1 we have

$$
\begin{aligned}
n^{-1/2} \hat{u}_t &\Rightarrow B_0(r) - B(\lambda, r)' \left(\int_0^1 B(\lambda, r)' B(\lambda, r) \, dr \right)^{-1} \left(\int_0^1 B(\lambda, r)' B_0(r) \, dr \right) \\
&\equiv Q(\lambda, r). \qquad (25)
\end{aligned}
$$

Note that we can define $\omega_{00.1} = \omega_{00} - \omega_{01}\Omega_{11}^{-1}\omega_{10}$, so that

$$\begin{aligned} Q(r,\lambda) &= \omega_{00.1}^{1/2}\left(W_0(r) - W(\lambda,r)'\left(\int_0^1 W(\lambda,r)'W(\lambda,r)\,dr\right)^{-1}\left(\int_0^1 W(\lambda,r)'W_0(r)\,dr\right)\right) \\ &= \omega_{00.1}^{1/2}W^*(\lambda,r) \end{aligned}$$

can be written in terms of uncorrelated Brownian motions $W_0(r)$ and $W(\lambda,r)$. Note that $W^*(\lambda,r)$ denotes the projection of the $W_0(r)$ Brownian motion onto the space spanned by the columns of $W(\lambda,r)$.

The ADF statistic is computed from the estimation of (13) as the t-ratio $t_\delta(\lambda)$ for testing the null hypothesis that $\delta = 0$. Note that the ADF-type regression equation can be expressed as

$$\Delta\hat{u}_t = \delta\hat{u}_{t-1} + \hat{\xi}_t'\varphi + \eta_t,$$

with $\hat{\xi}_t = (\Delta\hat{u}_{t-1}, \ldots, \Delta\hat{u}_{t-p})'$. Following Chang and Park (2002), we can define

$$\begin{aligned} A_n(\lambda) &= \sum_{t=1}^n \hat{u}_{t-1}\eta_t - \left(\sum_{t=1}^n \hat{u}_{t-1}\hat{\xi}_t'\right)\left(\sum_{t=1}^n \hat{\xi}_t\hat{\xi}_t'\right)^{-1}\left(\sum_{t=1}^n \hat{\xi}_t\eta_t\right) \\ B_n(\lambda) &= \sum_{t=1}^n \hat{u}_{t-1}^2 - \left(\sum_{t=1}^n \hat{u}_{t-1}\hat{\xi}_t'\right)\left(\sum_{t=1}^n \hat{\xi}_t\hat{\xi}_t'\right)^{-1}\left(\sum_{t=1}^n \hat{\xi}_t\hat{u}_{t-1}\right) \\ C_n(\lambda) &= \sum_{t=1}^n \eta_t^2 - \left(\sum_{t=1}^n \eta_t\hat{\xi}_t'\right)\left(\sum_{t=1}^n \hat{\xi}_t\hat{\xi}_t'\right)^{-1}\left(\sum_{t=1}^n \hat{\xi}_t\eta_t\right), \end{aligned}$$

so that $\hat{\delta} = B_n^{-1}(\lambda)A_n(\lambda)$, with the variance of the error term given by $\hat{\sigma}_n^2 = n^{-1}(C_n(\lambda)$ $- A_n^2(\lambda)B_n^{-1}(\lambda))$ and the variance of the estimated δ parameter given by $s_n^2\left(\hat{\delta}\right) = \hat{\sigma}_n^2(\lambda)$ $B_n^{-1}(\lambda)$. Using these elements, Chang and Park (2002) show that as $n \to \infty$

$$\begin{aligned} n^{-1}A_n(\lambda) &\Rightarrow \omega_{00.1}\int_0^1 W^*(\lambda,r)\,dW^*(r,\lambda) \\ n^{-2}B_n(\lambda) &\Rightarrow \omega_{00.1}\int_0^1 W^{*2}(r,\lambda) \\ n^{-1}C_n(\lambda) &\to {}^p\sigma_n^2, \end{aligned}$$

where $\to^p$denotes convergence in probability. Then, the $t_\delta(\lambda)$ statistic is computed as

$t_\delta(\lambda) = \hat{\sigma}_n^{-1}(\lambda) B_n^{-1/2}(\lambda) A_n(\lambda)$, which in the limit converges to

$$t_\delta(\lambda) \Rightarrow \left(\int_0^1 W^{*2}(r,\lambda)\right)^{-1/2} \left(\int_0^1 W^*(\lambda, r)\, dW^*(r,\lambda)\right),$$

for a given λ. Theorem 1 is proved.

A.2 Proof of Theorem 2

The proof follows Zivot and Andrews (1992), Gregory and Hansen (1996), and Perron (1997). As shown in the previous proof, the ADF statistic can be expressed as a composite functional g:

$$\inf_{\lambda\in\Lambda} \left[\hat{\sigma}_n^{-1}(\lambda) B_n^{-1/2}(\lambda) A_n(\lambda)\right] = g\left(A_n(\lambda), B_n(\lambda), \hat{\sigma}_n^2(\lambda)\right),$$

where

$$g = h^*\left[h\left[m_1(\lambda), m_2(\lambda), m_3(\lambda)\right]\right],$$

with $h^*(m) = \inf_{\lambda\in\Lambda} m(\lambda)$ for any real function $m = m(\cdot)$ on Λ. Furthermore, for any real functions $n_1(\lambda)$, $n_2(\lambda)$ and $n_3(\lambda)$ on Λ, $h\left[A_n(\lambda), B_n(\lambda), C_n(\lambda)\right] = B_n(\lambda)^{-1/2} A_n(\lambda) / C_n^{1/2}(\lambda)$. The weak convergence joint result for $A_n(\lambda)$, and $B_n(\lambda)$ has been shown in Lemma 1 above, while $n^{-1}C_n(\lambda) \to^p \sigma_n^2$. Continuity of h on Λ is proved in Zivot and Andrews (1992). Finally, Zivot and Andrews (1992) establish continuity of $h^*(m)$ for all real functions m on Λ. Therefore, the continuity of g follows from the continuity of a composition of continuous functions, so that the CMT can be used to obtain the result of Theorem 2.

Table 2: Critical values for Models 4 and 5

		Model 4								Model 5							
		$m_2 = 1$				$m_2 = 2$				$m_2 = 1$				$m_2 = 2$			
m_1	n	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10
0	50	-6.62	-6.23	-5.92	-5.56	-7.05	-6.67	-6.33	-5.99	-7.20	-6.68	-6.32	-5.94	-7.55	-7.17	-6.77	-6.37
	100	-6.18	-5.86	-5.56	-5.26	-6.55	-6.19	-5.92	-5.62	-6.50	-6.19	-5.90	-5.60	-6.84	-6.55	-6.28	-5.95
	250	-5.94	-5.66	-5.41	-5.18	-6.28	-5.93	-5.73	-5.45	-6.26	-5.94	-5.72	-5.45	-6.56	-6.27	-6.02	-5.75
	500	-5.90	-5.60	-5.37	-5.10	-6.17	-5.89	-5.67	-5.40	-6.14	-5.92	-5.68	-5.43	-6.45	-6.18	-5.94	-5.69
1	50	-7.00	-6.61	-6.30	-5.94	-7.43	-7.06	-6.67	-6.32	-7.48	-7.04	-6.68	-6.29	-7.93	-7.51	-7.15	-6.73
	100	-6.53	-6.16	-5.90	-5.59	-6.90	-6.52	-6.21	-5.92	-6.81	-6.49	-6.23	-5.90	-7.16	-6.84	-6.56	-6.24
	250	-6.22	-5.94	-5.72	-5.47	-6.57	-6.25	-6.00	-5.74	-6.53	-6.24	-5.99	-5.74	-6.80	-6.55	-6.27	-6.00
	500	-6.15	-5.91	-5.69	-5.41	-6.42	-6.13	-5.95	-5.70	-6.44	-6.18	-5.96	-5.71	-6.72	-6.44	-6.20	-5.96
2	50	-7.40	-6.94	-6.61	-6.24	-7.77	-7.35	-7.01	-6.62	-7.83	-7.34	-7.02	-6.62	-8.16	-7.81	-7.44	-7.01
	100	-6.80	-6.47	-6.21	-5.90	-7.12	-6.79	-6.50	-6.22	-7.12	-6.77	-6.52	-6.22	-7.43	-7.08	-6.82	-6.50
	250	-6.55	-6.23	-6.00	-5.74	-6.85	-6.53	-6.28	-5.99	-6.75	-6.52	-6.27	-5.98	-7.06	-6.77	-6.55	-6.25
	500	-6.44	-6.18	-5.94	-5.68	-6.73	-6.43	-6.23	-5.95	-6.71	-6.48	-6.22	-5.97	-6.94	-6.67	-6.44	-6.19
3	50	-7.71	-7.27	-6.93	-6.55	-8.12	-7.69	-7.35	-6.93	-8.17	-7.66	-7.33	-6.93	-8.49	-8.12	-7.75	-7.31
	100	-7.07	-6.77	-6.51	-6.18	-7.44	-7.10	-6.79	-6.48	-7.43	-7.05	-6.75	-6.47	-7.80	-7.37	-7.08	-6.75
	250	-6.80	-6.50	-6.26	-5.98	-7.08	-6.78	-6.52	-6.24	-7.00	-6.77	-6.51	-6.25	-7.31	-7.01	-6.78	-6.50
	500	-6.75	-6.42	-6.17	-5.95	-6.97	-6.65	-6.45	-6.19	-6.98	-6.66	-6.43	-6.18	-7.17	-6.88	-6.65	-6.41
4	50	-8.05	-7.63	-7.28	-6.90	-8.45	-8.03	-7.65	-7.24	-8.52	-8.02	-7.65	-7.24	-8.85	-8.46	-8.04	-7.63
	100	-7.36	-7.06	-6.77	-6.48	-7.72	-7.37	-7.07	-6.75	-7.75	-7.33	-7.04	-6.72	-8.00	-7.66	-7.31	-7.01
	250	-7.03	-6.76	-6.49	-6.24	-7.28	-7.00	-6.74	-6.48	-7.24	-6.95	-6.75	-6.47	-7.57	-7.25	-7.00	-6.72
	500	-6.95	-6.69	-6.45	-6.17	-7.17	-6.89	-6.66	-6.43	-7.23	-6.93	-6.67	-6.42	-7.42	-7.11	-6.88	-6.63

The indices m_1 and m_2 indicate the number of I(1) and I(2) variables, respectively. n indicates the sample size. The simulations were based upon 5,000 replications.

Table 3: Critical values for Models 6 and 7

		Model 6								Model 7							
		$m_2 = 1$				$m_2 = 2$				$m_2 = 1$				$m_2 = 2$			
m_1	n	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10
0	50	-6.93	-6.50	-6.16	-5.80	-7.68	-7.26	-6.92	-6.54	-6.54	-6.18	-5.88	-5.54	-7.01	-6.60	-6.28	-5.92
	100	-6.36	-6.06	-5.76	-5.49	-7.01	-6.66	-6.41	-6.10	-6.17	-5.80	-5.58	-5.29	-6.52	-6.20	-5.92	-5.63
	250	-6.16	-5.86	-5.64	-5.37	-6.71	-6.41	-6.15	-5.86	-5.94	-5.67	-5.43	-5.16	-6.20	-5.97	-5.73	-5.45
	500	-6.08	-5.77	-5.55	-5.29	-6.63	-6.28	-6.07	-5.81	-5.79	-5.55	-5.32	-5.10	-6.15	-5.84	-5.64	-5.39
1	50	-7.24	-6.88	-6.53	-6.18	-8.16	-7.68	-7.29	-6.88	-7.31	-6.89	-6.52	-6.21	-7.77	-7.27	-6.96	-6.59
	100	-6.77	-6.41	-6.13	-5.80	-7.37	-6.98	-6.71	-6.40	-6.81	-6.52	-6.23	-5.88	-7.13	-6.82	-6.53	-6.19
	250	-6.48	-6.18	-5.94	-5.68	-6.95	-6.67	-6.44	-6.16	-6.60	-6.30	-6.03	-5.76	-6.83	-6.54	-6.29	-6.00
	500	-6.35	-6.06	-5.86	-5.58	-6.83	-6.55	-6.31	-6.06	-6.38	-6.14	-5.95	-5.67	-6.66	-6.38	-6.18	-5.94
2	50	-7.62	-7.22	-6.87	-6.53	-8.51	-8.01	-7.63	-7.20	-7.97	-7.55	-7.19	-6.79	-8.38	-7.94	-7.55	-7.15
	100	-7.11	-6.71	-6.44	-6.12	-7.68	-7.28	-6.99	-6.68	-7.39	-7.01	-6.75	-6.44	-7.71	-7.34	-7.07	-6.73
	250	-6.69	-6.40	-6.21	-5.93	-7.21	-6.92	-6.68	-6.40	-7.18	-6.83	-6.54	-6.25	-7.37	-7.03	-6.75	-6.47
	500	-6.62	-6.33	-6.11	-5.88	-7.06	-6.80	-6.57	-6.32	-6.90	-6.66	-6.40	-6.16	-7.16	-6.87	-6.65	-6.39
3	50	-8.05	-7.56	-7.21	-6.84	-8.88	-8.36	-7.93	-7.52	-8.66	-8.15	-7.81	-7.35	-9.04	-8.57	-8.17	-7.73
	100	-7.38	-7.00	-6.71	-6.41	-7.94	-7.60	-7.26	-6.93	-7.92	-7.52	-7.25	-6.91	-8.19	-7.83	-7.53	-7.20
	250	-7.00	-6.71	-6.46	-6.22	-7.46	-7.16	-6.92	-6.64	-7.52	-7.23	-6.97	-6.67	-7.70	-7.42	-7.17	-6.91
	500	-6.88	-6.61	-6.39	-6.12	-7.27	-7.01	-6.79	-6.53	-7.37	-7.08	-6.88	-6.60	-7.57	-7.28	-7.06	-6.83
4	50	-8.50	-7.86	-7.54	-7.17	-9.21	-8.70	-8.28	-7.82	-9.24	-8.70	-8.32	-7.86	-9.68	-9.10	-8.67	-8.23
	100	-7.63	-7.28	-6.98	-6.68	-8.20	-7.83	-7.51	-7.19	-8.38	-8.06	-7.74	-7.38	-8.63	-8.30	-8.01	-7.65
	250	-7.28	-6.99	-6.72	-6.46	-7.71	-7.43	-7.17	-6.88	-7.88	-7.63	-7.38	-7.10	-8.07	-7.82	-7.59	-7.31
	500	-7.05	-6.84	-6.61	-6.36	-7.50	-7.24	-7.01	-6.75	-7.76	-7.47	-7.23	-6.99	-7.97	-7.64	-7.42	-7.18

The indices m_1 and m_2 indicate the number of I(1) and I(2) variables, respectively. n indicates the sample size. The simulations were based upon 5,000 replications.

Table 4: Critical values for Model 8

		$m_2 = 1$				$m_2 = 2$			
m_1	n	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10
0	50	-6.85	-6.48	-6.15	-5.79	-7.61	-7.18	-6.85	-6.47
	100	-6.48	-6.13	-5.86	-5.52	-7.08	-6.70	-6.43	-6.09
	250	-6.13	-5.87	-5.63	-5.36	-6.61	-6.36	-6.11	-5.86
	500	-6.06	-5.82	-5.55	-5.32	-6.55	-6.25	-6.06	-5.81
1	50	-7.52	-7.09	-6.77	-6.35	-8.24	-7.80	-7.41	-7.02
	100	-6.97	-6.66	-6.38	-6.06	-7.53	-7.21	-6.90	-6.59
	250	-6.67	-6.37	-6.14	-5.86	-7.14	-6.85	-6.56	-6.31
	500	-6.52	-6.31	-6.06	-5.80	-6.97	-6.73	-6.51	-6.25
2	50	-8.21	-7.74	-7.34	-6.89	-8.97	-8.44	-7.98	-7.55
	100	-7.50	-7.12	-6.84	-6.54	-8.01	-7.68	-7.40	-7.06
	250	-7.15	-6.86	-6.60	-6.30	-7.60	-7.29	-7.04	-6.75
	500	-6.94	-6.74	-6.49	-6.23	-7.43	-7.16	-6.92	-6.67
3	50	-8.74	-8.28	-7.88	-7.43	-9.51	-8.99	-8.54	-8.06
	100	-8.01	-7.62	-7.32	-7.00	-8.49	-8.10	-7.84	-7.51
	250	-7.55	-7.27	-7.02	-6.73	-8.03	-7.72	-7.44	-7.15
	500	-7.41	-7.12	-6.89	-6.63	-7.84	-7.52	-7.30	-7.03
4	50	-9.31	-8.78	-8.41	-7.98	-10.01	-9.49	-9.10	-8.59
	100	-8.47	-8.06	-7.78	-7.43	-8.98	-8.61	-8.25	-7.92
	250	-7.95	-7.67	-7.40	-7.16	-8.41	-8.10	-7.82	-7.55
	500	-7.87	-7.53	-7.31	-7.03	-8.18	-7.93	-7.67	-7.39

The indices m_1 and m_2 indicate the number of I(1) and I(2) variables, respectively. n indicates the sample size. The simulations were based upon 5,000 replications.

Table 5: Empirical size and power of standard multicointegration test

		Granger and Lee		Engsted, Gonzalo and Haldrup			
		Non-deterministics		Linear trend		Quadratic trend	
α_i	ρ_i	$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$
0	1	0.047	0.036	0.056	0.058	0.058	0.059
0.01	0	0.189	0.356	0.998	0.999	0.994	0.999
0.5	0	0.634	0.952	0.999	0.999	0.995	0.999
1	0	0.891	0.996	0.999	0.999	0.995	0.999
0.5	0.9	0.069	0.284	0.088	0.214	0.065	0.179
0.5	0.7	0.181	0.826	0.319	0.881	0.214	0.787
0.5	0.5	0.294	0.908	0.729	0.999	0.560	0.997
0.5	0	0.636	0.959	0.997	0.999	0.998	0.999

The nominal size is set at the 5% level of significance. The simulations were based upon 10,000 replications.

Table 6: Empirical size and power of $t_\delta^*(\lambda)$ statistic with one unknown structural break

		Empirical size	Empirical power: $(\alpha_i, \rho_i) = (0.5, 0.9)$					
		$(\alpha_i, \rho_i) = (0, 1)$	Small structural break (θ_1)			Large structural break (θ_2)		
	n		$\lambda = 0.25$	$\lambda = 0.5$	$\lambda = 0.75$	$\lambda = 0.25$	$\lambda = 0.5$	$\lambda = 0.75$
Model 4	50	0.037	0.057	0.053	0.064	0.070	0.086	0.080
	100	0.053	0.127	0.145	0.133	0.150	0.159	0.157
Model 5	50	0.041	0.042	0.039	0.038	0.051	0.082	0.055
	100	0.055	0.124	0.125	0.115	0.134	0.121	0.130
Model 6	50	0.032	0.078	0.100	0.162	0.118	0.363	0.465
	100	0.047	0.146	0.443	0.594	0.402	0.739	0.874
Model 7	50	0.039	0.116	0.122	0.131	0.370	0.507	0.444
	100	0.051	0.193	0.307	0.318	0.671	0.790	0.781
Model 8	50	0.031	0.097	0.256	0.277	0.428	0.629	0.683
	100	0.047	0.291	0.615	0.677	0.789	0.926	0.999

$\theta_1 = (1, 0.02, 0.01, 3, 0.06, 1.5, 1, 2, 0.1)'$ and $\theta_2 = (1, 0.02, 0.01, 9, 0.18, 1.5, 3, 2, 0.3)'$. The nominal size is set at the 5% level of significance. The simulations were based upon 10,000 replications.

Table 7: Behaviour of the AIC and BIC information criteria. Small magnitude of the structural break effect (θ_1)

	DGP		Model 2		Model 3		Model 4		Model 5		Model 6	
			$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$
$\lambda = 0.25$	Model 4	AIC	0.162	0.242	**0.283**	0.219	0.224	0.221	0.184	0.155	0.146	0.162
		BIC	0.346	0.532	0.247	0.160	0.202	0.163	0.146	0.098	0.058	0.046
	Model 5	AIC	0.156	0.061	0.304	0.453	0.177	0.052	0.236	0.385	0.126	0.048
		BIC	**0.305**	0.104	0.263	0.446	0.161	0.056	0.215	0.379	0.055	0.014
	Model 6	AIC	0.157	0.169	**0.268**	0.215	0.221	0.294	0.194	0.121	0.159	0.200
		BIC	**0.310**	**0.411**	0.241	0.153	0.205	0.271	0.173	0.092	0.070	0.072
	Model 7	AIC	0.085	0.005	0.133	0.048	0.109	0.007	0.443	0.878	0.229	0.061
		BIC	0.175	0.010	0.125	0.047	0.118	0.007	0.459	0.895	0.122	0.040
	Model 8	AIC	0.099	0.017	0.119	0.023	0.170	0.045	**0.338**	0.332	0.273	0.582
		BIC	0.201	0.058	0.106	0.018	0.179	0.061	**0.343**	0.426	0.170	0.436
$\lambda = 0.50$	Model 4	AIC	0.184	0.296	**0.273**	0.229	0.229	0.174	0.186	0.171	0.127	0.129
		BIC	0.349	0.577	0.232	0.153	0.210	0.122	0.160	0.118	0.048	0.029
	Model 5	AIC	0.196	0.067	0.262	0.441	0.171	0.077	0.234	0.366	0.136	0.048
		BIC	**0.350**	0.129	0.234	0.427	0.143	0.067	0.209	0.360	0.063	0.016
	Model 6	AIC	0.218	0.159	**0.282**	0.206	0.185	0.370	0.197	0.083	0.117	0.181
		BIC	**0.387**	0.314	0.234	0.171	0.161	0.370	0.166	0.075	0.051	0.069
	Model 7	AIC	0.041	0.000	0.063	0.012	0.041	0.003	0.477	0.782	0.377	0.202
		BIC	0.100	0.004	0.060	0.012	0.047	0.004	0.549	0.833	0.243	0.146
	Model 8	AIC	0.025	0.005	0.074	0.020	0.073	0.021	**0.425**	0.247	0.402	0.706
		BIC	0.078	0.020	0.068	0.019	0.073	0.018	**0.481**	0.331	0.299	0.611
$\lambda = 0.75$	Model 4	AIC	0.199	0.237	**0.287**	**0.239**	0.200	0.199	0.174	0.181	0.139	0.143
		BIC	0.368	0.508	0.247	0.156	0.174	0.153	0.144	0.132	0.066	0.050
	Model 5	AIC	0.175	0.071	0.315	0.452	0.140	0.067	0.246	0.366	0.123	0.043
		BIC	**0.308**	0.139	0.285	0.432	0.124	0.059	0.226	0.355	0.056	0.014
	Model 6	AIC	0.210	**0.289**	**0.281**	0.227	0.194	0.200	0.161	0.073	0.153	0.210
		BIC	**0.381**	**0.525**	0.242	0.154	0.163	0.190	0.135	0.050	0.078	0.080
	Model 7	AIC	0.055	0.003	0.096	0.034	0.059	0.007	0.472	0.862	0.317	0.093
		BIC	0.119	0.013	0.086	0.035	0.052	0.007	0.515	0.878	0.227	0.066
	Model 8	AIC	0.082	0.040	0.095	0.041	0.079	0.049	0.371	0.221	0.372	0.645
		BIC	0.144	0.097	0.088	0.026	0.082	0.055	**0.425**	0.275	0.260	0.546

Table 8: Behaviour of the AIC and BIC information criteria. Large magnitude of the structural break effect (θ_2)

			Model 2		Model 3		Model 4		Model 5		Model 6	
	DGP		$n=50$	$n=100$	$n=50$	$n=100$	$n=50$	$n=100$	$n=50$	$n=100$	$n=50$	$n=100$
$\lambda=0.25$	Model 4	AIC	0.239	0.290	0.198	0.257	0.194	0.183	0.169	0.110	0.199	0.159
		BIC	0.368	0.508	0.154	0.169	0.192	0.154	0.166	0.105	0.119	0.063
	Model 5	AIC	0.200	0.063	0.244	0.616	0.188	0.038	0.219	0.252	0.148	0.030
		BIC	**0.305**	0.112	0.206	0.596	0.191	0.028	0.210	0.249	0.087	0.014
	Model 6	AIC	0.126	0.051	0.112	0.022	0.416	0.592	0.077	0.014	0.268	0.320
		BIC	0.202	0.092	0.100	0.019	0.473	0.707	0.074	0.017	0.150	0.164
	Model 7	AIC	0.022	0.004	0.009	0.001	0.047	0.002	0.580	0.806	0.341	0.186
		BIC	0.036	0.005	0.006	0.001	0.041	0.003	0.646	0.838	0.270	0.152
	Model 8	AIC	0.049	0.006	0.022	0.006	0.085	0.028	0.234	0.099	0.609	0.860
		BIC	0.073	0.012	0.011	0.006	0.079	0.026	0.301	0.149	0.535	0.806
$\lambda=0.50$	Model 4	AIC	0.286	0.358	0.235	0.196	0.200	0.182	0.176	0.134	0.102	0.129
		BIC	0.465	0.605	0.180	0.119	0.166	0.129	0.149	0.098	0.039	0.048
	Model 5	AIC	**0.247**	0.046	0.243	0.531	0.191	0.103	0.210	0.237	0.108	0.082
		BIC	**0.427**	0.116	0.182	0.517	0.150	0.107	0.190	0.228	0.050	0.031
	Model 6	AIC	0.075	0.044	0.146	0.090	0.508	0.615	0.052	0.013	0.218	0.237
		BIC	0.131	0.079	0.128	0.083	0.592	0.726	0.040	0.012	0.108	0.099
	Model 7	AIC	0.009	0.001	0.007	0.000	0.013	0.000	0.557	0.718	0.413	0.280
		BIC	0.013	0.001	0.008	0.000	0.011	0.000	0.674	0.784	0.293	0.214
	Model 8	AIC	0.012	0.006	0.029	0.019	0.032	0.008	0.233	0.126	0.693	0.840
		BIC	0.018	0.014	0.029	0.017	0.034	0.010	0.276	0.155	0.642	0.803
$\lambda=0.75$	Model 4	AIC	0.243	0.288	0.227	0.216	0.216	0.208	0.154	0.074	0.159	0.213
		BIC	0.399	0.534	0.182	0.135	0.195	0.190	0.140	0.058	0.083	0.082
	Model 5	AIC	0.220	0.038	0.226	0.677	0.162	0.063	0.219	0.164	0.172	0.057
		BIC	**0.336**	0.089	0.205	0.660	0.147	0.062	0.213	0.160	0.098	0.028
	Model 6	AIC	0.171	0.055	0.116	0.032	0.390	0.568	0.022	0.013	0.300	0.331
		BIC	0.272	0.102	0.091	0.024	0.418	0.691	0.027	0.008	0.191	0.174
	Model 7	AIC	0.030	0.002	0.025	0.003	0.047	0.011	0.581	0.792	0.316	0.191
		BIC	0.049	0.004	0.017	0.004	0.045	0.012	0.624	0.828	0.264	0.151
	Model 8	AIC	0.068	0.038	0.033	0.014	0.147	0.072	0.130	0.073	0.621	0.802
		BIC	0.094	0.061	0.025	0.008	0.146	0.078	0.148	0.080	0.586	0.772

Table 9: Structural change and univariate unit root tests for Exports and Imports

	Imports	Exports
Perron-Yabu test	17.483	7.006
ADF-GLS	-1.705	-1.792
ZA	-5.765	-6.305
MZA	-5.657	-6.171
MSB	0.296	0.284
MZT	-1.673	-1.754

The critical value at the 5% level of significance for the Perron-Yabu test is 3.12, for the ADF-GLS and MZT tests is -3.36, for the ZA and MZA tests is -22.28, and for the MSB test is 0.15.

Table 10: Engle-Granger and Gregory-Hansen cointegration test

Gregory and Hansen (1996) ADF statistic		
	ADF	$\hat{T}_b$
Model C	-3.00	1984:Q3
Model C/T	-3.11	1986:Q4
Model C/S	-5.11*	1994:Q3
Carrion-i-Silvestre and Sansó (2006) SC statistic		
	$SC^{+}(\lambda)$	$\hat{T}_b$
Model An	0.053	1999:Q3
Model A	0.037	1999:Q3
Model B	0.046	1995:Q4
Model C	0.042	1997:Q3
Model D	0.122	1996:Q3
Model E	0.054	1995:Q4

* denotes rejection of the corresponding null hypothesis at 5% level of significance

Table 11: Testing for multicointegration with structural break

$$\sum_{t=1}^{T} M_t = \mu_0 + \mu_1 t + \mu_2 t^2 + \mu_3 DU_t + \mu_4 DT_t^* + \gamma_1 X_t + \gamma_2 X_t 1(t > T_b) + \beta_1 \sum_{t=1}^{T} X_t + \beta_2 \sum_{t=1}^{T} X_t 1(t > T_b) + u_t$$

	$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	AIC	BIC	$t_\delta^*(\lambda)$	$\hat{T}_b$
Model 4	383.28	-8.68	130.22	-7.15		-12.64		1.42		12.98	13.07	-4.56	1965:Q4
Model 5	199.60	0.87	1640.65	12.92	-0.16	-12.05		1.46		12.98	13.08	-4.58	1965:Q4
Model 6	-474.70	2.92	-300464.67	-1319.35		1.56		1.17	1.47	11.75	11.85	-5.89*	1992:Q4
Model 7	-52.76	-12.25	1505.33		-0.03	-4.73	-7.78	1.42		12.97	13.08	-4.65	1965:Q4
Model 8	-636.88	2.77	-210844.39	-946.60		5.45	-7.82	1.11	1.18	11.74	11.86	-5.89*	1992:Q2

* denotes rejection of the null hypothesis at 5% level of significance