Pieter Omtzigt, Paolo Paruolo

Impact factors

2002/4

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 October 2002
Università degli Studi dell'Insubria
Via Ravasi 2, 21100 Varese, Italy

Impact factors

Pieter Omtzigt, Paolo Paruolo*
Department of Economics, University of Insubria,
Via Ravasi 2, 21100 Varese, Italy

October 29, 2002

Abstract

In this paper we discuss sensitivity of forecast with respect to the information set considered in prediction; we define a sensitivity measure called impact factor, IF. We calculate this measure in VAR processes integrated of order 0, 1 and 2. For VAR processes this measure is a simple function of the impulse response coefficients. For integrated VAR systems this measure is shown to have a direct interpretation in terms of long-run forecasts. Various applications of this concept are reviewed, including one on the interpretation and effectiveness of economic policies and one on the sensitivity of forecasts with respect to data revisions. A unified approach to inference on the IF is given, showing under what circumstances standard asymptotic inference can be conducted also in systems integrated of order 1 and 2.

Keywords: Forecasting, Cointegration, Dynamic multipliers, (Generalized) Impulse Responses, VAR.

JEL classification: C32, C52, C53.

*Partial financial support from Italian MIUR grants ex 60% is gratefully acknowledged. Paper presented at the Econometric Society European Meeting 2002, Venice, Italy, 25-28 August 2002. We wish to thank Jan Magnus, Neil Ericsson and Søren Johansen for useful comments. The usual disclaimer applies.

Contents

1 Introduction

In this paper we discuss a measure of sensitivity of forecast with respect to the information set considered in prediction, called impact factor, IF. We calculate this measure in VAR processes integrated of order 0, 1 and 2. For VAR processes this measure is a simple function of the impulse response coefficients. For integrated VAR systems this measure is shown to have a direct interpretation in terms of long-run forecast of the levels of the process. Various applications of this concept are reviewed, including one on the interpretation and effectiveness of economic policies and one on the sensitivity of forecasts with respect to data revisions. A unified approach to inference on the IF is given, showing under what circumstances standard asymptotic inference can be conducted on the IF also in systems integrated of order 1 and 2.

Sensitivity indicators have long been advocated in econometrics; see Banerjee and Magnus (1999, 2000) for recent references. The concept of IF is also related to many standard econometric concepts, like dynamic multipliers and impulse responses. As a dynamic multiplier, IF measures the sensitivity of a function. However dynamic multipliers are defined only between some endogenous variables y and some exogenous variables x; impact factors, instead, are well defined for any dynamic systems, including VARs. Finally long-run multipliers are usually defined in terms of the static relation implied by a dynamic model for y and x, see e.g. Hendry (1995, p. 339), Gourieroux and Monfort (1995 p. 34-35), whereas impact factors measure the accumulated effects on forecasts of perturbations in past information.

Impact factors are functions of the impulse responses in case of VARs. While impulse responses are usually interpreted as measuring the effects of shocks, IFs are defined in terms of changes in observable variables; this difference in interpretation allows to view these measures also in the perspective of policy analysis and data revisions.

While the present approach is defined in terms of stationary processes, it is motivated and applied to non-stationary integrated systems. We consider I(1) and I(2) processes and compute impact factors for these processes. The present paper builds on ideas presented in Bedini and Mosconi (2000) for I(1) systems. They introduced the concept of 'long-run adjustment coefficients' with respect to the disequilibrium associated with an error correction term. We here offer different insights on the I(1) case and extend the concept to I(2) systems. For the I(1) case we show how the long-run adjustment coefficients is related to the forecast function, and more in general to the concept of IF. This concept is linked to the choice of state vector and the timing of variables, and we discuss the relation among different choices.

The rest of the paper is organized as follows. Section 2 reports relevant definitions and relates impact factors to impulse response. Section 3 discusses impact factors in I(1) and I(2) processes. Section 4 discusses two possible applications of this concept to the effectiveness of economic policies and to forecast sensitivity with respect to data revisions. Section 5 discusses the estimation of IF, while Section 6 reports an application on prices in Australia. Finally Section 7 reports conclusions. All proofs are placed in 3 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 $sp(H)$. $\|\cdot\|$ indicates a matrix norm and its associated vector norm. Moreover $P_H := H\bar{H}'$, $H_{AB} := \bar{A}'H\bar{B}$, $H_{AB.C} := H_{AB} - H_{AC}H_{CC}^{-1}H_{CB}$ while $H_A := H(A'H)^{-1}$. Finally $(\cdot)_{ij}$ indicates the ij-th element of the argument matrix, 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 Basic definition

Let $\{X_t\}_{t=-\infty}^{\infty}$ be a stationary p-variate time series, which contains the relevant information for the forecasting exercise. Let Y_t be a $n \times 1$ vector of variables of interest, which are to be forecast. Let $Y_{t+i|t}$ be the optimal forecast of Y_{t+i} based on available information up and including time t, indicated by $X_{-\infty}^t := (X_t, X_{t-1}, ...)$, deemed to be the relevant information set.

The forecast $Y_{t+i|t}$ is a function, $g_i^\circ(\cdot)$ say, of $X_{-\infty}^t$, $Y_{t+i|t} = g_i^\circ(X_{-\infty}^t)$. Under quadratic loss, for instance, one has $Y_{t+i|t} = E(Y_{t+i}|X_{-\infty}^t)$, the conditional expectation.[1] We wish to summarize the sensitivity of the forecast function with respect to its inputs. Let $\widetilde{X}_t$ be a vector containing the relevant part of the information set retained in the forecast function, i.e. $Y_{t+i|t} = g_i^\circ(X_{-\infty}^t) = g_i(\widetilde{X}_t)$ for some function $g_i(\cdot)$. $\widetilde{X}_t$ is thus a 'sufficient statistic' for the information contained in $X_{-\infty}^t$; we call $\widetilde{X}_t$ the FS statistic ('Forecast Sufficient'), and indicate its dimension with s.

Let $\widetilde{v} := \widetilde{X}_t^c - \widetilde{X}_t$ be a perturbation in the FS statistic which induces a change $e_i(\widetilde{v}, \widetilde{X}_t) := g_i(\widetilde{X}_t^c) - g_i(\widetilde{X}_t)$ in the forecast function at forecast horizons $i = 1, ..., \ell$. We consider the cumulated changes $\sum_{i=1}^{\ell} e_i(\widetilde{v}, \widetilde{X}_t)$ up to some finite forecast ℓ. If the sum converges for $\ell \to \infty$ we define the total effect, TE, of the perturbation as

$$\mathrm{TE}(\widetilde{v}, \widetilde{X}_t) := \sum_{i=1}^{\infty} e_i(\widetilde{v}, \widetilde{X}_t)$$

The quantity TE depends on $\widetilde{v}$ and possibly $\widetilde{X}_t$; we wish to find a sensitivity measure of TE with respect to (small) changes $\widetilde{v}$, for fixed $\widetilde{X}_t$. This reflects the fact that the actual forecast takes place for given $\widetilde{X}_t$ and the sensitivity is measured locally, i.e. around a specific value for $\widetilde{X}_t$. This local sensitivity measure becomes also a global one when TE only depends on $\widetilde{v}$ and not on $\widetilde{X}_t$.

TE as a function of the perturbation $\widetilde{v}$ may be approximated by Taylor expansion around $\widetilde{v} = 0$ for fixed $\widetilde{X}_t$; this gives

$$\mathrm{TE}(\widetilde{v}, x) = \mathrm{TE}(0, x) + F(0, x)\widetilde{v} + R(\widetilde{v}, x)$$

where R is a remainder term, which is of order $\|\widetilde{v}\|^2$ if TE is continuously differentiable up to order 2. By definition, $\mathrm{TE}(0, x) = 0$ because $e_i(0, \widetilde{X}_t) = 0$. Hence

$$\mathrm{TE}(\widetilde{v}, x) = F(0, x)\widetilde{v} + R(\widetilde{v}, x).$$

[1] Conditional expectations are defined up to a set of measure zero. In the following we will treat equalities concerning conditional expectations as a.s. equalities.

We call

$$F := F(0, \widetilde{X}_t) = \left.\frac{\partial \mathrm{TE}(\widetilde{v}, \widetilde{X}_t)}{\partial \widetilde{v}'}\right|_{\widetilde{v}=0}$$

the Impact Factor, IF. It represents the coefficient of the linear approximation of $\mathrm{TE}(\widetilde{v}, x)$ as a function of the perturbation $\widetilde{v}$ close to $\widetilde{v} = 0$. Under the usual regularity conditions, differentiation and summation within TE may be interchanged; in this case $F = \sum_{i=1}^{\infty} \partial e_i(\widetilde{v}, \widetilde{X}_t)/\partial \widetilde{v}'$.

Observe that F is by definition a $p \times s$ matrix, where each entry gives a particular IF. Specifically F_{ij} gives the IF of a perturbation in $\widetilde{X}_{jt}$, the j-th entry of $\widetilde{X}_t$, onto the forecast function of Y_{it}, the i-th element in Y_t. When y_t and x_t are subvectors of Y_t and $\widetilde{X}_t$ respectively we use the notation $F_{y,x} := F_{y_t, x_t}$ to indicate the corresponding submatrix of the IF matrix F.

2.1 Linear transformations

Under quite unrestrictive assumptions on the forecast function, the IF matrix F obeys a simple transformation rule under linear transformations of Y_t and/or $\widetilde{X}_t$. Let $Y_t^* := N_Y Y_t$, $\widetilde{X}_t^* := N_X \widetilde{X}_t$ be linear transformations the original variables, where the $N.$ matrices are square and non-singular. Let F^* be the IF for the starred variables; we here show that

$$F^* = N_Y F N_X^{-1} \tag{1}$$

when the forecast function is equivariant with respect to linear combinations of the forecasts, i.e. that $Y_{t+i|t}^* = N_Y Y_{t+i|t} = N_Y g_i(\widetilde{X}_t) =: g_i^*(\widetilde{X}_t)$. Conditional expectations e.g. possess this equivariant property.

The perturbation of the the input variables are simply related by $\widetilde{v}^* := \widetilde{X}_t^{*c} - \widetilde{X}_t^* = N_X(\widetilde{X}_t^c - \widetilde{X}_t) = N_X \widetilde{v}$; since N_X is nonsingular, $\widetilde{v} = N_X^{-1}\widetilde{v}^*$, and $\widetilde{v} = 0$ iff $\widetilde{v}^* = 0$. Let $\mathrm{TE}^*(\widetilde{v}^*, \widetilde{X}_t^*)$ be the total effect in terms of the starred variables; by the results above one has

$$\mathrm{TE}^*(\widetilde{v}^*, \widetilde{X}_t^*) = N_Y \mathrm{TE}(N_X^{-1}\widetilde{v}^*, N_X^{-1}\widetilde{X}_t^*).$$

Thus, applying the definition and the chain rule of differentiation,

$$\begin{aligned} F^* := & \left.\frac{\partial \mathrm{TE}^*(\widetilde{v}^*, \widetilde{X}_t^*)}{\partial \widetilde{v}^{*\prime}}\right|_{\widetilde{v}^*=0} = N_Y \left.\frac{\partial \mathrm{TE}(\widetilde{v}, \widetilde{X}_t)}{\partial \widetilde{v}'} \frac{\partial \widetilde{v}}{\partial \widetilde{v}^{*\prime}}\right|_{\widetilde{v}^*=0} = \\ = & N_Y \left.\frac{\partial \mathrm{TE}(\widetilde{v}, \widetilde{X}_t)}{\partial \widetilde{v}'}\right|_{\widetilde{v}=0} \cdot \frac{\partial N_X^{-1}\widetilde{v}^*}{\partial \widetilde{v}^{*\prime}} = N_Y F N_X^{-1} \end{aligned}$$

Hence one can derive from the IF matrix F all the IF implied by linear combinations of inputs and outputs applying the transformation (1). We also observe that one may be interested in just some linear combinations of Y_t and/or $\widetilde{X}_t$, and not the complete vector; this corresponds to selecting some rows of the $N.$ matrices in an appropriate way.[2]

We next specialize the notion of IF to the case of a linear forecast function.

[2] Note, however, that the complete specification of $\widetilde{X}_t$ is of interest in the interpretation of the IF, because the idea of perturbating just some linear combinations of $\widetilde{X}_t$ may be inconsistent with the specification of the other elements of the FS statistic.

2.2 Linear forecast function

When the forecast function is linear

$$g_i(\widetilde{X}_t) = a_i + B_i\widetilde{X}_t, \tag{2}$$

it is simple to note that $e_i(\widetilde{v}, \widetilde{X}_t) := g_i(\widetilde{X}_t^c) - g_i(\widetilde{X}_t) = B_i(\widetilde{X}_t^c - \widetilde{X}_t) = B_i\widetilde{v}$, which depends on $\widetilde{X}_t$ only through $\widetilde{v}$. Hence in this case, if B_i is summable, one finds TE$(\widetilde{v}, \widetilde{X}_t)$ =TE$(\widetilde{v}) = (\sum_{i=1}^{\infty} B_i)\widetilde{v}$, and

$$F = \sum_{i=1}^{\infty} B_i.$$

Observe that the remainder term R is zero because TE is a linear function of the perturbation $\widetilde{v}$ only. Here IF is a global sensitivity measure, since it is constant for all possible values of $\widetilde{X}_t$.

2.3 Stationary VARs

Let X_t be generated by a VAR $A(L)X_t = \mu^* D_t^* + \epsilon_t$, with deterministic component $\mu^* D_t^*$, and i.i.d. $N(0, \Omega)$ errors ϵ_t. Here and in the following we take $D_t^* := (t : 1 : d_t')'$, where $d_t := (d_{1,t} : ..d_{u-1,t})'$ is a vector of seasonal dummies 'orthogonal' to the constant, i.e. of the form $d_{i,t} = 1(t \text{ mod } u = i) - 1/u$, $1(\cdot)$ is the indicator function and u is the number of seasons.

The associated state space representation is $\widetilde{X}_t = A\widetilde{X}_{t-1} + u_t$ with state vector $\widetilde{X}_t := (X_t' : X_{t-1}' : ... : X_{t-k+1}')'$, companion matrix

$$A := \begin{pmatrix} A_1 & A_2 & \dots & A_k \\ I & & & \\ & \ddots & & \\ & & I & 0 \end{pmatrix}.$$

and $u_t := J(\mu^* D_t^* + \epsilon_t)$, $J := (I_p : 0_{p\times p(k-1)})'$, $X_t = J'\widetilde{X}_t$.

Let the variables to be forecast Y_t coincide with X_t; in this case the forecast function is $Y_{t+i|t} = E(Y_{t+i}|X_{-\infty}^t) = J'A^i\widetilde{X}_t + \sum_{j=0}^{i-1} J'A^jJ\mu^* D_{t+i-j}^*$. Note that $\widetilde{X}_t$ is the FS, and that $Y_{t+h|t} = g(\widetilde{X}_t)$ is a linear function of the FS, as in (2), with $a_i := \sum_{j=0}^{i-1} J'A^jJ\mu^* D_{t+i-j}^*$ and $B_i = J'A^i$. Hence $e_i = B_i\widetilde{v}$.

Assume also that the VAR process X_t is stationary, which implies that eigenvalue of A are less or equal to 1 in modulus. Thus

$$\text{TE} = \sum_{i=1}^{\infty} B_i\widetilde{v} = J'\left(\sum_{i=1}^{\infty} A^i\right)\widetilde{v} = J'((I-A)^{-1} - I)\widetilde{v}$$

where the series is convergent because of the stationarity assumption. In this case the IF is equal to $F := J'((I-A)^{-1} - I)$, a simple function of the companion matrix.

If the variables to be forecast are all the ones contained in the state vector, $Y_t = \widetilde{X}_t$, then the previous calculations reveal that TE $= ((I-A)^{-1} - I)\widetilde{v}$ and the IF is

$$F = (I-A)^{-1} - I. \tag{3}$$

In the present case of stationary VARs the possibility to consider all of the state vector as Y_t is not very interesting, because Y_t contains the same variables X_t at different lags. This possibility is instead of interest for non-stationary systems of order 1 and 2, considered in Section 3 below.

2.4 Impulse responses

Pesaran and Shin (1998) and Koop et al (1996) defined the scaled generalized impulse responses (GIR) for stationary VARs as

$$\psi^g(i) := J'A^iJ\Omega(diag(\Omega))^{-1/2} = J'A^iJ\Omega^*$$

where A is the companion matrix and $\Omega^* := \Omega(diag(\Omega))^{-1/2}$. This definition of impulse response does not depend on orthogonalization of shocks.

The cumulated GIR is

$$\Psi^g = \sum_{i=1}^{\infty} \psi^g(i) = \sum_{i=1}^{\infty} J'A^iJ\Omega^* = J'((I-A)^{-1} - I)J\Omega^*$$

which is proportional to the leading block of the IF matrix F in eq. (3). A similar derivation applies to the cumulated impulse responses, which converge to an expression similar to $J'((I-A)^{-1} - I)J\Omega^*$ with a different definition of the matrix Ω^*.

Unlike IF, impulse responses, IR, are usually interpreted as effects of shocks ϵ_t on the variables X_t. Nevertheless the algebra in IR is the same as in IF analysis. Thus, hopefully, the results presented below for IF in non-stationary VARs may be used also in association with impulse response analysis; see also Phillips (1998) on impulse responses in I(1) VARs.

2.5 Linearity and superposition

When the forecast function g is linear, the principle of superposition applies, see Kailath (1980); this property is reviewed in this subsection. If one considers various perturbations $\widetilde{v}_1$, ..., $\widetilde{v}_s$, their cumulated effect is equal to $\mathrm{TE}_s = F\sum_{i=1}^{s}\widetilde{v}_i$. This equals the effect TE:= $F\widetilde{v}$ of a single perturbation $\widetilde{v}$ defined as the sum of the individual perturbations, $\widetilde{v} := \sum_{i=1}^{s}\widetilde{v}_i$. Note that the IF is equal in both cases.

Consider next this equivalence specifically for VARs. Let perturbation $\widetilde{v}_j$ involve only the variables X_{t-j} at lag j, and consider various perturbations $\widetilde{v}_j$ of this sort at different lags j. The equivalence given by superposition simply says that the same IF matrix applies. In this sense, therefore, impact factors F are insensitive to the timing of the perturbations. Obviously this does not need to be the case for non-linear forecast functions.

In the rest of the paper we assume that the forecast function is the conditional expectation and that X_t is generated by a VAR.

3 Cointegrated systems

In this section we apply the definition of IF to the stationary subsystems of VAR integrated of order one and two, I(1) and I(2). We refer to Johansen (1996) for notation and definitions of I(1) and I(2) VAR systems.

3.1 Cointegrated I(1) VAR

Consider the following equilibrium correction (EC) form of the VAR:

$$\Delta X_t = \alpha\beta' X_{t-2} + \Gamma_1^* \Delta X_{t-1} + \Phi U_{t-1} + \mu_1 t + \mu D_t + \epsilon_t. \tag{4}$$

where $\Gamma_1^* := (\Gamma_1 + \Pi)$, $\Phi := (\Gamma_2 : ... : \Gamma_{k-1})$ and $U_{t-1} := (\Delta X'_{t-2} : ... : \Delta X'_{t-k+1})'$ is $m \times 1$, $m := p(k-2)$, and $\mu := (\mu_0 : \mu_d)$, $D_t := (1 : d'_t)'$.

This EC form presents the level term measured in $t-2$; this can always be accomplished by adding and subtracting appropriate terms, even in the case of $k = 1$, see Johansen (1996). This representation is chosen in order to simplify calculations in the following, and it is completely general, because results for any other EC formulation can be deduced from it, see the following Section 3.3.

We assume that the VAR process satisfies the following condition:

I(1) Assumption

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

I(1)_b : $\Pi := -A(1) = \alpha\beta'$, where α and β are $p \times p_0$ matrices of full rank $p_0 < p$.

I(1)_c : $\mu_1 = \alpha\beta'_0$ with β'_0 a $p_0 \times 1$ vector.

I(1)_d : $\alpha'_\perp \Gamma \beta_\perp$ has full rank $p - p_0$, where $\Gamma := -I + \sum_{i=1}^{k-1} \Gamma_i$.

These assumptions guarantee that ΔX_t and $\beta' X_t + \beta'_0 t$ are stationary processes, apart from the influence of initial values, and that X_t has at most a linear trend in all directions, see Johansen (1996).

The associated state space representation is $\widetilde{X}_t = A\widetilde{X}_{t-1} + u_t$ with $u_t := J(\mu^* D_t^* + \epsilon_t)$, $J := (I_p : 0)'$, and

$$\widetilde{X}_t := \begin{array}{cc} \begin{pmatrix} \Delta X_t \\ \beta' X_{t-1} \\ U_t \end{pmatrix} & \begin{array}{c} p \\ p_0 \\ m \end{array} \end{array} \qquad A := \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

$$A_{11} := \begin{array}{cc} \begin{array}{cc} p & p_0 \end{array} & \\ \begin{pmatrix} \Gamma_1^* & \alpha \\ \beta' & I_{p_0} \end{pmatrix} & \begin{array}{c} p \\ p_0 \end{array} \end{array} \qquad A_{12} := \begin{array}{cc} m & \\ \begin{pmatrix} \Phi \\ 0 \end{pmatrix} & \begin{array}{c} p \\ p_0 \end{array} \end{array} \tag{5}$$

$$A_{21} := \begin{array}{cc} \begin{array}{cc} p & p_0 \end{array} & \\ \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} & \begin{array}{c} p \\ m-p \end{array} \end{array} \qquad A_{22} := \begin{array}{cc} \begin{array}{cc} m-p & p \end{array} & \\ \begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix} & \begin{array}{c} p \\ m-p \end{array} \end{array}$$

where we have reported dimensions alongside blocks of the companion matrix.

The following proposition applies.

Proposition 1 (IF in I(1) systems) *Consider state space form* (5) *under the* $I(1)$ *assumption; then all eigenvalues of* A *are within the unit circle and the impact factor* $F := (I - A)^{-1} - I$ *has the following form: let*

$$B := \begin{pmatrix} C & (C\Gamma^\circ - I)\bar{\beta} \\ \bar{\alpha}'(\Gamma^\circ C - I) & \bar{\alpha}'(\Gamma^\circ C\Gamma^\circ - \Gamma^\circ)\bar{\beta} \end{pmatrix}$$

$c_1 := c_2 \otimes I_p$*, with* c_2 *a lower triangular matrix with ones on and below the main diagonal,* $\Gamma^\circ := -\Gamma$, $C = \beta_\perp (\alpha'_\perp \Gamma^\circ \beta_\perp)^{-1} \alpha'_\perp$, $\psi := (\psi_2 : \ldots : \psi_{k-1})$, $\psi_i = \sum_{j=i}^{k-1} \Gamma_j$*; then*

$$\begin{aligned} F + I &= \begin{pmatrix} B & B\begin{pmatrix} \psi \\ 0 \end{pmatrix} \\ (i_{k-2} \otimes I : 0)B & c_1 + i_{k-2} \otimes C\psi \end{pmatrix} \\ &= \begin{pmatrix} C & (C\Gamma^\circ - I)\bar{\beta} & C\psi \\ \bar{\alpha}'(\Gamma^\circ C - I) & \bar{\alpha}'(\Gamma^\circ C\Gamma^\circ - \Gamma^\circ)\bar{\beta} & \bar{\alpha}'(\Gamma^\circ C - I)\psi \\ i_{k-2} \otimes C & i_{k-2} \otimes (C\Gamma^\circ - I)\bar{\beta} & c_1 + i_{k-2} \otimes C\psi \end{pmatrix}. \end{aligned}$$

From this expression one can read the impact factors; in particular $F_{y,x}$ *equals*

1. $C - I$ *for* $y_t = x_t := \Delta X_t$
2. $(C\Gamma^\circ - I)\bar{\beta}$ *for* $y_t := \Delta X_t$, $x_t := \beta' X_{t-1}$
3. $\bar{\alpha}'(\Gamma^\circ C - I)$ *for* $y_t := \beta' X_{t-1}$, $x_t := \Delta X_t$
4. $\bar{\alpha}'(\Gamma^\circ C\Gamma^\circ - \Gamma^\circ)\bar{\beta}$ *for* $y_t = x_t := \beta' X_{t-1}$.

A special interpretation applies to the I(1) case. Consider $F_{y,x}$ for $y_t := \Delta X_t$, $x_t := \widetilde{X}_t$. The cumulated forecasts on the differences $\sum_{i=1}^{H} \Delta X_{t+i|t} = X_{t+H|t} - X_t$ give the forecast on the levels minus the initial value. Hence the total effect of a change in b_t is given by TE$= X^c_{\infty|t} - X_{\infty|t}$, where $X^c_{\infty|t}$ indicates the forecast on the level of X_∞ based on $\widetilde{X}^c_t$. Thus TE measures the change in the long-run forecast on the *levels*, and IF is a sensitivity measure of the level forecast with respect to changes in the FS variables.

This interpretation has been emphasized in Bedini and Mosconi (2000). In particular they focus on $F_{\Delta X_t, \beta' X_{t-1}} = (C\Gamma^\circ - I)\bar{\beta}$, which they call the *long-run adjustment coefficients* to disequilibrium errors. The approach of the present paper give a forecasting interpretation of the long-run adjustment coefficients, as well as of other IF.

3.2 Cointegrated I(2) VAR

Consider the equilibrium correction (EC) representation of the VAR suggested in Paruolo and Rahbek (1999) for I(2) systems:

$$\begin{aligned} \Delta^2 X_t &= \alpha(\beta' X_{t-1} + \delta\beta_2' \Delta X_{t-1}) + (\zeta_1 : \zeta_2)(\beta : \beta_1)' \Delta X_{t-1} + \\ &\quad + \Upsilon_1 \Delta^2 X'_{t-1} + \Phi W_{t-1} + \mu^* D^*_t + \epsilon_t \end{aligned} \tag{6}$$

where $W_{t-1} := (\Delta^2 X'_{t-2} : ... : \Delta^2 X'_{t-k+2})'$, of dimension $m \times 1$, $m := p(k-3)$, $\Phi := (\Upsilon_2 : ... : \Upsilon_{k-2})$. $\mu^* := (\mu_1 : \mu_0 : \mu_d)$, $D_t^* := (t : 1 : d_t')'$.

We first list some assumptions. Let $\phi := I - \sum_{i=1}^{k-2} \Upsilon_i$.

I(2) Assumption

I(2)_a : Assumptions I(1)_a, I(1)_b, I(1)_c hold.

I(2)_b : $P_{\alpha_\perp} \Gamma P_{\beta_\perp} = \alpha_1 \beta_1'$ 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)_c : $\alpha_2' \theta \beta_2$ has full rank $p_2 := p - p_0 - p_1$, where $\alpha_2 = (\alpha : \alpha_1)_\perp$, $\beta_2 = (\beta : \beta_1)_\perp$ and θ is defined as

$$\theta := \Gamma \bar{\beta} \bar{\alpha}' \Gamma + \phi. \tag{7}$$

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

Johansen's I(2) representation theorem, see Johansen (1992) or Johansen (1996) Theorem 4.6, establishes that under I(2)_a that necessary and sufficient conditions for

$$\Delta^2 X_t, \quad \beta' X_t + \delta \beta_2' \Delta X_t + \beta_0' t, \quad \beta_1' \Delta X_t \tag{8}$$

to be stationary, apart from initial values, and for X_t to have at most a linear trend in all directions are the conditions I(2)_b to d; see e.g. Paruolo (2002b) for a proof.[3] In the following 'I(2) assumption' and 'I(2) conditions' are used as synonyms.

The EC formulation in (6) imposes some of the I(2) restrictions; we refer to Paruolo and Rahbek (1999) for complete definitions of coefficients and background. As for the I(1) case we choose a specific timing of the EC terms in order to simplify later calculations. Again this is done without loss of generality, since results for any other EC formulation can be deduced from it, see again Section 3.3.

Proposition 7 in the Appendix shows that one of the many possible equivalent EC formulation of this system is

$$\begin{aligned} \Delta^2 X_t \quad = \quad & \alpha(\beta' X_{t-3} + \delta \beta_2' \Delta X_{t-2} + \beta_0' t) + (\zeta_1^* : \zeta_2)(\beta : \beta_1)' \Delta X_{t-2} + \\ & + \Upsilon^* \Delta^2 X_{t-1}' + \Phi W_{t-1} + \mu D_t + \epsilon_t, \end{aligned} \tag{9}$$

where we have imposed $\mu_1 = \alpha \beta_0'$. The timing of the EC terms $(\beta' X_{t-3} + \delta \beta_2' \Delta X_{t-2})$, $(\beta : \beta_1)' \Delta X_{t-2}$ is different from the one in (6) and $\zeta_1^* := \zeta_1 + 2\alpha$ and $\Upsilon_1^* := (\Upsilon_1 + \Gamma + \Pi)$. Note that this affects the definition only of ζ_1^* and Υ_1^* and not of the remaining coefficients. This timing can always be achieved, also for $k = 2$.[4] We summarize notation in Table 1.

[3] Note that the stationarity of the variables in (8) implies that also $\beta' \Delta X_t$ is stationary.

[4] Following the literature, we do not consider $k = 1$ in the I(2) case.

	symbol	dim		symbol	dim
Γ	$= \alpha\delta\beta_2' + \zeta_1\beta' + \zeta_2\beta_1'$	$p \times p$	Φ	$:= (\Upsilon_2 : ... : \Upsilon_{k-2})$	$p \times p(k-3)$
ζ_1^*	$:= \zeta_1 + 2\alpha$	$p \times p_0$	Υ_1^*	$:= (\Upsilon_1 + \Gamma + \alpha\beta')$	$p \times p$
ϕ	$:= I - \sum_{i=1}^{k-2} \Upsilon_i$	$p \times p$	ϕ^*	$:= \phi - \Gamma - \alpha\beta'$	$p \times p$
θ	$:= \zeta_1\bar{\alpha}'\Gamma + \phi$	$p \times p$	θ^*	$:= \zeta_1^*\bar{\alpha}'\Gamma + \phi^*$	$p \times p$
C_2	$:= \beta_2 \left(\alpha_2'\theta\beta_2\right)^{-1} \alpha_2'$	$p \times p$	τ	$:= (\beta : \beta_1)$	$p \times (p_0 + p_1)$
h	$:= \theta^* C_2 - I$	$p \times p$	q	$:= \bar{\alpha}'(I - \zeta_2\bar{\alpha}_1')h$	$p_0 \times p$
ψ_i	$:= \sum_{j=i}^{k-2} \Upsilon_j$	$p \times p$	ψ	$:= (\psi_2 : ... : \psi_{k-2})$	$p \times p(k-3)$

Table 1: Symbol definitions for the expression of the IF in the I(2) systems.

The system can be cast in the state space form $\widetilde{X}_t = A\widetilde{X}_{t-1} + u_t$ with $u_t := J(\mu^* D_t^* + \epsilon_t)$ and

$$\widetilde{X}_t := \begin{array}{c} \left(\begin{array}{c} \Delta^2 X_t \\ \beta'\Delta X_{t-1} \\ \beta_1'\Delta X_{t-1} \\ \beta' X_{t-2} + \delta\beta_2'\Delta X_{t-1} \\ W_t \end{array} \right) \end{array} \begin{array}{c} p \\ p_0 \\ p_1 \\ p_0 \\ m \end{array} \qquad A := \left(\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array} \right),$$

$$A_{11} := \begin{array}{cc} \begin{array}{cccc} p & p_0 & p_1 & p_0 \end{array} & \\ \left(\begin{array}{cccc} \Upsilon_1^* & \zeta_1^* & \zeta_2 & \alpha \\ \beta' & I_{p_0} & & \\ \beta_1' & & I_{p_1} & \\ \delta\beta_2' & I_{p_0} & & I_{p_0} \end{array} \right) & \begin{array}{c} p \\ p_0 \\ p_1 \\ p_0 \end{array} \end{array} \qquad A_{12} := \begin{array}{cc} m & \\ \left(\begin{array}{c} \Phi \\ 0 \end{array} \right) & \begin{array}{c} p \\ 2p_0 + p_1 \end{array} \end{array} \tag{10}$$

$$A_{21} := \begin{array}{cc} \begin{array}{cc} p & 2p_0 + p_1 \end{array} & \\ \left(\begin{array}{cc} I & 0 \\ 0 & 0 \end{array} \right) & \begin{array}{c} p \\ m-p \end{array} \end{array} \qquad A_{22} := \begin{array}{cc} \begin{array}{cc} m-p & p \end{array} & \\ \left(\begin{array}{cc} 0 & 0 \\ I & 0 \end{array} \right) & \begin{array}{c} p \\ m-p \end{array} \end{array}$$

where we have reported dimensions; 0 entries are not reported unless when needed for clarity. The following proposition applies.

Proposition 2 (IF in I(2) systems) *Consider the state space form* (10) *under the* $I(2)$ *assumption; then all eigenvalues of* A *are within the unit circle and the impact factor* $F := (I - A)^{-1} - I$ *has the following form: let*

$$B := \left(\begin{array}{ccc} C_2 & (C_2\phi^* - I)\bar{\tau} & -C_2\zeta_1 \\ -\delta\beta_2' C_2 & -\delta\beta_2' C_2 \phi^* \bar{\tau} & \delta\beta_2' C_2 \zeta_1 - I \\ -\alpha_1' h & -\alpha_1' h \phi^* \bar{\tau} & \alpha_1' h \zeta_1^* \\ -q & -q\phi^*\bar{\tau} & q\zeta_1^* \end{array} \right)$$

where $q := \bar{\alpha}'(I - \zeta_2\bar{\alpha}_1')h$, $h := I - \theta^* C_2$, $q := \bar{\alpha}'(I - \zeta_2\bar{\alpha}_1')h$, $\tau := (\beta, \beta_1)$, $\theta^* := \phi^* + \zeta_1^*\bar{\alpha}'\Gamma$, $C_2 := \beta_2 \left(\alpha_2'\theta\beta_2\right)^{-1} \alpha_2'$*; let also* $\psi := (\psi_2, ..., \psi_{k-2})$, $\psi_i = \sum_{j=i}^{k-2} \Upsilon_j$, $c_1 := c_2 \otimes I_p$, *where* c_2 *is a lower triangular matrix with ones on and below the main* *diagonal; then $F+I$ equals*

$$\begin{pmatrix} B & B\begin{pmatrix}\psi\\0\end{pmatrix} \\ (i_{k-2}\otimes I:0)B & c_1+i_{k-2}\otimes C_2\psi \end{pmatrix} \tag{11}$$

$$= \begin{pmatrix} C_2 & (C_2\phi^*-I)\bar{\tau} & -C_2\zeta_1 & C_2\psi \\ -\delta\beta_2'C_2 & -\delta\beta_2'C_2\phi^*\bar{\tau} & -I+\delta\beta_2'C_2\zeta_1 & -\delta\beta_2'C_2\psi \\ -\alpha_1'h & -\alpha_1'h\phi^*\bar{\tau} & \alpha_1'h\zeta_1 & -\alpha_1'h\psi \\ -q & -q\phi^*\bar{\tau} & q\zeta_1^* & -q\psi \\ i_{k-2}\otimes C_2 & i_{k-2}\otimes(C_2\phi^*-I)\bar{\tau} & -i_{k-2}\otimes C_2\zeta_1 & c_1+(i\otimes I_p)C_2\psi \end{pmatrix}.$$

From this expression one can read the impact factors; in particular $F_{y,x}$ equals

1. C_2-I *for* $y_t=x_t:=\Delta^2X_t$;
2. $-\delta\beta_2'C_2$ *for* $y_t:=\beta'\Delta X_t$, $x_t:=\Delta^2X_t$
3. $\bar{\alpha}_1'(\theta^*C_2-I)$ *for* $y_t:=\beta_1'\Delta X_t$, $x_t:=\Delta^2X_t$
4. $\bar{\alpha}'(I-\zeta_2\bar{\alpha}_1')(\theta^*C_2-I)$ *for* $y_t:=\beta'X_{t-2}+\delta\beta_2'\Delta X_{t-1}$, $x_t:=\Delta^2X_t$.

Again we note that IF of the type $F_{b'\Delta X,x}$ present the level interpretation given for I(1) systems: they measure the change in the long-run forecast of $b'X_t$ induced by a change in x_t. We observe that there are several long-run adjustment coefficients to various disequilibrium errors; they appear in the second and third column in formula (11). One can note that timing of the EC terms used in (9) is perhaps not the most natural. The following subsection discusses the relation among IF obtained for the various choices of timing of the EC terms, both for the I(1) and the I(2) cases.

3.3 Timing of the EC terms

The choice of timing of the EC terms in an EC formulation is arbitrary. It is well known that in the I(1) case the level term $\beta'X_{t-1}$ can be shifted to any lag j, $1\le j\le k$, by changing the definition of the coefficients to the variables ΔX_{t-1}, ..., ΔX_{t-k+1}. The same applies to the EC terms $(\beta,\beta_1)'\Delta X_{t-1}$ and $(\beta'X_{t-1}+\delta\beta_2'\Delta X_{t-1})$ in the I(2) systems: the level term X_{t-j} can be shifted to any lag j, $1\le j\le k$ and the differences ΔX_{t-j} to any lag j, $1\le j\le k-1$. The choices made in the previous sections were only done for ease of calculations.

Let Z_t and S_t be two possible choices of the state vector $\widetilde{X}_t$ corresponding to a specific timings of the EC terms. It is simple to see that they are connected by a linear map $Z_t=NS_t$, where N is square and non-singular, see examples in Appendix B. The two state vectors satisfy recursions $Z_t=A_ZZ_{t-1}+u_t$, and $S_t=A_SS_{t-1}+u_t$. Substituting $Z_t=NS_t$ in the first equation one sees that $NS_t=A_ZNS_{t-1}+u_t$ or $S_t=N^{-1}A_ZNS_{t-1}+N^{-1}u_t$, i.e. the companion matrices are related by $A_S=N^{-1}A_ZN$, or $NA_SN^{-1}=A_Z$. This implies a similar relation between the corresponding IF, which is a special case of the basic property (1), with $N_X=N_Y=N$.

Let F_Z and F_S indicate the IF calculated for state vectors Z_t and S_t. The following proposition applies.

Proposition 3 (timing and IF) *One has $F_Z = NF_SN^{-1}$ for $Z_t := NS_t$.*

The previous proposition shows that one can transform IF just as easily as one can redefine the timing of EC terms. A few leading examples of transformation N are described in the Appendix B, which collects also proofs of this subsection. Two remarks emerge from the analysis of these cases.

- The choice of timing of the EC term involves a transformation matrix N that contains either known elements (0 and 1s) or cointegrating parameters, β in the I(1) case and β, β_1, β_2 and δ in the I(2) case.
- The inverse N^{-1} of N is easily calculated, and often corresponds to a matrix with the same entries of N with same sign on the main diagonal and opposite sign in the rest of the matrix.

It is thus possible to calculate a single set of IF and deduce others possible choices from this set. The following proposition states which of these IF are invariant with respect to the choice of lag of the EC terms.

Proposition 4 (IF invariant w.r.t timing of EC terms) *1. In the $I(1)$ case, for any state space vector of the form*

$$\left(\Delta X_t' : X_{t-j}'\beta : U_t'\right)', \qquad j = 0, 1, \ldots, k$$

the IF F_{y_t,x_t} are invariant for $y_t = \Delta X_t$, U_t and $x_t = \Delta X_t$, $\beta' X_{t-j}$.

2. In the $I(2)$ case for any state space vector of the form

$$\left(\Delta^2 X_t' : \Delta X_{t-i}'\beta : \Delta X_{t-j}'\beta_1 : X_{t-l}'\beta + \Delta X_{t-m}'\beta_2\delta' : W_t'\right)',$$
$$i, j, m = 1, \ldots, k-2, \quad l = 1, \ldots, k-1$$

the IF F_{y_t,x_t} is invariant for $y_t = \Delta^2 X_t$, W_t and $x_t = \Delta^2 X_t$, $\beta_1' X_{t-j}$, $\beta' X_{t-l} + \delta\beta_2' \Delta X_{t-m}$.

This shows that some IF are invariant w.r.t choice of lags, while others are not. Note that in the I(1) case the long-run adjustment coefficient $F_{\Delta X_t, \beta X_{t-j}}$ is invariant. In the I(2) case the long-run adjustment coefficient for the multicointegration relation $F_{\Delta^2 X_t, \beta' X_{t-l} + \delta\beta_2' \Delta X_{t-m}}$ is also invariant. Note that the other long-run adjustment coefficient $F_{\Delta^2 X_t, \beta_1' X_{t-j}}$ is invariant, whereas $F_{\Delta^2 X_t, \beta' X_{t-j}}$ is not.

4 Areas of application

This section reports two possible areas of applicability interpretation of the IF. They regard the effectiveness of economic policy in the long run and the impact of data revisions on forecasts.

4.1 Policy effectiveness

The analysis of IF can be applied to policy analysis. Perturbations of the input variables $\widetilde{v}$ may be induced by policy interventions; in this case the IF captures the long-run response of the forecast function to policy interventions. The superposition principle for linear forecast functions applied here implies that one can restrict attention to single perturbations $\widetilde{v}$.

We observe that the perturbations $\widetilde{v}$ may involve variables at different points in time: for the policy intervention interpretation to apply, one needs to restrict attention to perturbations $\widetilde{v}$ that regard the most recent time subscript, i.e. of the form $\widetilde{v} = Jv$, where $J := (I_p : 0)'$. This type of perturbation corresponds to a factual experiment, in which some variables (instruments) are changed by the policy maker. We hence call this type of perturbation "factual".

On the contrary all perturbations $\widetilde{v}$ that are not of the form Jv are "counterfactual", in the sense that they cannot be obtained by actual policy actions, which affect variables at a single point in time. The counterfactual perturbations correspond to a thought experiment where variables at different lags are perturbed simultaneously. In the following we consider both factual and counterfactual perturbations.

If some perturbation induced by policy action does not affect the accumulated forecast on some "target" variables, this means that the policy is ineffective in the long run. If the system is I(1) and the target variable is the growth rate of some non-stationary variable, policy ineffectiveness is measured with respect to the long-run forecast of the level associated with the target variable.

Therefore it appears of importance to test if some IF are significantly different from zero. In this interpretation, insignificant IF would correspond to ineffectiveness of policies. Inference on the IF is treated in Section 5.

4.2 Data revisions

The perturbations $\widetilde{v}$ may be interpreted as induced by data revisions. Several macroeconomic indicators are first published in preliminary form, and next adjusted, e.g. on the basis of national account available at the end of the year. Because the data $\widetilde{X}_t$ containing preliminary data is fed into a forecast function in order to produce preliminary forecast of major macroeconomic aggregates, IF can be interpreted in this case as a sensitivity measure of the cumulated forecast profile to (small) revisions of the data.

Let $\widetilde{X}_t^c$ be the revised data. The TE can now be interpreted as the cumulated change in forecasts of Y_{t+h} due to the revision of preliminary figures. IF can thus be used to measure if the cumulated change in forecasts is significantly different from zero, and what sort of variability is induced in the forecast profile by the revisions of the data.

This interpretation can also be combined with the fact that, for cointegrated I(1) systems, IFs measure the sensitivity of the long-run forecast of the levels of the variables. A similar comment applies to growth rates in I(2) systems.

5 Inference on the IF

In this section we consider inference on IF in a unified framework for stationary, I(1) and I(2) systems. The approach is based on the observation that CI parameters are estimated super-efficiently. This implies that the inclusion of estimated CI parameters in the definition of regressors does not affect the limit distribution of the IF. Inference on the IF is associated with the one on the companion matrix A. This matrix is estimated below through a specific regression system, which is specified in the next subsection for the I(0), I(1) and I(2) cases. In Subsection 5.2, we then address the issue of inference on the IF F, which is calculated as $(I-A)^{-1}-I$.

5.1 Regression setup

In order to estimate the IF, one needs to estimate the companion matrix A. We define $G^* := J'A$ and $L := J_\perp' A$, where $J := (I_p : 0)'$ and $J_\perp = (0 : I)$. The matrix G^* contains the adjustment coefficients, while L contains only known values, 0 or 1, and CI parameters in the integrated cases. The matrix A is then reconstructed as $A = (G^{*\prime} : L')'$.

In the stationary case let $X_{0t} := J'\widetilde{X}_t = X_t$ be the regression dependent variable and $X_{1t} := (X_{t-1}' : ... : X_{t-k}')'$ be the matrix of stochastic regressors. For homogeneity with the integrated cases we assume that $\mu_1 = 0$, so that the system equations can be written as

$$X_{0t} = GX_{1t} + \mu D_t + \epsilon_t, \tag{12}$$

where $G := (A_1 :... : A_k)$. The likelihood analysis of the stationary VAR in (12) is simply performed by OLS. For later reference we also set $H := I$, $G^* := G$, $\widehat{X}_{1t} := X_{1t}$.

Consider now the integrated cases. The I(1) cointegration analysis with the deterministic specification used above is described in Johansen (1996), while the corresponding one for the I(2) model is described in Rahbek et al. (1999)[5].

Consider the I(1) case. Let $X_{0t} := J'\widetilde{X}_t = \Delta X_t$ be the regression dependent variable. The I(1) analysis permits to determine the CI rank p_0 and to estimate $\beta^* := (\beta' : \beta_0')'$. These estimates are at least T consistent, see Johansen (1996). The estimate of β^* permits to calculate the regressor vector $\widehat{X}_{1t}^* := (\Delta X_{t-1}' : (\widehat{\beta}' X_{t-2} + \widehat{\beta}_0' t)' : U_{t-1}')$, and eq. (4) can be rewritten as

$$X_{0t} = G^* \widehat{X}_{1t} + \mu D_t + \widehat{\epsilon}_t \tag{13}$$

where $G^* = (\Gamma_1^* : \alpha : \Phi)$ and $\widehat{\epsilon}_t := \epsilon_t - \alpha((\widehat{\beta}-\beta)' X_{t-2} + (\widehat{\beta}_0 - \beta_0)' t)'$ is the error term. Here and in the following we indicate with $\widehat{\ }$ quantities where the CI coefficients have been substituted with their estimators.

In the special case $k=1$ listed in Appendix C, G^* has reduced rank because of the reduced rank of $A = \widetilde{A}H'$, $G^* := J'A = J'\widetilde{A}H'$. In this case define $\widehat{X}_{1t} := \widehat{H}'\widehat{X}_{1t}^*$,

[5] The estimation of the cointegrating coefficients can be accomplished via likelihood techniques in I(1) and I(2) systems or via the 2SI2 procedure in I(2) systems, see Johansen (1995), Paruolo (1996), Rahbek et al. (1999).

$G := J'\widetilde{A}$; otherwise we let $H = I$ and define $G := G^*$, $\widehat{X}_{1t} := \widehat{X}^*_{1t}$. Eq (13) then reads

$$X_{0t} = G\widehat{X}_{1t} + \mu D_t + \widehat{\epsilon}_t. \tag{14}$$

Consider the I(2) case. We define $G^* := J'A$ and let $X_{0t} := J'\widetilde{X}_t = \Delta^2 X_t$ be the regression dependent variable. The I(2) analysis permits to determine the II p_0 and p_1 and to estimate $\beta^* := (\beta' : \beta_0')'$ and δ, β_1, β_2. These estimates are at least T consistent, see Johansen (1997) and Paruolo (2000). These estimates permits to calculate the regressor vector $\widehat{X}^*_{1t} := (\Delta^2 X'_{t-1} : \Delta X'_{t-2}(\widehat{\beta} : \widehat{\beta}_1) : (\widehat{\beta}' X_{t-3} + \widehat{\beta}_0' t + \widehat{\delta}\widehat{\beta}_2' \Delta X_{t-2})' : W'_{t-1})$, and eq. (9) can be rewritten as (13) where $G^* = (\Upsilon_1^* : \zeta_1^* : \zeta_2 : \alpha : \Phi)$ and the error term $\widehat{\epsilon}_t$ depends on ϵ_t and on the estimation error of the CI parameters.

In the special case $k = 2$ listed in Appendix C, G^* has reduced rank because of the reduced rank of $A = \widetilde{A}H'$, $G^* := J'A = J'\widetilde{A}H'$. In this case define $\widehat{X}_{1t} := \widehat{H}'\widehat{X}^*_{1t}$, $G := J'\widetilde{A}$; otherwise we let $H = I$ and define $G := G^*$, $\widehat{X}_{1t} := \widehat{X}^*_{1t}$. Eq (13) can then be transformed in (14) as in the I(1) case.

In all cases the matrix A is the reconstructed as

$$A = \begin{pmatrix} G^* \\ L \end{pmatrix} = \begin{pmatrix} GH' \\ L \end{pmatrix}.$$

5.2 Inference

Eq. (14) is the regression equation on which we base inference on the IF. For known CI coefficients the ML estimates of G and Ω are computed by OLS,

$$\widehat{G} = \widehat{S}_{01}\widehat{S}_{11}^{-1} \qquad \widehat{\Omega} = \widehat{S}_{00.1} := \widehat{S}_{00} - \widehat{S}_{01}\widehat{S}_{11}^{-1}\widehat{S}_{10},$$

where $S_{ij} := T^{-1}\sum_{t=1}^T R_{it}R'_{jt}$, $R_{it} := X_{it} - M_{iD}M_{DD}^{-1}D_t$, $M_{iD} := T^{-1}\sum_{t=1}^T X_{it}D'_t$, $M_{DD} := T^{-1}\sum_{t=1}^T D_tD'_t$, and $\widehat{\ }$ indicates quantities where the CI coefficients have been substituted with their estimators. Similarly $\widehat{H}$ and $\widehat{L}$ indicate the H and L matrices with CI coefficients have been substituted with their estimators.

The expressions of the regression estimators for the stationary case in (12) are identical, but obviously do not involve moments with pre-estimated CI coefficients. An analogous comment applies to the H and L matrices.

The corresponding estimate of A is

$$\widehat{A} = \begin{pmatrix} \widehat{G}\widehat{H}' \\ \widehat{L} \end{pmatrix}$$

and $\widehat{F} = (I - \widehat{A})^{-1} - I$. We next introduce some notation. Let $Z_{1t} := H'X^*_{1t}$ and $\Sigma := E((Z_{1t} - E(Z_{1t}))((Z_{1t} - E(Z_{1t}))'$

The following theorem states the relevant limit distributions for inference on the impact factors.

Theorem 5 (limit distribution of IF) *In the $I(1)$ and $I(2)$ cases the estimator $\widehat{H}$ and $\widehat{L}$ are superconsistent, i.e. $\widehat{H} - H$, $\widehat{L} - L \in O_p(T^{-1})$. In the $I(0)$, $I(1)$ and*

$I(2)$ cases the estimator of the adjustment coefficients $\widehat{G}$ is $T^{1/2}$-consistent and has a Gaussian limit distribution

$$T^{1/2}vec(\widehat{G}' - G') \xrightarrow{w} N(0, \Omega \otimes \Sigma^{-1}). \tag{15}$$

Moreover

$$T^{1/2}vec(\widehat{F}' - F') \xrightarrow{w} N\left(0, KJ\Omega J'K' \otimes K'H\Sigma^{-1}H'K\right) \tag{16}$$

where $K := (I-A)^{-1}$. The asymptotic covariance matrix of the impact factors can be estimated consistently by substituting parameter matrices with their regression-based consistent estimators, $\widehat{\Sigma} := \widehat{S}_{11}$, $\widehat{K} = \widehat{F} + I$, $\widehat{\Omega} = \widehat{S}_{00.1}$ within (16).

We observe that the asymptotic covariance matrix of F is singular. This singularity is due to several factors. The first source of singularity is due to the fact that L is known in the I(0) case and it is estimated superconsistenty in the integrated cases. This singularity is reflected in the matrix $J := (I : 0)'$ in the expression of the asymptotic covariance matrix. A similar phenomenon appears in connection with H for the special cases here H is not the identity matrix.

Other singularities are associated with the singularities of the matrix C in the I(1) case and of C_2 in the I(2) cases. Instead of focusing on these cases we refer to Paruolo (1997a,b) for inference on C and to Paruolo (2002a) for inference on C_2.

The results in the theorem allow to define Wald-type statistics for individual IF. For simple hypothesis the type $F_{ij} = c$, for instance, if the corresponding asymptotic variance σ^2 is non-zero, one can define an asymptotically $\chi^2(1)$ statistic $(\widehat{F}_{ij}-c)^2/\widehat{\sigma}^2$, or the corresponding $N(0,1)$ statistic $(\widehat{F}_{ij} - c)/\widehat{\sigma}$. These statistics are illustrated with an application in the next section.

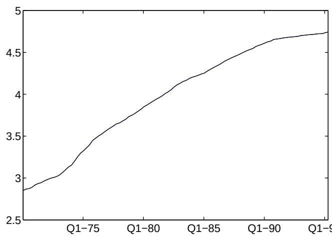
6 An application: price mark-up in Australia

As an example of IFs, we consider the data set analyzed by Banerjee et al. (2001).[6] It consists of three Australian macroeconomic data series: the consumer price deflator at factor cost (lpfc), unit labor costs in the non-farm sector (lulc) and import prices (lpm). All three variables are quarterly data measured in natural logs, and run from 1970Q1 to 1995Q2 for a total of 102 observations. The variables are graphed in levels and first differences in Fig. 1. The levels of the variables appear non-stationary, and also the differences show signs of possible non-stationarity.

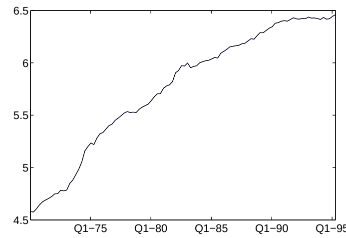
We include dummy variables to take account of a number of shocks to the economy, like the oil shocks. The dummies take value 1 in one quarter and zero otherwise; the quarters are 1974Q2, 1974Q3, 1982Q1, 1983Q2, 1985Q2 and 1986Q3.[7] We fit an unrestricted VAR in levels with $k = 2$ lags, seasonal dummies, a constant and a trend. We employ the package Me2 (Omtzigt, 2002), which performs maximum likelihood analysis also for the I(2) models.

[6]The data set is available at the data archive of the Journal of Applied Econometrics: http://qed.econ.queensu.ca/jae

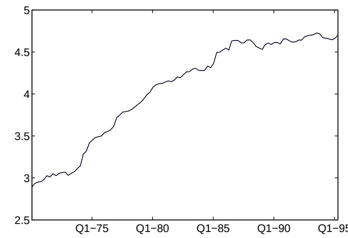
[7]Banerjee et al. (2001) condition on a number of stationary variables, but write on page 230 that "The cointegration results are essentially the same if the analysis is repeated with all the predetermined variables excluded." Our analysis finds the same selection of II and does not reject the nominal-to-real transformation, as in their paper.



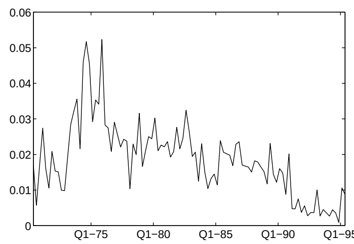
(a) lpfc: consumer prices



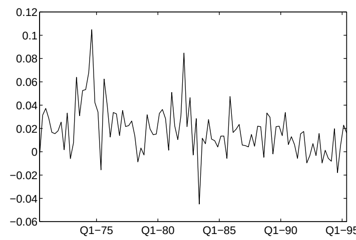
(b) lulc: unit labor cost



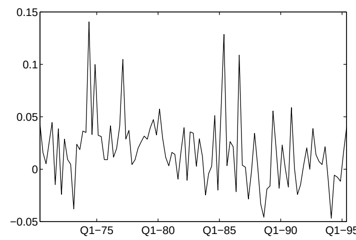
(c) lpm: import prices



(d) Δlpfc

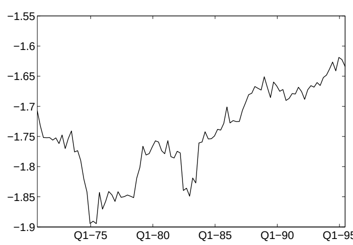


(e) Δlulc

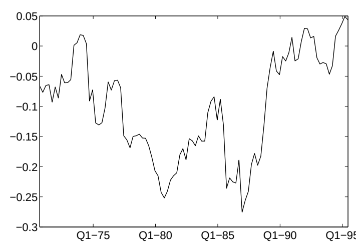


(f) Δlpm

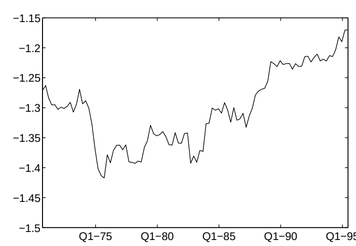
Figure 1: Australian data in levels and differences



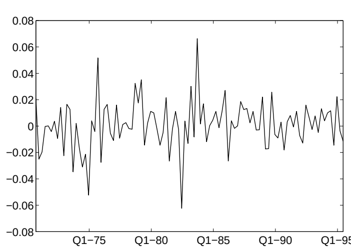
(a) $b_1'X_{t-1}$: markup on labor cost



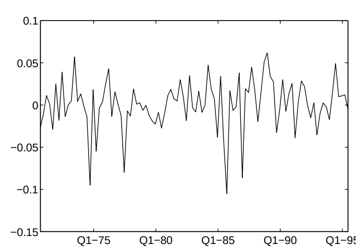
(b) $b_2'X_{t-1}$: markup on import prices



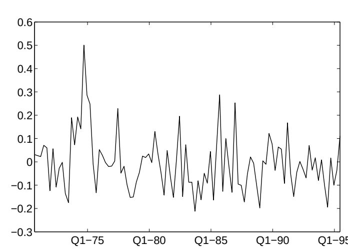
(c) $\beta' X_{t-1}$: combined markup



(d) $b_1'\Delta X_{t-1}$



(e) $b_2'\Delta X_{t-1}$



(f) mec_{t-1}, multicointegration relation

Figure 2: Cointegration and multicointegration relationships

$p_1{+}p_2$	p_0				
3	0	270.5 (87.6)	161.8 (68.2)	78.8 (53.2)	62.9 (42.7)
2	1		115.0 (47.6)	**34.1** (34.4)	30.2 (25.4)
1	2			74.7 (19.9)	11.4 (12.5)
p_2		3	2	1	0

Table 2: 2SI2 inference on the integration indices p_0, p_1. The first unrejected model is shown in boldface.

We next perform some mis-specification tests for normality and autocorrelation of the errors proposed by Doornik and Hansen (1994) and Doornik (1996). The normality test statistic is equal to 7.17 with a p-value of 0.31; the AR1 and AR4 test statistics are equal to 4.80 and 37.03, with p-values equal to 0.85 and 0.42. These results indicate that the model appears to be well specified.

6.1 Cointegration analysis

Since I(1) behavior of the growth rates implies that the levels are I(2), see Fig. 1, we leave open the possibility to select an I(2) model for the data. We first test for the number of unit roots allowing both I(1) and I(2) behavior, by selecting the integration indices of the system. This analysis considers all I(1) and I(2) submodels of the unrestricted VAR.

The selection of the integration indices is based on the 2SI2 estimator (Johansen 1995, Paruolo 1996, Rahbek et al. 1999); the test statistics for the specification $\mu_1 = \alpha\beta_0'$ are reported in table 2. Below each entry we report the 95% quantile of the asymptotic distribution, taken from Rahbek et al. (1999). We select $(p_0, p_1) = (1, 1)$, which corresponds to one I(1) trend and one I(2) trend. The roots of the characteristics polynomial are $1, 1, 1, 0.35, -0.21$ and 0.12 such that there is no trace of more non-stationary trends.[8] The same integration indices were selected by Banerjee et al. (2001).

We tested the nominal-to-real transformation (Kongsted 1998, 2002), i.e that lpfc−lulc (the markup of price on unit labor cost) and lpfc−lpm (the markup of price over import prices) are at most I(1). We used the likelihood ratio statistic; under the null the test has a $\chi^2(2)$-distribution, see Johansen (2002). The test statistic takes the value 1.242, with a p-value of 0.54, giving ample support to the transformation. This implies that $\beta = b\rho$, and $\beta_2 = b_\perp = (1 : 1 : 1)'$, where

$$b := \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The maximum likelihood, ML, estimates of the cointegration parameters are reported in Table 3. The $CI(2,1)$ relations, that is the cointegration relations from

[8] The roots of the unrestricted polynomial are 1.01, $0.87 \pm 0.04i$, 0.41, -0.21 and 0.15.

			lpfc	lupc	lpm	t
ρ	0.7296	b'	1	-1	0	0.00076
	0.2704		1	0	-1	-0.00030
δ	2.6679	β_2'	1	1	1	

Table 3: Estimates of the cointegration parameters under the nominal to real transformation; b is a basis of $sp(\tau)$, $\beta = b\rho$.

$I(2)$ to $I(1)$, are the two markups, pictured in figure 2; they are I(1). The combined mark-up on price β, obtained as a linear combination of the two, $\widehat{\beta} = b\widehat{\rho}$, is also $I(1)$:

$$\widehat{\beta}' X_{t-2} + \widehat{\beta}_0 t = \mathrm{lpfc}_{t-2} - 0.73\mathrm{lulc}_{t-2} - 0.27\mathrm{lpm}_{t-2} + 0.0005t.$$

The remaining relationship $\widehat{\beta}_1 = \bar{b}\widehat{\rho}_\perp$ is also I(1), where

$$\widehat{\beta}_1' X_{t-2} = -0.27\mathrm{lpfc}_{t-2} - 0.73\mathrm{lulc}_{t-2} + \mathrm{lpm}_{t-2}.$$

The fact that the combined mark-up $\widehat{\beta}' X_t$, is still I(1) by itself is consistent with imperfect competition theories, which predict that a high mark-up is associated with low inflation.[9] The combined markup $\widehat{\beta}' X_t$ next cointegrates with the I(1) trend in the first differences, represented by $\widehat{\beta}_2' \Delta X_t = (1 : 1 : 1)\Delta X_t$, proportional to the average inflation in the 3 series. This gives the following stationary multicointegration relationship

$$\begin{aligned} mec_t = \widehat{\beta}' X_{t-2} + \widehat{\delta}\widehat{\beta}_2' \Delta X_{t-1} + \widehat{\beta}_0' t &= \mathrm{lpfc}_{t-2} - 0.73\mathrm{lulc}_{t-2} - 0.27\mathrm{lpm}_{t-2} \\ &+ 2.67\left(\Delta\mathrm{lpfc}_{t-1} + \Delta\mathrm{lulc}_{t-1} + \Delta\mathrm{lpm}_{t-1}\right) + 0.0005t. \end{aligned} \tag{17}$$

6.2 Impact Factors

In Table 4 we report the impact factors of the restricted I(2) model, that is the model with the nominal-to-real transformation imposed. The first three columns in Table 4 are the impact factors which correspond to a factual experiment. The last three columns correspond to counterfactual experiments.

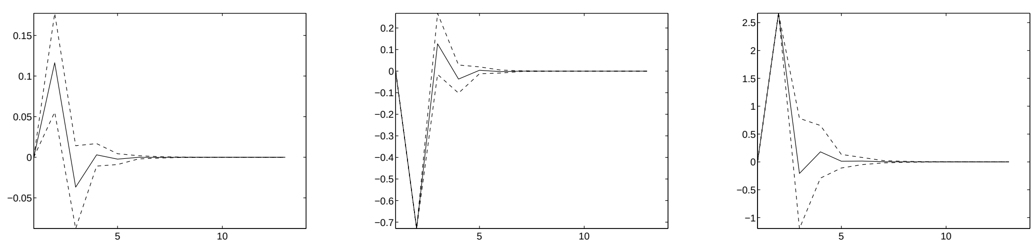
The consequences of a perturbation to the general price level can be read off from the first column in Table 4. Such a perturbation does not lead to significantly higher unit labor costs. The same insignificant effect is found for all values in the same column except the one for the multicointegration relationship (17), whose accumulated effect is significantly positive, in line with economic expectations.

Conversely a perturbation to unit labor cost and import prices have significant effects on price inflation, as can be seen from the first entries in the second and third columns. The impact factors of the second and third column are commented below in association with the IR and the accumulated IR, AIR, graphed in Fig. 3 and 4. Standard errors for IR are calculated as in Lütkepohl (1991). IFs appear in these

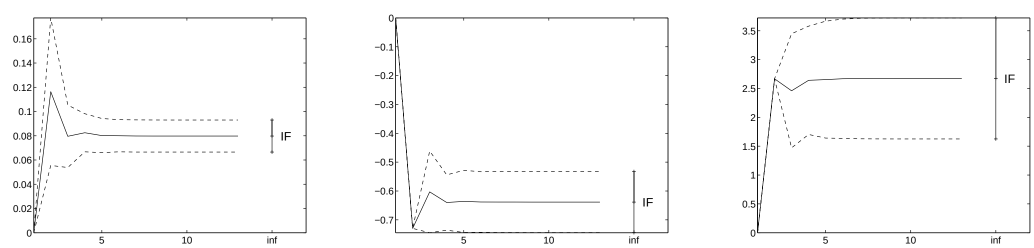
[9] For a full overview of the economic theory, we refer to Banerjee et al. (2001).

	$\Delta^2\text{lpfc}_t$	$\Delta^2\text{lulc}_t$	$\Delta^2\text{lpm}_t$	$\beta'\Delta X_{t-1}$	$\beta_1'\Delta X_{t-1}$	mec_t
$\Delta^2\text{lpfc}_t$	0.007 (0.20)	0.08 (12.06)	0.02 (3.76)	-0.77 (-27.85)	0.14 (16.09)	-0.11 (-26.32)
$\Delta^2\text{lulc}_t$	0.007 (0.20)	0.08 (12.06)	0.02 (3.76)	0.31 (11.13)	0.43 (49.52)	-0.11 (-26.32)
$\Delta^2\text{lpm}_t$	0.007 (0.20)	0.08 (12.07)	0.02 (3.76)	0.02 (0.78)	-0.65 (-74.11)	-0.11 (-26.32)
$\beta'\Delta X_{t-1}$	-0.057 (-0.20)	-0.64 (-12.06)	-0.17 (-3.76)	1.17 (5.31)	0.20 (2.83)	-0.11 (-3.18)
$\beta_1'\Delta X_{t-1}$	-0.211 (-0.30)	-0.59 (-4.52)	1.06 (9.83)	-0.01 (-0.02)	1.96 (11.50)	0.03 (0.40)
mec_t	14.422 (4.99)	2.67 (5.09)	1.23 (2.82)	9.85 (4.49)	-2.83 (-4.09)	3.29 (9.78)

Table 4: $K := F+I$: Impact factors ($+I$) in the Australian mark-up model. t-values are reported in brackets.

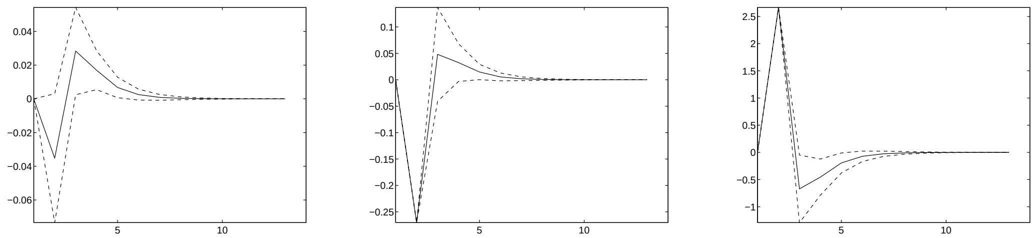


(a) IR: Effect on $\Delta^2\text{lpfc}_t$ (b) IR: Effect on $\beta' X_{t-1}$ (c) IR: Effect on mec_t

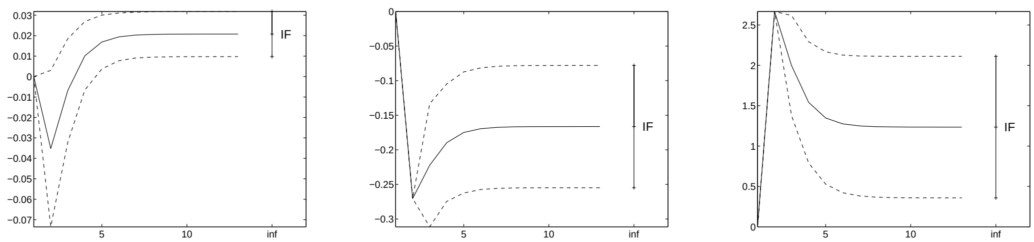


(d) AIR: Effect on $\Delta^2\text{lpfc}_t$ (e) AIR: Effect on $\beta' X_{t-1}$ (f) AIR: Effect on mec_t

Figure 3: Effect of perturbation to unit labor cost (lulc): Impulse Response functions (IR:top) and Accumulated Impulse Response functions (AIR: bottom) with 95% confidence intervals. Impact Factors with 95% confidence interval are given along AIR at horizon ∞.



(a) IR: Effect on $\Delta^2 \text{lpfc}_t$ (b) IR: Effect on $\beta' X_{t-1}$ (c) IR: Effect on mec_t



(d) AIR: Effect on $\Delta^2 \text{lpfc}_t$ (e) AIR: Effect on $\beta' X_{t-1}$ (f) AIR: Effect on mec_t

Figure 4: Effect of perturbation of import prices (lpm): Impulse Response functions (IR:top) and Accumulated Impulse Response functions (AIR: bottom) with 95% confidence intervals. Impact Factors with 95% confidence interval are given along AIR at horizon ∞.

graphs as the limit of the AIR at horizon ∞, indicate as 'inf'. Note that some of the IR have sometimes 0 standard errors for the first lead, because of the different timing of the variables in the state vector.

Figure 3(a) shows the effect on the the second difference of the price level, that is the acceleration rate of inflation. The initial impact is positive and followed by a small decline. The accumulated impulse response function shows the effect on the inflation rate. This effect converges rapidly to 0.08, the impact factor; this corresponds to a permanent increase in the inflation rate of 0.32% (due to a one percent perturbation of unit labor costs).

Graphs (b) and (e) show that an increase in unit labor costs leads a decline in the mark-up lpfc$-$lulc. Note that the combination of an increase in inflation and a decrease in this mark-up is completely in the line with the prediction of imperfect competition models. Graphs (c) and (f) show that influence on the multicointegration relation.

Figure 4 reports the effect of perturbation to import prices. The adjustment to the new equilibrium of the multicointegrating relation takes longer than for unit labor costs. Apart from the effect on relation $\beta_1' \Delta X_{t-1}$, the impact factors have the same sign as the impact factors above, but are 2 to 4 times smaller in magnitude. Labor costs thus have a greater impact on the forecast of price inflation than import prices, a reasonable finding.

7 Conclusions

In this paper we have defined impact factors as a sensitivity measure on forecasts, and discussed their relation to impulse responses. We have applied the definition to vector autoregressive processes, in the stationary, I(1) and I(2) cases. Not surprisingly, the impact factors are functions of the moving average total impact matrix of the stationary representation of the systems, which is singular in cointegrated processes. Inference on the impact factors can be addressed exploiting the results available for the MA impact matrix developed in Paruolo (1997a,b, 2002a).

An application to price mark-up in Australia shows, among other things, how perturbations to labor cost can have a permanent positive effect on inflation and a permanent negative effect on the mark-up. This is in line with imperfect competition models.

References

Banerjee, A., L. Cockeress and B. Russell (2001) An I(2) analysis of inflation and the markup, *Journal of Applied Econometrics* 16, 221-240.

Banerjee A. N. and J. Magnus (1999) The sensitivity of OLS when the variance matrix is (partially) unknown, *Journal of Econometrics* 92, 295-323

Banerjee A. N. and J. Magnus (2000) On the sensitivity of the usual t- and F-tests to covariance misspecification, *Journal of Econometrics* 95, 157-176

Bedini C. and R. Mosconi (2000) New tools for the dynamic analysis of cointegrated VAR models. Mimeo. Department of economics and production, Politecnico di Milano, Italy.

Doornik, J.A. and H. Hansen (1994) An omnibus test for univariate and multivariate normality, Working Paper, Nuffield College, Oxford, available at `http://www.nuff.ox.ac.uk/Users/Doornik/papers/`

Doornik, J.A. (1996) Testing vector autocorrelation and heteroscedasticity, Working Paper, Nuffield College, Oxford, available at `http://www.nuff.ox.ac.uk/Users/Doornik/papers/`

Faliva M. and M.G. Zoia (2002) On partitioned inversion formula having useful applications in econometrics, *Econometric theory* 18(2), 525-530.

Gourerioux C. and A. Monfort (1995) *Statistics and econometric models*, vol. 1, Cambridge, Cambridge University Press.

Hendry (1995), *Dynamic econometrics*. Oxford, Oxford University Press.

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. (2002) Testing hypotheses in the I(2) model, mimeo, University of Copenhagen.

Kailath, T. (1980) *Linear systems.* Englewood Cliffs, Prentice-hall.

Kongsted, H.C. (1998), An I(2) cointegration analysis of small country import price determination, Working Paper 98/22, Institute of Economics, University of Copenhagen, available at `http://www.econ.ku.dk/wpa/`

Kongsted, H.C. (2002), Testing the nominal-to-real transformation, Discussion papers 02/06, University of Copenhagen Institute of Economics, available at `http://www.econ.ku.dk/wpa/`

Koop G., M. H. Pesaran, S. M. Potter (1996) Impulse response analysis in nonlinear multivariate models, *Journal of Econometrics* 74, 119-147.

Lütkepohl H. (1991) *Introduction to Multiple Time Series Analysis*, Berlin, Springer-Verlag.

Omtzigt, P. (2002) *Topics in cointegration analysis*, PhD thesis, European University Institute, Florence.

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

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

Paruolo P. (1997b) Standard errors for the long-run variance matrix. *Econometric Theory* 13, Problem 97.1.1, 305, *Econometric Theory* 14, Solution 97.1.1, 152-153.

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

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) Testing for common trends in conditional I(2) VAR systems, Quaderno della Facoltà di Economia 2002/28, University of Insubria, Italy.

Paruolo P. (2002c) Common features and common I(2) trends in VAR systems, Quaderno della Facoltà di Economia 2002/29, University of Insubria, Italy.

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

Pesaran H. H. Y. Shin (1998) Generalized impulse response analysis in linear multivariate models, *Economics Letters* 58, 17-29.

Phillips P.C.B. (1998) Impulse response and forecast error variance asymptotics in nonstationary VARs, *Journal of Econometrics* 83, 21-56.

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

Appendix A: Derivation of the IF

In this appendix we report proofs of the propositions in the paper. The first lemma gives a well known result on the inversion of a partitioned matrix, see also Faliva and Zoia (2002).

Lemma 6 *Given the $p \times s$ matrices a, b of full column rank $s < p$, and Q any square $p \times p$ matrix, then a necessary and sufficient condition for the matrix*

$$S := \begin{pmatrix} Q & a \\ b' & 0 \end{pmatrix}$$

to be invertible is that $a_\perp' Q b_\perp$ be of full rank $p - s$; in this case

$$S^{-1} := \begin{pmatrix} Q & a \\ b' & 0 \end{pmatrix}^{-1} = \begin{pmatrix} R & (I - RQ)\bar{b} \\ \bar{a}'(I - QR) & \bar{a}'(QRQ - Q)\bar{b} \end{pmatrix}$$

where $R := b_\perp (a_\perp' Q b_\perp)^{-1} a_\perp'$.

Proof. We observe that S has the same rank as $K := J_1 S J_2$ for J_i invertible square matrices. Choose J_1 and J_2 as follows and calculate the resulting product $K := J_1 S J_2$

$$\begin{aligned} J_1 &:= \begin{pmatrix} & I_s \\ (\bar{a}, \bar{a}_\perp)' & \end{pmatrix}, \qquad J_2 := \begin{pmatrix} (\bar{b}, \bar{b}_\perp) & \\ & I_s \end{pmatrix} \\ K &:= J_1 S J_2 = \begin{pmatrix} I_s & & \\ Q_{a_\perp b} & Q_{a_\perp b_\perp} & \\ Q_{ab} & Q_{ab_\perp} & I_s \end{pmatrix} \end{aligned}$$

where we have used the notation $Q_{cd} := \bar{c}' Q \bar{d}$, $c, d = a,\ b,\ a_\perp,\ b_\perp$. $J_1 S J_2$ is block triangular and it is invertible iff $Q_{a_\perp b_\perp}$, or equivalently if $a_\perp' Q b_\perp$ is invertible. If this is the case, the inverse S^{-1} can be calculated as $S^{-1} = J_2 (J_1 S J_2)^{-1} J_1 = J_2 K^{-1} J_1$. By straightforward application of the partitioned inverse formula to K, one finds

$$K^{-1} := (J_1 S J_2)^{-1} = \begin{pmatrix} I_s & & \\ -Q_{a_\perp b_\perp}^{-1} Q_{a_\perp b} & Q_{a_\perp b_\perp}^{-1} & \\ -(Q_{ab} - Q_{ab_\perp} Q_{a_\perp b_\perp}^{-1} Q_{a_\perp b}) & -Q_{ab_\perp} Q_{a_\perp b_\perp}^{-1} & I_s \end{pmatrix}.$$

Finally calculating $S^{-1} = J_2 K^{-1} J_1$ one finds the results in the statement. ■

Proof. of Prop. 1. We apply partitioned inverses to the matrix $(I-A)$ partitioned conformably to the A_{ij} blocks in (5), using Lemma 6. Let

$$K := \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} := (I-A) = \begin{pmatrix} I-A_{11} & -A_{12} \\ -A_{21} & I-A_{22} \end{pmatrix}$$

and indicate by K^{ij} blocks of K^{-1} conformable with A_{ij}. Note that $K^{11} = K_{11.2}^{-1}$, where $K_{11.2} := K_{11} - K_{12}K_{22}^{-1}K_{21} = I - (A_{11} + A_{12}(I-A_{22})^{-1}A_{21})$, where

$$A_{12}(I-A_{22})^{-1}A_{21} = diag\left(\sum_{i=2}^{k-1}\Gamma_i, 0\right),$$

so that

$$K_{11.2} = I - (A_{11} + A_{12}(I-A_{22})^{-1}A_{21}) = \begin{pmatrix} I - \Gamma^\circ + \alpha\beta' & -\alpha \\ -\beta' & 0 \end{pmatrix}.$$

where $\Gamma^\circ := -\Gamma = I - \sum_{i=1}^{k-1}\Gamma_i$. Applying Lemma 6

$$K_{11.2}^{-1} = \begin{pmatrix} C & (C\Gamma^\circ - I)\bar{\beta} \\ \bar{\alpha}'(\Gamma^\circ C - I) & \bar{\alpha}'(\Gamma^\circ C\Gamma^\circ - \Gamma^\circ)\bar{\beta} \end{pmatrix},$$

where $C := \beta_\perp\left(\alpha_\perp'\Gamma^\circ\beta_\perp\right)^{-1}\alpha_\perp'$. The remaining blocks of K^{-1} can be expressed as $K^{21} = -K_{22}^{-1}K_{21}K_{11.2}^{-1}$, $K^{12} = -K_{11.2}^{-1}K_{12}K_{22}^{-1}$ and $K^{22} = K_{22}^{-1} + K_{22}^{-1}K_{21}K_{11.2}^{-1}K_{12}K_{22}^{-1}$, where $K_{11.2}^{-1}$ has already been calculated and $K_{22}^{-1} = c \otimes I$. Substituting one obtains the expression in the proposition. ■

The following EC formulation is convenient in the I(2) case.

Proposition 7 *An equivalent EC formulation of* (6) *is*

$$\begin{aligned} \Delta^2 X_t &= \alpha(\beta' X_{t-3} + \delta\beta_2'\Delta X_{t-2}) + (\zeta_1^* : \zeta_2)(\beta : \beta_1)'\Delta X_{t-2} + \\ &\quad + \Upsilon_1^*\Delta^2 X_{t-1} + \Phi W_{t-1} + \mu^* D_t^* + \epsilon_t \end{aligned} \tag{18}$$

where $\zeta_1^* := \zeta_1 + 2\alpha$ *and* $\Upsilon_1^* := (\Upsilon_1 + \Gamma + \Pi)$.

Proof. Adding and subtracting $\Pi(X_{t-1} - X_{t-3}) = \Pi\Delta X_{t-1} + \Pi\Delta X_{t-2}$ on the r.h.s. of (6) one obtains

$$\Delta^2 X_t = \Pi X_{t-3} + (\Gamma + \Pi)\Delta X_{t-1} + \Pi\Delta X_{t-2} + \Upsilon_1\Delta^2 X_{t-1} + \Phi W_t + \epsilon_t.$$

Further adding and subtracting $(\Gamma + \Pi)\Delta X_{t-2}$ on the r.h.s. yields

$$\begin{aligned} \Delta^2 X_t &= \Pi X_{t-3} + (\Gamma + 2\Pi)\Delta X_{t-2} + (\Upsilon_1 + \Gamma + \Pi)\Delta^2 X_{t-1} + \Phi W_t + \epsilon_t \\ &= \Pi X_{t-3} + \Gamma^*\Delta X_{t-2} + \Upsilon_1^*\Delta^2 X_{t-1} + \Phi W_t + \epsilon_t, \end{aligned} \tag{19}$$

where $\Gamma^* := \Gamma + 2\Pi$, $\Upsilon_1^* := \Upsilon_1 + \Gamma + \Pi$. In order to recover the EC terms within (19) note that $\Gamma^*\bar{\beta}_2 = (\Gamma + 2\Pi)\bar{\beta}_2 = \Gamma\bar{\beta}_2 = \alpha\delta$ and hence

$$\Gamma^* = \Gamma^*(P_\tau + P_{\beta_2}) = (\Gamma^*\bar{\tau})\tau' + (\Gamma^*\bar{\beta}_2)\beta_2' = \zeta^*\tau' + \alpha\delta\beta_2',$$

where $\zeta^* := \Gamma^*\bar{\tau}$, $\tau := (\beta, \beta_1)$ and we observe that $\zeta_2^* := \Gamma^*\bar{\beta}_1 = \Gamma\bar{\beta}_1 =: \zeta_2$. Substituting within (19) one finds (18). ∎

Proof. of Prop. 2. Let $m := p(k-3)$ be the dimension of W_t. In order to compute $(I-A)^{-1}$, we apply partitioned inverses to the matrix $(I-A)$ partitioned conformably to A_{ij} blocks in (5), using Lemma 6. As in the I(1) case let

$$K := \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} := (I-A) = \begin{pmatrix} I-A_{11} & -A_{12} \\ -A_{21} & I-A_{22} \end{pmatrix}$$

and indicate by K^{ij} blocks of K^{-1} conformable with A_{ij}. Note that $K^{11} = K_{11.2}^{-1}$, where $K_{11.2} := K_{11} - K_{12}K_{22}^{-1}K_{21} = I - (A_{11} + A_{12}(I-A_{22})^{-1}A_{21})$, where

$$A_{12}(I-A_{22})^{-1}A_{21} = diag\left(\sum_{i=2}^{k-1}\Upsilon_i, 0\right),$$

so that

$$K_{11.2} = I - (A_{11} + A_{12}(I-A_{22})^{-1}A_{21}) = \begin{pmatrix} \phi - \Gamma - \alpha\beta' & -\zeta_1^* & -\zeta_2 & -\alpha \\ -\beta' & & & \\ -\beta_1' & & & \\ -\delta\beta_2' & -I_{p_0} & & \end{pmatrix}.$$

where $\phi := I - \sum_{i=1}^{k-2}\Upsilon_i$. In order to calculate $K_{11.2}^{-1}$ we express it as $K_{11.2}^{-1} = (J_3K_{11.2})^{-1}J_3$ where

$$\begin{aligned} J_3 &:= \begin{pmatrix} I_p & & & \\ & & & I_{p_0} \\ & I_{p_0} & & \\ & & I_{p_1} & \end{pmatrix} \\ J_3K_{11.2} &= \left(\begin{array}{cc|cc} \phi^* & -\zeta_1^* & -\zeta_2 & -\alpha \\ -\delta\beta_2' & -I_{p_0} & & \\ \hline -\beta' & & & \\ -\beta_1' & & & \end{array}\right) = \begin{pmatrix} Q & a \\ b' & 0 \end{pmatrix}, \end{aligned}$$

where $\phi^* := \phi - \Gamma - \alpha\beta'$ and Q is $(p+p_0)\times(p+p_0)$. We now wish to apply Lemma 6, observing that $b_\perp = diag(\beta_2, I_{p_0})$ and $a_\perp = diag(\alpha_2, I_{p_0})$, because $\zeta_2 = \alpha\Gamma_{\alpha\beta_1} + \alpha_1 \in sp(\alpha : \alpha_1)$. Let $\theta^* := \phi^* + \zeta_1^*\bar{\alpha}'\Gamma$, $h := I - \theta^*C_2$ and recall that $\phi^* = \phi - \Gamma - \alpha\beta'$, $\zeta_1^* = \zeta_1 + 2\alpha$, $\bar{\tau} = (\bar{\beta} : \bar{\beta}_1)$. One finds

$$\begin{aligned} (J_3K_{11.2})^{-1} &= \begin{pmatrix} R & (I-RQ)\bar{b} \\ \bar{a}'(I-QR) & \bar{a}'(QRQ-Q)\bar{b} \end{pmatrix} = \\ &= \begin{pmatrix} C_2 & -C_2\zeta_1 & (C_2\phi^* - I)\bar{\tau} \\ -\delta\beta_2'C_2 & \delta\beta_2'C_2\zeta_1 - I & -\delta\beta_2'C_2\phi^*\bar{\tau} \\ -\alpha_1'h & \alpha_1'h\zeta_1^* & -\alpha_1'h\phi^*\bar{\tau} \\ -\bar{\alpha}'(I-\zeta_2\bar{\alpha}_1')h & \bar{\alpha}'(I-\zeta_2\bar{\alpha}_1')h\zeta_1^* & -\bar{\alpha}'(I-\zeta_2\bar{\alpha}_1')h\phi^*\bar{\tau} \end{pmatrix} \end{aligned}$$

where

$$\bar{a}' = \begin{pmatrix} -\alpha_1' & 0 \\ -\bar{\alpha}'(I-\zeta_2\bar{\alpha}_1') & 0 \end{pmatrix} \qquad \bar{b} = \begin{pmatrix} -\bar{\tau} \\ 0 \end{pmatrix}.$$

Thus $B := K_{11.2}^{-1} = (J_3 K_{11.2})^{-1} J_3$ corresponds to the expression found above for $(J_3 K_{11.2})^{-1}$ with the last 2 blocks of columns interchanged. The rest of the calculations are exactly the same as in the proof of Proposition 1; this completes the proof. ■

Appendix B: Timing and IF

In this appendix we illustrate various possible choices of lag for the EC terms, and report proofs of Section 3.3. In all cases below we adopt the following convention: the various subvectors of the state vector $Z_t := NS_t$ or S_t are numbered consecutively. Consider the i-th subvector of Z_t and the j-th subvector of S_t, of dimension n_i and n_j respectively; the elements of the transformation matrix N corresponding to these subvectors are indicated with the subscript ij, N_{ij}, of dimension $n_i \times n_j$. When not otherwise specified, elements of the N matrix are assumed to be zero.

1. I(1) case, EC in lag 1. Let S_t be the choice of state vector used above, $S_t := (S_{1t}' : S_{2t}' : S_{3t}')' = (\Delta X_t' : X_{t-1}'\beta : U_t')'$, and consider the following possible alternative choice of state vector $Z_t := (Z_{1t}' : Z_{2t}' : Z_{3t}')' = (\Delta X_t' : X_t'\beta : U_t')'$. It is simple to see that $Z_t = NS_t$ with $N_{ii} = I$, $i = 1, 2, 3$ and $N_{21} = \beta'$.

2. I(1) case, EC in lag j, where $1 < j \leq k$. Let S_t be the choice of state vector used above, and let $Z_{ht} := S_{ht}$, $h = 1, 3$ and $Z_{2t} := \beta' X_{t-j}$. It is simple to see that $Z_t = NS_t$ with $N_{ii} = I$, $i = 1, 2, 3$ and $N_{23} = (-i_{j-1}' \otimes \beta', 0)$, where i_j is an $j \times 1$ vector of ones.

3. I(2) case, level term in lag 1. Let S_t be the choice of state vector used above, $S_t := (S_{1t}' : ... : S_{5t}')' = (\Delta^2 X_t' : \Delta X_{t-1}'\beta : \Delta X_{t-1}'\beta_1 : X_{t-2}'\beta + \Delta X_{t-1}'\beta_2\delta' : W_t')'$ and consider the following possible alternative choice of state vector $Z_t := (Z_{1t}' : ... : Z_{5t}')'$ where $Z_{ht} := S_{ht}$, $h = 1, 2, 3, 5$ and $Z_{4t} := \beta' X_{t-1} + \delta\beta_2'\Delta X_{t-1}$. The only term that has been shifted is $X_s'\beta$ form $s = t-2$ to $s = t-1$. It is simple to see that $Z_t = NS_t$ with $N_{ii} = I$, $i = 1, ..., 5$ and $N_{42} = I_{p_0}$.

4. I(2) case, level term in lag s, where $2 < s \leq k$. Let S_t be the choice of state vector used above and consider the following possible alternative choice of state vector $Z_t := (Z_{1t}' : ... : Z_{5t}')'$ where $Z_{ht} := S_{ht}$, $h = 1, 2, 3, 5$ and $Z_{4t} := \beta' X_{t-s} + \delta\beta_2'\Delta X_{t-1}$ where the only term that has been shifted is $\beta' X_i$ form $i = t-2$ to $i = t-s$. It can be checked that $Z_t = NS_t$ with $N_{ii} = I$, $i = 1, ..., 5$, $N_{42} = -(s-2)I_{p_0}$, $N_{45} = (j' \otimes \beta')$, $j := (s-3,\ s-4,\ ...,\ 1,\ 0,\ ...,\ 0)'$.

5. I(2) case, differenced term in lag s, where $2 \leq s \leq k$. Let S_t be the choice of state vector used above and consider the following possible alternative choice of state vector $Z_t := (Z_{1t}' : ... : Z_{5t}')'$ where $Z_{ht} := S_{ht}$, $h = 1, 2, 3, 5$ and $Z_{4t} := \beta' X_{t-2} + \delta\beta_2'\Delta X_{t-s}$ where the only term that has been shifted is $\beta_2'\Delta X_i$ form $i = t-1$ to $i = t-s$. It can be checked that $Z_t = NS_t$ with $N_{ii} = I$, $i = 1, ..., 5$, $N_{45} = (-i_{s-2}' \otimes \delta\beta_2', 0)$.

Proof. of Prop. 3. By definition

$$\begin{aligned} F_Z := \ (I-A_Z)^{-1}-I = &\ (N(I-A_S)N^{-1})^{-1}-I = N(I-A_S)^{-1}N^{-1}-I = \\ = &\ N((I-A_S)^{-1}-I)N^{-1} =: NF_SN^{-1}. \end{aligned}$$

■

Proof. of Proposition 4. Let $Z_t = NS_t$ indicate the change of state vector, and let F_Z and $F := F_S$ indicate the corresponding IF. From Prop. 3 it follows that $F_Z = NFN^{-1}$. Hence $(F_Z)_{y,x} = \sum_i \sum_j N_{yi}F_{ij}N^{jx}$, where we use subscripts to indicate blocks. Blocks of N^{-1} are indicated with $N^{ij} := (N^{-1})_{ij}$. Thus if $N_{yi} = 0$, $N_{yy} = I$, $N_{jx} = 0$, $N_{xx} = I$, for $i \neq y$, $j \neq x$ one finds that $(F_Z)_{y,x} = F_{y,x}$, and that the IF are invariant.

For the first result we take $S_t := \left(\Delta X_t' : X_{t-1}'\beta : U_t'\right)'$ and $Z_t := \left(\Delta X_t : X_{t-j}'\beta : U_t'\right)'$, $j = 1, \ldots, k-1$ and note that $Z_t = NS_t$ with

$$N = \begin{pmatrix} I_p & & \\ & I_{p_0} & (-i_{j-1}' \otimes \beta' : 0) \\ & & I_m \end{pmatrix}, \qquad N^{-1} = \begin{pmatrix} I_p & & \\ & I_{p_0} & (i_{j-1}' \otimes \beta' : 0) \\ & & I_m \end{pmatrix}$$

It is thus immediate to note that F_{y_t,x_t} is invariant for $y_t = \Delta X_t$, U_t and $x_t = \Delta X_t$, $\beta' X_{t-j}$. When Z_t includes $\beta' X_t$, case $j = 0$ above, then the transformation matrix has a similar shape, but $N_{21} = \beta'$, $N_{23} = 0$, $N^{21} = -\beta'$, $N^{23} = 0$. The same conclusion thus applies.

For the I(2) results, we take $Z_t = NS_t$ with

$$S_t := S_t(i,j,l,m) := \left(\Delta^2 X_t' : \Delta X_{t-i}'\beta : \Delta X_{t-j}'\beta_1 : X_{t-l}'\beta + \Delta X_{t-m}'\beta_2\delta' : W_t'\right)'$$

and $Z_t := S_t(1,1,2,1)$. One finds

$$N = \begin{pmatrix} I & & & & \\ & I & & & N_{25} \\ & & I & & N_{35} \\ & N_{42} & & I & N_{45} \\ & & & & I \end{pmatrix}, \qquad N^{-1} = \begin{pmatrix} I & & & & \\ & I & & & -N_{25} \\ & & I & & -N_{35} \\ & -N_{42} & & I & Q \\ & & & & I \end{pmatrix},$$

where $Q := -N_{45} + N_{42}N_{25}$, where $N_{45} := N_{45a} + N_{45b}$,

$$N_{25} = \left(-i_{i-1}' \otimes \beta' : 0\right), N_{35} = \left(-i_{j-1}' \otimes \beta_1' : 0\right), N_{45a} = \left(-i_{m-1}' \otimes \delta\beta_2' : 0\right).$$

If $l = 1$, one has $N_{42} = I_{p_0}, N_{45b} = 0$ whereas if $l \geq 3$, $N_{42} = -(l-2)\, I_{p_0}$, $N_{45b} = (g \otimes \beta' : 0)$, $g := (l-3 : l-4 : \ldots : 1 : 0 : \ldots : 0)$.

From the expressions on N and N^{-1} we find that F_{y_t,x_t} is invariant for $y_t = \Delta^2 X_t$, W_t and $x_t = \Delta^2 X_t$, $\beta_1' X_{t-j}$, $\beta' X_{t-l} + \delta\beta_2' \Delta X_{t-m}$. We also observe that $F_{\Delta^2 X_t, \beta_1' X_{t-j}}$ can be simplified as follows

$$(C_2\phi^* - I)\bar{\beta}_1 = (C_2\phi - I)\bar{\beta}_1 - C_2\Gamma\bar{\beta}_1 - C_2\alpha\beta'\bar{\beta}_1 = (C_2\phi - I)\bar{\beta}_1,$$

where we have used that $C_2\Gamma\bar{\beta}_1$ contains the term $\alpha_2'\Gamma\beta_1$, which equals zero by assumption I(2)_b. ■

Appendix C: Inference on the IF

In this appendix we illustrate how the state space representations used for $k = 1$ in the I(1) and $k = 2$ for the I(2) cases are not minimal; proofs of Section 5 are also provided.

The non-minimality of the state space vectors does not affect the derivations of IF, although it is relevant for inference. We thus show how the companion matrices A can be rank-decomposed in $A = \widetilde{A}H'$. In case of no rank reduction of the matrix A, we take $H = I$ in the decomposition $A = \widetilde{A}H$, i.e. $A = \widetilde{A}$.

In the I(1) state space formulation, when $k = 1$ the companion matrix A reduces to the block A_{11} in formula (5), where, moreover, $\Gamma_1 = 0$, i.e. $\Gamma_1^* = \Pi = \alpha\beta'$. It is simple to see that A has in this case reduced rank, since

$$A := \begin{pmatrix} \alpha\beta' & \alpha \\ \beta' & I \end{pmatrix} = \begin{pmatrix} \alpha \\ I \end{pmatrix} \begin{pmatrix} \beta' & I \end{pmatrix} =: \widetilde{A}H', \tag{20}$$

where $\widetilde{A} := (\alpha' : I)'$, $H := (\beta' : I)'$ are $p + p_0 \times p_0$ matrices with full column rank p_0.

For the state space representation of I(2) systems, when $k = 2$ the companion matrix A reduces to the block A_{11} in formula (10) above, where, moreover, $\Upsilon_1 = 0$, i.e. $\Upsilon_1^* = \Gamma + \Pi = \Gamma + \alpha\beta'$. It can be checked that, similarly to the I(1) case, A has in this case reduced rank:

$$\begin{aligned} A &= \begin{pmatrix} \Gamma + \alpha\beta' & \Gamma\bar{\beta} + 2\alpha & \Gamma\bar{\beta}_1 & \alpha \\ \beta' & I_{p_0} & & \\ \beta_1' & & I_{p_1} & \\ \delta\beta_2' & I_{p_0} & & I_{p_0} \end{pmatrix} = \qquad (21) \\ &= \begin{pmatrix} \Gamma\bar{\beta} + \alpha & \Gamma\bar{\beta}_1 & \alpha \\ I_{p_0} & & \\ & I_{p_1} & \\ & & I_{p_0} \end{pmatrix} \begin{pmatrix} \beta' & I_{p_0} & & \\ \beta_1' & & I_{p_1} & \\ \delta\beta_2' & I_{p_0} & & I_{p_0} \end{pmatrix} =: \widetilde{A}H', \end{aligned}$$

where

$$\widetilde{A} := \begin{pmatrix} \Gamma\bar{\beta} + \alpha & \Gamma\bar{\beta}_1 & \alpha \\ I_{p_0} & & \\ & I_{p_1} & \\ & & I_{p_0} \end{pmatrix} = \begin{pmatrix} \zeta_1 + \alpha & \zeta_2 & \alpha \\ I_{p_0} & & \\ & I_{p_1} & \\ & & I_{p_0} \end{pmatrix}, \quad H := \begin{pmatrix} \beta & \beta_1 & \beta_2\delta' \\ I_{p_0} & & I_{p_0} \\ & I_{p_1} & \\ & & I_{p_0} \end{pmatrix}$$

are $(p + 2p_0 + p_1) \times (2p_0 + p_1)$ matrices with full column rank $(2p_0 + p_1)$.

Proof. of Theorem 5. $\widehat{H} - H$, $\widehat{L} - L \in O_p(T^{-1})$ because they are functions of the cointegrating coefficients, which are at least T-consistent. Result (15) follows by standard regression arguments, after observing that, due to superconsistency, one can substitute the estimated cointegration coefficients with their true values, see Paruolo (2002c) for a detailed proof of $\widehat{S}_{ij} - S_{ij} = O_p(T^{-1})$. In fact one has $\widehat{G} = \widehat{S}_{01}\widehat{S}_{11}^{-1} = S_{01}S_{11}^{-1} + O_p(T^{-1}) = G + S_{\epsilon 1}S_{11}^{-1} + O_p(T^{-1})$, from which $T^{1/2}(\widehat{G} - G) = T^{1/2}S_{\epsilon 1}S_{11}^{-1} + o_p(1)$, and $T^{1/2}vec(\widehat{G} - G)' \xrightarrow{w} N(0, \Omega \otimes \Sigma^{-1})$.

In order to show (16) note that differentiating F one has $\mathrm{d}F = K\mathrm{d}AK$, so that

$$T^{1/2}(\widehat{F} - F) = T^{1/2}K(\widehat{A} - A)K + o_p(1). \tag{22}$$

Because $\widehat{H} - H$, $\widehat{L} - L \in O_p(T^{-1})$ one has

$$T^{1/2}(\widehat{A} - A) = \begin{pmatrix} T^{1/2}(\widehat{G} - G)H' \\ 0 \end{pmatrix} + o_p(1) = JT^{1/2}(\widehat{G} - G)H' + o_p(1)$$

Substituting in (22) one finds $T^{1/2}(\widehat{F} - F) = T^{1/2}KJ(\widehat{G} - G)H'K + o_p(1)$. Transposing and vectorizing one obtains (16) from (15). ■