Cointegration with structural breaks: from the single equation analysis to the multivariate approach with application to US money demand

Philippe Andrade
THEMA - Université Paris X Nanterre
CREST
Ufr Segmi, 200 av. de la République, 92001 Nanterre Cedex, France.
andrade@u-paris10.fr

Catherine Bruneau
THEMA - Université Paris X Nanterre
cbruneau@u-paris10.fr

January 2000

Abstract

In this paper we propose a multivariate analysis of a cointegrated vectorial autoregressive model with structural breaks affecting the cointegrating vectors.

These changes are previously recognized using a single-equation methodology. Asymptotic properties of the breaks dates estimators allow us to implement a full information maximum likelihood analysis with the breaks identified considered as fixed. Thus, a VECM is estimated, providing a relevant framework to test for non-causality and neutrality tests, as well as impulse response analysis of the dynamics. The methodology is applied to the trivariate system analysed by Gregory and Hansen (1996), with a money variable, an interest rate and the output for the US over the 1960-1990 period.

Keywords: Cointegration, structural breaks, full information maximum likelihood analysis,.causality, neutrality, impulses-responses.

JEL classification: C32.

1 Introduction

The econometric literature has witnessed recently an upsurge of interest in testing for structural breaks. In a non-stationary framework Gregory and Hansen (1996) propose a test of no-cointegration against the alternative of cointegration with a single structural break of unknown timing in the cointegrating vector. More recently, Bai and Perron (1998) consider issues related to multiple structural changes in the linear regression model estimated by least squares. We follow them to propose an estimation of structural breaks in cointegration vectors in a multivariate framework along the lines of Johansen (1988, 1991).

More precisely, our methodology is based on a two-steps approach. First we implement the estimation procedure of Gregory and Hansen (1996) to recognize the possible breaks of the system. Second, we use the results of Bai and Perron (1998) to show that the estimators of the breaks converge sufficiently rapidly so that we can consider them as known, when we implement the test procedure, as proposed by Andrade and Bruneau (1999) to estimate the cointegration rank of the system.. We refer to Bruneau (2000) to write the different equivalent multivariate representations of the dynamics. In particular, we use the VECM representation to implement non-causality tests in the Granger sense, following Toda and Phillips (1993). Moreover, we propose impulse-response analysis of the dynamics, along the lines of Bruneau and Nicolaï (1995), King and Watson (1993), and Lütkepohl and Reimers (1992), by focusing on the long-run dynamic multipliers which are involved in the characterization of neutrality or persistent causality properties.

The paper is structured as follows. In section 2 we give the assumptions of the analysis and we write the related multivariate model with the different equivalent parametrizations, VAR, VECM, RVAR, MA. In Section 3 we recall briefly the results obtained by Andrade and Bruneau (1999) concerning the rate of convergence of the estimators of the dates of breaks for cointegrated $I(1)$ variables. In Section 4 we present the likelihood ratio approach to estimate the cointegration rank with the necessary extensions of Johansen's procedure to estimate the vector error correcting model with the structural breaks previously recognized. In section 5 we apply the methology to the triavriate system analysed by Gregory and Hansen (1996), with a money stock variable, an interest rate and the output for the US over the 1960-1990 period. Section 6 concludes.

2 The data generating process and the representation of the dynamics

In this section we characterize the processes we want to study.

Let us consider a $p-$dimensional $I(1)$ process X_t whose components define r cointegrating vectors with possible structural breaks. In what follows, for sake of simplicity, we focus on the simplified case, where all cointegration relationships display a single structural break at the same date, which is supposed to be given and to occur at $t_0 = [\tau_0 T]$, with $0 < \tau_0 < 1$ the relative timing of the break and T the sample size. Notice that we are not interested here in investigating the case where the cointegration rank differ from one regime to another.

Thus, β_t is the $p \times r$ matrix of the cointegrating vectors. Accordingly, the related cointegrating vectors $z_t = \beta_t' X_t$ define r $I(0)$ processes of dimension p.

Let us suppose that the dynamics of the previous process is described by a VECM model of order k :

$$\Delta X_t = \sum_{i=1}^{k-1} \Gamma_{ti} \Delta X_{t-i} + \Pi_t X_{t-k} + \Phi_t + \varepsilon_t,$$

with $p \times p$ matrices Γ_{ti} supposed time dependent, as well as the $p \times r$ matrices $\Pi_t = \alpha_t \beta_t'$, where α_t and β_t are of full column rank r. More precisely, they evolve in the following way:

$$\begin{array}{rcl} \Gamma_{ti} & = & \Gamma_i^{(1)} 1_{t \leq t_0} + \Gamma_i^{(2)} 1_{t > t_0} \\ \Pi_t & = & \Pi^{(1)} 1_{t \leq t_0} + \Pi^{(2)} 1_{t > t_0} \\ \alpha_t & = & \alpha^{(1)} 1_{t \leq t_0} + \alpha^{(2)} 1_{t > t_0} \\ \beta_t & = & \beta^{(1)} 1_{t \leq t_0} + \beta^{(2)} 1_{t > t_0} \end{array}$$

The deterministic part Φ_t also shifts over time in the same way:

$$\Phi_t = \alpha^{(1)} \mu^{(1)} 1_{t \leq t_0} + \alpha^{(2)} \mu^{(2)} 1_{t > t_0}$$

where $\mu^{(1)}, i = 1, 2$ denote two $r-$dimensional column vectors. Indeed, we suppose that the system has a cointegrated "constant" term for each of the two regimes $t \leq t_0$ and $t > t_0$.

The innovation process $\varepsilon_t = X_t - EL(X_t/\underline{X_{t-1}})$ is supposed to be a gaussian white noise $N(0, \Omega)$, with a constant variance matrix Ω.

Now, we are looking for a VAR representation of the dynamics, equivalent to the previous ECM representation:

$$A_t(L)X_t = \Phi_t + \varepsilon_t$$

If such a VAR specification exists, it has to obey:

$$\begin{aligned}
A_t(L) &= \sum_{i=0}^{k} A_{t,i}L^i \\
\text{with } A_{t,0} &= I_d \\
A_{t,i} &= I_d - \sum_{j=1}^{i}\left[1_{t\leq t_0}\Gamma_j^{(1)} + 1_{t>t_0}\Gamma_j^{(2)}\right], i = 1, ..., k-1 \\
\text{and } A_{tk} &= 1_{t\leq t_0}\left(-\alpha^{(1)}\beta^{(1)\prime} + \Gamma_{k-1}^{(1)}\right) \\
&\quad +1_{t>t_0}\left(-\alpha^{(2)}\beta^{(2)\prime} + \Gamma_j^{(2)}\right)
\end{aligned}$$

A useful alternative representation of VAR in levels and ECM is the restricted VAR representation (RVAR) (see Campbell and Shiller, 1988, Mellander *et alii*, 1990 and Warne, 1993):

$$B_t(L)Y_t = M_t\Phi_t + \eta_t,$$

where:

$$Y_t = D_{\perp}(L)M_tX_t, \text{ with } M_t = \begin{bmatrix} S_{p-r,t} \\ \beta_t' \end{bmatrix}, \eta_t = M_t\varepsilon_t,$$

$$D(L) = \begin{bmatrix} I_{p-r} & 0 \\ 0 & (1-L)I_r \end{bmatrix}, D_{\perp}(L) = \begin{bmatrix} (1-L)I_{p-r} & 0 \\ 0 & I_r \end{bmatrix}.$$

The $(p-r, p)$ matrix $S_{p-r,t}$ is a selection matrix such that the $p \times p$ matrix $M_t = \begin{bmatrix} S_{p-r,t} \\ \beta_t' \end{bmatrix}$ is of full rank p.

It is easy to see that if $\{X_t\}$ is cointegrated of order $(1, 1)$, with r cointegration relationships, then the following relation holds between the parameters of the ECM and the RVAR models:

$$B_t(L) = M_t\left[\Gamma_t(L)M_t^{-1}D(L) - \alpha_t^*L^k\right]$$

where $\underset{(p,p)}{\alpha_t^*} = \left[\begin{array}{cc} 0_{(p,p-r)} & \alpha_t \end{array} \right]$, with the time dependent operator $B_t(L)$ such that:

$$\begin{aligned} B_t(L) &= 1_{t\leq t_0} M^{(1)} \left[\Gamma^{(1)}(L) \left(M^{(1)} \right)^{-1} D(L) - \alpha^{(1)*} L^k \right] \\ &\quad + 1_{t>t_0} M^{(2)} \left[\Gamma^{(2)}(L) \left(M^{(2)} \right)^{-1} D(L) - \alpha^{(2)*} L^k \right] \end{aligned}$$

with:

$$\begin{aligned} \Gamma^{(h)}(L) &= I_d - \sum_{i=1}^{k-1} \Gamma_i^{(h)} L^i \ , \ h = 1,2 \\ M^{(h)} &= \left[\begin{array}{c} S_{p-r,t} \\ \beta^{(h)\prime} \end{array} \right], \ h = 1,2 \end{aligned}$$

Given the RVAR representation, we also get a generalized MA representation of the dynamics with time dependent dynamic multipliers:

$$\Delta X_t = C_t(L) \left[\Phi_t + \varepsilon_t \right]$$

by simply inverting the $B_t(L)$ matrix of polynomials as:

$$C_t(L) = M_t^{-1} D(L) B_t(L)^{-1} M_t \tag{1}$$

which proves that the Wold decomposition is generalized as follows:

$$\begin{aligned} C_t(L) &= 1_{t\leq t_0} \left(M^{(1)} \right)^{-1} D(L) \left(B^{(1)}(L) \right)^{-1} M^{(1)} \\ &\quad + 1_{t>t_0} \left(M^{(2)} \right)^{-1} D(L) \left(B^{(2)}(L) \right)^{-1} M^{(2)} \end{aligned}$$

Notice that the deterministic part disappears, because $\Phi_t = \alpha^{(1)} \mu^{(1)} 1_{t\leq t_0} + \alpha^{(2)} \mu^{(2)} 1_{t>t_0}$, while $C_t(L)$can be decomposed into:

$$\begin{aligned} C_t(L) &= \left[1_{t\leq t_0} C^{(1)}(1) + 1_{t\leq t_0} C^{*(1)}(L)(1-L) \right] \\ &\quad + \left[1_{t>t_0} C^{(2)}(1) + 1_{t>t_0} C^{*(2)}(L)(1-L) \right] \end{aligned}$$

with $C^{(h)}(1)\alpha^{(h)} = 0$,and $(1-L)\alpha^{(h)} = 0$, for $h = 1,2$.

Then from (1), the long-run dynamic multipliers $C_t(1)$ shift at the date of the structural break t_0 as following:

$$C_t(1) = 1_{t\leq t_0} C^{(1)}(1) + 1_{t>_0} C^{(2)}(1)$$

with the property:

$$\beta^{(i)\prime} C^{(i)}(1) = 0 \text{ for } i = 1, 2$$

It is easy to express the the long-run dynamic multipliers by focusing on $B_t(1)$ and more precisely on $F_t(1) = M_t^{-1} B_t(1)$ which obeys:

$$\begin{aligned} F_t(1) &= \left[\Gamma_t(1) M_t^{-1} D(1) - \alpha_t^*\right] \\ &= 1_{t \le t_0} \left[\Gamma^{(1)}(1) \left(M^{(1)}\right)^{-1} D(1) - \alpha^{(1)*}\right] + 1_{t>_0} \left[\Gamma^{2)}(1) \left(M^{(2)}\right)^{-1} D(1) - \alpha^{(2)*}\right] \end{aligned}$$

so that, one can write:

$$\begin{aligned} C_t(1) &= M_t^{-1} D(1) F_t(1)^{-1} \\ &= 1_{t \le t_0} \left(M^{(1)}\right)^{-1} D(1) \left(F^{(1)}(1)\right)^{-1} + 1_{t>_0} \left(M^{(2)}\right)^{-1} D(1) \left(F^{(2)}(1)\right)^{-1} \end{aligned}$$

More generally, the common trend representation of the dynamics can be written as:

$$\begin{aligned} X_t &= 1_{t \le t_0} C^{(1)}(1) \sum_{s=1}^{t_0} \varepsilon_s + 1_{t>t_0} C^{(2)}(1) \sum_{s=t_0+1}^{t} \varepsilon_s \\ &\quad + 1_{t=t_0} \left[C^{(1)*}(L) - C^{(2)*}(L)\right] \varepsilon_t + 1_{t<t_0} C^{(1)*}(L) \varepsilon_t + 1_{t>t_0} C^{(2)*}(L) \varepsilon_t \end{aligned}$$

with the common trend characterized as the sum of two random walks:

$$T_t = 1_{t \le t_0} C^{(1)}(1) \sum_{s=1}^{t_0} \varepsilon_s + 1_{t>t_0} C^{(2)}(1) \sum_{s=t_0+1}^{t} \varepsilon_s$$

and the cyclical part defined as:

$$\begin{aligned} C_t^*(L) \varepsilon_t &= 1_{t=t_0} \left[C^{(2)*}(L) - C^{(1)*}(L)\right] \varepsilon_t \\ &\quad + 1_{t<t_0} C^{(1)*}(L) \varepsilon_t + 1_{t>t_0} C^{(2)*}(L) \varepsilon_t \end{aligned}$$

Here we write the model in the simplified case where all cointegration relationships display a single structural break at the same date. The general case is analysed in a companion paper Bruneau (2000). In what follows, we estimate the different parameters of interest for the different equivalent representations of the dynamics. First, we recall the results obtained in Andrade and Bruneau (1999), concerning the rate of convergence of the estimators of the parameters characterizing the structural break.

3 Single equation and estimation of the dates of shifts

Andrade and Bruneau (1999) proved that the estimate of the date of the structural break, identified along the lines of Gregory and Hansen (1996), converge at the rate $O(T^2)$. Accordingly, in the multivariate analysis of the cointegrated dynamics, we can suppose that the date of the shift is known.

Let us partition X_t into $(y_t \quad x_t)'$, with y_t a univariate process and $x_t = (1, x_{1t}, ..., x_{p-1,t})'$, a $p-$vector. We choose to normalize the first row of β_t' to one, so that we associate the cointegration relationship with the following regression:

$$\begin{aligned} y_t &= x_t'\delta_t + z_t \\ &= \delta_{0t} + \sum_{j=1}^{p-1} \delta_{jt} x_{jt} + z_t \end{aligned} \tag{2}$$

where $\delta_t = \delta^{(1)} 1_{t \leq t_0} + \delta^{(2)} 1_{t > t_0}$.

The multiple regression (2) also expresses in matrix form:

$$Y = \bar{X}\delta + Z$$

where $\bar{X} = diag(X_1, X_2)$ with $X_1 = (x_1, \ldots, x_{[\tau T]})'$, $X_2 = (x_{[\tau T]+1}, \ldots, x_T)'$ and $\delta = (\delta^{(1)\prime} \ \delta^{(2)\prime})'$ denoting the parameters of the cointegration relationship, with τ such that $0 < \tau < 1$. In the following we denote τ_0 the value of τ associated with the structural break $t_0 = [\tau_0 T]$, wich allows the cointegration property.

We recall the necessary assumptions given in Andrade and Bruneau (1998):

Assumption A1: Defining $T_0 = 1, T_1 = t_0$ and $T_2 = T$, we assume that, for each $i = 1, 2$, the matrix $X_i^{0\prime} X_i^0 / (T_i - T_{i-1})^2$, $i = 1, 2$, converges in probability to some nonrandom positive definite matrix, not necessarily the same for all i.

Assumption A2: There exists an $l_0 > 0$ such that for all $l > l_0$, the minimum eigenvalues of $A_{il} = l^{-2} \sum_{t=T_i+1}^{T_i+l} x_t x_t'$, for $i = 0, 1$ and the minimum eigenvalues of $A_{il}^* = l^{-2} \sum_{t=T_i-l}^{T_i} x_t x_t'$,for $i = 1, 2$, are bounded away from zero.

Assumption A3: The matrix $B_{kl} = \sum_{t=k}^{l} x_t x_t'$ is invertible for $l - k \geq p$ the dimension of x_t.

Assumption A1 is standard for multiple non-stationary regressions. Assumption A2 states that there are enough observations near the true break point so it can be identified. Assumption A3 is imposed because Gregory and Hansen procedure estimates the break point by referring to a global least squares criterium.

Gregory and Hansen test the null of no cointegration against the alternative of cointegration with one unknown structural break in the cointegrating vector as in (2). Identification of the breaks is achieved by minimizing the Phillips and Perron (1988) statistics of the z_t, where z_t is the residuals from an OLS estimation of equation (2). Therefore when the null is rejected, the procedure provides OLS estimates of the δ_j's $(j = 1, 2)$.

First, one can prove, the following lemma:

Lemma 1 *Under A1-A3,* $T^{-2} \sum_{t=1}^{T} y_t d_t' = o_p(1)$, *where* $d_t = (\hat{\delta}_k - \delta_j^0)' x_t$ *with* $\hat{\delta}_k$ *the OLS estimator of* $\delta_k, k = 1, 2$. *Under A1-A4,* $\hat{\tau}_{0p} \tau_0$ *implies that:*

$$\lim_{T \to \infty} \sup P \left(T^{-2} \sum_{t=1}^{T} d_t d_t' > C \left\| \delta_1^{(0)} - \delta_2^0 \right\|^2 \right) > \epsilon_0,$$

for some $p \times p$ *matrix* C *positive definite and some* $\epsilon_0 > 0$.

Next, we prove the proposition which establishes the consistency of the estimators proposed by Gregory and Hansen:

Proposition 2 *Under A1-A3, the OLS estimator of* τ_0, $\hat{\tau}_0$, *obeys:* $\hat{\tau}_0 \to_p \tau_0$.

Finally the rate of convergence of the estimate of the date of the break is characterized as follows:

Proposition 3 *Under A1-A3, for every* $\eta > 0$, *there exists a* $C < \infty$, *such that for a large* T, $P(T^2 \left| \hat{\tau}_0 - \tau_0 \right| > C) < \eta$.

The test procedure proposed by Gregory and Hansen is based on a single-equation analysis. Therefore, this may induce some bias in the estimation of the δ_j's in (2) as it may exist several cointegration vectors and non-exogeneity effects. Indeed, for a cointegrated rank $r > 1$, the long-term parameters estimated from a single-equation analysis will be in fact a nonlinear combination of the true long-term parameters of the system (Johansen, 1992). Nevertheless, in what follows, we use the results of the single equation analysis to implement the multivariate test procedure which identifies the cointegration rank of the trivariate dynamics analysed by Gregory and Hansen (1996). More precisely, we extend the Johansen procedure, by supposing that all possible cointegration relationships display a single break at the same date t_0 estimated by the single equation analysis of Gregory and Hansen. From proposition 2, the estimated date of break can be considered as known in the test procedure, because the convergence rate of the estimator $\hat{\tau}_0$ is T^2. We choose to exclude the cases where cointegration relationships between two of the three variables could be estimated with several structural breaks. Accordingly, we test for the cointegration rank with the single structural break at $[\hat{\tau}_0 T]$ but check whether the cointegration relationship estimated by Gregory and Hansen belongs to the cointegration space. Before empirically implementing this procedure, we recall the principles of the test to identify the cointegration rank of the system, as described in details in Andrade and Bruneau (1999).

4 FIML approach

To save space, we refer closely in this section to Andrade and Bruneau (1999) and Bruneau (2000), without giving any proof of the different propositions.

4.1 Statistical analysis

Considering the VECM representation of the previous section, let us introduce the following notations

$$\begin{aligned} Z_{0t} &= \Delta X_t, \\ Z_{1t} &= \begin{bmatrix} X_{t-k} \\ X_{t-k} \end{bmatrix}, \\ \Pi_{\tau_0} &= \begin{bmatrix} \Pi^{(1)} 1_{t \leq t_0} & \Pi^{(2)} 1_{t > t_0} \end{bmatrix}, \end{aligned}$$

$$
\begin{aligned}
Z_{2t} &= \begin{bmatrix} \vdots \\ \Delta X_{t-i} \\ \Delta X_{t-i} \\ \vdots \end{bmatrix}, \quad i = 1, \ldots, k-1 \\
\Psi_{\tau_0} &= \begin{bmatrix} \cdots \Gamma_i^{(1)} 1_{t>t_0} & \Gamma_i^{(2)} 1_{t>t_0} & \cdots \end{bmatrix}, \quad i = 1, \ldots, k-1
\end{aligned}
$$

so that Z_{0t} Z_{1t} and Π_{τ_0} are repectively of dimension $p \times p$, $2p \times p$ and $p \times 2p$, while Z_{2t} and Ψ_{τ_0} are respectively $2p(k-1) \times p$ and $p \times 2p(k-1)$ matrices. This model can be rewritten as:

$$
Z_{0t} = \Pi_{\tau_0} Z_{1t} + \Psi_{\tau_0} Z_{2t\tau} + \varepsilon_t, \quad t = 1, \ldots, T \tag{3}
$$

with

$$
\begin{aligned}
\Pi_{\tau_0} &= \alpha_{\tau_0} \beta'_{\tau_0}, \\
\alpha_{\tau_0} &= \begin{bmatrix} \alpha^{(1)} 1_{t \leq t_0} & \alpha^{(2)} 1_{t>t_0} \end{bmatrix}, \\
\beta_{\tau_0} &= \begin{bmatrix} \beta^{(1)} 1_{t \leq t_0} & 0 \\ 0 & \beta^{(2)} 1_{t>t_0} \end{bmatrix},
\end{aligned}
$$

where $\alpha^{(1)}$, $\alpha^{(2)}$, $\beta^{(1)}$ and $\beta^{(2)}$ are $p \times r$ matrices and $1_{t \leq t_0} = 1$ if $t \leq t_0$, and 0 otherwise.

When the date of shift, τ_0, is known, the log-likelihood function to maximize is:

$$
\begin{aligned}
&\log \mathcal{L}(\Psi_{\tau_0}, \alpha_{\tau_0}, \beta_{\tau_0}, \Omega) = -Tp \log \pi - \frac{1}{2} T \log |\Omega| \\
&-\frac{1}{2} \sum_{t=1}^{T} (Z_{0t} - \alpha_{\tau_0} \beta_{\tau_0} Z_{1t} + \Psi_{\tau_0} Z_{2t})' \Omega^{-1} (Z_{0t} - \alpha_{\tau_0} \beta_{\tau_0} Z_{1t} + \Psi_{\tau_0} Z_{2t}),
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
&\log \mathcal{L}(\alpha^{(1)}, \alpha^{(2)}, \beta^{(1)}, \beta_2, \Omega) = -Tp \log \pi - \frac{1}{2} T \log |\Omega| \\
&-\frac{1}{2} \sum_{t=1}^{T} (R_{0t} - \alpha_{\tau_0} \beta_{\tau_0} R_{1t})' \Omega^{-1} (R_{0t} - \alpha_{\tau_0} \beta_{\tau_0} R_{1t}),
\end{aligned}
$$

where R_{0t} and R_{1t} are the residuals obtained by regressing Z_{0t} and Z_{1t} on Z_{2t}.

Introducing the product moment matrices of the residuals:

$$\begin{aligned}
S_{00} &= T^{-1}\sum_{t=1}^{[\tau_0 T]} R_{0t}R_{0t}' + T^{-1}\sum_{t=[\tau_0 T]+1}^{T} R_{0t}R_{0t}' = S_{00}^{(1)} + S_{00}^{(2)},\\
S_{01\tau_0} &= T^{-1}\left[\begin{array}{cc} \sum_{t=1}^{[\tau_0 T]} R_{0t}R_{1t}' & \sum_{t=[\tau_0 T]+1}^{T} R_{0t}R_{1t}' \end{array}\right] = \left[\begin{array}{cc} S_{01}^{(1)} & S_{01}^{(2)} \end{array}\right],\\
S_{10\tau_0} &= \left[\begin{array}{c} S_{10}^{(1)} \\ S_{10}^{(2)} \end{array}\right],\\
S_{11\tau_0} &= T^{-1}\left[\begin{array}{cc} \sum_{t=1}^{[\tau_0 T]} R_{1t}R_{1t}' & 0 \\ 0 & \sum_{t=[\tau_0 T]+1}^{T} R_{1t}R_{1t}' \end{array}\right] = \left[\begin{array}{cc} S_{11}^{(1)} & 0 \\ 0 & S_{11}^{(2)} \end{array}\right],
\end{aligned}$$

one finally gets

$$\begin{aligned}
\mathcal{L}^{-2/T}(\alpha_{\tau_0}(\beta_{\tau_0}), \beta_{\tau_0}, \hat{\Omega}(\beta_{\tau_0})) &= (2\pi e)^{2p}\left|\hat{\Omega}(\beta_{\tau_0})\right|,\\
\text{where, } \left|\hat{\Omega}(\beta_{\tau_0})\right| &= \left|S_{00} - S_{01\tau_0}\beta_{\tau_0}(\beta_{\tau_0}S_{11\tau_0}\beta_{\tau_0})^{-1}\beta_{\tau_0}S_{10\tau_0}\right|,
\end{aligned}$$

which also expresses after concentrating all parameters except the β-parameters.

$$\begin{aligned}
&\mathcal{L}^{-2/T}(\beta^{(1)}, \beta^{(2)})\\
= &\left|S_{00}^{(1)} + S_{00}^{(2)} - S_{01}^{(1)}\beta^{(1)}(\beta^{(1)}S_{11}^{(1)}\beta^{(1)})^{-1}\beta^{(1)}S_{10}^{(1)} - S_{01}^{(2)}\beta^{(2)}(\beta^{(2)\prime}S_{11}^{(2)}\beta^{(2)})^{-1}\beta^{(2)\prime}S_{10}^{(2)}\right|.
\end{aligned}$$

Suppose we have a consistent estimator β_1^* of the $p\times r$ matrix $\beta^{(1)}$ (under the constraint $\left\{rank(\beta^{(1)}) \leq r\right\}$) and thus let us consider the test statistic defined as:

$$2Log\frac{Max_{\widetilde{\beta_2^*}/\beta_1^*}\mathcal{L}(\beta_1^*, \widetilde{\beta_2^*})}{Max_{\beta_2^*/\beta_1^*}\mathcal{L}(\beta_1^*, \beta_2^*)}$$

where $\widetilde{\beta_2^*}$ (resp.β_2^*) denotes a $p\times p$ (resp. $p\times r$) matrix and both maxima are computed for the given $p\times r$ matrix β_1^*.

This statistic has the following property:

Proposition 4 *Under the null* $\left\{rank(\beta^{(1)}) \leq r \text{ and } rank(\beta^{(2)}) \leq r\right\}$

$$2Log\frac{Max_{\widetilde{\beta_2^*}/\beta_1^*}\mathcal{L}(\beta_1^*, \widetilde{\beta_2^*})}{Max_{\beta_2^*/\beta_1^*}\mathcal{L}(\beta_1^*, \beta_2^*)} \underset{T\to\infty}{\simeq} -T\sum_{i=r+1}^{p} Log(1-\lambda_i^*(\beta^{(1)}))$$

for the $\lambda_1^*(\beta^{(1)}) \geq \lambda_2^*(\beta^{(1)}) \geq ... \geq \lambda_p^*(\beta^{(1)})$ *solutions of the equation:*

$$\left|\lambda S_{11}^{(2)} - S_{10}^{(2)} \left(S_{00}^{(2)}(\beta^{(1)})\right)^{-1} S_{01}^{(2)}\right| = 0.$$

with

$$S_{00}^{(2)}(\beta^{(1)}) = S_{00}^{(1)} + S_{00}^{(2)} - S_{01}^{(1)}\beta^{(1)}(\beta^{(1)}S_{11}^{(1)}\beta^{(1)})^{-1}\beta^{(1)}S_{10}^{(1)}$$

A sketch of the proof goes as follows. $\mathcal{L}^{-2/T}(\beta^{(1)}, \widetilde{\beta_2^*})$ is obtained by solving a minimization problem which is equivalent to the one considered by Johansen (1988):

$$Min_{\beta_2}\left|S_{00}^{(2)}(\beta_1^*) - S_{01}^{(2)}\beta^{(2)}(\beta^{(2)\prime}S_{11}^{(2)}\beta^{(2)})^{-1}\beta^{(2)\prime}S_{10}^{(2)}\right|$$

where $S_{00}^{(2)}(\beta_1^*) = S_{00}^{(1)} + S_{00}^{(2)} - S_{01}^{(1)}\beta_1^*(\beta_1^{*\prime}S_{11}^{(1)}\beta_1^*)^{-1}\beta_1^{*\prime}S_{10}^{(1)}$.

Thus, as β_1^* is a consistent estimator of $\beta^{(1)}$, the solution β_2^* of the previous minimization problem must tend to the solution of the minimization problem:

$$Min_{\beta_2}\left|S_{00}^{(2)}(\beta^{(1)}) - S_{01}^{(2)}\beta^{(2)}(\beta^{(2)\prime}S_{11}^{(2)}\beta^{(2)})^{-1}\beta^{(2)\prime}S_{10}^{(2)}\right|$$

So we are led to solve the following eigenvalue problem (see lemma A.8 of Johansen, 1995)

$$\left|\lambda S_{11}^{(2)} - S_{10}^{(2)} \left(S_{00}^{(2)}(\beta^{(1)})\right)^{-1} S_{01}^{(2)}\right| = 0.$$

By noting the n eigenvalues $\lambda_1^*(\beta^{(1)}) \geq \lambda_2^*(\beta^{(1)}) \geq ... \geq \lambda_p^*(\beta^{(1)})$, and the associated n eigenvectors $v_i, 1 \leq i \leq p$, which are normalized such that:

$$v_i' S_{11}^{(2)} v_j = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ otherwise} \end{cases}, \quad 1 \leq i, j \leq p.$$

one gets for the factor $\frac{\mathcal{L}^{-2/T}(\beta_1^*, \widetilde{\beta_2^*})}{\mathcal{L}^{-2/T}(\beta_1^*, \beta_2^*)}$ the expression:

$$\frac{\mathcal{L}^{-2/T}(\beta_1^*, \widetilde{\beta_2^*})}{\mathcal{L}^{-2/T}(\beta_1^*, \beta_2^*)} \underset{T\to\infty}{\simeq} \prod_{i=r+1}^{p}(1 - \lambda_i^*(\beta^{(1)})).$$

It is easy to prove that the following dual property holds, for a consistent estimator β_2^* of the $p \times r$ matrix β_2. Finally, the test statistic which is proposed to identify the cointegration rank of the system with the structural break at t_0 is the following one:

$$2Log\frac{Max_{\widetilde{\beta_2^*}/\beta_1^*}\mathcal{L}(\beta_1^*,\widetilde{\beta_2^*})}{Max_{\beta_2^*/\beta_1^*}\mathcal{L}(\beta_1^*,\beta_2^*)}+2Log\frac{Max_{\widetilde{\beta_1^*}/\beta_2^*}\mathcal{L}(\widetilde{\beta_1^*},\beta_2^*)}{Max_{\beta_1^*/\beta_2^*}\mathcal{L}(\beta_1^*,\beta_2^*)}$$

One can give an asymptotically equivalent distribution of this statistic, under the null hypothesis, according to the proposition:

Proposition 5 *Under the null* $\left\{rank(\beta^{(1)})\leq r \;\; and \; rank(\beta^{(2)})\leq r\right\}$

$$2Log\frac{Max_{\widetilde{\beta_2^*}/\beta_1^*}\mathcal{L}(\beta_1^*,\widetilde{\beta_2^*})}{Max_{\beta_2^*/\beta_1^*}\mathcal{L}(\beta_1^*,\beta_2^*)}+2Log\frac{Max_{\widetilde{\beta_1^*}/\beta_2^*}\mathcal{L}(\widetilde{\beta_1^*},\beta_2^*)}{Max_{\beta_1^*/\beta_2^*}\mathcal{L}(\beta_1^*,\beta_2^*)}$$
$$\underset{T\to\infty}{\simeq} -T\left[\sum_{i=r+1}^{p}Log(1-\lambda_i^*(\beta^{(1)}))+\sum_{i=r+1}^{p}Log(1-\mu_i^*(\beta^{(1)}))\right]$$

for the $\lambda_1^*(\beta^{(1)})\geq\lambda_2^*(\beta^{(1)})\geq...\geq\lambda_p^*(\beta^{(1)})$ *and* $\mu_1^*(\beta^{(2)})\geq\mu_2^*(\beta^{(2)})\geq ...\geq\mu_p^*(\beta^{(2)})$

solutions of the equations:

$$\left|\lambda S_{11}^{(2)}-S_{10}^{(2)}\left(S_{00}^{(2)}(\beta^{(1)})\right)^{-1}S_{01}^{(2)}\right|=0.$$

with

$$S_{00}^{(2)}(\beta^{(1)})=S_{00}^{(1)}+S_{00}^{(2)}-S_{01}^{(1)}\beta^{(1)}(\beta^{(1)}S_{11}^{(1)}\beta^{(1)})^{-1}\beta^{(1)}S_{10}^{(1)}$$

and

$$\left|\mu S_{11}^{(1)}-S_{10}^{(1)}\left(S_{00}^{(1)}(\beta_2)\right)^{-1}S_{01}^{(1)}\right|=0$$

with

$$S_{00}^{(1)}(\beta_2)=S_{00}^{(1)}+S_{00}^{(2)}-S_{01}^{(2)}\beta_2(\beta_2'S_{11}^{(2)}\beta_2)^{-1}\beta_2'S_{10}^{(2)}.$$

The next step is to show that the previous test statistic has an asymptotic distribution which is free of nuisance parameters under the null hypothesis.

4.1.1 Asymptotic properties

Starting from the generalized Wold decomposition of the first differentiated process:

$$\begin{aligned}\Delta X_t &= C_1(L)\varepsilon_t, \quad \text{for } t \leq [\tau T],\\ \Delta X_t &= C_2(L)\varepsilon_t, \quad \text{for } t > [\tau T],\end{aligned}$$

let us introduce the following notations which are similar to the ones which are used by Johansen (1995):

$$\begin{aligned}M_{ij}^{(1)} &= \frac{1}{T}\sum_{t=1}^{[\tau_0 T]} \Delta X_{t-i}\Delta X'_{t-j},\ M_{ij}^{(2)} = \frac{1}{T}\sum_{t=[\tau_0 T]+1}^{T} \Delta X_{t-i}\Delta X'_{t-j}\\ M_{1i}^{(1)} &= \frac{1}{T}\sum_{t=1}^{[\tau_0 T]} X_{t-k}\Delta X'_{t-i},\ M_{1i}^{(2)} = \frac{1}{T}\sum_{t=[\tau_0 T]+1}^{T} X_{t-k}\Delta X'_{t-i}\\ M_{11}^{(1)} &= \frac{1}{T}\sum_{t=1}^{[\tau_0 T]} X_{t-k}X'_{t-k},\ M_{11}^{(2)} = \frac{1}{T}\sum_{t=[\tau_0 T]+1}^{T} X_{t-k}X'_{t-k},\\ i,j &= 0,...k-1,\end{aligned}$$

and their theoretical counterparts:

$$\begin{aligned}\mu_{ij}^{(1)} &= E(\Delta X_{t-i}\Delta X'_{t-j}1_{t\leq t_0}),\ \mu_{ij}^{(2)} = E(\Delta X_{t-i}\Delta X'_{t-j}1_{t>t_0}),\\ \mu_{1i}^{(1)} &= \sum_{j=k-i}^{\infty} E(X_{t-k}\Delta X'_{t-j}1_{t\leq t_0}),\ \mu_{1i}^{(2)} = \sum_{j=k-i}^{\infty} E(X_{t-k}\Delta X'_{t-j}1_{t>t_0}),\\ i,j &= 0,...,k-1.\end{aligned}$$

Furthermore, letting $C_1 = C_1(1),\ C_2 = C_2(1)$ and $W(r)$ be a brownian motion of dimension p and covariance matrix Ω, we have the following lemma:

Lemma 6 *As* $T \longrightarrow \infty$,

$$T^{-1/2}X_{[Tr]} \xrightarrow{w} C_1 W(r), \quad 0 \leq r \leq 1,\ r \leq \tau_0, \tag{4}$$

$$T^{-1/2}X_{[Tr]} \xrightarrow{w} C_2 W(r), \quad 0 \leq r \leq 1,\ r > \tau_0, \tag{5}$$

$$M_{ij}^{(1)} \xrightarrow{a.s} \tau_0\mu_{ij}^{(1)}, \quad i,j = 0,...k-1, \tag{6}$$

$$M_{ij}^{(2)} \xrightarrow{a.s} (1-\tau_0)\mu_{ij}^{(2)}, \quad i,j = 0,...k-1, \tag{7}$$

$$M_{1i}^{(1)} \xrightarrow{w} C_1\int_0^{\tau_0} W dW' C_1' + \mu_{1i}^{(1)}, \quad i = 0,...,k-1, \tag{8}$$

$$M_{1i}^{(2)} \xrightarrow{w} C_2\int_{\tau_0}^{1} W dW' C_2' + \mu_{1i}^{(2)}, \quad i = 0,...,k-1, \tag{9}$$

$$T^{-1}M_{11}^{(1)} \xrightarrow{w} C_1 \int_0^{\tau_0} W(u)W'(u)duC_1', \quad (10)$$

$$T^{-1}M_{11}^{(2)} \xrightarrow{w} C_2 \int_{\tau_0}^{1} W(u)W'(u)duC_2', \quad (11)$$

$$\beta^{(1)}M_{11}^{(1)}\beta^{(1)} \xrightarrow{a.s.} \beta^{(1)}\mu_{pj}^{(1)}\beta^{(1)}, \quad (12)$$

$$\beta^{(2)\prime}M_{11}^{(2)}\beta^{(2)} \xrightarrow{a.s.} \beta^{(2)\prime}\mu_{pj}^{(2)}\beta^{(2)}. \quad (13)$$

Proof. (4) and (5) are application of the standard multivariate invariance principle. (8) and (9) are direct applications of equation A.4 of Gregory and Hansen (1996a). (10) and (11) hold from, respectively, (4) and (5) and the continuous mapping theorem, since $\int_0^{\tau_0} WW'$ and $\int_{\tau_0}^1 WW'$ are continuous with respect to τ_0. Finally, (6), (7), (12) and (13) come from the stationarity properties of the ΔX_t's and of $\beta^{(1)\prime}X_{t-k}.1_{t\leq t_0}$, (resp., $\beta^{(2)\prime}X_{t-k}.1_{t>t_0}$(since $\beta^{(i)\prime}C_i = 0$, $i = 1, 2$).

■

Now let us define the theoretical counterparts of the product moment of the residuals

$$Var\begin{bmatrix} Z_{0t} \\ \beta_{\tau_0}'Z_{1t} \end{bmatrix} |Z_{2t\tau_0}] = Var\begin{bmatrix} R_{0t} \\ \beta_{\tau_0}'R_{1t} \end{bmatrix} = \begin{bmatrix} \Sigma_{00} & \Sigma_{0\beta_{\tau_0}} \\ \Sigma_{\beta_{\tau_0}0} & \Sigma_{\beta_{\tau_0}\beta_{\tau_0}} \end{bmatrix}$$

where:

$$\begin{aligned}
\Sigma_{00} &= \tau_0\Sigma_{00}^{(1)} + (1-\tau_0)\Sigma_{00}^{(2)} \\
\Sigma_{00}^{(1)} &= \lim_{T\to\infty} [\tau_0 T]^{-1} \sum_{t=0}^{[\tau_0 T]} R_{0t}R_{0t}', \\
\Sigma_{00}^{(2)} &= \lim_{T\to\infty} [(1-\tau_0)T]^{-1} \sum_{t=[\tau_0 T]+1}^{T} R_{0t}R_{0t}', \\
\Sigma_{0\beta_{\tau_0}} &= \begin{bmatrix} \Sigma_{0\beta^{(1)}} & \Sigma_{0\beta^{(2)}} \end{bmatrix}, \\
\Sigma_{\beta_{\tau_0}0} &= \Sigma_{0\beta_{\tau_0}}', \\
\Sigma_{\beta_{\tau_0}\beta_{\tau_0}} &= \begin{bmatrix} \Sigma_{\beta^{(1)}\beta^{(1)}} & 0 \\ 0 & \Sigma_{\beta^{(2)}\beta^{(2)}} \end{bmatrix},
\end{aligned}$$

Let $B_T^{(1)}$ and $B_T^{(2)}$ be two $p \times (p-r)$ matrices such that:

$$B_T^{(i)} = \kappa_i(\kappa_i'\kappa_i)^{-1}, \; i = 1, 2$$

where $\kappa_1$and κ_2 are orthogonal to, respectively, $\beta^{(1)}$ and $\beta^{(2)}$. Let also G_1 and G_2 be two brownian motions such that

$$G_i = \kappa_i(\kappa_i'\kappa_i)^{-1}C_iW, \; i=1,2.$$

With this notations, we have the following lemma:

Lemma 7 *As $T \longrightarrow \infty$,we have*

$$S_{00}^{(1)} \xrightarrow{a.s} \tau_0\Sigma_{00}^{(1)}, \; S_{00}^{(2)} \xrightarrow{a.s} (1-\tau_0)\Sigma_{00}^{(2)}, \tag{14}$$

$$\beta^{(1)}S_{10}^{(1)} \xrightarrow{a.s} \tau_0\Sigma_{\beta^{(1)}0}, \; \beta_2'S_{10}^{(2)} \xrightarrow{a.s} (1-\tau_0)\Sigma_{\beta_2 0}, \tag{15}$$

$$\beta^{(1)}S_{11}^{(1)}\beta^{(1)} \xrightarrow{a.s} \tau_0\Sigma_{\beta^{(1)}\beta^{(1)}}, \; \beta_2'S_{11}^{(2)}\beta_2 \xrightarrow{a.s} (1-\tau_0)\Sigma_{\beta_2\beta_2}, \tag{16}$$

$$T^{-1}B_T^{(1)\prime}S_{11}^{(1)}B_T^{(1)} \xrightarrow{w} \int_0^{\tau_0} G_1(u)G_1'(u)du, \tag{17}$$

$$T^{-1}B_T^{(2)\prime}S_{11}^{(2)}B_T^{(2)} \xrightarrow{w} \int_{\tau_0}^{1} G_2(u)G_2'(u)du, \tag{18}$$

$$B_T^{(1)\prime}\left[S_{10}^{(1)} - S_{11}^{(1)}\alpha_1\beta^{(1)}\right] \xrightarrow{w} \int_0^{\tau_0} G_1 dW', \tag{19}$$

$$B_T^{(2)\prime}\left[S_{10}^{(2)} - S_{11}^{(2)}\alpha_2\beta_2'\right] \xrightarrow{w} \int_{\tau_0}^{1} G_2 dW', \tag{20}$$

$$B_T^{(1)\prime}S_{11}^{(1)}\beta^{(1)} \sim O_p(1), \; B_T^{(2)\prime}S_{11}^{(2)}\beta_2 \sim O_p(1), \tag{21}$$

$$\Sigma_{00}^{(i)} = \alpha_i\Sigma_{\beta_i 0} + \Omega, \quad i=1,2, \tag{22}$$

$$\Sigma_{0\beta_i} = \alpha_i\Sigma_{\beta_i\beta_i} \quad i=1,2, \tag{23}$$

$$\Sigma_{00}^{(i)} = \alpha_i\Sigma_{\beta_i\beta_i}\alpha_i' + \Omega \quad i=1,2, \tag{24}$$

Proof. (14) to (16) are immediate since ΔX_t, $\beta^{(1)}X_{t-k}.1_{t\leq t_0}$ and $\beta^{(2)\prime}X_{t-k}.1_{t>t_0}$ are stationary. (17) to (21) come from Lemma 2 and the continuous mapping theorem (see lemma 10.3 in Johansen (1995, pp.146-148). (22) to (24) follow directly from (3). ■

Lemma 8 *The asymptotic limit of $S_{00}^{(1)} + S_{00}^{(2)}$ when T goes to infinity is:*

$$\tau_0\alpha_1\Sigma_{\beta_i 0} + (1-\tau_0)\alpha_2\Sigma_{\beta_2 0} + \Omega, \tag{25}$$

Accordingly we have the following theorem:

Theorem 9 *Under the null hypothesis*

$$H_0 : \left\{rank(\Pi^{(1)}) \leq r\right\} \cup \left\{rank(\Pi^{(2)}) \leq r\right\},$$

with $\Pi^{(1)}$ *and* $\Pi^{(2)}$ *the matrices defined in (3) against the alternative that*

$$H_a : \left\{rank(\Pi^{(1)}) \leq p\right\} \cup \left\{rank(\Pi^{(2)}) \leq p\right\},$$

and for consistent constrained estimators β_1^* *and* β_2^**, one has the limit property:*

$$\begin{aligned}
&2Log\frac{Max_{\widetilde{\beta_1^*}/\beta_2^*}\mathcal{L}(\widetilde{\beta_1^*},\beta_2^*)}{Max_{\beta_1^*/\beta_2^*}\mathcal{L}(\beta_1^*,\beta_2^*)} + 2Log\frac{Max_{\widetilde{\beta_2^*}/\beta_1^*}\mathcal{L}(\beta_1^*,\widetilde{\beta_2^*})}{Max_{\beta_2^*/\beta_1^*}\mathcal{L}(\beta_1^*,\beta_2^*)} \underset{T\to\infty}{\to} \\
&tr\left\{\int_0^{\tau_0} FdB'\left(\int_0^{\tau_0} FF'du\right)^{-1}\int_0^{\tau_0}(FdB')'\right\} \\
&+tr\left\{\int_{\tau_0}^{1} FdB'\left(\int_{\tau_0}^{1} FF'du\right)^{-1}\int_{\tau_0}^{1}(FdB')'\right\}
\end{aligned}$$

where B *is a standard brownian motion of dimension* $p-r$ *and* F *is such that:* $F_i(u) = B_i(u), i = 1, p-r$ *and* $F_{p-r+1}(u) = 1$*, because the constant term is cointegrated.*

Thus, the relevant critical values to test whether the cointegration rank is smaller than r, depend on the dimension of the system, p, as of the relative date(s) of shift, τ_0.

In what follows, we propose to estimate $\beta^{(1)}$and $\beta^{(2}$ by implementing a sequential procedure suggested by Johansen (1995), for testing linear constraints on the β parameters of the type: $\beta = (H_1\beta^{(1)}, H_2\beta^{(2)})$. Indeed, this procedure has been proved to provide consistent estimators of the β parameters (see Boswijk, 1995) in specific cases.

Let us first notice that the minimization problem:

$$Min_{\left\{\beta^{(1)};\,\beta_2\right\}}\left|S_{00}^{(2)}(\beta^{(1)}) - S_{01}^{(2)}\beta^{(2)}(\beta^{(2)\prime}S_{11}^{(2)}\beta^{(2)})^{-1}\beta^{(2)\prime}S_{10}^{(2)}\right|$$

(resp.$Min_{\beta^{(1)}}\left|S_{00}^{(1)}(\beta_2^*) - S_{01}^{(1)}\beta^{(1)}(\beta^{(1)}S_{11}^{(2)}\beta^{(1)})^{-1}\beta^{(1)}S_{10}^{(1)}\right|$) is solved as the eigenvalue problem examined in Johansen (1995):

$$\left|\lambda H_2' S_{11.\beta^{(1)}}H_2 - H_2' S_{10\beta^{(1)}}\left(S_{00.\beta^{(1)}}\right)^{-1} S_{01\beta^{(1)}}H_2\right| = 0$$

for $\beta^{(1)} = \beta^{(1)}$, (resp.$\left|\lambda H_1' S_{11.\beta^{(2)}} H_1 - H_1' S_{10\beta^{(2)}} \left(S_{00.\beta^{(2)}}\right)^{-1} S_{01\beta^{(2)}} H_1\right| = 0$, for $\beta^{(2)} = \beta_2^*$), with the $2p \times 2p$ matrices H_1 and H_2 defined as:

$$H_1 = \begin{bmatrix} I_{dp} & 0 \\ 0 & 0 \end{bmatrix} \text{and } H_2 = \begin{bmatrix} 0 & 0 \\ 0 & I_{dp} \end{bmatrix}$$

Indeed, the problem at hand when maximizing the likelihood, for a given (unknown) rank r, with a structural break is equivalent to solving, for each r, a Johansen's problem, with the $2p \times 2r$ dimensional β matrix, constrained as:

$$Vec(\beta) = H \begin{bmatrix} Vec(\beta^{(1)}) \\ Vec(\beta^{(2)}) \end{bmatrix}$$

with $\beta^{(i)}$, $i = 1, 2$, denoting two $p \times r$ dimensional matrices, and H, a known $4pr \times 2pr$ matrix of full column rankdefined as:

$$H = \begin{bmatrix} I_{dpr} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & I_{dpr} \end{bmatrix}$$

Notice that the α matrix is also constrained in the same way:

$$Vec(\alpha) = H \begin{bmatrix} Vec(\alpha_1) \\ Vec(\alpha_2) \end{bmatrix}$$

Under such constraints, it can be proved (Boswijk, 1995)[1] that the following sequential estimation procedure provide consistent estimators of the β parameters.

More precisely, the sequential estimation procedure (Johansen,1995) goes as follows:

1) estimate the $2p \times 2p$ matrix $\varphi = (\varphi_1, \varphi_2)$ by solving the eigenvalue problem:

$$\min_{\left(\beta^{(1)};\beta_2\right)} \Phi(\varphi_1, \varphi_2) = Min_{\{\beta\}} \det(I_{d,2r} - \begin{bmatrix} \varphi_1 & \varphi_2 \end{bmatrix}' V \begin{bmatrix} \varphi_1 & \varphi_2 \end{bmatrix})$$

[1] Recently, Hansen (1999) extends this procedure.

under the constraint:

$$\left[\begin{array}{cc} \varphi_1 & \varphi_2 \end{array}\right]' \left[\begin{array}{cc} S_{kk}^{(1)} & 0 \\ 0 & S_{kk}^{(2)} \end{array}\right] \left[\begin{array}{cc} \varphi_1 & \varphi_2 \end{array}\right] = Id_{2r}$$

with

$$V = \left[\begin{array}{cc} S_{k0}^{(1)}(S_{00})^{-1}S_{0k}^{(1)} & S_{k0}^{(1)}(S_{00})^{-1}S_{0k}^{(2)} \\ S_{k0}^{(2)}(S_{00})^{-1}S_{0k}^{(1)} & S_{k0}^{(2)}(S_{00})^{-1}S_{0k}^{(2)} \end{array}\right]$$

2) Construct an initial estimate of $\beta^{(1)}$ by solving

$$\left|\lambda\widehat{\varphi}'\widehat{\varphi} - \widehat{\varphi}' H_1(H_1'H_1)^{-1}H_1'\widehat{\varphi}\right| = 0$$

as indicated in Johansen.

3) For fixed $\beta^{(1)}$,estimate β_2 by solving the eigenvalue problem

$$\left|\lambda S_{11}^{(2)} - S_{10}^{(2)}\left(S_{00}^{(2)}(\beta_1^*)\right)^{-1} S_{01}^{(2)}\right| = 0.$$

with $S_{00}^{(2)}(\beta_1^*) = S_{00}^{(1)} + S_{00}^{(2)} - S_{01}^{(1)}\beta_1^*(\beta_1'^* S_{11}^{(1)}\beta_1^*)^{-1}\beta_1'^* S_{10}^{(1)}$

4) For fixed β_2^*, estimate β_1 by solving the eigenvalue problem:

$$\left|\mu S_{11}^{(1)} - S_{10}^{(1)}\left(S_{00}^{(1)}(\beta_2^*)\right)^{-1} S_{01}^{(1)}\right| = 0.$$

with $S_{00}^{(1)}(\beta_2^*) = S_{00}^{(1)} + S_{00}^{(2)} - S_{01}^{(2)}\beta_2^*(\beta_2'^* S_{11}^{(2)}\beta_2^*)^{-1}\beta_2'^* S_{10}^{(2)}$

5) Continue with 3) and 4) until convergence.

In the next section, we implement the procedure in order to test the cointegration rank of the triavariate case analyzed by Gregory and Hansen (1996). So we can replace the single equation analysis of this system by a multivariate analysis of the dynamics, taking into account simultaneously, the cointegration properties and the existence of one structural break in the cointegration relationships.

5 Application: US money demand

The components of the trivariate dynamic system examined by Gregory and Hansen are:

$$X_t = (m_t - p_t, i_t, y_t)'$$

where m_t denotes the logarithm of the M1-money stock, p_t is the implicit output deflator, y_t is the logarithm of the real output, while i_t is a nominal interest rate (6 months commercial paper rate). The series are from the Fed of St-Louis (FRED database) and of monthly frequency. Implementing LM tests onto the VAR in level lead us to retain an order of 2 lags for the system.

The empirical study goes as follows. First we test for the cointegration rank according to the procedure proposed in the previous sections, with the structural break identified by means of Gregory and Hansen's test. Next, we estimate the VECM representation of the dynamics, and implement several causality - in the Granger sense - tests. Finally, following Lütkepohl and Reimers (1992), we do not infer any long run causality properties from the estimates of the parameters of the cointegration relationships. We prefer to we compute long run dynamic multipliers in order to investigate neutrality properties first, in the spirit of King and Watson (1993), and second, along the lines of Bruneau and Nicolaï (1995).

5.1 Estimation of the cointegration rank

The implementation of the previous rank test procedure leads us to conclude that we can not reject a cointegration rank of 1, with a structural break at the date identified in the single equation analysis of Gregory and Hansen, that is in april 1975.[2]

Cointegration with structural breaks
Gregory and Hansen's test

Z^*_α (C/S)	90 %	95 %	date of break
-54.18	-58.33	-52.85	75:4

[2]This corresponds to the results Gregory and Hansen found since their date of break is 75:2 using quaterly data.

Cointegration with structural breaks

The trace test

H_0	trace	90 %	95 %
$r = 0$	60.21	58.11	62.03
$r \leq 1$	26.10	31.98	34.83
$r \leq 2$	5.55	13.00	14.98

Moreover, we test the null hypothesis which states that the cointegration vector estimated by the single equation analysis of Gregory and Hansen belongs to the cointegration space, by implementing a standard likelihood ratio test . We do not reject this null hypothesis.

LR test	p-value
0.994	0.998

This justifies the choice expressed before, concerning the specification of the cointegration space with just one structural break. In general, the exogeneity topic should be examined when comparing the information provided by a single equation analysis compared to the one obtained in a multivariate approach. Perhaps, we could estimate a "second" cointegration relationship with multiple structural breaks. Such an estimation could be justified by economic arguments. However, we have thus to extend the Gregory and Hansen procedure to the case of several structural breaks (see Andrade, 1998) and this would complicate the present study. We prefer to base the further analysis on the multivariate results and keep the two cointegration relationships with the single structural break estimated by Gregory and Hansen (1996). However, in the general case, it would be necessary to implement weak exogenity tests in order to trust the results obtained in the single equation analysis, concerning in particular the estimation of the date of the break. This is left for further research.

We can have a look at the cointegrating vectors and propose some comments about the corresponding targets and their evolution over time.

On the first period we obtain

$$m_t - p_t = 7.769 - 0222 \times y_t + 0.084 \times i_t$$

and on the second

$$m_t - p_t = -1.171 + 0.937 \times y_t - 0.027 \times i_t$$

Although the second regime long-term equilibrium relationship is consistent with the standard monetary demand theory, this is not the case for the first one. Therefore, we prefer to investigate the properties of the long-run dynamics multipliers as suggested by Lütkepohl and Reimers (1992). However, even if we cannot interpret it by simply observing the cointegrating vectors, when can already see that the structure of the economy did clearly change over the two periods identified.

Moreover, we verify the significance of the long-run error correcting mechanism, described by the α vector. Indeed, we check that, over each period, at least one of the variable is significantly affected by the error mecanism of the VECM.

Error correcting mechanisms

Coefficient	1st period	2nd period
α_1	-0.00073 $(-2.003)^*$	-0.00164 $(-4.502)^*$
α_2	-0.03444 (-1.468)	-0.04498 $(-1.918)^*$
α_3	-0.00049 $(-5.360)^*$	-0.00014 (-1.485)

The number in brakets give the t-stat of the coefficients.

* indicates significance at the 5 % level.

5.2 VECM and Causal links in the Granger sense

We can easily test for non-causality from component X_j to component X_i, in the Granger sense, by implementing likelihood ratio tests which are, under the null of non-causality, asymptotically distributed as χ^2 distribution with k degrees of freedom.

Toda and Phillips (1995) formulate the null hypothesis of non-causality onto the VECM as

$$\{\Gamma_{ij,1} = ... = \Gamma_{ij,k-1} = \Pi_{ij} = 0\}$$

For the case of study, we obtain the following results:

Non-Causality Test - First Period

		Causality from		
		$m_t - p_t$	i_t	y_t
to	$m_t - p_t$	.	10.755 (0.005)	0.698 (0.705)
	i_t	0.976 (0.614)	.	0.709 (0.701)
	y_t	0.076 (0.963)	12.610 (0.002)	.

The table reports value of the χ^2 statistic.
The number in brakets are the significance level.

On the first subsample, significant causal links semerge from the interest rate to the money stock and the output.

Non-Causality Test - Second period

		Causality from		
		$m_t - p_t$	i_t	y_t
to	$m_t - p_t$	.	47.227 (0.000)	18.270 (1.000)
	i_t	13.895 (0.001)	.	14.723 (0.001)
	y_t	0.015 (0.992)	5.411 (0.067)	.

The table reports value of the χ^2 statistic.
The number in brakets are the significance level.

On the second period, still the interest rate causes real money stock as real output, but the converse also holds. Money as output cause the interest rate.

Notice that the results of the previous test do not allow us to distinguish short-run from long-run causal links as pointed out by Bruneau and Jondeau (2000). That is the reason why we turn to impulse response analysis of the dynamics and focus on long run dynamic multipliers which can be used to characterize long run causality.

5.3 Long run dynamic multipliers, neutrality and long run causality properties

In this section, we focus on the effects of shocks, in the spirit of Sims (1980, 1981). More precisely, we are interested in the long run effects of the stochastic shocks which are characterized as canonical or orthogonalized statistical innovations. In what follows, we first focus on "canonical" long run dynamic multipliers in order to test for neutrality properties (in a statistical sense) as investigated by King and Watson (1993).

5.3.1 (Statistical) Neutrality properties

A variable X_j is said (statistically) neutral for the X_i variable, if the corresponding long-run dynamic multiplier $C_{ij}(1)$ is equal to zero. Once the VECM has been estimated, we have the expression of the long run dynamic multipliers as functions of the parameters of the VECM. Moreover, the asymptotic normality of the estimators holds. Accordingly, a neutrality test is just a standard likelihood ratio test, with a $\chi^2(1)$ asymptotic distribution of the test statistic, if one limits the analysis to each sub-periods, or a $\chi^2(2)$ asymptotic distribution, if the neutrality property is analysed over the whole period. The results are the following:

Neutrality Test - First period

		Long-run Neutrality from		
		$m_t - p_t$	i_t	y_t
over	$m_t - p_t$	.	-0.019 (2.597)	-0.027 (0.258)*
	i_t	-18.158 (2.793)	.	7.701 (4.078)
	y_t	-0.134 (1.899)*	-0.003 (2.320)	.

The number in brakets give the t-stat of the coefficients.
* indicates significance at the 5 % level.

Neutrality Test - Second period

		Long-run Neutrality from		
		$m_t - p_t$	i_t	y_t
over	$m_t - p_t$	.	-0.012 (3.788)	0.527 (2.539)
	i_t	-15.018 (2.162)	.	-9.967 (1.050)*
	y_t	-0.027 (1.457)*	-0.000 (0.925)*	.

The number in brakets give the t-stat of the coefficients.
* indicates significance at the 5 % level.

Over the two subsamples, money is neutral over output. If the interest rate significantly influences money and output over the first subperiod (we cannot accept the nul of neutrality), it does not influence output over the second regime. Finally, output is neutral with regards to money over the first regime, but not over the second one and conversely neutral over the interest rate over the second subsample but not over the first one.

5.3.2 Long run causality along the lines of Bruneau and Nicolaï (1995)

The notion of statitical neutrality has been questioned because there are instantaneous correlations between the variables, and, accordingly, between the canonical innovations ε_i, $i = 1, ...3$. As a consequence, it is not easy to comment separately the effects of the differents innovations.

Sims (1981) propose to orthogonalized the innovations by introducing a Choleski decomposition of the variance of the canonical innovations. Thus, an implicit causal ordering has to be chosen a priori, from the most exogeneous variable to most endogeneous one. For example, in Lütkepohl and Reimers (1992), the variables are ordered as following: $X_t = (i_t, y_t, m_t - p_t)'$. Once a causal ordering has been chosen, a Choleski decomposition of the variance matrix of the innovations Ω provides a inferior triangular matrix P such that:

$$PP' = \Omega$$

Thus, one can characterize the effects of the orthogonalized innovations:

$$\eta_t = P^{-1}\varepsilon_t$$

by focusing on the long run dynamic multipliers $[C_t(1)P]_{ij}$.

However, in these impulse response analysis, the results depend on the causal ordering of the variables which is chosen a priori. One can prefer to characterize long run causal links along the lines of Bruneau and Nicolaï (1995). Indeed, the properties which are investigated do not rely on any causal ordering and can be interpreted as long run prediction improvement.

If one chooses as the first compent the presumed causal variable and if one computes the matrix P associated with the corresponding Choleski decomposition of the variance matrix of the canonical innovations, one can prove that the nullity of the long run dynamic multiplier $[C(1)P]_{i1}$ can be interpreted as a non-causality property in the long run, from the presumed causal variable to the variable X_i as follows:

Proposition 10 *: The nullity of the long run dynamic multiplier* $[C(1)P]_{i1}$*, for the choleski decompition* $PP' = \Omega$*, is equivalent to the non-causality* *property from the first component to the X_j variable, as charaterized by:*

$$\lim_{H \to +\infty} EL(X_{it+H}/X_{1t}; \underline{X_{1t-1}}, ..., \underline{X_{pt-1}}) = \lim_{H \to +\infty} EL(X_{it+H}/\underline{X_{1t-1}}, ..., \underline{X_{pt-1}})$$

where EL denotes the linear regresion operator and $\underline{X_{jt-1}}$ is the set of the lagged variables $X_{jt-k}, k \geq 1$.

Implementing this procedure of test leads to the following results:

Long-Run Non-Causality Test - First period

		Persistent Causality from		
		$m_t - p_t$	i_t	y_t
to	$m_t - p_t$	.	-0.008 (1.946)	-0.0000 (0.474)*
	i_t	-0.126 (2.666)	.	0.013 (2.259)
	y_t	-0.001 (0.447)*	-0.001 (1.944)	.

The number in brakets give the t-stat of the coefficients.
* indicates significance at the 5 % level.

Long-Run Non-Causality Test - Second period

		Persistent Causality from		
		$m_t - p_t$	i_t	y_t
to	$m_t - p_t$	.	-0.006 (2.310)	0.001 (2.339)
	i_t	-0.104 (3.508)	.	-0.017 (2.355)
	y_t	-0.0001 (0.162)*	-0.0002 (2.305)	.

The number in brakets give the t-stat of the coefficients.
* indicates significance at the 5 % level.

Therefore, over the two regimes, money persistantly causes the interest rate but not ouput. By contrast, still over both regimes, the interest rate persistently causes money as output. Lastly, output cuses over the long-run the interest rate over the first subsample and money and the interest rate over the first one.

The results may be summarized as follows. On the first regime there seems to be an unilateral causal links from the interest rate to the other variables of the system. By contrats the second one is characterized by cross causality relations, which may be interpreted as monetary policy reaction functions.

6 Conclusion

This paper extends the test of cointegration with one structural break of Gregory and Hansen (1996) in a single equation approach framework to a multivariate one. This allows first to verify this hypothesis in a generalized framework and to perform analysis of the long run properties of the dynamics. Our results show that once the structural break hypothesis is accepted, the dynamics exhibit strong different properties over the two regimes identified in term of long-run equilibrium as long-run dynamics multipliers. Allowing for different ranks over these two periods presents an interesting development and is dedicated for further research.

References

[1] P. Andrade. Cointegration and structural breaks: Extending the Gregory and Hansen's test to the case of multiple breaks. mimeo, THEMA, Université Paris X Nanterre, 1998.

[2] P. Andrade and C. Bruneau. Estimation of long-run relationships with structural breaks in a multivariate framework. Manuscript prepared for the ESEM99 conference, THEMA, Université Paris X Nanterre, 1999.

[3] J. Bai and P. Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66:47–78, 1998.

[4] H. P. Boswijk. Identifian=bility of cointegrated systems. Working Paper, Tinbergen Institute, 1995.

[5] C. Bruneau and Nicola

[6] J. Y. Campell and R. J. Shiller. Interpreting cointegrated models. *Journal of Economic Dynamics and Control*, 12:505–522, 1988.

[7] A. W. Gregory and B. E. Hansen. Residuals-based tests for cointegration in models with regime shifts. *Journal of Econometrics*, 70:99–126, 1996.

[8] P. R. Hansen. Structural breaks in the cointegrated vector autoregressive model. Manuscrpit, University of California, San Diego, 1999.

[9] S. Johansen. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12:231–254, 1988.

[10] S. Johansen. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59:1551–1580, 1991.

[11] S. Johansen. Cointegration in partial systems and the efficiency of single-equation analysis. *Journal of Econometrics*, 52:389–402, 1992.

[12] R. G. King and M. W. Watson. Testing for long-run neutrality. *NBER working paper*, (4156), 1993.

[13] H. Lütkepohl and H.-E. Reimers. Impulse response analysis of cointegrated systems. *Journal of Economic Dynamics and Control*, 16:53–78, 1992.

[14] E. Mellander, A. Vredin, and A. Warne. Stochastic trends and economic fluctuations in a small open economy. *Journal of Applied Econometrics*, 7:369–394, 1992.

[15] P. C. Phillips and P. Perron. Testing for a unit root in time series regression. *Biometrika*, 75:335–346, 1988.

[16] C. A. Sims. Macroeconomics and reality. *Econometrica*, 48:1–48, 1980.

[17] C. A. Sims. An autoregressive index model for the us, 1948-1975. In J. Kmenta and J. Ramsey, editors, *Large-scale Macroeconometrics Models*, pages 283–327. North Holland, Amsterdam, 1981.

[18] H. Y. Toda and P. C. Phillips. Vector autoregression and causality. *Econometrica*, 61:1367–1394, 1993.

[19] A. Warne. A common trends model: Identification, estimation and inference. *Institute for International Economic Studies' Seminar Paper*, (555), 1993.