

Rocco Mosconi, Paolo Paruolo

Identification of cointegrating  
relations in I(2) Vector  
Autoregressive models

2010/07



UNIVERSITÀ DELL'INSUBRIA  
FACOLTÀ DI ECONOMIA

<http://eco.uninsubria.it>

I punti di vista espressi nei quaderni della Facoltà di Economia riflettono unicamente le opinioni degli autori, e non rispecchiano necessariamente quelli della Facoltà di Economia dell'Università dell'Insubria.

The views expressed in the working papers reflect the opinions of the authors only, and not necessarily the ones of the Economics Faculty of the University of Insubria.

© Copyright Massimo Franchi and Paolo Paruolo  
Printed in Italy in December 2010  
Università degli Studi dell'Insubria  
Via Monte Generoso 71, 21100 Varese, Italy

All rights reserved. No part of this paper may be reproduced in any form without permission from the Author.

# IDENTIFICATION OF COINTEGRATING RELATIONS IN I(2) VECTOR AUTOREGRESSIVE MODELS

ROCCO MOSCONI<sup>°</sup> AND PAOLO PARUOLO<sup>‡</sup>

**ABSTRACT.** This paper discusses identification within a new parametrization for I(2) systems, where the integral and proportional control cointegrating relations are not necessarily orthogonal. The new parametrization, while equivalent to previously proposed ones, gives more flexibility in choosing the variables to include in first differences in the integral and proportional control term. We discuss the joint identification of the cointegrating relations, providing rank and order conditions. We discuss likelihood estimation, and propose a simple alternating algorithm for likelihood-maximization, under the cases of under- exact- and over-identification. An illustration on US consumption is also presented.

---

*Date:* May 26, 2011.

*JEL classification:* C32, C51.

*Key words and phrases.* Identification, proportional and integral control, multi-cointegration, I(2), VAR.

<sup>°</sup> Politecnico di Milano, (I) (corresponding author).

<sup>‡</sup> Università dell'Insubria, Varese (I).

*Corresponding address:* Rocco Mosconi, Dipartimento di Ingegneria Gestionale, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milano, Italy; Phone: +39 0223992747; Fax : +39 02700423151. Email: rocco.mosconi@polimi.it.

## CONTENTS

1. Introduction	1
2. Representation	2
2.1. $I(2)$ processes	2
2.2. Error correction formulation	3
2.3. Deterministic components	4
3. Identification	5
3.1. The identification problem	5
3.2. General linear restrictions	6
3.3. Equation-by-equation restrictions	7
3.4. Rank and order conditions	8
3.5. Rank of the Jacobian	9
3.6. Identification strategy	11
4. Likelihood analysis	12
4.1. An alternating maximization algorithm	12
4.2. Restrictions on the cointegrating coefficients	13
4.3. Restrictions on the adjustment coefficients	13
4.4. Asymptotics	14
5. An application to US consumption	14
5.1. Routine tests	15
5.2. Number of $CI(2,2)$ relations	18
5.3. Testing for the absence of breaks	19
5.4. Nominal to real transformation	20
5.5. Identified relations	21
6. Conclusions	25
Acknowledgements	26
References	26
A. Appendix	27
A.1. Representation	28
A.2. Connection matrices	28
A.3. Rank condition	29
A.4. Jacobian	30
A.5. Standard errors	33

## 1. INTRODUCTION

Cointegrated (CI) vector autoregressive systems (VAR), with their dual error-correction representation (ECM), provide interesting statistical counterparts to the economic concepts of dynamic equilibrium and adjustment towards it. Over the last decades, these characteristics have first catalyzed interest on systems integrated of order 1,  $I(1)$ , see Engle and Granger (1987), and subsequently on systems integrated of order 2,  $I(2)$ .

$I(2)$  systems have been studied quite extensively, see e.g. Boswijk (2000, 2010), Johansen (1992, 1995b, 1997, 2006), Kitamura (1995), Kurita, Nielsen, and Rahbek (2011), Paruolo (2000), Paruolo and Rahbek (1999), Rahbek, Kongsted, and Jørgensen (1999), Stock and Watson (1993). Their ECM representation includes two types of error correction terms. The first one consists of linear combinations of levels and differences of the processes, and it has been named the ‘integral control term’, see Haldrup and Salmon (1998), or ‘multi-cointegration’, see Granger and Lee (1989), Engsted and Johansen (1999); it is a special case of ‘polynomial-cointegration’ as introduced by Engle and Yoo (1991). The second type of error correction terms consists of linear combinations of the first differences of the process only, and it has been called the ‘proportional control term’.

The presence of integral and proportional control terms gives a richer, albeit admittedly more complicated, structure for the cointegrating relations. The first differences of the integral-control terms, in fact, qualify as proportional-control terms, which potentially include also additional cointegrating relations. This gives rise to an identification problem, which shares some similarity to the one encountered for cointegrating relations in  $I(1)$  systems and in classical systems of simultaneous equations.

Stock and Watson (1993) considered a linear process setup and stated a (recursive) triangular form for the integral and proportional control terms; this is a special case of a just-identified, reduced-form system. Boswijk (2000) discussed the relation between the triangular form and the just-identified case of  $I(2)$  VAR systems. In this paper we concentrate attention to  $I(2)$  VAR systems, discussing not only the just-identified case but also the under- and over-identified ones. However, our identification results hold for the more general class of linear processes employed e.g. in Stock and Watson (1993).

The identification of the linear combination of the levels that enter the integral control term has been implicitly treated in Johansen (1995a, 1996). When linear combinations of the levels are (over-)identified, Rahbek, Kongsted, and Jørgensen (1999), Kurita, Nielsen, and Rahbek (2011) discussed a just-identified version of the associated coefficients that multiply the differences of the process. The identification of the proportional control equations has not been addressed in the literature.

In this paper we discuss the joint identification of integral and proportional control cointegrating relations. In view of the discussion of identification, we first introduce a new parametrization for the  $I(2)$  model, based on a different decomposition of the matrices that appear in the ECM representation. This allows to define the CI relations that span the proportional control terms as

containing the same linear combinations that multiply the levels in the integral control term and other additional CI relations, that are not necessarily orthogonal to them.

We discuss the joint identification of the CI relations that appear in the integral and proportional control terms, providing rank and order conditions for general linear identifying restrictions, as well as for the leading special case of linear equation-by-equation constraints. The new parametrization, while equivalent to previously proposed ones, gives more flexibility in choosing the variables to include in first differences in the integral and proportional control term, because the cointegrating relations are not partially-identified via orthogonality.

We work within the likelihood-based inference framework for VAR processes with I(2) variables developed by Johansen (1997). We discuss likelihood estimation of the proposed parametrization, and we show that one can devise a simple alternating algorithm for the maximization of likelihood, which works under the cases of under-, exact-, and over-identification.

An application on the US consumption data also analyzed in Kurita, Nielsen, and Rahbek (2011) illustrates the identification issues discussed in the paper. For the empirical analysis we adopt the ‘star-specification’ of deterministic linear terms introduced in Rahbek, Kongsted, and Jørgensen (1999), and later extended by Kurita, Nielsen, and Rahbek (2011) along the lines of Johansen, Mosconi, and Nielsen (2000) to allow for the presence of broken linear trends in all directions of the process.

The rest of the paper is organized as follows. Section 2 presents the new ECM formulation and states the I(2) conditions. Section 3 discusses identification in the proposed parametrization. Section 4 presents the statistical analysis based on the likelihood. Section 5 reports the application to US consumption, while Section 6 concludes. Proofs are placed in the Appendix.

In the following  $a := b$  and  $b =: a$  indicate that  $a$  is defined by  $b$ ;  $(a : b)$  indicates the matrix obtained by horizontally concatenating  $a$  and  $b$ . For any full column rank matrix  $H$ ,  $\text{col}(H)$  is the linear span of the columns of  $H$ ,  $\bar{H}$  indicates  $H(H'H)^{-1}$  and  $H_{\perp}$  indicates a basis of  $\text{col}^{\perp}(H)$ , the orthogonal complement of  $\text{col}(H)$ . Moreover  $P_H := H\bar{H}'$ ,  $H_A := H(A'H)^{-1}$ . Finally  $(\cdot)_{ij}$  indicates the  $ij$ -th element of the argument matrix,  $\text{vec}$  is the column stacking operator,  $\otimes$  is the Kronecker product,  $\text{blkdiag}(A_1, \dots, A_n)$  a block-diagonal matrix with  $A_1, \dots, A_n$  as diagonal blocks.

## 2. REPRESENTATION

This section introduces notation, states the I(2) conditions and present a general ECM formulation, which is later used as a basis for the new parametrization. We consider a  $p$ -dimensional VAR process  $X_t$  generated by  $\Pi(L)X_t = \varepsilon_t$  where  $\Pi(z) = I - \sum_{j=1}^k \Pi_j z^j$  is a  $p \times p$  matrix polynomial,  $L$  is the lag operator and  $\varepsilon_t$  is an i.i.d.  $p \times 1$  vector with expectation 0 and  $p \times p$  positive definite covariance matrix  $\Omega$ . Deterministic components are added to this formulation in Subsection 2.3 below.

**2.1. I(2) processes.** A useful re-writing of the VAR equations is given by the following form, see Johansen (1992)

$$\Delta^2 X_t = \Pi X_{t-1} + \Gamma \Delta X_{t-1} + \Upsilon \Delta^2 \mathbb{X}_{t-1} + \varepsilon_t, \quad (1)$$

where  $\Upsilon := (\Upsilon_1 : \dots : \Upsilon_{k-2})$ ,  $\Delta^2 \mathbb{X}_{t-1} := (\Delta^2 X'_{t-1} : \dots : \Delta^2 X'_{t-k+2})'$ ,  $\Pi := -\Pi(1)$ ,  $\Gamma := \dot{\Pi}(1) - \Pi(1)$  with  $\dot{\Pi}(z) := d\Pi(z)/dz$ . We assume that the roots of  $\det \Pi(z)$  are outside the unit disk or at the point  $z = 1$ . For these processes, Johansen's representation theorem, see Johansen (1992), provides the following necessary and sufficient conditions on  $\Pi(z)$  for  $X_t$  to be integrated of order 2.

**Definition 1 (I(2) conditions).** *The following three statements on (1) are defined as the I(2) conditions:*

- I2.1  $\Pi = \alpha\beta'$ , with  $\alpha, \beta$  full-column rank  $p \times r$  matrices,  $0 \leq r < p$ ;
- I2.2  $P_{\alpha_{\perp}} \Gamma P_{\beta_{\perp}} = \alpha_1 \beta_1'$ , with  $\alpha_1, \beta_1$  full-column rank  $p \times s$  matrices,  $0 \leq s < p - r$ ;
- I2.3  $(I - P_{(\alpha:\alpha_1)}) \left( \Gamma \bar{\alpha}' \Gamma + I - \sum_{i=1}^{k-2} \Upsilon_i \right) (I - P_{(\beta:\beta_1)})$  is of rank  $p - r - s$ .

The class of processes (1) under the conditions I2.1 and I2.2 with unrestricted remaining parameters is called the 'I(2) model', see Johansen (1997).

Equalities in conditions I2.1 and I2.2 define rank decompositions of the matrices on the l.h.s.; we allow the ranks to be 0, in which case  $\alpha, \beta$  or  $\alpha_1, \beta_1$  are understood to equal 0 vectors. Remark that the I(2) conditions define two partitions of  $\mathbb{R}^p$  into orthogonal linear subspaces, spanned by blocks of  $(\alpha : \alpha_1 : \alpha_2)$  and  $(\beta : \beta_1 : \beta_2)$ , where  $\alpha_2$  and  $\beta_2$  are a basis of  $\text{col}^{\perp}(\alpha : \alpha_1)$  and  $\text{col}^{\perp}(\beta : \beta_1)$  respectively.

Johansen's representation theorem states that under the I(2) conditions,  $\beta' X_t - \bar{\alpha}' \Gamma \Delta X_t, \tau' \Delta X_t$  and  $\Delta^2 X_t$  are I(0), where  $\tau$  is any basis of  $\text{col}(\beta : \beta_1)$ . The relations  $\beta' X_t - \bar{\alpha}' \Gamma \Delta X_t$  are called polynomially cointegrated or multi-cointegrated (and correspond to an integral-control term), while  $\tau' \Delta X_t$  (which correspond to the proportional control term) describe the cointegrating relations of the differences, the so-called CI(2,1) relations in the terminology of Engle and Granger (1987).

**2.2. Error correction formulation.** In this section we discuss a decomposition of  $\Gamma$  in (1) under the I(2) conditions. Assume that a matrix  $\gamma$  is given, of dimension  $p \times s$ , such that

$$\text{col}(\beta : \gamma) = \text{col}(\beta : \beta_1), \quad (2)$$

The following theorem presents a class of decompositions based on any  $\gamma$  satisfying (2).

**Theorem 2 (ECM formulation).** *Under the I(2) conditions and requirement (2) one has:*

$$\Gamma = \alpha v' + \xi \gamma' + \varsigma \beta' = (\alpha : \xi : \varsigma) (v : \gamma : \beta)' \quad (3)$$

with

$$\text{col}(\alpha : \xi) = \text{col}(\alpha : \alpha_1), \quad (4)$$

and

$$\begin{aligned} v &:= \beta_2 \delta' + \gamma b' + \beta a', \\ \xi &:= \Gamma \beta_1 (\gamma' \beta_1)^{-1} - \alpha b, \\ \varsigma &:= \Gamma (I - \beta_1 (\gamma' \beta_1)^{-1} \gamma') \bar{\beta} - \alpha a, \end{aligned} \quad (5)$$

for arbitrary matrices  $a, b$  of dimensions  $r \times r$  and  $r \times s$  respectively (not necessarily of full rank) and  $\delta := \bar{\alpha}' \Gamma \bar{\beta}_2$ .

The resulting ECM formulation from Theorem 2 is:

$$\Delta^2 X_t = \left[ \alpha (\beta' : v') \begin{pmatrix} X_{t-1} \\ \Delta X_{t-1} \end{pmatrix} \right] + \left[ (\xi : \varsigma) \begin{pmatrix} \gamma' \\ \beta' \end{pmatrix} \Delta X_{t-1} \right] + \Upsilon \Delta^2 \mathbb{X}_{t-1} + \varepsilon_t. \quad (6)$$

The terms in the square brackets on the r.h.s. describe the integral- and proportional-control mechanisms. Several other formulations for the ECM in I(2) VAR have been proposed, see e.g. Johansen (1997) and Paruolo and Rahbek (1999). In the proof of Theorem 2 we show the equivalence of the present ECM formulation with the one in Paruolo and Rahbek (1999). We refer to Paruolo and Rahbek (1999) and Boswijk (2000) for the connections of the latter formulation with the one in Johansen (1997) and with the triangular form of Stock and Watson (1993).

**2.3. Deterministic components.** The ECM formulation in (6) can be extended in order to allow for the presence of deterministic (broken) linear trends in all directions of  $X_t$  as in Kurita, Nielsen, and Rahbek (2011). Define  $T_i$  be the last observations in sub-period  $i = 1, \dots, q$ , with  $T_0 := 0 < T_1 < \dots < T_q := T$ . Let  $D_{i,t} := (t - T_{i-1})1(t > T_{i-1})$ ,  $D_t = (D_{1,t} : \dots : D_{q,t})'$  and  $X_t^* := (X_t' : D_t)'$ , so that  $\Delta X_t^* := (\Delta X_t' : \Delta D_t)'$ , where  $\Delta D_{1t} = 1$ ,  $\Delta D_{it} = 1(t > T_{i-1})$ . Finally, let  $\mathbb{D}_t$  contain  $\Delta^2 D_{it-j}$  for  $i = 1, \dots, q-1$  and  $j = 1, \dots, k$ , as well as other impulse dummies, where  $\Delta^2 D_{it} = 1(t = T_{i-1} + 1)$ . Remark that  $D_{1t} = t$  and  $\Delta D_{1t} = 1$ , so that  $X_t^*$  and  $\Delta X_t^*$  contain  $t$  and 1 respectively.

The extended ‘‘star’’ formulation is given by

$$\Delta^2 X_t = \left[ \alpha (\beta^{*'} : v^{*'}) \begin{pmatrix} X_{t-1}^* \\ \Delta X_{t-1}^* \end{pmatrix} \right] + \left[ (\xi : \varsigma) \begin{pmatrix} \gamma^{*'} \\ \beta^{*'} \end{pmatrix} \Delta X_{t-1}^* \right] + (\Upsilon : \mu) \begin{pmatrix} \Delta^2 \mathbb{X}_{t-1} \\ \mathbb{D}_{t-1} \end{pmatrix} + \varepsilon_t, \quad (7)$$

where  $\beta^{*'} = (\beta' : \beta'_D)$ ,  $\gamma^{*'} = (\gamma' : \gamma'_D)$ ,  $v^{*'} = (v' : v'_D)$ . Eq. (7) gives rise to a VAR of the type

$$\Delta^2 X_t = \Pi X_{t-1} + \Gamma \Delta X_{t-1} + \Upsilon \Delta^2 \mathbb{X}_{t-1} + \mu_1 D_t + \mu_0 \Delta D_t + \mu \mathbb{D}_t + \varepsilon_t \quad (8)$$

Rahbek (1997) modified the I(2) conditions in order to define necessary and sufficient conditions on  $\mu_0$ ,  $\mu_1$  for  $X_t$  not to contain (broken) quadratic trends, but just (broken) linear trends. The extended I(2) conditions in his Theorem 4.1 can be are stated as in the following definition.

**Definition 3 (I(2) conditions with deterministic (broken) linear trends).** *The I(2) conditions with deterministic (broken) linear trends on eq. (8) consist of I2.1, I2.2, I2.3 from the I(2) conditions and of the following two extra conditions*

$$\text{I2.4 } \mu_1 = \alpha \beta'_D,$$

$$\text{I2.5 } \bar{\alpha}'_2 \mu_0 = \bar{\alpha}'_2 \Gamma \bar{\beta} \beta'_D.$$

Rahbek (1997) proved that necessary and sufficient conditions for  $X_t$  in (8) not to contain quadratic trends for the case  $D_t = t$  are given by the conditions in Definition 3. As argued in Kurita, Nielsen, and Rahbek (2011), his proof is identical for the present case where  $D_t$  contains broken linear trends.

We next verify that the ECM formulation (7) satisfies I2.4 and I2.5. One finds  $\mu_1 = \alpha \beta'_D$  and  $\mu_0 = (\alpha : \xi : \varsigma) (v_D : \gamma_D : \beta_D)'$ ; hence one sees that ECM formulation (7) satisfies condition I2.4.

Regarding condition I2.5, observe that, by (5) in Theorem 2, one has

$$\bar{\alpha}'_2 \varsigma = \bar{\alpha}'_2 \left( \Gamma(I - \beta_1 (\gamma' \beta_1)^{-1} \gamma') \bar{\beta} - \alpha a \right) = \bar{\alpha}'_2 \Gamma \bar{\beta}$$

and that by (4) in Theorem 2, one has  $\bar{\alpha}'_2 \mu_0 = \bar{\alpha}'_2 \varsigma \beta'_D = \bar{\alpha}'_2 \Gamma \bar{\beta} \beta'_D$  as requested by condition I2.5. Hence we conclude that the extended star formulation (7) does not generate quadratic trends in  $X_t$ . For simplicity of exposition, in the rest of the paper we omit the deterministic terms, and return to the extended formulation (7) in the empirical application.

### 3. IDENTIFICATION

The ECM formulation (6) in Theorem 2 can be used to parametrize the I(2) model. Recall that the I(2) model is the class of VAR processes which satisfy conditions I2.1 and I2.2. Consider the class of processes in (6) indexed by the parameters  $\alpha$ ,  $\xi$ ,  $\varsigma$ ,  $\beta$ ,  $v$ ,  $\gamma$ ,  $\Upsilon$ ,  $\Omega$ , where these parameters are unrestricted, with the exception of  $\Omega$ , which is assumed to be a symmetric p.d. matrix. The processes in this class satisfy conditions I2.1 and I2.2, because  $\Pi = \alpha \beta'$  and  $P_{\alpha_\perp} \Gamma P_{\beta_\perp} = \alpha_1 \beta'_1$ . Conversely, given any process that satisfies I2.1 and I2.2, one can apply Theorem 2 to obtain processes of the form (6). Hence the present class of processes is a reparametrization of the I(2) model.

In this section we discuss the identification of parametrization (6) under linear restrictions. Extending the approach proposed by Johansen (1995a) for the I(1) model, we focus attention on (economically-motivated) identifying restrictions specified equation-by-equation on the cointegrating vectors. This approach is grounded in a large econometric literature on identification for system of simultaneous linear equations, see Sargan (1988) and references therein.

We first state the identification problem and define the relevant linear restrictions; next we give rank and order conditions for the identification problem using the implicit form of the restrictions. We then connect the rank condition to the rank of the relevant Jacobian matrix, see e.g. Rothenberg (1971).

**3.1. The identification problem.** Within the proposed parametrization in eq. (6), we focus our attention on

$$\Phi := (\Pi : \Gamma) = \eta \zeta' := \begin{pmatrix} \alpha & \xi & \varsigma \end{pmatrix} \begin{pmatrix} \beta & 0 & 0 \\ v & \gamma & \beta \end{pmatrix}'. \quad (9)$$

The CI parameters are contained in  $\zeta$ ; its first  $r$  columns represents the integral control (multi-cointegrating) relations, while the last  $s + r$  columns represent the proportional control relations.

We note that the parameters  $\eta$  and  $\zeta$  are not identified because they enter the likelihood only through their product  $\Phi = \eta \zeta'$ ; this makes  $(\eta, \zeta)$  and  $(\eta^\circ, \zeta^\circ)$  observationally equivalent parameters when  $(\eta^\circ, \zeta^\circ)$  are defined by  $\eta \zeta' = (\eta Q^{-1'}) (Q' \zeta') =: \eta^\circ \zeta^{\circ'}$  with  $\zeta^\circ := \zeta Q$ ,  $\eta^\circ := \eta Q^{-1'}$ . Here  $Q$  is

any square, non-singular, lower block-triangular  $(2r + s)$ -dimensional matrix of the form

$$Q = \begin{pmatrix} Q_{11} & 0 & 0 \\ Q_{21} & Q_{22} & 0 \\ Q_{31} & Q_{32} & Q_{11} \end{pmatrix}. \quad (10)$$

For this choice of  $Q$ , in fact, we note that  $\zeta^\circ$  and  $\zeta$  have the same upper-right zero blocks

$$\begin{pmatrix} \beta^\circ & 0 & 0 \\ v^\circ & \gamma^\circ & \beta^\circ \end{pmatrix} = \begin{pmatrix} \beta & 0 & 0 \\ v & \gamma & \beta \end{pmatrix} Q.$$

Moreover, note that

$$\beta^\circ = \beta Q_{11}, \quad \gamma^\circ = \gamma Q_{22} + \beta Q_{32} \quad (11)$$

which implies that  $\text{col } \beta = \text{col } \beta^\circ$  and  $\text{col } (\beta : \gamma) = \text{col } (\beta^\circ : \gamma^\circ)$ , thanks to the fact that  $Q_{11}$  and  $Q_{22}$  are non-singular, due to the invertibility of  $Q$ .

Recall that in the classic identification problem in system of simultaneous linear equations the matrix  $Q$  is square and non-singular, while (10) requires  $Q$  to be lower block-triangular, with the upper-left  $r \times r$  block equal to the lower-right  $r \times r$  block.

**3.2. General linear restrictions.** Denote

$$\zeta_{2p \times (2r+s)} =: \begin{pmatrix} \zeta^1 & : & \zeta^2 & : & \zeta^3 \end{pmatrix}_{\substack{2p \times r \\ 2p \times s \\ 2p \times r}} = \begin{pmatrix} \beta & 0 & 0 \\ v & \gamma & \beta \end{pmatrix}, \quad (12)$$

Let also  $\zeta^{12} := (\zeta^1 : \zeta^2)$ . We consider linear restrictions on  $\zeta$ ; these restrictions can be stated terms of the vector

$$\theta_\zeta := \text{vec } (v : \gamma : \beta),$$

because it contains the same nonzero elements as  $\zeta$  and  $\zeta^{12}$ . In the following we let  $f := p(2r + s)$  indicate the number of rows in  $\theta_\zeta$ .

General linear restrictions on  $\theta_\zeta$  are formulated as

$$\theta_\zeta = h_\zeta + H_\zeta \phi_\zeta, \quad \text{or} \quad R'_\zeta \theta_\zeta = c_\zeta \quad (13)$$

$$\substack{f \times 1 \\ f \times n_\zeta} \quad \substack{m_\zeta \times f \\ m_\zeta \times 1}$$

with  $R_\zeta := H_{\zeta^\perp}$ ,  $c_\zeta := R'_\zeta h_\zeta$ , where by the Rank-Nullity Theorem one has  $n_\zeta + m_\zeta = f$ .  $\phi_\zeta$  are the unrestricted parameters, the matrices  $H_\zeta$ ,  $h_\zeta$  represent the design matrices of the restrictions in explicit form, with  $h_\zeta$  indicating their normalization. The  $R_\zeta$ ,  $c_\zeta$  design matrices express the same restrictions in implicit form. Both the explicit and the implicit forms are used in the following.

The restrictions in (13) can be equivalently stated in terms of  $\zeta$  or  $\zeta^{12}$ , using the fact that

$$\text{vec } \zeta = A \text{vec } \zeta^{12}, \quad \text{vec } \zeta^{12} = B \theta_\zeta \quad (14)$$

where  $A$  and  $B$  are appropriate 0-1 matrices, described in detail in Appendix A.2. It is simple to see that pre-multiplying  $\theta_\zeta = h_\zeta + H_\zeta \phi_\zeta$  in (13) by  $B$ , one directly obtains

$$h_\dagger = B h_\zeta, \quad H_\dagger = B H_\zeta \quad (15)$$

in

$$\text{vec } \zeta^{12} = h_{\dagger} + H_{\dagger} \phi_{\zeta}, \quad \iff \quad R'_{\dagger} \text{vec } \zeta^{12} = c_{\dagger}. \quad (16)$$

From (15) one deduces  $R_{\dagger} = (B_{\perp} : \bar{B}R_{\zeta})$ ,  $c_{\dagger} = R'_{\dagger}h_{\dagger} = (0' : c'_{\zeta})'$ . A similar construction holds for  $\text{vec } \zeta$ .

Conversely, if (16) is given, note that (15) can be inverted pre-multiplying both sides by  $\bar{B}'$ , obtaining  $h_{\zeta} = \bar{B}'h_{\dagger}$ ,  $H_{\zeta} = \bar{B}'H_{\dagger}$ . From this one also obtains  $R_{\zeta}$  as  $H_{\zeta\perp}$  and  $c_{\zeta}$  as  $R'_{\zeta}h_{\zeta}$ . This shows that (13) and (16) are equivalent, and if one of them is given, the other one is uniquely determined and vice versa. For equation-by-equation restrictions, the formulation (16) in terms of  $\text{vec } \zeta^{12}$  appears somehow more natural. We hence employ both (13) and (16), using the relations (15) to convert between the two, when necessary.

**3.3. Equation-by-equation restrictions.** We indicate the  $i$ -th column of  $\zeta$  as  $\zeta_i$ ,  $i = 1, \dots, 2r+s$ . Consider the important special case of equation-by-equation linear restrictions on the  $r+s$  columns  $\zeta_i$ ,  $i = 1, \dots, r+s$  in  $\zeta^{12}$ , of the type

$$\zeta_i = h_i + H_i \phi_i, \quad \iff \quad R'_i \zeta_i = c_i \quad (17)$$

where  $R_i = H_{i\perp}$ ,  $c_i = R'_i h_i$ . Consider the integral control relations  $\zeta_i$  in  $\zeta^1$  for  $i = 1, \dots, r$ . The design matrices  $h_i$ ,  $H_i$  for  $i = 1, \dots, r$  can be partitioned as follows without loss of generality, see Lemma 7(i) in Appendix A.4:

$$h_i = \begin{pmatrix} h_i^{\beta} \\ h_i^{\nu} \end{pmatrix}, \quad H_i = \begin{pmatrix} H_i^{\beta} & 0 \\ H_i^{\beta\nu} & H_i^{\nu} \end{pmatrix}, \quad i = 1, \dots, r. \quad (18)$$

Here  $H_i^{\beta}$  is of dimension  $p \times n_i^{\beta}$  and  $H_i^{\nu}$  is of dimension  $p \times n_i^{\nu}$ , where  $n_i = n_i^{\beta} + n_i^{\nu}$ .

The first block of  $n_i^{\beta}$  columns in  $H_i$  places restrictions on  $\beta_i$  and possibly (but not necessarily) joint restrictions on  $\beta_i$  and  $\nu_i$ . The second block of  $n_i^{\nu}$  columns places restrictions on  $\nu_i$  only. We allow  $n_i, n_i^{\beta}, n_i^{\nu} \geq 0$ ; when  $n_i^{\beta} = 0$  the first block of columns of  $H_i$  in (18) is absent, and similarly when  $n_i^{\nu} = 0$  the second block of columns in (18) is missing; finally  $n_i = 0$  corresponds to the case of known  $\zeta_i$  and no parameter in  $\zeta_i$  to be estimated. When  $n_i^{\beta} = 0$  or  $n_i = 0$ , we assume that the normalization vector  $h_i$  has non-zero entries corresponding to entries in  $\beta_i$ , i.e.  $h_i^{\beta} \neq 0$ , in order not to contradict the fact that  $\beta$  has full column rank  $r$ .

We also note that the diagonal blocks  $H_i^{\beta}$  and  $H_i^{\nu}$  in (18) must be of full column rank, in order for  $H_i$  to be of full column rank. Therefore one has that  $R_i$  has the form

$$R_i = \begin{pmatrix} R_i^{\beta} & -\bar{H}_i^{\beta} H_i^{\beta\nu'} R_i^{\nu} \\ 0 & R_i^{\nu} \end{pmatrix}$$

where  $R_i^x := H_{i\perp}^x$ ,  $x = \beta, \nu$ , see Lemma 7(ii) in Appendix A.4.

Next consider the proportional control relations; the restrictions on  $\zeta_i$  in  $\zeta^2$  of the type (17), with  $i = r+1, \dots, r+s$  must have  $h_i$  and  $H_i$  in the column space of  $J_2 := (0_{p,p} : I_p)'$  i.e.

$$h_i = \begin{pmatrix} 0_{p,1} \\ h_i^{\gamma} \end{pmatrix}, \quad H_i = \begin{pmatrix} 0_{p,n_i} \\ H_i^{\gamma} \end{pmatrix}, \quad i = r+1, \dots, r+s. \quad (19)$$

Observe that the  $2p \times (2p - n_i)$  matrix  $R_i$  equals  $R_i = \text{blkdiag}(I_p, R_i^\gamma)$  where  $R_i^\gamma := H_{i\perp}^\gamma$ . When  $n_i = 0$ , we assume  $h_i^\gamma \neq 0$ , in order not to contradict the fact that  $\gamma$  has full column rank  $s$ .

Stacking the implicit form in (17) for  $i = 1, \dots, r + s$ , we observe that one obtains restrictions of the form (16) with

$$R_{\dagger} = \begin{pmatrix} R_{11} & 0 \\ 0 & R_{22} \end{pmatrix} = \begin{pmatrix} \text{blkdiag}(R_1, \dots, R_r) & 0 \\ 0 & \text{blkdiag}(R_{r+1}, \dots, R_{r+s}) \end{pmatrix}, \quad (20)$$

with corresponding  $H_{\dagger} = \text{blkdiag}(H_1, \dots, H_r, H_{r+1}, \dots, H_{r+s})$ .

Using the relations  $H_{\zeta} = \bar{B}' H_{\dagger}$ ,  $h_{\zeta} = \bar{B}' h_{\dagger}$  and the form of  $B$  given in (58) in Appendix A.2, one finds the following expressions for  $H_{\zeta}$ ,  $h_{\zeta}$  in terms of the blocks of the  $H_i$  matrices in (18):

$$H_{\zeta} = \begin{pmatrix} \text{blkdiag}((H_1^{\beta v} : H_1^v), \dots, (H_r^v : H_r^{\beta v})) & 0 \\ 0 & \text{blkdiag}(H_{r+1}^\gamma, \dots, H_{r+s}^\gamma) \\ \text{blkdiag}((H_1^\beta : 0), \dots, (H_r^\beta : 0)) & 0 \end{pmatrix}, \quad (21)$$

$$h_{\zeta} = (h_1^{v'} : \dots : h_r^{v'} : h_{r+1}^{\gamma'} : \dots : h_{r+s}^{\gamma'} : h_1^{\beta'} : \dots : h_r^{\beta'})'.$$

We are now in the position to discuss identification.

**3.4. Rank and order conditions.** We first state rank and order conditions for general linear restrictions in Theorem 4, and subsequently for the case of equation-by-equation restrictions in Theorem 5. The proofs of this subsection are based on techniques similar to the ones employed in the classical treatment of identification in Sargan (1988), and they are reported in Appendix A.3.

**Theorem 4 (Identification, general linear case).** *A necessary and sufficient condition (rank condition) for the restrictions (13) or (16) to identify  $\eta$  and  $\zeta$  in (9) is given by*

$$\text{rank}(R_{\dagger}' C_{\dagger}) = r^2 + (r + s)^2. \quad (22)$$

where  $\text{rank}(R_{\dagger}' C_{\dagger}) = \text{rank}(R_{\zeta}' C_{\zeta})$  with

$$C_{\dagger} := \text{blkdiag}(I_r \otimes \zeta, I_s \otimes \zeta^{23}) \quad (23)$$

$$C_{\zeta} := ((L_1 \otimes (v : \gamma : \beta) + L_3 \otimes (\beta L_1')) : (L_2 \otimes (\gamma : \beta))), \quad (24)$$

with  $(L_1 : L_2 : L_3) := I_{2r+s}$ , where  $L_1, L_3$  have  $r$  columns and  $L_2$  has  $s$  columns.

*A necessary but not sufficient condition (order condition) for (22) is that the number of columns  $m_{\zeta}$  of  $R_{\zeta}$ , i.e. the number of restrictions on  $\theta_{\zeta}$ , satisfies*

$$m_{\zeta} \geq r^2 + (r + s)^2 \quad (25)$$

where we note that  $m_{\dagger} \geq m_{\zeta}$ , so that if (25) is satisfied, then also  $m_{\dagger} \geq r^2 + (r + s)^2$ .

Recall that  $m_{\dagger} = f - n_{\dagger} = p(2r + s) - n_{\dagger}$  so that eq. (25) can be equivalently restated as a condition on the number of free parameters in  $\phi_{\zeta}$ , namely

$$n_{\zeta} \leq p(2r + s) - r^2 - (r + s)^2. \quad (26)$$

We next state the rank and order conditions for the case of equation-by-equation restrictions.

**Theorem 5 (Identification, equation-by-equation restrictions).** *Let the restrictions be given as in (20) where  $R_i$  are specified as in (18)-(19); then a necessary and sufficient condition (rank condition) for these restrictions to identify  $\eta$  and  $\zeta$  in (9) is*

$$\text{rank}(R'_{11}(I_r \otimes \zeta)) = r(2r + s) \quad \text{and} \quad \text{rank}(R'_{22}(I_s \otimes \zeta^{23})) = s(r + s). \quad (27)$$

*One can have identification of single equations (columns) in  $\zeta$ ; more precisely the  $i$ -th column of  $\zeta$ ,  $1 \leq i \leq r$  (and in  $\zeta^1 := (\beta' : v)'$ ) is identified if and only if*

$$\text{rank}(R'_i \zeta) = 2r + s, \quad 1 \leq i \leq r. \quad (28)$$

*Similarly the  $i$ -th column in  $\zeta$ ,  $r + 1 \leq i \leq r + s$  (i.e. column number  $i - r$  in  $\zeta^2 := (0 : \gamma)'$ ) is identified if and only if:*

$$\text{rank}(R_i^{\gamma'}(\gamma : \beta)) = r + s, \quad r + 1 \leq i \leq r + s. \quad (29)$$

*The rank conditions (28) for  $1 \leq i \leq r$  and (29) for  $r + 1 \leq i \leq r + s$  are equivalent to (27).*

*A necessary but not sufficient condition (order condition) for (28) is*

$$2p - n_i \geq 2r + s, \quad 1 \leq i \leq r. \quad (30)$$

*Similarly, a necessary but not sufficient condition (order condition) for (29) is*

$$p - n_i \geq r + s, \quad r + 1 \leq i \leq r + s. \quad (31)$$

*The number of over-identifying restrictions in (18) (if the rank conditions (28) is satisfied) is  $2(p - r) - n_i - s$  for  $1 \leq i \leq r$ , and it is equal to  $p - n_i - r - s$  in (19) (if the rank conditions (29) is satisfied). In case of the extended star specification (7),  $p$  in (30), (31) and in the following lines should be replaced by  $p + q$ .*

In practice conditions (28) and (29) can be controlled prior to estimation by generating random numbers for the parameters and substituting them into (28) and (29), as suggested by Boswijk and Doornik (2004). Alternatively one could modify the generic identification approach of Johansen (1995a), Theorem 3, to obtain conditions that do not depend on specific parameter values.

**3.5. Rank of the Jacobian.** It is well known that the regression model (1) with  $\Phi$  unrestricted is globally identified. The I(2) model is a sub-model of it, corresponding to the specification (9) of  $\Phi$ . Hence an equivalent approach to the identification problem is the one based on the Jacobian, see e.g. Rothenberg (1971, Theorem 6). This approach has been applied to the I(1) model by Boswijk and Doornik (2004), and we state it here for the I(2) model.

The I(2) model parameters that appear in  $\Phi$  are  $\theta_\eta := \text{vec}(\alpha : \xi : \varsigma)$ , and  $\theta_\zeta := \text{vec}(v : \gamma : \beta)$ , both of dimension  $f \times 1$ . The restrictions (13) are here supplemented with the following linear restrictions on  $\theta_\eta$

$$\theta_\eta = h_\eta + H_\eta \phi_\eta. \quad (32)$$

Let now  $\theta_\Phi = (\theta'_\eta : \theta'_\zeta)'$ ,  $\phi_\Phi = (\phi'_\eta : \phi'_\zeta)'$ . The number of identified parameters in the model is associated with the rank of the Jacobian  $\mathcal{J}_\Phi(\phi_\Phi) := \partial\Phi^v/\partial\phi'_\Phi$  of the transformation  $\Phi(\phi_\Phi)$ , where  $\Phi^v := \text{vec } \Phi$ . In the present case one finds

$$\begin{aligned} \mathcal{J}_\Phi(\phi_\Phi) &= \begin{pmatrix} \frac{\partial\Phi^v}{\partial\theta'_\eta} H_\eta : \frac{\partial\Phi^v}{\partial\theta'_\zeta} H_\zeta \end{pmatrix}, \\ \frac{\partial\Phi^v}{\partial\theta'_\eta} &= \zeta \otimes I_p, \quad \frac{\partial\Phi^v}{\partial\theta'_\zeta} = (J_2 \otimes \alpha : J_2 \otimes \xi : J_1 \otimes \alpha + J_2 \otimes \varsigma) K. \end{aligned} \quad (33)$$

Here  $(J_1 : J_2) := I_{2p}$ , with  $J_1$  and  $J_2$  of dimensions  $2p \times p$  and  $K := \text{blkdiag}(\mathcal{K}_{pr}, \mathcal{K}_{ps}, \mathcal{K}_{pr})$ , where  $\mathcal{K}_{mn}$  is a commutation matrix, which satisfies  $\mathcal{K}_{mn} \text{vec}(A) = \text{vec}(A')$  when  $A$  is  $m \times n$ . In the following we consider both the case of unrestricted and restricted  $\theta_\zeta$  (respectively  $\theta_\eta$ ), where the former is indicated as  $H_\zeta = I$  (respectively  $H_\eta = I$ ), implicitly assuming in this case that  $h_\eta = h_\zeta = 0$ .

**Theorem 6 (Rank of the Jacobian).** (i) *When  $H_\eta = I$  and  $H_\zeta = I$  in (13) and (32) the Jacobian matrix  $\mathcal{J} := \mathcal{J}_\Phi(\phi_\Phi)$  has rank equal to*

$$\text{rank } \mathcal{J} = 2p(2r + s) - (2r^2 + s^2 + 2rs). \quad (34)$$

(ii) *Let  $H_\eta = I$  and  $H_\zeta$  be of full column rank, with  $n_\zeta$  columns (and hence  $R_\zeta$  is  $f \times m_\zeta$ , with  $m_\zeta = f - n_\zeta$ ), and assume that the rank condition in Theorem 4 is satisfied; then the Jacobian matrix  $\mathcal{J}$  has full column rank, equal to*

$$\text{rank } \mathcal{J} = 2p(2r + s) - m_\zeta = p(2r + s) + n_\zeta. \quad (35)$$

*In case of the extended star specification (7), we note that  $2p$  in (34) and (35) should be replaced by  $2p + q$ .*

The proof of this theorem is reported in Appendix A.4.

The approach based on the Jacobian is very versatile: one possible use is to check whether the system is identified, which is done by checking if the Jacobian (33) has full column rank.<sup>1</sup> However, it is possible to use it to count the number of restrictions implied by (13) and (32) also when they do not identify all the parameters (i.e. when the column rank of  $\mathcal{J}$  is not full). In fact, the number of restrictions is equal to  $2p(2r + s) - (2r^2 + s^2 + 2rs) - \text{rank}(\mathcal{J})$ , whether the system is identified or not. The latter use will be extensively illustrated in the application in Section 5. Notice however that, in the ‘equation by equation’ identification scheme, the approach based on the Jacobian does not override Theorem 5, which turns out to be useful to spot which equations are responsible for possible lack of identification.

---

<sup>1</sup>Also in this case this can be done prior to estimation by generating random numbers for the parameters and substituting them into (33).

**3.6. Identification strategy.** The applied I(2) literature is mainly focused on the analysis of the multicointegration vectors, i.e. the integral control relations. In line with this literature, in this subsection we suggest an identification strategy which first identifies the integral control relations, and then moves on to the identification of the proportional control relations, which is a complementary, but equally important, goal. Of course alternative specification strategies are possible, and – as always – the final selection may be to some extent path-dependent. We describe the identification strategy via the following 3 steps; these steps are later illustrated in the empirical section.

*3.6.1. Step 1: just-identifying the integral control relations.* As a first step, it seems natural to set up just-identifying restrictions for the multicointegration relationships, i.e. of the columns of  $\zeta^1$  in (12). As clarified above, see formula (30), the necessary order condition for just-identification of the  $i$ -th multicointegration relationship,  $i = 1, \dots, r$ , is that  $2r + s$  restrictions (including the restrictions possibly placed on the deterministic components and the normalization) are placed on the  $i$ -th column of  $\zeta^1$ .

When selecting these restrictions on  $\zeta_i^1$ , one has to keep in mind that: (i) at least one variable has to enter in levels, in order to be consistent with the choice that  $\beta$  has rank  $r$  and (ii) at least  $r + s$  out of the  $2r + s$  restrictions must be placed on  $v$ . The latter requirement is needed to prevent that, by combining the  $r + s$  columns of  $(\beta : \gamma)$  in the proportional control terms, one obtains a vector consistent with the restrictions placed on  $v$ : this would make the  $i$ -th multicointegration vector  $\zeta_i^1$  non-unique.

One advantage of the proposed parametrization is that  $v$  is not required to be orthogonal to  $\beta$  (or  $\tau$ ), so that one can set up  $2r + s$  identifying restrictions focusing exclusively on the variables (or linear combinations thereof) which economic theory suggests to be involved in (or excluded from) the relation of interest. Once each of the multicointegration relations fulfills the order condition, the rank condition can be checked for  $i = 1, \dots, r$  using (28) with  $\gamma$  unrestricted.

*3.6.2. Step 2: just-identifying the proportional control relations.* One can then just-identify the proportional control relations, noting that they are spanned by  $\beta$  and  $\gamma$ , where  $\beta$  has been identified in step 1. The order condition on  $\gamma$ , given in (31), is that at least  $r + s$  restrictions are placed on each CI vector, and one might follow the same approach illustrated above, placing exactly  $r + s$  restrictions on each column of  $\gamma$  (including the restrictions possibly placed on the deterministic components and the normalization) and then checking the rank conditions (28) for  $i = 1, \dots, r$  and (29) for  $i = r + 1, \dots, r + s$  hold, using the specification for  $\beta, v$  in step 1.

However, since  $\beta$  and  $v$  have been already identified in step 1, when setting up the restrictions for the  $i$ -th column of  $\gamma$ , one has to avoid choices which would be inconsistent with the specification of  $\beta$  or  $v$ , leading to violation of conditions (28) or (29). For example, if

$$\text{col}(h_j^\beta : H_j^\beta) \subseteq \text{col}(h_{r+i}^\gamma : H_{r+i}^\gamma)$$

for some  $j$ , then the  $j$ -th column of  $\beta$  would fulfill all the restrictions placed on the  $i$ -th column of  $\gamma$ , which would therefore be unidentified. Similarly one has to avoid to restrict the  $i$ -th column of  $\gamma$  in the same subspace as the  $j$ -th column of  $v$ .

As this example shows, in order to achieve identification, some separation among the columns of  $\gamma$ ,  $v$  and  $\beta$  is needed. Notice however that the proposed parametrization does not require (nor excludes) that  $\gamma$ ,  $v$  and  $\beta$  are mutually orthogonal. This leaves more flexibility when searching for theoretically grounded integral and proportional control specifications.

**3.6.3. Step 3: setting up over-identifying restrictions.** A system with just-identified CI relations devised as suggested in steps 1 and 2 can then be estimated, and asymptotic  $t$ -tests can be computed for the parameters, see Appendix A.5. In the light of both economic theory and empirical evidence, one can then proceed to formulate and test overidentifying restrictions, which are theoretically grounded and data-consistent.

#### 4. LIKELIHOOD ANALYSIS

This section describes likelihood-related statistical calculations on the reference model (6); the modifications needed for the extended model with deterministic components (7) are straightforward and they are omitted for readability. Let  $\theta$  be the parameter vector that contains all the model parameters  $\theta_\zeta$ ,  $\theta_\eta$ ,  $\Upsilon$  and  $\Omega$ . The Gaussian log-likelihood function is proportional to  $2\ell(\theta) = -T \ln |\Omega| - T \text{tr}(\Omega^{-1} M_{\varepsilon\varepsilon})$ , where  $M_{\varepsilon\varepsilon} := T^{-1} \sum_{t=1}^T \varepsilon_t \varepsilon_t'$ .

Since the log-likelihood score is bi-linear in the parameters  $\theta_\zeta$  and  $\theta_\eta$ , see (9), we present here an iterative maximization of  $\ell(\theta)$  that alternates the maximization with respect to  $\theta_\zeta$ ,  $\theta_\eta$  and  $\Omega$ . Obviously, one can employ alternative numerical maximization schemes of the quasi-Newton or Newton-Raphson types. The present alternating algorithm has the advantage, however, that each step is explicit and no line search is needed to guarantee an increase of the likelihood function. Moreover, the parameters  $\theta_\zeta$  are identified for fixed  $\theta_\eta$  and viceversa, which ‘makes switching algorithms particularly suited for partially identified models’, as stated by Boswijk and Doornik (2004) in the discussion of similar algorithms in the I(1) model.

For ease of exposition, in the remainder of this section we assume that also  $\Upsilon = 0$ ; the extension to  $\Upsilon \neq 0$  is obtained replacing  $\Delta^2 X_t$ ,  $\Delta X_{t-1}$ , or  $X_{t-1}$  with the residuals of the same variables regressed on  $\Delta^2 \mathbb{X}_{t-1}$ .

**4.1. An alternating maximization algorithm.** Unrestricted ML estimates may be obtained using the following algorithm, which is similar to the one proposed for the alternative parametrization in Johansen (1997)

- Adjustment Coefficients-step (AC-step): for fixed  $\theta_\zeta := \text{vec}(v : \gamma : \beta)$ , (6) is a linear regression model, and therefore maximization with respect to  $\theta_\eta := \text{vec}(\alpha : \xi : \varsigma)$  is obtained by OLS if there are no restrictions on  $\theta_\eta$ .
- Cointegrating Coefficients-step (CC-step): for fixed  $\theta_\eta$  and  $\Omega$ , maximization with respect to  $\theta_\zeta$  is obtained by GLS in a properly defined regression model, as shown e.g. in Paruolo

(2001). In fact, by transposing and vectorizing both sides of model (6) one obtains

$$\Delta^2 X_t = Z_t \theta_\zeta + \varepsilon_t, \quad (36)$$

where  $Z_t := Z_t(\theta_\eta, \Omega) = ((\alpha \otimes \Delta X'_{t-1}) : (\xi \otimes \Delta X'_{t-1}) : (\varsigma \otimes \Delta X'_{t-1} + \alpha \otimes X'_{t-1}))$ , maximization with respect to  $\theta_\zeta$  is obtained by GLS

$$\theta_\zeta(\theta_\eta, \Omega) = \left( \sum_{t=1}^T Z_t' \Omega^{-1} Z_t \right)^{-1} \sum_{t=1}^T Z_t' \Omega^{-1} \Delta^2 X_t.$$

- $\Omega$ -step: for fixed  $\theta_\zeta$ , and  $\theta_\eta$ , the log-likelihood is maximized with respect to  $\Omega$  by setting it equal to  $M_{\varepsilon\varepsilon}$  with  $\varepsilon_t$  computed for the current value of  $\theta_\zeta$  and  $\theta_\eta$ .

The AC, CC and  $\Omega$  steps are repeated until convergence. The order of the 3 steps is arbitrary; one may also consider an intermediate  $\Omega$  step between the AC and CC steps. Suitable initialization for  $\nu$ ,  $\gamma$  and  $\beta$  may be obtained using the two-stage I(2) estimator (2SI2), see Johansen (1995b), along the lines of Section 3 in Johansen (1995a).

In the following it is shown how linear restrictions that are separable between the parameters that appear in the AC and the CC step can be incorporated in this algorithm.

**4.2. Restrictions on the cointegrating coefficients.** The CC-step may be easily modified in order to derive restricted estimates under linear constraints of the type (13). In fact, substituting (13) into equation (36), one finds  $\Delta^2 X_t = Z_t(H_\zeta \phi_\zeta + h_\zeta) + \varepsilon_t$ , from which it follows directly that the CC step for  $\phi_\zeta$  are given by

$$\phi_\zeta(\theta_\eta, \Omega) = \left( \sum_{t=1}^T H_\zeta' Z_t' \Omega^{-1} Z_t H_\zeta \right)^{-1} \sum_{t=1}^T H_\zeta' Z_t' \Omega^{-1} (\Delta^2 X_t - Z_t h_\zeta)$$

and obviously  $\theta_\zeta(\theta_\eta, \Omega) = H_\zeta \phi_\zeta(\theta_\eta, \Omega) + h_\zeta$ .

**4.3. Restrictions on the adjustment coefficients.** The algorithm may be also modified to incorporate restrictions on the adjustment coefficients  $\theta_\eta$  of the type (32). For fixed  $\theta_\zeta$  and  $\Omega$ , restricted estimates of  $\theta_\eta$ , and an updated estimate of  $\Omega$  are obtained as follows. By vectorizing both sides of model (6) we obtain:

$$\Delta^2 X_t = W_t \theta_\eta + \varepsilon_t, \quad (37)$$

where  $W_t = ((\Delta X'_{t-1} \nu + X'_{t-1} \beta) \otimes I_p : \Delta X'_{t-1} \gamma \otimes I_p : \Delta' X_{t-1} \beta \otimes I_p)$ . Linear restrictions on  $\theta_\eta$  can be accommodated substituting (32) into (37) obtaining  $\Delta^2 X_t = W_t(H_\eta \phi_\eta + h_\eta) + \varepsilon_t$ . The likelihood function can be maximized with respect to  $\phi_\eta$  using GLS as follows

$$\phi_\eta(\theta_\zeta, \Omega) = \left( \sum_{t=1}^T H_\eta' W_t' \Omega^{-1} W_t H_\eta \right)^{-1} \sum_{t=1}^T H_\eta' W_t' \Omega^{-1} (\Delta^2 X_t - W_t h_\eta).$$

Obviously  $\theta_\eta(\theta_\zeta, \Omega) = H_\eta \phi_\eta(\theta_\zeta, \Omega) + h_\eta$ . It is easily shown that, in the absence of constraints on  $\theta_\eta$ , this procedure coincides with OLS.

Observe that the lack of identification illustrated in Section 3 is not critical for the functioning of this algorithm; in fact,  $\theta_\eta$  is identified when  $\theta_\zeta$  is fixed, and vice versa. Numerically, the lack

of identification has been found in our experience to be an advantage. The intuition behind this empirical evidence is that the unidentified parameters leave to the algorithm more freedom to find the steepest path when climbing the likelihood function.

A related issue concerns normalizations. In the application we have found that the algorithm is much faster if the normalization on  $\theta_\zeta$  is dropped, replacing  $H_\zeta$  with  $(h_\zeta : H_\zeta)$ , and adding one extra coefficient to  $\phi_\zeta$ . The normalization may be then performed algebraically on the estimated parameters. Similarly in the equation-by-equation approach to identification we suggest to remove the  $r + s$  normalizations to facilitate convergence, reintroducing them algebraically after estimation. Remark that removing the normalizations is possible when they are all placed either on  $\zeta$  or  $\eta$ , but this is not possible when they are placed on both.

**4.4. Asymptotics.** The asymptotic analysis of the I(2) VAR model has been presented in Johansen (1997, 2006), Boswijk (2000, 2010), Kurita, Nielsen, and Rahbek (2011). As emphasized there, not all inference is LAMN, and the asymptotic distribution for Wald and LR tests on some hypotheses on the CI parameters are still unknown.

Given that the focus of the present paper is on the identification of the I(2) VAR model, we decided to simply apply asymptotic results derived elsewhere to the present situation, and leave the full discussion of the asymptotic distribution of Wald and LR tests on all possible hypotheses to future research.

Because the asymptotic distribution for the Wald test of many hypotheses is  $\chi^2$  with degrees of freedom equal to the number of restrictions, see Johansen (2006), Boswijk (2010), Kurita, Nielsen, and Rahbek (2011) in the application we report results *as if* LAMN inference applied in all cases; see Appendix A.5 for details on the chosen calculations for standard errors of parameters.

## 5. AN APPLICATION TO US CONSUMPTION

In this section we illustrate how the proposed parametrization can be used to identify cointegrating relations in the I(2) model. The illustration is based on the US quarterly consumption data analyzed in Kurita, Nielsen, and Rahbek (2011), henceforth KNR. The 5-dimensional vector:

$$X_t = (c_t : y_t : w_t : R_t : p_t)'$$

is considered for the period 1964:2-2008:3;  $c_t$  is nominal private consumption,  $y_t$  is nominal disposable income after tax,  $w_t$  is nominal wealth including financial wealth and housing equity,  $R_t$  is the nominal annual bond yield, while  $p_t$  represents the price level measured as the consumption deflator. All variables except for  $R_t$  are in logs.<sup>2</sup>

Following KNR, we set  $k = 3$  and allow for one deterministic break in the trend slope in 1981:2,<sup>3</sup> reflecting the shift in policy focus following the stagflation period in the late 70's. We do not include the nine impulse dummies proposed in KNR. The analysis is therefore carried on within model (7)

---

<sup>2</sup> $R_t$  is divided by 4 to be comparable to a quarterly inflation rate,  $\Delta p_t$ . See KNR for details on the data.

<sup>3</sup>In the notation of the present paper  $T_1$ , the last observation of the first period, corresponds to 1981:1.

$r$	$-2 \log Q(H_{r,s}   H_p)$					
0	442.0 [.00]	342.4 [.00]	260.8 [.00]	202.4 [.00]	167.7 [.00]	154.6 [.00]
1	248.5 [.00]	180.1 [.00]	129.3 [.01]	104.0 [.02]	94.3 [.01]	
2	124.8 [.01]	85.4 [.10]	67.2 [.12]	58.8 [.05]		
3	54.3 [.27]	42.5 [.17]	31.5 [.17]			
4	22.1 [.27]	12.4 [.31]				
$s$	0	1	2	3	4	5

TABLE 1. Likelihood ratio tests for the cointegration ranks  $(r, s)$ . The numbers in brackets are tail probabilities derived from the Gamma approximation of the simulated distribution.

with  $q = 2$ . With the present parametrization and the algorithm described in Subsection 4.1, we reproduce exactly the LR tests computed in KNR, which are reported for completeness in Table 1.

The models corresponding to different choices of  $r$  and  $s$  are only partially nested: following the usual (row wise) testing sequence, and taking the usual 5% significance level, one would select  $r = 2$  and  $s = 1$ , which corresponds to  $p - r - s = 2$ , i.e. 2 I(2) common trends. KNR argued in favor of  $r = 2$  and  $s = 2$  instead, corresponding to a unique I(2) common trend, employing both economic rationale and further statistical analysis. In the following we adhere to their choice, in order to make results in present paper more comparable with KNR.

To simplify later exposition, the following default values for the design matrices  $H_i$  in (18), (19) are employed,

$$\begin{aligned}
H_i^\beta &= I_{p+q}, & h_i^\beta &= 0_{p+q,1} & i &= 1, \dots, r, \\
H_i^v &= I_{p+q}, & H_i^{\beta v} &= 0_{p+q,p+q}, & h_i^v &= 0_{p+q,1} & i &= 1, \dots, r, \\
H_i^\gamma &= I_{p+q}, & h_i^\gamma &= 0_{p+q,1} & i &= r + 1, \dots, r + s,
\end{aligned} \tag{38}$$

corresponding to unrestricted estimates of the parameters in  $\theta_\zeta$ . It is easily checked that (38) implies  $H_\zeta = I$  and  $h_\zeta = 0$ , so that the conditions of Theorem 6 hold, and the Jacobian (33) has dimension  $70 \times 72$  and rank  $52 = (2p + q)(2r + s) - (2r^2 + s^2 + 2rs)$ . This corresponds to 72 parameters, 52 of which are identified; one would hence need 20 (appropriate) additional restrictions for identification.

**5.1. Routine tests.** Within the present parametrization it is simple to perform some ‘routine’ tests on individual linear combinations of variables, say  $y_t = \omega' X_t$  with  $\omega$  known  $p$ -dimensional vector ( $\omega$  might be, for example, the  $i$ -th elementary vector). In the following we illustrate tests for trend stationarity, I(1)-ness, and weak exogeneity.

**Trend-stationarity.** Trend-stationarity of  $y_t$  implies that one column of  $\beta$  (say the first) is equal to  $\omega$ , and the corresponding vector of  $v$  is equal to zero; all of the other cointegration vectors

are unconstrained. This is specified by setting

$$h_1^\beta = (\omega' : 0_{1,q})', \quad H_1^\beta = (0_{q,p} : I_q)', \quad H_1^{\beta v} = 0_{p+q,q}, \quad H_1^v = (0_{q,p} : I_q)' \quad (39)$$

in (18), while all the other restriction matrices are set as in (38).

In the light of Theorem 5 the number of restrictions for (39) is  $2(p-r) - s$ ; in fact the matrix  $H_1$  has  $n_1 = 2q$  columns, and  $R_1 = H_{1\perp}$  fulfills condition (28) because there are no other restrictions; the first multicointegration vector is therefore over-identified. Hence, by Theorem 5, the number of overidentifying restrictions is equal to  $2(p+q-r) - n_1 - s$ . Since  $n_1 = 2q$ , the number of restrictions is therefore  $2(p-r) - s = 4$ .

Resorting to the Jacobian approach, one obtains the same result. In fact inserting the  $H_\zeta$  matrix corresponding to (39) into (33), see (21), one finds that the Jacobian has dimension  $70 \times 62$  and rank 48. Comparing this number with the rank of the Jacobian derived under (38), the difference is  $52 - 48 = 4$ , which gives that the number of restrictions is 4.

The limit distribution for this test cannot be found in the literature; here we act as if the test is  $\chi^2$  distributed, see Section 4.4.

**I(1)-ness.** We wish to test if the  $i$ -th variable is I(1); a linear combination  $y_t := \omega' X_t$  is I(1) if  $\omega$  is in  $\text{col } \tau$ . Because  $\text{col } \tau = \text{col}(\beta : \gamma)$  in the present parametrization, we control this hypothesis using a sequence of two tests: we first check that  $\omega$  is in  $\text{col } \beta$ , and if this is not the case, we test that  $\omega$  is in  $\text{col } \gamma$ . This procedure is justified as follows. Decompose the hypothesis  $\omega \in \text{col } \tau$  into the following disjoint and exclusive sub-hypotheses:

$$\begin{aligned} H_{01} : & \quad \omega \in \text{col } \beta \\ H_{02} : & \quad \omega \in \text{col } \tau, \quad \omega \notin \text{col } \beta. \end{aligned}$$

The first sub-hypothesis  $H_{01}$  is an hypothesis on  $\beta$ , while by eq. (11), the second sub-hypothesis can be stated as an hypothesis on  $\gamma$  of the type  $\gamma = (\omega : \phi)$  with  $\phi$  unrestricted. This is obtained thanks the possibility to replace the first column of  $\gamma$  with any linear combination in  $\text{col } \tau = \text{col}(\beta : \gamma)$ .

It is interesting to ascertain which of the two types of I(1)-ness applies to  $y_t$ . In fact, if  $\omega$  is in  $\text{col } \beta$  (i.e. if  $H_{01}$  holds), then  $\omega' X_t$  could be a candidate linear combination, together with some linear combination  $v_1$  of  $\Delta X_t$ , for one multicointegration relation; if  $\omega$  is not in  $\text{col } \beta$  (i.e. if  $H_{02}$  holds), this possibility is excluded.

We next describe the specification of the hypotheses  $H_{01}$  and  $H_{02}$  in terms of the  $H_i^x$  matrices. The hypothesis  $\omega \in \text{col } \beta$  is similar to (39) but less restrictive, since it does not involve  $v$ ; in fact the restrictions are defined by

$$h_1^\beta = (\omega' : 0_{1,q})', \quad H_1^\beta = (0_{q,p} : I_q)', \quad H_1^{\beta v} = 0_{p+q,q} \quad (40)$$

in (18), while all the other restriction matrices are set as in (38). Because  $r$  constraints can be imposed on the first column of  $\beta$  for just-identification, the number of restrictions for (40) is  $p-r$ , in accordance with Theorem 5. This number of restrictions is also confirmed using the Jacobian, which has dimensions  $70 \times 67$  and rank 49, implying that the number of restrictions is  $52 - 49 = 3$ .

Variable	trend-stationarity	I(1)- $\beta$	I(1)- $\gamma$	weak-exogeneity
$c_t$	20.35 [.00]	12.77 [.01]	7.79 [.01]	45.96 [.00]
$y_t$	16.31 [.00]	12.44 [.01]	8.09 [.00]	80.69 [.00]
$w_t$	21.90 [.00]	18.47 [.00]	3.70 [.05]	46.20 [.00]
$R_t$	16.14 [.00]	16.13 [.00]	0.54 [.46]	81.32 [.00]
$p_t$	28.76 [.00]	14.25 [.00]	5.80 [.02]	15.47 [.02]
$c_t - p_t$	24.00 [.00]	19.71 [.00]	0.42 [.52]	55.52 [.00]
$y_t - p_t$	23.34 [.00]	19.41 [.00]	0.61 [.44]	84.41 [.00]
$w_t - p_t$	21.19 [.00]	10.32 [.02]	0.11 [.74]	44.95 [.00]
$c_t - y_t$	24.89 [.00]	15.24 [.00]	0.89 [.35]	91.55 [.00]

TABLE 2. Routine tests.  $p$ -values in square brackets, calculated from a  $\chi_g^2$  distribution.  $g$  equals 4 for trend-stationarity, 3 for I(1)- $\beta$ , 1 for I(1)- $\gamma$  and 6 for weak-exogeneity.

Johansen (2006, Proposition 7) proved that the LR test  $-2 \log LR$  is asymptotically  $\chi^2$  distributed, with degrees of freedom equal to the number of restrictions, given by  $p - r$ .

If the test (40) is rejected, one can consider the hypothesis  $H_{02}$ , which is specified choosing

$$h_{r+1}^\gamma = (\omega' : 0_{1,q})', \quad H_{r+1}^\gamma = (0_{q,p} : I_q)' \quad (41)$$

in (19), while all the other restriction matrices are set as in (38). In the light of Theorem 5, the number of restrictions for (41) is  $p - r - s$ . In fact the matrix  $H_{r+1}^\gamma$  has  $n_{r+1} = q$  columns, and  $R_{r+1}^\gamma = H_{r+1}^\perp$  fulfills condition (29) because there are no other restrictions on  $\gamma$  or  $\beta$ ; this implies that the first column of  $\gamma$  is therefore over-identified. Hence the number of overidentifying restrictions is equal to  $p + q - n_{r+1} - r - s$ ; because  $n_{r+1} = q$ , the number of restrictions is  $p - r - s = 1$ . This number is confirmed calculating the Jacobian which has dimension  $70 \times 67$  and rank 51, giving  $52 - 51 = 1$  restriction. Johansen (2006, Proposition 11) derived that, when  $\omega \notin \text{col } \beta$ , the LR test  $-2 \log LR$  of  $H_{02}$  is asymptotically  $\chi^2$  distributed, with degrees of freedom equal to the number of restrictions, given by  $p - r - s$ .

The hypothesis that  $y_t = \omega' X_t$  is I(1) is not rejected if either (40) or (41) are not rejected.

**Weak exogeneity.** As shown in Paruolo and Rahbek (1999),  $\omega' X_t$  is weakly exogenous for all cointegration parameters iff  $\omega \in \text{col}^\perp(\eta)$ , i.e.  $\text{col}(\eta) \subseteq \text{col}^\perp(\omega)$ , which may be written as  $\theta_\eta = H_\eta \varphi_\eta$ , see (32), with

$$H_\eta = I_{2r+s} \otimes \omega_\perp, \quad (42)$$

and  $H_\zeta = I_{(p+q)(2r+s)}$ . The number of restrictions for (42) is  $2r + s$ . In fact, the Jacobian computed using (33) has dimension  $70 \times 66$  and rank 46, which shows that there are  $52 - 46 = 6$  restrictions. The interpretation is that (42) leaves the need for 20 identifying restrictions unchanged, but reduces the number of parameters in  $\eta$  by  $2r + s$ . As shown in Johansen (2006, Theorem 4), inference in this case is asymptotically  $\chi^2$ .

Table 2 reports the results of the routine tests on each of the variables and some linear combinations of interest. As expected, trend stationarity is strongly rejected for all nominal and real variables, and for the log of the consumption to income ratio. Moreover, none of the  $\omega$ 's considered here seems to be in  $\text{col } \beta$ , which means that each multicointegration relation involves more than a single real variable in levels.

The nominal interest rate  $R_t$ , the real variables, and the log of the consumption to income ratio seem to be  $I(1)$  with the corresponding  $\omega$ 's in  $\text{col } \gamma$ : this might appear inconsistent, since  $\gamma$  is an  $s$ -dimensional matrix ( $s = 2$  here); however, in the light of the discussion in Section 3, this means that, starting from any basis of  $\text{col } (\gamma : \beta)$  it is possible to find two matrices  $Q_{22}$  and  $Q_{32}$  (with  $Q_{22}$  non singular) such that  $\omega$  is a column of  $\gamma^\circ = \gamma Q_{22} + \beta Q_{32}$ .

Remark that the results of the tests (40) and (41) for  $R_t$  imply that the nominal interest rate is  $I(1)$ , but cannot be reduced to  $I(0)$  by linearly combining it with first differences, and therefore the ex-post real interest rate must also be  $I(1)$  (this will also be tested for below). Finally, the weak exogeneity tests strongly reject the null for all tested  $\omega$ 's.

**5.2. Number of CI(2,2) relations.** One may be interested in finding if there exist  $I(0)$  linear combinations involving only the levels. The hypothesis that there are  $j$  CI(2,2) relations, i.e. that there are  $j$  columns of  $v$  equal to  $0_{p,1}$ , may be specified by choosing

$$H_i^v = (0_{q,p} : I_q)' \quad i = 1, \dots, j, \quad j \leq r \quad (43)$$

in (18), while all the other restriction matrices are set as in (38).

Notice that, when  $j \leq \max(2r + s - p, 0)$ , eq. (43) does not impose any restriction. In fact, in the light of the discussion in Section 3, given the  $p \times (2r + s)$  matrix  $(v : \gamma : \beta)$  with  $2r + s > p$ , it is always possible to find three matrices  $Q_{11}$ ,  $Q_{21}$  and  $Q_{31}$  (with  $Q_{11}$  non singular) such that  $v^\circ = vQ_{11} + \gamma Q_{21} + \beta Q_{31}$  has  $2r + s - p$  columns equal to zero. The corresponding columns in  $\beta^\circ = \beta Q_{11}$  are such that  $\beta^{\circ'} X_t$  is stationary without the help of  $\Delta X_t$ .

Conversely, when  $j > \max(2r + s - p, 0)$ , eq. (43) provides binding restrictions; using the rank of the Jacobian, it is always possible to find the number of restrictions corresponding to any value of  $j$ . In case  $j = r$ , it can be shown that the number of implied restrictions is equal to  $(p - r - s)r$ . In fact, because  $\beta$  and  $\gamma$  are not restricted in any way under (43), one can choose  $a$  and  $b$  equal to 0 in Theorem 2, so that  $v = 0$  if and only if  $\delta = 0$ , where the number of elements in  $\delta$  is equal to  $(p - r - s)r$ . In the present empirical application, when we set  $j = 2$ , one is in the case  $j = r$ , and the number of restrictions is hence  $(p - r - s)r = 2$ . Coherently with the discussion above, the rank of the Jacobian (33) is found to be 52 when  $j = 1$  and 50 when  $j = 2$ , indicating that there are no restrictions in the former case and 2 restrictions in the latter case.<sup>4</sup>

<sup>4</sup>When the model is estimated under (43) with  $j = 1$ , the likelihood is obviously exactly the same as in the unconstrained model, and the first column of  $v$  is set to zero. The coefficients of the identified CI(2,2) relation are found in the first column of  $\beta$ ; normalizing it on  $c_t$  gives

$$c_t - 1.01w_t = -2.09y_t + 1.61p_t - 22.09R_t + 0.03t - 0.022D_{2t} + 28.1 + 1.6\Delta D_{2t} \sim I(0)$$

Hypotheses	degrees of freedom $g$	test statistic	$p$ -value
(44)-(45)-(46)	6	39.57	$5.5 \cdot 10^{-7}$
(44)-(45)	4	17.82	0.0013
(44)	2	10.77	0.0046

TABLE 3. Tests for the absence of breaks.  $p$ -values are calculated from a  $\chi^2$  distribution with  $g$  degrees of freedom.

The limit distribution for this LR test for  $j = r$  can be derived using the results in Boswijk (2010), who proved that the LR test for the related hypothesis  $\delta = 0$  is  $\chi^2$  distributed. In fact, as shown above, one has  $v = 0$  if and only if  $\delta = 0$ , and the LR for these two hypotheses are the same. Hence the LR test for  $j = r$  is asymptotically  $\chi^2$  distributed with  $(p - r - s)r$  degrees of freedom.

When  $j = 2$  the LR test takes on the value 21.5, with a  $p$ -value equal to  $2.14 \cdot 10^{-5}$  using a  $\chi^2_2$  distribution. Hence first differences are definitely needed in order to achieve stationarity for at least one of the two relations in  $\beta$ .

**5.3. Testing for the absence of breaks.** In this subsection we discuss tests for the absence of breaks. The restriction matrices corresponding to absence of breaks are given by

$$H_i^\beta = (I_{p+1} : 0_{p+1, q-1})', \quad H_i^{\beta v} = 0_{p+q, p+1}, \quad i = 1, \dots, r, \quad (44)$$

$$H_i^\gamma = (I_{p+1} : 0_{p+1, q-1})', \quad i = r + 1, \dots, r + s \quad (45)$$

$$H_i^v = (I_{p+1} : 0_{p+1, q-1})', \quad i = 1, \dots, r \quad (46)$$

in (18) - (19), while all the other restriction matrices are set as in (38).

The number of degrees of freedom for the joint restriction (44)-(45)-(46) is  $(2r + s)(q - 1) = 6$ , since the same  $(q - 1)$  restrictions are placed on each of the columns of  $\beta$ ,  $v$  and  $\gamma$ . The Jacobian computed as in (33) under (44)-(45)-(46) is  $70 \times 66$  and has rank 46, which confirms that there are  $6 = 52 - 46$  restrictions.

It may be interesting to test less restrictive versions of the hypothesis of absence of breaks. Relaxing (46), i.e. replacing  $H_i^v$  with  $I_{p+q}$ , one allows for a break in the intercept in the multicointegration vectors, maintaining the exclusion of the break in the slope for the I(1) directions. The same argument exposed above shows that the number of degrees of freedom for the joint restriction (44)-(45) is  $(r + s)(q - 1) = 4$ . This is confirmed by the Jacobian, which is  $70 \times 68$  and it has rank 48, which gives  $4 = 52 - 48$  restrictions.

An even less restrictive version of the hypothesis corresponds to the restriction (44) only. This allows for breaks in the trend slope in the I(1) directions in general, but excludes it from the integral control ECM, where only a change in the intercept is allowed for. The number of degrees of freedom in this case is  $r(q - 1) = 2$ ; this is confirmed by Jacobian, whose rank is equal to 50.

---

which seems to suggest that the consumption to wealth ratio is negatively related to real income and the interest rate, and grows over time (somewhat more slowly after the break).

Corollary 3.1 in KNR proves that the LR (and Wald tests) of (44)-(45) and of (44) alone are asymptotically  $\chi^2$  with degrees of freedom equal to the number of restrictions. The corresponding result for the joint restriction (44)-(45)-(46) has not been discussed in the literature, and we act as if the LR test is asymptotically  $\chi^2$ . The empirical results are reported in Table 3, which shows that, whichever version of the test for the absence of breaks one considers, one obtains a rejection and a strong support for the presence of a break.<sup>5</sup>

**5.4. Nominal to real transformation.** In the case  $r = s = 2$ , it is natural to test if the unique I(2) common trend is a nominal trend, shared by all nominal variables (i.e.  $c_t$ ,  $y_t$ ,  $w_t$  and  $p_t$ ). This is a test of the ‘nominal-to-real transformation’, see Kongsted (2005). If this is the case, coherently with the findings of the routine tests, the real variables  $c_t - p_t$ ,  $y_t - p_t$  and  $w_t - p_t$  should be I(1), as well as the nominal interest rate  $R_t$ . This specifies completely  $\tau$ , where  $\text{col}(\tau) = \text{col}(\beta : \gamma)$ , as  $H := (e_1 - e_5 : e_2 - e_5 : e_3 - e_5 : e_4)$ , where  $e_i$  is the  $i$ -th column in  $I_5$ . This can be tested by taking

$$H_i^\beta = \text{blkdiag}(H, I_q), \quad H_i^{\beta v} = 0_{p+q, r+s+q} \quad i = 1, \dots, r \quad (47)$$

$$H_i^\gamma = \text{blkdiag}(H, I_q) \quad i = r + 1, \dots, r + s \quad (48)$$

in (18), while all the other restriction matrices are set as in (38). The total number of restrictions in (47)-(48) is given by  $(p - r - s)(r + s) = 4$ .<sup>6</sup> In fact, the restriction (47) reduces the number of parameters in  $\beta$  by  $(p - r - s)r$ , while the restriction (48) reduces the number of parameters in  $\gamma$  by  $(p - r - s)s$ ; these restrictions leave the number of restrictions needed for just-identification unchanged. The Jacobian computed as in (33) is  $70 \times 68$  and has rank 48, which confirms that there are  $4 = 52 - 48$  restrictions.

Johansen (2006, Proposition 10) showed that the LR test for hypothesis like  $\tau = H$  is asymptotically  $\chi^2$  distributed with degrees of freedom equal to the number of restrictions. As in KNR, the LR test for (47)-(48) is equal to 5.27 with a  $p$ -value of 0.26 in the asymptotic  $\chi_4^2$  distribution, and therefore the hypothesis is not rejected.

It is usually argued in the literature that when  $\tau = H$  is known, analyzing the original variables within the I(2) model would be equivalent to analyzing the transformed vector  $W_t = (X_t' H : \Delta X_t' H_\perp)'$  using the I(1) model. In this case one can take  $\tau_\perp = H_\perp = \frac{1}{4}(1 : 1 : 1 : 1 : 0)'$ , so that  $\pi_t := \tau_\perp' \Delta X_t = \frac{1}{4}(\Delta c_t + \Delta y_t + \Delta w_t + \Delta p_t)$  can be interpreted as an average inflation rate, and  $W_t$  is equal to  $(c_t - p_t : y_t - p_t : w_t - p_t : R_t : \pi_t)'$ .

We remark here that this statement is correct in the absence of deterministic components, but it does not extend to the case when linear deterministic terms are included. In fact, the I(1) model for  $W_t$  with trend coefficients constrained in  $\text{col}(\alpha)$  and unconstrained intercept, does not correspond to the I(2) model discussed here, since it would imply that  $\pi_t$  has a linear trend, so that

<sup>5</sup>KNR perform only the test for (44)-(45), and find the same test statistic reported here. It may be interesting to remark that (45) implies zero restrictions unless (44) is also placed jointly; similarly (46) implies zero restrictions unless one considers it jointly with both (44) and (45). Both statements are supported by the computations of the ranks of the appropriate Jacobians.

<sup>6</sup>Notice that (48) alone would not imply any restriction.

$\sum_{i=1}^t \pi_i = \frac{1}{4}(c_t + y_t + w_t + p_t)$  has a quadratic trend, which is excluded in the I(2) model discussed here. This is why we do not switch to the I(1) model for identifying the cointegrating vectors even if the nominal to real transformation is not rejected.

**5.5. Identified relations.** We illustrate here the identification strategy illustrated in Section 3.6. The first step is to set up just-identifying restrictions on the multicointegration vectors. Actually, we slightly modify this first step: in fact, since the nominal to real transformation is clearly accepted, in the following we maintain the corresponding (overidentifying) restriction. In practice, in the light of the analysis illustrated above, with no loss of generality we rotate the system as follows

$$Y_t = (c_t - p_t : y_t - p_t : w_t - p_t : R_t : p_t)' := (c_t^* : y_t^* : w_t^* : R_t : p_t)'$$

and maintain the restriction that the last row of  $\tau$  (and therefore of  $\beta$  and  $\gamma$ ) is zero. Notice that this does not modify the number of restrictions to be placed on each cointegration vector in order to achieve identification: the number of parameters in each multicointegration vector is  $14 = 2(p + q)$ , we set to zero the coefficient of  $p_t$  in both vectors, which leaves 13 parameters, and we still need  $6 = 2r + s$  restrictions to achieve just-identification (including the restrictions possibly placed on the deterministic components and the normalization).

Recall that at least  $4 = r + s$  of such restrictions have to be placed on the parameters in  $v$ , which otherwise would not be identified. We propose the following just-identifying scheme as a starting point:

$$c_t^* = a_1 y_t^* + a_2 w_t^* + a_3 t + a_4 D_{2t} + a_5 \Delta p_t + a_6 + a_7 \Delta D_{2t} + u_{1t} \quad (49)$$

$$R_t = b_1 c_t^* + b_2 y_t^* + b_3 t + b_4 D_{2t} + b_5 \Delta p_t + b_6 + b_7 \Delta D_{2t} + u_{2t} \quad (50)$$

We denote (49) more compactly as  $z_{1t} = f_{1t} + u_{1t}$  and (50) as  $z_{2t} = f_{2t} + u_{2t}$ , with  $z_{1t} := c_t^* - a_1 y_t^* - a_2 w_t^* - a_5 \Delta p_t$ ,  $f_{1t} := a_3 t + a_4 D_{2t} + a_6 + a_7 \Delta D_{2t}$ ,  $z_{2t} := R_t - b_1 c_t^* - b_2 y_t^* - b_5 \Delta p_t$ ,  $f_{2t} := b_3 t + b_4 D_{2t} + b_6 + b_7 \Delta D_{2t}$ . Here  $f_{it}$  denotes the deterministic component and  $u_{it}$  the stationary error in equation  $i$ ,  $i = 1, 2$ .

The first relation corresponds to a consumption function for real consumption in terms of real income and real wealth, with inflation reflecting a deviation from price homogeneity in the short run. Out of the 6 identification restrictions, 2 involve the levels (the normalization placed on the parameter of  $c_t^*$ , and the exclusion of  $R_t$ ), and 4 involve the first differences (the exclusion of  $\Delta c_t^*$ ,  $\Delta y_t^*$ ,  $\Delta w_t^*$  and  $\Delta R_t$ : according to the preliminary evidence these are stationary variables, and therefore they are the natural candidates to be excluded from the multicointegration vectors); no restrictions are placed on the deterministic components.

The second multicointegration vector is identified as a relation for the nominal interest rate, which is related to the inflation rate, real consumption and real income. Also in this case, out of the 6 identification restrictions 2 involve the levels (the normalization placed on the parameter of  $R_t$ , and the exclusion of  $w_t^*$ ), and 4 involve the first differences (the exclusion of  $\Delta c_t^*$ ,  $\Delta y_t^*$ ,  $\Delta w_t^*$  and  $\Delta R_t$ ). Notice that the restrictions on the first differences are the same as in the first equation, but this is not a problem for identification, as far as no linear combination of the columns of  $\beta$

LHS var	Coeff	RHS var	Just Identified Model			Over Identified Model			
			Estimate	Std.Err.	t-stat	Estimate	Std.Err.	t-stat	
eq. (49)	$a_1$	$y_t^*$	0.638	0.065	9.82	0.656	0.039	16.82	
	$a_2$	$w_t^*$	0.088	0.032	2.75	0.118	0.019	6.21	
	$a_3$	$t$	0.0031	0.0005	6.20	0.0028	0.00035	8.00	
	$c_t^*$	$D_{2t}$	-0.0008	0.0003	-2.67	-0.0008	0.00012	-6.67	
	$a_5$	$\Delta p_t$	-3.03	0.656	-4.62	-2.98	0.56	-5.32	
	$a_6$	1	3.64	0.802	4.54	2.88	0.73	3.94	
	$a_7$	$\Delta D_{2t}$	0.036	0.019	1.89				
eq. (50)	$b_1$	$c_t^*$	0.474	0.186	2.55	0.246	0.044	5.59	
	$b_2$	$y_t^*$	-0.405	0.139	-2.91	-0.246	0.044	5.59	
	$b_3$	$t$	-0.0007	0.0006	-1.17				
	$R_t$	$D_{2t}$	-0.0003	0.0001	-3.33	-0.0004	0.00005	-8.00	
	$b_5$	$\Delta p_t$	1.66	0.498	3.00	1			
	$b_6$	1	-0.889	0.783	-1.14	0.070	0.012	5.83	
	$b_7$	$\Delta D_{2t}$	0.053	0.009	5.89	0.047	0.0044	10.68	
eq. (51)	$\Delta y_t^*$	$c_1$	1	0.007	0.001	7.00	0.007	0.0007	10.00
		$c_2$	$\Delta D_{2t}$	0.0009	0.0016	0.56			
eq. (52)	$\Delta (c_t^* - y_t^*)$	$d_1$	1	0.0009	0.0003	3.00			
		$d_2$	$\Delta D_{2t}$	-0.0001	0.0005	-0.20	0.00079	0.00030	2.63

TABLE 4. Unrestricted (just-identified) and restricted (over-identified) estimates.

and  $\gamma$  is restricted in the same way. Also for the second equation no restrictions are placed on the deterministic components.

There are some theoretically relevant hypotheses, which are expected to hold on the interest rate relation: the coefficient of  $\Delta p_t$  is expected to be 1 ( $b_5 = 1$  would imply that the equation may be interpreted as a relation for the real interest rate). Moreover, the coefficient of  $y_t^*$  is expected to be equal to the coefficient of  $c_t^*$  with opposite sign ( $b_2 = -b_1$  would imply that the interest rate depends on the consumption income ratio, which may be seen as a measure of the business cycle). However these and other overidentifying restrictions possibly suggested by the data will be analyzed later.

The main difference of the present specification with KNR lies with the presence of  $\Delta p_t$  as the single variable in first differences in the integral control term in place of the average inflation term  $\pi_t := \tau'_\perp X_t = \frac{1}{4}(\Delta c_t + \Delta y_t + \Delta w_t + \Delta p_t)$  that appears in KNR. The present specification also relaxes all the over-identifying restrictions placed in KNR, except for the price homogeneity assumption.

The second step is to set up just identifying restrictions on the proportional control relations. From Theorem 5, the order condition, is that at least  $4 = r + s$  restrictions are placed on the  $7 = p + q$  coefficients of each cointegration vector. Recall that we maintain the nominal to real

(price homogeneity) assumption, and therefore the row of  $\gamma$  corresponding to  $\Delta p_t$  is set to zero; this is an additional overidentified restriction which does not contribute to identification since it holds for all columns of  $\beta$  and  $\gamma$ . Therefore, the number of free parameters in each equation has to be 2. Coherently with the findings of the routine tests illustrated above, we propose the following just-identifying scheme:

$$\Delta y_t^* = c_1 + c_2 \Delta D_{2t} + u_{3t} \quad (51)$$

$$\Delta(c_t^* - y_t^*) = d_1 + d_2 \Delta D_{2t} + u_{4t} \quad (52)$$

We denote (51) more compactly as  $z_{3t} = f_{3t} + u_{3t}$  and (52) as  $z_{4t} = f_{4t} + u_{4t}$ , with  $z_{3t} := \Delta y_t^*$ ,  $f_{3t} := c_1 + c_2 \Delta D_{2t}$ ,  $z_{4t} := \Delta(c_t^* - y_t^*)$ ,  $f_{4t} := d_1 + d_2 \Delta D_{2t}$ . The first equation implies that the real income is I(1), with a (possibly broken) drift (i.e. the first difference is stationary with a possibly changing mean): the purpose of the statistical analysis is therefore to analyze whether there is a significant change in the growth rate of real income after the break. The 4 identifying restrictions are given by the normalization placed on the parameter of  $\Delta y_t^*$ , and the exclusion of  $\Delta c_t^*$ ,  $\Delta w_t^*$  and  $\Delta R_t$ . The second equation is very similar, and aimed at analyzing the trend behavior of the I(1) consumption to income ratio.

The proposed identifying scheme with respect to  $Y_t$  corresponds to

$$\beta^* = \begin{pmatrix} -1 & b_1 \\ a_1 & b_2 \\ a_2 & 0 \\ 0 & -1 \\ 0 & 0 \\ a_3 & b_3 \\ a_4 & b_4 \end{pmatrix}, \quad v^* = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ a_5 & b_5 \\ a_6 & b_6 \\ a_7 & b_7 \end{pmatrix}, \quad \gamma^* = \begin{pmatrix} 0 & -1 \\ -1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ c_1 & d_1 \\ c_2 & d_2 \end{pmatrix}.$$

The restriction matrices are therefore specified as follows:

$$\begin{aligned} h_1^\beta &= -e_1, & H_1^\beta &= (e_2 : e_3 : e_6 : e_7), \\ h_1^v &= 0_{7,1}, & H_1^{\beta v} &= 0_{7,4}, & H_1^v &= (e_5 : e_6 : e_7), \\ h_2^\beta &= -e_4, & H_2^\beta &= (e_1 : e_2 : e_6 : e_7), \\ h_1^v &= 0_{7,1}, & H_2^{\beta v} &= 0_{7,4}, & H_2^v &= (e_5 : e_6 : e_7), \\ h_1^\gamma &= -e_2, & H_1^\gamma &= (e_6 : e_7), & h_2^\gamma &= -e_1 + e_2, & H_2^\gamma &= (e_6 : e_7), \end{aligned} \quad (53)$$

where  $e_i$  is the  $i$ -th column of  $I_7$ .

To verify if the proposed structure identifies the system, following Theorem 5 one can check if the rank condition (27) holds; in this case the two ranks in (27) are  $12 = r(2r + s)$  and  $8 = s(r + s)$  respectively, which proves that the structure is identified.<sup>7</sup> An alternative and equivalent approach is to check if the Jacobian, computed as in (33), has full column rank. In our case, the Jacobian is a

<sup>7</sup>If condition (27) does not hold, one can check equation by equation the rank conditions (28) and (29) to detect which equation is not identified.

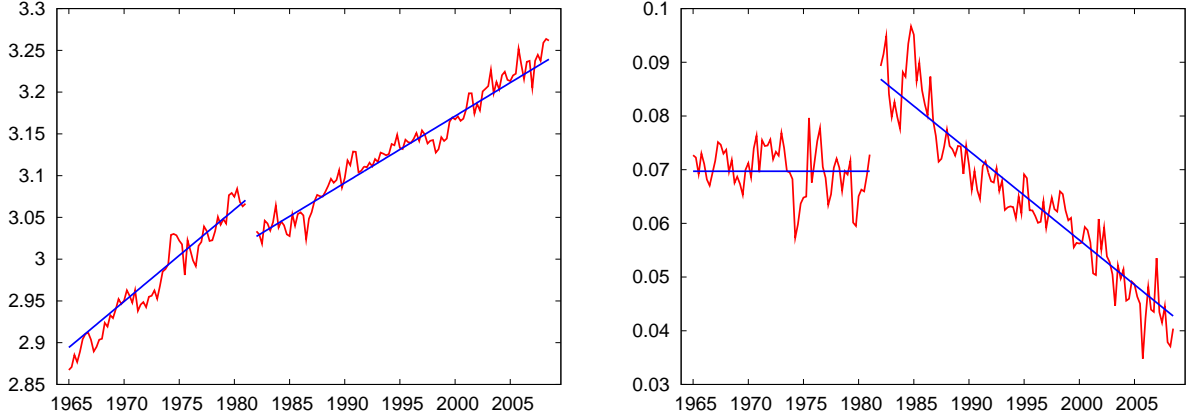


FIGURE 1. Plots of the integral control terms and their deterministic components. Parameters are taken from the over-identified panel in Table 4. Left panel:  $z_{1t}$  and  $f_{1t}$  in eq. (49), right panel:  $z_{2t}$  and  $f_{2t}$  in eq. (50).

$70 \times 48$  matrix and has rank 48, which confirms that the structure is identified. Notice that the rank is exactly the same as the rank of the Jacobian corresponding to (47)-(48); eq. (53) places some additional just-identifying restrictions on (47)-(48), without adding overidentifying restrictions.

The next step is to estimate the just identified system. Maximum likelihood estimates are given in the left side of Table 4, along with standard errors (computed as illustrated in Appendix A.5) and asymptotic  $t$ -tests. Since the system is just identified, the likelihood is exactly equal to the likelihood of the model estimated under (47)-(48) in Subsection 5.4.

Notice that some of the parameters of the deterministic components are not significant; moreover, the  $t$ -test for the hypothesis  $b_1 + b_2 = 0$  is equal to 1.23, while the  $t$ -test for the null hypothesis  $b_5 = 1$  is equal to 1.33. Therefore, we gradually introduce 6 overidentifying restrictions ( $a_7 = b_1 + b_2 = b_3 = c_2 = d_1 = 0$  and  $b_5 = 1$ ). The restriction  $d_1 = 0$  is not obvious from the unrestricted estimates, which would rather suggest  $d_2 = 0$  implying the same constant (around 0.0008) in both periods. Here we have decided to test  $d_1 = 0$  and restrict it accordingly, motivated by the plot of  $c_t^* - y_t^*$  (see Figure 2); the data do not reject the restriction.

The overidentified structure is described by the following restriction matrices:

$$\begin{aligned}
 h_1^\beta &= -e_1, & H_1^\beta &= (e_2 : e_3 : e_6 : e_7), \\
 h_1^v &= 0_{7,1}, & H_1^{\beta v} &= 0_{7,4}, & H_1^v &= (e_5 : e_6), \\
 h_2^\beta &= -e_4, & H_2^\beta &= (e_1 - e_2 : e_7), \\
 h_2^v &= e_5, & H_2^{\beta v} &= 0_{7,2}, & H_2^v &= (e_6 : e_7), \\
 h_1^\gamma &= -e_2, & H_1^\gamma &= e_6, & h_2^\gamma &= -e_1 + e_2, & H_2^\gamma &= e_7.
 \end{aligned} \tag{54}$$

The structure (54) identifies the system; the rank condition (27) holds, and the Jacobian in (33) has 42 columns and full column rank. ML estimates for the over-identified structure (54) are reported in the right panel of Table 4. The LR test for the 6 overidentifying restrictions is equal

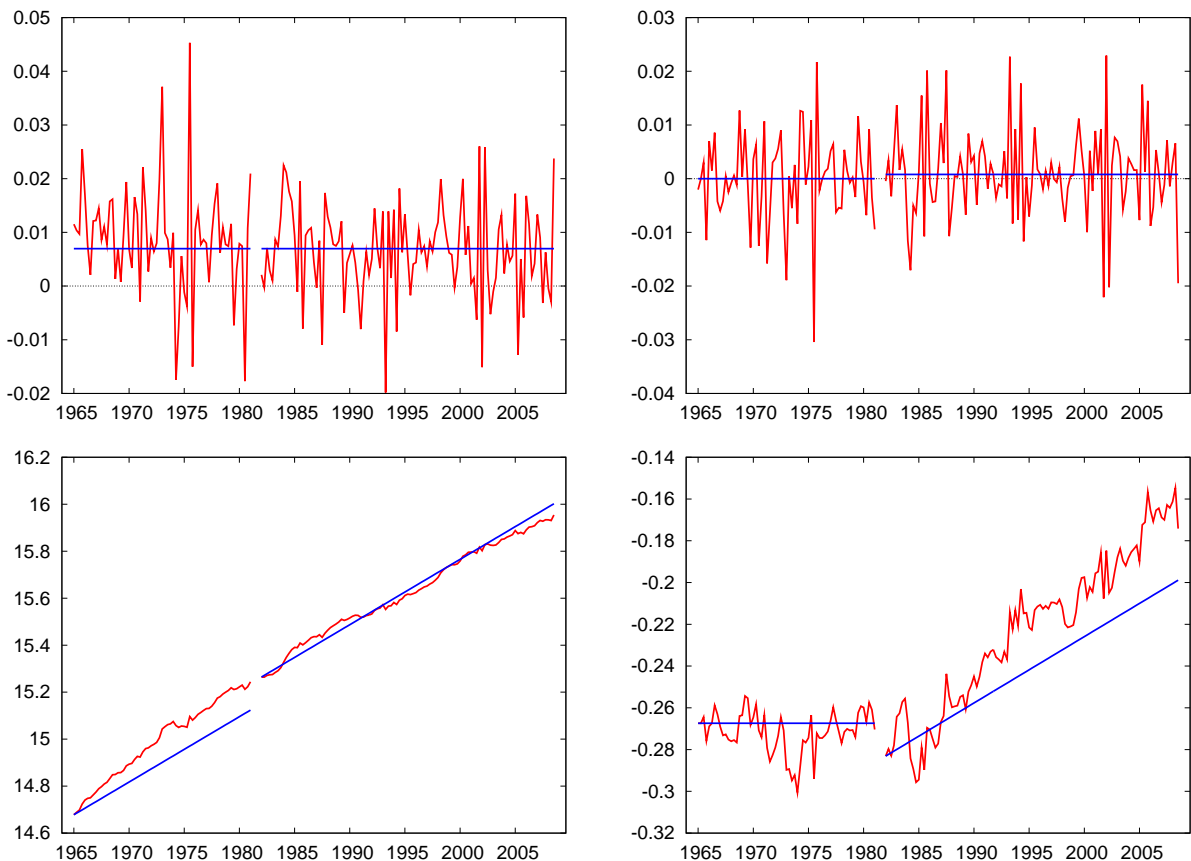


FIGURE 2. Plots of the proportional control terms and their deterministic components. Parameters are taken from the over-identified panel in Table 4. The upper-left panel plots  $z_{3t}$  and  $f_{3t}$  in eq. (51), the upper-right panel reports  $z_{4t}$  and  $f_{4t}$  in eq. (52); the lower-left panel plots the cumulation of  $z_{3t}$  and  $f_{3t}$  in eq. (51), the lower-right panel plots the cumulation of  $z_{4t}$  and  $f_{4t}$  in eq. (52). The initial values for the deterministic components  $f_{jt}$  in the cumulative plots are set equal to the first observation of the series in each sub-sample.

to 8.71, with a  $p$ -value of 0.19 using a  $\chi^2_6$  distribution, while the test for these restrictions plus the 4 restrictions giving price homogeneity is equal to 13.98, corresponding to a  $p$ -value of 0.17 using a  $\chi^2_{10}$  distribution. These limit distributions have not yet been derived in the literature, see Section 4.4. Unpublished preliminary Monte Carlo simulations show that departures from  $\chi^2$  inference appear minor.

Figure 1 graphs the integral control relations, while Figure 2 reports the proportional control relations based on the restricted estimates.

## 6. CONCLUSIONS

This paper introduces a new ECM representation of  $I(2)$  systems, which is used to discuss the identification, estimation and testing of cointegrating relations. The new parametrization allows

the econometrician to choose which variables to include in differences in the cointegrating relations. Because the cointegrating relations that appear in the integral and the proportional control terms of the ECM are inter-related, the identification of the cointegrating relations needs to be solved jointly.

We provide order and rank conditions for general linear hypotheses on the cointegrating vectors, and we specialize results for the case of equation-by-equation restrictions. We connect the order and rank conditions with the rank of the relevant Jacobian, which can be used to control identification. We adopt a likelihood-based approach to inference, and we present a switching algorithm for likelihood maximization based on the present parametrization. An application on US consumption illustrates the use of the proposed parametrization.

#### ACKNOWLEDGEMENTS

We wish to thank Andrea Colombo, who significantly contributed to an early version of this paper. Partial financial support from Italian MUR Grant Cofin2006-13-1140 (both authors), University of Insubria FAR funds (second author) is gratefully acknowledged.

#### REFERENCES

- Boswijk, P., 2000, Mixed Normality and Ancillarity in  $I(2)$  Systems. *Econometric Theory*, 16, 878–904.
- Boswijk, P., and J. A. Doornik, 2004, Identifying, estimating and testing restricted cointegrated systems: An overview. *Statistica Neerlandica*, 58, 440–465.
- Boswijk, P. H., 2010, Mixed normal inference on multicointegration. *Econometric Theory*, 26, 1565–1576.
- Engle, R., and C. Granger, 1987, Co-integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55, 251–276.
- Engle, R., and S. Yoo, 1991, Cointegrated economic time series: an overview with new results. in *Long-run Economic Relations: Readings in Cointegration*, Engle and Granger eds., Oxford University Press.
- Engsted, T., and S. Johansen, 1999, Granger’s representation theorem and multi-cointegration. in *Causality and Forecasting. Festschrift for Clive W. Granger.*, in R. F. Engle and H. White, eds. Oxford University Press.
- Granger, C. W. J., and T. H. Lee, 1989, Investigation of Production, Sales and Inventory Relationships Using Multicointegration and Non-Symmetric Error Correction Models. *Journal of Applied Econometrics*, 4, pp. S145–S159.
- Haldrup, N., and M. Salmon, 1998, Representation of  $I(2)$  cointegrated systems using the Smith-McMillan form. *Journal of Econometrics*, 44, 303–325.
- Johansen, S., 1992, A representation of vector autoregressive processes integrated of order 2. *Econometric Theory*, 8, 188–202.

- , 1995a, Identifying restrictions of linear equations with applications to simultaneous equations and cointegration. *Journal of Econometrics*, 69, 111–132.
- , 1995b, A Statistical Analysis of Cointegration for I(2) variables. *Econometric Theory*, 11, 25–59.
- , 1996, *Likelihood-based Inference in Cointegrated Vector Auto-Regressive Models*. Oxford University Press.
- , 1997, A Likelihood Analysis of the I(2) Model. *Scandinavian Journal of Statistics*, 24, 433–462.
- , 2006, Statistical analysis of hypotheses on the cointegrating relations in the I(2) model. *Journal of Econometrics*, 132, 81–115.
- 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.
- Kitamura, Y., 1995, Estimation of cointegrated systems with I(2) processes. *Econometric Theory*, 11, 1–24.
- Kongsted, H., 2005, Testing the Nominal-to-Real Transformation. *Journal of Econometrics*, 124, 205–225.
- Kurita, T., H. Nielsen, and A. C. Rahbek, 2011, An I(2) Cointegration Model with Piecewise Linear Trends: Likelihood Analysis and Application. *The Econometrics Journal*, 14, forthcoming.
- Magnus, J., and H. Neudecker, 2007, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 3rd ed. Wiley.
- Paruolo, P., 2000, Asymptotic efficiency of the two stage estimator in I(2) systems. *Econometric Theory*, 16, 524–550.
- , 2001, On Likelihood Maximizing Algorithms for I(2) VAR Models. Manuscript, Università dell’Insubria, Varese, Italy.
- , 2006, A likelihood ratio test for the rank of a cointegration submatrix. *Oxford Bulletin of Economics and Statistics*, 68, 921–948.
- Paruolo, P., and A. Rahbek, 1999, Weak exogeneity in I(2) VAR systems. *Journal of Econometrics*, 93(2), 281 – 308.
- Rahbek, A. C., 1997, Representation of Cointegrated I(2) and I(1) Processes with deterministic trends. Mimeo, University of Copenhagen.
- Rahbek, A. C., H. Kongsted, and C. Jørgensen, 1999, Trend-Stationarity in the I(2) Cointegration Model. *Journal of Econometrics*, 90, 265–89.
- Rothenberg, T. J., 1971, Identification in Parametric Models. *Econometrica*, 39, pp. 577–591.
- Sargan, J. D., 1988, *Lectures on Advanced Econometrics*. Basil Blackwell.
- Stock, J., and M. Watson, 1993, A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica*, 61, 783–820.
- Zhang, F., 2005, *The Schur Complement and Its Applications*. Springer.

**A.1. Representation.** *Proof* of Theorem 2. By (2) one has  $\gamma = \beta g'_0 + \beta_1 g'$ , with  $g'_0 := \bar{\beta}' \gamma$ ,  $g' := \bar{\beta}'_1 \gamma$  where the latter is of dimension  $s \times s$  and of full rank. In Paruolo and Rahbek (1999) eq. (2.10), the following decomposition of  $\Gamma$  was discussed

$$\Gamma = \alpha \delta \beta'_2 + \zeta_2 \beta'_1 + \zeta_1 \beta' = (\alpha : \zeta_2 : \zeta_1) (\beta_2 \delta' : \beta_1 : \beta)' = FG', \text{ say,}$$

where  $(\zeta_1 : \zeta_2) = \Gamma (\bar{\beta} : \bar{\beta}_1)$  and  $\delta = \bar{\alpha}' \Gamma \bar{\beta}_2$ . Inserting  $I = N^{-1}N$  between  $F$  and  $G'$  with

$$N := \begin{pmatrix} I_r & & \\ & g & g_0 \\ & & I_r \end{pmatrix}, \quad N^{-1} = \begin{pmatrix} I_r & & \\ & g^{-1} & -g^{-1}g_0 \\ & & I_r \end{pmatrix}$$

one finds  $\Gamma = (FN^{-1})(NG')$  with  $NG' = (\beta_2 \delta' : \gamma : \beta)'$  and

$$\begin{aligned} FN^{-1} &= (\alpha : \zeta_2 g^{-1} : \zeta_1 - \zeta_2 g^{-1} g_0) = (\alpha : \Gamma \beta_1 (\gamma' \beta_1)^{-1} : \Gamma (I - \beta_1 (\gamma' \beta_1)^{-1} \gamma') \bar{\beta}) \\ &=: (\alpha : \xi^\circ : \varsigma^\circ), \text{ say.} \end{aligned}$$

Adding and subtracting  $\alpha(a\beta' + b\gamma')$  one finds

$$\Gamma = (\alpha : \xi^\circ - \alpha b : \varsigma^\circ - \alpha a) (\beta_2 \delta' + \gamma b' + \beta a' : \gamma : \beta)' =: (\alpha : \xi : \varsigma) (v : \gamma : \beta)'.$$

This completes the proof.  $\square$

**A.2. Connection matrices.** Recall that  $(J_1 : J_2) = I_{2p}$ , where  $J_1$  and  $J_2$  contain  $p$  columns each, and that  $(L_1 : L_2 : L_3) := I_{2r+s}$  where  $L_1$  and  $L_3$  have  $r$  columns and  $L_2$  has  $s$  columns. Let also  $L_{ij} := (L_i : L_j)$ ,  $i, j = 1, 2, 3$ . and define  $(U_1 : U_2) := I_{r+s}$ , with  $U_1, U_2$  with  $r$  and  $s$  columns respectively. We consider also the analogue of (16) for  $\zeta$ :

$$\text{vec } \zeta = h + H\phi_\zeta, \quad \iff \quad R' \text{vec } \zeta = c. \quad (55)$$

From (14) one finds that the design matrices in (13), (16) and (55) are related in the following way

$$h_\dagger = Bh_\zeta, \quad H_\dagger = BH_\zeta, \quad R_\dagger = (B_\perp : \bar{B}R_\zeta), \quad c_\dagger = (0' : c'_\zeta)', \quad (56)$$

$$h = ABh_\zeta, \quad H = ABH_\zeta, \quad R = (A_\perp : \bar{A}B_\perp : \bar{A}\bar{B}R_\zeta), \quad c = (0' : 0' : c'_\zeta)'. \quad (57)$$

Here we note that  $A_\perp$  within  $R$  describes the restrictions that link  $\zeta^{12}$  and  $\zeta^3$ , and that  $B_\perp$  within  $R$  and  $R_\dagger$  describes the 0-restriction on the upper-right block of  $\zeta^{12}$ . Hence the relevant restrictions for identification are contained in  $R_\zeta$ .

In order to find the  $B$  matrix, write  $\zeta^{12} = J_2(v : \gamma) + J_1\beta U'_1$  and take vecs. One finds  $\text{vec } \zeta^{12} = B\theta_\zeta$  with

$$B := (I_{r+s} \otimes J_2 : U_1 \otimes J_1), \quad (58)$$

where  $B$  is a  $2p(r+s) \times f$  matrix with orthogonal columns.

In order to find the  $A$  matrix, observe that  $\text{vec } \zeta = (\text{vec } \zeta^{12'} : \text{vec } \zeta^3)'$ , and that  $\zeta^3 = J_2 J'_1 \zeta^{12} U_1$ . Taking vecs one finds  $\text{vec } \zeta^3 = (U'_1 \otimes J_2 J'_1) \text{vec } \zeta^{12}$  from which one finds  $\text{vec } \zeta = A \text{vec } \zeta^{12}$  with

$$A := \begin{pmatrix} I_{2p(r+s)} \\ U'_1 \otimes J_2 J'_1 \end{pmatrix} = (L_{12} \otimes I_{2p}) + (L_3 \otimes J_2)(U'_1 \otimes J'_1). \quad (59)$$

Using (58) and (59) one finds that  $\bar{A}$  in (57) is given by

$$\bar{A} = (L_{12} \otimes I_{2p}) + \frac{1}{2}(L_3 \otimes J_2 - L_1 \otimes J_1)(U'_1 \otimes J'_1), \quad (60)$$

and that  $\bar{A}\bar{B}$  in (57) is given by

$$\bar{A}\bar{B} = \left( L_{12} \otimes J_2 : \frac{1}{2}(L_1 \otimes J_1 + L_3 \otimes J_2) \right). \quad (61)$$

**A.3. Rank condition.** *Proof* of Theorem 4. We wish to show that, if  $\zeta$  and  $\zeta^\circ := \zeta Q$  both satisfy (13) or (16), then  $Q = I_{2r+s}$  iff (22) holds. Observe that (13), (16) and (55) are equivalent under (56), (57). Hence consider (55) for  $\zeta$  and  $\zeta^\circ$  in implicit form, i.e.  $R' \text{vec } \zeta = c$  and  $R' \text{vec } \zeta^\circ = c$ . Because  $A_\perp$  and  $\bar{A}B_\perp$  within  $R$ , see (57), describe the restrictions that link  $\zeta^{12}$  and  $\zeta^3$ , and the 0-restriction on the upper-right block of  $\zeta^{12}$ , we restrict attention to the last block of  $R$ , i.e. consider  $R'_\zeta \bar{B}' \bar{A}' \text{vec } \zeta = c_\zeta$  and  $R'_\zeta \bar{B}' \bar{A}' \text{vec } \zeta^\circ = c_\zeta$ . Subtracting these two equations and setting  $G := I_{2r+s} - Q$ , one finds

$$0 = R'_\zeta \bar{B}' \bar{A}' \text{vec } (\zeta G) = R'_\zeta \bar{B}' \bar{A}' (I_{2r+s} \otimes \zeta) \text{vec } G \quad (62)$$

Let  $G_0 := L'_{23} G L_{12}$  be the lower-left block of  $G$  and  $G_{11} = I_r - Q_{11}$ , where  $G$  is partitioned conformably with  $Q$ . Because of (10), one observes that

$$G = L_1 G_{11} L'_1 + L_3 G_{11} L'_3 + L_{23} G_0 L'_{12}.$$

Applying vecs, one finds  $\text{vec } G = (N_1 : N_0)(g'_1 : g'_0) =: Ng$  where  $g_1 := \text{vec } G_{11}$ ,  $g_0 := \text{vec } G_0$ ,  $N_1 := L_1 \otimes L_1 + L_3 \otimes L_3$  and  $N_0 := L_{12} \otimes L_{23}$ . We observe that  $G = 0$  iff  $g := (g'_1 : g'_0) = 0$ , i.e. iff  $G_{11} = 0$  and  $G_0 = 0$ .

Substituting  $\text{vec } G = Ng$  into (62) one finds that  $0 = R'_\zeta \bar{B}' \bar{A}' (I_{2r+s} \otimes \zeta) Ng$  implies  $g = 0$  iff  $R'_\zeta \bar{B}' \bar{A}' (I_{2r+s} \otimes \zeta) N$  has full column rank. Substituting (61) in the last expression one finds

$$\bar{B}' \bar{A}' (I \otimes \zeta) N = ((L_1 \otimes v + L_3 \otimes \beta) : L_{12} \otimes (\gamma : \beta)),$$

so that, rearranging the columns of this matrix, one obtains the rank condition (22) with the matrix  $C_\zeta$  defined in (24). Note that one can rearrange the columns of  $C_\zeta$ , because this does not alter the rank of  $R'_\zeta C_\zeta$ . The order condition simply requires the number of rows to be greater or equal to the number of columns in these equations.

Next recall that from (56) one has  $R_\dagger = (B_\perp : \bar{B}R_\zeta)$ , where  $B'_\perp (I \otimes \zeta) N = 0$  by construction, so that  $\text{rank}(R'_\dagger (I \otimes \zeta) N) = \text{rank}(R'_\zeta \bar{B}' (I \otimes \zeta) N)$ . Substituting from (60) one obtains

$$\bar{A}' (I \otimes \zeta) N = ((U_1 \otimes \zeta L_1) : (I_{r+s} \otimes \zeta L_{23})).$$

so that, rearranging the columns of this matrix, one obtains  $\text{rank} \left( R'_\zeta C_\zeta \right) = \text{rank} \left( R'_\dagger C_\dagger \right)$  with the matrix  $C_\dagger$  defined in (23). The number of columns  $m_\dagger$  in  $R_\dagger$  is greater or equal to the one in  $R_\zeta$  as follows directly from  $R_\dagger = (B_\perp : \bar{B}R_\zeta)$ .  $\square$

*Proof* of Theorem 5. Condition (27) is a restatement of condition (22) in Theorem 4. This is obtained by substituting the special form of  $R_\dagger := \text{blkdiag}(R_{11}, R_{22})$ . The conditions on single equations are obtained selecting the elements in  $f$  which correspond to the stated equations, see the proof of Theorem 4.  $\square$

**A.4. Jacobian.** We report two results that are used in the proof of Theorem 6; they are stated in Lemma 7 and Theorem 8.

Consider a full column rank  $(b + c) \times n$  matrix  $F$ ; we partition  $F$  in the first block of  $b$  rows  $S$  and the remaining block of  $c$  rows  $V$ ,  $F = (S' : V)'$ .

**Lemma 7 (Bases and orthogonal complements).** *One has  $\text{col } F = \text{col } F^\circ$  for*

$$F^\circ = \begin{pmatrix} G & 0 \\ N & M \end{pmatrix}, \quad (63)$$

with the following properties:

- (i)  $\text{col } G = \text{col } S$ ,  $M$  is full column rank and  $N = M_\perp Q$ , with  $Q := \bar{M}'_\perp N$  not necessarily of full column rank, with rank-decomposition  $Q = gk'$  with  $g$  and  $k$  of full column rank;
- (ii) any of the following two matrices is a basis of  $\text{col}^\perp F$ :

$$\begin{pmatrix} G_\perp & -\bar{G}Q' \\ 0 & \bar{M}_\perp \end{pmatrix}, \quad \begin{pmatrix} G_\perp & -\bar{G}k & 0 \\ 0 & \bar{M}_\perp \bar{g} & \bar{M}_\perp g_\perp \end{pmatrix}. \quad (64)$$

*Proof* of Lemma 7. (i): Rank-decompose  $S = GW'$  with  $G, W$  of full column rank. Post-multiplying  $F$  by the full rank, square matrix  $Z := (\bar{W} : W_\perp)$ , one obtains a lower block triangular matrix

$$FZ = \begin{pmatrix} G & 0 \\ Y & M \end{pmatrix},$$

where  $Y := V\bar{W}$ ,  $M := VW_\perp$ . Because  $\text{col}(FZ) = \text{col } F$  implies that  $FZ$  has full column rank  $n$  and because  $G$  is of full column rank by construction, one has that also  $M$  must have full column rank. Next define  $F^\circ = FZP$  with

$$P := \begin{pmatrix} I & 0 \\ -\bar{M}'Y & I \end{pmatrix}$$

which has the form (63) with  $N = (I - M\bar{M}')Y = M_\perp \bar{M}_\perp Y = M_\perp Q$ , as required.

(ii): The first matrix in (64) is of full column rank, it is orthogonal to  $F^\circ$  and it is of the right dimensions; hence it is a basis of  $\text{col}^\perp F$ . The second matrix in (64) is equal to the first, post-multiplied by  $\text{blkdiag}(I, (\bar{g} : g_\perp))$ , which is nonsingular by construction; hence also the second matrix is a basis of  $\text{col}^\perp F$ .  $\square$

The next result is an extension of Theorem 2 in Paruolo (2006).

**Theorem 8 (A rank equality).** *Let  $R$  and  $C$  be  $p \times d$  and  $p \times r$  respectively, both of full column rank; then*

$$\text{rank}(R'_\perp C_\perp) = p - r - d + \text{rank}(R'C).$$

*Proof* of Theorem 8. We follow the proof of Theorem 2 in Paruolo (2006). Define  $F := (R : C_\perp)$ ,  $N := (C : \bar{C}_\perp)$ ,  $G := (\bar{R} : R_\perp)$ , and observe that  $N$  and  $G$  are square and of full rank. Hence  $\text{rank } F = \text{rank}(N'F) = \text{rank}(G'F)$ . Note that  $N'F$  is lower block triangular with diagonal blocks equal to  $I_{p-r}$  and  $R'C$ ; this implies  $\text{rank}(N'F) = p - r + \text{rank}(R'C)$ . Similarly  $G'F$  is upper block

triangular diagonal blocks equal to  $I_d$  and  $R'_\perp C_\perp$ , and hence  $\text{rank}(G'F) = d + \text{rank}(R'_\perp C_\perp)$ . Setting  $\text{rank}(N'F) = \text{rank}(G'F)$ , one obtains the statement in the theorem.  $\square$

We next turn to the proof of Theorem 6.

*Proof of Theorem 6.* Let  $F := \partial\Phi^v/\partial\theta'_\eta$  and  $GK := \partial\Phi^v/\partial\theta'_\zeta$  in (33), where  $G := (G_1 : G_2 : G_3)$  is partitioned conformably with  $\theta_\zeta := \text{vec}(v : \gamma : \beta)$ .

*Proof of (i).* We want to find the rank  $\lambda$  of  $(F : GK)$ , i.e.  $\lambda := \dim \text{col}(F : GK) = \dim \text{col}(F : G)$ ; to this end, we compute  $w := \dim \text{col}^\perp(F : G) = \dim(\text{col}^\perp(F) \cap \text{col}^\perp(G))$  and then use the rank-nullity theorem to obtain  $\lambda = 2p^2 - w$ .

We define the rank decompositions of  $\delta' := (\bar{\alpha}'\Gamma\bar{\beta}_2)'$  and  $\bar{\alpha}'_2\Gamma\bar{\beta}$ ; in particular  $\delta' := (\bar{\alpha}'\Gamma\bar{\beta}_2)' = gk'$  where  $g, k$  are of full column rank  $n_g$ , and similarly  $\bar{\alpha}'_2\Gamma\bar{\beta} = uv'$ , where  $u, v$  are of full column rank  $n_u$ .

One can replace  $F, G$  with different bases of the same spaces  $F^\circ, G^\circ$ . Using Lemma 7(i), one finds

$$\begin{aligned} F^\circ &:= (F_1^\circ : F_2^\circ : F_3^\circ) = ((J_1\beta + J_2\beta_2gk') \otimes I_p : J_2\beta_1 \otimes I_p : J_2\beta \otimes I_p) \\ G^\circ &:= (G_1^\circ : G_2^\circ : G_3^\circ) = (J_2 \otimes \alpha : J_2 \otimes \alpha_1 : (J_1 \otimes \alpha + J_2 \otimes \alpha_2uv')). \end{aligned}$$

Applying Lemma 7(ii), one finds that  $F_\perp$  and  $G_\perp$  can be chosen equal to

$$\begin{aligned} F_\perp &:= (F_{\perp 1} : F_{\perp 2} : F_{\perp 3}) = (J_1\beta_\perp \otimes I_p : (J_2\bar{\beta}_2\bar{g} - J_1\bar{\beta}k) \otimes I_p : J_2\bar{\beta}_2g_\perp \otimes I_p), \\ G_\perp &:= (G_{\perp 1} : G_{\perp 2} : G_{\perp 3}) = (J_1 \otimes \alpha_\perp : (J_2 \otimes \bar{\alpha}_2\bar{u} - J_1 \otimes \bar{\alpha}v) : J_2 \otimes \bar{\alpha}_2u_\perp). \end{aligned}$$

We next show that a basis  $c$  of  $\text{col}(F_\perp) \cap \text{col}(G_\perp)$  can be chosen equal to  $c := (c_1 : c_2 : c_3 : c_4 : c_5)$ , where

$$\begin{aligned} c_1 &:= J_1\beta_\perp \otimes \alpha_\perp, \quad c_2 := (J_2\bar{\beta}_2\bar{g} - J_1\bar{\beta}k) \otimes \bar{\alpha}_2u_\perp, \quad c_3 := J_2\bar{\beta}_2g_\perp \otimes \bar{\alpha}_2u_\perp, \\ c_4 &:= J_2\bar{\beta}_2g_\perp \otimes \bar{\alpha}_2\bar{u} - J_1\bar{\beta}_2g_\perp \otimes \bar{\alpha}v, \quad c_5 := -J_1\bar{\beta}k \otimes \bar{\alpha}_2\bar{u} + J_2\bar{\beta}_2\bar{g} \otimes \bar{\alpha}_2\bar{u} - J_1\bar{\beta}_2\bar{g} \otimes \bar{\alpha}v. \end{aligned}$$

The orthogonality of  $c$  and  $F^\circ : G^\circ$  can be verified directly. One observes that  $c$  has full column rank, and its number of columns is

$$\begin{aligned} w &= (p-r)^2 + n_g(p_2 - n_u) + (p_2 - n_g)(p_2 - n_u) + (p_2 - n_g)n_u + n_gn_u = \\ &= (p-r)^2 + p_2(p_2 - n_u) + p_2n_u = (p-r)^2 + p_2^2 \end{aligned}$$

where we use the shorthand  $p_2 := p - r - s$ . Hence  $\lambda = 2p^2 - w = 2p^2 - (p-r)^2 - (p-r-s)^2$ , as had to be shown.

*Proof of (ii).* Let  $d := r^2 + (r+s)^2$ ,  $j := 2(p-r) - s$ . We wish to show that in this case a basis of the orthogonal complement of  $\mathcal{J}$  is given by  $c := (c_1 : c_2 : \dots : c_5 : c_6)$ , where  $c_1, \dots, c_5$  are defined as in (i) and  $c_6$  is an additional matrix orthogonal to  $\mathcal{J}$ . We wish to show that  $c'_6 = M'_\perp Z^* \mathcal{K}_{r+s,j} V'$  where  $M := Z^* \mathcal{K}_{r+s,j} V' G K H_\zeta$  is  $(f-d) \times (f-m_\zeta)$  and of full column rank, thanks to the rank condition.

The matrices that enter the definition of  $M$  are defined as follows:  $V = (\zeta_\perp \otimes a)$ ,  $Z^* := \text{blkdiag}(Z^{-1}, I_s)$ ,  $Z := I + (Y_1 \otimes \ell)$ ,  $\ell := a_\ell b'_\ell$ ,  $a_\ell := (0_{s+p_2, p_2} : I_{p_2})'$ ,  $b_\ell := (0_{s+p_2, s+p_2} : \beta'_\perp \beta_2)'$ ,

$Y := (Y'_1 : Y'_2)' := a' \zeta$  with  $Y_1$  of dimension  $r \times r$ ,  $a' := ((\bar{\alpha} : \bar{\alpha}_1)'(\alpha : \xi))^{-1} (\bar{\alpha} : \bar{\alpha}_1)'$ , where

$$(\bar{\alpha} : \bar{\alpha}_1)'(\alpha : \xi) = \begin{pmatrix} I & \bar{\alpha}'\xi \\ 0 & (\gamma'\bar{\beta}_1)^{-1} \end{pmatrix}$$

is invertible, see Theorem 2. This implies that  $a'(\alpha : \xi) = I_{r+s} =: (U_1 : U_2)$ .

Note that  $F_\perp$  can also be chosen equal to  $(\zeta_\perp \otimes I_p)$ . By Lemma 7(ii), eq. (64),  $\zeta_\perp$  can be chosen as

$$\zeta_\perp = \begin{pmatrix} \beta_\perp & -\bar{\beta}\delta \\ 0_{p,p-r} & \bar{\beta}_2 \end{pmatrix}. \quad (65)$$

Next compute  $V'G = (b'_2 \otimes U_1 : b'_2 \otimes U_2 : b'_2 \otimes Y + b'_1 \otimes U_1)$ , where  $b_1 := J'_1 \zeta_\perp = (\beta_\perp : -\bar{\beta}\delta)$  and  $b_2 := J'_2 \zeta_\perp = (0_{p,p-r} : \beta_2)$  are  $p \times j$ . Pre-multiplying by  $\mathcal{K}_{r+s,j}$  and post-multiplying by  $K$ , by the properties of commutation matrices, see e.g. Magnus and Neudecker (2007) Theorem 3.9(b), one finds

$$\mathcal{K}_{r+s,j} V'GK = (U_1 \otimes b'_2 : U_2 \otimes b'_2 : Y \otimes b'_2 + U_1 \otimes b'_1) = (I_{r+s} \otimes b'_2 : Y \otimes b'_2 + U_1 \otimes b'_1).$$

Define  $N := (D \otimes I_p : L_2 \otimes I_p)$ ,  $D := (L_{3,1} : L_{1,1} : L_{3,2} : L_{1,2} : \cdots : L_{3,r} : L_{1,r})$ , where  $L_{i,t}$  denotes column  $t$  in  $L_i$ ; in words,  $D$  contains all the paired columns of  $L_3$  and  $L_1$ . Note that  $N$  is an orthogonal matrix, so that  $N^{-1} = N'$ , and define  $\tilde{H}_\zeta := N'H_\zeta$  and  $\tilde{R}_\zeta := N'R_\zeta$ , where  $\tilde{R}'_\zeta \tilde{H}_\zeta = R'_\zeta N N' H_\zeta = R'_\zeta H_\zeta = 0$ .

Rewriting  $Z^{*-1}M$  as  $\mathcal{K}_{r+s,j} V'GKN\tilde{H}_\zeta$ , one finds  $Z^{*-1}M = \text{blkdiag}(Q, I_s \otimes b'_2)\tilde{H}_\zeta$ , with  $Q := (I_r \otimes \zeta'_\perp) + (Y_1 \otimes (b'_2 : 0))$ . Note that, using (65), one can write  $(b'_2 : 0) = a_\ell b'_\ell \zeta'_\perp$ , and hence  $Y_1 \otimes (b'_2 : 0) = (Y_1 \otimes a_\ell b'_\ell)(I_r \otimes \zeta'_\perp)$ . This implies  $Q = Z(I_r \otimes \zeta'_\perp)$ , where  $Z := I_{rj} + (Y_1 \otimes a_\ell b'_\ell)$ . We next wish to show that  $Z$  is invertible; to this end, we partition  $Y_1$  as follows

$$Y_1 =: \begin{pmatrix} Y^{(1)} & y^{(2)} \\ y^{(3)} & y_{rr} \end{pmatrix}$$

where  $Y^{(1)}$  is  $(r-1) \times (r-1)$ ,  $y^{(2)}$  is  $(r-1) \times 1$ ,  $y^{(3)}$  is  $1 \times (r-1)$  and  $y_{rr}$  is a scalar. Partition also  $Z$  conformably as follows

$$Z = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}$$

with  $Z_{22} := I_j + y_{rr}\ell$ ,  $Z_{11} := I_{(r-1)j} + Y^{(1)} \otimes \ell$ ,  $Z_{12} := y^{(2)} \otimes \ell$ ,  $Z_{21} := y^{(3)} \otimes \ell$ .

Note that  $Z_{22}$  is lower block triangular with diagonal blocks equal to identity matrices; this implies that  $Z_{22}$  is invertible, and one finds  $Z_{22}^{-1} = I_j - y_{rr}\ell$ . One can then apply the Guttman rank additivity formula, see Zhang (2005) formulas (0.9.2) and (6.0.4), which states that the rank of  $Z$  equals the rank of  $Z_{22}$ , which is equal to  $j$ , plus the rank of the Schur complement of  $Z_{22}$ , i.e. of  $Z_{11.2} := Z_{11} - Z_{12}Z_{22}^{-1}Z_{21}$ . Because  $\ell^2 = a_\ell b'_\ell a_\ell b'_\ell = 0$  due to the orthogonality of  $a_\ell$  and  $b_\ell$ ,

$b'_\ell a_\ell = 0$ , one finds

$$\begin{aligned} Z_{12}Z_{22}^{-1}Z_{21} &= (y^{(2)} \otimes \ell)(I - y_{rr}a_\ell b'_\ell)(y^{(3)} \otimes \ell) = \\ &= (y^{(2)}y^{(3)} \otimes \ell^2) - (y^{(2)} \otimes \ell)(1 \otimes y_{rr}\ell)(y^{(3)} \otimes \ell) \\ &= (y^{(2)}y^{(3)} \otimes \ell^2) - (y^{(2)}y^{(3)} \otimes y_{rr}\ell^3) = 0. \end{aligned}$$

This implies that  $Z_{11.2} = Z_{11}$ , which has the same form of the original matrix  $Z$ . Repeating this argument on  $Z_{11}$ , one finds that the rank of  $Z$  is full, i.e. that  $Z$  is invertible. Hence, pre-multiplying by  $Z^*$  one finds

$$M = \text{blkdiag}(I_r \otimes \zeta'_\perp, I_s \otimes b'_2)\tilde{H}_\zeta. \quad (66)$$

By the rank condition (22),  $\text{rank}(R'_\zeta C_\zeta) = d$  where  $R_\zeta$  and  $C_\zeta$  have dimensions  $f \times m_\zeta$  and  $f \times d$  respectively; the corresponding order condition is  $m_\zeta \geq d$ . Computing  $\tilde{C}_\zeta := N' C_\zeta$  one finds  $\tilde{C}_\zeta = \text{blkdiag}(I_r \otimes \zeta, I_s \otimes (\gamma : \beta))$ , with  $\tilde{C}_{\zeta\perp} = \text{blkdiag}(I_r \otimes \zeta_\perp, I_s \otimes \beta_2)$ , so that one can write (66) as  $M = \tilde{C}'_{\zeta\perp} \tilde{H}_\zeta$ . From Theorem 8 one has

$$\text{rank}(C'_{\zeta\perp} H_\zeta) = f - d - m_\zeta + d = p(2r + s) - m_\zeta$$

where  $M := C'_{\zeta\perp} H_\zeta = \tilde{C}'_{\zeta\perp} \tilde{H}_\zeta$  has dimension  $(f - d) \times (f - m_\zeta)$ , where  $(f - d) \geq (f - m_\zeta)$  by the order condition. This implies that  $M_\perp$  has  $f - d - f + m_\zeta = m_\zeta - d$  columns, which is equal to the number of over-identifying restrictions. One can show that  $c_6$  is linearly independent from  $(c_1 : \dots : c_5)$  and hence this implies that the null space of the Jacobian  $\mathcal{J}$  has dimension  $w = (p - r)^2 + p_2^2 + m_\zeta - d$ . This implies

$$\begin{aligned} \lambda &= 2p^2 - w = 2p^2 - (p - r)^2 - (p - r - s)^2 - m_\zeta + r^2 + (r + s)^2 \\ &= 2p(2r + s) - m_\zeta = p(2r + s) + n_\zeta. \end{aligned}$$

Finally we observe that in the case of the extended star parametrization, the number of parameters in  $\Phi$  is not equal to  $2p^2$  but it equals  $2p(p + q)$ . The remaining calculations are unaffected by this change; this completes the proof of the theorem.  $\square$

**A.5. Standard errors.** The I(2) model is a sub-model of the regression model

$$Y_t = \Psi U_t + \varepsilon_t,$$

where  $\Psi = \Psi(\phi)$  is a be a smooth function of the vector of parameters  $\phi$ . Specifically,  $Y_t := \Delta^2 X_t$ ,  $Z_{1t} := (X_{t-1}^{*'} : \Delta X_{t-1}^{*'})'$ ,  $Z_{2t} := (\Delta^2 \mathbb{X}'_{t-1} : \mathbb{D}'_t)'$ ,  $U_t := (Z'_{1t} : Z'_{2t})'$ ,  $\Psi := (\Phi : \Upsilon)$ ,  $\Phi := (\Pi : \Gamma)$ . Let  $\ell(\phi, \Omega) := \log L(\phi, \Omega)$  be the Gaussian log-likelihood function. The second derivatives with respect to  $\Psi^v := \text{vec } \Psi$  are  $-T M_{UU} \otimes \Omega^{-1}$ , see e.g. Johansen (2006) eq. (13), where  $M_{UU} = T^{-1} \sum_{t=1}^T U_t U_t'$ . Because the parameters in  $\phi$  and in  $\Omega$  are asymptotically independent, one finds that minus the hessian with respect to  $\phi$  equals

$$\mathcal{I}_{\phi, T} := -\frac{\partial^2 \ell(\phi, \Omega)}{\partial \phi \partial \phi'} = T \mathcal{J}(\phi)' (M_{UU} \otimes \Omega^{-1}) \mathcal{J}(\phi), \quad \mathcal{J}(\phi) := \frac{\partial \Psi^v(\phi)}{\partial \phi'}$$

When we are interested in a subset of parameters  $v'\phi$ , and when LAMN inference applies, one finds that the ‘variance’ of  $v'\phi$  can be estimated by  $v' \mathcal{I}_{\phi, T}^{-1} v$ . Johansen (2006) and Boswijk (2010) show

that not all inference on the cointegration parameters is LAMN in the I(2) model. Because preliminary Monte Carlo simulations show that departures from LAMN are minor, we report standard errors for all cointegration parameters using the square root of  $v' \mathcal{I}_{\phi, T}^{-1} v$ .

In order to describe  $\mathcal{J}$  in more detail, call  $\phi_{\Upsilon}$  the free parameters in  $\Upsilon$ , which are assumed to be functionally independent from  $\phi_{\Phi}$  in  $\phi = (\phi'_{\Phi} : \phi'_{\Upsilon})' = (\phi'_{\eta} : \phi'_{\zeta} : \phi'_{\Upsilon})'$ . Because  $\Psi := (\Phi : \Upsilon)$ , and the assumption of functional independence, one has

$$\mathcal{J}(\phi) = \frac{\partial \Psi^v(\phi)}{\partial \phi'} = \text{blkdiag} \left( \frac{\partial \Phi^v(\phi)}{\partial \phi'_{\Phi}}, \frac{\partial \Upsilon^v(\phi)}{\partial \phi'_{\Upsilon}} \right)$$

where  $\partial \Phi^v(\phi) / \partial \phi'_{\Phi}$  is given in eq. (33).