Journal of Econometrics 178 (2014) 310–315

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

An asymptotic invariance property of the common trends under linear transformations of the data[☆]

Søren Johansen[a,b,*], Katarina Juselius[a]

[a] Department of Economics, University of Copenhagen, Øster Farimagsgade 5, DK-1353 Copenhagen K, Denmark
[b] CREATES, Denmark

ARTICLE INFO

Article history:
Available online 5 September 2013

JEL classification:
C32

Keywords:
Cointegration vectors
Common trends
Prediction errors

ABSTRACT

It is well known that if X_t is a nonstationary process and Y_t is a linear function of X_t, then cointegration of Y_t implies cointegration of X_t. We want to find an analogous result for common trends if X_t is generated by a finite order VAR with i.i.d. $(0, \Omega_x)$ errors ε_{xt}. We first show that Y_t has an infinite order VAR representation in terms of its white noise prediction errors, ε_{yt}, which are a linear process in ε_{xt}, the prediction error for X_t. We then apply this result to show that the limit of the common trends for Y_t generated by ε_{yt}, are linear functions of the common trends for X_t, generated by ε_{xt}.

We illustrate the findings with a small analysis of the term structure of interest rates.

1. Introduction and motivation

It is well known that if X_t is a p-dimensional $I(1)$ process and if the m-dimensional linear transformation $Y_t = a'X_t$, $m < p$, is cointegrated, that is, $\beta_y' Y_t$ is stationary for some $\beta_y \neq 0$, then X_t is cointegrated with cointegration vector $a\beta_y$, because $\beta_y' a' X_t = \beta_y' Y_t$ is stationary. Thus cointegration in the small system, Y_t, implies cointegration in the large system, X_t, but not necessarily the other way.

We want to investigate if a similar result holds for common trends. We discuss this in the context of an $I(1)$ cointegrated vector autoregressive process X_t, generated by

$$\Delta X_t = \alpha_x \beta_x' X_{t-1} + \sum_{i=1}^{k} \Gamma_{xi} \Delta X_{t-i} + \varepsilon_{xt}, \tag{1}$$

where ε_{xt} is i.i.d. $(0, \Omega_x)$ and α_x and β_x are $p \times r$. For $I(1)$ processes the solution of (1) is given by the Granger representation

$$X_t = C_x \sum_{i=1}^{t} \varepsilon_{xi} + \sum_{n=0}^{\infty} C_{xn}^* \varepsilon_{xt-n} + A_x, \tag{2}$$

see Johansen (1996, Theorem 4.2). The common trends are $\alpha_{x\perp}' \sum_{i=1}^{t} \varepsilon_{xi}$, where $(\alpha_{x\perp}, \alpha_x)$ is a full rank matrix and $\alpha_x' \alpha_{x\perp} = 0$. This formulation covers the case $r = p$, where $C_x = 0$ so that X_t is stationary, and $r = 0$, where C_x has full rank, so that X_t is not cointegrating.

The m-dimensional $Y_t = a'X_t$ has a vector moving average (VMA) representation in terms of the p-dimensional i.i.d. sequence ε_{xt},

$$Y_t = a' C_x \sum_{i=1}^{t} \varepsilon_{xi} + \sum_{n=0}^{\infty} a' C_{xn}^* \varepsilon_{xt-n} + a' A_x. \tag{3}$$

We define the $\mathcal{L}_2$ space $\mathcal{L}_{2t}^x = \mathcal{L}_2(\beta_x' X_t, \Delta X_s, s \leq t)$ spanned by $\{\beta_x' X_t, \Delta X_s, s \leq t\}$ and the linear projection $\mathcal{P}_{t-1}^x$ onto $\mathcal{L}_{2,t-1}^x$. Then $\mathcal{P}_{t-1}^x(\Delta X_t) = \alpha_x \beta_x' X_{t-1} + \sum_{i=1}^{k} \Gamma_{xi} \Delta X_{t-i}$ and the prediction error of ΔX_t with respect to $\mathcal{L}_{2,t-1}^x$ is $\varepsilon_{xt} = \Delta X_t - \mathcal{P}_{t-1}^x(\Delta X_t)$. Similarly we define the prediction error of ΔY_t with respect to $\mathcal{L}_{2,t-1}^y$ as $\varepsilon_{yt} = \Delta Y_t - \mathcal{P}_{t-1}^y(\Delta Y_t)$.

If Y_t were a finite order CVAR, with m-dimensional prediction errors $\varepsilon_{yt} = \Delta Y_t - \mathcal{P}_{t-1}^y(\Delta Y_t)$, we would find the corresponding Granger representation

$$Y_t = C_y \sum_{i=1}^{t} \varepsilon_{yi} + \sum_{n=0}^{\infty} C_{yn}^* \varepsilon_{yt-n} + A_y, \tag{4}$$

and it is tempting to conclude that the nonstationary part of the two representations (3) and (4) should be the same

$$a' C_x \sum_{i=1}^{t} \varepsilon_{xi} = C_y \sum_{i=1}^{t} \varepsilon_{yi}, \tag{5}$$

☆ We would like to thank the referee for a careful reading and for many useful comments. The first author wants to thank Søren Tolver Jensen for very helpful discussions on prediction theory and CREATES — Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

* Corresponding author at: Department of Economics, University of Copenhagen, Øster Farimagsgade 5, DK-1353 Copenhagen K, Denmark.

E-mail addresses: Soren.Johansen@econ.ku.dk (S. Johansen), Katarina.Juselius@econ.ku.dk (K. Juselius).

0304-4076/$ – see front matter
http://dx.doi.org/10.1016/j.jeconom.2013.08.029

and therefore the common trends, $\alpha_{y\perp}'\sum_{i=1}^{t}\varepsilon_{yi}$, of Y_t based on $\mathcal{L}_{2,t-1}^{y}$ are a linear function of the common trends, $\alpha_{x\perp}'\sum_{i=1}^{t}\varepsilon_{xi}$, of X_t based on $\mathcal{L}_{2,t-1}^{x}$.

It turns out that this simple argument is almost correct, but that (5) only holds in the sense that the difference is stationary, that is, the two random walk components of Y_t are cointegrated, see Fig. 2. Their difference does not go to zero with t, only by normalizing by $T^{-1/2}$ and passing to the limit do we find a relation between the limiting Brownian motions. Note that this does not give any relation between the cyclic components of (3) and (4).

In general Y_t, however, is not a finite order VAR process and therefore the residuals estimated by fitting a finite order VAR do not estimate prediction errors of ΔY_t with respect to $\mathcal{L}_{2,t-1}^{y}$, moreover one cannot conclude (5) from (3) and (4).

In order to sort this out, we therefore first apply the prediction theory of stationary processes to find an infinite order VAR representation of Y_t in terms of its white noise prediction errors ε_{yt}, and a corresponding VMA representation, or Granger representation of Y_t, in terms of ε_{yt}. Note, however, that ε_{yt} is a white noise and need not be i.i.d. as we have assumed about ε_{xt}. The exception is of course if ε_{xt} is i.i.d. Gaussian, then also ε_{yt} are i.i.d. Gaussian.

By applying the Functional Limit Theorem to the two cointegrating common trends we can then deduce from (3) and (4) that there is a linear mapping from the limiting common trends, $\alpha_{x\perp}'W_x(u)$, of the large system onto those of the small system: $\alpha_{y\perp}'W_y(u)$.

We illustrate the ideas and findings in an empirical analysis of monthly US interest rates 1987:1 to 2006:1.

2. The process X_t

Let X_t be given by (1) and define the generating matrix polynomial of a complex argument z:

$$\Pi_x(z) = (1-z)I_p - \alpha_x\beta_x'z - \sum_{i=1}^{k}\Gamma_{xi}(1-z)z^i,$$

and assume that α_x and β_x are $p\times r_x$ of rank $r_x \le p$.

Under the conditions that the roots of $\det \Pi_x(z) = 0$ satisfy either $|z| > 1$ or $z = 1$, we define $1+\delta > 1$ as the absolute value of the smallest root different from 1:

$$1+\delta = \min\{|z| : \det(\Pi_x(z)) = 0, z \neq 1\}, \tag{6}$$

and $\Gamma_x = I_p - \sum_{i=1}^{k}\Gamma_{xi}$ and assume $\det(\alpha_{x\perp}'\Gamma_x\beta_{x\perp}) \neq 0$ so that X_t is $I(1)$ and

$$C_x = \beta_{x\perp}(\alpha_{x\perp}'\Gamma_x\beta_{x\perp})^{-1}\alpha_{x\perp}'$$

is well defined. Under these assumptions the polynomial $\Pi_x(z)$ can be inverted in the sense that

$$(1-z)\Pi_x^{-1}(z) = (1-z)\frac{\operatorname{adj}(\Pi_x(z))}{\det(\Pi_x(z))} = C_x + (1-z)C_x^*(z), \tag{7}$$

and $C_x^*(z)$ are rational functions on $\{z : |z| < 1+\delta\}$ satisfying $\beta'C_x^*(1)\alpha = -I_{r_x}$, see Johansen (2009, Theorem 3). These results can be translated into the Granger representation (2).

3. The process $Y_t = a'X_t$

In general $Y_t = a'X_t$, where a is $p\times m$, of rank $m < p$, is not a finite order autoregressive process, but if β_y are the cointegrating relations for Y_t, the processes

$$U_{1t} = \beta_y'Y_t = \beta_y'a'\sum_{n=0}^{\infty}C_{xn}^*\varepsilon_{xt-n}, \tag{8}$$

$$U_{2t} = \beta_{y\perp}'\Delta Y_t = \beta_{y\perp}'a'\left(C_x\varepsilon_{xt} + \sum_{n=0}^{\infty}C_{xn}^*\Delta\varepsilon_{xt-n}\right), \tag{9}$$

of dimensions r_y and $m - r_y$ respectively are stationary linear processes in the p-dimensional prediction errors ε_{xt}. We define the $m\times p$ matrix function

$$\Phi(z) = \begin{pmatrix}\beta_y'a'C_x^*(z)\\ \beta_{y\perp}'a'(C_x + (1-z)C_x^*(z))\end{pmatrix}, \tag{10}$$

and note that $\Phi(L)\varepsilon_{xt} = U_t = (U_{1t}', U_{2t}')'$, so that the spectral density of U_t is

$$\phi_u(\lambda) = \frac{1}{2\pi}\Phi(e^{i\lambda})\Omega_x\Phi'(e^{-i\lambda}).$$

We first show that U_t is an invertible linear process in its prediction errors $\varepsilon_{ut} = U_t - \mathcal{P}_{t-1}^{u}(U_t)$, where $\mathcal{P}_{t-1}^{u}(U_t)$ is the linear projection onto $\mathcal{L}_2^u = \mathcal{L}_2(U_s, s < t)$.

Lemma 1. *The rational function $\Phi(z)$ is of rank m for $|z| < 1+\delta$, see (6). It follows that there exists an $m\times m$ function $A(z) = \sum_{n=0}^{\infty}A_nz^n$ of full rank for $|z| < 1+\delta$ with real exponentially decreasing coefficients, $A_0 = I_m$, and an $m\times m$ positive definite symmetric matrix Ω_u, so that the spectral density for U_t has the representation*

$$\phi_u(\lambda) = \frac{1}{2\pi}A(e^{i\lambda})\Omega_uA'(e^{-i\lambda}). \tag{11}$$

Moreover, we find the prediction error decomposition (VMA)

$$U_t = \sum_{n=0}^{\infty}A_n\varepsilon_{ut-n}, \tag{12}$$

in terms of the white noise $\varepsilon_{ut} = U_t - \mathcal{P}_{t-1}^{u}(U_t)$, which inverted gives the infinite order VAR representation of U_t:

$$\varepsilon_{ut} = \sum_{n=0}^{\infty}B_nU_{t-n}. \tag{13}$$

Here the prediction error ε_{ut} is a white noise process with $Var(\varepsilon_{ut}) = \Omega_u$, $A_0 = I_m$, and $A = \sum_{n=0}^{\infty}A_n$ has full rank. The function $B(z) = \sum_{n=0}^{\infty}B_nz^n = A(z)^{-1}$ is defined for $|z| < 1+\delta$ with exponentially decreasing coefficients, $B_0 = I_m$, and $B = \sum_{n=0}^{\infty}B_n = A^{-1}$.

Proof. To prove that $rank(\Phi(z)) = m$ for $|z| < 1+\delta$, we assume we have z_0 with $|z_0| < 1+\delta$, and $rank(\Phi(z_0)) < m$. Then we can find $v = (v_1', v_2')' \in \mathbb{R}_m$ so that

$$v_1'\beta_y'a'C_x^*(z_0) + v_2'\beta_{y\perp}'a'(C_x + (1-z_0)C_x^*(z_0)) = 0. \tag{14}$$

We show that $v = 0$.

Case 1: If $z_0 = 1$, we multiply (14) by α_x from the right and find $v_1'\beta_y'a'C_x^*(1)\alpha_x = 0$ because $C_x\alpha_x = 0$. Because $a\beta_y$ are cointegrating relations for X_t we have $a\beta_y = \beta_x\kappa_1$ for some matrix κ_1 of rank r_y, and $0 = v_1'\beta_y'a'C_x^*(1)\alpha_x = v_1'\kappa_1'\beta_x'C_x^*(1)\alpha_x = -v_1'\kappa_1'$ because $\beta_x'C_x^*(1)\alpha_x = -I_{r_x}$, so that $v_1 = 0$ and therefore from (14), $v_2'\beta_{y\perp}'a'C_x = 0$. But then $a\beta_{y\perp}v_2$ is a cointegrating vector for X_t and $\beta_{y\perp}v_2$ a cointegrating vector for Y_t, which implies that $v_2 = 0$, and hence $v = 0$.

Case 2: If $z_0 \neq 1$, then $(1-z_0) \neq 0$, and because $\beta_y'a'C_x = 0$ we find

$$0 = v'\Phi(z_0) = [(1-z_0)^{-1}v_1'\beta_y'a' + v_2'\beta_{y\perp}'a'] \times [C_x + (1-z_0)C_x^*(z_0)].$$

Now $C_x + (1-z_0)C_x^*(z_0) = (1-z_0)\Pi_x(z_0)^{-1}$ has full rank because $\det(\Pi_x(z_0)) \neq 0$, and therefore

$$(1-z_0)^{-1}v_1'\beta_y'a' + v_2'\beta_{y\perp}'a' = 0.$$

But $\beta_y'a'$ and $\beta_{y\perp}'a'$ are linearly independent which implies that $v_1 = 0$ and $v_2 = 0$.

This proves that $v = 0$, and $rank(\Phi(z)) = m$ for $|z| < 1+\delta$.