Paolo Paruolo

Testing for common trends in conditional I(2) VAR models

2002/28

UNIVERSITÀ DELL'INSUBRIA
FACOLTÀ DI ECONOMIA

http://eco.uninsubria.it

In questi quaderni vengono pubblicati i lavori dei docenti della Facoltà di Economia dell'Università dell'Insubria. La pubblicazione di contributi di altri studiosi, che abbiano un rapporto didattico o scientifico stabile con la Facoltà, può essere proposta da un professore della Facoltà, dopo che il contributo sia stato discusso pubblicamente. Il nome del proponente è riportato in nota all'articolo. 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.

These Working papers collect the work of the Faculty of Economics of the University of Insubria. The publication of work by other Authors can be proposed by a member of the Faculty, provided that the paper has been presented in public. The name of the proposer is reported in a footnote. 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.

Printed in Italy in December 2002
Università degli Studi dell'Insubria
Via Ravasi 2, 21100 Varese, Italy

Testing for common trends in conditional I(2) VAR models

Paolo Paruolo*

Department of Economics, University of Insubria

Via Ravasi 2, 21100 Varese, Italy

December 2002

Abstract

This paper presents cointegration tests in the integration indices (II) in cointegrated (CI) vector autoregressive processes (VAR). The statistical analysis is performed under the assumption that some variables are weakly exogenous with respect to the (multi-)cointegration parameters, a condition that corresponds to no integral and proportional feedback in the marginal system (NF). The specification of the deterministic components is chosen so as to allow for a linear trend in all possible directions. The asymptotic distribution is derived both under correct specification of the weak exogeneity assumptions and under mis-specification. Tables of limit distributions are obtained by simulation. It is found that some types of mis-specification modify the asymptotic distributions of the tests considerably. However, the asymptotics are unaffected by misspecification provided the adjustment coefficients in the conditional system are of full rank.

Keywords: Cointegration rank test, Common trends, VAR, I(2), 2SI2, Conditional systems

*Partial financial support from Italian MIUR grants Cofin2002 and ex 60% is gratefully acknowledged. Invited paper at the Latin American Meeting of the Econometric Society 2002, São Paolo, Brasil, 24-27 July, 2002.

Contents

1 Introduction

The determination of CI rank is a key aspect in the analysis of I(1) systems. Likelihood ratio (LR) tests for the estimation of CI rank in VAR systems are widely used, see Johansen (1996). Of similar importance is the determination of integration indices (II) in I(2) systems, see Johansen (1995), Paruolo (1996), Rahbek et al. (1999).

Sometimes the analysis of a complete system is not deemed feasible or appropriate, e.g. due to a limited number of degrees of freedom in medium– to large–scale systems. In this case one can turn to the analysis of 'conditional' or 'partial' systems, where some variables are not modelled. The LR–based determination of CI rank in conditional I(1) systems has been addressed in Harbo et al. (1998). Paruolo and Rahbek (1999) have discussed weak exogeneity with respect to the cointegration parameters in conditional I(2) systems.

The present paper addresses the issue of determination of II in conditional I(2) systems, on the basis of the Two Stage I(2) analysis (2SI2) of conditional I(2) systems proposed by Johansen (1995). The 2SI2 analysis is based on repeated application of the reduced rank regression (RRR) technique of Anderson (1951).

The statistical analysis is performed under the assumption that some variables are weakly exogenous with respect to the (multi-)cointegration parameters, i.e. there is no integral and proportional feedback in the marginal system (NF). This assumption implies a lower bound on the number of I(2) common trends, equal to the number of un-modelled variables.

The specification of the deterministic components is chosen so as to allow for a linear trend in all possible directions, along the lines of Rahbek et al. (1999). This choice allows to make inference on the II before testing hypothesis on the linear trend. In this paper we also include seasonal dummy variables.

The asymptotic distribution is derived both under correct specification of the NF assumptions and under misspecification. It is found that some types of misspecification modify the asymptotic distributions of the tests considerably. However, the asymptotics are unaffected by misspecification provided the adjustment coefficients in the conditional system are of full rank, i.e. there is enough information conveyed by the equilibrium correction coefficients on the CI relations in the conditional system.

Tables of limit distributions are obtained by simulation, making use of response surface analysis and of the Gamma approximation suggested in Doornik (1998) for multivariate unit root distributions. We adopt the sample-size response surface proposed in MacKinnon et al. (1999) on the Monte Carlo (MC) moments, in order to estimate the first two moments of the limit distributions. We select here a maximal sample size equal to 5000 and run 10^6 replications for each design, which provide accurate MC estimates of moments. Tables of limit quantiles are provided. As a by-product, these tables also give better estimates of limit quantiles for the full-system analysis described in Rahbek et al. (1999) than previously published ones, because the present ones are based on more replications and larger maximal sample size.

The plan of the rest of the paper is the following: Section 2 reviews basic properties of I(2) system and defines the conditional I(2) model. Section 3 discusses the modification of the 2SI2 analysis for conditional systems. Section 4 presents the asymptotic results under correct specification, and describes the MC esperiment.

Section 5 investigates the effects of misspecification, while Section 6 reports an empirical illustration of the proposed II estimator. Section 7 summarizes and concludes. Proofs and tables of quantiles are placed in four Appendices.

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 matrices H, A, B, $sp(H)$ is the linear span of the columns of H, $\bar{H}$ indicates $H(H'H)^{-1}$ and $H_\perp$ indicates a basis of $sp(H)^\perp$, the orthogonal complement of $sp(H)$. Moreover $P_H := H\bar{H}'$. Finally vec is the column stacking operator, $\otimes$ is the Kronecker product (i.e. $A \otimes B$ is the matrix with generic block $a_{ij}B$, where $A := (a_{ij})$) and $\overset{w}{\to}$ indicates weak convergence.

2 I(2) representation

In this paper we consider k-th order vector autoregressive (VAR(k)) systems of the type

$$\Delta^2 X_t = \Pi X_{t-1} + \Gamma \Delta X_{t-1} + \sum_{i=1}^{k-2} \Upsilon_i \Delta^2 X_{t-i} + \mu_0 + \mu_1 t + \varkappa d_t + \epsilon_t$$

for $k \geq 2$. Here X_t and ϵ_t are $p \times 1$ and ϵ_t is i.i.d. $N(0, \Omega)$, $d_t := (d_{1,t} : ... : d_{n-1,t})'$ is a vector of seasonal dummies 'orthogonal' to the constant, i.e. of the form $d_{i,t} = 1(t \bmod n = i) - 1/n$, $1(\cdot)$ is the indicator function and n is the number of seasons. These dummies have the property that $\left|\sum_{t=1}^{j} d_{i,t}\right| < 1$, although $\left|\sum_{j=1}^{h}\sum_{t=1}^{j} d_{i,t}\right| = O(h)$, i.e. they generate a linear trend if cumulated twice. This sort of deterministics is similar to the one in Rahbek et al. (1999) and leads to a simpler statistical analysis with respect to the models in Paruolo (1996). This approach permits to postpone tests on the deterministic trend to a later stage of the analysis.

In the next three subsections we summarize results on the representation of I(2) systems under the NF condition.

2.1 Common trends representation

In this subsection we report Johansen's I(2) representation theorem, see Johansen (1992) and Rahbek (1997), for the present choice of deterministics. We first list some assumptions. Let $\Upsilon := \sum_{i=1}^{k-2} \Upsilon_i$.

I(2) conditions

I(2)_a Every root z of the characteristic polynomial of X_t satisfies $z = 1$ or $|z| > 1$.

I(2)_b $\Pi = \alpha\beta'$, where α and β are $p \times p_0$ matrices of full rank $p_0 < p$.

I(2)_c $P_{\alpha_\perp}\Gamma P_{\beta_\perp} = \alpha_1\beta_1'$ where α_1 and β_1 are $p \times p_1$ matrices of full rank $p_1 < p - p_0$, or, equivalently, $\alpha_\perp'\Gamma\beta_\perp = \xi\eta'$ where $\xi = \alpha_\perp'\alpha_1$ and $\eta = \beta_\perp'\beta_1$ are $p - p_0 \times p_1$ matrices of full rank $p_1 < p - p_0$

I(2)_d $\alpha_2'\theta\beta_2$ has full rank $p - p_0 - p_1$, where $\alpha_2 = (\alpha : \alpha_1)_\perp$, $\beta_2 = (\beta : \beta_1)_\perp$ and θ is defined as

$$\theta := (\Gamma - \Pi)\bar{\beta}\bar{\alpha}'(\Gamma - \Pi) + I - \Upsilon. \tag{1}$$

I(2)_e $\mu_1 = \alpha\beta_0'$, with β_0' a $p_0 \times 1$ vector

I(2)_f $\alpha_\perp'\mu_0 = \xi\eta_0' + \alpha_\perp'\Gamma\bar{\beta}\beta_0'$, with η_0' a $p_1 \times 1$ vector.

In the following 'I(2) assumption' and 'I(2) conditions' are used as synonyms. Johansen's I(2) representation theorem, see Johansen (1992) or (1996) Theorem 4.6, establishes under I(2)_a that necessary and sufficient conditions for

$$\Delta^2 X_t, \qquad \beta' X_t + \delta\beta_2'\Delta X_t, \qquad \beta_1'\Delta X_t \tag{2}$$

to be I(0), apart from deterministic components and initial values, are the conditions I(2)_b to d.

In presence of a constant and trend, it can be proved, see Rahbek (1997), that under I(2)_a, X_t is I(2) and presents linear trends in all directions iff the conditions I(2)_b to f hold. For completeness we restate this result as a proposition in the present setting, which includes dummies. The proof is reported in the Appendix.

Proposition 1 (I(2) representation theorem) *Assume $I(2)$_a holds; then necessary and sufficient conditions for X_t to be $I(2)$ and not to generate cubic or quadratic trends are given by $I(2)$_b to f. Under the $I(2)$ assumption X_t has a common trends representation*

$$X_t = C_2 \sum_{s=1}^{t}\sum_{i=1}^{s} \epsilon_i + C_1 \sum_{i=1}^{t} \epsilon_i + C_0(L)\epsilon_t + m_0 + m_1 t + A + Bt + m(L)d_t,$$

where m_0, m_1 do not depend on initial values while A and B do, $(\beta : \beta_1)'B = 0$, $\beta' A + \bar{\alpha}'\Gamma\bar{\beta}_2\beta_2' B = 0$. Moreover $m(z)$ is a polynomial of degree n,

$$C_2 = \beta_2(\alpha_2'\theta\beta_2)^{-1}\alpha_2', \tag{3}$$

C_1 is given in (27) in the Appendix and m_1 does not vanish when pre-multiplied by β, β_1 or β_2; in particular $\beta' m_1 = -\beta_0'$, $\beta_1' m_1 = -\eta_0' := \bar{\alpha}_1'\mu_0 - \bar{\alpha}_1'\Gamma\bar{\beta}\beta_0'$.

Remarks

1. The exclusion of cubic and quadratic trends in X_t reflects common understanding of economic phenomena.

2. A linear trend is present in all directions, given that $(\beta : \beta_1 : \beta_2)$ span all $\mathbb{R}^p$. The linear trend has coefficient $-\beta_0'$ in the β direction and coefficient $-\eta_0'$ in the β_1 direction.

3. The inclusion of the dummies d_t does not change the main features of the system. In fact d_t generates a linear trend when cumulated twice in the direction of $sp(\beta_2)$. In this direction the linear trend has non vanishing coefficients due to the other deterministic components, and it is stochastically dominated by the I(2) common trends. For details we refer to the proof of Proposition 1 in the Appendix.

4. C_2 is the matrix of loadings of the I(2) common stochastic trends into variables of the system. This matrix can be decomposed in the product of $\beta_2(\alpha_2'\theta\beta_2)^{-1}$ and α_2'; the linear combinations $\alpha_2'\epsilon_i$ are the innovations that form the common I(2) trends. Inference on α_2 and C_2 is addressed in Paruolo (2002a). The matrix C_2 also plays a central role in the calculation of impact factors, defined in Omtzigt and Paruolo (2002).

Under the I(2) assumption the system can be rewritten imposing condition $I(2)\_e$ as follows

$$\begin{aligned}\Delta^2 X_t &= \alpha(\beta' : \beta_0')\begin{pmatrix} X_{t-1} \\ t \end{pmatrix} + \Gamma\Delta X_{t-1} + \Psi K_t + \mu D_t + \epsilon_t = \\ &= \alpha\beta^{*\prime}X_{t-1}^* + \Gamma\Delta X_{t-1} + \Psi K_t + \mu D_t + \epsilon_t,\end{aligned} \tag{4}$$

where $\beta^{*\prime}X_{t-1}^* := \beta' X_{t-1} + \beta_0' t$, $\Psi := (\Upsilon_1 : \ ... \ : \Upsilon_{k-2})$, $K_t := (\Delta^2 X_{t-1}' : ... : \Delta^2 X_{t-k+2}')'$, $\mu := (\varkappa : \mu_0)$, $D_t := (d_t' : 1)'$.

2.2 No feedback condition

This subsection reports conditions under which some variables are weakly exogeneous with respect to the CI parameters β, β_1, δ, see (2).[1] Let $a'X_t$ be the m modelled variables, where a is a known $p \times m$ selection matrix, let $b'X_t$ be the (candidate) $p-m$ weakly exogenous variables, where b is a known $p \times p-m$ selection matrix and $(a : b)$ is $p \times p$ and full rank. The complete system is decomposed in the marginal model of $b'X_t$ given the past and the model of $a'X_t$ conditional on $b'X_t$ and the past. Without loss of generality, one can choose $b_\perp$ such that a has representation $a = b_\perp + ba_0$. [2]

The following examples illustrate two possible cases.

Example 2 *As an illustration of the possible choices of a and b, consider a 4-variable system consisting of the logarithms of nominal money $m1_t$, prices pr_t, real income inc_t and of the opportunity cost of holding money R_t^*, $X_t := (m1_t : pr_t : inc_t : R_t^*)'$, see Section 6 below. Assume that one wishes to model real money $m1_t - pr_t$, income inc_t and prices pr_t conditional on interest rates R_t^*. To this end one could choose $b = (0 : 0 : 0 : 1)'$ and $a = (a_1 : a_2 : a_3)$, $a_1 = (1 : -1 : 0 : 0)'$, $a_2 = (0 : 1 : 0 : 0)'$, $a_3 = (0 : 0 : 1 : 0)'$ such that the 'endogenous' variables are $a'X_t = (m1_t - pr_t : inc_t : pr_t)'$.*

Example 3 *Another possible choice is $a = (I_m : 0)$, $b = (0 : I_{p-m})'$, which corresponds to selecting the m first variables as endogenous and the last $p - m$ variables of X_t, i.e. $(X_{m+1} : \ldots : X_p)'$ as (candidate weakly) exogenous.*

The marginal system equations for $b'X_t$ given the past are given by

$$b'\Delta^2 X_t = \alpha_b\beta^{*\prime}X_{t-1}^* + b'\Gamma\Delta X_{t-1} + b'\Psi K_t + b'\mu D_t + \epsilon_{bt}, \tag{5}$$

[1] Recall that β_2 is a function of $(\beta : \beta_1)$, and thus it is not a separate parameter.

[2] In fact, from the requirement that $(a : b)$ has full rank, $b_\perp^0 = P_{b_\perp}a$ has full rank m; thus $b_\perp^0$ is a basis of span$(b_\perp)$, and we can choose $b_\perp = b_\perp^0$. It is simple to verify by orthogonal projections that for this choice $a = b_\perp + ba_0$, where $a_0 = \overline{b}'a$.

where $\alpha_b := b'\alpha$ and $\epsilon_{bt} = b'\epsilon_t$ is i.i.d. $N(0, b'\Omega b)$. The conditional system equations for $a'X_t$ given $b'X_t$ and the past, indicated as $a'X_t|b'X_t$, are given by

$$\begin{aligned} a'\Delta^2 X_t &= \omega_a b'\Delta^2 X_t + \alpha_a \beta^{*\prime} X_{t-1}^* + \Gamma_a \Delta X_{t-1} + \Psi_a K_t + \mu_a D_t + \epsilon_{at}, & (6) \\ (\alpha_a &: \Pi_a : \Gamma_a : \Psi_a : \mu_a : \varkappa_a : \epsilon_{at}) := (a' - \omega_a b')(\alpha : \Pi : \Gamma : \Psi : \mu : \varkappa : \epsilon_t) & (7) \\ \omega_a &: = a'\Omega b(b'\Omega b)^{-1}, \end{aligned}$$

where ϵ_{at} is i.i.d. $N(0, \Omega_{aa.b})$, $\Omega_{aa.b} := a'\Omega a - a'\Omega b(b'\Omega b)^{-1} b'\Omega a$ and ϵ_{at} is independent of ϵ_{bt} by standard properties of the Gaussian distribution.

We note the following simple equality.

Proposition 4 *One has*

$$a' - \omega_a b' = b'_\perp - \omega_{b_\perp} b';$$

hence the conditional system $a'X_t|b'X_t$ is identical to the conditional system $b'_\perp X_t|b'X_t$, with model equation

$$b'_\perp \Delta^2 X_t = \omega_{b_\perp} b'\Delta^2 X_t + \alpha_a \beta^{*\prime} X_{t-1}^* + \Gamma_a \Delta X_{t-1} + \Psi_a K_t + \mu_a D_t + \epsilon_{at}, \qquad (8)$$

where the coefficients and error term in (8) *are defined as in* (7) *with $b_\perp$ in place of a.*

In the rest of the paper we take $a = b_\perp$ and (8) as the reference conditional model equations, on the basis of Proposition 4.

In the following we wish to address the estimation of (p_0, p_1) through the analysis of the conditional system (6) (or (8)) alone, disregarding information from the marginal model (5). The resulting loss of information is irrelevant if the marginal system does not have information on the cointegrating coefficients β, β_1, δ, i.e. if $b'X_t$ is weakly exogeneous.

Paruolo and Rahbek (1999) stated the conditions for $b'X_t$ to be weakly exogenous for the cointegrating parameters in (4) with no deterministic terms; it is simple to see that the proof of their Theorem 3.1 is unaffected by the presence of a linear trend in (4), so that the same conditions apply here.[3] In particular the conditions can be written as follows:

$$b'(\alpha : \alpha_1 : \Gamma\overline{\beta}) = 0. \qquad (9)$$

These restrictions correspond to the situation in which the marginal equations of $b'X_t$ have zero adjustment coefficients to the ECM terms. This situation can be described as 'no feedback' in the equations of $b'X_t$, abbreviated as NF.

Paruolo and Rahbek (1999) also noted that a necessary condition for NF is $m \geq p_0 + p_1$, i.e. that the number of modelled variables is greater or equal to the number of CI relations in the system. In the following we call this the 'order condition' for NF. Because $p_2 := p - p_0 - p_1$, this condition corresponds to $p - m \leq p_2$, i.e. that there are at least $p - m$ I(2) trends in the system. Note that $p - m \leq p_2$ does not imply in general that $b'X_t$ is I(2), see Proposition 1, unless $k = 2$.

Obviously one can perform the conditional statistical analysis also when the conditions (9) are not met. In Section 4 we discuss how the violation of these conditions affects the asymptotic distributions, assuming the order condition holds.

[3] For the definition of weak exogeneity see Engle, Hendry and Richard (1983).

2.3 Conditional model

In this subsection we illustrate the effects of the NF condition (9) on the formulation of the I(2) condition. The ensuing representation results are used to define the conditional model parameters. The next proposition states the relation between the marginal- (α_b), conditional- (α_a) and full-system (α) adjustment coefficients. In the following we define $a_1 := b_\perp \alpha_{a\perp}$ and $c := (b_\perp - b\omega'_{b_\perp})\alpha_{a\perp} = a_1 - b\omega'_{b_\perp}\alpha_{a\perp}$ where $\alpha_{a\perp} := (\alpha_a)_\perp$ and $\alpha_{b\perp} := (\alpha_b)_\perp$.

Proposition 5 *The following equalities hold in general*

$$\begin{aligned}\alpha &= \left(\bar{b}_\perp : \Omega b(b'\Omega b)^{-1}\right)\begin{pmatrix}\alpha_a \\ \alpha_b\end{pmatrix} \\ \alpha_\perp &= \left((b_\perp - b\omega'_{b_\perp}) : b\right)\begin{pmatrix}\alpha_{a\perp} & 0 \\ 0 & \alpha_{b\perp}\end{pmatrix} =: (c : b\alpha_{b\perp}),\end{aligned} \tag{10}$$

where $c := (b_\perp - b\omega'_{b_\perp})\alpha_{a\perp}$, $\alpha_{a\perp} := (\alpha_a)_\perp$ *and* $\alpha_{b\perp} := (\alpha_b)_\perp$.

Assume that the NF condition (9) *holds (and hence that the order condition* $m \geq p_0 + p_1$ *is satisfied); then*

i. α *can be represented as* $\alpha = \bar{b}_\perp \alpha_a$ *while* $\alpha_\perp$ *can be chosen as*

$$\alpha_\perp = (b_\perp \alpha_{a\perp} : b) =: (a_1 : b), \tag{11}$$

where $a_1 := b_\perp \alpha_{a\perp}$;

ii. *with the choice* (11) *the condition* $I(2)\_c$ *can be stated as*

$$\alpha'_\perp \Gamma \beta_\perp = \begin{pmatrix} a'_1 \Gamma \beta_\perp \\ 0 \end{pmatrix} = \begin{pmatrix} \xi_1 \\ 0 \end{pmatrix} \eta', \tag{12}$$

with ξ_1 *of dimension* $m - p_0 \times p_1$ *and of full rank* p_1;

iii. *with the choice* (11) *the condition* $I(2)\_f$ *can be stated as*

$$\alpha'_\perp \mu_0 = \begin{pmatrix} \xi_1 \eta'_0 + a'_1 \Gamma \bar{\beta} \beta'_0 \\ 0 \end{pmatrix}; \tag{13}$$

iv. *for the choices* (11) *and* (12) *one can choose* $\alpha_1 = \bar{a}_1 \xi_1$, $\alpha_2 = (a_1 \xi_{1\perp} : b)$.

We observe that Γ_a in (8) is restricted by the I(2) condition. In fact consider $\alpha'_{a\perp}\Gamma_a = \alpha'_{a\perp}(b'_\perp - \omega_{b_\perp} b')\Gamma$; under NF $b'\Gamma = 0$ and hence $\alpha'_{a\perp}\Gamma_a = a'_1\Gamma$. Post-multiplying by β_2, one sees that $\alpha'_{a\perp}\Gamma_a\beta_2 = a'_1\Gamma\beta_2 = 0$ by (12). This shows that a non-linear restriction holds among the coefficients in (8).

In order to define the conditional model parameters, observe that the second reduced rank condition (12) restricts only (a part of) $\alpha'_\perp\Gamma$, while $\alpha'\Gamma$ is unrestricted, where by orthogonal projections, $\Gamma = \bar{\alpha}_\perp \alpha'_\perp \Gamma + \alpha\bar{\alpha}'\Gamma$. We thus define $\nu := \bar{\alpha}'\Gamma$ as an unrestricted parameter.[4]

[4] Note that ν contains the multi(CI) parameter $\delta := \bar{\alpha}'\Gamma\bar{\beta}_2 = \nu\bar{\beta}_2$.

Next consider $\alpha_\perp'\Gamma$, and decompose it into $\alpha_\perp'\Gamma\beta_\perp$ and $\alpha_\perp'\Gamma\bar{\beta}$ again by orthogonal projections. For the choice (11) one sees that $\alpha_\perp'\Gamma\beta_\perp$ contains the ξ_1 and η parameters, see (12). Next note that $\alpha_\perp'\Gamma\bar{\beta} = (a_1 : b)'\Gamma\bar{\beta}$, where $b'\Gamma\bar{\beta} = 0$ by the NF condition (9) and $\gamma := a_1'\Gamma\bar{\beta}$ is another unrestricted parameter. Summarizing, the parameters in Γ under the I(2) and NF conditions are ν, γ, ξ_1 and η, which are unrestricted.

Under correct specification of the NF condition (9), the conditional model is defined by (6) or (8) and the Gaussianity assumption on the error term. The parameters are α_a, β, β_0, ξ_1, η, η_0, ν, γ, $\omega_{b_\perp}$ Ψ_a, μ_a, $\Omega_{aa.b}$. The marginal model consists of the equations (5) and the Gaussianity of the error term, with parameters $b'\Psi$, $b'\mu$, Ω_{bb}.

The parameter space in the statistical analysis is assumed to be unrestricted except for the positive definiteness of the variance-covariance matrices.

3 The 2SI2 analysis of conditional I(2) systems

In this section we summarize the 2SI2 analysis of Johansen (1995), which has already been adapted to the analysis of conditional systems in Paruolo and Rahbek (1999). This procedure makes repeated use of reduced rank regression, RRR, see Anderson (1951).

3.1 First stage

The first stage consists of the likelihood analysis of (6) or equivalently (8) for unrestricted Γ_a. This is achieved through a RRR of $a'Z_{0t} := a'\Delta^2 X_t$ on $Z_{2t} := X_{t-1}^*$ corrected for $Z_{1t} := \Delta X_{t-1}$, $b'Z_{0t} := b'\Delta^2 X_t$ and $Z_{4t} := (Z_{3t}' : 1)'$, where $Z_{3t} := (K_t' : d_t')'$. This stage gives estimates of $\alpha = \bar{b}_\perp\alpha_a$ and β^* which are fixed in the second stage.

In fact for unrestricted Γ_a, the Gaussian likelihood is maximized with respect to β^* by solving the eigenvalue problem

$$\left|\lambda S_{22.(1,b)} - S_{2a.(1,b)}S_{aa.(1,b)}^{-1}S_{a2.(1,b)}\right| = 0, \tag{14}$$

where $S_{ij.h} := S_{ij} - S_{ih}S_{hh}^{-1}S_{hj}$, $S_{ij} := M_{ij} - M_{i4}M_{44}^{-1}M_{4j}$, $M_{yz} := T^{-1}\sum_{t=1}^{T} y_t z_t'$, $y_t, z_t = Z_{jt}$, $j = 0, 1, 2, 4$, and the subscripts a and b correspond to the variables $a'Z_{0t}$ and $b'Z_{0t}$ respectively. Let the eigenvalues of (14) be $\lambda_1 \geq ... \geq \lambda_m \geq \lambda_{m+1} = ...\lambda_{p+1} = 0$. The associated RRR test within the conditional model for $rank(\Pi_a) \leq r$ versus $rank(\Pi_a) \leq m$ when $m > r$ is

$$Q^{p-m}(r) := -T\sum_{i=r+1}^{m} \ln(1 - \lambda_i). \tag{15}$$

When $r \geq m$ then $Q^{p-m}(r) := 0$. Here and in the following the superscript $p - m$ indicates that the test is performed assuming $p - m$ weakly exogenous variables.

3.2 Second stage

The second stage corresponds to the likelihood analysis of the system for fixed α and β. The first stage of the analysis provides estimates of $\widehat{\alpha} = \bar{b}_\perp\widehat{\alpha}_a$ and $\widehat{\beta}^*$ for any given value of p_0; correspondingly one also finds $\widehat{\alpha}_\perp = (b_\perp \widehat{\alpha}_{a\perp} : b) =: (\widehat{a}_1 : b)$ and $\widehat{\beta}_\perp$, see (11). In the second stage α_a, $\alpha_{a\perp}$, β^* and $\beta_\perp$are fixed at the values $\widehat{\alpha}_a$, $\widehat{\alpha}_{a\perp}$, $\widehat{\beta}^*$ and $\widehat{\beta}_\perp$ found in the first stage. In the description of the second step we omit for brevity the hat over estimated quantities from the first step; in the Appendix we resort to the full notation when needed for clarity.

The second stage consists of two separate steps, only one of which is relevant here. Recall that $a_1 := b_\perp \alpha_{a\perp}$; pre-multiplying (8) by $\alpha'_{a\perp}$ and inserting $I_p = P_\beta + P_{\beta_\perp}$, one obtains

$$a'_1 \Delta^2 X_t = \omega_{a_1} b' \Delta^2 X_t + \Gamma_{a_1}(P_\beta + P_{\beta_\perp})\Delta X_{t-1} + \Psi_{a_1} K_t + \mu_{a_1} D_t + \epsilon_{a_1 t}.$$

where the coefficients and error term are defined as in (7) with a_1 in place of a. Because of (11) and (13) this equals

$$\begin{aligned} a'_1 \Delta^2 X_t \;&=\; \omega_{a_1} b' \Delta^2 X_t + \xi_1 (\eta' : \eta'_0) \begin{pmatrix} \bar{\beta}'_\perp \Delta X_{t-1} \\ 1 \end{pmatrix} + \gamma (\beta' : \beta'_0) \begin{pmatrix} \Delta X_{t-1} \\ 1 \end{pmatrix} + \\ &\quad + \Psi_{a_1} K_t + \varkappa_{a_1} d_t + \epsilon_{a_1 t}. \end{aligned}$$

where $\gamma := \Gamma_{a_1} \bar{\beta}$. Setting $\eta^* := (\eta' : \eta'_0)'$, one has

$$(a'_1 Z_{0t}) = \omega_{a_1} b' Z_{0t} + \xi_1 \eta^{*\prime} \begin{pmatrix} \bar{\beta}'_\perp \Delta X_{t-1} \\ 1 \end{pmatrix} + \gamma \left(\beta^{*\prime} \Delta X^*_{t-1} \right) + \Psi_{a_1} K_t + \varkappa_{a_1} d_t + \epsilon_{a_1 t}. \quad (16)$$

where variables in parenthesis are observable, given that α, β^*, $\alpha_\perp$, and $\beta_\perp$ are fixed from the previous stage.

The LR tests on the rank of $\xi_1 \eta^{*\prime}$, when the quantities in parenthesis are fixed, is found by RRR, with associated eigenvalue problem

$$\left| \rho S^*_{\beta_\perp \beta_\perp . (\beta, b)} - S^*_{\beta_\perp a_1 . (\beta, b)} S^{*-1}_{a_1 a_1 . (\beta, b)} S^*_{a_1 \beta_\perp . (\beta, b)} \right| = 0, \quad (17)$$

where the subscripts β, $\beta_\perp$, a_1, b refer respectively to the variables $\beta^{*\prime} \Delta X^*_{t-1} := \beta' \Delta X_{t-1} + \beta'_0$, $(\Delta X'_{t-1} \bar{\beta}_\perp : 1)'$, $a'_1 Z_{0t}$, $b' Z_{0t}$ and the moment matrices S^*_{ij} are defined as $S^*_{ij} := M_{ij} - M_{i3} M^{-1}_{33} M_{3j}$, i.e. they are not corrected for the mean.

Let the eigenvalues of (17) be $\rho_1 \geq ... \geq \rho_{m-r} \geq \rho_{m-r+1} = ... = \rho_{p-r+1} = 0$. The associated test statistic within the conditional model for $rank(\xi_1 \eta^{*\prime}) \leq s$ versus $rank(\xi_1 \eta^{*\prime}) \leq m - r$ for $m - r > s$ is

$$Q^{p-m}(s|r) := -T \sum_{i=s+1}^{m-r} \ln(1 - \rho_i). \quad (18)$$

When $s \geq m - r$, $Q^{p-m}(s|r) := 0$. The joint test on the integration indices is based on

$$Q^{p-m}(r, s) := Q^{p-m}(r) + Q^{p-m}(s|r). \quad (19)$$

The limit distribution of $Q^{p-m}(r, s)$ under correct specification is derived in Section 4. In the next section we discuss the properties of the II selection procedure based on $Q^{p-m}(r, s)$.

p_0	p_1+p_2		(p_0,p_1,p_2)			
0	4	$(0,0,p)$	$(0,1,p-1)$	...	$(0,m-1,p_{22}+1)$	$(0,m,p_{22})$
1	3		$(1,0,p-1)$	...	$\vdots$	$(1,m-1,p_{22})$
...	$\vdots$				$\vdots$	$\vdots$
				$\ddots$	$\vdots$	$\vdots$
$m-1$	$p_{22}+1$				$(m-1,0,p_{22}+1)$	$(m-1,1,p_{22})$
	p_2	p	$p-1$		$p_{22}+1$	p_{22}
	p_{21}	m	$m-1$		1	0

Table 1: Values of (p_0,p_1,p_2) in the sequence of tests for the selection procedure, which starts from the upper left corner and proceeds row-wise to the lower right corner, left to right. $p_{22} := p-m$.

3.3 II selection procedure

On the basis of the tests described in the previous subsections, one can construct a selection procedure for the II along the lines of the ones proposed for the full system analysis. Let $p_{22} := p-m$ and $c(p_1,p_2,p_{22})$ be the $(1-\varsigma)$–quantile of the limit distribution of $Q^{p_{22}}(p_0,p_1)$ which is described in the following Theorem 6. Let $R_{rs} := \{Q^{p-m}(r,s) > c(s,p-r-s,p-m)\}$ be the rejection region of the test $Q^{p-m}(r,s)$ based on this limit distribution. A selection procedure analogous to the one performed for the complete system considers the sequence of values for (r,s) reported in Table 1, starting from the upper left hand corner and proceeding row-wise from left to right to the lower right hand corner, see Johansen (1995) or Paruolo (1996).

The selection procedure estimates the II as the values $(\widehat{p}_0,\widehat{p}_1)$ corresponding to the first test that does not imply a rejection in the sequence. If all the tests reject, then $(\widehat{p}_0,\widehat{p}_1)$ is set equal to the most general model against which other submodels are compared, i.e. $(\widehat{p}_0,\widehat{p}_1)=(m,0)$. Formally

$$\{(\widehat{p}_0,\widehat{p}_1)=(r,s)\} = (\cap_{i=0}^{r-1}\cap_{j=0}^{m-i} R_{ij}) \cap_{j=0}^{s-1} R_{rj} \cap R_{rs}^c$$

and $\{(\widehat{p}_0,\widehat{p}_1)=(m,0)\} = \cap_{r=i}^{m-1}\cap_{j=0}^{m-i} R_{ij}$, where $R_\cdot^c$ is the complement of the $R_\cdot$ rejection region.

Note that if all tests reject, then the chosen II are $(p_0,p_1,p_2)=(m,0,p-m)$, i.e. there are at least as many I(2) trends as un-modelled variables. This reflects the fact that the order condition $m \geq p_0+p_1$, i.e. $p-m \leq p_2$, is a maintained assumption. The reference unrestricted model differs from the one of a full system analysis. In the latter case when all tests reject, the selected values for the II is $(p_0,p_1,p_2)=(p,0,0)$, and one concludes that the system is completely stationary. This difference reflects the structure of the tests, where the excluded values of the II in the conditional test sequence are $p_2 = 0,1,...,p-m-1$.

4 Asymptotics under correct specification

Let the I(2) assumptions hold. In this section we discuss the asymptotic distribution of the test $Q^{p-m}(r,s)$ under correct specification of the NF assumption (9), which implies the validity of the order condition $p-m \leq p_2$. Because b is a $p \times p-m$ full column rank matrix, under correct specification of (9) $b \in sp(\alpha_2)$ and one can choose b as part of the basis of $sp(\alpha_2)$, i.e. $\alpha_2 := (\alpha_{21} : \alpha_{22}) := (a_1\xi_{1\perp} : b)$, see Proposition 5.*iv*. We hence define $\alpha_{21} := a_1\xi_{1\perp}$ of dimension $p \times p_{21}$ and $\alpha_{22} := b$ of dimension $p \times p_{22}$, where $p_2 = p_{21} + p_{22}$. In words $p_{22} := p - m$ represents the lower bound of I(2) common trends, while $p_{21} := p_2 - p_{22}$ indicates the extra number of I(2) trends.

4.1 Limit distribution

Let $W(u)$ be a Brownian motion with covariance Ω, which is the weak limit of the partial sums of ϵ_t, $T^{-1/2}\sum_{i=1}^{[Tu]}\epsilon_i \xrightarrow{w} W(u)$, where here and in the following $[\cdot]$ is the integer part. In the following the argument u from processes like $W(u)$ is suppressed unless needed for clarity. Define also

$$
\begin{aligned}
B_1 \ &: = \Omega_{\alpha_1\alpha_1.\alpha_2}^{-1/2}(\alpha_1' - \Omega_{\alpha_1\alpha_2}\Omega_{\alpha_2\alpha_2}^{-1}\alpha_2')W \\
B_2 \ &: = \begin{pmatrix} B_{21} \\ B_{22} \end{pmatrix} := \begin{pmatrix} \Omega_{\alpha_{21}\alpha_{21}.\alpha_{22}}^{-1/2} & -\Omega_{\alpha_{21}\alpha_{21}.\alpha_{22}}^{-1/2}\Omega_{\alpha_{21}\alpha_{22}}\Omega_{\alpha_{22}\alpha_{22}}^{-1} \\ 0 & \Omega_{\alpha_{22}\alpha_{22}}^{-1/2} \end{pmatrix}\begin{pmatrix} \alpha_{21}' \\ \alpha_{22}' \end{pmatrix} W \\
B \ &: = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \qquad B^* := \begin{pmatrix} B_1 \\ B_{21} \end{pmatrix}.
\end{aligned}
$$

Note that $B := (B_1' : B_2')'$ is a $p_1 + p_2$ dimensional standard Brownian motion, where B_1 has dimension p_1, B_2 has dimension p_2, and it is composed of B_{21} of dimension $p_{21} := m - p_0 - p_1$ and of B_{22} of dimension $p_{22} := p - m$. Here we have used the notation $\Omega_{ff} := f'\Omega f$ and $\Omega_{ff.g} := \Omega_{ff} - \Omega_{fg}\Omega_{gg}^{-1}\Omega_{gf}$.

Let $(a(u)|b(u))$ indicate $a(u) - \int_0^1 a(s)b'(s)ds(\int_0^1 b(s)b'(s)ds)^{-1}b(u)$ for continuous time processes and define

$$
\begin{aligned}
F(u) \ &: = \begin{pmatrix} F_1(u) \\ F_2(u) \\ F_3(u) \end{pmatrix} := \left(\begin{array}{c} B_1(u) \\ \int_0^u B_2(s)ds \\ u \end{array} \middle| \, G(u) \right), \\
G(u) \ &: = \begin{pmatrix} G_1(u) \\ G_2(u) \end{pmatrix} := \begin{pmatrix} B_2(u) \\ 1 \end{pmatrix}.
\end{aligned}
$$

Let also

$$h(a, b) := \int (da)b'(\int bb'du)^{-1}\int b(da)',$$

where in the following integrals are understood from 0 to 1 unless otherwise stated. The following theorem holds, where all limits are for $T \to \infty$.

Theorem 6 *Under correct specification of the NF condition* (9)*, the eigenvalue equation* (15) *has asymptotically* p_0 *non-vanishing roots* λ_1*, ...,* λ_{p_0}*; this implies that* $Q^{p_{22}}(j) \to \infty$ *for* $j < p_0$*. Moreover*

$$Q^{p_{22}}(p_0) \xrightarrow{w} Q_\infty^{p_{22}}(p_0) := tr\,(h(B^*, F))\,. \tag{20}$$

Similarly the eigenvalue equation (18) *has asymptotically* p_1 *non-vanishing roots* ρ_1*, ...,* ρ_{p_1}*; this implies that* $Q^{p_{22}}(j|p_0) \to \infty$ *for* $j < p_1$*. Moreover*

$$Q^{p_{22}}(p_1|p_0) \xrightarrow{w} Q_\infty^{p_{22}}(p_1|p_0) := tr\,(h(B_{21}, G))\,. \tag{21}$$

The convergences in (20) *and* (21) *hold jointly, and hence,*

$$Q^{p_{22}}(p_0,p_1) \xrightarrow{w} Q^{p_{22}}_{\infty}(p_0,p_1) := Q^{p_{22}}_{\infty}(p_0) + Q^{p_{22}}_{\infty}(p_1|p_0). \tag{22}$$

Finally $Q^{p_{22}}(i,j)$ *diverges if* $i < p_0$ *and if* $i = p_0$ *and* $j < p_1$.

As in the case for the full system analysis, also the limit distribution of $Q^{p_{22}}(p_0,p_1)$ depend on both p_0 and p_1. Moreover it depends on the dimension $p_{22} := p-m$ of the marginal system. This limit distribution has not been tabulated before. In the next subsection the quantiles of this distribution are approximated by simulation, using a combination of response surface analysis and the Gamma approximation proposed in Doornik (1998) for multivariate unit root distributions.

The fact that $Q^{p_{22}}(i,j)$ diverges if $i < p_0$ and if $i = p_0$ and $j < p_1$ implies the following properties for the estimator $(\widehat{p}_0, \widehat{p}_1)$.

Corollary 7 *Under the same hypotheses as Theorem* 6, *the following holds:*

i. $\Pr((\widehat{p}_0,\widehat{p}_1) = (i,j)) \to 0$ *if* $i < p_0$ *and if* $i = p_0$ *and* $j < p_1$

ii. $\Pr((\widehat{p}_0,\widehat{p}_1) = (p_0,p_1)) \to (1-\varsigma)$, *where* ς *is the asymptotic significance level.*

Thus a smaller value of the II will never be selected in the limit under correct specification; moreover the correct selection will be made 95% of the times with a 5% significance level. Letting the size of the test shrink with sample size, as suggested e.g. in Hendry and Krolzig (2002), would make $(\widehat{p}_0,\widehat{p}_1)$ a consistent estimator.

In the case of incorrect specification, see Section 5 below, all the properties in Corollary 7 may be lost.

4.2 Quantiles of the limit distribution

In order to estimate the quantiles of the limit distribution, a Monte Carlo experiment was performed. We selected $p_2 = 1, ..., 6$, and $p_1 \leq 8 - p_2$, considering all possible cases for $p_{21} := p_2 - p_{22}$, the number of I(2) in the conditional system, $0 \leq p_{21} \leq p_2$. The simulations were performed in Gauss 3.6.

When $p_{21} = p_2$, i.e. $p_{22} = 0$, the conditional and the full systems coincide, and so do the present limit distributions Q^0_{∞} with the ones labelled $S^{\infty}_{r,s}$ in Rahbek et al. (1999). When $p_{21} = 0$, instead, the maintained lower bound p_{22} and the number of I(2) trend coincide, $p_2 = p_{22}$. In this case the relevant test statistic is $Q^{p-m}(p_0, m-p_0) := Q^{p_2}(p_0)$, which corresponds to the last column in Table 1.

Doornik (1998) suggested the use of a Gamma approximation for the limit distributions for CI tests also in the I(2) case, reporting a very good fit. The same approximation was adopted here. This approximation can be summarized as follows. Let y be a Gamma random variable with distribution function

$$\Gamma(u;\kappa_1,\kappa_2) := \frac{\kappa_1^{\kappa_2}}{\Gamma(\kappa_2)} \int_0^u t^{\kappa_2-1} e^{-\kappa_1 t} dt$$

with moments $\varphi_1 := E(y) = \kappa_2/\kappa_1$, $\varphi_2 := V(y) = \kappa_2/\kappa_1^2$, where $\Gamma(\cdot)$ is the Gamma function. Note that $\kappa_1 = \varphi_1/\varphi_2$, $\kappa_2 = \varphi_1^2/\varphi_2$, so that the two parameters κ_1, κ_2 are invertible functions of the mean and variance of the distribution. Monte Carlo is used to estimate the moments φ_1 and φ_2 of the limit distribution, which is then approximated by $\Gamma(y; \varphi_1/\varphi_2, \varphi_1^2/\varphi_2)$.

In order to estimate the first two moments φ_1 and φ_2 of the limit distribution, we designed the following experiment. A set $\mathcal{T}$ of 11 sample sizes T_i was selected, $\mathcal{T} := \mathcal{T}_1 \cup \mathcal{T}_2$, $\mathcal{T}_1 := \{[5000/j], j = 1, .., 6\}$, $\mathcal{T}_2 := \{[500/i], i = 1, .., 5\}$, i.e. $\mathcal{T} := \{100,$ 125, 166, 250, 500, 833, 1000, 1250, 1666, 2500, 5000$\}$. As in Johansen et al. (2000), the sample sizes in $\mathcal{T}_i$ were chosen in order to be equally spaced for T^{-1}, which took values in $\mathcal{T}_2^{-1} := \{i/500,\ i = 1, ..., 5\}$ and $\mathcal{T}_1^{-1} := \{j/5000,\ j = 1, ..., 6\}$. The choice of $\mathcal{T}_1^{-1}$ and $\mathcal{T}_2^{-1}$ allows to have two clusters of equally spaced observations for T^{-1}.

For each Data Generating Process (DGP), we approximated $B(u)$ with a Gaussian random walk $B_{T_i}(u) := T_i^{-1/2} \sum_{t=1}^{[T_i u]} \epsilon_t$, where ϵ_t was drawn from a $N(0, I_{p_1+p_2})$ random number generator. This allowed to construct $B_{T_i}^*$, G_{T_i}, F_{T_i} and $Q_{T_i} := h(B_{T_i}^*, F_{T_i}) + h(B_{21,T_i}, G_{T_i})$ with B_{T_i} in place of B. Let $Q_{T_i}^{(j)}$ indicate the j-th replicate of Q_{T_i} and n_i the number of replication; n_i was equal to 10^6 for all DGPs. We next calculated the average and variance of $Q_{T_i}^{(j)}$ across j, $\widehat{\varphi}_{1,i} := n_i^{-1} \sum_{j=1}^{n_i} Q_{T_i}^{(j)}$, $\widehat{\varphi}_{2,i} := n_i^{-1} \sum_{j=1}^{n_i} (Q_{T_i}^{(j)})^2 - \widehat{\varphi}_{1,i}^2$. The maximum sample size equal to 5000 was selected after some preliminary pilot study, which revealed slow convergence of moments to a limit for high values of p_1 and p_2.

Following MacKinnon et al. (1999), we estimated the following response surface regression of Monte Carlo moments $\widehat{\varphi}_{s,i}$, $s = 1$, 2, on negative powers of sample size T_i,

$$\widehat{\varphi}_{s,i} = \varphi_s + \sum_{t=1}^{u} \theta_{st} T_i^{-\frac{t}{2}} + e_{s,i}, \qquad i = 1, ..., 11 \tag{23}$$

where u was set to 4 and then reduced to 3 because the coefficients of T_i^{-2} were not significant. The regression intercept provides estimates $\widehat{\varphi}_s$ of φ_s. Unlike in previous studies, we considered the non integer power $T_i^{-1/2}$ and $T_i^{-3/2}$ and found them to be significant.[5]

Each regression had 11 observations, 4 regressors, 7 degrees of freedom. Fig. 1 present results for one of the worst regressions, with high values of p_1 and p_2. It can be seen that there is no apparent pattern of heteroskedasticity in the residuals, and that the regressions have a very good fit.

Summary statistics across different regressions (23) are reported in Table 2. It can be seen that both the regression standard error (SE) and the intercept SE for the mean are lower than for the variance. Usually lower values of p_1, p_2 were associated with smaller standard errors. A log-transformation on the moments did not provide any improvement of the fit, and was discarded.[6]

The estimates of κ_1 and κ_2 were obtained as $\widehat{\kappa}_1 = \widehat{\varphi}_1/\widehat{\varphi}_2$, $\widehat{\kappa}_2 = \widehat{\varphi}_1^2/\widehat{\varphi}_2$, and the $(1-\varsigma)$ quantiles of the limit distribution were estimated as $\Gamma^{-1}(1-\varsigma; \widehat{\kappa}_1, \widehat{\kappa}_2)$. Selected quantiles of the limit distribution of $Q_\infty^{p_{22}}(p_0, p_1)$ and $Q_\infty^{p_{22}}(p_0)$ are reported in Appendices A.3 and A.4.

The table in Appendix A.3 reports quantiles of the limit distribution of $Q_\infty^{p_{22}}(p_0, p_1)$. When $p_2 = p_{21}$ the limit distribution $Q_\infty^0(p_0, p_1)$ coincides with the one for the full systems analysis. For sample size $T_i = 500$ close to the one in Rahbek et al. (1999) of 400, we obtained similar results for the moments of the distributions. The table

[5] Previous studies also considered shorter maximal sample sizes, $\max_i T_i = 2000$ or less.

[6] More elaborate response surfaces will be discussed elsewhere, see Paruolo (2003).

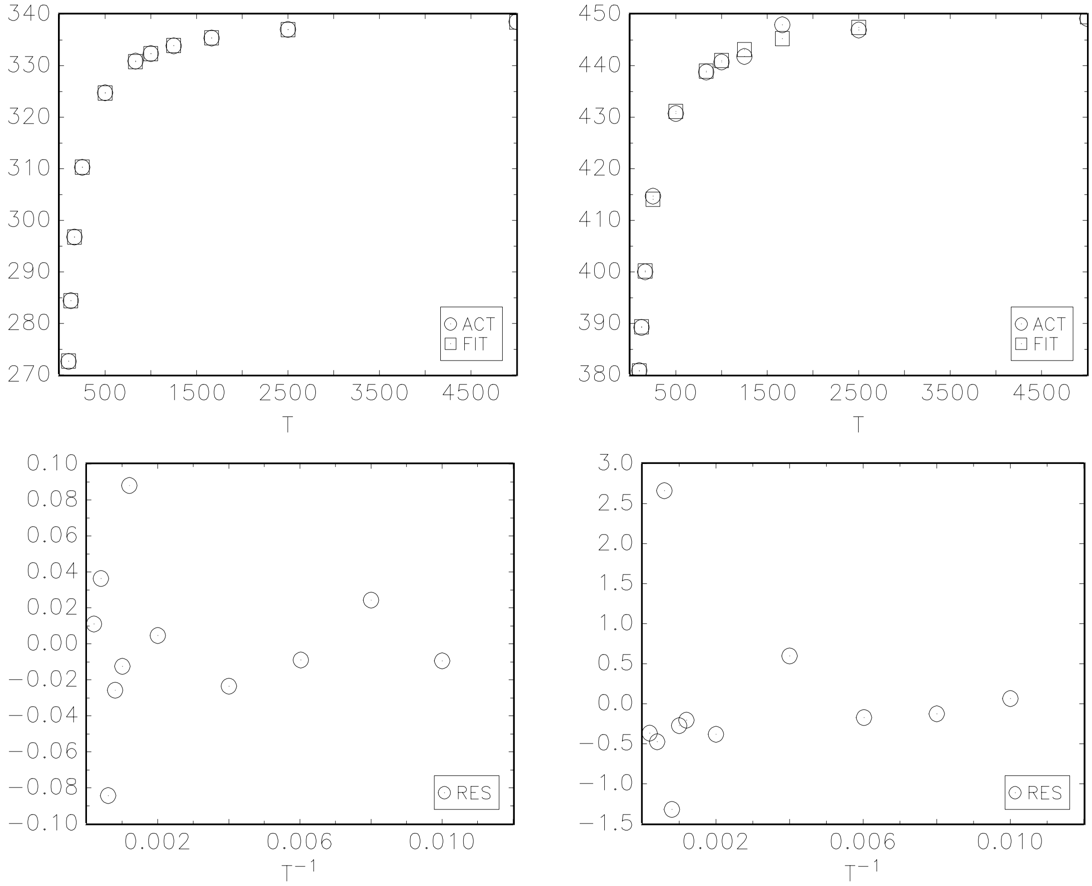

Figure 1: Response surface regressions (23) for φ_1 (upper left and lower left panels) and φ_2 (upper right and lower right panels) for the limit distribution $Q^0_\infty(p_0, p_1)$, $p_1 = 2$, $p_2 = p_{21} = 6$. Upper graphs report actual (ACT) and fitted (FIT) values as a function of T, lower graphs plot residuals (RES) as functions of T^{-1}.

	intercept	standard error	regression	standard error
	φ_1	φ_2	φ_1	φ_2
min	0.019	0.281	0.007	0.109
max	0.176	4.915	0.068	1.912
mean	0.094	1.898	0.036	0.738
stdev	0.035	1.075	0.014	0.418

Table 2: Summary statistics across experiments on residuals of response surface regressions (23) for φ_1 and φ_2 .

in Appendix A.3 provides better estimates of the limit quantiles of $Q^0_\infty(p_0, p_1)$ for the full system analysis than the published ones, because figures are based on more replications and on the response surface (23).

The table in Appendix A.4 lists quantiles the limit distribution $Q^{p_{22}}_\infty(p_0)$, which is relevant when $p_{21} = 0$ and $p_2 = p_{22}$. In all cases, a comparison of the quantiles obtained through the Gamma approximation and the estimated Monte Carlo quantiles revealed a very close fit.

5 Effects of incorrect specification

Consider the parameter space of the full model; let $\mathcal{D}$ indicate this parameter space constrained to satisfy the I(2) assumptions. We discuss several possible failures of the NF conditions (9) by considering classes of processes within $\mathcal{D}$; these classes are indicated by $\mathcal{D}_i \subset \mathcal{D}$, $i = 1, 2, 3$. In the following a superscript c indicates set-complementation with respect to $\mathcal{D}$. In all cases we assume that the order condition $m \geq p_0 + p_1$ is satisfied, because failure of this condition would prevent any procedure from selecting the correct II with positive limit probability.

The classes $\mathcal{D}_i$ are defined as follows.

$$\begin{aligned}
\mathcal{D}_1 &= \mathcal{D} \cap \{\alpha_a \text{ not of full rank } p_0\}, \\
\mathcal{D}_2 &= \mathcal{D}_1^c \cap \{\xi_1 \text{ not of full rank } p_1\}, \\
\mathcal{D}_3 &= \mathcal{D}_1^c \cap \mathcal{D}_2^c \cap \{b'\Gamma\beta \neq 0\}.
\end{aligned} \tag{24}$$

The first class $\mathcal{D}_1$ corresponds to the case in which the chosen equations $(a'X_t)$ are the 'wrong ones' for the first stage, in the sense that α_a is of reduced rank. This case gives very serious implications for the II selection procedure: it is shown that p_0 is underestimated with positive probability in the limit, and that the estimator of $sp(\beta)$ is (quite obviously) inconsistent. This is treated in Subsection 5.1.

The second class considers the situation where there is enough information in the conditional system for the first stage, but not enough for the second stage, in the sense that α_a is of full rank ($\mathcal{D}_1^c$), and ξ_1 is of deficient rank. This case is treated in Subsection 5.2. The last case correponds to the situation where also ξ_1 is of full rank, but the marginal system $b'X_t$ still contains adjustment to the ECM terms in $\beta'\Delta X_t$; this is treated in Subsection 5.3.

5.1 Wrong endogenous variables for integral control ECM

In this subsection we discuss case $\mathcal{D}_1$, where α_a is not of full rank p_0. Although α is of rank p_0, the transformation α_a may result in a singular matrix with rank inferior to p_0. This situation is impossible if $\alpha_b = 0$, because this implies α_a of full rank p_0, see (10). The following example illustrates the case.

Example 8 *Consider the bivariate case $p = 2$ with $p_0 = 1$, and the following values*

$$\alpha = \begin{pmatrix} 1/2 \\ 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad b_\perp = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \Omega = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}.$$

One can easily verify that $\alpha_a = 0$ and $\alpha_b = 1$.

When α_a is of deficient rank, it admits a rank-decomposition of the type $\alpha_a = \alpha_{01}q'$ with matrices α_{01} and q of full column rank $p_{01} < p_0$. The following proposition collects some implications of this type of misspecification for the limit distribution of $Q^{p_{22}}(i)$.

Proposition 9 *Let $m > p_0$ and $rank(\alpha_a) = p_{01} < p_0$. Then $Q^{p_{22}}(j)$ diverges for $j < p_{01}$ and $Q^{p_{22}}(j) \xrightarrow{w} \sum_{i=j-p_{01}+1}^{m-p_{01}} \psi_i$, where $\psi_1 \geq ... \geq \psi_{m-p_{01}}$ are the ordered eigenvalues of the matrix*

$$J := z'z + h(\tilde{B}, F).$$

Here F is defined as in the case of correct specification, $\tilde{B} := (g'\Omega g)^{-1/2} g'W$, $g := (b_\perp - b\omega'_{b_\perp})\alpha_{01\perp}$. The Brownian motions in $\tilde{B}$ and F may have any correlation structure, according to how g and α_1, α_2 are related. Finally z is a $(p_0 - p_{01}) \times (m - p_{01})$ matrix of standard normal variates independent of $h(\tilde{B}, F)$. In particular $Q^{p_{22}}(p_{01}) \xrightarrow{w} tr(J) = \chi^2((p_0 - p_{01})(m - p_{01})) + tr(h(\tilde{B}, F))$, where the χ^2 and the multivariate Dickey-Fuller type components $tr(h(\tilde{B}, F))$ are independent.

This proposition implies that there is positive limit probability to select smaller values of p_0, between p_{01} and p_0, and Corollary 7.1 does not hold for $p_{01} \leq i < p_0$. This follows from the fact that $Q^{p_{22}}(i)$ does not diverge in this case. Moreover the limit distribution of $Q^{p_{22}}(p_{01})$ presents nuisance parameters, the correlation between the Brownian motions $\tilde{B}$ and the Brownian motions B on which F is built. These correlations depend on the relation among $g := (b_\perp - b\omega'_{b_\perp})\alpha_{01\perp}$ and α_1, α_2.

The limit distribution of $Q^{p_{22}}(p_{01})$ is also expected to be shifted to the right with respect to the one of $Q^{p_{22}}(p_0)$ under correct specification. This is because of the extra χ^2 component and because the number of Brownian motions in $\tilde{B}$ is $m - p_{01}$ instead of $m - p_0$ as in B^* for case of correct specification, see Theorem 6.[7]

It can be shown that the estimate of α_{01} and $\beta^*_{01} := \beta^* q$ from the first stage are $T^{1/2}$ and T consistent when one selects p_{01} components from the first stage. For this choice the following proposition states implication on $Q^{p_{22}}(j|p_{01})$ in the second stage.

Proposition 10 *Let $m > p_0 + p_1$ and $rank(\alpha_a) = p_{01} < p_0$. Then $Q^{p_{22}}(j|p_{01})$ diverges for j less than the rank s of the matrix*

$$((I_{m-p_{01}} : 0)\xi\eta' : \alpha'_{01\perp}\Gamma_a\bar{\beta}_{02}), \tag{25}$$

where both $0 \leq s \leq m - p_{01}$ and $0 \leq p_1 \leq m - p_{01}$.

The previous proposition clarifies that many possible situations may arise according to the rank of the matrix (25). If in particular $s > p_1$, the II selection procedure would overestimate p_1 with limit probability one, even when the first integration index has been chosen equal to $rank(\alpha_a) = p_{01}$, the number of ECM terms present in the conditional system.

Summarizing, if the selected equations in the conditional system do not have a full rank loading matrix on the polynomial CI disequilibrium errors, the procedure selects a smaller value of polynomial CI relations than the true value with positive limit probability. In this case there is no control on the limit probability of correct selection of p_{01} because the relevant limit distribution is different from the case of correct specification of NF.

[7] Since the null limit distribution and the one under this type of misspecification also differ because of non unit correlation between the components in $\tilde{B}$ and in B, this expectation may not hold for some particular correlation stuctures.

5.2 Wrong endogenous variables for proportional control ECM

In this subsection we analyze case $\mathcal{D}_2$, where α_a is of full rank p_0 and ξ_1 is not of full rank p_1, under the order condition $m \geq p_0 + p_1$. Recall that $\alpha_\perp' \Gamma \beta_\perp = \xi\eta'$, with ξ, η matrices of dimension $(p - p_0) \times p_1$ of full rank p_1. Recall also that by (10), $a_1 := \alpha_\perp (I_{m-p_0} : 0)'$, and that hence $\xi_1 := (I_{m-p_0} : 0)\xi$. Although ξ is of rank p_1, the transformation $(I_{m-p_0} : 0)$ may select a singular block with rank inferior to p_1, as the following example shows.

Example 11 *Consider the case $m - p_0 = 1$, and $\xi = (0 : 1/2)'$. One can easily verify that $\xi_1 = 0$ and $\xi_2 := (0 : 1)\xi = 1/2$.*

We recall that if the NF condition $b'\alpha_1 = 0$ holds, then ξ_1 must have full rank, see Proposition 5. In general however, ξ_1 may have rank $p_{11} < p_1$; in this case it admits a representation of the type $\xi_1 = \xi_{11} q'$ with matrices ξ_{11} and q of full column rank p_{11}.

The following proposition collects some implications of this type of misspecification for the limit distribution of $Q^{p_{22}}(i|p_0)$. Recall that if the first stage is correctly specified, i.e. $rank(\alpha_a) = p_0$, then the estimates of α and β from the first stage are $T^{1/2}$ and T consistent respectively.

Proposition 12 *Let $m > p_0 + p_1$, $rank(\alpha_a) = p_0$ and $rank(\xi_1) = p_{11} < p_1$. Then $Q^{p_{22}}(j|p_0)$ diverges for $j < p_{11}$ and $Q^{p_{22}}(j|p_0) \xrightarrow{w} \sum_{i=j-p_{11}+1}^{m-p_0-p_{11}} \psi_i$, where $\psi_1 \geq ... \geq \psi_{m-p_0-p_{11}}$ are the eigenvalues of the matrix*

$$J := z'z + h(\tilde{B}, G).$$

Here G is defined as in the case of correct specification, $\breve{B} := (\ell'\Omega\ell)^{-1/2}\ell' W$, $\ell := a_1 \xi_{11\perp}$. The Brownian motions in $\breve{B}$ and G may have any correlation structure, according to how ℓ and α_1, α_2 are related. Finally z is a $(p_1 - p_{11}) \times (m - p_0 - p_{11})$ matrix of standard normal variates independent of $h(\tilde{B}, G)$. In particular $Q^{p_{22}}(p_{11}|p_0) \xrightarrow{w} tr(J) = \chi^2((p_1 - p_{11})(m - p_0 - p_{11})) + tr(h(\breve{B}, G))$, where the χ^2 and the multivariate Dickey-Fuller type components $tr(h(\breve{B}, G))$ are independent.

Again we observe that there is positive limit probability to select smaller values of p_1, between p_{11} and p_1, i.e. Corollary 7.1 does not hold for $p_{11} \leq j < p_1$. This follows from the fact that $Q^{p_{22}}(j|p_0)$ does not diverge in this case.

The limit distribution of $Q^{p_{22}}(p_{11}|p_0)$ depends on nuisance parameters, the correlations between $\breve{B}$ and the Brownian motions that enter G; these correlation may well be different from 0 or 1 as in the case of correct specification.

The limit distribution of $Q^{p_{22}}(p_{11}|p_0)$ is expected to be shifted to the right with respect to the one of $Q^{p_{22}}(p_1|p_0)$ under correct specification. This is because of the extra χ^2 component and because the number of Brownian motions in $\breve{B}$ is $m - p_0 - p_{11}$ instead of $m - p_0 - p_1$ as in B^* for case of correct specification, see Theorem 6.[8]

The same final remark as in Section 5.1 applies here: if the selected equations in the conditional system do not have a full rank loading matrix on $\beta_1' \Delta X_{t-1}$, $\widehat{p}_1$ will select a smaller value than the true value p_1 with positive limit probability.

[8] Since the null limit distribution and the one under this type of mispecification also differ because of non unit correlation between the components in $\tilde{B}$ and in B, this expectation may not hold for some particular correlation structure.

p_0	p_1+p_2	$Q^1(p_0,p_1)$			$Q^1(p_0)$
0	4	173.4 (109.7)	123.4 (88.1)	89.7 (70.7)	84.5 (57.0)
1	3		91.4 (63.1)	**43.6** (47.6)	39.3 (35.7)
2	2			21.7 (28.1)	13.8 (18.1)
	p_2	4	3	2	1
	p_{21}	3	2	1	0

Table 3: $Q^1(p_0,p_1)$ statistics for the UK money data, $p=4$, $m=3$, $p_{22}=p-m=1$ $p_{21}:=3-p_0-p_1$. Numbers in parenthesis are the 5% asymptotic critical values, taken from the Appendix. The bold entry corresponds to the first test that does not reject.

5.3 Adjustment to changes of integral control ECM in the marginal equations

In this subsection we discuss case 3, $\mathcal{D}_3$, where $b'\Gamma\beta \neq 0$, when both α_a and ξ_1 have full rank. It is simple to observe that the proof of Theorem 6 does not require the condition $b'\Gamma\beta = 0$ to hold, and hence this type of misspecification has no effect on the asymptotics. This is parallel to the findings in Paruolo and Rahbek (1999) for the aspect of estimation efficiency.

6 An empirical illustration

In order to illustrate the above procedures we considered the data on UK money, income, prices and interest rates analyzed in Doornik et al. (1998) and Rahbek et al. (1999) among others. The quarterly seasonally adjusted data cover the period 1963:1 to 1989:2, and include the log of nominal money (M1), indicated with $m1_t$, the log of total final expenditure at 1985 prices, inc_t, the log of the implicit price deflator, pr_t, and a measure of the opportunity cost of holding money given by R_t^*, the difference between the 3-month local authority interest rate and the learning-adjusted retail sight-deposit interest rate. The data consists of the time series of $X_t := (m1_t : pr_t : inc_t : R_t^*)'$. For ease of comparison with previous analyses, the lag length has been set equal to $k=5$ and the period 1964:3 to 1989:2 has been selected as the estimation period, for a total of 100 effective observations.

We select the marginal equation for R_t^* and the conditional system of $(m1_t : pr_t : inc_t)'$ conditional on R_t^*; hence $p=4$, $m=3$, $p-m=1$. The choice of R_t^* for the marginal system is consistent with earlier empirical analysis of the data and the economic intuition that the interest rates may be driven by international factors, which are external to the UK money market.

The 2SI2 analysis proposed in this paper was then applied to the data, obtaining the results in Table 3. It can be seen that the II selected by the conditional procedure leads to $(\widehat{p}_0, \widehat{p}_1, \widehat{p}_2) = (1,1,2)$. The same selection has been obtained in the full model analysis in Rahbek et al. (1999) and in Paruolo (1996), who however considered a model with no linear trend in the polynomial CI relations.

We next tested the hypothesis that $\beta = (1 : -1 : -1 : \star)'$ and $\beta_0 = 0$, where $\star$ indicates a generic coefficient; this hypothesis imposes the coefficients predicted by a relation involving inverse velocity of money circulation and no trend in the (multi)-CI relation. The RRR test of this hypothesis in the first stage of the 2SI2 analysis was equal to 1.5235, which compared with a $\chi^2(3)$ produced an asymptotic p-value equal to 0.6769, thus giving ample support to the hypothesis. The model thus reduces to the one analyzed in Paruolo (1996). For the analysis of serial correlation common features of this system we refer to Paruolo (2002b).

7 Summary and conclusions

In this paper we have illustrated how the 2SI2 procedure can be modified to address inference on the integration indices in conditional systems, under the assumption of no integral and proportional feedback to the marginal system (NF), which is also a weak exogeneity condition with respect to the cointegration parameters. The limit distributions have been derived; they are free of nuisance parameters. Their quantiles have been tabulated by a combination of response surface and of the Gamma approximation.

The effects of possible misspecification of the NF condition have also been investigated. It is found that various types of misspecification have implications of varying degree of importance on the properties of the selection procedure of II. Misspecification may lead to a limit distribution which is not free of nuisance parameters, and even more importantly, to the selection of fewer (multi)cointegrating relations with positive limit probability.

Moreover, similarly to the discussion of efficiency of estimation performed in Paruolo and Rahbek (1999), if the adjustment matrices of the conditional system are of full rank, the asymptotic distributions are unaffected by misspecification of the NF condition.

A Appendices

In the appendices we report proofs on the I(2) representation theorem in Appendix A.1; Appendix A.2 collects proofs of the asymptotic results, while Appendices A.3, A.4 list quantiles of the limit distributions.

A.1 I(2) representation

Proof. of proposition 1. Consider $A(L)X_t = e_t$ with $e_t := \mu_0 + \mu_1 t + \varkappa d_t + \epsilon_t$. We wish to show that under the assumption that the AR polynomial $A(z)$ has stable roots or roots at $z = 1$, necessary and sufficient conditions for X_t to be I(2) with no cubic or quadratic trends are the following ones, which are a restatement of conditions $I(2)\_b$ to f:

$$\begin{array}{ll} b) & A(1) = -\alpha\beta' \\ c) & P_{\alpha_\perp} A^1(1) P_{\beta_\perp} = \alpha_1 \beta_1' \\ d) & rank(\alpha_2'\theta\beta_2) = p - p_0 - p_1 \\ e) & \mu_1 = \alpha\beta_0' \\ f) & \bar{\alpha}_2'\mu_0 = \bar{\alpha}_2' A^1(1)\bar{\beta}\beta_0'. \end{array} \qquad (26)$$

Here and in the following $A^1(z) := dA(z)/dz = -\Pi + \Gamma$, and more in general $A^j := (-1)^{j+1}\frac{1}{j!}d^jA(z)/dz^j|_{z=1}$. Recall also that $\alpha_2 = (\alpha : \alpha_1)_\perp$, $\beta_2 = (\beta : \beta_1)_\perp$.

Johansen (1996) has shown that under $I(2)\_a$, $b)$ and $c)$ are necessary and sufficient for X_t to be at least $I(2)$, and that $b)$, $c)$ and $d)$ are necessary and sufficient for X_t to be $I(2)$. We here repeat the argument of the proof in Johansen (1992) to find full expressions for the matrix C_1, and give the proof that $I(2)\_e$ and f are necessary and sufficient for X_t not to contain quadratic or cubic trends.

Expand the AR polynomial $A(z)$ around $z = 1$, $A(L) = -(\alpha\beta' + A^1\Delta + A^2\Delta^2 + A^3\Delta^3) + A^*(L)\Delta^4$; $A^*(L)$ is a remainder polynomial of degree $k-4$. Define also $N := (\alpha : \alpha_1 : \alpha_2)$, $J := (\beta : \beta_1 : \beta_2)$. Rewrite the AR equations as $\bar{N}'A(L)\bar{J}J'X_t = \bar{N}'e_t$. Define $(X'_{0t} : X'_{1t} : X'_{2t})' := J'X_t$ and let $A^h_{ij} := \bar{\alpha}'_i A^h \bar{\beta}_j$, where the subscripts i and j refer to $\bar{\alpha}_i$ and $\bar{\beta}_j$, $i, j = 0, 1, 2$; here the subscript 0 is associated with α and β. Note that by (26.c) $A^1_{11} = I$ while A^1_{12} and $A^1_{21}A^1_{22}$ vanish. With this notation one finds, by (26.c)

$$\begin{aligned} -\bar{N}'A(L)\bar{J} &= \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} A^1_{00} & A^1_{01} & A^1_{02} \\ A^1_{10} & I & 0 \\ A^1_{20} & 0 & 0 \end{pmatrix}\Delta + \\ &+ \sum_{i=2}^{3} \begin{pmatrix} A^i_{00} & A^i_{01} & A^i_{02} \\ A^i_{10} & A^i_{11} & A^i_{12} \\ A^i_{20} & A^i_{21} & A^i_{22} \end{pmatrix}\Delta^i - A^*(L)\Delta^4 \end{aligned}$$

Note that the lowest power of Δ at which X_{it} appears in the equations is i, for $i = 0, 1$; X_{2t} appears in first differences in the first set of equations. Hence, rewriting equations in terms of $\breve{X}_t := (X'_{0t} : \Delta X'_{1t} : \Delta X'_{2t})$ one finds the AR representation $B(L)Z_t = \bar{N}'e_t$ where

$$\begin{aligned} -B(L) &= \begin{pmatrix} I & A^1_{01} & A^1_{02} \\ 0 & I & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} A^1_{00} & A^2_{01} & A^2_{02} \\ A^1_{10} & A^2_{11} & A^2_{12} \\ A^1_{20} & A^2_{21} & A^2_{22} \end{pmatrix}\Delta \\ &+ \begin{pmatrix} A^2_{00} & A^3_{01} & A^3_{02} \\ A^2_{10} & A^3_{11} & A^3_{12} \\ A^2_{20} & A^3_{21} & A^3_{22} \end{pmatrix}\Delta^2 - B^*(L)\Delta^3 \end{aligned}$$

and $B^*(L)$ is a remainder polynomial. Define next the new variables $\mathring{X}_t$,

$$\begin{aligned} \mathring{X}_t \quad : &= \begin{pmatrix} Y_t \\ \Delta X_{1t} \\ \Delta X_{2t} \end{pmatrix} := L\breve{X}_t := \begin{pmatrix} I & 0 & A^1_{02} \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} X_0 \\ \Delta X'_{1t} \\ \Delta X'_{2t} \end{pmatrix} \text{ where} \\ L^{-1} &= \begin{pmatrix} I & 0 & -A^1_{02} \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix}. \end{aligned}$$

Inserting $I = L^{-1}L$ in $B(L)\breve{X}_t$ one finds $B(L)L^{-1}L\breve{X}_t =: J(L)\mathring{X}_t$, where

$$-J(L) = \begin{pmatrix} I & A^1_{01} & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{pmatrix} + \sum_{j=1}^{2} \begin{pmatrix} A^j_{00} & A^{j+1}_{01} & A^{j+1.1}_{02} \\ A^j_{10} & A^{j+1}_{11} & A^{j+1.1}_{12} \\ A^j_{20} & A^{j+1}_{21} & A^{j+1.1}_{22} \end{pmatrix}\Delta^j - J^*(L)\Delta^3.$$

Here we have used the notation $A_{i2}^{j+1.1} := A_{i2}^{j+1} - A_{i0}^{j} A_{02}^{1}$ and $J^{*}(L)$ is a remainder polynomial. The variable ΔX_{2t} enters the equations only in first differences, and hence one can define yet another set of variables $\widetilde{X}_t := (Y_t' : \Delta X_{1t}' : \Delta^2 X_{2t}')'$ which satisfy the AR equations $D(L)\widetilde{X}_t = \bar{N}' e_t$ with

$$
\begin{aligned}
-D(L) \quad &: = -D(1) + \dot{D}(1)\Delta - D^{*}(L)\Delta^2 = \\
&= \begin{pmatrix} I & A_{01}^{1} & A_{02}^{2.1} \\ 0 & I & A_{12}^{2.1} \\ 0 & 0 & A_{22}^{2.1} \end{pmatrix} + \begin{pmatrix} A_{00}^{1} & A_{01}^{2} & A_{02}^{3.1} \\ A_{10}^{1} & A_{11}^{2} & A_{12}^{3.1} \\ A_{20}^{1} & A_{21}^{2} & A_{22}^{3.1} \end{pmatrix} \Delta - D^{*}(L)\Delta^2
\end{aligned}
$$

and $D^{*}(L)$ is a remainder polynomial. Note that $-A_{ij}^{2.1} = \theta_{ij} := \bar{\alpha}_i' \theta \bar{\beta}_j$. The matrix $D(1)$ is upper block triangular, and it is invertible iff θ_{22} is of full rank. In this case $\widetilde{X}_t = D^{-1}(L)\bar{N}' e_t$. Expand $F(z) := D^{-1}(z)$ around $z = 1$ as $F(z) = F(1) - \dot{F}(1)\Delta + F^{*}(L)\Delta^2$, where $\dot{F}(z) := dF(z)/dz = dD^{-1}(z)/dz = -D^{-1}(z)\dot{D}(z)D^{-1}(z)$ and $\dot{D}(z) := dD(z)/dz$. Substituting one finds $F(z) = D^{-1}(1) + D^{-1}(1)\dot{D}(1)D^{-1}(1)\Delta + F^{*}(L)\Delta^2$, and hence

$$
\begin{aligned}
\widetilde{X}_t &= D^{-1}(1)\bar{N}' e_t + D^{-1}(1)\dot{D}(1)D^{-1}(1)\bar{N}'\Delta e_t + F^{*}(L)\bar{N}'\Delta^2 e_t \\
&= : (H^0 + H^1\Delta + H^{*}(L)\Delta^2) e_t.
\end{aligned}
$$

Let H_0^j, H_1^j, H_2^j indicate the blocks of rows in H^j that correspond to $\bar{\alpha}$, $\bar{\alpha}_1$, $\bar{\alpha}_2$, respectively. Observe that

$$
\begin{aligned}
D(1) &= \begin{pmatrix} -I & -A_{01}^{1} & \theta_{02} \\ 0 & -I & \theta_{12} \\ 0 & 0 & \theta_{22} \end{pmatrix} \qquad D^{-1}(1) = \begin{pmatrix} -I & A_{01}^{1} & (\theta_{02} - A_{01}^{1}\theta_{12})\theta_{22}^{-1} \\ 0 & -I & \theta_{12}\theta_{22}^{-1} \\ 0 & 0 & \theta_{22}^{-1} \end{pmatrix} \\
H^0 \quad : &= \begin{pmatrix} H_0^0 \\ H_1^0 \\ H_2^0 \end{pmatrix} := D^{-1}(1)\bar{N}' = \begin{pmatrix} -\bar{\alpha}' + A_{01}^{1}\bar{\alpha}_1' + (\theta_{02} - A_{01}^{1}\theta_{12})\theta_{22}^{-1}\bar{\alpha}_2' \\ -\bar{\alpha}_1' + \theta_{12}\theta_{22}^{-1}\bar{\alpha}_2' \\ \theta_{22}^{-1}\bar{\alpha}_2' \end{pmatrix} \\
H_1^1 &= (-A_{10}^{1} + \theta_{12}\theta_{22}^{-1}A_{20}^{1})(-\bar{\alpha}' + A_{01}^{1}\bar{\alpha}_1' + (\theta_{02} - A_{01}^{1}\theta_{12})\theta_{22}^{-1}\bar{\alpha}_2') + \\
&\quad + (-A_{11}^{2} + \theta_{12}\theta_{22}^{-1}A_{21}^{2})(-\bar{\alpha}_1' + \theta_{12}\theta_{22}^{-1}\bar{\alpha}_2') + (-A_{12}^{3.1} + \theta_{12}\theta_{22}^{-1}A_{22}^{3.1})\theta_{22}^{-1}\bar{\alpha}_2' \\
H_2^1 &= \theta_{22}^{-1}(A_{20}^{1}(-\bar{\alpha}' + A_{01}^{1}\bar{\alpha}_1' + (\theta_{02} - A_{01}^{1}\theta_{12})\theta_{22}^{-1}\bar{\alpha}_2') + A_{21}^{2}(-\bar{\alpha}_1' + \theta_{12}\theta_{22}^{-1}\bar{\alpha}_2') + \\
&\quad + A_{22}^{3.1}\theta_{22}^{-1}\bar{\alpha}_2' \\
&= -\theta_{22}^{-1}A_{20}^{1}\bar{\alpha}' + \theta_{22}^{-1}(A_{20}^{1}A_{01}^{1} - A_{21}^{2})\bar{\alpha}_1' + \\
&\quad + \theta_{22}^{-1}(A_{20}^{1}(\theta_{02} - A_{01}^{1}\theta_{12}) + A_{21}^{2}\theta_{12} + A_{22}^{3.1})\theta_{22}^{-1}\bar{\alpha}_2'
\end{aligned}
$$

where we have reported only the last 2 set of rows H_1^1 and H_2^1 of

$$
H^1 := D^{-1}(1)\dot{D}(1)D^{-1}(1)\bar{N}'
$$

for brevity. In order to find the common trends representation, note that $X_t = \bar{\beta}X_{0t} + \bar{\beta}_1 X_{1t} + \bar{\beta}_2 X_{2t}$, or, in terms of the elements of $\widetilde{X}_t$,

$$
\begin{aligned}
X_t &= \bar{\beta}Y_t - \bar{\beta}A_{02}^{1}\Delta X_{2t} + \bar{\beta}_1 X_{1t} + \bar{\beta}_2 X_{2t} = \\
&= \bar{\beta}Y_t - \bar{\beta}A_{02}^{1}(\sum_{i=1}^{t}\Delta^2 X_{2i} + \Delta X_{20}) + \bar{\beta}_1(\sum_{i=1}^{t}\Delta X_{1i} + X_{10}) + \\
&\quad + \bar{\beta}_2(\sum_{j=1}^{t}\sum_{i=1}^{j}\Delta^2 X_{2i} + t\Delta X_{20} + X_{20})
\end{aligned}
$$

so that

$$\begin{aligned} X_t &= \bar\beta(H_0^0+H_0^1\Delta)e_t-\bar\beta A_{02}^1\sum_{i=1}^t(H_2^0+H_2^1\Delta)e_i+\bar\beta_1\sum_{i=1}^t(H_1^0+H_1^1\Delta)e_i+ \\ &\quad +\bar\beta_2\sum_{j=1}^t\sum_{i=1}^j(H_2^0+H_2^1\Delta)e_i+C^*(L)e_t+A+Bt= \\ &=: C_2\sum_{j=1}^t\sum_{i=1}^j e_i+C_1\sum_{i=1}^t e_i+C_0e_t+C^*(L)\Delta e_t \end{aligned}$$

where

$$\begin{aligned} C_2 &:= \bar\beta_2H_2^0=\bar\beta_2\theta_{22}^{-1}\bar\alpha_2' \\ C_1 &:= -\bar\beta A_{02}^1H_2^0+\bar\beta_1H_1^0+\bar\beta_2H_2^1=-\bar\beta A_{02}^1\theta_{22}^{-1}\bar\alpha_2'+\bar\beta_1(-\bar\alpha_1'+\theta_{12}\theta_{22}^{-1}\bar\alpha_2')+ \\ &\quad +\bar\beta_2\theta_{22}^{-1}(-A_{20}^1\bar\alpha'+(A_{20}^1A_{01}^1-A_{21}^2)\bar\alpha_1'+ \\ &\quad +(A_{20}^1(\theta_{02}-A_{01}^1\theta_{12})+A_{21}^2\theta_{12}+A_{22}^{3.1})\theta_{22}^{-1}\bar\alpha_2'), \\ C_0 &:= \bar\beta H_0^0-\bar\beta A_{02}^1H_2^1+\bar\beta_1H_1^1+\bar\beta_2H_2^2 \end{aligned} \tag{27}$$

$C^*(L)$ has exponentially decreasing coefficients and A, B depend on initial conditions that satisfy $(\beta:\beta_1)'B=0$, $\beta'A+\bar\alpha'\Gamma\bar\beta_2\beta_2'B=0$.

We now turn to the analysis of the deterministic components. Observe that

$$\begin{aligned} \Delta(\mu_0+\mu_1t+\varkappa d_t) &= \mu_1+\varkappa\Delta d_t, \\ \sum_{i=1}^t(\mu_0+\mu_1 i+\varkappa d_i) &= \frac{1}{2}\mu_1t^2+(\mu_0+\frac{1}{2}\mu_1)t+\varkappa\sum_{i=1}^t d_i =: f_2t^2+f_1t+\varkappa\sum_{i=1}^t d_i \\ \sum_{j=1}^t\sum_{i=1}^j(\mu_0+\mu_1 i+\varkappa d_i) &= \frac{1}{6}\mu_1t^3+\frac{1}{2}(\mu_0+\mu_1)t^2+(\frac{1}{2}\mu_0+\frac{1}{3}\mu_1+\varkappa\iota(\frac{1}{2}-\frac{1}{2n}))t+\varkappa r_t \\ &=: g_3t^3+g_2t^2+g_1t+\varkappa r_t \end{aligned}$$

where $\|\Delta d_t\|$, $\left\|\sum_{i=1}^t d_i\right\|$ and $\|r_t\|$ are all bounded by $2p$ and ι is a vector of ones, where $\|\cdot\|$ is Euclidean norm. Here $r_t:=\sum_{j=1}^t\sum_{i=1}^j d_i-\iota(\frac{1}{2}-\frac{1}{2n})t$. Note that the dummies generate a linear trend when cumulated twice, but otherwise generate bounded deterministic components.

Hence

$$X_t=C_2\sum_{j=1}^t\sum_{i=1}^j\epsilon_i+C_1\sum_{i=1}^t\epsilon_i+C_0(L)\epsilon_t+m_3t^3+m_2t^2+m_1t+m^*(L)D_t$$

where

$$m_3:=C_2g_3 \qquad m_2:=C_2g_2+C_1f_2 \qquad m_1:=C_2g_1+C_1f_1+C_0\mu_1.$$

The series $m^*(z)$ has exponentially decreasing coefficients. Because D_t is periodic with period n, one can collect coefficients of all lags with $t \bmod n=j$ for $j=1$, to $n-1$ (which are summable) and substitute $m^*(L)$ with a polynomial indicated here as $m(L)$, $m_0:=m(1)$.

Thus there is no cubic trend in the system iff $C_2 g_3 = \frac{1}{6} C_2 \mu_1 = 0$, i.e. iff $\alpha_2' \mu_1 = 0$. Similarly the quadratic trend is absent iff $C_2 g_2 + C_1 f_2 = 0$, i.e. $C_2(\mu_0 + \mu_1) = -C_1 \mu_1$ or $-C_2 \mu_1 = C_1 \mu_1 + C_2 \mu_0$, i.e.,

$$C_1 \mu_1 = -C_2 \mu_0 \tag{28}$$

Multiplying this equation by β' one obtains $\alpha_2' \mu_1 = 0$, which is identical to the no-cubic-trend condition. Observe that, substituting $\alpha_2' \mu_1 = 0$, one finds

$$\begin{array}{rcl} C_1 \mu_1 & = & (-\bar{\beta}_1 \bar{\alpha}_1' + \bar{\beta}_2 \theta_{22}^{-1} (-A_{20}^1 \bar{\alpha}' + (A_{20}^1 A_{01}^1 - A_{21}^2) \bar{\alpha}_1') \mu_1 \\ C_2 \mu_0 & = & \bar{\beta}_2 \theta_{22}^{-1} \bar{\alpha}_2' \mu_0 \end{array}$$

Multiplying eq. (28) by $(\beta : \beta_1 : \beta_2)'$ one has

$$\begin{pmatrix} 0 \\ -\bar{\alpha}_1' \mu_1 \\ \theta_{22}^{-1} (-A_{20}^1 \bar{\alpha}' + (A_{20}^1 A_{01}^1 - A_{21}^2) \bar{\alpha}_1') \mu_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -\theta_{22}^{-1} \bar{\alpha}_2' \mu_0 \end{pmatrix}$$

where the second set of equations is equivalent to $\alpha_1' \mu_1 = 0$; substituting in the last set of equations one finds $A_{20}^1 \bar{\alpha}' \mu_1 = \bar{\alpha}_2' \mu_0$, condition $(26.f)$. This shows that in order for X_t not to contain cubic or quadratic trends the conditions $(26.e)$ and $(26.f)$ are necessary and sufficient.

It can now be checked that $\beta' X_t$ has trend coefficient $\beta' m_1 = -\beta_0'$; in fact

$$\begin{array}{rcl} \beta' m_1 & = & \beta' C_1 f_1 + \beta' C_0 \mu_1 = -A_{02}^1 \theta_{22}^{-1} \bar{\alpha}_2' (\mu_0 + \dfrac{1}{2} \mu_1) + H_0^0 \mu_1 - A_{02}^1 H_2^1 \mu_1 = \\ & = & -A_{02}^1 \theta_{22}^{-1} A_{20}^1 \beta_0' - \beta_0' + A_{02}^1 \theta_{22}^{-1} A_{20}^1 \beta_0' = -\beta_0' \end{array}$$

where we have used $\mu_1 = \alpha \beta_0'$ and $A_{20}^1 \bar{\alpha}' \mu_1 = \bar{\alpha}_2' \mu_0$. Similarly the trend coefficient of $\beta_1' X_t$ is

$$\begin{array}{rcl} \beta_1' m_1 & = & \beta_1' (C_1 (\mu_0 + \dfrac{1}{2} \mu_1) + C_0 \mu_1) = (-\bar{\alpha}_1' + \theta_{12} \theta_{22}^{-1} \bar{\alpha}_2')(\mu_0 + \dfrac{1}{2} \mu_1) + H_1^1 \mu_1 = \\ & = & -\bar{\alpha}_1' \mu_0 + \theta_{12} \theta_{22}^{-1} A_{20}^1 \beta_0' - (-A_{10}^1 + \theta_{12} \theta_{22}^{-1} A_{20}^1) \beta_0' = \\ & = & -\bar{\alpha}_1' (\mu_0 - A^1 \bar{\beta} \beta_0') =: -\eta_0'. \end{array} \tag{29}$$

We wish to show that $(26.f)$ plus (29) is equivalent to the condition $I(2)\_f$. Stacking $(26.f)$ and (29) together one has

$$(\bar{\alpha}_1 : \bar{\alpha}_2)' (\mu_0 - A^1 \bar{\beta} \beta_0') =: \begin{pmatrix} \eta_0' \\ 0 \end{pmatrix}. \tag{30}$$

Next note that $(\alpha_1 : \alpha_2) = \alpha_\perp \varsigma$, for some nonsingular ς, so that $(\bar{\alpha}_1 : \bar{\alpha}_2) = \alpha_\perp \varsigma (\varsigma' \alpha_\perp' \alpha_\perp \varsigma)^{-1}$. Eq. (30) can thus we written as $\varsigma' \alpha_\perp' (\mu_0 - A^1 \bar{\beta} \beta_0') = \varsigma' \alpha_\perp' \alpha_1 \eta_0'$. Finally recall that $\xi := \alpha_\perp' \alpha_1$ and pre-multiply the previous equation by ς'^{-1}; one finds $\alpha_\perp' (\mu_0 - A^1 \bar{\beta} \beta_0') = \xi \eta_0'$, which is condition $I(2)\_f$. The vice versa is similar. ■

Proof. of Proposition 4. We have chosen $a = b_\perp + b a_0$. Substituting in $a' - \omega_a b'$ one finds

$$\begin{array}{rcl} a' - \omega_a b' & = & a'(I - \Omega b (b' \Omega b)^{-1} b') = (b_\perp' + a_0' b')(I - \Omega b (b' \Omega b)^{-1} b') = \\ & = & b_\perp' (I - \Omega b (b' \Omega b)^{-1} b') = b_\perp' - \omega_{b_\perp} b'. \end{array}$$

Symbol	definition	symbol	definition
$Z_{0t} :=$	$\Delta^2 X_t$	$Z_{1t} :=$	ΔX_{t-1}
$Z_{2t} :=$	X^*_{t-1}	$Z_{3t} :=$	$(K'_t : d'_t)'$
$Z_{4t} :=$	$(Z'_{3t} : 1)'$	$Y_t :=$	$\beta^{*\prime} X^*_{t-1} + \bar{\alpha}'\Gamma\bar{\beta}_2\beta'_2\Delta X_{t-1}$
$U_t :=$	$\beta_1^{*\prime}\Delta X^*_{t-1}$	$V_t :=$	$(U'_t : Z'_{0t}b)'$
$u_t :=$	$(Z'_{0t}b : \Delta X'_{t-1}\beta : K'_t : d'_t)'$	$y_t :=$	$(s'_t : U'_t)'$
$v_t :=$	$(u'_t : 1)'$	$x_t :=$	$(y'_t : v'_t)'$
$s_t :=$	$(Z'_{0t}a : Y'_t : \epsilon'_{at})'$	$z_t :=$	$(U'_t : v'_t)'$

Table 4: Variable definitions

Hence eq. (6) is equal to (8). ■

Proof. of Proposition 5. Consider the following projection identity

$$S - Sg(g'Sg)^{-1}g'S = g_\perp(g'_\perp S^{-1} g_\perp)g'_\perp, \tag{31}$$

see Srivastava and Khatri (1979) p. 19. Let

$$A := \left(\bar{b}_\perp : \Omega b(b'\Omega b)^{-1}\right) \qquad B := \left(\Omega^{-1}\bar{b}_\perp(\bar{b}'_\perp\Omega^{-1}\bar{b}_\perp)^{-1} : b\right).$$

By projection identity (31), one has $AB' = I$ i.e. $B = A'^{-1}$, and it shows that $(\alpha'_a : \alpha'_b)' = B'\alpha$. This proves eq. (10). When (9) holds, $\alpha_b = 0$, and substituting one finds the results in *i*. Given the specification of $\alpha_\perp$, eq. (12) in *ii* follows because $b'\Gamma = 0$. *iii* follows analogously. Finally *iv* conbines the choices in *i* and *ii*. ■

A.2 Asymptotics

We here provide proofs of the asymptotics. Many results are identical to the ones stated in Rahbek et al. (1999), Johansen (1995), Paruolo and Rahbek (1999); hence we here only report the main steps in the proofs.

In order to study the asymptotic behavior of X^*_{t-1}, define

$$\begin{aligned} (\beta^* \ &: \ \beta^*_1 : \beta^*_2 : \beta^*_3) := \left(\begin{pmatrix} \beta \\ \beta_0 \end{pmatrix} : \begin{pmatrix} \beta_1 \\ \eta_0 \end{pmatrix} : \begin{pmatrix} \beta_2 \\ 0 \end{pmatrix} : \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right), \\ B_{1T} \ &: \ = (\beta^*_1 : T^{-1}\beta^*_2 : T^{-1/2}\beta^*_3). \end{aligned} \tag{32}$$

Let $W(u)$ be a Brownian motion with covariance Ω; note that $T^{-1/2}\sum_{i=1}^{[Tu]}\epsilon_i \xrightarrow{w} W(u)$, $u \in [0,1]$ and that by Johansen's I(2) representation theorem, reported in Proposition 1, and the continuous mapping theorem one finds

$$\begin{aligned} T^{-1/2}B'_{1T}X^*_{[Tu].(1,b)} \xrightarrow{w} F^*(u) \ &: \ = \left(\begin{array}{c|c} \beta'_1 C_1 W(u) & \\ \beta'_2 C_2 \int_0^u W(s)ds & G^*(u) \\ u & \end{array} \right), \\ G^*(u) \ &: \ = \begin{pmatrix} \beta'_2 C_2 W(u) \\ 1 \end{pmatrix}. \end{aligned} \tag{33}$$

In the following the argument u is suppressed unless needed for clarity, and integrals are understood to be from 0 to 1 when not otherwise indicated.

Let $Y_t := \beta^{*\prime}X^*_{t-1} + \bar{\alpha}'\Gamma\bar{\beta}_2\beta'_2\Delta X_{t-1}$, $U_t := \beta'_1\Delta X_{t-1}$, $V_t := (U'_t : Z'_{0t}b)'$. Let also M with no subscript be the sample moment matrix of $x_t := (y'_t : v'_t)'$, where $y_t := (s'_t$

$: U_t')'$, $s_t := (Z_{0t}'a\!: Y_t'\!: \epsilon_{at}')'$, $v_t := (u_t' : 1)'$, $u_t := (Z_{0t}'b : \Delta X_{t-1}'\beta : K_t'\!: d_t')'$. Correspondingly let $\Sigma^0 := E(x_t x_t')$. Blocks of composite matrices are indicated with subscripts a, b, Y, β, K, U, D, d, ϵ_a, x, y, v, u, s in an obvious notation. These definitions are summarized in Table 4.

Let Σ be the second moment matrix of s_t corrected for $z_t := (U_t' : v_t')'$, i.e. $\Sigma := \Sigma_{ss.z}^0 := \Sigma_{ss}^0 - \Sigma_{sz}^0 \Sigma_{zz}^{0-1} \Sigma_{zs}^0$; here and in the following $A_{ij.h} := A_{ij} - A_{ih} A_{hh}^{-1} A_{hj}$ with A any moment matrix is the moment matrix of i and j 'corrected for h'. Note that $\Sigma = (\Sigma_{ij})$ is the population second moment matrix of the relevant stationary processes contained in the matrices $S_{ij.(1,b)}$ that appear in the first stage eigenvalue problem (14).

Similarly let $\Sigma^* := \Sigma_{yy.u}^0$ be the second moment matrix of y_t corrected for u_t, where u_t contains the same set of variables as v_t except for the constant. $\Sigma^* = (\Sigma_{ij}^*)$ is the population second moment matrix of the relevant stationary processes contained in the matrices $S_{ij.(\beta,b)}^*$ that appear in the second stage eigenvalue problem (17).

We summarize some basic results in the following lemma.

Lemma 13 (First stage) *Regardless of the correct specification of condition* (9) *the following holds:*

$$\Sigma_{aY} \Sigma_{YY}^{-1} = \alpha_a, \tag{34}$$

$$(\beta^*, B_{1T}/\sqrt{T})' S_{22.(1,b)} (\beta^*, B_{1T}/\sqrt{T}) \xrightarrow{w} diag(\Sigma_{YY}, \int F^* F^{*\prime} du), \tag{35}$$

$$\alpha_{a\perp}' S_{a2.(1,b)} (T^{1/2}\beta^*, B_{1T}) \xrightarrow{w} (z^* : c' \int (dW) F^{*\prime}), \tag{36}$$

where $c' := \alpha_{a\perp}'(b_\perp' - \omega_{b_\perp} b')$, $F^*(u)$ *is defined in* (33), $vec(z^*) \sim N(0, \Sigma_{YY}^{-1} \otimes a_1' \Omega a_1)$, z^* *is independent of* W *and vec is the column stacking operator.*

Moreover if α_a *is of full column rank, then*

$$\Sigma_{aa}^{-1} - \Sigma_{aa}^{-1} \alpha_a (\alpha_a' \Sigma_{aa}^{-1} \alpha_a)^{-1} \alpha_a' \Sigma_{aa}^{-1} = \alpha_{a\perp} (\alpha_{a\perp}' \Omega_{aa.b} \alpha_{a\perp})^{-1} \alpha_{a\perp}'. \tag{37}$$

Proof. As shown in eq. (2.11) in Paruolo and Rahbek (1999), the conditional system equations (8) can be written as $(g_1' : g_2')(y_t' : v_t')' = 0$ with $g_1' := (-I : \alpha_a : I : \Gamma_a \bar{\beta}_1)$, $g_2' := (\omega_a : \Gamma_a \bar{\beta} : \Psi_a : \mu_a)$, which implies $g'\Sigma^0 = 0$. Correcting Σ^0 for z_t one finds $\Sigma^0 - \Sigma_{xz}^0 \Sigma_{zz}^{0-1} \Sigma_{zx}^0 = diag(\Sigma_{ss.z}^0 : 0) =: diag(\Sigma : 0)$, so that $g'\Sigma^0 = 0$ implies $g_1'\Sigma := g_1' \Sigma_{ss.z}^0 = 0$. This equality can be rewritten as $\Sigma_{aj} = \alpha_a \Sigma_{Yj} + \Sigma_{\epsilon_a j}$, or, setting $j = Y$, ϵ_a, a,

$$\Sigma_{aY} = \alpha_a \Sigma_{YY} + \Sigma_{\epsilon_a Y}, \qquad \Sigma_{a\epsilon_a} = \alpha_a \Sigma_{Y\epsilon_a} + \Sigma_{\epsilon_a \epsilon_a}, \qquad \Sigma_{aa} = \alpha_a \Sigma_{Ya} + \Sigma_{\epsilon_a a}. \tag{38}$$

In the first equality one has $\Sigma_{\epsilon_a Y} = 0$ because $\Sigma_{\epsilon_a j} := \Sigma_{\epsilon_a j}^0 - \Sigma_{\epsilon_a z}^0 \Sigma_{zz}^{0-1} \Sigma_{zj}^0 = 0$ for $j = Y, U, \beta, K, D$. In fact $\Sigma_{\epsilon j}^0 = 0$ and $\Sigma_{\epsilon_a b}^0 = (b_\perp' - \omega_{b_\perp} b') \Sigma_{\epsilon 0}^0 b = (b_\perp' - \omega_{b_\perp} b') \Omega b = 0$. The first equality is thus $\Sigma_{aY} = \alpha_a \Sigma_{YY}$ which gives (34).

In the second equality this implies also that $\Sigma_{Y\epsilon_a} = 0$ and $\Sigma_{\epsilon_a \epsilon_a} = \Omega_{aa.b}$. Hence $\Sigma_{a\epsilon_a} = \Omega_{aa.b}$, and substituting this and (34) in the third equality one finds $\Sigma_{aa} = \alpha_a \Sigma_{YY} \alpha_a' + \Omega_{aa.b}$. When a_a is of full column rank, applying the projection identity (31) to the l.h.s. of (37), one finds, substituting $\alpha_a \Sigma_{YY} \alpha_a' + \Omega_{aa.b}$ for Σ_{aa},

$$\Sigma_{aa}^{-1} - \Sigma_{aa}^{-1} \alpha_a (\alpha_a' \Sigma_{aa}^{-1} \alpha_a)^{-1} \alpha_a' \Sigma_{aa}^{-1} = \alpha_{a\perp} (\alpha_{a\perp}' \Sigma_{aa} \alpha_{a\perp})^{-1} \alpha_{a\perp}' = \alpha_{a\perp} (\alpha_{a\perp}' \Omega_{aa.b} \alpha_{a\perp})^{-1} \alpha_{a\perp}'$$

which gives (37).

Eq. (35) follows by (33) and the continuous mapping theorem for the non-stationary part and the ergodicity of the stationary AR processes in x_t. Finally consider (36) and note that $S_{a2.(1,b)}\beta^* = S_{aY.(V,\beta)} + o_p(1)$. Moreover $S_{aY.(V,\beta)} = \alpha_a S_{YY.(V,\beta)} + S_{\epsilon_a Y.(V,\beta)}$ similarly to the first equality in (38), so that $\alpha'_{a\perp} S_{aY.(V,\beta)} = \alpha'_{a\perp} S_{\epsilon_a Y.(V,\beta)} = a'_1 S_{\epsilon_a Y.(V,\beta)}$. Hence by the central limit theorem for martingale differences $T^{1/2}\alpha'_{a\perp} S_{a2.(1,b)}\beta^* \xrightarrow{w} z$. Similarly $\alpha'_{a\perp} S_{a2.(1,b)} B_{1T} = \alpha'_{a\perp} S_{\epsilon_a 2.(1,b)} B_{1T} \xrightarrow{w} c' \int (dW) F^{*\prime}$ by weak converge results to stochastic integrals in Chan and Wei (1988). The independence of z^* and W is proved as e.g. in Johansen (1996), proof of Theorem 13.5.

Note that in these derivations no use was made of the NF conditions, so that the above relations hold in general, except for the case of (37) where α_a of full rank is needed. ■

In order to discuss asymptotics for the estimators of β and α, for any $l \times j$ matrix $\widehat{\varsigma}$, $j \leq l$, we use the identified version given by $\widetilde{\varsigma} := \widehat{\varsigma}(\bar{\varsigma}'\widehat{\varsigma})^{-1}$, where ς is of the same dimension of $\widehat{\varsigma}$ and is such that $\bar{\varsigma}'\widehat{\varsigma}$ is square and full rank. E.g. we normalize $\widehat{\beta}_\perp$ as $\widetilde{\beta}_\perp := \widehat{\beta}_\perp(\bar{\beta}'_\perp \widehat{\beta}_\perp)^{-1}$.

Lemma 14 *Under the NF condition* (9) $(\widetilde{\beta}^* - \beta^*)$ *and* $(\widetilde{\beta}_\perp - \beta_\perp)$ *are* $O_p(T^{-1})$, *while* $(\widetilde{\alpha}_a - \alpha_a)$ *and* $(\widetilde{\alpha}_{a\perp} - \alpha_{a\perp})$ *are* $O_p(T^{-1/2})$. *The same order in probability apply to* β_{01} *and* α_{01} *when* α_a *is of deficient rank,* $\alpha_a = \alpha_{01}q'$, $\beta_{01} := \beta q$.

Proof. The proof is identical to the one of Theorem 5.1 in Paruolo and Rahbek (1999) apart from the deterministic terms. For the present choice of deterministic trends one has

$$(\widetilde{\beta}^* - \beta^*) = T^{-1}B_{1T}\left(\int F^* F^*\right)^{-1}\int F^* dW \Omega^{-1} a_1 (a_1'\Omega^{-1}a_1)^{-1} + o_p(T^{-1}).$$

where B_{1T} is defined in (32), which proves $(\widetilde{\beta}^* - \beta^*) = O_p(T^{-1})$. The process F^* and β_i^* in B_{1T} reflect the presence of linear trend in all directions, see Rahbek et al. (1999), eq. (B.8).

In order to show that α_a is $T^{1/2}$-consistent, observe that $\widetilde{\alpha}_a - \alpha_a = S_{\epsilon_a Y.(V,\beta)} S^{-1}_{YY.(V,\beta)} + O_p(T^{-1})$ as in the proof of Lemma B.3 in Paruolo and Rahbek (1999), where the first term on the r.h.s. multiplied by $T^{1/2}$ satisfies the central limit theorem for martingale differences. This shows $(\widetilde{\alpha}_a - \alpha_a) = O_p(T^{-1/2})$.

The above results also imply $(\widetilde{\beta}_\perp - \beta_\perp) = O_p(T^{-1})$, and $(\widetilde{\alpha}_\perp - \alpha_\perp) = O_p(T^{-1/2})$ given the relation between $(\widetilde{\varsigma}_\perp - \varsigma_\perp)$ and $(\widetilde{\varsigma} - \varsigma)$ for $\varsigma = \alpha,\ \beta$, discussed e.g. in Paruolo (1997) Section 4. ■

In order to study the asymptotic behavior of $\bar{\beta}'_\perp \Delta X^*_{t-1}$, define the basis

$$(\eta^* : \eta_1^* : \eta_2^*) := \left(\begin{pmatrix} \eta \\ \eta_0 \end{pmatrix} : \begin{pmatrix} \beta'_\perp \beta_\perp \eta_\perp \\ 0 \end{pmatrix} : \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right).$$

Define the matrices $H_{1T} := (\eta_1^* : T^{1/2}\eta_2^*)$, $H_T := (\eta^* : T^{-1/2}H_{1T})$ and observe that by the I(2) representation theorem $T^{-1/2}H'_{1T}\bar{\beta}'_\perp \Delta X_{[Tu].(\beta,b)} \xrightarrow{w} G^*(u)$, where the $G^*(u)$ process is defined in (33) above.

Lemma 15 (Second stage) *Regardless of the correct specification of the NF condition* (9) *the following holds:*

$$\Sigma^*_{a_1U}\Sigma^{*-1}_{UU} = \xi_1, \tag{39}$$

$$(\eta^*, T^{-1/2}H_{1T})' S^*_{\beta_\perp\beta_\perp.(\beta,b)}(\eta^*, T^{-1/2}H_{1T}) \xrightarrow{w} diag(\Sigma^*_{UU}, \int G^* G^{*\prime} du), \tag{40}$$

$$\xi_{1\perp}' S^*_{a_1\beta_\perp.(\beta,b)}(T^{1/2}\eta^*, H_{1T}) \xrightarrow{w} (z^\circ : \ell' \int (dW) G^{*\prime}), \qquad \ell' := \xi_{1\perp}' c', \tag{41}$$

where β, $\beta_\perp$, a_1 *above are the true values,* $vec(z^\circ) \sim N(0, \Sigma^{*-1}_{UU} \otimes \ell'\Omega\ell)$, z° *is independent of* W, *vec is the column stacking operator and the convergences in* (40) (41) *and in* (35) (36) *hold jointly.*

Finally if ξ_1 *is of full column rank, then*

$$\Sigma^{*-1}_{a_1a_1} - \Sigma^{*-1}_{a_1a_1}\xi_1(\xi_1'\Sigma^{*-1}_{a_1a_1}\xi_1)^{-1}\xi_1'\Sigma^{*-1}_{a_1a_1} = \xi_{1\perp}(\xi_{1\perp}'\Omega_{a_1a_1.b}\xi_{1\perp})^{-1}\xi_{1\perp}'. \tag{42}$$

Proof. Eq. (16) for the true values of the parameters can be rewritten as $g_1'y_t + g_2'u_t = 0$ where $g_1' := (\alpha_{a\perp}' : 0 : \alpha_{a\perp}' : \xi_1)$, $g_2' := (\omega_{a_1} : \gamma : \Psi_{a_1} : \varkappa_{a_1} : 0)$, $g' := (g_1' : g_2')$. This implies $g'\Sigma^0 = 0$; correcting for u_t one finds $g_1'\Sigma^* = 0$, which can be written as

$$\Sigma^*_{a_1j} = \xi_1\Sigma^*_{Uj} + \Sigma^*_{\epsilon_{a_1}j}$$

where $j = U$, a_1, ϵ_{a_1}. Setting $j = U$ and observing that $\Sigma_{\epsilon_{a_1}U} = 0$ by the same arguments in Lemma (13), eq. (39) follows. Similarly by setting $j = a_1$ and $j = \epsilon_{a_1}$ one finds $\Sigma^*_{a_1a_1} = \xi_1\Sigma^*_{UU}\xi_1' + \Omega_{a_1a_1.b}$ and hence

$$\xi_{1\perp}'\Sigma^*_{a_1a_1}\xi_{1\perp} = \xi_{1\perp}'\Omega_{a_1a_1.b}\xi_{1\perp}. \tag{43}$$

When ξ_1 is of full column rank, applying the projection identity (31) one finds $\Sigma^{*-1}_{a_1a_1} - \Sigma^{*-1}_{a_1a_1}\xi_1(\xi_1'\Sigma^{*-1}_{a_1a_1}\xi_1)^{-1}\xi_1'\Sigma^{*-1}_{a_1a_1} = \xi_{1\perp}(\xi_{1\perp}'\Sigma^*_{a_1a_1.b}\xi_{1\perp})^{-1}\xi_{1\perp}'$ and substituting (43) one obtains (42).

Next note that by ergodicity $\eta^{*\prime}S^*_{\beta_\perp\beta_\perp.(\beta,b)}\eta^* = S^*_{UU.(\beta,b)} \xrightarrow{p} \Sigma^*_{UU}$, while by convergence results to stochastic integrals in Chan and Wei (1988), $\eta^{*\prime}S^*_{\beta_\perp\beta_\perp.(\beta,b)}H_{1T} = O_p(1)$, and $T^{-1}H_{1T}'S^*_{\beta_\perp\beta_\perp.(\beta,b)}H_{1T} = (\beta_2^*T^{-1/2} : \beta_3^*)'S^*_{11.(\beta,b)}(\beta_2^*T^{-1/2} : \beta_3^*) \xrightarrow{w} \int G^*G^{*\prime}du$, which gives (40). Finally $\xi_{1\perp}'S^*_{a_1\beta_\perp.(\beta,b)} = \xi_{1\perp}'S^*_{\epsilon_{a_1}\beta_\perp.(\beta,b)}$ so that $T^{1/2}\xi_{1\perp}'S^*_{a_1\beta_\perp.(\beta,b)}\eta^* = T^{1/2}\xi_{1\perp}'S^*_{\epsilon_{a_1}U.(\beta,b)} \xrightarrow{w} z^\circ$ by the central limit theorem for martingale differences and $\xi_{1\perp}'S^*_{a_1\beta_\perp.(\beta,b)}H_{1T} \xrightarrow{w} \xi_{1\perp}'\alpha_{a\perp}'(a' - \omega_a b')\int(dW)G^* =: \xi_{1\perp}'c'\int(dW)G^*$. The independence of z° and W is proved as e.g. in Johansen (1996), proof of Theorem 13.5. ■

We now state the proof of Theorem 6.

Proof. of Theorem 6.

First stage. Note that (14) has only m positive roots, and that by the NF assumption, $m \geq p_0 + p_1$. Define the square nonsingular matrix $D_T := (\beta^*, B_{1T}/\sqrt{T})$, and pre- and post–multiply (14) by D_T' and its transpose. From Lemma 13 one has that $D_T'S_{22.(1,b)}D_T \xrightarrow{p} diag(\Sigma_{YY}, \int F^*F^{*\prime}du)$ and

$$D_T'S_{2a.(1,b)}S^{-1}_{aa.(1,b)}S_{a2.b}D_T \xrightarrow{p} diag(\Sigma_{Ya}\Sigma^{-1}_{aa}\Sigma_{aY}, 0).$$

Hence if α_a has full rank p_0, $(\lambda_1, ..., \lambda_{p_0})$ converge by (34) to the eigenvalues of the non-random matrix

$$\Sigma_{YY}\alpha_a'\Sigma_{aa}^{-1}\alpha_a, \tag{44}$$

while $(\lambda_{p_0+1}, ..., \lambda_m) = o_p(1)$.

Let now $\psi := T\lambda$ so that $\lambda = \psi/T$, define $S(\lambda) := \lambda S_{22.(1,b)} - S_{2a.(1,b)}S_{aa.(1,b)}^{-1}S_{a2.(1,b)}$, and pre- and post- multiply $S(\lambda)$ by $(\beta^* : B_{1T})'$ and its transpose. Applying standard properties of determinants one finds from (14)

$$0 = |D_T'S(\lambda)D_T| = |\beta^{*\prime}S(\lambda)\beta^*| \left|B_{1T}'\left(S(\lambda) - S(\lambda)\beta^*(\beta^{*\prime}S(\lambda)\beta^*)^{-1}\beta^{*\prime}S(\lambda)\right)B_{1T}\right| \tag{45}$$

From the results in Lemma 13 one has $\beta^{*\prime}S(\lambda)\beta^* = -\beta^{*\prime}N\beta^* + o_p(1)$ where $N := S_{2a.(V,\beta)}S_{aa.(V,\beta)}^{-1}S_{a2.(V,\beta)}$ so that the first factor in (45) has no roots in the limit. We also note that, from Lemma 13, $B_{1T}'S(\lambda)\beta^* = -B_{1T}'N\beta^* + o_p(1)$ and $B_{1T}'S(\lambda)B_{1T} = B_{1T}'(\psi S_{22.(1,b)}/T - N)B_{1T}$. Thus the second factor in (45) is of the form

$$\left|B_{1T}'\left(\psi S_{22.(1,b)}/T - N_1\right)B_{1T}\right| = o_p(1), \tag{46}$$

where $N_1 := (N - N\beta^*(\beta^{*\prime}N\beta^*)^{-1}\beta^{*\prime}N)$. From (34), (37) one finds

$$\begin{aligned} B_{1T}'N_1B_{1T} &= B_{1T}'S_{2a.(1.b)}(\Sigma_{aa}^{-1} - \Sigma_{aa}^{-1}\alpha_a(\alpha_a'\Sigma_{aa}^{-1}\alpha_a)^{-1}\alpha_a'\Sigma_{aa}^{-1})S_{a2.(1,b)}B_{1T} + o_p(1) = \\ &= B_{1T}'S_{2a.(1.b)}\alpha_{a\perp}(\alpha_{a\perp}'\Omega_{aa.b}\alpha_{a\perp})^{-1}\alpha_{a\perp}'S_{a2.(1,b)}B_{1T} + o_p(1) \\ &\overset{w}{\to} \int F^*(dW)'c(c'\Omega c)^{-1}c'\int(dW)F^{*\prime} \end{aligned}$$

where $c' := \alpha_{a\perp}'(b_\perp' - \omega_{b_\perp}b')$. Moreover from Lemma 13 one has $T^{-1}B_{1T}'S_{22.(1,b)}B_{1T} \overset{w}{\to} \int F^*F^{*\prime}du$. Thus ψ_1, ..., ψ_{m-p_0} in (46) converge to the eigenvalues of

$$\int F^*(dW')c(c'\Omega c)^{-1}c'\int(dW)F^{*\prime}\left(\int F^*F^{*\prime}\right)^{-1}. \tag{47}$$

Under correct specification of the NF conditions (9), one has $\alpha = \bar{b}_\perp\alpha_a$, $\alpha_1 = \bar{a}_1\xi_1$, $\alpha_2 := (\alpha_{21} : \alpha_{22}) := (a_1\xi_{1\perp} : b)$, where $a_1 := b_\perp\alpha_{a\perp}$. Define the following square nonsingular matrix $f := ((a_1'a_1)^{-1}\xi_1 : \xi_{1\perp})$; note that $d := (d_1 : d_2) := cf = (\alpha_1 : \alpha_{21})$ and

$$c(c'\Omega c)^{-1}c' = cf(f'c'\Omega cf)^{-1}f'c' = d(d'\Omega d)^{-1}d'.$$

Hence normalizing processes in (47) one finds the results for $Q_\infty^{p_{22}}(p_0)$.

Second stage. Consider the eigenvalue equation (17); pre and post multiplying by the square and nonsingular matrix $\overline{\beta}_\perp'\widehat{\beta}_\perp$ and its transpose, one sees that $\overline{\overline{\beta}}_\perp$ can be substituted by its identifies version $\overline{\overline{\beta}}_\perp$ without changing the equation. Similarly one can insert similar square matrices in order to substitute $\widehat{\beta}^*$ with $\widetilde{\beta}^*$, $\widehat{a}_1$ with $\widetilde{a}_1$ without changing the equation. Next note that

$$\begin{aligned} \overline{\overline{\beta}}_\perp' S_{11.(\widetilde{\beta},b)}^*\overline{\overline{\beta}}_\perp &= \overline{\beta}_\perp'S_{11.(\beta,b)}^*\overline{\beta}_\perp + o_p(1) \\ \overline{\overline{\beta}}_\perp' S_{1\widetilde{a}_1.(\widetilde{\beta},b)}^* &= \overline{\beta}_\perp'S_{1a_1.(\beta,b)}^* + o_p(1) \\ S_{\widetilde{a}_1\widetilde{a}_1.(\widetilde{\beta},b)}^* &= S_{a_1a_1.(\beta,b)}^* + o_p(1), \end{aligned}$$

i.e. the estimated values from the first stage can be substituted with the true values; this follows from Lemma 15 above and results like Lemma 18, Propositions 19 and 20 in Paruolo (2002b).

Let now $S(\rho) := \rho S^*_{\beta_\perp\beta_\perp.(\beta,b)} - S^*_{\beta_\perp a_1.(\beta,b)} S^{*-1}_{a_1a_1.(\beta,b)} S^*_{a_1\beta_\perp.(\beta,b)}$. Pre and post multiplying the eigenvalue equation by $E_T := (\eta^* : T^{-1/2}H_{1T})'$ and its transpose one sees that

$$|E_T' S(\rho) E_T| \xrightarrow{w} \left|\rho\Sigma^*_{UU} - \Sigma^*_{Ua_1}\Sigma^{*-1}_{a_1a_1}\Sigma^*_{a_1U}\right| \left|\rho\int G^*G^*du\right|.$$

The first factor is the only one to have non-zero roots in the limit. Thus under correct specification of the NF assumption one has that the first p_1 roots $(\rho_1, ..., \rho_{p_1})$ converge to the eigenvalues of the non-random matrix

$$\Sigma^*_{UU}\xi_1'\Sigma^{*-1}_{a_1a_1}\xi_1, \tag{48}$$

while $(\rho_{p_1+1}, ..., \rho_{m-p_0-p_1}) = o_p(1)$, where we have used Lemma 15.

Let now $\psi := T\rho$ so that $\rho = \psi/T$, and pre- and post- multiply $S(\rho)$ by $D_T := (\eta^* : H_{1T})'$ and its transpose. Applying standard properties of determinants one finds from (17)

$$0 = |D_T' S(\rho) D_T| = |\eta^{*\prime} S(\rho)\eta^*| \left|H_{1T}' \left(S(\rho) - S(\rho)\eta^*(\eta^{*\prime}S(\rho)\eta^*)^{-1}\eta^{*\prime}S(\rho)\right) H_{1T}\right| \tag{49}$$

From the results in Lemma 15 one has $\eta^{*\prime}S(\rho)\eta^* = -\eta^{*\prime}N\eta^* + o_p(1)$ where $N := S^*_{\beta_\perp a_1.(\beta,b)} S^{*-1}_{a_1a_1.(\beta,b)} S^*_{a_1\beta_\perp.(\beta,b)}$ so that the first factor in (49) has no roots in the limit. We also note that, from Lemma 15, $H_{1T}'S(\rho)\eta^* = -H_{1T}'N\eta^* + o_p(1)$ and $H_{1T}'S(\rho)H_{1T} = H_{1T}'(\psi S^*_{\beta_\perp\beta_\perp.(\beta,b)}/T - N)H_{1T}$. Thus the second factor in (49) is of the form

$$\left|H_{1T}'\left(\psi S^*_{\beta_\perp\beta_\perp.(\beta,b)}/T - N_1\right)H_{1T}\right| = o_p(1),$$

where $N_1 := (N - N\eta^*(\eta^{*\prime}N\eta^*)^{-1}\eta^{*\prime}N)$.

Let $K := \Sigma^{*-1}_{a_1a_1} - \Sigma^{*-1}_{a_1a_1}\xi_1(\xi_1'\Sigma^{*-1}_{a_1a_1}\xi_1)^{-1}\xi_1'\Sigma^{*-1}_{a_1a_1}$. From (39), (42) one finds

$$\begin{aligned} H_{1T}'N_1H_{1T} &= H_{1T}'S^*_{\beta_\perp a_1.(\beta,b)}KS^*_{a_1\beta_\perp.(\beta,b)}H_{1T} + o_p(1) = \\ &= H_{1T}'S^*_{\beta_\perp a_1.(\beta,b)}\xi_{1\perp}(\xi_{1\perp}'\Omega_{a_1a_1.b}\xi_{1\perp})^{-1}\xi_{1\perp}'S^*_{a_1\beta_\perp.(\beta,b)}H_{1T} + o_p(1) \\ &\xrightarrow{w} \int G^*(dW)'\ell\Omega_{\ell\ell.b}{}^{-1}\ell'\int(dW)G^{*\prime} \qquad \ell := c\xi_{1\perp}. \end{aligned}$$

Moreover from Lemma 15 one has $T^{-1}H_{1T}'S^*_{\beta_\perp\beta_\perp.(\beta,b)}H_{1T} \xrightarrow{w} \int G^*G^{*\prime}du$. Thus ψ_1, ..., $\psi_{m-p_0-p_1}$ in converge to the eigenvalues of

$$f'\int(dW)G^{*\prime}\left(\int G^*G^{*\prime}\right)^{-1}\int G^*(dW')f,$$

where $f' := \Omega_{\ell\ell.b}{}^{-1/2}\ell'$, $\ell' := \xi_{1\perp}'c'$. Under correct specification of the NF conditions (9), one has $\alpha = \bar{b}_\perp\alpha_a$, $\alpha_1 = \bar{a}_1\xi_1$, $\alpha_2 := (\alpha_{21} : \alpha_{22}) := (a_1\xi_{1\perp} : b)$, where $a_1 := b_\perp\alpha_{a\perp}$. Hence $\ell = \alpha_{21} - \alpha_{22}\Omega^{-1}_{\alpha_{22}\alpha_{22}}\Omega_{\alpha_{22}\alpha_{21}}$, $\Omega_{\ell\ell.b} = \Omega_{\alpha_{21}\alpha_{21}.\alpha_{22}}$, $f'W = B_{21}$. Normalizing processes one obtains the results for $Q^{p_{22}}_\infty(p_1|p_0)$. The weak convergence results hold jointly, and one can apply the continuous mapping theorem to prove that $Q^{p_{22}}_\infty(p_0, p_1) = Q^{p_{22}}_\infty(p_0) + Q^{p_{22}}_\infty(p_1|p_0)$. This completes the proof. ■

Proof. of Corollary 7. Apply arguments in Paruolo (2001) to $Q^{p_{22}}(i,j)$, using the results in Theorem 6 ■

Proof. of Proposition 9. When $rank(\alpha_a) = p_{01} < p_0$, then α_a can be rank-decomposed as $\alpha_a = \alpha_{01}q'$, where α_{01} and q are of rank p_{01}. Let $\beta_{01}^* := \beta^* q$, $\beta_{02}^* = \beta^* q_\perp$, and define Y_1 and Y_2 accordingly. It then follows as in (44) in the proof of Theorem 6 that the first p_{01} eigenvalues λ_1,..., $\lambda_{p_{01}}$ converge to the eigenvalues of $\Sigma_{YY}^{-1}\Sigma_{YY_1}\alpha_{01}'\Sigma_{aa}^{-1}\alpha_{01}\Sigma_{Y_1Y}$, while the remaining $m - p_{01}$ converge to 0. Define next $B_T := (T^{1/2}\beta_{02}^* : B_{1T})$ and carry on with the proof using β_{01}^* and B_T in place of β^* and B_{1T}. It then follows by the same arguments that $T\lambda_{p_{01}+1}$, ..., $T\lambda_{m-p_{01}}$ converge to the eigenvalues of the weak limit of the matrix

$$(T^{-1}B_T' S_{22.(1.b)}B_T)^{-1}B_T' S_{2a.(1.b)}\alpha_{01\perp}(\alpha_{01\perp}'\Omega_{aa.b}\alpha_{01\perp})^{-1}\alpha_{01\perp}' S_{a2.(1,b)}B_T$$

where, by (36) one finds $T^{-1}B_T' S_{22.(1.b)}B_T \xrightarrow{w} diag(\Sigma_{Y_2Y_2}, \int F^*F^*du)$ and

$$B_T' S_{2a.(1.b)}\alpha_{01\perp} \xrightarrow{w} (\breve{z} : \int F^*(dW)'g),$$

where $g := (b_\perp - b\omega_{b_\perp}')\alpha_{01\perp}$ and $vec(\breve{z}) \sim N(0, g'\Omega g \otimes \Sigma_{Y_2Y_2})$. The independence of $\breve{z}$ and W is proved as e.g. in Johansen (1996), proof of Theorem 13.5. Hence $T\lambda_{p_{01}+1}$, ..., $T\lambda_{m-p_{01}}$ converge to the eigenvalues of the matrix $J := z'z + h(\widetilde{B}, F)$ where $\widetilde{B}$ is the standardization of $g'W$ and z is a standard normal matrix obtained by standardization of $\breve{z}$. The covariance of $\widetilde{B}$ and B depends on the relation between g and α_1, α_2. ■

Proof. of Proposition 10. Under the assumption of the proposition, the first stage estimate of α_a and β^* have p_{01} columns and are consistent for α_{01} and $\beta_{01}^* := \beta^* q$. Define $\beta_{02} := \bar{\beta}q_\perp$, and note that $\beta_{01\perp}$ can be chosen equal to $(\beta_\perp : \beta_{02})$. With this choice, pre-multiply the equation of the second stage by $\alpha_{01\perp}$ in place of $\alpha_{a\perp}$ and insert $P_{\beta_{01}} + P_{\beta_{01\perp}} = I$ in place of $P_\beta + P_{\beta_\perp} = I$. One finds, in an obvious notation,

$$\begin{aligned}\alpha_{01\perp}'a'R_{0t.(b,\beta_{01})} &= ((I:0)\xi\eta' : \alpha_{01\perp}'\Gamma_a\bar{\beta}_{02})\begin{pmatrix}\beta_\perp' R_{1t.(b,\beta_{01})}\\ \beta_{02}' R_{1t.(b,\beta_{01})}\end{pmatrix} + \\ &+(I \quad : \quad 0)\eta_0' R_{\iota.(b,\beta_{01})} + \alpha_{01\perp}' R_{\epsilon_a t.(b,\beta_{01})}.\end{aligned}$$

where R are residuals and ι stands for the constant. The matrix $((I:0)\xi\eta' : \alpha_{01\perp}'\Gamma_a\bar{\beta}_{02})$ is of dimension $(m - p_{01}) \times (p - p_{01})$ which may have any rank from 0 to full rank $m - p_{01}$. The number of non vanishing eigenvalues of (17) is equal to the rank of the coefficient matrix. This completes the proof. ■

Proof. of Proposition 12. When $rank(\xi_1) = p_{11} < p_1$, then ξ_1 can be rank-decomposed as $\xi_1 = \xi_{11}q'$, where ξ_{11} and q are of rank p_{11}. Let $\beta_{11}^* := \beta_1^* q$, $\beta_{12}^* = \beta_1^* q_\perp$, $\eta_{01}^* := \eta^* q$, $\eta_{02}^* := \eta^* q_\perp$ and define U_1 and U_2 accordingly. It then follows as in (48) in the proof of Theorem 6 that the first p_{11} eigenvalues ρ_1,..., $\rho_{p_{11}}$ converge to the eigenvalues of $\Sigma_{UU}^{*-1}\Sigma_{UU_1}^*\xi_{11}'\Sigma_{a_1a_1}^{*-1}\xi_{11}\Sigma_{U_1U}$, while the remaining $m - p_0 - p_{11}$ converge to 0. Define next $H_T := (T^{1/2}\eta_{02}^* : H_{1T})$ and carry on with the proof using η_{01}^* and H_T in place of η^* and H_{1T}. It then follows by the same arguments that $T\rho_{p_{11}+1}$, ..., $T\rho_{m-p_0-p_{11}}$ converge to the eigenvalues of the weak limit of the matrix

$$(T^{-1}H_T' S_{\beta_\perp\beta_\perp.(\beta.b)}^* H_T)^{-1}H_T' S_{\beta_\perp a_1.(\beta.b)}^*\xi_{11\perp}(\xi_{11\perp}'\Omega_{a_1a_1.b}\xi_{11\perp})^{-1}\xi_{11\perp}' S_{a_1\beta_\perp.(\beta.b)}^* H_T$$

where, by (41) one finds $T^{-1}H_T' S^*_{\beta_\perp \beta_\perp.(\beta.b)} H_T \xrightarrow{w} diag(\Sigma^*_{U_2U_2}, \int G^* G^* du)$ and

$$H_T' S^*_{\beta_\perp a_1.(\beta.b)} \xi_{11\perp} \xrightarrow{w} (\widetilde{z} : \int G^*(dW)' g),$$

where $g := a_1 \xi_{11\perp}$ and $vec(\widetilde{z}) \sim N(0, \Omega_{gg.b} \otimes \Sigma^*_{U_2U_2})$. The independence of $\widetilde{z}$ and W is proved as e.g. in Johansen (1996), proof of Theorem 13.5. Hence $T\rho_{p_{11}+1}$, ..., $T\rho_{m-p_0-p_{11}}$ converge to the eigenvalues of the matrix $z'z + h(\widetilde{B}, G)$ where $\widetilde{B}$ is the standardization of $g'W$ and z is a standard normal matrix obtained by standardization of $\widetilde{z}$. The covariance of $\widetilde{B}$ and B depends on the relation between g and α_1, α_2.
■

A.3 Tables of $Q_{\infty}^{p_{22}}(p_0, p_1)$

Estimates of the quantiles of the limit distribution $Q_{\infty}^{p_{22}}(p_0, p_1)$ based on the response surface (23), for $p_2 = 1, ..., 6$, $p_1 \leq 8 - p_2$, and all positive values of p_{21}, $1 \leq p_{21} < p_2$, where $p_{21} + p_{22} = 0$, $p_{22} := p - m$. Entries in *italics* correspond to the case $p_{21} = p_2$, i.e. $p_{22} = 0$, where the distribution coincides with the one in Rabhek et al. (1999) for the full system analysis. The case $p_{21} = 0$ is treated in the next Appendix.

p_2	p_1	p_{21}	κ_1	κ_2	mean	var	90%	95%	99%
6	2	6	*0.753*	*255.898*	*339.7*	*450.8*	*367.1*	*375.3*	*391.0*
	2	5	0.700	205.553	293.7	419.5	320.2	328.1	343.4
	2	4	0.653	161.672	247.8	379.7	273.0	280.7	295.3
	2	3	0.604	122.020	201.9	333.9	225.6	232.8	246.8
	2	2	0.568	88.553	155.9	274.6	177.5	184.2	197.1
	2	1	0.544	59.836	110.0	202.1	128.5	134.4	145.7
	1	6	*0.757*	*222.245*	*293.6*	*388.0*	*319.2*	*326.8*	*341.4*
	1	5	0.694	173.247	249.7	359.8	274.3	281.7	295.9
	1	4	0.645	132.704	205.8	319.2	229.0	236.1	249.7
	1	3	0.594	96.138	161.9	272.6	183.4	190.0	202.8
	1	2	0.561	66.183	118.0	210.3	136.9	142.8	154.3
	1	1	0.533	39.493	74.0	138.8	89.5	94.4	104.2
	0	6	*0.762*	*191.684*	*251.7*	*330.5*	*275.3*	*282.3*	*295.9*
	0	5	0.699	146.525	209.7	300.0	232.2	239.0	252.1
	0	4	0.645	108.286	167.8	260.1	188.8	195.2	207.6
	0	3	0.594	74.750	125.8	211.8	144.8	150.7	162.2
	0	2	0.556	46.677	83.9	151.0	100.0	105.1	115.1
	0	1	0.524	21.997	42.0	80.2	53.8	57.8	65.6
5	3	5	*0.753*	*225.470*	*299.6*	*398.2*	*325.5*	*333.2*	*348.0*
	3	4	0.685	176.593	257.7	375.9	282.8	290.4	304.9
	3	3	0.636	137.172	215.7	339.1	239.6	246.8	260.8
	3	2	0.599	104.054	173.7	290.0	195.9	202.6	215.8
	3	1	0.565	74.512	131.8	233.1	151.7	157.9	169.9
	2	5	*0.758*	*193.936*	*255.7*	*337.1*	*279.5*	*286.6*	*300.3*
	2	4	0.686	147.955	215.7	314.6	238.7	245.7	259.1
	2	3	0.635	111.700	175.8	276.6	197.4	204.0	216.8
	2	2	0.593	80.546	135.8	229.0	155.5	161.6	173.5
	2	1	0.554	53.140	95.9	173.0	113.1	118.5	129.1
	1	5	*0.761*	*164.122*	*215.7*	*283.4*	*237.5*	*244.1*	*256.8*
	1	4	0.681	121.106	177.7	260.8	198.7	205.1	217.4
	1	3	0.626	87.484	139.7	223.2	159.2	165.2	176.8
	1	2	0.584	59.428	101.8	174.4	119.0	124.4	135.0
	1	1	0.545	34.839	63.9	117.1	78.1	82.6	91.7
	0	5	*0.767*	*137.881*	*179.7*	*234.3*	*199.6*	*205.6*	*217.2*
	0	4	0.685	98.495	143.8	209.8	162.6	168.4	179.6
	0	3	0.626	67.535	107.8	172.1	124.9	130.3	140.7
	0	2	0.580	41.674	71.9	123.9	86.4	91.1	100.3
	0	1	0.539	19.360	35.9	66.7	46.7	50.3	57.6

p_2	p_1	p_{21}	κ_1	κ_2	mean	var	90%	95%	99%
4	4	4	*0.724*	*191.057*	*263.9*	*364.5*	*288.7*	*296.1*	*310.3*
	4	3	0.674	152.344	225.9	335.0	249.6	256.8	270.6
	4	2	0.624	117.367	188.0	301.1	210.5	217.4	230.7
	4	1	0.591	88.645	150.0	253.7	170.7	177.1	189.5
	3	4	*0.729*	*161.671*	*221.9*	*304.5*	*244.5*	*251.3*	*264.5*
	3	3	0.672	124.952	185.9	276.6	207.5	214.1	226.8
	3	2	0.620	92.993	150.0	241.9	170.2	176.5	188.5
	3	1	0.581	66.179	114.0	196.3	132.2	137.9	149.1
	2	4	*0.734*	*135.045*	*183.9*	*250.3*	*204.4*	*210.6*	*222.7*
	2	3	0.669	100.220	149.9	224.2	169.4	175.4	186.9
	2	2	0.608	70.561	116.0	190.7	134.0	139.6	150.5
	2	1	0.563	46.163	82.0	145.6	97.8	102.8	112.6
	1	4	0.745	111.668	149.8	201.1	168.3	173.9	184.8
	1	3	0.667	78.569	117.9	176.8	135.2	140.5	151.0
	1	2	0.597	51.346	86.0	143.9	101.7	106.6	116.3
	1	1	0.549	29.587	53.9	98.3	67.0	71.2	79.7
	0	4	0.764	91.597	119.9	156.8	136.2	141.2	150.9
	0	3	0.676	60.746	89.9	133.0	104.9	109.7	118.9
	0	2	0.602	36.091	60.0	99.7	73.1	77.3	85.6
	0	1	0.543	16.288	30.0	55.2	39.8	43.1	49.9
3	5	3	*0.704*	*163.556*	*232.4*	*330.2*	*256.0*	*263.1*	*276.7*
	5	2	0.658	130.485	198.4	301.6	220.9	227.8	241.0
	5	1	0.609	100.144	164.4	269.7	185.7	192.3	205.0
	4	3	*0.704*	*135.379*	*192.2*	*272.9*	*213.7*	*220.2*	*232.7*
	4	2	0.650	104.153	160.2	246.5	180.6	186.9	199.0
	4	1	0.599	76.831	128.2	214.0	147.3	153.2	164.7
	3	3	*0.709*	*110.809*	*156.2*	*220.3*	*175.5*	*181.4*	*192.8*
	3	2	0.647	81.638	126.3	195.3	144.5	150.1	161.0
	3	1	0.593	57.068	96.2	162.3	112.9	118.1	128.3
	2	3	*0.712*	*88.419*	*124.2*	*174.5*	*141.4*	*146.7*	*157.0*
	2	2	0.640	61.625	96.2	150.3	112.2	117.2	127.0
	2	1	0.581	39.668	68.2	117.4	82.4	87.0	95.9
	1	3	*0.719*	*69.026*	*96.0*	*133.6*	*111.1*	*115.8*	*124.9*
	1	2	0.638	44.704	70.1	109.8	83.8	88.1	96.7
	1	1	0.572	25.178	44.0	77.1	55.6	59.4	67.0
	0	3	*0.737*	*53.040*	*71.9*	*97.6*	*84.9*	*88.9*	*96.9*
	0	2	0.639	30.652	48.0	75.1	59.4	63.1	70.4
	0	1	0.563	13.496	24.0	42.6	32.6	35.6	41.7

p_2	p_1	p_{21}	κ_1	κ_2	mean	var	90%	95%	99%
2	6	2	*0.697*	*142.516*	*204.4*	*293.2*	*226.6*	*233.4*	*246.3*
	6	1	0.656	114.416	174.5	266.0	195.7	202.1	214.6
	5	2	*0.696*	*115.780*	*166.3*	*239.0*	*186.4*	*192.6*	*204.4*
	5	1	0.647	89.602	138.4	213.8	157.4	163.3	174.7
	4	2	*0.698*	*92.412*	*132.3*	*189.5*	*150.3*	*155.8*	*166.4*
	4	1	0.639	67.989	106.4	166.5	123.2	128.5	138.7
	3	2	*0.699*	*71.566*	*102.4*	*146.5*	*118.2*	*123.1*	*132.6*
	3	1	0.632	49.561	78.4	124.2	93.0	97.6	106.7
	2	2	*0.696*	*53.134*	*76.4*	*109.8*	*90.1*	*94.4*	*102.8*
	2	1	0.613	33.358	54.4	88.6	66.7	70.7	78.6
	1	2	*0.687*	*37.239*	*54.2*	*79.0*	*65.9*	*69.6*	*77.0*
	1	1	0.590	20.187	34.2	58.0	44.3	47.6	54.4
	0	2	*0.696*	*25.096*	*36.1*	*51.9*	*45.6*	*48.7*	*54.9*
	0	1	0.583	10.538	18.1	31.0	25.5	28.1	33.5

A.4 Tables of $Q_\infty^{p_2}(p_0)$

Quantiles of the limit distribution $Q_\infty^{p_2}(p_0)$ on the response surface (23). The same design of Monte Carlo experiment as for the limit distribution of $Q_\infty^{p_{22}}(p_0, p_1)$ was used, see previous Appendix. This distribution is the relevant one when $p_{21} = 0$, $p_{22} = p_2$, i.e. there are no more I(2) trends in addition to the maintained lower bound $p_{22} := p - m$.

p_2	p_1	κ_1	κ_2	mean	var	90%	95%	99%
6	2	0.524	33.530	64.0	122.2	78.5	83.2	92.5
	1	0.505	15.173	30.0	59.5	40.3	43.7	50.8
5	3	0.540	48.585	90.0	166.6	106.8	112.2	122.7
	2	0.523	29.295	56.0	107.1	69.6	74.1	82.9
	1	0.514	13.349	26.0	50.5	35.4	38.7	45.3
4	4	0.573	64.173	112.0	195.6	130.3	136.0	147.1
	3	0.561	43.739	78.0	139.2	93.5	98.4	108.1
	2	0.540	25.947	48.0	88.8	60.4	64.5	72.6
	1	0.519	11.391	22.0	42.4	30.6	33.7	39.9
3	5	0.584	76.092	130.4	223.4	149.9	155.9	167.7
	4	0.572	55.007	96.2	168.4	113.2	118.5	129.0
	3	0.562	37.236	66.2	117.8	80.5	85.0	94.1
	2	0.544	21.862	40.2	73.9	51.5	55.3	62.9
	1	0.528	9.534	18.0	34.2	25.8	28.6	34.3
2	6	0.625	90.311	144.4	230.9	164.2	170.3	182.1
	5	0.617	68.085	110.4	178.9	127.8	133.2	143.8
	4	0.599	48.131	80.3	134.1	95.5	100.3	109.7
	3	0.586	31.831	54.3	92.8	67.0	71.1	79.2
	2	0.564	18.201	32.3	57.3	42.3	45.7	52.5
	1	0.529	7.493	14.2	26.8	21.1	23.6	28.9
1	7	0.655	101.087	154.4	236.0	174.4	180.5	192.4
	6	0.642	77.348	120.5	187.8	138.4	143.9	154.7
	5	0.628	56.814	90.4	143.9	106.1	111.0	120.7
	4	0.618	39.784	64.4	104.2	77.8	82.0	90.5
	3	0.604	25.561	42.3	70.1	53.4	57.0	64.2
	2	0.587	14.219	24.2	41.3	32.8	35.7	41.6
	1	0.552	5.619	10.2	18.4	15.9	18.1	22.7

References

Anderson, T.W. (1951), Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Annals of Mathematical Statistics* 22, 327–351. Correction, *Annals of Statistics* 8, 1980, 1400.

Boswijk H. P. and J. A. Doornik (1999), Distribution approximation for cointegration tests with stationary exogenous regressors, paper available at http://www.nuff.ox.ac.uk/Users/Doornik/

Chan N. H. and C. Z. Wei (1988), "Limiting distributions of least-squares estimates of unstable autoregressive processes", *Annals of Statistics* 16, 367–401.

Doornik J. A. (1998), "Approximations to the asymptotic distributions of cointegration tests", *Journal of Economic Surveys* 12, 573-93.

Doornik J., Hendry D., Nielsen B. (1998), Inference in cointegrating models: UK M1 revisited. *Journal of Economic Surveys* 12, 533-40.

Engle, R.F., D.F. Hendry & J.F. Richard (1983), Exogeneity, *Econometrica* 51, 277-304.

Harbo I., S. Johansen, B. Nielsen & A. Rahbek (1998), Asymptotic inference on cointegrating rank in partial systems, *Journal of Business and Economic Statistics*, 16, 388-399.

Hendry D. and M. H. Krolzig (2002), The properties of automatic Gets modeling, Mimeo, Nuffield College, Oxford.

Johansen S. (1992), A Representation of Vector Autoregressive Processes Integrated of Order 2. *Econometric Theory* 8, 188-202.

Johansen S. (1995), A Statistical Analysis of Cointegration for I(2) variables. *Econometric Theory* 11, 25-59.

Johansen S. (1996), *Likelihood–based inference in cointegrated vector autoregressive models*, second printing. Oxford: Oxford University Press.

Johansen S. (1997), A Likelihood Analysis of the I(2) Model. *Scandinavian Journal of Statistics* 24, 433-462.

Johansen S., R. Mosconi and B. Nielsen (2000), Cointegration analysis in presence of structural breaks in the deterministic trend, *Econometrics Journal*, 3, 216-249.

MacKinnon J. G., A. A. Huag and L. Michelis (1999), Numerical distribution functions of likelihood ratio tests for cointegration, *Journal of Applied Econometrics*, 14, 563-577.

Omtzigt P. and P. Paruolo (2002), Impact factors, WP Università dell'Insubria Varese, Italy, 2002/4, available at: http://eco.uninsubria.it/dipeco/Quaderni/files/QF2002_4.pdf

Paruolo P. (1996), On the determination of integration indices in I(2) systems. *Journal of Econometrics* 72, 313-356.

Paruolo P. (1997), Asymptotic inference on the moving average impact matrix in cointegrated I(1) VAR systems. *Econometric Theory* 13, 79-118.

Paruolo P. (2000), Asymptotic efficiency of the two stage estimator in I(2) systems. *Econometric Theory* 16, 524-550.

Paruolo P. (2001), The power of lambda max, *Oxford Bullettin of Economics and Statistics* 63, 3, 305–349.

Paruolo P. (2002a), Asymptotic inference on the moving average impact matrix in cointegrated I(2) VAR systems. *Econometric Theory* 18, 673-690.

Paruolo P. (2002b), Common features and common I(2) trends in VAR systems, Quaderno della Facoltà di Economia 2002/29, University of Insubria, Italy – invited paper at the conference "Common features in Rio", Rio de Janeiro 28-31 July 2002. Available at:
http://eco.uninsubria.it/dipeco/Quaderni/files/QF2002_29.pdf

Paruolo, P. (2003), Response surface analysis for quantile estimation, in progress.

Paruolo, P. and A. Rahbek (1999), Weak exogeneity in I(2) VAR systems. *Journal of Econometrics* 93, 281-308.

Rahbek A. (1997), Representation of cointegrated I(2) and I(1) processes with deterministic trends, manuscript, University of Copenhagen.

Rahbek A., Kongsted H.C., and C. Jørgensen (1999), Trend-Stationarity in the I(2) Cointegration Model. *Journal of Econometrics* 90, 265-89.

Srivastava M.S. and C.G. Khatri (1979), *An introduction to multivariate statistics.* New York: Elsevier North Holland.

Stock J. and Watson M. (1988), Testing for common trends. *Journal of the American Statistical Association* 83, 1097-1107.