Time Varying Cointegration

Herman J. Bierens* and Luis F. Martins†

January 18, 2009

Abstract

In this paper we propose a time varying cointegration vector error correction model in which the cointegrating relationship varies smoothly over time. The Johansen's setup is a special case of our model. A likelihood ratio test for time-invariant cointegration is defined and its asymptotic distribution is derived. The limiting law is a chi-square. We apply our test to the purchasing power parity hypothesis of international prices and nominal exchange rates, and find evidence of time-varying cointegration.

Keywords: Time Varying Error Correction Model; Chebyshev Polynomials; Likelihood Ratio; Power; Trace Statistic

J.E.L. Classification: C32.

*Department of Economics and CAPCP, Pennsylvania State University, 608 Kern Graduate Building, University Park, PA 16802, USA. E-mail: hbierens@psu.edu. Support for research within the Center for the Study of Auctions, Procurements, and Competition Policy (CAPCP) at Penn State has been provided by a gift from the Human Capital Foundation

†UNIDE and Department of Quantitative Methods, ISCTE - Business School, Av. das Forças Armadas, 1649-026 Lisbon, Portugal. E-mail: luis.martins@iscte.pt. Financial support under grants SFRH/BD/814/2000 and PTDC/ECO/68367/2006 from the Fundação para a Ciência e Tecnologia is gratefully acknowledged. This paper is a substantial further elaboration of Martins' Ph.D. thesis at the Pennsylvania State University. See Martins (2005). The authors are grateful to Pentti Saikkonen and two anonymous referees for helpful comments. A previous version of this paper was presented by Martins at the Econometric Society World Congress 2005 in London.

1 Introduction

The literature growth on cointegration has been impressive, especially since Engel and Granger's (1987) and Johansen's (1988) seminal papers. In their standard approach, it is assumed that the cointegrating vector(s) do not change over time. As in the structural break literature, the assumption of fixed cointegrating vector(s) is quite restrictive.

Researchers became concerned with the effects that structural changes may have on econometric models. Previous literature on structural change and cointegration has focused on developing procedures to detect breaks or to estimate the temporal location of eventual shifts. Papers addressing these issues in a single-equation framework include Hansen (1992); Quintos and Phillips (1993); Hao (1996); Andrews et al. (1996); Bai et al. (1998); and Kuo (1998), among others (see also Maddala and Kim 1998 for a general survey). In the context of a system of equations, which is the focus of our analysis, the main contributions are those by Seo (1998), which extended the tests of Hansen (1992). Hansen and Johansen (1999) and Quintos (1997) propose fluctuation tests (based on recursive sequences of eigenvalues and cointegrating vectors) for parameter constancy in cointegrated VAR's, but they do not parameterize the shifts. Regarding time-varying ECM's, Hansen (2003) generalizes reduced-rank methods to cointegration under sudden regime shifts with a known number of break points. Andrade et al. (2005) study a similar model to Hansen (2003) and develop tests on the cointegration rank and on the cointegration space under known and unknown break dates. Structural breaks are also the concern of Lütkepohl et al. (2003); Inoue (1999); and Johansen et al. (2000), who analyze the effects of breaks in the deterministic trend.

Unlike previous studies, we explicitly model the changes in the long run component of a time varying vector error-correction model and discuss its estimation by maximum likelihood. In order to allow for more flexibility in the specification of this long-run economic relationship, we must take into account the possibility that the cointegration vectors may be unknown vector-valued functions of time. Because we are dealing with long-term concepts such as equilibrium and cointegration, we argue that the cointegration vector may be subject to a smooth and gradual change over time instead of a sudden jump (as in Hansen 2003). In this context, structural change is associated with smooth time variations of the cointegrating vector.

A cointegrating regression, in the spirit of Engle and Granger (1987), with parameters that vary with time was proposed by Park and Hahn (1999). The cointegration vector is modeled as a smooth function and semi-nonparametrically parameterized by a Fourier expansion. Based on this model, they derive the asymptotic properties of the estimator and of several residual-based specification test statistics. This type of approach is also related to recent papers that specify the long run relationship to be non-linear, such as threshold cointegration of Blake and Fomby (1997) and hidden cointegration of Granger and Yoon (2002). In Blake and Fomby, the equilibrium error follows a threshold autoregression that is mean-reverting outside a given range and has a unit root inside the range. In Granger and Yoon, there are hitherto unnoticed cointegrating relationships among integrated variables. Moreover, De Jong (2001) extends Engle and Granger's approach by specifying a nonlinear model in the second step; Harris et al. (2002) present the idea of stochastic cointegration in which some or all of the variables are heteroskedastistically integrated; and Juhl and Xiao (2005) propose a semiparametric ECM (partially linear) in which the nonparametric term is a function of stationary covariates.

Also, the Markov-switching approach of Hall et al. (1997), and the smooth transition model of Saikkonen and Choi (2004) provide an interesting way of modeling shifts in the cointegration vector. The first authors considering sudden shifts between two states, while the latter authors permits a gradual shift between regimes. Lütkepohl et al. (1999) and Terasvirta and Eliasson (2001) propose money demand functions modeled by single-equation ECM's in which a smooth transition stationary term is added. The transition function is driven by one of the processes of the long-run relationship. Our approach offers more flexibility, in the sense that the cointegrating relationship is not confined to regimes and is allowed to take different values at each point in time.

We explicitly model the changes in the long run relationship. The ECM is specified with a cointegrating vector indexed by time and we approximate this vector by a linear combination of orthogonal Chebyshev time polynomials so that the resulting ECM only has time invariant coefficients. This way, we can apply the usual ML technique and define a LR statistic that tests for standard cointegration. This test is a simple hypothesis test on parameters of our model. We show that its limiting law is a chi-square. The novelty of our approach is that the limiting random variables are blocks of weighted Wiener processes where the weights are trigonometric functions on the $[0,1]$ support.

The remainder of the paper is organized as follows. In Section 2 we introduce the time varying ECM using Chebyshev time polynomials. In Section 3 we propose a likelihood ratio test to distinguishes Johansen's standard cointegration from our time-varying alternative, for the case without drift, and show that the asymptotic null distribution is chi-square. In Section 4 the asymptotic power of the test is derived analytically and via Monte Carlo simulations. In Section 5 we show that our results carry over to the drift case. In Section 6 we illustrate the merits of our approach by testing for TV cointegration of international prices and nominal exchange rates. In Section 7 we make concluding remarks. The proofs of the lemmas and theorems can be found in either the appendix at the end, or in the separate appendix Bierens and Martins (2009).

As to some notations, "$\Rightarrow$" denotes weak convergence, "$\xrightarrow{d}$" denotes convergence in distribution, and $\mathbf{1}(.)$ is the indicator function.

2 Definitions and Representations

For the $k \times 1$ vector time series Y_t we assume that for some t there are fixed $r < k$ linearly independent columns of the time-varying (TV) $k \times r$ matrix $\beta_t = (\beta_{1t}, \beta_{2t}, ..., \beta_{rt})$. The columns form the basis of the time-varying space of cointegrating vectors, $S_t^c = \text{span}(\beta_{1t}, \beta_{2t}, ..., \beta_{rt}) \subset \mathbb{R}^k$, $t = 1, 2, ...$. In this context, the cointegration space "wiggles" over time according to a certain law. The remaining $k-r$ orthogonal vectors, expressed in the $k \times (k-r)$ matrix $\beta_{t_\perp}$, are such that $\beta_{t_\perp}^{'} Y_{t-1}$ does not represent a cointegrating relationship. The matrices β_t will be modeled using Chebyshev time polynomials.

2.1 Time Varying VECM Representation

Consider the time-varying VECM(p) with Gaussian errors, without intercepts and time trends,

$$\triangle Y_t = \Pi_t^{'} Y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \triangle Y_{t-j} + \varepsilon_t, \ t = 1, ..., T, \tag{1}$$

where $Y_t \in \mathbb{R}^k$, $\varepsilon_t \sim i.i.d.\ N_k[0, \Omega]$ and T is the number of observations. Our objective is to test the null hypothesis of time-invariant (TI) cointegration,

$\Pi_t^{'} = \Pi^{'} = \alpha\beta^{'}$, $\text{rank}\left(\Pi_t^{'}\right) = r < k$, where α and β are fixed $k \times r$ matrices, against time varying (TV) cointegration of the type

$$\Pi_t^{'} = \alpha\beta_t^{'},$$

with $\text{rank}\left(\Pi_t^{'}\right) = r < k$ for $t = 1, ..., T$, under the maintained hypothesis that the cointegration rank is $r < k$, where α is a fixed $k \times r$ matrix, but now the β_t's are time-varying $k \times r$ matrices of cointegrating vectors. In both cases Ω and the Γ_j's are fixed $k \times k$ matrices.

Admittedly, this form of TV cointegration is quite restrictive, as only the β_t's are assumed to be time dependent. A more general form of TV cointegration is the case $Y_t = C_t Z_t$, where C_t is a sequence of nonsingular $k \times k$ matrices and $Z_t \in \mathbb{R}^k$ is a time-invariant cointegrated $I(1)$ process with a VECM(p) representation. Then Y_t has a VECM(p) representation, but where all the parameters are functions of t.

2.2 Chebyshev Time Polynomials

Chebyshev time polynomials $P_{i,T}(t)$ are defined by

$$\begin{aligned} P_{0,T}(t) &= 1,\ P_{i,T}(t) = \sqrt{2}\cos\left(i\pi\left(t-0.5\right)/T\right), \\ t &= 1,2,...,T,\ i = 1,2,3,... \end{aligned} \tag{2}$$

See for example Hamming (1973). Bierens (1997) used them in his unit root test against nonlinear trend stationarity. Chebyshev time polynomials are orthonormal, in the sense that for all integers i, j, $\frac{1}{T}\sum_{t=1}^{T} P_{i,T}(t) P_{j,T}(t) = \mathbf{1}(i = j)$. Due to this orthonormality property, any function $g(t)$ of discrete time, $t = 1, ..., T$ can be represented by

$$g(t) = \sum_{i=0}^{T-1} \xi_{i,T} P_{i,T}(t),\ \text{ where } \xi_{i,T} = \frac{1}{T}\sum_{t=1}^{T} g(t) P_{i,T}(t). \tag{3}$$

In the expression (3), $g(t)$ is decomposed linearly in components $\xi_{i,T}P_{i,T}(t)$ of decreasing smoothness, as illustrated Figure 1, where we present the plots of $\cos\left[i\pi\left(t-0.5\right)/T\right]$, $T = 100$, for $i = 1, 3, 5$. Therefore, if $g(t)$ is smooth (to be made more precise in Lemma 1 below), it can be approximated quite well by

$$g_{m,T}(t) = \sum_{i=0}^{m} \xi_{i,T} P_{i,T}(t) \tag{4}$$

for some fixed natural number $m < T - 1$.

In the following lemma we set forth conditions such that for real valued functions $g\left(t\right)$,

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \left(g\left(t\right) - g_{m,T}\left(t\right)\right)^2 \to 0 \text{ for } m \to \infty, \tag{5}$$

and we determine the rate of convergence involved.

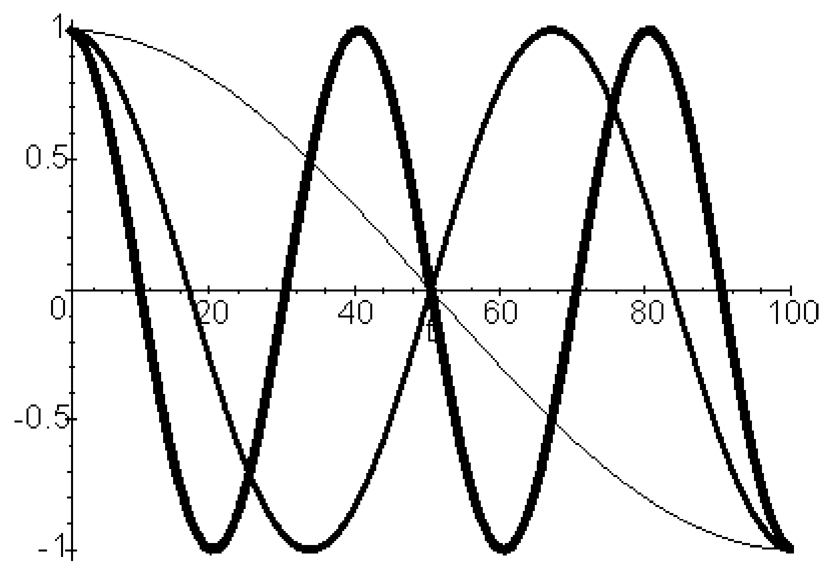

Figure 1. Thin $(i = 1)$; medium $(i = 3)$; thick $(i = 5)$.

Lemma 1. *Let $g\left(t\right) = \varphi\left(t/T\right)$, where $\varphi\left(x\right)$ is continuous function on $[0, 1]$.*[1] *Then (5) holds. Moreover, if $\varphi(x)$ is $q \geq 2$ times differentiable, where q is even, with $\varphi^{(q)}(x) = d^q\varphi(x)/\left(dx\right)^q$ satisfying $\int_0^1 \left(\varphi^{(q)}(x)\right)^2 dx < \infty$, then for $m \geq 1$,*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \left(g\left(t\right) - g_{m,T}\left(t\right)\right)^2 \leq \frac{\int_0^1 \left(\varphi^{(q)}(x)\right)^2 dx}{\pi^{2q}\left(m+1\right)^{2q}}, \tag{6}$$

[1] Therefore, $\beta(x)$ is uniformly continuous on $[0, 1]$.

Proof: Appendix

Consequently, we may without loss of generality write β_t for $t = 1, ..., T$ as $\beta_t = \sum_{i=0}^{T-1} \xi_{i,T} P_{i,T}(t)$, where the $P_{i,T}(t)$'s are Chebyshev time polynomials and $\xi_{i,T} = \frac{1}{T}\sum_{t=1}^{T} \beta_t P_{i,T}(t)$, $i = 0, ..., T-1$, are unknown $k \times r$ matrices. Then the null hypothesis of TI cointegration corresponds to $\xi_{i,T} = O_{k\times r}$ for $i = 1, ..., T-1$, and the alternative of TV cointegration corresponds to $\lim_{T\to\infty} \xi_{i,T} \neq O_{k\times r}$ for some $i \geq 1$. To make the latter operational, we will confine our analysis to TV alternatives for which $\lim_{T\to\infty} \xi_{i,T} \neq O_{k\times r}$ for some $i = 1, ..., m$, and $\xi_{i,T} = O_{k\times r}$ for all $i > m$, where m is chosen in advance. Effectively this means that under the alternative β_t is specified as

$$\beta_t = \beta_m(t/T) = \sum_{i=0}^{m} \xi_{i,T} P_{i,T}(t) \tag{7}$$

for some fixed m. Because low-order Chebyshev polynomials are rather smooth functions of t, we allow β_t to change gradually over time under the alternative of TV cointegration, contrary to Hansen's (2003) sudden change assumption.

Assuming that the elements of $\beta_m(x)$ are square-integrable on $[0, 1]$ we have for $x \in [0, 1]$,

$$\beta_m(x) = \lim_{T\to\infty} \beta_{[xT]} = \xi_0 + \sum_{i=1}^{m} \xi_i \sqrt{2}\cos(i\pi x)$$

where $\xi_0 = \int_0^1 \beta_m(x)\,dx$, $\xi_i = \int_0^1 \sqrt{2}\cos(i\pi x)\,\beta_m(x)\,dx$ for $i \geq 1$. More generally, $\beta_m(x)$ can be interpreted as an approximation of a square-integrable vector or matrix valued function $\beta(x)$ on $[0, 1]$ of the type

$$\beta(x) = \xi_0 + \sum_{i=1}^{\infty} \xi_i \sqrt{2}\cos(i\pi x)$$

where $\xi_0 = \int_0^1 \beta(x)\,dx$, $\xi_i = \int_0^1 \sqrt{2}\cos(i\pi x)\,\beta(x)\,dx$ for $i \geq 1$. However, we will not use this interpretation. We will derive our results under the assumption that (7) holds exactly.

This specification of the matrix of time varying cointegrating vectors is related to the approach of Park and Hahn (1999). They consider a time varying cointegrating relationship of the form

$$Z_t = \alpha_t' X_t + U_t$$

where $Z_t \in \mathbb{R}$, X_t is a k-variate $I(1)$ process and U_t is a stationary process. Thus, with $Y_t = (Z_t, X_t')'$ and and $\beta_t = (1, -\alpha_t')'$, $\beta_t' Y_t = U_t$ is stationary. Park and Hahn (1999) assume that the elements of α_t are of the form $\varphi(t/T)$, where $\varphi(x)$ has a Fourier flexible functional form. In particular,

$$\lim_{m\to\infty} \int_0^1 (\varphi(x) - \varphi_m(x))^2 \, dx = 0 \tag{8}$$

where

$$\varphi_m(x) = \overline{\varphi} + \sum_{i=1}^{m} \varsigma_i \sqrt{2} \cos(2i\pi x) + \sum_{j=1}^{m} \eta_j \sqrt{2} \sin(2j\pi x), \tag{9}$$

with Fourier coefficients

$$\begin{aligned} \overline{\varphi} &= \int_0^1 \varphi(x)\, dx, \ \varsigma_i = \int_0^1 \varphi(x) \sqrt{2} \cos(2i\pi x)\, dx, \\ \eta_j &= \int_0^1 \varphi(x) \sqrt{2} \sin(2j\pi x)\, dx. \end{aligned}$$

Moreover, they assume that $\varphi(x)$ is $q \geq 1$ times differentiable, with uniformly bounded q-th derivative, $\sup_{0\leq x\leq 1} \left|\varphi^{(q)}(x)\right| < \infty$.

As is well-known,[2] the functions $\sqrt{2}\cos(2i\pi x)$, $\sqrt{2}\sin(2j\pi x)$ for $i, j = 1, 2, 3, ...$, together with the constant 1, form a complete orthonormal sequence in the Hilbert space $L^2[0,1]$ of square-integrable real functions on $[0,1]$. However, the functions $\sqrt{2}\cos(i\pi x)$ for $i = 1, 2, 3, ...$ together with the constant 1 also form a complete orthonormal sequence in $L^2[0,1]$, so that for any function $\varphi(x) \in L^2[0,1]$,

$$\lim_{m\to\infty} \int_0^1 (\varphi(x) - \varphi_m(x))^2 \, dx = 0, \tag{10}$$

where

$$\varphi_m(x) = \xi_0 + \sum_{j=1}^{m} \xi_j \sqrt{2} \cos(j\pi x) \tag{11}$$

with Fourier coefficients $\xi_0 = \int_0^1 \varphi(x)\, dx$, $\xi_i = \int_0^1 \varphi(x) \sqrt{2} \cos(i\pi x)\, dx$, $i = 1, 2, 3, ...$ See the proof of Lemma 1 in the Appendix.

[2] See for example Young (1988).

On the other hand, Park and Hahn (1999) allow m to converge to infinity with T, whereas we only consider the fixed m case. The latter does not matter under the null hypothesis of time invariant cointegration, but it may affect the power of our test. This power issue will be addressed in section 4.3 via a Monte Carlo analysis.

2.3 Modeling TV Cointegration via Chebyshev Time Polynomials

Substituting $\Pi_t' = \alpha\beta_t' = \alpha\left(\sum_{i=0}^{m}\xi_i P_{i,T}(t)\right)'$ in (1) yields

$$\triangle Y_t = \alpha\left(\sum_{i=0}^{m}\xi_i P_{i,T}(t)\right)' Y_{t-1} + \sum_{j=1}^{p-1}\Gamma_j \triangle Y_{t-j} + \varepsilon_t$$

for some $k \times r$ matrices ξ_i, which can be written more conveniently as

$$\triangle Y_t = \alpha\xi' Y_{t-1}^{(m)} + \Gamma X_t + \varepsilon_t, \tag{12}$$

where $\xi' = \left(\xi_0', \xi_1', ..., \xi_m'\right)$ is an $r \times (m+1)k$ matrix of rank r, $Y_{t-1}^{(m)}$ is defined by

$$Y_{t-1}^{(m)} = \left(Y_{t-1}', P_{1,T}(t)\, Y_{t-1}', P_{2,T}(t)\, Y_{t-1}', ..., P_{m,T}(t)\, Y_{t-1}'\right)' \tag{13}$$

and

$$X_t = \left(\triangle Y_{t-1}', ..., \triangle Y_{t-p+1}'\right)'. \tag{14}$$

The null hypothesis of TI cointegration corresponds to $\xi' = (\beta', O_{r,k.m})$, where β is the $k \times r$ matrix of TI cointegrating vectors, so that then $\xi' Y_{t-1}^{(m)} = \beta' Y_{t-1}^{(0)}$, with

$$Y_{t-1}^{(0)} = Y_{t-1}. \tag{15}$$

This suggests to test the null hypothesis via a likelihood ratio test $LR^{tvc} = -2\left[\widehat{l}_T(r,0) - \widehat{l}_T(r,m)\right]$, where $\widehat{l}_T(r,0)$ is the log-likelihood of the VECM(p) (12) in the case $m = 0$, so that $Y_{t-1}^{(m)} = Y_{t-1}$, and $\widehat{l}_T(r,m)$ is the log-likelihood of the VECM(p) (12) in the case where $Y_{t-1}^{(m)}$ is given by (13), where in both cases r is the cointegration rank.

3 Testing TI Cointegration Against TV Cointegration

3.1 ML Estimation and the LR Test

Similar to Johansen (1988), the log-likelihood $\widehat{l}_T(r,m)$, given r and m takes the form

$$\widehat{l}_T(r,m) = -0.5T.\sum_{j=1}^{r}\ln\left(1-\widehat{\lambda}_{m,j}\right) - 0.5T.\ln\left(\det\left(S_{00,T}\right)\right)$$

plus a constant, where $1 > \widehat{\lambda}_{m,1} \geq \widehat{\lambda}_{m,2} \geq ... \geq \widehat{\lambda}_{m,r} \geq ... \geq \widehat{\lambda}_{m,(m+1)k}$ are the ordered solutions of the generalized eigenvalue problem

$$\det\left[\lambda S_{11,T}^{(m)} - S_{10,T}^{(m)}S_{00,T}^{-1}S_{01,T}^{(m)}\right] = 0 \tag{16}$$

with

$$\begin{aligned} S_{00,T} &= \frac{1}{T}\sum_{t=1}^{T}\triangle Y_t \triangle Y_t' \qquad (17)\\ &- \left(\frac{1}{T}\sum_{t=1}^{T}\triangle Y_t X_t'\right)\left(\frac{1}{T}\sum_{t=1}^{T}X_t X_t'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}X_t \triangle Y_t'\right), \\ S_{11,T}^{(m)} &= \frac{1}{T}\sum_{t=1}^{T}Y_{t-1}^{(m)}Y_{t-1}^{(m)\prime} \qquad (18)\\ &- \left(\frac{1}{T}\sum_{t=1}^{T}Y_{t-1}^{(m)}X_t'\right)\left(\frac{1}{T}\sum_{t=1}^{T}X_t X_t'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}X_t Y_{t-1}^{(m)\prime}\right), \\ S_{01,T}^{(m)} &= \frac{1}{T}\sum_{t=1}^{T}\triangle Y_t Y_{t-1}^{(m)\prime} \qquad (19)\\ &- \left(\frac{1}{T}\sum_{t=1}^{T}\triangle Y_t X_t'\right)\left(\frac{1}{T}\sum_{t=1}^{T}X_t X_t'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}X_t Y_{t-1}^{(m)\prime}\right), \\ S_{10,T}^{(m)} &= \left(S_{01,T}^{(m)}\right)'. \qquad (20) \end{aligned}$$

Note that $\widehat{\lambda}_{m,k+1} = ... = \widehat{\lambda}_{m,(m+1)k} \equiv 0$, because the rank of $\widehat{S}_{10}^{(m)}\widehat{S}_{00}^{-1}\widehat{S}_{01}^{(m)}$ is k.

Similar to Johansen (1988) it can be shown that, given r and m, the normalized maximum likelihood estimator $\widehat{\xi}$ of ξ is $\widehat{\xi} = (\widehat{q}_1, ..., \widehat{q}_r)$,where $\widehat{q}_1, ..., \widehat{q}_r$ are the r columns of the matrix of generalized eigenvectors associated with the r largest solutions for λ of (16).

As is well-known, in the TI case $m = 0$, the log-likelihood $\widehat{l}_T(r, 0)$ takes the form

$$\widehat{l}_T(r,0) = -0.5T. \sum_{j=1}^{r} \ln\left(1 - \widehat{\lambda}_{0,j}\right) - 0.5T. \ln\left(\det\left(S_{00,T}\right)\right)$$

plus the same constant as before, where $1 > \widehat{\lambda}_{0,1} \geq \widehat{\lambda}_{0,2} \geq ... \geq \widehat{\lambda}_{0,r}$ are the r largest solutions of the generalized eigenvalue problem

$$\det\left[\lambda S_{11,T}^{(0)} - S_{10,T}^{(0)} S_{00,T}^{-1} S_{01,T}^{(0)}\right] = 0.$$

The matrices $S_{11,T}^{(0)}, S_{01,T}^{(0)}$ and $S_{10,T}^{(0)}$ are defined by (18), (19) and (20), respectively, with $Y_{t-1}^{(m)}$ replaced by $Y_{t-1}^{(0)} = Y_{t-1}$. Therefore, given m and r, the LR test of the null hypothesis of standard (TI) cointegration against the alternative of TV cointegration takes the form

$$LR_T^{tvc} = -2\left[\widehat{l}_T(r,0) - \widehat{l}_T(r,m)\right] = T\sum_{j=1}^{r} \ln\left(\frac{1 - \widehat{\lambda}_{0,j}}{1 - \widehat{\lambda}_{m,j}}\right). \tag{21}$$

3.2 Data-Generating Process under the Null Hypothesis

For $m = 0$ we have the standard cointegration case:

Assumption 1. $\triangle Y_t = C(L)\, U_t = \sum_{j=0}^{\infty} C_j U_{t-j}$, *where* $U_t \sim i.i.d.\ N_k[0, I_k]$. *The elements of the* $k \times k$ *matrices* C_j *decrease exponentially to zero as* $j \to \infty$.

We can write $\triangle Y_t$ as

$$\triangle Y_t = C(1)U_t + (1 - L)D(L)U_t \tag{22}$$

where

$$D(L) = \frac{C(L) - C(1)}{1 - L}. \tag{23}$$

This is the well-known Beveridge-Nelson decomposition, which implies that

$$Y_t = C(1)\sum_{j=1}^{t} U_j + V_t + Y_0 - V_0 \tag{24}$$

where $V_t = D(L)U_t$ is a zero-mean stationary Gaussian process.

Assumption 2. *The matrix $C(1)$ is singular, with rank $1 \leq r < k$: There exists a $k \times r$ matrix β such that $\beta' C(1) = O_{r,k}$. Moreover, the $r \times k$ matrix $\beta' D(1)$ has rank r.*

For the time being we will also assume that

Assumption 3. $U_t = 0$ *for* $t < 1$,

so that $Y_0 = V_0 = 0$ in (24).

Admittedly, Assumption 3 is too restrictive, but is made to focus on the main issues. For the same reason we do not yet consider the more realistic case of drift in Y_t. Once we have completed the asymptotic analysis for the case under review, we will show what happens if there is drift in Y_t and Assumption 3 is dropped.

Assumption 4. *Under Assumptions 1-3, Y_t has the VECM(p) representation*

$$\triangle Y_t = \alpha\beta' Y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \triangle Y_{t-j} + \varepsilon_t,\ t \geq 1, \tag{25}$$

where $\varepsilon_t \sim$ i.i.d. $N_k[0, \Omega]$, with Ω non-singular.

Due to Assumption 3, there is no vector of constants in this model. Moreover, note that $\varepsilon_t = C_0 U_t$, so that $\Omega = C_0 C_0'$.

Similar to (12), model (25) can be written more conveniently as

$$\triangle Y_t = \alpha\beta' Y_{t-1} + \Gamma X_t + C_0 U_t,\ t \geq 1, \tag{26}$$

and replacing ε_t by $C_0 U_t$ in (12) the time-varying VECM(p) model becomes

$$\triangle Y_t = \alpha \xi' Y_{t-1}^{(m)} + \Gamma X_t + C_0 U_t, \tag{27}$$

where under the null hypothesis,

$$\xi = \begin{pmatrix} \beta \\ O_{m.k \times r} \end{pmatrix} \tag{28}$$

To exclude the case that $\beta' Y_{t-1}$ and X_t are multicollinear, we need to assume that

Assumption 5. Var$\begin{pmatrix} \beta' Y_{t-1} \\ X_t \end{pmatrix}$ *is nonsingular.*

3.3 Asymptotic Null Distribution

The asymptotic results in the standard cointegration case hinge on the following well-known convergence results. Under Assumptions 1-2,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T} U_t Y_{t-1}' \xrightarrow{d} \int_0^1 (dW)\, W' C(1)'. \\
&\frac{1}{T}\sum_{t=1}^{T} (\triangle Y_{t-\ell})\, Y_{t-1}' \xrightarrow{d} C(1)\left(\int_0^1 (dW)\, W'\right) C(1)' + M_\ell,\ \ \ell \geq 0, \\
&\frac{1}{T^2}\sum_{t=1}^{T} Y_t Y_{t-1}' \xrightarrow{d} C(1)\left(\int_0^1 (W(x))\, W'(x) dx\right) C(1)'
\end{aligned}$$

where W is a k-variate standard Wiener process, and the M_ℓ's are non-random $k \times k$ matrices. See Phillips and Durlauf (1986) and Phillips (1988).

We need to generalize these results to the case where Y_{t-1} is replaced by $Y_{t-1}^{(m)}$:

Lemma 2. *Under Assumptions 1-2,*

$$\frac{1}{T}\sum_{t=1}^{T} U_t \left(Y_{t-1}^{(m)}\right)' \xrightarrow{d} \int_0^1 (dW)\, \widetilde{W}_m' \left(C(1)' \otimes I_{m+1}\right). \tag{29}$$

$$\frac{1}{T}\sum_{t=1}^{T}(\triangle Y_{t-\ell})\left(Y_{t-1}^{(m)}\right)' \overset{d}{\to} C(1)\int_0^1 (dW)\,\widetilde{W}_m'\,(C(1)'\otimes I_{m+1}) \quad (30)$$
$$+M_\ell^*,\quad \ell\geq 0\ ,$$

$$\frac{1}{T^2}\sum_{t=1}^{T}\left(Y_{t-1}^{(m)}\right)\left(Y_{t-1}^{(m)}\right)' \overset{d}{\to} \quad (31)$$
$$(C(1)\otimes I_{m+1})\int_0^1 \widetilde{W}_m(x)\widetilde{W}_m'(x)dx\,(C(1)'\otimes I_{m+1})$$

where W is a k-variate standard Wiener process,

$$\widetilde{W}_m(x)=\left(W'(x)\,,\sqrt{2}\cos(\pi x)W'(x)\,,...,\sqrt{2}\cos(m\pi x)W'(x)\right)',$$

and the M_ℓ^*'s are $k\times k(m+1)$ non-random matrices.

Proof: Appendix.

The result (30) implies that $(1/T)\sum_{t=1}^{T}(\triangle Y_{t-\ell})\left(Y_{t-1}^{(m)}\right)'=O_p(1)$. The latter is what is needed for our analysis. Therefore, the question how the matrices M_ℓ^* look like is not relevant.

Note that

$$\int_0^1 (dW)\,\widetilde{W}_m'=\left(\int_0^1 (dW(x))\,W'(x),\sqrt{2}\int_0^1\cos(1\pi x)\,dW(x)W'(x),\right.$$
$$\left.\sqrt{2}\int_0^1\cos(2\pi x)\,dW(x)W'(x),...,\sqrt{2}\int_0^1\cos(m\pi x)\,dW(x)W'(x)\right)$$

In the Appendix we define the proper meaning of the random matrices

$$\int_0^1\cos(\ell\pi x)\,dW(x)W'(x) \quad (32)$$

for $\ell=1,2,3,....$ In particular, if $W(x)$ is univariate then

$$\int_0^1\cos(\ell\pi x)\,W(x)dW(x)=\frac{(-1)^\ell}{2}W^2(1)+\frac{\ell\pi}{2}\int_0^1\sin(\ell\pi x)\,W^2(x)dx.$$

Using Lemma 2 (together with rather long list of auxiliary lemmas), the following results can be shown.

Lemma 3. *Under Assumptions 1-5 the r largest ordered solutions $\widehat{\lambda}_{m,1} \geq \widehat{\lambda}_{m,2} \geq ... \geq \widehat{\lambda}_{m,r}$ of generalized eigenvalue problem* (16) *converge in probability to constants $1 > \overline{\lambda}_1 \geq ... \geq \overline{\lambda}_r > 0$, which do not depend on m. Thus, these probability limits are the same as in the standard TI cointegration case.*

Proof: Appendix.

As is well-known (see Johansen 1988), in the standard TI cointegration case $m = 0$ and under Assumptions 1-5, $T\left(\widehat{\lambda}_{0,r+1}, \widehat{\lambda}_{0,r+2}, ..., \widehat{\lambda}_{0,k}\right)'$ converges in distribution to the vector or ordered solutions $\rho_{0,1} \geq \rho_{0,2} \geq .... \geq \rho_{0,k-r}$ of

$$\det\left[\rho \int_0^1 W_{k-r}(x)W'_{k-r}(x)dx - \int_0^1 W_{k-r}dW'_{k-r} \int_0^1 \left(dW_{k-r}\right) W'_{k-r}\right] = 0, \tag{33}$$

where

$$W_{k-r}(x) = \left(\alpha'_\perp \Omega \alpha_\perp\right)^{-1/2} \alpha'_\perp C_0 W(x) \tag{34}$$

is a $k-r$ variate standard Wiener process.[3] This result is based on the fact that one can choose an orthogonal complement $\beta_\perp$ of β such that

$$\frac{1}{T}\beta'_\perp S^{(0)}_{11,T}\beta_\perp \xrightarrow{d} \int_0^1 W_{k-r}(x)W'_{k-r}(x)dx$$
$$\left(\alpha'_\perp \Omega \alpha_\perp\right)^{-1/2} \alpha'_\perp C_0 S^{(0)}_{01,T}\beta_\perp \xrightarrow{d} \left(\int_0^1 \left(dW_{k-r}\right) W'_{k-r}\right)$$

One would therefore expect that this result can be generalized to the TV cointegration case simply by replacing $W_{k-r}(x)$ in (33) with

$$\begin{aligned}\widetilde{W}_{k-r,m}(x) &= \left(W'_{k-r}(x), \sqrt{2}\cos(\pi x)W'_{k-r}(x), ..., \sqrt{2}\cos(m\pi x)W'_{k-r}(x)\right)' \\ &= \left(\left(\alpha'_\perp \Omega \alpha_\perp\right)^{-1/2} \alpha'_\perp C_0 \otimes I_{m+1}\right)\widetilde{W}_m(x) \end{aligned} \tag{35}$$

while leaving dW_{k-r} as is. However, that is not the case!

Lemma 4. *Under Assumptions 1-5,*

$$T\left(\widehat{\lambda}_{m,r+1}, \widehat{\lambda}_{m,r+2}, ..., \widehat{\lambda}_{m,k}\right)' \xrightarrow{d} \left(\rho_{m,1}, ..., \rho_{m,k-r}\right)',$$

[3] Because $\alpha'_\perp C_0 C'_0 \alpha_\perp = \alpha'_\perp \Omega \alpha_\perp$.

where $\rho_{m,1} \geq \rho_{m,2} \geq .... \geq \rho_{m,k-r}$ are the $k-r$ largest solutions of generalized eigenvalue problem

$$0 = \det\left[\rho\begin{pmatrix} \int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}'_{k-r,m}(x)dx & O_{(k-r)(m+1),r.m} \\ O_{r.m,(k-r)(m+1)} & I_{r.m}\end{pmatrix} - \begin{pmatrix} \int_0^1 \widetilde{W}_{k-r,m}dW'_{k-r} \\ V \end{pmatrix}\left(\int_0^1 (dW_{k-r})\widetilde{W}'_{k-r,m}, V'\right)\right] \quad (36)$$

with V is an $r.m \times (k-r)$ random matrix with i.i.d. $N[0,1]$ elements. Moreover, V is independent of W_{k-r} and $\widetilde{W}_{k-r,m}$.

Proof: Appendix.

The reason for this unexpected result is the following. Under the null hypothesis (28), any orthogonal complement of the matrix ξ of TV cointegrating vectors takes the form

$$\xi_\perp = \left(\beta_\perp \otimes I_{m+1}, \begin{pmatrix} O_{k,m.r} \\ \beta \otimes I_m \end{pmatrix}\right) \times R$$

where R is a nonsingular $(k(m+1)-r) \times (k(m+1)-r)$ matrix. We need to choose R such that $\frac{1}{T}\xi'_\perp S^{(m)}_{11,T}\xi_\perp$ converges in distribution to a nonsingular matrix.[4] A suitable version of $\xi_\perp$ that delivers this result is

$$\xi_{\perp,T} = \left(\beta_\perp \otimes I_{m+1}, \begin{pmatrix} O_{k,m.r} \\ \sqrt{T}\left(\beta\Sigma_{\beta\beta}^{-1/2} \otimes I_m\right) \end{pmatrix}\right) \quad (37)$$

where $\Sigma_{\beta\beta} = p\lim_{T\to\infty}(1/T)\sum_{t=1}^T \beta' Y_{t-1}Y'_{t-1}\beta$. Then $\frac{1}{T}\xi'_{\perp,T}S^{(m)}_{11,T}\xi_{\perp,T}$ converges in distribution to the first matrix in (36). The matrix V involved is now due to

$$\left(\alpha'_\perp\Omega\alpha_\perp\right)^{-1/2}\alpha'_\perp C_0 S^{(m)}_{01,T}\begin{pmatrix} O_{k,m.r} \\ \sqrt{T}\left(\beta\Sigma_{\beta\beta}^{-1/2} \otimes I_m\right) \end{pmatrix} \xrightarrow{d} V'.$$

Under standard cointegration, the ML estimator $\widehat{\beta}$ of β, normalized as $\widetilde{\beta} = \widehat{\beta}\left(\beta'\widehat{\beta}\right)^{-1}\beta'\beta$, satisfies

$$T\left(\widetilde{\beta} - \beta\right) \xrightarrow{d} \beta_\perp\left(\int_0^1 W_{k-r}W'_{k-r}\right)^{-1}\left(\int_0^1 W_{k-r}d\underline{W}'_\alpha\right)\left(\alpha'\Omega^{-1}\alpha\right)^{-1/2}$$

[4] So that Lemma 2 in Andersson et al. (1983) can be applied.

where $\underline{W}_\alpha$ is an r-variate standard Wiener process which is independent of W_{k-r}. See Johansen (1988). In our case, however, the corresponding result is again quite different:

Lemma 5 *Let $\widehat{\xi}$ be the ML estimator of ξ, and denote $\widetilde{\xi} = \widehat{\xi}\left(\xi'\widehat{\xi}\right)^{-1}\xi'\xi$. Let $\xi_{\perp,T}$ be the orthogonal complement of ξ defined by* (37). *We can always write $\widetilde{\xi} - \xi = \xi_{\perp,T}U_{m,T}$, where*

$$U_{m,T} = \left(\xi'_{\perp,T}\xi_{\perp,T}\right)^{-1}\left(\xi'_{\perp,T}\widehat{\xi}\right)\left(\xi'\widehat{\xi}\right)^{-1}(\xi'\xi).$$

Under Assumptions 1-5,

$$T.U_{m,T} \xrightarrow{d} \begin{pmatrix} \left(\int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}'_{k-r,m}(x)dx\right)^{-1}\int_0^1 \widetilde{W}_{k-r,m}d\underline{W}'_\alpha \\ \underline{V}_\alpha \end{pmatrix} \times \left(\alpha'\Omega^{-1}\alpha\right)^{-1/2} \tag{38}$$

where $\underline{W}_\alpha$ is an r-variate standard Wiener process, $\underline{V}_\alpha$ is a $k.m \times r$ matrix with independent $N[0,1]$ distributed elements, and $\underline{V}_\alpha$, $\underline{W}_\alpha$ and $\widetilde{W}_{k-r,m}$ are independent. Consequently,

$$\begin{pmatrix} T.I_k & O_{k,k.m} \\ O_{k.m,k} & \sqrt{T}I_{k.m} \end{pmatrix}\left(\widetilde{\xi} - \xi\right) \xrightarrow{d} \begin{pmatrix} (\beta_\perp, O_{k,k.m})\left(\int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}'_{k-r,m}(x)dx\right)^{-1}\int_0^1 \widetilde{W}_{k-r,m}d\underline{W}'_\alpha \\ \left(\beta\Sigma_{\beta\beta}^{-1/2} \otimes I_m\right)\underline{V}_\alpha \end{pmatrix} \times \left(\alpha'\Omega^{-1}\alpha\right)^{-1/2},$$

Proof: Appendix.

The test for standard cointegration is based on a simple hypothesis, $\xi' = \left(\beta', O_{k.m,r}\right)$. The chi-square asymptotic distribution of the likelihood ratio statistic, derived in the Appendix, follows from the previous four Lemmas and the Taylor expansion around the MLE of a function of the type

$$f_{m,T}(x) = T.\ln\left(\frac{\det\left(x'\left(S_{11,T}^{(m)} - S_{10,T}^{(m)}S_{00,T}^{-1}S_{01,T}^{(m)}\right)x\right)}{\det\left(x'S_{11,T}^{(m)}x\right)}\right).$$

Then, we simply apply the Taylor expansion, derived in Johansen (1988), to the decomposition of the LR statistic LR_T^{tvc}

$$f_{m,T}(\widetilde{\xi}) - f_{0,T}(\widetilde{\beta}) = \left(f_{m,T}(\widetilde{\xi}) - f_{0,T}(\beta)\right) - \left(f_{0,T}(\widetilde{\beta}) - f_{0,T}(\beta)\right),$$

where similar to $\widetilde{\xi}$ defined in Lemma 5, $\widetilde{\beta} = \widehat{\beta}\left(\beta'\widehat{\beta}\right)^{-1}\beta'\beta$. It follows then from Lemma 5 that under the null hypothesis,

$$\begin{aligned} & f_{m,T}(\widehat{\xi}) - f_{0,T}(\beta) \xrightarrow{d} \operatorname{trace}\left(\underline{V}'_{\alpha}\underline{V}_{\alpha}\right) \\ & +\operatorname{trace}\left[\left(\int_0^1 d\underline{W}_{\alpha}\widetilde{W}'_{k-r,m}\right)\left(\int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}'_{k-r,m}(x)dx\right)^{-1}\right. \\ & \qquad \left.\times\left(\int_0^1 \widetilde{W}_{k-r,m}d\underline{W}'_{\alpha}\right)\right] \sim \chi^2_{r.m.r} + \chi^2_{r(m+1)(k-r)}, \end{aligned}$$

where the two chi-square distribution are independent, whereas it has been shown by Johansen (1988) that

$$\begin{aligned} & T\left(\widehat{f}_0\left(\widetilde{\beta}\right) - \widehat{f}_0\left(\beta\right)\right) \\ & \xrightarrow{d} \operatorname{trace}\left[\left(\int_0^1 d\underline{W}_{\alpha}W'_{k-r}\right)\left(\int_0^1 W_{k-r}(x)W'_{k-r}(x)dx\right)^{-1}\right. \\ & \qquad \left.\times\left(\int_0^1 W_{k-r}d\underline{W}'_{\alpha}\right)\right] \sim \chi^2_{r(k-r)}. \end{aligned}$$

It follows now straightforwardly that:

Theorem 1 *Given $m \geq 1$ and $r \geq 1$, under the null hypothesis of standard cointegration the LR statistics LR_T^{tvc} defined in* (21) *is asymptotically χ^2_{mkr} distributed.*

The limiting law is, not surprisingly, chi-square in which m, r and of course k are not allowed to be zero. Moreover, for a given pair (k,r), the distribution becomes less skewed as m increases. The optimal choice for m can be compared to the optimal choice of the order of an autoregressive process. With respect to the AR process researchers usually employ some sort of information criteria on model selection such as the Akaike (1974), Hannan-Quinn (1979) or Schwarz (1978) information criteria. The results in section 4 suggest that the Hannan-Quinn and Schwarz information criteria can be used to estimate m consistently if m is finite, but a formal proof is beyond the scope of this paper.

3.4 Empirical Size

To check how close the asymptotic critical values based on the χ^2 distribution are to the ones based on the small sample null distribution, we have applied our test to 10,000 replications of the bivariate cointegrated vector time series process $Y_t = (Y_{1,t}, Y_{2,t})'$, where $Y_{1,t} = Y_{2,t} + U_{1,t}$, $Y_{2,t} = Y_{2,t-1} + U_{2,t}$ with $U_t = (U_{1,t}, U_{2,t})'$ drawn independently from the bivariate standard normal distribution, for various values of T and m. The results are given in Table 1. In each entry, $q_{1-\alpha}$ stands for the empirical $1-\alpha$ quantile. Thus, they are the empirical critical values. The values in parenthesis are the acceptance frequencies based on the χ^2 critical values. The case $T = 324$ is included because this the sample size of the empirical application in section 6.

For large T and small m the right tail of the distribution is very well approximated by the asymptotic one. For smaller T the test suffers from size distortion. For example, for $T = 100$ and 5% asymptotic size the nominal size is of 3% for $m = 1$; 2% for $m = 3$ and 1.3% for $m = 5$. Thus, by using the asymptotic critical values the test tends to over-reject the correct null hypothesis of standard cointegration. As expected, for $T = 500$, the empirical and the asymptotic distributions almost coincide.

4 The LR Test under the Alternative of TV Cointegration

4.1 The Data Generating Process under TV Cointegration

A time-varying cointegrated data generating process Y_t with VECM(p) representation (1) can be constructed, for example, as follows. Let

$$Y_t = AZ_t = \alpha Z_{1,t} + \gamma Z_{2,t}$$

Table 1: Empirical Distribution of the LR TVC Statistic

		$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=10$	$m=15$	$m=25$
$T=100$	$q_{0.90}$	5.320 (0.930)	9.195 (0.943)	12.787 (0.953)	16.320 (0.961)	19.854 (0.969)	38.759 (0.992)	60.610 (0.999)	119.797 (0.999)
	$q_{0.95}$	7.027 (0.970)	11.159 (0.975)	15.111 (0.980)	18.829 (0.984)	22.643 (0.987)	42.543 (0.997)	65.650 (0.999)	127.696 (0.999)
	$q_{0.99}$	10.426 (0.994)	15.271 (0.995)	19.973 (0.997)	24.160 (0.997)	28.643 (0.998)	49.833 (0.999)	76.269 (0.999)	143.749 (0.999)
$T=200$	$q_{0.9}$	4.880 (0.912)	8.313 (0.919)	11.607 (0.928)	14.693 (0.934)	17.792 (0.941)	33.273 (0.968)	49.068 (0.984)	85.794 (0.998)
	$q_{0.95}$	6.406 (0.959)	10.065 (0.960)	13.595 (0.965)	16.896 (0.968)	20.395 (0.974)	37.000 (0.988)	53.424 (0.994)	91.731 (0.999)
	$q_{0.99}$	9.666 (0.992)	14.188 (0.993)	17.834 (0.993)	21.993 (0.995)	25.364 (0.995)	43.754 (0.998)	62.006 (0.999)	102.433 (0.999)
$T=324$	$q_{0.90}$	4.790 (0.908)	8.149 (0.913)	11.181 (0.917)	14.059 (0.919)	17.050 (0.926)	31.247 (0.947)	45.621 (0.966)	76.331 (0.990)
	$q_{0.95}$	6.275 (0.956)	10.015 (0.959)	13.197 (0.959)	16.400 (0.963)	19.452 (0.965)	34.608 (0.977)	49.515 (0.986)	81.177 (0.996)
	$q_{0.99}$	9.530 (0.991)	14.173 (0.993)	18.042 (0.993)	21.193 (0.993)	24.749 (0.994)	40.850 (0.996)	56.899 (0.997)	91.638 (0.999)
$T=500$	$q_{0.90}$	4.658 (0.902)	7.945 (0.906)	10.952 (0.910)	13.861 (0.914)	16.616 (0.916)	30.083 (0.931)	43.651 (0.948)	71.478 (0.975)
	$q_{0.95}$	6.088 (0.952)	9.709 (0.954)	13.119 (0.958)	16.138 (0.959)	19.075 (0.960)	33.377 (0.969)	47.753 (0.979)	76.213 (0.990)
	$q_{0.99}$	9.092 (0.989)	13.898 (0.992)	17.313 (0.991)	20.767 (0.992)	24.063 (0.992)	39.983 (0.994)	55.072 (0.996)	84.851 (0.998)

where $A = (\alpha, \gamma)$ is a nonsingular $k \times k$ matrix, with α the matrix of the first r columns of A and γ the matrix of the remaining $k-r$ columns of A, with $Z_{1,t} \in \mathbb{R}^r$ and $Z_{2,t} \in \mathbb{R}^{k-r}$ $I(1)$ processes generated by

$$Z_{1,t} = \sum_{j=1}^{p} D_j Z_{1,t-j} + B_2(t/T) Z_{2,t-1} + \sum_{j=1}^{p-1} C_{12,j} \Delta Z_{2,t-j} + U_{1,t} \quad (39)$$

$$\Delta Z_{2,t} = \sum_{j=1}^{p-1} C_{22,j} \Delta Z_{2,t-j} + U_{2,t} \quad (40)$$

where

Assumption 6. *$U_t = (U'_{1,t}, U'_{2,t})' \sim$ i.i.d. $N_k[0, V_u]$. The matrix valued lag polynomials $D(L) = I_r - \sum_{j=1}^{p} D_j L^j$ and $C_{22}(L) = I_{k-r} - \sum_{j=1}^{p-1} C_{22,j} L^j$ are invertible, with inverses $D(L)^{-1} = \sum_{j=0}^{\infty} \Pi_j L^j$ and $C_{22}(L)^{-1} = \sum_{j=0}^{\infty} \Gamma_j L^j$ satisfying $\Pi_j \to O$, $\Gamma_j \to O$ exponentially as $j \to \infty$. The elements of $B_2(\tau)$ are continuously differentiable function on an open interval containing $[0,1]$ with bounded derivatives, and*

$$B_2(\tau) = B_2(0) \text{ for } \tau < 0, B_2(\tau) = B_2(1) \text{ for } \tau > 1 \quad (41)$$

Note that the nonstationarity of $Z_{1,t}$ is due to the dependence of $Z_{1,t}$ on $B_2(t/T) Z_{2,t-1}$.

As is well-known, we can rewrite model (39) as

$$\Delta Z_{1,t} = B_1 Z_{1,t-1} + B_2(t/T) Z_{2,t-1} + \sum_{j=1}^{p-1} C_{11,j} \Delta Z_{1,t-j} + \sum_{j=1}^{p-1} C_{12,j} \Delta Z_{2,t-j} + U_{1,t} \quad (42)$$

where $B_1 = \sum_{j=1}^{p} D_j - I_r$ is nonsingular, [5] hence

$$\Delta Z_t = \begin{pmatrix} B_1 & B_2(t/T) \\ O_{k-r,r} & O_{k-r,k-r} \end{pmatrix} Z_{t-1} + \sum_{j=1}^{p-1} C_j \Delta Z_{t-j} + U_t \quad (43)$$

[5] Because the invertibility of $D(L)$ implies that all the roots of the polynomial $\det\left(I_r - \sum_{j=1}^{p} D_j x^j\right)$ lie outside the complex unit circle.

where

$$C_j = \begin{pmatrix} C_{11,j} & C_{12,j} \\ O_{k-r,r} & C_{22,j} \end{pmatrix},$$

and thus

$$\begin{aligned} \Delta Y_t &= A\Delta Z_t \\ &= A\begin{pmatrix} B_1 & B_2(t/T) \\ O_{k-r,r} & O_{k-r,k-r} \end{pmatrix} A^{-1}Y_{t-1} + \sum_{j=1}^{p-1} AC_j A^{-1}\Delta Y_{t-j} + AU_t \\ &= \alpha\beta_t' Y_{t-1} + \sum_{j=1}^{p-1}\Gamma_j \Delta Y_{t-j} + AU_t \end{aligned} \quad (44)$$

say, where

$$\beta_t' = (B_1, B_2(t/T))\, A^{-1},\ \Gamma_j = AC_jA^{-1}. \quad (45)$$

A more general TV model can be formulated by allowing B_1 to be a function of t/T as well, and by including lagged $\Delta Z_{1,t}$'s in the equation for $\Delta Z_{2,t}$. However, that will make the power analysis too complicated.

Under Assumption 6 the process Y_t is time-varying cointegrated, in the sense that with β_t define in (45),

$$\beta_t' Y_{t-1} = \varepsilon_t + O_p(1),$$

where ε_t is a strictly stationary zero-mean Gaussian process, and the $O_p(1)$ term is uniform in $t = 1, 2, ..., T$. This follows from the result (48) in the following lemma.

Lemma 6. *Under Assumption 6 we can write*

$$\Delta Z_{1,t} = \sum_{j=0}^{t-1}\Pi_j\left(B_2((t-j)/T) - B_2(0)\right)\Delta Z_{2,t-1-j} + V_t + O_p\left(1/\sqrt{T}\right), \quad (46)$$

uniformly in $t = 1, ..., T$, *where* V_t *is a strictly stationary zero-mean Gaussian process. Moreover, denote* $\sum_{j=0}^{\infty} Q_j L^j = D(L)^{-1}C_{11}(L)$, *where* $C_{11}(L) = I_r - \sum_{j=1}^{p-1} C_{11,j}L^j$. *Then*

$$\begin{aligned} B_1 Z_{1,t-1} + B_2(t/T)\, Z_{2,t-1} &= \sum_{j=0}^{t-1} Q_j\left(B_2((t-j)/T) - B_2(0)\right)\Delta Z_{2,t-1-j} \\ &\quad + R_t + O_p\left(1/\sqrt{T}\right) \end{aligned} \quad (47)$$

uniformly in $t = 1, ..., T$, where R_t is a strictly stationary zero-mean Gaussian process. Consequently,

$$B_1 Z_{1,t-1} + B_2 (t/T) Z_{2,t-1} = R_t + O_p(1) \tag{48}$$

uniformly in $t = 1, ..., T$.

Proof: Appendix

4.2 Power of the LR test

To study the power of our test, we will adopt the VECM(p) model (44) with β_t defined by (7) as the data generating process. Moreover, to keep the power analysis tractable we will focus on the case $p = 1,\ k = 2,\ r = 1,\ V_u = I_2,\ A = I_2$. Thus,

$$Y_t = Z_t = (Z_{1,t}, Z_{2,t})'$$

where $Z_{1,t} \in \mathbb{R}$ and $Z_{2,t} \in \mathbb{R}$ are assumed to be generated by

$$\begin{aligned} \Delta Z_{1,t} &= b_1 Z_{1,t-1} + b_2 (t/T) Z_{2,t-1} + U_{1,t} \\ \Delta Z_{2,t} &= U_{2,t} \\ U_t &= (U_{1,t}, U_{2,t})' \sim \text{ i.i.d. } N_2[0, I_2], \end{aligned} \tag{49}$$

Next, suppose that for some $m > 0$,

$$b_1^{-1} b_2 (t/T) = \sum_{j=0}^{m} \rho_j P_{j,T}(t),\ \ \rho' = (\rho_0, \rho_1, ..., \rho_m)$$

Then

$$(b_1, b_2 (t/T)) = b_1 \sum_{j=0}^{m} \varsigma_j' P_{j,T}(t),$$

where $\varsigma_0' = (1, \rho_0)$, and $\varsigma_j' = (0, \rho_j)$ for $j \geq 1$. Hence,

$$\begin{aligned} \Delta Z_{1,t} &= b_1 \left(Z_{1,t-1} + \sum_{j=0}^{m} \rho_j P_{j,T}(t) Z_{2,t-1} \right) + U_{1,t} \\ &= b_1 \sum_{j=0}^{m} \varsigma_j' P_{j,T}(t) Z_{t-1} + U_{1,t} = b_1 \varsigma' Z_{t-1}^{(m)} + U_{1,t} \\ \Delta Z_{2,t} &= U_{2,t} \end{aligned}$$

where

$$\varsigma' = (1, \rho_0, 0, \rho_1, 0, \rho_2, \ldots, 0, \rho_m) \tag{50}$$

and

$$Z_{t-1}^{(m)} = \begin{pmatrix} Z_{1,t-1}^{(m)} \\ Z_{2,t-1}^{(m)} \end{pmatrix} = Z_{t-1} \otimes \widehat{p}_m (t/T) \tag{51}$$

with

$$Z_{i,t-1}^{(m)} = \left(Z'_{i,t-1}, P_{1,T}(t) Z'_{i,t-1}, P_{2,T}(t) Z'_{i,t-1}, ..., P_{m,T}(t) Z_{i,t-1} \right)', \ i = 1, 2,$$

and

$$\widehat{p}_m (t/T) = (1, P_{1,T}(t), ..., P_{m,T}(t))'$$

We can now write the model in VECM(1) form as

$$\Delta Z_t = \delta \varsigma' Z_{t-1}^{(m)} + U_t \tag{52}$$

where

$$\delta = \begin{pmatrix} b_1 \\ 0 \end{pmatrix} \tag{53}$$

In the sequel we will refer to this model, together with the applicable parts of Assumptions 1-2, as $H_1^{(m)}(p = 1)$.

Under $H_1^{(m)}(p = 1)$ the matrices $S_{00,T}$, $S_{11,T}^{(m)}$ and $S_{01,T}^{(m)}$ become

$$S_{00,T} = \frac{1}{T} \sum_{t=1}^{T} \triangle Z_t \triangle Z_t' \tag{54}$$

$$S_{11,T}^{(m)} = \frac{1}{T} \sum_{t=1}^{T} Z_{t-1}^{(m)} Z_{t-1}^{(m)\prime} \tag{55}$$

$$S_{01,T}^{(m)} = \frac{1}{T} \sum_{t=1}^{T} \triangle Z_t Z_{t-1}^{(m)\prime} \tag{56}$$

respectively. The maximum log-likelihood in the standard case with $r = 1$ is

$$\begin{aligned} \widehat{l}_T(1,0) = & -\frac{1}{2} T. \ln \left(1 - \max_{\beta} \frac{\beta' S_{10,T}^{(0)} S_{00,T}^{-1} S_{01,T}^{(0)} \beta}{\beta' S_{11,T}^{(0)} \beta} \right) \\ & -\frac{1}{2} T. \ln (\det (S_{00,T})) - T.k \ln \left(\sqrt{2\pi} \right) - \frac{1}{2} kT \end{aligned}$$

and in the TV case

$$\begin{aligned}
\widehat{l}_T(1,m) &= -\frac{1}{2}T.\ln\left(1-\max_{\xi}\frac{\xi' S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(m)}\xi}{\xi' S_{11,T}^{(m)}\xi}\right) \\
&\quad -\frac{1}{2}T.\ln\left(\det\left(S_{00,T}\right)\right) - T.k\ln\left(\sqrt{2\pi}\right) - \frac{1}{2}kT
\end{aligned}$$

Thus,

$$p\lim_{T\to\infty} T^{-1}\left(\widehat{l}_T(1,m) - \widehat{l}_T(1,0)\right) > 0 \tag{57}$$

if $p\lim_{T\to\infty}\widehat{\lambda}_{\max}^{(0)} < p\lim_{T\to\infty}\widehat{\lambda}_{\max}^{(m)}$, where

$$\widehat{\lambda}_{\max}^{(0)} = \max_{\beta}\frac{\beta' S_{10,T}^{(0)} S_{00,T}^{-1} S_{01,T}^{(0)}\beta}{\beta' S_{11,T}^{(0)}\beta},\ \widehat{\lambda}_{\max}^{(m)} = \max_{\xi}\frac{\xi' S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(m)}\xi}{\xi' S_{11,T}^{(m)}\xi}.$$

Note that $\lambda_{\max}^{(m)}$ is the maximal solution of (16). Because

$$p\lim_{T\to\infty}\widehat{\lambda}_{\max}^{(m)} = p\lim_{T\to\infty}\frac{\varsigma' S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(m)}\varsigma}{\varsigma' S_{11,T}^{(m)}\varsigma}$$

where ς is defined by (50), the consistency of our test against the alternative $H_1^{(m)}(p=1)$ follows from the following theorem.

Theorem 2. *Under* $H_1^{(m)}(p=1)$,

$$p\lim_{T\to\infty}\frac{\varsigma' S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(m)}\varsigma}{\varsigma' S_{11,T}^{(m)}\varsigma} \in (0,1)\,,\ p\lim_{T\to\infty}\frac{\beta' S_{10,T}^{(0)} S_{00,T}^{-1} S_{01,T}^{(0)}\beta}{\beta' S_{11,T}^{(0)}\beta} = 0$$

for all nonzero vectors $\beta\in\mathbb{R}^2$, *hence* (57) *holds.*

The proof of Theorem 2 is not too difficult but tedious and lengthy. This proof is therefore given in a separate appendix to this paper, Bierens and Martins (2009). It is our conjecture that the result of Theorem 2 carry over to more general alternatives, but verifying this analytically proved to be too tedious an exercise. The same applies to the local power of the test. It is our conjecture that along the lines of the proof of Theorem 2 it can be shown that the test has nontrivial local power.

4.3 Empirical Power

The assumption that the time varying cointegrating vector can be exactly represented by a fixed number of Chebyshev polynomials is quite restrictive. Therefore, in this subsection we check via a limited Monte Carlo study how the test performs if this assumption is not true, and we compare our results with the tests of Park and Hahn (1999).

The data generating process we have used is

$$\begin{aligned} Z_{1,t} &= 0.75Z_{1,t-1} - 0.5f\left(t/T\right)Z_{2,t-1} - 0.25Z_{1,t-2} + U_{1,t} \\ \Delta Z_{2,t} &= U_{2,t}, t = 1, ..., T \end{aligned}$$

where $Z_{1,t} \in \mathbb{R}$, $Z_{2,t} \in \mathbb{R}$ and the error vectors $(U_{1,t}, U_{2,t})'$ are independently $N_2\left[0, I_2\right]$ distributed. The number of replications is $10,000$. For the function f we have chosen the following S-shaped function on $[0,1]$:

$$f\left(x\right) = 12\int_0^x y\left(1-y\right)dy - 1 = 6x^2 - 4x^3 - 1.$$

Note that $f\left(t/T\right)$ cannot be represented by a fixed number m of Chebyshev polynomials.

The results are presented in Table 2. In this table, α_{asy} indicates the asymptotic size, so that the rejection rates involved are with respect to the asymptotic critical values, whereas α_{real} is the empirical size, so that rejection rates involved are with respect to the empirical critical values in Table 1. These results suggests that the choice of m is not critical for the power.

Park and Hahn (1999) propose two types of test for TV cointegration, with statistics given by

$$\widehat{\tau}_1 = \frac{\sum_{t=1}^T \widehat{u}_t^2 - \sum_{t=1}^T \widehat{s}_t^2}{\widehat{\omega}_{T\kappa}^2}, \ \widehat{\tau}_2 = \frac{\sum_{t=1}^T \left(\sum_{i=1}^t \widehat{u}_i\right)^2}{T^2\widehat{\omega}_{T\kappa}^2},$$

where the $\widehat{u}_t$'s are the residuals of the regression of $Z_{1,t}$ on $Z_{2,t}$, the $\widehat{s}_t$'s are the residuals of the regression of $Z_{1,t}$ on $Z_{2,t}$ and t, $t^2, .., t^s$, and $\widehat{\omega}_{T\kappa}^2 = \frac{1}{T}\sum_{|k|<\ell_T} g\left(k/\ell_T\right)\sum_{t=k+1}^T \widehat{u}_{\kappa,t}\widehat{u}_{\kappa,t-k}$ is a long-run variance estimator, where the $\widehat{u}_{\kappa,t}$'s are the residuals of the regression of $Z_{1,t}$ on $\varphi_i\left(t/T\right)Z_{2,t}$ for $i = 1, ..., K$, with the φ_i's Fourier and other functions. As to the latter, we consider two cases, indicated by c:

$$\begin{aligned} c = 1 : \varphi_1(r) &= \cos(2\pi r),\ \varphi_2(r) = \sin(2\pi r),\ \varphi_3(r) = 1,\ \varphi_4(r) = r \\ c = 2 : \varphi_1(r) &= \cos(2\pi r),\ \varphi_2(r) = \sin(2\pi r),\ \varphi_3(r) = \cos(4\pi r), \\ \varphi_4(r) &= \sin(4\pi r),\ \varphi_5(r) = 1,\ \varphi_6(r) = r,\ \varphi_7(r) = r^2 \end{aligned}$$

Table 2: Power of the LR TVC test

$T = 100$	$\alpha_{asy} = 0.10$	$\alpha_{asy} = 0.05$	$\alpha_{asy} = 0.01$	$\alpha_{\text{real}} = 0.05$
$m = 1$	0.999	0.999	0.997	0.999
$m = 3$	0.998	0.997	0.992	0.997
$m = 5$	0.998	0.997	0.986	0.996
$m = 10$	0.998	0.997	0.987	0.994
$m = 15$	0.999	0.998	0.994	0.995
$T = 200$	$\alpha_{asy} = 0.10$	$\alpha_{asy} = 0.05$	$\alpha_{asy} = 0.01$	$\alpha_{\text{real}} = 0.05$
$m = 1$	1.000	1.000	1.000	1.000
$m = 3$	1.000	1.000	1.000	1.000
$m = 5$	1.000	1.000	1.000	1.000
$m = 10$	1.000	1.000	1.000	1.000
$m = 15$	1.000	1.000	1.000	1.000

Table 3: Power of the Park-Hahn tests

$\widehat{\tau}_1, \alpha_{asy} = 0.05$	$T = 100$	$T = 200$
$s = 4, c = 2$	0.993	0.998
$s = 1, c = 2$	0.877	0.936
$s = 4, c = 1$	0.991	0.998
$\widehat{\tau}_2, \alpha_{asy} = 0.05$	$T = 100$	$T = 200$
$c = 2$	0.999	1.000
$c = 1$	0.998	1.000

The statistic $\widehat{\tau}_1$ also depends on the polynomial order s. We consider the cases $s = 1$ and $s = 4$. Note that the test $\widehat{\tau}_2$ is in essence the well-known KPSS test. See Kwiatkowski et al. (1992). Finally, we use for g the Barlett kernel, the truncation lag is $\ell_T = [T^{1/3}]$, the number of replications is $10,000$, and $T = 100, 200$. The results are presented in Table 3.

These results show that all three tests perform very well.

5 The Drift Case

The Assumptions 1-2 imply that ΔY_t and $\beta' Y_t$ are zero-mean stationary processes. However, for most cointegrated macroeconomic time series ΔY_t and $\beta' Y_t$ are nonzero-mean stationary processes, which correspond to the following modification of Assumption 1:

Assumption 1*. *$\triangle Y_t = C(L)(U_t + \mu) = \sum_{j=0}^{\infty} C_j (U_{t-j} + \mu)$, where $\mu \neq 0$ is a vector of imbedded drift parameters, and U_t and $C(L)$ are the same as in Assumption 1.*

Then similar to (24) we can write

$$Y_t = C(1) \sum_{j=1}^{t} U_j + C(1)\mu.t + V_t + Y_0 - V_0 \tag{58}$$

Under Assumption 2,

$$\beta' Y_t = \beta' V_t + \beta' (Y_0 - V_0).$$

Thus, Assumption 2 can be adopted without modifications, but Assumption 3 needs to be dropped as otherwise $\beta' (Y_0 - V_0) = 0$. However, due to the drift we now need to include a vector of intercepts in VECM (25), as in Johansen (1991):

Assumption 4*. *ΔY_t has the VECM(p) representation*

$$\triangle Y_t = \gamma_0 + \alpha \beta' Y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \triangle Y_{t-j} + C_0 U_t \tag{59}$$

Moreover, with $X_t = (\triangle Y_{t-1}, ..., \triangle Y_{t-p+1})'$, Assumption 5 still applies. These modified Assumptions 1-5 will be referred to as "the drift case".

The corresponding time-varying VECM(p) is now

$$\triangle Y_t = \gamma_0 + \alpha \xi' Y_{t-1}^{(m)} + \sum_{j=1}^{p-1} \Gamma_j \triangle Y_{t-j} + C_0 U_t \tag{60}$$

To re-derive our previous results for this drift case, we need some additional notation. First, let

$$\overline{\mu} = \left(\mu' C_0 \alpha_\perp \left(\alpha'_\perp \Omega \alpha_\perp\right)^{-1} \alpha'_\perp C'_0 \mu\right)^{-1/2} \left(\alpha'_\perp \Omega \alpha_\perp\right)^{-1/2} \alpha'_\perp C'_0 \mu$$

which is a vector in $\mathbb{R}^{k-r}$. Note that $\overline{\mu}'\overline{\mu} = 1$ by normalization. Let $\overline{\mu}_\perp$ be an orthogonal complement of $\overline{\mu}$, normalized such that $\overline{\mu}'_\perp \overline{\mu}_\perp = I_{k-r-1}$. Then

Lemma 7. *In the drift case,*

$$\begin{aligned} \left(\overline{\mu}'_\perp \otimes I_{m+1}\right) \left(\beta'_\perp \otimes I_{m+1}\right) \frac{1}{\sqrt{T}} Y_{[xT]}^{(m)} &\Rightarrow p(x) \otimes \underline{W}_{k-r-1}(x) \\ \left(\overline{\mu}' \otimes I_{m+1}\right) \left(\beta'_\perp \otimes I_{m+1}\right) \frac{1}{T} Y_{[x.T]}^{(m)} &\Rightarrow p(x) \otimes x \end{aligned}$$

for $x \in [0,1]$, *where* $p(x) = \left(1, \sqrt{2}\cos(\pi x), ..., \sqrt{2}\cos(m\pi x)\right)'$ *and*

$$\underline{W}_{k-r-1} = \overline{\mu}'_\perp \left(\alpha'_\perp \Omega \alpha_\perp\right)^{-1/2} \alpha'_\perp C'_0 W \tag{61}$$

is a $(k-r-1)$*-variate standard Wiener process.*

Next, let

$$M_T = \left(T^{-1/2}\overline{\mu}, \overline{\mu}_\perp\right).$$

Redefine the orthogonal complement $\xi_{\perp,T}$ of ξ in (37) as

$$\xi_{\perp,T} = \left(\left(\beta_\perp \otimes I_{m+1}\right)\left(M_T \otimes I_{m+1}\right), \begin{pmatrix} O_{k,m.r} \\ \sqrt{T}\left(\beta \Sigma_{\beta\beta}^{-1/2} \otimes I_m\right) \end{pmatrix} \right) \tag{62}$$

and redefine $\widetilde{W}_{k-r,m}$ in (35) as

$$\begin{aligned} \widetilde{W}_{k-r,m}(x) &= p(x) \otimes \begin{pmatrix} \underline{W}_{k-r-1}(x) \\ x \end{pmatrix} \\ &\quad - \int_0^1 p(y) \otimes \begin{pmatrix} \underline{W}_{k-r-1}(y) \\ y \end{pmatrix} dy, \end{aligned} \tag{63}$$

where $\underline{W}_{k-r-1}$ is defined by (61). Then

Theorem 3. *With* $\xi_{\perp,T}$ *in* (37) *replaced by* (62) *and* $\widetilde{W}_{k-r,m}$ *in* (35) *replaced by* (63) *the results of Lemmas 3-5 and Theorem 1 carry over.*

The proofs of Lemma 7 and Theorem 3 are not too difficult but rather lengthy. These proofs are therefore given in a separate appendix to this paper, Bierens and Martins (2009).

6 An Empirical Application

The validity of the PPP hypothesis has generated a great deal of controversy, intimately related to the type of method applied. Recently, Falk and Wang (2003) found that PPP hypothesis holds against some economies but not for all. Their work is based on Caner's (1998) concept of cointegration where the VECM errors follow a stable distribution.

A reason why linear VECM models may be unable to detect long run PPP is the presence of transaction costs in equilibrium models of real exchange rate determination, which imply a nonlinear adjustment process in the PPP relationship. Michael et al. (1997) successfully fit an exponential smooth transition autoregressive model, thus capturing the implied nonlinearities.

We propose an alternative framework where the cointegrating vectors fluctuate over the sample. We test the constant cointegration hypothesis against the time varying cointegration, for

$$Y_t = \left(\ln S_t^f, \ln P_t^n, \ln P_t^f\right)',$$

where P_t^n and P_t^f are the price indices in the domestic and foreign economies, respectively, and S_t^f is the nominal exchange rate in home currency per unit of the foreign currency. Since the log-prices are unit root with drift processes the tests will be conducted under the "drift-case" assumptions. The time-varying cointegrating relation is $\beta_t' Y_t = e_t$, where the process e_t represents the short run deviations from the PPP due to disturbances in the economic system (real or monetary shocks), and β_t is a 3×1 unknown deterministic function of time. Using Chebyshev time polynomials, $P_{i,T}(t)$, β_t will be approximated by $\beta_t(m) = \sum_{i=0}^{m} \xi_i P_{i,T}(t)$, where the ξ_i's are the Fourier coefficients.

We use the same data as Falk and Wang (2003), downloaded from the Journal of Applied Econometrics data archive website. The domestic country is the US and the bilateral relationship of study is with Canada, France, Germany, Italy, Japan, and the U.K.. The data are monthly and cover the period from January 1973 to December 1999, so that the time series involved have length 324. Falk and Wang (2003) find support for the presence of unit roots in all series. By means of the standard Johansen's approach, they find support of the PPP hypothesis at the 5% level in eight of the twelve cases. With one cointegrating vector, Belgium, Denmark, France, Japan, Netherlands, Norway, Spain, and UK are found to have PPP with the US. At the 10% level, Italy and Sweden were added to the list. Therefore, Canada and Germany were the only countries for which US has not had price parity according to the standard approach.

The asymptotic p-values[6] of our test are presented in Table 4, for different combinations of the order m of the Chebyshev polynomial and the lag order p. Because the results did not vary much with p we only report the results for $p = 1$, $p = 6$, $p = 10$ (Falk and Wang's lag), and $p = 18$. Moreover, we have computed the p-values for m ranging from 1 to 25, although we only present the nonzero p-values.

We find that, regardless of the lag order, the p-values are zero for any m larger than four. Hence, there is strong evidence of a time varying type of cointegration between international prices and nominal exchange rates for all cases. Thus, our results refute Falk and Wang's findings of standard PPP for all countries except Canada and Germany.

The plots of the time-varying coefficients β_{1t}, β_{2t} and β_{3t} in the cointegrating PPP relation $\beta_t' Y_t = \beta_{1t} \ln S_t^f + \beta_{2t} \ln P_t^n + \beta_{3t} \ln P_t^f$ are presented in the separate appendix Bierens and Martins (2009). The patterns of these parameters suggest that, approximately, $\beta_{2t} + \beta_{3t} = \delta$ for some constant δ. This is related to the symmetry assumption in the standard PPP theory, where $\beta_3 + \beta_2 = 0$. However, δ seems to be positive for Canada and the UK, and negative for the other countries. Moreover, the variation of β_{1t} is minor compared with the variation of β_{2t} and β_{3t}, suggesting that β_{1t} may be constant.

[6] $P\left(\chi^2_{(mrk)} \geq LR^{tvc}\right)$.

Table 4: LR TVC Standard Cointegration Tests: Asymptotic p-values

	m	$p=1$	$p=6$	$p=10$	$p=18$
Can	≥ 1	0.000	0.000	0.000	0.000
Fra	≥ 1	0.000	0.000	0.000	0.000
Ger	1	0.158	0.461	0.033	0.010
	≥ 2	0.000	0.000	0.000	0.000
Ita	≥ 1	0.000	0.000	0.000	0.000
Jap	1	0.662	0.383	0.001	0.019
	2	0.102	0.457	0.001	0.004
	3	0.029	0.013	0.000	0.000
	4	0.037	0.014	0.000	0.000
	≥ 5	0.000	0.000	0.000	0.000
$U.K.$	1	0.119	0.053	0.242	0.023
	2	0.194	0.039	0.011	0.001
	3	0.004	0.000	0.000	0.000
	≥ 4	0.000	0.000	0.000	0.000

7 Conclusion

In Johansen's standard approach it is assumed that the cointegrating vector is constant over time. This assumption may be restrictive in practice due to changes in taste, technology, or economic policies. We propose a generalization of the standard approach by allowing the cointegration vector to be time-varying and we approximate it using orthogonal Chebyshev time polynomials. In time-varying cointegration, the long run relationship is a dependent, heterogeneously distributed process.

We propose a cointegration model that captures smooth time transitions on the cointegrating vector - the time varying error correction model - and estimate it by maximum likelihood. To distinguish our model from the time invariant Johansen's specification, we construct a likelihood ratio test for the null hypothesis of standard cointegration. The limiting law appears to be chi-square. To illustrate the practical relevancy of our approach we applied out test to international prices and nominal exchange rates. We find evidence of time-varying cointegration between these series.

There are issues that merit further research. In particular, the analytical study of the power of the test against local alternatives deserves attention. Moreover, a natural extension of our approach is to include deterministic components such time trends and/or seasonal dummy variables, and to allow for other time varying parameters.

References

Akaike, H. (1974), A New Look at the Statistical Model Identification, *I.E.E.E. Transactions on Automatic Control* **19**, 716-723.

Andersson, S. A., H. K. Brons and S. T. Jensen (1983), Distribution of Eigenvalues in Multivariate Statistical Analysis. *Annals of Statistics* **11**, 392-415.

Andrade, P., C. Bruneau and S. Gregoir (2005), Testing for the Cointegration Rank when some Cointegrating Directions are Changing. *Journal of Econometrics* **124**, 269-310.

Andrews, D. W. K., I. Lee and W. Ploberger (1996), Optimal changepoint tests for normal linear regression. *Journal of Econometrics* **70**, 9-38.

Bai, J., R. L. Lumsdaine and J. H.Stock (1998), Testing for and Dating Breaks in Multivariate Time Series. *Review of Economic Studies* **65**, 395-432.

Bierens, H. J. (1994), *Topics in Advanced Econometrics: Estimation, Testing and Specification of Cross-Section and Time Series Models.* Cambridge University Press.

Bierens, H. J. (1997), Testing the Unit Root with Drift Hypothesis Against Nonlinear Trend Stationarity, with an Application to the US Price Level and Interest Rate. *Journal of Econometrics* **81**, 29-64.

Bierens, H. J. (2007), Weak Convergence to the Matrix Stochastic Integral $\int_0^1 BdB'$ in the Gaussian Case, with Application to Likelihood-Based Cointegration Analysis. Lecture notes, downloadable from http://econ.la.psu.edu/~hbierens/LECNOTES.HTM

Bierens, H. J. and L. Martins (2009), Separate Appendix to "Time Varying Cointegration", downloadable from http://econ.la.psu.edu/~hbierens/TVCOINT_APPENDIX.PDF

Blake, N. S. and T. B. Fomby (1997), Threshold Cointegration. *International Economic Review* **38**, 627-645.

Caner, M. (1998), Tests for Cointegration with Infinite Variance Errors. *Journal of Econometrics* **86**, 155-175.

De Jong, R. M. (2001), Nonlinear Estimation Using Estimated Cointegrating Relations. *Journal of Econometrics* **101**, 109-122.

Engle, R. F., and C. W. J. Granger (1987), Cointegration and Error Correction: Representations, Estimation and Testing. *Econometrica* **55**, 251-276.

Falk, B. and C. Wang (2003), Testing Long Run PPP with Infinite Variance Returns. *Journal of Applied Econometrics* **18**, 471-484.

Granger, C.W. J. and G. Yoon (2002), Hidden Cointegration. Working Paper, Department of Economics, UCSD.

Hall, S. G., Z. Psaradakis and M. Sola (1997), Cointegration and Changes in Regime: the Japanese Consumption Function. *Journal of Applied Econometrics* **12**, 151-168.

Hamming, R. W. (1973), *Numerical Methods for Scientists and Engineers.* Dover, New York.

Hannan, E. J., and B. G. Quinn (1979), The Determination of the Order of an Autoregression, *Journal of the Royal Statistical Society B* **41**, 190-195.

Hansen, B. E. (1992), Tests for Parameter Instability in Regressions with $I(1)$ Processes. *Journal of Business and Economic Statistics* **10**, 321-335.

Hansen, P. R. (2003), Structural Changes in the Cointegrated Vector Autoregressive Model. *Journal of Econometrics* **114**, 261-295.

Hansen, H. and S. Johansen (1999), Some Tests for Parameter Constancy in Cointegrated VAR-models. *Econometrics Journal* **2**, 306-333.

Hao, K. (1996), Testing for Structural Change in Cointegrated Regression Models: Some Comparisons and Generalizations. *Econometric Reviews* **15**, 401-429.

Harris, D., B. McCabe and S. Leybourne (2002), Stochastic Cointegratione: Estimation and Inference. *Journal of Econometrics* **111**, 363-384.

Inoue, A. (1999), Tests of Cointegration Rank with a Trend Break. *Journal of Econometrics* **90**, 215-237.

Johansen, S. (1988), Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control* **12**, 231-254.

Johansen, S. (1991), Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* **59**, 1551-1580.

Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Johansen, S., R. Mosconi and B. Nielsen (2000), Cointegration Analysis in the Presence of Structural Breaks in the Deterministic Trend. *Econometrics*

Journal **3**, 216-249.

Juhl, T. and Z. Xiao (2005), Testing for Cointegration using Partially Linear Models. *Journal of Econometrics* **124**, 363-394.

Kronmal, R. and M. Tarter (1968), The Estimation of Densities and Cumulatives by Fourier Series Methods. *Journal of the Americal Statistical Association* **63**, 925-952.

Kuo, B. S. (1998), Test for Partial Parameter Instability in Regressions with I(1) processes. *Journal of Econometrics* **86**, 337-368.

Kwiatkowski, D., P. Phillips, P. Schmidt, and Y. Shin (1992), Testing the Null of Stationarity Against the Alternative of a Unit Root. *Journal of Econometrics* **54**, 159-178.

Lütkepohl, H., P. Saikkonen and C. Trenkler (2003), Comparison of Tests for the Cointegrating Rank of a VAR Process with a Structural Shift. *Journal of Econometrics* **113**, 201-229.

Lütkepohl, H., T. Terasvirta and J. Wolters (1999), Investigating Stability and Linearity of a German M1 Money Demand Function. *Journal of Applied Econometrics* **14**, 511-525.

Maddala, G. S. and I-M. Kim (1998), *Unit Roots, Cointegration and Structural Change.* Cambridge University Press.

Martins, L. F. (2005), *Structural Changes in Nonstationary Time Series Econometrics.* Ph.D. Dissertation, Pennsylvania State University, downloadable from http://etda.libraries.psu.edu/theses/approved/WorldWideFiles/ETD-794/LuisPhDThesisFinal.pdf.

McLeish, D. L. (1974), Dependent Central Limit Theorems and Invariance Principles. *Annals of Probability* **2**, 620-628.

Michael, P., A. R. Nobay and D. A. Peel (1997), Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation. *Journal of Political Economy* **105**, 862-879.

Oksendal, B. (2003), *Stochastic Differential Equations.* Springer-Verlag.

Park, J. Y. and S. B. Hahn (1999), Cointegrating Regressions with Time Varying Coefficients. *Econometric Theory* **15**, 664–703.

Phillips, P.C.B. (1988), Weak Convergence to the Matrix Stochastic Integral $\int_0^1 BdB'$. *Journal of Multivariate Time Series Analysis* **24**, 252-264.

Phillips, P. C. B. and S. N. Durlauf (1986), Multiple Time Series Regression with Integrated Processes. *Review of Economic Studies* **53**, 473-496.

Quintos, C. E. (1997), Stability Tests in Error Correction Models. *Journal of Econometrics* **82**, 289-315.

Quintos, C. E. and P. C. B. Phillips (1993), Parameter Constancy in Cointegrating Regressions. *Empirical Economics* **18**, 675-706.

Saikkonen, P. and I. Choi (2004), Cointegrating Smooth Transition Regressions. *Econometric Theory* **20**, 301–340.

Seo, B. (1998), Tests for Structural Change in Cointegrated Systems. *Econometric Theory* **14**, 222-259.

Schwarz, G. (1978), Estimating the Dimension of a Model, *Annals of Statistics* **6**, 461-464.

Terasvirta, T. and A. C. Eliasson (2001), Non-Linear Error Correction and the UK Demand for Broad Money, 1878-1993. *Journal of Applied Econometrics* **16**, 277-288.

Young, N. (1988), *An Introduction to Hilbert Space.* Cambridge University Press.

8 Appendix: Proofs

8.1 Proof of Lemma 1

The continuity of $\varphi(x)$ on $[0,1]$ implies that $\varphi(x)$ is uniformly continuous on $[0,1]$ and therefore bounded on $[0,1]$: $\max_{0\leq x\leq 1}|\varphi(x)| < \infty$. Consequently, $\varphi(x)$ is square integrable on $[0,1]$, and thus is an element of the Hilbert space $L^2[0,1]$ of square integrable real functions on $[0,1]$, with inner product $\langle f,g\rangle = \int_0^1 f(x)g(x)dx$ and associated norm $||f|| = \sqrt{\langle f,f\rangle}$ and metric $||f-g||$.

We first show that the sequence

$$\kappa_j(x) = \begin{cases} 1 & \text{for} \quad j=0, \\ \sqrt{2}\cos(j\pi x) & \text{for} \quad j=1,2,3,.... \end{cases} \tag{64}$$

is a complete orthonormal sequence in $L^2[0,1]$. This result is well-known in the statistics literature (see for example Kronmal and Tarter 1968), but to the best of our knowledge has not been used in the econometrics literature. Therefore, we will prove it here, as follows.

Recall[7] that the functions

$$1,\ \sqrt{2}\cos(i\pi x),\ \sqrt{2}\sin(j\pi x),\ i,j=1,2,3,....,\ x\in[-1,1],$$

form a complete orthonormal sequence in $L^2[-1,1]$ with respect to the uniform density on $[-1,1]$, hence every real function $\psi \in L^2[-1,1]$ satisfies

$$\lim_{\substack{m\to\infty \\ n\to\infty}} \frac{1}{2}\int_{-1}^{1}(\psi(x)-\psi_{m,n}(x))^2 dx \tag{65}$$

where

$$\psi_{m,n}(x) = \underline{\omega}_0 + \sum_{i=1}^{m}\omega_i\sqrt{2}\cos(i\pi x) + \sum_{j=1}^{n}\varpi_j\sqrt{2}\sin(j\pi x) \tag{66}$$

with Fourier coefficients $\underline{\omega}_0 = \frac{1}{2}\int_{-1}^{1}\psi(x)dx$, $\omega_k = \frac{1}{2}\int_{-1}^{1}\sqrt{2}\cos(k\pi x)\psi(x)dx$ and $\varpi_k = \frac{1}{2}\int_{-1}^{1}\sqrt{2}\sin(k\pi x)\psi(x)dx$. Now let $\varphi(u) \in L^2[0,1]$ be arbitrary,

[7] See for example Young (1988)

and let $\psi(x) = \varphi(|x|)$. Then $\psi(x) \in L^2(-1,1)$, with Fourier coefficients

$$\begin{aligned} \underline{\omega}_0 &= \frac{1}{2}\int_{-1}^{1} \varphi(|x|)\,dx = \int_0^1 \varphi(u)\,du \\ \omega_k &= \frac{1}{2}\int_{-1}^{1} \sqrt{2}\cos(k\pi x)\,\varphi(|x|)\,dx = \int_0^1 \sqrt{2}\cos(k\pi u)\,\varphi(u)\,du \\ \varpi_k &= \frac{1}{2}\int_{-1}^{1} \sqrt{2}\sin(k\pi x)\,\varphi(|x|)\,dx = 0 \end{aligned}$$

Hence it follows from (65) and (66) that

$$\begin{aligned} &\lim_{n\to\infty}\int_0^1 \left(\varphi(u) - \underline{\omega}_0 - \sum_{k=1}^{n}\omega_k\sqrt{2}\cos(k\pi u)\right)^2 du \qquad (67) \\ &= \frac{1}{2}\lim_{n\to\infty}\int_{-1}^{1}\left(\varphi(|x|) - \underline{\omega}_0 - \sum_{k=1}^{n}\omega_k\sqrt{2}\cos(k\pi x)\right)^2 dx \\ &= 0 \end{aligned}$$

This proves the completeness of (64).

Next, let $t_x = [xT]+1$ for an $x \in [0,1)$, where $[xT]$ is the largest integer $\leq xT$. Then

$$g_m(t_x) = g_m([xT]+1) = \xi_{0,T} + \sqrt{2}\sum_{i=1}^{m}\xi_{i,T}\cos\left[i\pi\left(\frac{[xT]+1}{T} - \frac{1}{2T}\right)\right]$$

where

$$\begin{aligned} \xi_{0,T} &= \frac{1}{T}\sum_{t=1}^{T}\varphi(t/T) = \int_0^1 \varphi\left(\frac{[yT]+1}{T}\right)dy \\ \xi_{i,T} &= \frac{1}{T}\sum_{t=1}^{T}\varphi(t/T)\sqrt{2}\cos[i\pi(t-0.5)/T] \\ &= \int_0^1 \varphi\left(\frac{[yT]+1}{T}\right)\sqrt{2}\cos\left[i\pi\left(\frac{[yT]+1}{T} - \frac{1}{2T}\right)\right]dy \end{aligned}$$

Hence by bounded convergence,

$$\varphi_m(x) = \lim_{T\to\infty} g_m([xT]+1) = \xi_0 + \sum_{i=1}^{m}\xi_i\sqrt{2}\cos(i\pi x)\,, \text{ where}$$

$$\xi_0 = \lim_{T\to\infty} \xi_{0,T} = \int_0^1 \varphi(y)\, dy,$$
$$\xi_i = \lim_{T\to\infty} \xi_{i,T} = \int_0^1 \varphi(y)\sqrt{2}\cos(i\pi y)dy,\ i \geq 1.$$

It is now easy to verify that

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} (g(t) - g_{m,T}(t))^2 = \int_0^1 (\varphi(x) - \varphi_m(x))^2\, dx \tag{68}$$

Note that

$$\int_0^1 (\varphi_m(x) - \varphi(x))^2\, dx = \int_0^1 \varphi(x)^2\, dx - \sum_{i=1}^{m} \xi_i^2 > 0,$$

hence $\sum_{i=1}^{\infty} \xi_i^2 \leq \int_0^1 \varphi(x)^2\, dx < \infty$. However, due to the completeness of (64), we also have $\sum_{i=1}^{\infty} \xi_i^2 = \int_0^1 \varphi(x)^2\, dx$. See for example Young (1988, Theorem 4.15, p.37). Thus,

$$\lim_{m\to\infty} \int_0^1 (\varphi(x) - \varphi_m(x))^2\, dx = 0. \tag{69}$$

Combining (68) and (69), the first part of Lemma 1 follows.

To prove the second part of Lemma 1, suppose that $\varphi(x)$ is q times differentiable, where $q \geq 2$ is *even*, and that $\varphi^{(q)}(x) = d^q\varphi(x)/(dx)^q$ is square-integrable. Then $\varphi^{(q)}(x) \in L^2[0,1]$:

$$\lim_{m\to\infty} \int_0^1 \left(\varphi^{(q)}(x) - \sum_{i=1}^{m}(-1)^{q/2}\pi^q i^q \xi_i \sqrt{2}\cos(i\pi x)\right)^2 dx,$$

where $\int_0^1 \left(\varphi^{(q)}(x)\right)^2 dx = \pi^{2q}\sum_{i=1}^{\infty} i^{2q}\xi_i^2 < \infty$. Now for $m \geq 1$,

$$\int_0^1 (\varphi_m(x) - \varphi(x))^2\, dx = \sum_{i=m+1}^{\infty} \xi_i^2 \leq \sum_{i=m+1}^{\infty} \xi_i^2 \left(\frac{i}{m+1}\right)^{2q}$$
$$\leq \frac{1}{\pi^{2q}(m+1)^{2q}}\pi^{2q}\sum_{i=1}^{\infty} \xi_i^2 i^{2q} = \frac{\int_0^1 \left(\varphi^{(q)}(x)\right)^2 dx}{\pi^{2q}(m+1)^{2q}}.$$

8.2 Proof of Lemma 2

To prove Lemma 2 we need two auxiliary lemmas:

Lemma A.1. *Under Assumptions 1-2,*

$$\frac{1}{T}\sum_{t=1}^{T}(\triangle Y_{t-\ell})\,Y_{t-1}' \xrightarrow{d} C(1)\left(\int_0^1 (dW)\,W'\right)C(1)' + M_\ell,\ \ \ell \geq 0,$$
$$\frac{1}{T}\sum_{t=1}^{[xT]}(\triangle Y_{t})\,Y_{t-1}' \xrightarrow{d} C(1)\left(\int_0^x (dW)\,W'\right)C(1)' + xM_0,$$
$$\frac{1}{T}\sum_{t=1}^{T}U_t Y_{t-1}' \xrightarrow{d} \int_0^1 (dW)\,W'C(1)'.$$

where W is a k-variate standard Wiener process, and the M_ℓ's are non-random $k\times k$ matrices.

Proof: Phillips and Durlauf (1986) and Phillips (1988).

Lemma A.2. *Let η_t be an arbitrary sequence in $\mathbb{R}^n$, and let $F(x)$ be an arbitrary differentiable function on $[0,1]$, with derivative $f(x)$. Then*

$$\sum_{t=1}^{T}\eta_t F(t/T) = \sum_{t=1}^{T}\eta_t F(1) - \int_0^1 f(x)\left(\sum_{t=1}^{[xT]}\eta_t\right)dx$$

Proof: Bierens (1994), Lemma 9.6.3, page 200.

Proof of Lemma 2: It follows from (22) and (23) that

$$Y_t = Y_0 - V_0 + C(1)\sum_{j=1}^{t}U_j + V_t$$

Assumption 3 implies that $Y_0 - V_0 = 0$, but there is no need to impose this restriction here. Now

$$\frac{1}{T}\sum_{t=1}^{T}P_{j,T}(t)\,U_t Y_{t-1}' = \sqrt{2}\frac{1}{T}\sum_{t=1}^{T}\cos\left(j.\pi\,(t-0.5)\,/T\right)U_t' Y_{t-1} \quad (70)$$

$$
\begin{aligned}
&= \sqrt{2}\frac{1}{T}\sum_{t=1}^{T}\cos\left(j.\pi\left(t-0.5\right)/T\right)U_t\sum_{j=1}^{t-1}U_j'C(1)' \\
&+ \sqrt{2}\frac{1}{T}\sum_{t=1}^{T}\cos\left(j.\pi\left(t-0.5\right)/T\right)U_tV_{t-1}' \\
&+ \sqrt{2}\frac{1}{T}\sum_{t=1}^{T}\cos\left(j.\pi\left(t-0.5\right)/T\right)U_t\left(Y_0-V_0\right)'
\end{aligned}
$$

It is easy to verify that the last two terms are of order $O_p\left(1/\sqrt{T}\right)$. Moreover, it follows from Lemma A.2 that

$$
\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\cos\left(j.\pi\left(t/T-0.5/T\right)\right)U_t\sum_{j=1}^{t-1}U_j' \\
&\quad = \cos\left(j.\pi\left(1-0.5/T\right)\right)\frac{1}{T}\sum_{t=1}^{T}U_t\sum_{j=1}^{t-1}U_j' \\
&\quad + j.\pi\int_0^1\sin\left(j.\pi\left(x-0.5/T\right)\right)\left(\frac{1}{T}\sum_{t=1}^{[xT]}U_t\sum_{j=1}^{t-1}U_j'\right)dx
\end{aligned}
$$

which by Lemma A.1 and the continuous mapping theorem converges in distribution to

$$
\begin{aligned}
&\cos\left[j.\pi\right]\int_0^1\left(dW\right)W' + j.\pi\int_0^1\sin\left(j.\pi x\right)\left(\int_0^x\left(dW\right)W'\right)dx \\
&= \cos\left[j.\pi\right]\int_0^1\left(dW\right)W' - \int_0^1\frac{d\cos\left(j.\pi x\right)}{dx}\left(\int_0^x\left(dW\right)W'\right)dx \\
&= \int_0^1\left(dW(x)\right)\cos\left(j.\pi x\right)W(x)'
\end{aligned}
$$

The latter follows via integration by parts. Part (29) now follows easily from these results.

To prove (30), observe that

$$
\frac{1}{T}\sum_{t=1}^{T}\cos\left(j.\pi\left(t-0.5\right)/T\right)\left(\triangle Y_{t-\ell}\right)Y_{t-1}'
$$

$$= \frac{1}{T}\sum_{t=1}^{T}\cos\left(j.\pi\left(t-0.5\right)/T\right)\left(\triangle Y_{t-\ell}\right)\left(Y_{t-1}-Y_{t-1-\ell}\right)'$$
$$+\frac{1}{T}\sum_{t=1}^{T}\cos\left(j.\pi\left(t-0.5\right)/T\right)\left(\triangle Y_{t-\ell}\right)Y_{t-1-\ell}'$$

Again, it follows from Lemmas A.1-A.2 that

$$\frac{1}{T}\sum_{t=1}^{T}\cos\left(j.\pi\left(t-0.5\right)/T\right)\left(\triangle Y_{t-\ell}\right)Y_{t-\ell-1}$$
$$= \cos\left(j.\pi\left(1-0.5/T\right)\right)\frac{1}{T}\sum_{t=1}^{T}\left(\triangle Y_{t-\ell}\right)Y_{t-\ell-1}$$
$$+ j.\pi\int_0^1 \sin\left(j.\pi\left(x-0.5/T\right)\right)\left(\frac{1}{T}\sum_{t=1}^{[xT]}\left(\triangle Y_{t-\ell}\right)Y_{t-1}\right)dx$$
$$\overset{d}{\to} \cos\left(j.\pi\right)\left(C(1)\left(\int_0^1 \left(dW\right)W'\right)C(1)'+M_0\right)$$
$$+ j.\pi\int_0^1 \sin\left(j.\pi x\right)\left(C(1)\left(\int_0^x \left(dW\right)W'\right)C(1)'+xM_0\right)dx$$
$$= C(1)\left(\int_0^1 \left(dW\right)\cos\left(j.\pi.x\right)W'\right)C(1)'+\int_0^1 \cos\left(j.\pi.x\right)dxM_0$$
$$= C(1)\left(\int_0^1 \left(dW\right)\cos\left(j.\pi.x\right)W'\right)C(1)'$$

Moreover, by stationarity,

$$\frac{1}{T}\sum_{t=1}^{T}\cos\left(j.\pi\left(t-0.5\right)/T\right)\left(\triangle Y_{t-\ell}\right)\left(Y_{t-1}-Y_{t-1-\ell}\right)'$$

converges in probability to a matrix $M_{j,\ell}$. The result (30) follows now easily.

Finally, it follows from Lemma A.2 that

$$\frac{1}{T^2}\sum_{t=1}^{T}P_{i,T}\left(t\right)P_{j,T}\left(t\right)Y_{t-1}Y_{t-1}' \tag{71}$$
$$= 2\frac{1}{T^2}\sum_{t=1}^{T}\cos\left(i.\pi\left(t-0.5\right)/T\right)\cos\left(j.\pi\left(t-0.5\right)/T\right)Y_{t-1}Y_{t-1}'$$

$$
\begin{aligned}
&= 2\cos\left(i.\pi\left(1-0.5/T\right)\right)\cos\left(j.\pi\left(1-0.5/T\right)\right)\frac{1}{T^2}\sum_{t=1}^{T}Y_{t-1}Y_{t-1}'\\
&- 2\int_0^1 \frac{d}{dx}\left(\cos\left(i.\pi\left(x-0.5/T\right)\right)\cos\left(j.\pi\left(x-0.5/T\right)\right)\right)\\
&\qquad\times\left(\frac{1}{T^2}\sum_{t=1}^{[xT]}Y_{t-1}Y_{t-1}'\right)dx
\end{aligned}
$$

As is well known, under Assumption 1,

$$
\frac{1}{T^2}\sum_{t=1}^{[xT]}Y_{t-1}Y_{t-1}' \Rightarrow C(1)\int_0^x W(y)W(y)'dyC(1)'
$$

hence by the continuous mapping theorem,

$$
\begin{aligned}
&\frac{1}{T^2}\sum_{t=1}^{T}P_{i,T}\left(t\right)Y_{t-1}P_{j,T}\left(t\right)Y_{t-1}'\\
&\xrightarrow{d} 2\cos\left(i.\pi\right)\cos\left(j.\pi\right)C(1)\int_0^1 W(x)W(x)'dxC(1)'\\
&- 2C(1)\int_0^1\frac{d}{dx}\left(\cos\left(i.\pi x\right)\cos\left(j.\pi x\right)\right)\int_0^x W(y)W(y)'dyC(1)'\\
&= 2C(1)\int_0^1\cos\left(i.\pi x\right)W(x)\cos\left(j.\pi x\right)W(x)'dxC(1)'
\end{aligned}
$$

where again the equality follows via integration by parts. This proves (31).

8.3 The Stochastic Integral $\int_0^1\cos\left(\ell\pi x\right)W(x)dW'(x)$

The matrix $\int_0^1\cos\left(\ell\pi x\right)W(x)dW'(x)$, $\ell = 1, 2, ...$, is a $k \times k$ matrix of random variables whose (i,j)-th element is the scalar integral $\int_0^1\cos\left(\ell\pi x\right)W_i(x)dW_j(x)$. We first claim that for an arbitrary entry (i,j),

$$
C(x,\omega) = \cos\left(\ell\pi x\right)W(x,\omega) \in \mathcal{V}^{k\times 1}(0,s) \text{ for } s = 1,
$$

where $\mathcal{V}^{k\times 1}(0,s)$ is the class of integrand functions V for which the Ito integral $\int_0^s VdW'$ is defined: $(x,\omega) \to C(x,\omega)$ is $\mathcal{B}\times\mathcal{F}$ measurable, $C(x,\omega)$ is

$\mathcal{F}_x$ adapted, and

$$E\left[\int_0^s C(x)^2 dx\right] = \frac{1}{4}s^2 + \frac{\sin(2\ell\pi s)}{4\ell\pi}s + \frac{[\cos(2\ell\pi s) - 1]}{2(2\ell\pi)^2} < \infty$$
$$\left(= \frac{1}{4}, \text{ if } s = 1, \ell \geq 1\right).$$

Therefore, because $C \in \mathcal{V}(0, S)$, the Ito stochastic integral of C from 0 to s is defined as

$$I[C](\omega) = \int_0^s C(x,\omega) dW(x,\omega) = \lim_{n\to\infty} \int_0^s \phi_n(x,\omega) dW(x,\omega),$$

with limit in $L^2(P)$, where $\{\phi_n\}$ is a sequence of simple functions such that

$$\lim_{n\to\infty} E\left[\int_0^s (C(x) - \phi_n(x))^2 dx\right] = 0.$$

This condition is satisfied by taking

$$\phi_n(x,\omega) = \sum_{j=0}^{n} \cos(\ell\pi s_j) W(s_j,\omega).\mathbf{1}(s_j \leq x < s_{j+1}) \tag{72}$$

and

$$0 = s_0 \leq s_1 \leq ... \leq s_{n-1} \leq s_n = s.$$

For the chosen $\{\phi_n\}$ in (72),

$$I[C](\omega) = \lim_{s_{j+1}-s_j\to 0} \sum_{j=0}^{n} \cos(\ell\pi s_j) W(s_j,\omega)(W(s_{j+1},\omega) - W(s_j,\omega)).$$

Moreover, by the one-dimensional Ito formula (see Oksendal 2003, page 44)[8] we have

$$I[C](\omega) = \frac{\cos(\ell\pi s)}{2} W^2(s) - \frac{\sin(\ell\pi s)}{2\ell\pi} + \frac{\ell\pi}{2}\int_0^s \sin(\ell\pi x) W^2(x) dx$$

where $\int_0^s \sin(\ell\pi x) W^2(x)dx$ is a random variable with expectation

$$E\left(\int_0^s \sin(\ell\pi x) W^2(x) dx\right) = \frac{\sin(\ell\pi s)}{(\ell\pi)^2} - \frac{\cos(\ell\pi s)}{\ell\pi}s.$$

[8] In Oksendal's (2003) notation, $X_t \equiv B_t$; $g(t,x) = \cos(\ell\pi t)x^2/2$; $Y_t = \cos(\ell\pi t)B_t^2/2$. The result follows after some manipulations.

Therefore,

$$E\left(\int_0^1 \cos(\ell\pi x)\, W(x,\omega) dW(x,\omega)\right) = 0.$$

The *quadratic variation process* of

$$C(t) = \int_0^t \cos(\ell\pi x)\, W(x)\, dW'(x),$$

a $k \times k$ matrix-valued martingale in continuous time with respect to $\mathcal{F}_t^{(k)}$, is now

$$\begin{aligned} \langle C \rangle (t) &= \int_0^t \operatorname{Var}\left[\cos(\ell\pi x)\, W(x)\, d^{'} W(x) \Big| \mathcal{F}_x \right] \\ &= \int_0^t \cos^2(\ell\pi x)\, W(x)\, W(x)^{'}\, dx \otimes I_k. \end{aligned}$$

8.4 Proof of Lemmas 3 and 4

8.4.1 Auxiliary Results

The proofs of Lemmas 3 and 4 share the following auxiliary results.

Lemma A.3. *Under Assumptions 1-2 the following probability limits exist*:

$$\begin{aligned} \Sigma_{\beta\beta} &= p \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \beta' Y_{t-1} Y_{t-1}' \beta, \\ \Sigma_{X\beta} &= p \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} X_t Y_{t-1}' \beta, \\ \Sigma_{XX} &= p \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} X_t X_t^{'} \end{aligned}$$

Moreover, under the additional Assumption 5, Σ_{XX} is nonsingular and the matrix

$$\Sigma_{\beta\beta}^{*} = \Sigma_{\beta\beta} - \Sigma_{X\beta}' \Sigma_{XX}^{-1} \Sigma_{X\beta}$$

is nonsingular. Furthermore, under Assumptions 1,2 and 5,

$$\begin{aligned}
\Sigma_{\beta\otimes I_{m+1},\beta\otimes I_{m+1}} &= p\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\left(\beta'\otimes I_{m+1}\right)Y_{t-1}^{(m)}Y_{t-1}^{(m)'}\left(\beta\otimes I_{m+1}\right) \\
&= \Sigma_{\beta\beta}\otimes I_{m+1}, \\
\Sigma_{\beta,\beta\otimes I_{m+1}} &= p\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\beta^{'}Y_{t-1}Y_{t-1}^{(m)'}\left(\beta\otimes I_{m+1}\right) \\
&= \left(\Sigma_{\beta\beta},O_{r,r.m}\right), \\
\Sigma_{X,\beta\otimes I_{m+1}} &= p\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}X_tY_{t-1}^{(m)'}\left(\beta\otimes I_{m+1}\right) \\
&= \left(\Sigma_{X\beta},O_{k(p-1),r.m}\right).
\end{aligned}$$

Consequently,

$$\Sigma_{\beta,\beta\otimes I_{m+1}}-\Sigma_{\beta X}\Sigma_{XX}^{-1}\Sigma_{X,\beta\otimes I_{m+1}}=\left(\Sigma_{\beta\beta}^{*},O_{r,r.m}\right)$$

and

$$\begin{aligned}
\Sigma_{\beta\otimes I_{m+1},\beta\otimes I_{m+1}}^{*} &= \Sigma_{\beta\otimes I_{m+1},\beta\otimes I_{m+1}}-\Sigma_{X,\beta\otimes I_{m+1}}^{\prime}\Sigma_{XX}^{-1}\Sigma_{\beta\otimes I_{m+1},X} \\
&= \begin{pmatrix} \Sigma_{\beta\beta}^{*} & O_{r,r.m} \\ O_{r.m,r} & \Sigma_{\beta\beta}\otimes I_m \end{pmatrix},
\end{aligned}$$

The latter is a nonsingular matrix.

Proof: The existence of the probability limits $\Sigma_{\beta\beta}$, $\Sigma_{X\beta}$ and Σ_{XX} follows straightforwardly from Assumptions 1-2, and the nonsingularity of Σ_{XX} follows straightforwardly from Assumption 5. Note that $\Sigma_{\beta\beta}^{*}$ is the variance matrix of the residual ς_t of the linear projection of $\beta^{'}Y_{t-1}$ on X_t: $\beta^{'}Y_{t-1}=\Pi_0X_t+\varsigma_t$, say. Therefore, if this variance matrix were singular then there exists a vector δ such that $\delta'\beta^{'}Y_{t-1}=\delta'\Pi_0X_t$ a.s. Assumption 5 excludes this.

As to the probability limit $\Sigma_{\beta\otimes I_{m+1},\beta\otimes I_{m+1}}$, observe that similar to (71),

$$\frac{1}{T}\sum_{t=1}^{T}P_{i,T}\left(t\right)P_{j,T}\left(t\right)\beta'Y_{t-1}Y_{t-1}'\beta$$

$$= 2\frac{1}{T}\sum_{t=1}^{T}\cos\left(i.\pi\left(t-0.5\right)/T\right)\cos\left(j.\pi\left(t-0.5\right)/T\right)\beta' Y_{t-1}Y_{t-1}'\beta$$

$$= 2\cos\left(i.\pi\left(1-0.5/T\right)\right)\cos\left(j.\pi\left(1-0.5/T\right)\right)\frac{1}{T}\sum_{t=1}^{T}\beta' Y_{t-1}Y_{t-1}'\beta$$

$$-2\int_0^1 \frac{d}{dx}\left(\cos\left(i.\pi\left(x-0.5/T\right)\right)\cos\left(j.\pi\left(x-0.5/T\right)\right)\right)$$

$$\times\left(\frac{1}{T}\sum_{t=1}^{[xT]}\beta' Y_{t-1}Y_{t-1}'\beta\right)dx$$

for $i, j \geq 1$. We have already established that $\frac{1}{T}\sum_{t=1}^{T}\beta' Y_{t-1}Y_{t-1}'\beta = \Sigma_{\beta\beta} + o_p(1)$. Moreover, it is not hard to verify that

$$p\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{[xT]}\beta' Y_{t-1}Y_{t-1}'\beta = x.\Sigma_{\beta\beta}$$

pointwise in $x \in [0,1]$. It follows therefore by bounded convergence and integration by parts that

$$p\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}P_{i,T}\left(t\right)P_{j,T}\left(t\right)\beta' Y_{t-1}Y_{t-1}'\beta$$

$$= 2\left(\cos\left(i.\pi\right)\cos\left(j.\pi\right)-\int_0^1 x\frac{d}{dx}\left(\cos\left(i.\pi x\right)\cos\left(j.\pi x\right)\right)dx\right)\Sigma_{\beta\beta}$$

$$= 2\int_0^1\left(\cos\left(i.\pi x\right)\cos\left(j.\pi x\right)\right)dx.\Sigma_{\beta\beta}$$

$$= \left(\int_0^1\cos\left(\left(i+j\right)\pi x\right)dx + \int_0^1\cos\left(\left(i-j\right)\pi x\right)dx\right)\Sigma_{\beta\beta}$$

$$= \left(\frac{\sin\left(\left(i+j\right)\pi\right)}{\left(i+j\right)\pi}+\frac{\sin\left(\left(i-j\right)\pi\right)}{\left(i-j\right)\pi}\right)\Sigma_{\beta\beta}$$

$$= \begin{cases}\Sigma_{\beta\beta} & \text{if } i=j,\\ O_{r,r} & \text{if } i\neq j.\end{cases}$$

Similarly, for $i = 0$ and $j \geq 1$,

$$p\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}P_{j,T}\left(t\right)P_{0,T}\left(t\right)\beta' Y_{t-1}Y_{t-1}'\beta \tag{73}$$

$$= \sqrt{2}\left(\cos(j.\pi) - \int_0^1 x\frac{d}{dx}(\cos(j.\pi x))\,dx\right)\Sigma_{\beta\beta}$$
$$= \sqrt{2}\int_0^1 \cos(j.\pi x)\,dx.\Sigma_{\beta\beta} = \sqrt{2}\frac{\sin(j\pi)}{j\pi}\Sigma_{\beta\beta} = O_{r,r.m}$$

Hence, $\Sigma_{\beta\otimes I_{m+1},\beta\otimes I_{m+1}} = \Sigma_{\beta\beta}\otimes I_{m+1}$. Moreover, note that $\Sigma_{\beta,\beta\otimes I_{m+1}}$ is the matrix formed by the first r rows of $\Sigma_{\beta\otimes I_{m+1},\beta\otimes I_{m+1}}$. Thus, $\Sigma_{\beta,\beta\otimes I_{m+1}} = (\Sigma_{\beta\beta}, O_{r,r.m})$. The result for $\Sigma_{X,\beta\otimes I_{m+1}}$ follows by replacing $P_{0,T}(t)\,\beta' Y_{t-1}$ in (73) by X_t.

Finally, since $\Sigma^*_{\beta\beta}$ is nonsingular, so is $\Sigma_{\beta\beta}$, and therefore $\Sigma^*_{\beta\otimes I_{m+1},\beta\otimes I_{m+1}}$ is nonsingular.

Lemma A.4. *Let $\alpha_\perp$ be an orthogonal complement of α. Then under Assumptions 1-5,*

$$S_{00,T}^{-1} = \begin{pmatrix} (\alpha_\perp'\Omega\alpha_\perp)^{-1/2}\alpha_\perp' \\ (\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1} \end{pmatrix}' \begin{pmatrix} I_{k-r} & O_{k-r,r} \\ O_{r,k-r} & \left((\alpha'\Omega^{-1}\alpha)^{-1} + \Sigma^*_{\beta\beta}\right)^{-1} \end{pmatrix}$$
$$\times \begin{pmatrix} (\alpha_\perp'\Omega\alpha_\perp)^{-1/2}\alpha_\perp' \\ (\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1} \end{pmatrix} + o_p(1).$$

Proof: This is a standard result. See for example Johansen (1995).

Lemma A.5. *Let ξ be given by* (28). *Under Assumptions 1-5,*

$$\begin{aligned} N_T &= S_{00,T}^{-1} - S_{00,T}^{-1}S_{01,T}^{(m)}\xi\left(\xi' S_{10,T}^{(m)}S_{00,T}^{-1}S_{01,T}^{(m)}\xi\right)^{-1}\xi' S_{10,T}^{(m)}S_{00,T}^{-1} \\ &= S_{00,T}^{-1} - S_{00,T}^{-1}S_{01,T}^{(0)}\beta\left(\beta' S_{10,T}^{(0)}S_{00,T}^{-1}S_{01,T}^{(0)}\beta\right)^{-1}\beta' S_{10,T}^{(0)}S_{00,T}^{-1} \\ &= \alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp' + o_p(1) \end{aligned}$$

Proof: Johansen (1995, Lemma 10.1).

Lemma A.6. *There exists an orthogonal complement $\beta_\perp'$ of β such that*

$$\beta_\perp' C(1) = (\alpha_\perp'\Omega\alpha_\perp)^{-1/2}\alpha_\perp' C_0'.$$

Proof: This is a standard result. See Johansen (1995), or Lemma 2 in Bierens (2007).

Lemma A.7. *Let $\beta_\perp$ be the orthogonal complement of β defined in Lemma A.6. Let Assumptions 1-5 hold. Then*

$$\left(\alpha_\perp'\Omega\alpha_\perp\right)^{-1/2}\alpha_\perp' S_{01,T}^{(m)}\left(\beta_\perp'\otimes I_{m+1}\right)\overset{d}{\to}\int_0^1\left(dW_{k-r}\right)\widetilde{W}_{k-r,m}' \tag{74}$$

and

$$\sqrt{T}\left(\alpha_\perp'\Omega\alpha_\perp\right)^{-1/2}\alpha_\perp' S_{01,T}^{(m)}\left(\beta\otimes I_{m+1}\right)\overset{d}{\to} Z \tag{75}$$

jointly, where Z is a $(k-r)\times r(m+1)$ random matrix. In particular, the $k-r$ columns of Z' are independent

$$N_{r(m+1)}\left[0,\begin{pmatrix}\Sigma_{\beta\beta}^* & O_{r,r.m}\\ O_{r.m,r} & \Sigma_{\beta\beta}\otimes I_m\end{pmatrix}\right] \tag{76}$$

distributed. Moreover,

$$\left(\alpha'\Omega^{-1}\alpha\right)^{-1}\alpha'\Omega^{-1}S_{01,T}^{(m)}\left(\beta_\perp'\otimes I_{m+1}\right)\overset{d}{\to} M \tag{77}$$

where M is a $r\times(k-r)(m+1)$ random matrix, and

$$\left(\alpha'\Omega^{-1}\alpha\right)^{-1}\alpha'\Omega^{-1}S_{01,T}^{(m)}\left(\beta\otimes I_{m+1}\right)=\left(\Sigma_{\beta\beta},O_{r,r.m}\right)+o_p(1) \tag{78}$$

Proof: Substituting (59) in the expression for $S_{01,T}^{(m)}$ yields

$$\begin{aligned} S_{01,T}^{(m)} &= \alpha\frac{1}{T}\sum_{t=1}^{T}\beta' Y_{t-1}Y_{t-1}^{(m)\prime} \qquad (79)\\ &-\alpha\left(\frac{1}{T}\sum_{t=1}^{T}\beta' Y_{t-1}X_t'\right)\left(\frac{1}{T}\sum_{t=1}^{T}X_tX_t'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}X_tY_{t-1}^{(m)\prime}\right)\\ &+C_0\frac{1}{T}\sum_{t=1}^{T}U_tY_{t-1}^{(m)\prime}\\ &-C_0\left(\frac{1}{T}\sum_{t=1}^{T}U_tX_t'\right)\left(\frac{1}{T}\sum_{t=1}^{T}X_tX_t'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}X_tY_{t-1}^{(m)\prime}\right)\end{aligned}$$

hence

$$\left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1/2}\alpha_{\perp}'S_{01,T}^{(m)} = \left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1/2}\alpha_{\perp}'C_0\frac{1}{T}\sum_{t=1}^{T}U_tY_{t-1}^{(m)\prime} - \left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1/2}\alpha_{\perp}'C_0\left(\frac{1}{T}\sum_{t=1}^{T}U_tX_t'\right)\left(\frac{1}{T}\sum_{t=1}^{T}X_tX_t'\right)^{-1} \times\left(\frac{1}{T}\sum_{t=1}^{T}X_tY_{t-1}^{(m)\prime}\right). \tag{80}$$

Proof of (74). It follows now straightforwardly from (80) and Lemma A.1 that

$$\begin{aligned}\left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1/2}\alpha_{\perp}'S_{01,T}^{(m)} &= \alpha_{\perp}'C_0\frac{1}{T}\sum_{t=1}^{T}U_tY_{t-1}^{(m)\prime} + o_p(1)\\ &\xrightarrow{d}\left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1/2}\alpha_{\perp}'C_0\int_0^1(dW)\widetilde{W}_m'\left(C(1)\otimes I_{m+1}\right),\end{aligned}$$

hence

$$\begin{aligned}&\left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1/2}\alpha_{\perp}'S_{01,T}^{(m)}\left(\beta_{\perp}'\otimes I_{m+1}\right)\\ &\xrightarrow{d}\left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1/2}\alpha_{\perp}'C_0\int_0^1(dW)\widetilde{W}_m'\left(\left(C_0'\alpha_{\perp}\left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1/2}\right)\otimes I_{m+1}\right)\\ &=\int_0^1\left(dW_{k-r}\right)\widetilde{W}_{k-r,m}'\end{aligned}$$

where W_{k-r} and $\widetilde{W}_{k-r,m}$ are defined by (34) and (35), respectively.

Proof of (75). It follows from (80),

$$\begin{aligned}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}U_tY_{t-1}^{(m)\prime}\left(\beta\otimes I_{m+1}\right) &= O_p(1),\\ \frac{1}{\sqrt{T}}\sum_{t=1}^{T}U_tX_t' &= O_p(1)\end{aligned}$$

and Lemma A.3 that

$$\sqrt{T}\left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1/2}\alpha_{\perp}'S_{01,T}^{(m)}\left(\beta\otimes I_{m+1}\right) \tag{81}$$

$$= \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1/2} \alpha_\perp' C_0 \frac{1}{\sqrt{T}} \sum_{t=1}^{T} U_t Y_{t-1}^{(m)\prime} \left(\beta \otimes I_{m+1}\right)$$

$$- \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1/2} \alpha_\perp' C_0 \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} U_t X_t' \right) \left(\frac{1}{T} \sum_{t=1}^{T} X_t X_t' \right)^{-1}$$

$$\times \left(\frac{1}{T} \sum_{t=1}^{T} X_t Y_{t-1}^{(m)\prime} \left(\beta \otimes I_{m+1}\right) \right)$$

$$= \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1/2} \alpha_\perp' C_0 \frac{1}{\sqrt{T}} \sum_{t=1}^{T} U_t Y_{t-1}^{(m)\prime} \left(\beta \otimes I_{m+1}\right)$$

$$- \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1/2} \alpha_\perp' C_0 \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} U_t X_t' \right) \left(\Sigma_{XX}^{-1} \Sigma_{X\beta}, O_{k(p-1),r.m}\right)$$

$$+ o_p(1)$$

$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \vartheta_t \left(Y_{t-1}^{(m)\prime} \left(\beta \otimes I_{m+1}\right) - \left(X_t' \Sigma_{XX}^{-1} \Sigma_{X\beta}, 0_{r.m}' \right) \right) + o_p(1)$$

where $\vartheta_t = \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1/2} \alpha_\perp' C_0 U_t \sim$ i.i.d. $N_{k-r}\left[0, I_{k-r}\right]$. The result (75) follows now from McLeish's (1974) martingale difference central limit theorem. The joint convergence of (74) and (75) is obvious.

Proof of (76). Let $\vartheta_{i,t}$ be component i of ϑ_t. Then

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \vartheta_{it} \left(\left(\beta' \otimes I_{m+1}\right) Y_{t-1}^{(m)} - \begin{pmatrix} \Sigma_{X\beta}' \Sigma_{XX}^{-1} X_t \\ O_{r.m,1} \end{pmatrix} \right)$$

converges in distribution to column i of Z'. The result (76) then follows from McLeish's (1974) central limit theorem and Lemma A.3, and the independence of the columns of Z' follows from the independence of the components of ϑ_t.

Proof of (77) and (78). It follows from (79) and Lemmas A.4-A.3 that

$$\left(\alpha' \Omega^{-1} \alpha\right)^{-1} \alpha' \Omega^{-1} S_{01,T}^{(m)} = \frac{1}{T} \sum_{t=1}^{T} \beta' Y_{t-1} Y_{t-1}^{(m)\prime} \tag{82}$$

$$- \left(\frac{1}{T} \sum_{t=1}^{T} \beta' Y_{t-1} X_t' \right) \left(\frac{1}{T} \sum_{t=1}^{T} X_t X_t' \right)^{-1} \left(\frac{1}{T} \sum_{t=1}^{T} X_t Y_{t-1}^{(m)\prime} \right)$$

$$
\begin{aligned}
&+ \left(\alpha'\Omega^{-1}\alpha\right)^{-1}\alpha'\Omega^{-1}C_0\frac{1}{T}\sum_{t=1}^{T}U_tY_{t-1}^{(m)\prime} \\
&- \left(\alpha'\Omega^{-1}\alpha\right)^{-1}\alpha'\Omega^{-1}C_0\left(\frac{1}{T}\sum_{t=1}^{T}U_tX_t'\right)\left(\frac{1}{T}\sum_{t=1}^{T}X_tX_t'\right)^{-1} \\
&\quad \times\left(\frac{1}{T}\sum_{t=1}^{T}X_tY_{t-1}^{(m)\prime}\right) \\
&= \frac{1}{T}\sum_{t=1}^{T}\beta'Y_{t-1}Y_{t-1}^{(m)\prime} - \Sigma_{X\beta}'\Sigma_{XX}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}X_tY_{t-1}^{(m)\prime}\right) \\
&+ \left(\alpha'\Omega^{-1}\alpha\right)^{-1}\alpha'\Omega^{-1}C_0\frac{1}{T}\sum_{t=1}^{T}U_tY_{t-1}^{(m)\prime} + o_p(1)
\end{aligned}
$$

Now (77) follows straightforwardly from (82) and Lemma A.4, and (78) follows straightforwardly from (82) and Lemma A.3.

Lemma A.8. *Under Assumptions 1-5,*

$$
\begin{aligned}
&\left(\beta'\otimes I_{m+1}, T^{-1/2}\beta_\perp'\otimes I_{m+1}\right)S_{10,T}^{(m)}S_{00,T}^{-1}S_{01,T}^{(m)}\left(\beta\otimes I_{m+1}, T^{-1/2}\beta_\perp'\otimes I_{m+1}\right) \\
&\overset{d}{\to}\begin{pmatrix} \Sigma_{\beta\beta}^*\left(\left(\alpha'\Omega^{-1}\alpha\right)^{-1}+\Sigma_{\beta\beta}^*\right)^{-1}\Sigma_{\beta\beta}^* & O_{r,k-r+k.m} \\ O_{k-r+k.m,r} & O_{k-r+k.m,k-r+k.m}\end{pmatrix}.
\end{aligned}
$$

Proof: This result follows straightforwardly from Lemmas A.4 and A.7.

Lemma A.9. *Under Assumptions 1-5,*

$$
\begin{aligned}
&\left(T^{-1/2}\beta_\perp'\otimes I_{m+1}, \beta'\otimes I_{m+1}\right)S_{11,T}^{(m)}\left(T^{-1/2}\beta_\perp'\otimes I_{m+1}, \beta\otimes I_{m+1}\right) \\
&\overset{d}{\to}\begin{pmatrix} \int_0^1\widetilde{W}_{k-r,m}(x)\widetilde{W}_{k-r,m}'(x)dx & O_{(k-r)(m+1),r(m+1)} \\ O_{r(m+1),(k-r)(m+1)} & \begin{pmatrix}\Sigma_{\beta\beta}^* & O_{r,r.m} \\ O_{r.m,r} & \Sigma_{\beta\beta}\otimes I_m\end{pmatrix}\end{pmatrix}
\end{aligned}
$$

Proof: It follows from Lemmas A.1 and A.3 that

$$
S_{11,T}^{(m)} = \frac{1}{T}\sum_{t=1}^{T}Y_{t-1}^{(m)}Y_{t-1}^{(m)\prime} - \left(\frac{1}{T}\sum_{t=1}^{T}Y_{t-1}^{(m)}X_t'\right)\Sigma_{XX}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}X_tY_{t-1}^{(m)\prime}\right)
$$

$$
\begin{aligned}
& +o_p(1) \\
= \quad & \frac{1}{T}\sum_{t=1}^{T} Y_{t-1}^{(m)} Y_{t-1}^{(m)\prime} + O_p(1)
\end{aligned}
$$

Moreover, it follows from Lemma A.1, part (31), that

$$
\begin{aligned}
& \frac{1}{T}\left(\beta_\perp' \otimes I_{m+1}\right) S_{11,T}^{(m)} \left(\beta_\perp' \otimes I_{m+1}\right) \\
& \quad \xrightarrow{d} \left(\left(C_0'\alpha_\perp \left(\alpha_\perp'\Omega\alpha_\perp\right)^{-1/2}\right) \otimes I_{m+1}\right)' \int_0^1 \widetilde{W}_m(x)\widetilde{W}_m'(x)dx \\
& \qquad\qquad \times \left(\left(C_0'\alpha_\perp \left(\alpha_\perp'\Omega\alpha_\perp\right)^{-1/2}\right) \otimes I_{m+1}\right) \\
& \quad = \int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}_{k-r,m}'(x)dx
\end{aligned}
$$

Furthermore, it is not hard to verify from Lemma 2 that

$$
S_{11,T}^{(m)} \left(\beta \otimes I_{m+1}\right) = O_p(1)
$$

hence

$$
p\lim_{T\to\infty} \frac{1}{\sqrt{T}} \left(\beta_\perp' \otimes I_{m+1}\right) S_{11,T}^{(m)} \left(\beta \otimes I_{m+1}\right) = O_{k-r,r}
$$

Finally, it follows from Lemma A.3 that

$$
\begin{aligned}
& p\lim_{T\to\infty} \left(\beta' \otimes I_{m+1}\right) S_{11,T}^{(m)} \left(\beta \otimes I_{m+1}\right) \qquad (83) \\
& \quad = p\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \left(\beta' \otimes I_{m+1}\right) Y_{t-1}^{(m)} Y_{t-1}^{(m)\prime} \left(\beta \otimes I_{m+1}\right) \\
& \qquad - \left(p\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \left(\beta' \otimes I_{m+1}\right) Y_{t-1}^{(m)} X_t'\right) \Sigma_{XX}^{-1} \\
& \qquad\quad \times \left(p\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} X_t Y_{t-1}^{(m)\prime} \left(\beta \otimes I_{m+1}\right)\right) \\
& \quad = \Sigma_{\beta\otimes I_{m+1},\beta\otimes I_{m+1}} - \Sigma_{\beta\otimes I_{m+1}X}\Sigma_{XX}^{-1}\Sigma_{X,\beta\otimes I_{m+1}} \\
& \quad = \begin{pmatrix} \Sigma_{\beta\beta}^{*} & O_{r,r.m} \\ O_{r.m,r} & \Sigma_{\beta\beta}\otimes I_m \end{pmatrix}
\end{aligned}
$$

With these results at hand we are now able to prove our main results.

8.4.2 Proof of Lemma 3

Combining the results of Lemmas A.8 and A.9 it follows from Lemma 2 in Andersson et al. (1983) that

Lemma A.10. *Under Assumptions 1-5 the ordered solutions* $\lambda_{1,T} \geq \lambda_{2,T} \geq ... \geq \lambda_{(m+1)k,T}$ *of generalized eigenvalue problem*

$$\det\left(\lambda S_{11,T}^{(m)} - S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(m)}\right) = 0$$

converge in distribution to the ordered solutions $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{(m+1)k}$ *of*

$$\det\left[\lambda\begin{pmatrix} \begin{pmatrix} \Sigma_{\beta\beta}^{*} & O_{r,r.m} \\ O_{r.m,r} & \Sigma_{\beta\beta}\otimes I_m \end{pmatrix} & O_{r(m+1),(k-r)(m+1)} \\ O_{(k-r)(m-1),r(m+1)} & \int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}_{k-r,m}'(x)dx \end{pmatrix} - \begin{pmatrix} \Sigma_{\beta\beta}^{*}\left(\left(\alpha'\Omega^{-1}\alpha\right)^{-1} + \Sigma_{\beta\beta}^{*}\right)^{-1}\Sigma_{\beta\beta}^{*} & O_{r,k(m+1)-r} \\ O_{k(m+1)-r,r} & O_{k(m+1)-r,k(m+1)-r} \end{pmatrix}\right] = 0$$

Obviously, all but r solutions are zero, and the non-zero solutions are the solutions of eigenvalue problem

$$\det\left(\lambda\Sigma_{\beta\beta}^{*} - \Sigma_{\beta\beta}^{*}\left(\left(\alpha'\Omega^{-1}\alpha\right)^{-1} + \Sigma_{\beta\beta}^{*}\right)^{-1}\Sigma_{\beta\beta}^{*}\right) = 0$$

This is the same result as in the standard TI cointegration case!

8.4.3 Proof of Lemma 4

To derive the limiting distribution of $T\left(\lambda_{r+1,T}, \lambda_{r+2,T}, ..., \lambda_{k,T}\right)'$, we follow a similar procedure as in Johansen (1995, p.159). Let

$$\begin{aligned} S(\lambda) &= \lambda S_{11,T}^{(m)} - S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(m)} \\ \xi_{\perp,T} &= \left(\beta_\perp \otimes I_{m+1}, \begin{pmatrix} O_{k,m.r} \\ \sqrt{T}\left(\beta\Sigma_{\beta\beta}^{-1/2}\otimes I_m\right) \end{pmatrix}\right) \qquad (84) \\ \rho &= T.\lambda = O_p(1) \end{aligned}$$

The reason for the factor $\sqrt{T}$ in (84) is to prevent $T^{-1}\xi_{\perp,T}' S_{11,T}^{(m)} \xi_{\perp,T}$ from converging to a singular matrix, because otherwise we cannot apply Lemma

2 in Andersson et al. (1983), and the reason for the normalization of β by $\Sigma_{\beta\beta}^{-1/2}$ will become clear below. Then

$$\det\left(\begin{pmatrix}\xi' \\ \xi'_{\perp,T}\end{pmatrix} S(\lambda)(\xi,\xi_{\perp,T})\right) = \det\begin{pmatrix}\xi' S(\lambda)\xi & \xi' S(\lambda)\xi_{\perp,T} \\ \xi'_{\perp,T} S(\lambda)\xi & \xi'_{\perp,T} S(\lambda)\xi_{\perp,T}\end{pmatrix}$$
$$= \det\left(\xi' S(\lambda)\xi\right)\det\left(\xi'_{\perp,T}\left(S(\lambda) - S(\lambda)\xi\left(\xi' S(\lambda)\xi\right)^{-1}\xi' S(\lambda)\right)\xi_{\perp,T}\right)$$

where ξ is defined by (28).

It follows from Lemma A.9 that $\beta' S_{11,T}^{(0)}\beta = O_p(1)$ and $\xi'_{\perp,T} S_{11,T}^{(m)}\xi = O_p(1)$, whereas by assumption, $\rho = O_p(1)$. Therefore,

$$\begin{aligned}\xi' S(\lambda)\xi &= \lambda\xi' S_{11,T}^{(m)}\xi - \xi' S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(m)}\xi \\ &= \lambda\beta' S_{11,T}^{(0)}\beta - \beta' S_{10,T}^{(0)} S_{00,T}^{-1} S_{01,T}^{(0)}\beta \\ &= \rho\frac{1}{T}\beta' S_{11,T}^{(0)}\beta - \beta' S_{10,T}^{(0)} S_{00,T}^{-1} S_{01,T}^{(0)}\beta \\ &= -\beta' S_{10,T}^{(0)} S_{00,T}^{-1} S_{01,T}^{(0)}\beta + o_p(1)\end{aligned}$$

and

$$\begin{aligned}\xi'_{\perp,T} S(\lambda)\xi &= \lambda\xi'_{\perp,T} S_{11,T}^{(m)}\xi - \xi'_{\perp,T} S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(m)}\xi \\ &= \lambda\xi'_{\perp,T} S_{11,T}^{(m)}\xi - \xi'_{\perp,T} S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(0)}\beta \\ &= \rho\frac{1}{T}\xi'_{\perp,T} S_{11,T}^{(m)}\xi - \xi'_{\perp,T} S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(0)}\beta \\ &= -\xi'_{\perp,T} S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(0)}\beta + o_p(1)\end{aligned}$$

Combining these results it follows that

$$\begin{aligned}&\xi'_{\perp,T}\left(S(\lambda) - S(\lambda)\xi\left(\xi' S(\lambda)\xi\right)^{-1}\xi' S(\lambda)\right)\xi_{\perp,T} \\ &= \rho\frac{1}{T}\xi'_{\perp,T} S_{11,T}^{(m)}\xi_{\perp,T} - \xi'_{\perp,T} S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(m)}\xi_{\perp,T} \\ &- \xi'_{\perp,T} S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(0)}\beta\left(\beta' S_{10,T}^{(0)} S_{00,T}^{-1} S_{01,T}^{(0)}\beta\right)^{-1} \\ &\qquad\times\beta' S_{10,T}^{(0)} S_{00,T}^{-1} S_{01,T}^{(0)}\xi_{\perp,T} + o_p(1) \\ &= \rho\frac{1}{T}\xi'_{\perp,T} S_{11,T}^{(m)}\xi_{\perp,T} - \xi'_{\perp,T} S_{10,T}^{(m)} N_T S_{01,T}^{(m)}\xi_{\perp,T} + o_p(1)\end{aligned}$$

where N_T is defined in Lemma A.5. Since by Lemma A.7, $S_{01,T}^{(m)}\xi_{\perp,T} = O_p(1)$, it follows now from Lemmas A.5-A.7 that

$$\begin{aligned}
&\xi'_{\perp,T}\left(S(\lambda) - S(\lambda)\xi\left(\xi' S(\lambda)\xi\right)^{-1}\xi' S(\lambda)\right)\xi_{\perp,T} \\
&\quad = \rho\frac{1}{T}\xi'_{\perp,T}S_{11,T}^{(m)}\xi_{\perp,T} \\
&\quad - \xi'_{\perp,T}S_{10,T}^{(m)}\alpha_\perp\left(\alpha'_\perp\Omega\alpha_\perp\right)^{-1}\alpha'_\perp S_{01,T}^{(m)}\xi_{\perp,T} + o_p(1)
\end{aligned}$$

Next, observe from Lemma A.9 that

$$\begin{aligned}
&\frac{1}{T}\begin{pmatrix} I_{(m+1)(k-r)} & O_{(m+1)(k-r),r.m} \\ O_{r.m,(m+1)(k-r)} & \Sigma_{\beta\beta}^{1/2}\otimes I_m \end{pmatrix}\xi'_{\perp,T}S_{11,T}^{(m)}\xi_{\perp,T}\begin{pmatrix} I_{(m+1)(k-r)} & O_{(m+1)(k-r),r.m} \\ O_{r.m,(m+1)(k-r)} & \Sigma_{\beta\beta}^{1/2}\otimes I_m \end{pmatrix} \\
&\quad \xrightarrow{d} \begin{pmatrix} \int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}'_{k-r,m}(x)dx & O_{(k-r)(m+1),r.m} \\ O_{r.m,(k-r)(m+1)} & \overline{\Psi} \end{pmatrix}
\end{aligned}$$

where by Lemma A.9,

$$\begin{aligned}
\overline{\Psi} &= p\lim_{T\to\infty}\begin{pmatrix} O_{k,m.r} \\ \beta\otimes I_m \end{pmatrix}' S_{11,T}^{(m)}\begin{pmatrix} O_{k,m.r} \\ \beta\otimes I_m \end{pmatrix} \\
&= \begin{pmatrix} O_{r,r.m} \\ I_{r.m} \end{pmatrix}' p\lim_{T\to\infty}\left(\beta'\otimes I_{m+1}\right)S_{11,T}^{(m)}\left(\beta\otimes I_{m+1}\right)\begin{pmatrix} O_{r,r.m} \\ I_{r.m} \end{pmatrix} \\
&= \Sigma_{\beta\beta}\otimes I_m.
\end{aligned}$$

Hence

$$\frac{1}{T}\xi'_{\perp,T}S_{11,T}^{(m)}\xi_{\perp,T} \xrightarrow{d} \begin{pmatrix} \int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}'_{k-r,m}(x)dx & O_{(k-1)(m+1),r.m} \\ O_{r.m,(k-1)(m+1)} & I_{r.m} \end{pmatrix} \tag{85}$$

Moreover, it follows from Lemma A.7 that

$$\left(\alpha'_\perp\Omega\alpha_\perp\right)^{-1/2}\alpha'_\perp S_{01,T}^{(m)}\left(\beta_\perp\otimes I_{m+1}\right) \xrightarrow{d} \int_0^1 \left(dW_{k-r}\right)\widetilde{W}'_{k-r,m} \tag{86}$$

and

$$\begin{aligned}
&\left(\alpha'_\perp\Omega\alpha_\perp\right)^{-1/2}\alpha'_\perp S_{01,T}^{(m)}\begin{pmatrix} O_{k,m.r} \\ T^{1/2}\beta\otimes I_m \end{pmatrix} \\
&\quad = \left(\alpha'_\perp\Omega\alpha_\perp\right)^{-1/2}\alpha'_\perp S_{01,T}^{(m)}\left(T^{1/2}\beta\otimes I_{m+1}\right)\begin{pmatrix} O_{r,r.m} \\ I_{r.m} \end{pmatrix} \\
&\quad \xrightarrow{d} Z_2
\end{aligned} \tag{87}$$

where

$$Z_2 = Z \begin{pmatrix} O_{r,r.m} \\ I_{r.m} \end{pmatrix} \tag{88}$$

with Z defined in Lemma A.7. It follows now from (76) that the columns of Z_2' are independently $N_{r.m}\left[0, \Sigma_{\beta\beta} \otimes I_m\right]$ distributed. Thus, the columns of

$$V = \left(\Sigma_{\beta\beta}^{-1/2} \otimes I_m\right) Z_2' \tag{89}$$

are independent $N_{r.m}\left[0, I_{r.m}\right]$ distributed, and it follows from (86) and (87) that

$$\begin{aligned} &\xi_{\perp,T}' S_{10,T}^{(m)} \alpha_\perp \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1} \alpha_\perp' S_{01,T}^{(m)} \xi_{\perp,T} \\ &\stackrel{d}{\to} \begin{pmatrix} \int_0^1 \widetilde{W}_{k-r,m} dW_{k-r}' \\ V \end{pmatrix} \left(\int_0^1 \left(dW_{k-r}\right) \widetilde{W}_{k-r,m}', V'\right) \end{aligned} \tag{90}$$

Lemma 4 now follows from (85), (90), Lemma 2 in Anderson et al. (1983), and the next lemma:

Lemma A.11. *Under Assumptions 1-5, V is independent of W_{k-r} and $\widetilde{W}_{k-r,m}$.*

Proof: It follows from (81), (88) and (89) that

$$\begin{aligned} \widehat{V}_i &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \vartheta_{i,t} \left(\Sigma_{\beta\beta}^{-1/2} \otimes I_m\right) \left(O_{r.m,r}, I_{r.m}\right) \left(\beta' \otimes I_{m+1}\right) Y_{t-1}^{(m)} \\ \stackrel{d}{\to} V_i &\sim N_{r.m}\left[0, I_{r.m}\right], \end{aligned}$$

where V_i is column i of V, and $\vartheta_{i,t}$ is component i of $\vartheta_t = \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1/2} \alpha_\perp' C_0 U_t$. Moreover, note that

$$\widehat{W}_{i,k-r}(x) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[xT]} \vartheta_{it} \Rightarrow W_{i,k-r}(x)$$

where $W_{i,k-r}$ is component i of W_{k-r}. To prove that V_i and $W_{j,k-r}(x)$ are independent, consider the empirical process

$$\widetilde{V}_i(x) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[xT]} \vartheta_{i,t} \left(\Sigma_{\beta\beta}^{-1/2} \otimes I_m\right) \left(O_{r.m,r}, I_{r.m}\right) \left(\beta' \otimes I_{m+1}\right) Y_{t-1}^{(m)}$$

Clearly, $\widetilde{V}_i \Rightarrow \overline{V}_i$, where $\overline{V}_i(.)$ is a $r.m$ variate standard Wiener process, and $V_i = \overline{V}_i(1)$. It suffices to show that for all $x, y \in [0,1]$ and $i, j = 1, ..., k-r$, $E\left[\widetilde{V}_i(x)\widehat{W}_{j,k-r}(y)\right] \to 0$. This is trivial for $i \neq j$. For $i = j$,

$$
\begin{aligned}
E\left[\widetilde{V}_i(x)\widehat{W}_{i,k-r}(y)\right] &= \left(\Sigma_{\beta\beta}^{-1/2} \otimes I_m\right) \\
&\quad \times \frac{1}{T} \sum_{t=1}^{\min([xT],[yT])} E\left[(O_{r.m,r}, I_{r.m})\left(\beta' \otimes I_{m+1}\right) Y_{t-1}^{(m)}\right] \\
&= \left(\Sigma_{\beta\beta}^{-1/2} \otimes I_m\right) \\
&\quad \times \frac{1}{T} \sum_{t=1}^{\min([xT],[yT])} \left(P_{1,T}(t)\, E\left[Y_{t-1}'\beta\right], ..., P_{m,T}(t)\, E\left[Y_{t-1}'\beta\right]\right)' \\
&= 0
\end{aligned}
$$

because by Assumptions 1-3, $E\left[Y_{t-1}'\beta\right] = 0$. This proves the independence of V and W_{k-r}. The proof of the independence of V and $\widetilde{W}_{k-r,m}$ is similar.

8.5 Proof of Lemma 5

Let $\widehat{\xi} = \left(\widehat{\xi}_1, ..., \widehat{\xi}_r\right)$ be the ML estimator of ξ, where $\widehat{\xi}_i$, $i = 1, ..., r$, are the eigenvectors associated with the r largest eigenvalues $\widehat{\lambda}_{m,i}$,

$$S_{10,T}^{(m)} S_{00,T}^{-1} S_{01,T}^{(m)} \widehat{\xi}_i = \widehat{\lambda}_{m,i} S_{11,T}^{(m)} \widehat{\xi}_i, \ \ i = 1, .., r. \tag{91}$$

If we normalize $\widehat{\xi}$ as

$$\widetilde{\xi} = \widehat{\xi}\left(\xi'\widehat{\xi}\right)^{-1} \xi'\xi$$

then similar to Johansen (1988) we can write

$$\widetilde{\xi} - \xi = \xi_{\perp,T} U_{m,T} \tag{92}$$

where $\xi_{\perp,T}$ is defined by (84), and

$$U_{m,T} = \left(\xi_{\perp,T}'\xi_{\perp,T}\right)^{-1}\left(\xi_{\perp,T}'\widehat{\xi}\right)\left(\xi'\widehat{\xi}\right)^{-1}\left(\xi'\xi\right).$$

Similar to Johansen (1988) we can expand $T.U_T$ as

$$T.U_{m,T} = \left(T^{-1}\xi'_{\perp,T}\widehat{S}^{(m)}_{11,T}\xi_{\perp,T}\right)^{-1}\xi'_{\perp,T}\left(\widehat{S}^{(m)}_{10,T} - \widehat{S}^{(m)}_{11,T}\xi\alpha'\right) \times \Omega^{-1}\alpha\left(\alpha'\Omega^{-1}\alpha\right)^{-1} + o_p(1)$$

Moreover, similar to Johansen (1988) it can be shown that

$$\widehat{S}^{(m)}_{10,T} - \widehat{S}^{(m)}_{11,T}\xi\alpha' = \frac{1}{T}\sum_{t=1}^{T} Y^{(m)}_{t-1}U'_tC'_0 - \left(\frac{1}{T}\sum_{t=1}^{T} Y^{(m)}_{t-1}X'_t\right)\left(\frac{1}{T}\sum_{t=1}^{T} X_tX'_t\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} X_tU'_tC'_0\right)$$

Thus,

$$\left(\beta'_\perp \otimes I_{m+1}\right)\left(\widehat{S}^{(m)}_{10,T} - \widehat{S}^{(m)}_{11,T}\xi\alpha'\right) = \left(\frac{1}{T}\sum_{t=1}^{T}\left(\beta'_\perp \otimes I_{m+1}\right)Y^{(m)}_{t-1}U'_tC'_0\right) + o_p(1)$$

and by Lemma A.3,

$$\sqrt{T}\left(\beta' \otimes I_{m+1}\right)\left(\widehat{S}^{(m)}_{10,T} - \widehat{S}^{(m)}_{11,T}\xi\alpha'\right) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\beta' \otimes I_{m+1}\right)Y^{(m)}_{t-1}U'_tC'_0 - \begin{pmatrix} \Sigma'_{X\beta}\Sigma^{-1}_{XX} \\ O_{r.m,k(p-1)} \end{pmatrix}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_tU'_tC'_0 + o_p(1)$$

Hence,

$$\begin{aligned}
&\xi'_{\perp,T}\left(\widehat{S}^{(m)}_{10,T} - \widehat{S}^{(m)}_{11,T}\xi\alpha'\right) \\
&= \begin{pmatrix} \left(\beta'_\perp \otimes I_{m+1}\right)\left(\widehat{S}^{(m)}_{10,T} - \widehat{S}^{(m)}_{11,T}\xi\alpha'\right) \\ \left(O_{m.r,k}, \sqrt{T}\left(\Sigma^{-1/2}_{\beta\beta}\beta' \otimes I_m\right)\right)\left(\widehat{S}^{(m)}_{10,T} - \widehat{S}^{(m)}_{11,T}\xi\alpha'\right) \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{T}\sum_{t=1}^{T}\left(\beta'_\perp \otimes I_{m+1}\right)Y^{(m)}_{t-1}U'_tC'_0 \\ \left(O_{m.r,k}, I_{r.m}\right)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\Sigma^{-1/2}_{\beta\beta}\beta' \otimes I_{m+1}\right)Y^{(m)}_{t-1}U'_tC'_0 \end{pmatrix} + o_p(1) \\
&\quad + o_p(1)
\end{aligned}$$

Similar to Johansen (1988) it follows now that

$$\frac{1}{T}\sum_{t=1}^{T}(\beta_\perp' \otimes I_{m+1})\, Y_{t-1}^{(m)}U_t'C_0'\Omega^{-1}\alpha\left(\alpha'\Omega^{-1}\alpha\right)^{-1/2} \xrightarrow{d} \int_0^1 \widetilde{W}_{k-r,m}d\underline{W}_\alpha'$$

where

$$\underline{W}_\alpha = \left(\alpha'\Omega^{-1}\alpha\right)^{-1/2}\alpha'\Omega^{-1}C_0 W$$

is an r-variate standard Wiener process, which is independent of $\widetilde{W}_{k-r,m}$. Moreover, similar to parts (75) and (76) of Lemma A.7 it follows that

$$\left(O_{m.r,k}, I_{r.m}\right)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\Sigma_{\beta\beta}^{-1/2}\beta' \otimes I_{m+1}\right)Y_{t-1}^{(m)}U_t'C_0'\Omega^{-1}\alpha\left(\alpha'\Omega^{-1}\alpha\right)^{-1/2}$$
$$\xrightarrow{d} \underline{V}_\alpha$$

where $\underline{V}_\alpha$ is an $r.m \times r$ matrix with independent $N[0,1]$ distributed elements, which is also independent of $\widetilde{W}_{k-r,m}$. However, similar to Lemma A.11 it can be shown that $\underline{V}_\alpha$, $\underline{W}_\alpha$ and $\widetilde{W}_{k-r,m}$ are independent. Therefore, it follows from Lemma A.9 that (38) holds.

Denoting $\widetilde{\xi} = \left(\widetilde{\xi}_0', \widetilde{\xi}_m'\right)$, where $\widetilde{\xi}_0$ is a $k \times r$ matrix and $\widetilde{\xi}_m$ a $k.m \times r$ matrix, it follows now from (84), (92) and (38) that jointly,

$$T\left(\widehat{\xi}_0 - \beta\right) \xrightarrow{d} (\beta_\perp, O_{k,k.m})\left(\int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}_{k-r,m}'(x)dx\right)^{-1}$$
$$\times \int_0^1 \widetilde{W}_{k-r,m}d\underline{W}_\alpha'\left(\alpha'\Omega^{-1}\alpha\right)^{-1/2}$$
$$\sqrt{T}\widetilde{\xi}_m \xrightarrow{d} \left(\beta\Sigma_{\beta\beta}^{-1/2} \otimes I_m\right)\underline{V}_\alpha\left(\alpha'\Omega^{-1}\alpha\right)^{-1/2}.$$

8.6 Proof of Theorem 1

Consider the likelihood-ratio statistic $f_{m,T}\left(\widetilde{\xi}\right) - f_{0,T}\left(\widetilde{\beta}\right)$, where $\widetilde{\xi} = \widehat{\xi}\left(\xi'\widehat{\xi}\right)^{-1}\xi'\xi$, $\widetilde{\beta} = \widehat{\beta}\left(\beta'\widehat{\beta}\right)^{-1}\beta'\beta$ and

$$f_{0,T}(\beta) = T.\ln\left(\frac{\det\left(\beta'\left(S_{11,T}^{(0)} - S_{10,T}^{(0)}S_{00,T}^{-1}S_{01,T}^{(0)}\right)\beta\right)}{\det\left(\beta' S_{11,T}^{(0)}\beta\right)}\right),$$

$$f_{m,T}(\xi) = T.\ln\left(\frac{\det\left(\xi'\left(S_{11,T}^{(m)} - S_{10,T}^{(m)}S_{00,T}^{-1}S_{01,T}^{(m)}\right)\xi\right)}{\det\left(\xi' S_{11,T}^{(m)}\xi\right)}\right),$$

Recall from Lemma 5 that $\widetilde{\xi} = \xi + \xi_{\perp,T}U_{m,T}$, where $U_{m,T} = O_p(T^{-1})$. It follows from Lemma 7 in Johansen (1988), page 249[9] that under the null hypothesis $\xi = (\beta', O_{r,k.m})'$,

$$\begin{aligned}
f_{m,T}\left(\widetilde{\xi}\right) &= f_{m,T}(\xi + \xi_{\perp,T}U_{m,T}) \\
&= f_{0,T}(\beta) + T.\text{trace}\left\{\left(\beta'\left(S_{11,T}^{(0)} - S_{10,T}^{(0)}S_{00,T}^{-1}S_{01,T}^{(0)}\right)\beta\right)^{-1}\right. \\
&\times\left(U'_{m,T}\xi'_{\perp,T}\left(S_{11,T}^{(m)} - S_{10,T}^{(m)}S_{00,T}^{-1}S_{01,T}^{(m)}\right)\xi_{\perp,T}U_T\right. \\
&-U'_{m,T}\xi'_{\perp,T}\left(S_{11,T}^{(m)} - S_{10,T}^{(m)}S_{00,T}^{-1}S_{01,T}^{(m)}\right)\xi \\
&\times\left(\beta'\left(S_{11,T}^{(0)} - S_{10,T}^{(0)}S_{00,T}^{-1}S_{01,T}^{(0)}\right)\beta\right)^{-1} \\
&\left.\left.\times\xi'\left(S_{11,T}^{(m)} - S_{10,T}^{(m)}S_{00,T}^{-1}S_{01,T}^{(m)}\right)\xi_{\perp,T}U_{m,T}\right)\right\} \\
&-T.\text{trace}\left\{\left(\beta' S_{11,T}^{(0)}\beta\right)^{-1}\left(U'_T\xi'_{\perp,T}S_{11,T}^{(m)}\xi_{\perp,T}U_{m,T}\right.\right. \\
&\left.\left.-U'_{m,T}\xi'_{\perp,T}S_{11,T}^{(m)}\xi\left(\beta' S_{11,T}^{(0)}\beta\right)^{-1}\xi' S_{11,T}^{(m)}\xi_{\perp,T}U_{m,T}\right\}\right. \\
&+O\left(T.\left\|\xi_{\perp,T}U_{m,T}\right\|^3\right)
\end{aligned}$$

where for a matrix,$\|.\|$ denotes the maximum absolute value of its elements. Since

$$\begin{aligned}
&U_{m,T} = O_p\left(T^{-1}\right),\ \xi_{\perp,T}U_{m,T} = O_p\left(T^{-1/2}\right), \\
&\xi'_{\perp,T}\left(S_{11,T}^{(m)} - S_{10,T}^{(m)}S_{00,T}^{-1}S_{01,T}^{(m)}\right)\xi = O_p(1) \\
&\xi'_{\perp,T}S_{11,T}^{(m)}\xi = O_p(1)
\end{aligned}$$

and by Johansen (1995, Lemma 10.1),

$$\left(\beta' S_{11,T}^{(0)}\beta\right)^{-1} - \left(\beta'\left(S_{11,T}^{(0)} - S_{10,T}^{(0)}S_{00,T}^{-1}S_{01,T}^{(0)}\right)\beta\right)^{-1} = -\alpha'\Omega^{-1}\alpha + o_p(1)$$

[9] See also Johansen (1995), equation A.11 on page 224.

it follows now from (85) and Lemma 5 that

$$
\begin{aligned}
& f_{m,T}\left(\widetilde{\xi}\right) - f_{0,T}\left(\beta\right) \\
& = \operatorname{trace}\left[\left(\alpha'\Omega^{-1}\alpha\right)\left(T.U'_{m,T}\right)\xi'_{\perp,T}\left(\frac{1}{T}S^{(m)}_{11,T}\right)\xi_{\perp,T}\left(T.U_{m,T}\right)\right] + o_p(1) \\
& \xrightarrow{d} \operatorname{trace}\left(\underline{V}'_{\alpha}\underline{V}_{\alpha}\right) \\
& + \operatorname{trace}\left[\left(\int_0^1 d\underline{W}_{\alpha}\widetilde{W}'_{k-r,m}\right)\left(\int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}'_{k-r,m}(x)dx\right)^{-1}\right. \\
& \qquad \left.\times\left(\int_0^1 \widetilde{W}_{k-r,m}d\underline{W}'_{\alpha}\right)\right]
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
& f_{0,T}\left(\widetilde{\beta}\right) - f_{0,T}\left(\beta\right) \\
& \xrightarrow{d} \operatorname{trace}\left[\left(\int_0^1 d\underline{W}_{\alpha}W'_{k-r}\right)\left(\int_0^1 W_{k-r}(x)W'_{k-r}(x)dx\right)^{-1}\left(\int_0^1 W_{k-r}d\underline{W}'_{\alpha}\right)\right]
\end{aligned}
$$

Johansen (1995, page 192) has shown that, with $V_{\alpha} = \left(\alpha'\Omega^{-1}\alpha\right)^{-1}\alpha'\Omega^{-1/2}W$,

$$
\begin{aligned}
& \operatorname{trace}\left[\left(\alpha'\Omega^{-1}\alpha\right)\left(\int_0^1 dV_{\alpha}W'_{k-r}\right)\left(\int_0^1 W_{k-r}(x)W'_{k-r}(x)dx\right)^{-1}\right. \\
& \qquad \left.\times\left(\int_0^1 W_{k-r}dV'_{\alpha}\right)\right] \sim \chi^2_{r(k-r)}
\end{aligned}
$$

In our notation, $\underline{W}_{\alpha} = \left(\alpha'\Omega^{-1}\alpha\right)^{-1/2}\alpha'\Omega^{-1}C_0W$ is a r-variate standard Wiener process, which is distributed as $\left(\alpha'\Omega^{-1}\alpha\right)^{1/2}V_{\alpha}$ with V_{α} as in Johansen(1995). Thus,

$$
\begin{aligned}
& \operatorname{trace}\left[\left(\int_0^1 d\underline{W}_{\alpha}W'_{k-r}\right)\left(\int_0^1 W_{k-r}(x)W'_{k-r}(x)dx\right)^{-1}\right. \\
& \qquad \left.\times\left(\int_0^1 W_{k-r}d\underline{W}'_{\alpha}\right)\right] \sim \chi^2_{r(k-r)}.
\end{aligned} \tag{93}
$$

Similarly, it follows that

$$
\operatorname{trace}\left[\left(\int_0^1 d\underline{W}_{\alpha}\widetilde{W}'_{k-r,m}\right)\left(\int_0^1 \widetilde{W}_{k-r,m}(x)\widetilde{W}'_{k-r,m}(x)dx\right)^{-1}\right. \tag{94}
$$

$$\times \left(\int_0^1 \widetilde{W}_{k-r,m} d\underline{W}'_\alpha \right) \bigg] \sim \chi^2_{r(m+1)(k-r)}$$

because $\underline{W}_\alpha$ and $\widetilde{W}_{k-r,m}$ are independent. Then the difference of (94) and (93) is $\chi^2_{r.m.(k-r)}$ distributed, which follows from the following easy result:

If $Z = \begin{pmatrix} Y \\ X \end{pmatrix} \sim N_{p+q}[0, \Sigma]$, where $Y \in \mathbb{R}^p$, $X \in \mathbb{R}^q$ and

$\Sigma = \begin{pmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{pmatrix}$, then $Z'\Sigma^{-1}Z - X'\Sigma^{-1}_{XX}X \sim \chi^2_p$.

Moreover, since $\underline{V}_\alpha$ is *a* $r.m \times r$ matrix with independent $N[0,1]$ distributed elements, it follows that

$$\text{trace}\left(\underline{V}'_\alpha \underline{V}_\alpha\right) \sim \chi^2_{r.m.r}. \tag{95}$$

Furthermore, since $\underline{V}_\alpha$ and $\underline{W}_\alpha$ are independent, (95) is independent of (93) and (94), conditional on $\widetilde{W}_{k-r,m}$. Hence, the likelihood-ratio statistic $T\left(\widehat{f}_1\left(\widetilde{\xi}\right) - \widehat{f}_0(\beta)\right) - T\left(\widehat{f}_0\left(\widetilde{\beta}\right) - \widehat{f}_0(\beta)\right)$ converges in distribution to (95) plus (94) minus (93), resulting in a χ^2_{mkr} distribution.

8.7 Proof of Lemma 6

To prove (46), let $b_{2,i,j}(\tau)$ be element (i,j) of $B_2(t/T)$ with derivative $b'_{2,i,j}(\tau)$ and let $Z_{2,j,t-1}$ be component j of $Z_{2,t-1}$. Then by the mean value theorem,

$$\begin{aligned}
&\Delta\left(b_{2,i,j}(t/T)Z_{2,j,t-1}\right) \\
&\quad = \left(b_{2,i,j}(t/T) - b_{2,i,j}((t-1)/T)\right) Z_{2,j,t-1} + b_{2,i,j}((t-1)/T)\Delta Z_{2,j,t-1} \\
&\quad = b'_{2,i,j}((t-\lambda_{t,i,j,T})/T)Z_{2,j,t-1}/T + b_{2,i,j}((t-1)/T)\Delta Z_{2,j,t-1}
\end{aligned}$$

for some $\lambda_{t,i,j,T} \in [0,1]$. Denote by $\Psi_{t,T}$ be the matrix with elements $\Psi_{i,j,t,T} = b'_{2,i,j}((t-\lambda_{t,i,j,T})/T)$. Then

$$\begin{aligned}
\Delta\left(B_2(t/T)Z_{2,t-1}\right) &= \Psi_{t,T}Z_{2,t-1}/T + B_2(t/T)\Delta Z_{2,t-1} \\
&= B_2(t/T)\Delta Z_{2,t-1} + O_p\left(1/\sqrt{T}\right) \qquad (96)
\end{aligned}$$

where the latter follows from the fact that $\Psi_{t,T}$ is uniformly bounded and that $Z_{2,t-1}/\sqrt{T} = O_p(1)$.

Next, observe from (39) and (96) that

$$\begin{aligned}\Delta Z_{1,t} &= \sum_{j=1}^{p} D_j \Delta Z_{1,t-j} + \Delta\left(B_2\left(t/T\right) Z_{2,t-1}\right) + \sum_{j=1}^{p-1} C_{12,j}\Delta^2 Z_{2,t-j} + \Delta U_{1,t} \\ &= \sum_{j=0}^{t-1} \Pi_j \Delta\left(B_2\left((t-j)/T\right) Z_{2,t-1-j}\right) + \sum_{j=t}^{\infty} \Pi_j B_2\left(0\right)\Delta Z_{2,t-1-j} \\ &\quad + \sum_{i=1}^{p-1} C_{12,i} \sum_{j=0}^{\infty} \Pi_j \Delta^2 Z_{2,t-i-j} + \sum_{j=0}^{\infty} \Pi_j \Delta U_{1,t-j} \\ &= \sum_{j=0}^{t-1} \Pi_j \left(B_2\left((t-j)/T\right) - B_2\left(0\right)\right)\Delta Z_{2,t-1-j} \\ &\quad + V_t + O_p\left(1/\sqrt{T}\right) \end{aligned} \tag{97}$$

where

$$V_t = \sum_{j=0}^{\infty} \Pi_j B_2\left(0\right)\Delta Z_{2,t-1-j} + \sum_{i=1}^{p-1} C_{12,i} \sum_{j=0}^{\infty} \Pi_j \Delta^2 Z_{2,t-i-j} + \sum_{j=0}^{\infty} \Pi_j \Delta U_{1,t-j}$$

This proves (46).

Finally, it follows from (42) and (97) that

$$\begin{aligned}& B_1 Z_{1,t-1} + B_2\left(t/T\right) Z_{2,t-1} \\ &\quad = \Delta Z_{1,t} - \sum_{j=1}^{p-1} C_{11,j}\Delta Z_{1,t-j} - \sum_{j=1}^{p-1} C_{12,j}\Delta Z_{2,t-j} - U_{1,t} \\ &\quad = \sum_{j=0}^{t-1} \Pi_j \left(B_2\left((t-j)/T\right) - B_2\left(0\right)\right)\Delta Z_{2,t-1-j} \\ &\quad - \sum_{i=1}^{p-1} C_{11,i} \sum_{j=0}^{t-1-i} \Pi_j \left(B_2\left((t-j-i)/T\right) - B_2\left(0\right)\right)\Delta Z_{2,t-i-j-1} \\ &\quad + V_t - \sum_{i=1}^{p-1} C_{11,i} V_{t-i} - \sum_{j=1}^{p-1} C_{12,j}\Delta Z_{2,t-j} - U_{1,t} + O_p\left(1/\sqrt{T}\right) \\ &\quad = \sum_{j=0}^{t-1} Q_j \left(B_2\left((t-j)/T\right) - B_2\left(0\right)\right)\Delta Z_{2,t-1-j} + R_t + O_p\left(1/\sqrt{T}\right)\end{aligned}$$

say, where

$$R_t = V_t - \sum_{i=1}^{p-1} C_{11,i} V_{t-i} - \sum_{j=1}^{p-1} C_{12,j} \Delta Z_{2,t-j} - U_{1,t}.$$