Testing Weak Exogeneity in Cointegrated System

February 6, 2004

Pu Chen, Hsiao Chiying
Faculty of Economics
Bielefeld University
33501 Bielefeld, Germany
January 2004

Abstract

This paper develops a limiting theory for Wald tests of weak exogeneity in error correction models (ECMs). It is well known that Wald statistics on cointegrated systems may involve nonstandard distribution and nuisance parameters, if $I(1)$ variables are not negligible in the statistics. To overcome this problem we construct a new statistic that takes only the $I(0)$ components of a Wald statistic into account and thus results in a valid χ^2 criterion. Applying this procedure to test weak exogeneity in ECMs we obtain a simple and direct χ^2 test.

Keyword : Error correction model, Exogeneity, Wald testest, VAR, Cointegration.

JEL Classification: C32, C12

1 Introduction

Vector error correction models (VECM) have now become standard tools to explore the relation among I(1) variables in econometrics. Research interest has also been paid to partial systems of VECM[1] that are conditioned on a subset of the variables. The motivation for such a partial model rather than a full system is manifold: One can decrease the dimension of the system analyzed; the results are sometimes easier to interpret; there are explicit structures in the partial system that helps to understand the data; and sometimes economists are particularly interested in the parameters of a partial model conditioned on some other variables. In these cases one would like to model a partial system.

However, valid inference based on a partial system can only be conducted when the conditioning variables are weakly exogenous[2] for the parameters of the partial system. Standard procedures[3] to test weak exogeneity of the conditioning variables have to be based on the estimated cointegration vectors. This implies that cointegration analysis of the whole system has to be done before weak exogeneity can be tested. I. Habro (1998) suggest to carry out the full system reduced rank regression first to get a valid estimate of the cointegration vectors, and then test weak exogeneity.

In this paper we present two procedures to test weak exogeneity in a cointegrated system without estimating the cointegration vectors. In section 2 we review weak exogeneity in VECM. In section 3 we develop the test procedures. In section 4 we outline some potential applications.

2 Weak Exogeneity in VECM

2.1 Condition for Weak Exogeneity of y_{2t}

We present a cointegration system (CIS) of y_t with h cointegration relations in a VECM:

$$\Delta y_t = J_1 \Delta y_{t-1} + J_2 \Delta y_{t-2} + J_{k-1} \Delta y_{t-k+1} + J_k y_{t-1} + u_t \tag{2.1}$$

where y_t is an $n \times 1$ vector of variables, J_i $(i = 1, .., k-1)$ are $n \times n$ matrices of parameters; $J_k = BA'$, B and A are $h \times n$ vectors of parameters; u_t is $n \times 1$ vector of residuals with $u_t \sim iid\ N(0, \Sigma_u)$.

[1]see I. Habro (1998)
[2]For a detailed discussion about exogeneity see Engle, Hendry, and Richard (1983)
[3]see Johansen (1992)

Following I. Habro (1998) we partition y_t into $(y'_{1t}, y'_{2t})'$, where y_{1t} and y_{2t} are $g \times 1$ and $(n-g) \times 1$ vectors respectively, and $g \geq h$. Partitioning the parameter matrices conformably we have:

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} J_{1,1} \\ J_{1,2} \end{pmatrix} \Delta y_{t-1} + ... + \begin{pmatrix} J_{k-1,1} \\ J_{k-1,2} \end{pmatrix} \Delta y_{t-k+1} + \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} A' y_{t-1} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}, \quad (2.2)$$

where $E \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \begin{pmatrix} u_{1t} & u_{2t} \end{pmatrix}' = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}$.

We transform (2.2) by premultiplying it with

$$\begin{pmatrix} I & -W_{12}W_{22}^{-1} \\ 0 & I \end{pmatrix} \quad (2.3)$$

and obtain:

$$\Delta y_{1t} = J^*_{0,1} \Delta y_{2t} + J^*_{1,1} \Delta y_{t-1} + ... + J^*_{k-1,1} \Delta y_{t-k+1} + B^*_1 A' y_{t-1} + u^*_{1t} \quad (2.4)$$

$$\Delta y_{2t} = J_{1,2} \Delta y_{t-1} + ... + J_{k-1,2} \Delta y_{t-k+1} + B_2 A' y_{t-1} + u_{2t} \quad (2.5)$$

where

$$E \begin{pmatrix} u^*_{1t} \\ u_{2t} \end{pmatrix} \begin{pmatrix} u^*_{1t} & u_{2t} \end{pmatrix}' = \begin{pmatrix} W^*_{11} & 0 \\ 0 & W_{22} \end{pmatrix},$$

$$J^*_{0,1} = W_{12} W_{22}^{-1},$$

$$J^*_{i,1} = J_{i,1} - W_{12} W_{22}^{-1} J_{i,2} \qquad \text{for } i = 1, ...k-1,$$

$$B^*_1 = B_1 - W_{12} W_{22}^{-1} B_2.$$

For weak exogeneity of y_{2t} for the parameter in the partial system (2.4) we have the following theorem:

Proposition 2.1 (Weak exogeneity of y_{2t} for the partial VECM) *The variable y_{2t} is weakly exogenous for the parameters in (2.4) if and only if $B_2 = 0$.*

Proof: See Johansen (1992) □

Comments: $B_2 = 0$ implies that the cointegrated variables $A'y_{t-1}$ do not appear in the regression equation of (2.5), i.e. we do not need to consider the marginal process (2.5) to estimate the cointegration relations in (2.4). This is essentially the meaning of weak exogeneity.

Testing weak exogeneity of y_{2t} results in testing $H_0 : B_2 = 0$ in the regression equation (2.5). A standard procedure is to estimate first the cointegration matrix $\hat{A}$ by applying reduced rank regression in (2.1), then carry out a $F - test$ to (2.5) using $\hat{A}'y_{t-1}$ as regressors.

2.2 Implication of Weak Exogeneity of y_{2t} in the VECM

Comparing (2.1) with (2.2) we have

$$J_k = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} = \begin{pmatrix} B_1 A' \\ B_2 A' \end{pmatrix} = \begin{pmatrix} B_1 A_1' & B_1 A_2' \\ B_2 A_1' & B_2 A_2' \end{pmatrix},$$

where A_1 and A_2 are $h \times h$ and $(n-h) \times h$ matrices; B_1 and B_2 are $g \times h$ and $(n-g) \times h$ matrices respectively. $B_2 = 0$ implies $J_{2\cdot} = (J_{21}, J_{22}) = 0$. And $J_{2\cdot} = 0$ implies $B_2 = 0$. Hence we can test $B_2 = 0$ by testing $J_{2\cdot} = 0$.

On the other hand if A_1' is invertible, we have $B_2 = J_{21} A_1^{-1'}$, then $J_{21} = 0$ implies $B_2 = 0$. In this case we can test $B_2 = 0$ by testing $J_{21} = 0$. In following we present two procedure to test the hypothesis $H_0 : J_{2\cdot} = 0$ and $H_0 : J_{21} = 0$ respectively.

3 Test of Weak Exogeneity

3.1 Test of $J_{21} = 0$ in case of invertable A_1

The technique used here is basically adopted from Toda and Phillips (1993), where they look at the Wald statistic for the null hypothesis on the parameter of the VAR in level. Following Toda and Phillips (1993) it is not difficult to conclude that the Wald statistic for testing $H_0 : J_{21} = 0$ is $\chi^2((n-g)h)$ distributed. Following are the technical details:

Let $\Phi := (J_1, J_2, ...J_k)$, $x_t = \begin{pmatrix} \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-k+1} \\ y_{t-1} \end{pmatrix}$,

$\Delta Y' = (\Delta y_1, \Delta y_2, ...\Delta y_T)$, $X' = (x_1, x_2, ...x_T)$, and $U' = (u_1, u_2, ...u_T)$ we can write the VECM (2.1):

$$\Delta y_t = \Phi x_t + u_t \tag{3.6}$$

The OLS of (3.6) is:

$$\hat{\Phi} = \Delta Y' X (X'X)^{-1} \tag{3.7}$$

The hypothesis $J_{21} = 0$ can be formulated as

$$H_0 : S_1' \Phi S = 0 \qquad or \qquad (S_1' \otimes S')\text{vec}(\Phi) = 0 \tag{3.8}$$

with $S_1 = \begin{pmatrix} 0_{g\times(n-g)} \\ I_{n-g} \end{pmatrix}$ is a $n\times(n-g)$ matrix, $S = (e_k \otimes S_2)$, $e_k' = (0, ..., 0, 1)$ is a $k \times 1$ vector with only the last element equal to one, S_2 is an $n \times h$ matrix $S_2 = \begin{pmatrix} I_h \\ 0 \end{pmatrix}$.[4] $vec(\Phi)$ stack rows of matrix Φ into a column vector. We have $vec(J_{21}) = (S_1' \otimes S')vec(\Phi)$. $S_1' \otimes S'$ is an $(n-g)h \times nnk$ matrix, i.e. we are testing $(n-g)h$ restrictions on the parameter matrix Φ.

Define an invertible $nk\times nk$ matrix $H = \left(\begin{bmatrix} I_{k-1} \\ 0_{1\times(k-1)} \end{bmatrix} \otimes I_n, e_k \otimes A, e_k \otimes A_\perp \right)$, where $A_\perp$ is an $n \times (n-h)$ matrix with full column rank and $A'A_\perp = 0$. Let $z_t = H'x_t$ and $Z' = H'X'$. We obtain for z_t:

$$z_t = \begin{pmatrix} \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-k+1} \\ A' y_{t-1} \\ A_\perp' y_{t-1} \end{pmatrix}.$$

Let z_{1t} denote the $I(0)$ part of z_t and z_{2t} denote the $I(1)$ part. Then $z_{1t} = (\Delta y_{t-1}', \cdots, \Delta y_{t-k+1}, (A' y_{t-1})')'$ and $z_{2t} = A_\perp' y_{t-1}$. Following Lemma 2 in Toda and Phillips (1993) we have:

[4] S_1 and S pick out the rows and columns of the parameters in Φ that are to be tested under H_0.

$$\frac{1}{T}\sum_{t=1}^{T} z_{1t}z_{1t}' \xrightarrow{P} \Sigma_1 \tag{3.9}$$

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_{1t}u_t' \xrightarrow{L} N_0 \tag{3.10}$$

$$\frac{1}{T}\sum_{t=1}^{T} z_{2t}u_t' \xrightarrow{L} \int_0^1 B_{2t}dB_{0t}' \tag{3.11}$$

$$\frac{1}{T}\sum_{t=1}^{T} z_{2t}z_{1t}' \xrightarrow{L} \int_0^1 B_{2t}dB_{1t}' + \Sigma_{21} + \Lambda_{21} \tag{3.12}$$

$$\frac{1}{T^2}\sum_{t=1}^{T} z_{2t}z_{2t}' \xrightarrow{L} \int_0^1 B_{2t}B_{2t}'dt \tag{3.13}$$

$$\frac{1}{T}\sum_{t=1}^{T} z_{2t}u_t' \left(\frac{1}{T}\sum_{t=1}^{T} z_{2t}u_t'\right)^{-1} \xrightarrow{L} \int B_{2t}dB_{0t}' \left(\int B_{2t}B_{2t}'dt\right)^{-1} \tag{3.14}$$

where Σ_1,Σ_{21}, Λ_{21} are matrices of constants; N_0 is a $(n(k-1)+h) \times n$ normally distributed random matrix; B_{it}, $i = 1, 2, 3$ are Brownian motions.[5]

We have the Wald statistic for the hypothesis H_0 in (3.8):

$$\begin{aligned} Fl &= \left[((S_1' \otimes S')vec(\hat{\Phi}))' \left[(S_1' \otimes S')(\hat{\Sigma}_u \otimes (X'X)^{-1}(S_1 \otimes S)\right]^{-1} (S_1' \otimes S)vec(\hat{\Phi})\right] \\ &= tr\left[(S_1'\hat{\Phi}S)(S'(X'X)^{-1}S)^{-1}(S'\hat{\Phi}'S_1)(S_1'\hat{\Sigma}_u S_1)^{-1}\right], \end{aligned} \tag{3.15}$$

where $\hat{\Sigma}_u$ is the consistent OLS estimator of the covariance matrix of the residuals. Using (3.7) we have under H_0:

$$S_1'\hat{\Phi}S = S_1'U'X(X'X)^{-1}S.$$

Inserting this into (3.15) we get:

$$Fl = tr\left[S_1'U'X(X'X)^{-1}S(S'(X'X)^{-1}S)^{-1}S'(X'X)^{-1}X'US_1(S_1'\hat{\Sigma}_u S_1)^{-1}\right] \tag{3.16}$$

[5]See Lemma 2 in Toda and Phillips (1993) for details. The last equation is not listed in the Lemma 2 of Toda and Phillips (1993). It can be easily conducted from the third and the fifth equations above.

Replacing $X' = H^{-1'}Z'$ in (3.16) we get:

$$Fl = tr\left[S_1'U'Z(Z'Z)^{-1}H'S(S'H(Z'Z)^{-1}H'S)^{-1}S'H(Z'Z)^{-1}Z'US_1(S_1'\hat{\Sigma}_u S_1)^{-1}\right]. \tag{3.17}$$

For any full rank $h \times h$ matrix K_T we have:

$$Fl = tr\left[S_1'U'Z(Z'Z)^{-1}H'SK_T(K_T'S'H(Z'Z)^{-1}H'SK_T)^{-1}K_T'S'H(Z'Z)^{-1}Z'US_1(S_1'\hat{\Sigma}_u S_1)^{-1}\right] \tag{3.18}$$

We choose the scaling matrix Υ_T:

$$\Upsilon_T = \begin{pmatrix} \sqrt{T}I_{n(k-1)+h} & 0 \\ 0 & TI_{n-h} \end{pmatrix}$$

Inserting the scaling matrix into (3.18) we get:

$$\begin{aligned} Fl &= tr(S_1'U'Z\Upsilon_T^{-1}(\Upsilon_T^{-1}Z'Z\Upsilon_T^{-1})^{-1}\Upsilon_T^{-1}H'SK_T(K_T'S'H\Upsilon_T^{-1}(\Upsilon_T^{-1}Z'Z\Upsilon_T^{-1})^{-1}\Upsilon_T^{-1}H'SK_T)^{-1} \\ &\quad K_T'S'H\Upsilon_T^{-1}(\Upsilon_T^{-1}Z'Z\Upsilon_T^{-1})^{-1}\Upsilon_T^{-1}Z'US_1(S_1'\hat{\Sigma}_u S_1)^{-1}) \end{aligned} \tag{3.19}$$

To see the asymptotical distribution of Fl we have the following proposition.

Proposition 3.1

$$\Upsilon_T^{-1}Z'Z\Upsilon_T^{-1} \xrightarrow{L} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \int B_2B_2' \end{pmatrix}$$

$$\Upsilon_T^{-1}Z'U \xrightarrow{L} \begin{pmatrix} N_0 \\ \int B_2dB_0' \end{pmatrix},$$

where Σ_1 is an $m \times m$ constant matrix, $m = n(k-1) + h$.

Proof: These results follow directly from Lemma 2 of Toda and Phillips (1993).

Notice that

$$\begin{aligned} S'H &= \left(e_k' \otimes \begin{pmatrix} I_h & 0 \end{pmatrix}_{h\times n}\right)\left(\begin{pmatrix} I_{k-1} \\ 0 \end{pmatrix}_{k\times(k-1)} \otimes I_n, e_k \otimes A, e_k \otimes A_\perp\right) \\ &= \left(0 \otimes \begin{pmatrix} I_h & 0 \end{pmatrix}, 1 \otimes A_h, 1 \otimes A_{\perp h}\right) \\ &= (0_{h\times(n(k-1))}, A_h, A_{\perp h}) \end{aligned}$$

where A_h and $A_{\perp h}$ are the first h rows of A and $A_\perp$ respectively.

Choosing $K_T = (\sqrt{T} I_h)$ we have

$$K_T' S' H \Upsilon_T^{-1} \rightarrow (0, A_h, 0) = (A_h^*, 0),$$

where A_h^* denotes the $h \times (n(k-1)+h)$ matrix $(0, A_h)$.

Taking limit and inserting the results above into (3.19) we get:

$$\begin{aligned}
Fl \;\xrightarrow{L}\; & tr\Big(S_1',\ \big(N_0', (\int B_2 dB_0')'\big) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \int B_2 B_2' \end{pmatrix}^{-1} \begin{pmatrix} A_h^{*'} \\ 0 \end{pmatrix} \qquad (3.20)\\
& \left[(A_h^*, 0) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \int B_2 B_2' \end{pmatrix}^{-1} \begin{pmatrix} A_h^{*'} \\ 0 \end{pmatrix} \right]^{-1} \\
& (A_h^*, 0) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \int B_2 B_2' \end{pmatrix}^{-1} \begin{pmatrix} N_0 \\ \int B_2 dB_0 \end{pmatrix} S_1 (S_1' \Sigma_u S_1)^{-1}\Big) \\
= \; & tr(S_1' N_0' \Sigma_1^{-1} A_h^{*'} (A_h^* \Sigma_1^{-1} A_h^{*'})^{-1} A_h^* \Sigma_1^{-1} N_0 S_1 (S_1' \Sigma_u S_1)^{-1}) \\
= \; & tr(vec(A_h^* \Sigma_1^{-1} N_0 S_1)' ((A_h^* \Sigma_1^{-1} A_h^{*'}) \otimes (S_1' \Sigma_u S_1))^{-1} vec(A_h^* \Sigma_1^{-1} N_0 S_1)).
\end{aligned}$$

We have

$$vec(A_h^* \Sigma_1^{-1} N_0 S_1) = A_h^* \Sigma_1^{-1} \otimes S_1' vec(N_0) \sim N(0, A_h^* \Sigma_1^{-1} A_h^{*'} \otimes S_1' \Sigma_u S_1)$$

Therefore, for the asymptotic distribution of the Wald statistic in (3.20) we have the following theorem.

Theorem 3.2 *If $Rank(A_1) = h$ then the Wald statistic in (3.15) has asymptotically a $\chi^2((n-g)h)$ distribution.*

Proof: See the discussion above. □

Comments: $Rank(A_1) = h$ means that the first h elements in y_t should be sufficiently cointegrated such that the $h \times h$ matrix A_1 has full rank. Then the Wald test statistic will have a $\chi^2((n-g)h)$ distribution. A similar result is obtained in Toda and Phillips (1993) for testing of Granger causality in levels vector autoregressions (VAR's) with cointegrated relations. If the first h elements of y_t are insufficiently cointegrated such that A_1 is not invertable, then the inverse matrix in the second line of (3.20) does not exist. consequently we may not be able to apply this theorem to test weak exogeneity of y_{2t}. We turn to these cases in the next section.

3.2 Testing $J_{2\cdot} = 0$

The basic problem in testing $J_{2\cdot} = 0$ is that the corresponding Wald statistic has a nonstandard distribution and depends on nuisance parameters in general[6], therefore it is difficulty to conduct a reliable statistic to test $J_{2\cdot} = 0$. Our situation does not seem so hopeless, because we do not actually want to test $J_{2\cdot} = 0$ but to test weak exogeneity of y_{2t}.

Under the assumption that the cointegration system has h cointegrating relations and y_{2t} is weakly exogenous, we have $J_{2\cdot} = B_2A'$ and $rank(A) = h$ i.e. there exits a $h \times h$ submatrix in A' with rank h. Hence there exits a corresponding $(n-g) \times h$ submatrix in $J_{2\cdot}$ whose Wald statistic will have a standard $\chi^2((n-g)h)$ distribution, as shown in the last subsection. We could test weak exogeneity of y_{2t} by looking at the Wald statistics of a certain $(n-g) \times h$ submatrix of $J_{2\cdot}$, if we knew that the corresponding submatrix of A would have rank h. This gives us a hint that we do not need to look at every component of $J_{2\cdot}$, it is sufficient to look at those components of $J_{2\cdot}$ that correspond to $I(0)$ combinations of y_t. In other words, we need only to look at the Wald statistic of $J_{2\cdot}$ in its $I(0)$ directions but not the I(1) direction that would have resulted in a nonstandard distribution. In following we construct a statistic that modifies the Wald statistic of $J_{2\cdot}$ by looking only at its $I(0)$ directions.

To prepare the main presentation we provide two auxiliary lemmas first.

Lemma 3.3 *Let Σ_x be an $h \times h$ full rank positive definite matrix and A be an $n \times h$ matrix with $rank(A) = h$ $(n > h)$. Let $\Sigma_y = A\Sigma_x A'$, P_y is the matrix of eigenvectors of Σ_y and Λ_h is the diagonal matrix of non-zero eigenvalues of Σ_y. Then:*

$$\Sigma_x^{-1} = A'P_y \begin{pmatrix} \Lambda_h^{-1} & 0 \\ 0 & 0 \end{pmatrix} P_y'A \,.$$

Proof: Using the definition of eigenvector and eigenvalue we have:

$$P_y'\Sigma_y P_y = P_y'A\Sigma_x A'P_y = \Lambda_y = \begin{pmatrix} \Lambda_h & 0 \\ 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} \Lambda_h^{-1} & 0 \\ 0 & 0 \end{pmatrix} P_y'A\Sigma_x A'P_y \begin{pmatrix} \Lambda_h^{-1} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_h & 0 \\ 0 & 0 \end{pmatrix} .$$

Now we define $P_{yh}^* := P_y \begin{pmatrix} \Lambda_h^{-\frac{1}{2}} \\ 0 \end{pmatrix}_{n\times h}$. Then

$$\begin{pmatrix} P_{yh}^{*\prime} \\ 0 \end{pmatrix}_{n\times n} A\Sigma_x A' \begin{pmatrix} P_{yh}^* & 0 \end{pmatrix}_{n\times n} = \begin{pmatrix} P_{yh}^{*\prime}A\Sigma_x A'P_{yh}^* & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_h & 0 \\ 0 & 0 \end{pmatrix}_{n\times n} .$$

[6] See **?** for detailed discussion.

Note that $P_{yh}^{*'}A$ is a $h \times h$ matrix with rank h. Thus we can invert it and get

$$\Sigma_x = (P_{yh}^{*'}A)^{-1}(A'P_{yh}^{*})^{-1} .$$

Therefore

$$\begin{aligned}\Sigma_x^{-1} &= A'P_{yh}^{*}P_{yh}^{*'}A = A'P_y \begin{pmatrix} \Lambda_h^{-\frac{1}{2}} \\ 0 \end{pmatrix}_{n\times h} \begin{pmatrix} \Lambda_h^{-\frac{1}{2}} & 0 \end{pmatrix}_{h\times n} P_y'A \\ &= A'P_y \begin{pmatrix} \Lambda_h^{-1} & 0 \\ 0 & 0 \end{pmatrix} P_y'A.\end{aligned}$$

□

Lemma 3.4 *Let* $\hat{\Sigma}_y \xrightarrow{P} \Sigma_y$ *and* $\hat{P}_y$ *is the matrix of the eigenvectors of* $\hat{\Sigma}_y$. $\hat{\Lambda}_h$ *is the matrix of the* h *largest eigenvalues of* $\hat{\Sigma}_y$. *Then* $\hat{P}_y \xrightarrow{P} P_y$ *and* $\hat{\Lambda}_h \xrightarrow{P} \Lambda_h$.

Proof:

Because eigenvalues are continuous function of the corresponding matrix, we have:

$$\hat{\Sigma}_y \xrightarrow{P} \Sigma_y \Rightarrow \hat{P}_y \xrightarrow{P} P_y$$

$$\hat{\Sigma}_y \xrightarrow{P} \Sigma_y \Rightarrow \hat{\Lambda}_y \xrightarrow{P} \Lambda_y \Rightarrow \hat{\Lambda}_h \xrightarrow{P} \Lambda_h$$

□

For the further calculations we introduce the following notations. We write $\{X_t\}_{t>0} = o_p(T^{-\alpha})$ if $\operatorname*{plim}_{T\to\infty} \frac{X_T}{T^{-\alpha}} = 0$ for the random sequence $\{X_t\}_{t>0}$. And we write $\{X_t\}_{t>0} = O_p(T^{-\alpha})$ if there exists a random variable X such that $\frac{X_T}{T^{-\alpha}} \xrightarrow{L} X$ for the random sequence $\{X_t\}_{t>0}$.

Lemma 3.5

$$o_p(1)O_p(1) = o_p(1).$$

Especially, for $\alpha > 0$, *we have*

$$T^{-\alpha}O_p(1) = o_p(1).$$

Proof:

The first equality follows directly from Slutzky theorem. The second equality is a special case of the first with $T^{-\alpha} \to 0 \Rightarrow T^{-\alpha} \xrightarrow{P} 0$. □

Example 1:

From Proposition 3.1 we have:

$$\frac{1}{T}\sum_{t=1}^{T} z_{1t}z_{1t}' \xrightarrow{P} \Sigma_1.$$

We can rewrite this equation as follows:

$$\frac{1}{T}\sum_{t=1}^{T} z_{1t}z_{1t}' - \Sigma_1 = o_p(1).$$

Example 2

$$\begin{aligned}
& \begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_{1t}u_t' \\ \frac{1}{T}\sum_{t=1}^{T} z_{2t}u_t' \end{pmatrix}' \begin{pmatrix} \frac{1}{T}\sum_{t=1}^{T} z_{1t}z_{1t}' & \frac{1}{T^{\frac{2}{3}}}\sum_{t=1}^{T} z_{1t}z_{2t}' \\ \frac{1}{T^{\frac{2}{3}}}\sum_{t=1}^{T} z_{2t}'z_{1t} & \frac{1}{T^2}\sum_{t=1}^{T} z_{2t}z_{2t}' \end{pmatrix}^{-1} \\
= & \begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_{1t}u_t' \\ \frac{1}{T}\sum_{t=1}^{T} z_{2t}u_t' \end{pmatrix}' \begin{pmatrix} \Sigma_1 + o_p(1) & o_p(1) \\ o_p(1) & \frac{1}{T^2}\sum_{t=1}^{T} z_{2t}z_{2t}' \end{pmatrix}^{-1} \\
= & \begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_{1t}u_t' \\ \frac{1}{T}\sum_{t=1}^{T} z_{2t}u_t' \end{pmatrix}' \begin{pmatrix} (\Sigma_1^{-1} + o_p(1) & o_p(1) \\ o_p(1) & (\frac{1}{T^2}\sum_{t=1}^{T} z_{2t}z_{2t}')^{-1} + o_p(1) \end{pmatrix} \\
= & \begin{pmatrix} (\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_{1t}u_t')'\Sigma_1^{-1} + o_p(1), & (\frac{1}{T}\sum_{t=1}^{T} z_{2t}u_t')'(\frac{1}{T^2}\sum_{t=1}^{T} z_{2t}z_{2t}')^{-1} + o_p(1) \end{pmatrix} \\
= & \begin{pmatrix} (\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_{1t}u_t')'\Sigma_1^{-1} + o_p(1), & O_p(1) + o_p(1) \end{pmatrix}
\end{aligned}$$

The second equality follows from Lemma 2 of Toda and Phillips (1993). The third equality follows from the fact that:

$$\begin{pmatrix} \Sigma_1 + o_p(1) & o_p(1) \\ o_p(1) & \frac{1}{T^2}\sum_{t=1}^{T} z_{2t}z_{2t}' \end{pmatrix} \begin{pmatrix} \Sigma_1^{-1} + o_p(1) & o_p(1) \\ o_p(1) & (\frac{1}{T^2}\sum_{t=1}^{T} z_{2t}z_{2t}')^{-1} + o_p(1) \end{pmatrix} = I + o_p(1).$$

The last equality follows from Lemma 2 of Toda and Phillips (1993).

For the OLS estimation of the VECM (3.6) we have:

$$
\begin{aligned}
& \hat{\Phi} - \Phi \\
= \; & U'X(X'X)^{-1} \\
= \; & U'(ZH^{-1})(H^{-1'}Z'ZH^{-1})^{-1} \\
= \; & U'Z(Z'Z)^{-1}H' \\
= \; & U'Z\Upsilon_T^{-1}(\Upsilon_T^{-1}Z'Z\Upsilon_T^{-1})^{-1}\Upsilon_T^{-1}H' \\
= \; & \begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^T z_{1t}u_t' \\ \frac{1}{T}\sum_{t=1}^T z_{2t}u_t' \end{pmatrix}' \begin{pmatrix} \frac{1}{T}\sum_{t=1}^T z_{1t}z_{1t}' & \frac{1}{T^{\frac{3}{2}}}\sum_{t=1}^T z_{1t}z_{2t}' \\ \frac{1}{T^{\frac{3}{2}}}\sum_{t=1}^T z_{2t}'z_{1t} & \frac{1}{T^2}\sum_{t=1}^T z_{2t}z_{2t}' \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{\sqrt{T}}I_{n(k-1)+h} & 0 \\ 0 & \frac{1}{\sqrt{T}}A' \\ 0 & \frac{1}{T}A_\perp' \end{pmatrix} \\
= \; & \begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^T u_t z_{1t}'\Sigma_1^{-1} + o_p(1), & O_p(1) + o_p(1) \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{T}}I_{n(k-1)+h} & 0 \\ 0 & \frac{1}{\sqrt{T}}A' \\ 0 & \frac{1}{T}A_\perp' \end{pmatrix} \\
= \; & [\frac{1}{\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t z_{1t}'\Sigma_1^{-1}\right)_{n(k-1)} + o_p(\frac{1}{\sqrt{T}}), \\
& + \frac{1}{\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t z_{1t}'\Sigma_1^{-1}\right)_h A' + o_p(\frac{1}{\sqrt{T}}) + O_p(\frac{1}{T})A_\perp' + o_p(\frac{1}{T})A_\perp']
\end{aligned}
$$

Here $\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t z_{1t}'\Sigma_1^{-1}\right)_{n(k-1)}$ and $\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t z_{1t}'\Sigma_1^{-1}\right)_h$ are the first $(n(k-1))$ and last h columns of the matrix $\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t z_{1t}'\Sigma_1^{-1}\right)$ respectively.

For testing weak exogeneity we are only interested in $\hat{J}_k - J_k$ i.e. the last n columns of $\hat{\Phi} - \Phi$. Looking at the last n columns of the last equation only, we have:

$$
\hat{J}_k - J_k = \left[\frac{1}{\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t z_{1t}'\Sigma_1^{-1}\right)_h A' + o_p(T^{-1/2}) + O_p(\frac{1}{T})A_\perp' + o_p(\frac{1}{T})A_\perp'\right].
$$

It follows that

$$
\sqrt{T}(\hat{J}_k - J_k) = \left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t z_{1t}'\Sigma_1^{-1}\right)_h A' + o_p(1) + O_p(\frac{1}{\sqrt{T}})A_\perp' + o_p(\frac{1}{\sqrt{T}})A_\perp'\right]. \tag{3.21}
$$

Then,

$$
\sqrt{T}(\hat{J}_k - J_k)\begin{pmatrix} A' \\ A_\perp' \end{pmatrix}^{-1} = \left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t z_{1t}'\Sigma_1^{-1}\right)_h + o_p(1), O_p(\frac{1}{\sqrt{T}}) + o_p(\frac{1}{\sqrt{T}})\right]. \tag{3.22}
$$

Now we look only at some h columns of $\sqrt{T}(\hat{J}_k - J_k)$ denoted by $\sqrt{T}(\hat{J}_{k,h} - J_{k,h})$. Similar to (3.21) we have

$$\sqrt{T}(\hat{J}_{k,h} - J_{k,h}) = \left[\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t z'_{1t} \Sigma_1^{-1} \right)_h A'_{h*} + o_p(1) + O_p(\frac{1}{\sqrt{T}}) A'_{\perp h*} + o_p(\frac{1}{\sqrt{T}}) A'_{\perp h*} \right]. \tag{3.23}$$

where A_{h*} and $A_{\perp h*}$ denote the h selected rows of the A and $A_{\perp}$ matrix respectively. Because A has rank h there exists at least one submatrix A_{h*} that is invertible. From now on we denote such invertible submatrix by A_{h*} and the corresponding h columns of J_k by $J_{k,h}$. According to this definition we have:

$$\sqrt{T}(\hat{J}_{k,h} - J_{k,h}) A'^{-1}_{h^*} = \left[\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t z'_{1t} \Sigma_1^{-1} \right)_h + o_p(1) + O_p(\frac{1}{\sqrt{T}}) + o_p(\frac{1}{\sqrt{T}}) \right] \tag{3.24}$$

For simplicity of presentation but without loss of generality we consider testing the weak exogeneity of the last variable Y_{nt}. In this case i.e. $n - g = 1$. We have the following hypothesis:

$$H_0 : J_{k,nh^*} = 0 \qquad\qquad H_1 : J_{k,nh^*} \neq 0$$

where J_{k,nh^*} denotes a $1 \times h$ submatrix of of the last row of J_k. The problem of carrying out the test is that we do not know A, henceforth we do not know the position of the J_{k,nh^*} that corresponds to an invertible A_{h^*}. Consequently we can not calculate the Wald statistic, although we know that this Wald statistic would have a $\chi^2(h)$ distribution. We solve this problem by calculating a statistic that is asymptotically equivalent to the Wald statistic of $\hat{J}_{k,nh^*}$. For this reason we look at the Wald statistic of $\hat{J}_{k,nh^*} - J_{k,nh^*}$. Under H_0 we have:

$$
\begin{aligned}
& Wald(\hat{J}_{k,nh^*} - J_{k,nh^*}) \\
= \; & Wald(\hat{J}_{k,nh^*}) \qquad (3.25)\\
= \; & Wald(\sqrt{T}(\hat{J}_{k,nh^*})) \\
= \; & \sqrt{T}(\hat{J}_{k,nh^*})Var^{-1}(\sqrt{T}(\hat{J}_{k,nh^*}))\sqrt{T}(\hat{J}_{k,nh^*})' \\
= \; & \sqrt{T}(\hat{J}_{k,nh^*})A'^{-1}_{h*}Var^{-1}\left((T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})_{nh}\right)A^{-1}_{h*}\sqrt{T}(\hat{J}_{k,nh^*})' \\
& +\sqrt{T}(\hat{J}_{k,nh^*})op(1)\sqrt{T}(\hat{J}_{k,nh^*})' \\
= \; & ((T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})_{nh} + o_p(1))\left(Var\left((T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})_{nh}\right)\right)^{-1} \\
& ((T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})'_{nh} + o_p(1)) + o_p(1) \\
= \; & (T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})_{nh}\left(Var\left((T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})_{nh}\right)\right)^{-1}(T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})'_{nh} + o_p(1)
\end{aligned}
$$

Here $(T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})_{nh}$ denotes the last row of $(T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})_h$.
Let $\hat{\Sigma}_{J_{k,n}}$[7] be a consistent estimator of $\sqrt{T}$ times the covariance matrix of the OLS estimator of the last row of J_k: $Var(\sqrt{T}(\hat{J}_{k,n} - J_{k,n}))$ and $\hat{P}_{J_{k,n}}$ be the matrix of the eigenvectors such that

$$\hat{P}'_{J_{k,n}}\hat{\Sigma}_{J_{k,n}}\hat{P}_{J_{k,n}} = \hat{\Lambda}.$$

We choose the h greatest eigenvalues of $\hat{\Lambda}$ and denote it as $\hat{\Lambda}_h$. Let

$$\hat{P}^*_{J_{k,n}} = \hat{P}_{J_{k,n}}\begin{pmatrix}\hat{\Lambda}_h^{-\frac{1}{2}} \\ 0\end{pmatrix}.$$

According to (3.21) we have:

$$Var(\sqrt{T}(\hat{J}_{k,n} - J_{k,n})) = A \;\; Var\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t z'_{1t}\Sigma_1^{-1}\right)_{nh}\right)A' + o_p(1) \quad (3.26)$$

Using Lemma 3.3 and Lemma 3.4 we have:

$$A'\hat{P}_{J_{k,n}}\begin{pmatrix}\hat{\Lambda}_h^{-1} & 0 \\ 0 & 0\end{pmatrix}\hat{P}'_{J_{k,n}}A = Var^{-1}\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t z'_{1t}\Sigma_1^{-1}\right)_{nh}\right) + o_p(1) \quad (3.27)$$

[7] A ready candidate of the consistent estimator is $T\hat{\sigma}_n^2(X'X)^{-1}_{nn}$, where $(X'X)^{-1}_{nn}$ denoted the low right $n \times n$ block of $(X'X)^{-1}$ and $\hat{\sigma}_n^2$ is the OLS estimator of the variance of the residual in the last equation of the VECM.

Now we look at following statistic:

$$
\begin{aligned}
& \sqrt{T}(\hat{J}_{k,n} - J_{k,n})\hat{P}^{*}_{J_{k,n}}\hat{P}^{*\prime}_{J_{k,n}}\sqrt{T}(\hat{J}_{k,n} - J_{k,n})' \\
= & \sqrt{T}(\hat{J}_{k,n} - J_{k,n})\hat{P}_{J_{k,n}}\begin{pmatrix} \hat{\Lambda}_h^{-1} & 0 \\ 0 & 0 \end{pmatrix}\hat{P}'_{J_{k,n}}\sqrt{T}(J_{k,n} - J_{k,n})' \\
= & \sqrt{T}(\hat{J}_{k,n} - J_{k,n})\begin{pmatrix} A' \\ A'_\perp \end{pmatrix}^{-1}\begin{pmatrix} A' \\ A'_\perp \end{pmatrix}\hat{P}_{J_{k,n}}\begin{pmatrix} \hat{\Lambda}_h^{-1} & 0 \\ 0 & 0 \end{pmatrix}\hat{P}'_{J_{k,n}}\begin{pmatrix} A & A_\perp \end{pmatrix} \\
& \begin{pmatrix} A & A_\perp \end{pmatrix}^{-1}\sqrt{T}(\hat{J}_{k,n} - J_{k,n})' \\
= & \left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t z'_{1t}\Sigma_1^{-1}\right)_{nh} + o_p(1), O_p(\frac{1}{\sqrt{T}}) + o_p(\frac{1}{\sqrt{T}})\right] \\
& \begin{pmatrix} A'\hat{P}_{J_{k,n}}\begin{pmatrix} \hat{\Lambda}_h^{-1} & 0 \\ 0 & 0 \end{pmatrix}\hat{P}'_{J_{k,n}}A & A'\hat{P}_{J_{k,n}}\begin{pmatrix} \hat{\Lambda}_h^{-1} & 0 \\ 0 & 0 \end{pmatrix}\hat{P}'_{J_{k,n}}A_\perp \\ A'_\perp\hat{P}_{J_{k,n}}\begin{pmatrix} \hat{\Lambda}_h^{-1} & 0 \\ 0 & 0 \end{pmatrix}\hat{P}'_{J_{k,n}}A & A'_\perp\hat{P}_{J_{k,n}}\begin{pmatrix} \hat{\Lambda}_h^{-1} & 0 \\ 0 & 0 \end{pmatrix}\hat{P}'_{J_{k,n}}A_\perp \end{pmatrix} \\
& \left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t z'_{1t}\Sigma_1^{-1}\right)_{nh} + o_p(1), O_p(\frac{1}{\sqrt{T}}) + o_p(\frac{1}{\sqrt{T}})\right]' \\
= & (T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})_{nh}\left(Var\left((T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})_{nh}\right)\right)^{-1}(T^{-1/2}\sum u_t z'_{1t}\Sigma_1^{-1})'_{nh} + o_p(1)
\end{aligned}
$$

Comparing the equation above with (3.25) we get:

$$
Wald(\hat{J}_{h,k} - J_{h,k}) - \sqrt{T}(\hat{J}_{k,n} - J_{k,n})\hat{P}^{*}_{J_{k,n}}\hat{P}^{*\prime}_{J_{k,n}}\sqrt{T}(\hat{J}_{k,n} - J_{k,n}) \xrightarrow{P} 0 \quad (3.28)
$$

This implies that although we cannot calculate the Wald statistic for $\hat{J}_{k,nh^*}$ we are able to calculate an asymptotically equivalent statistic:

$$
\sqrt{T}(\hat{J}_{k,n} - J_{k,n})\hat{P}^{*}_{J_{k,n}}\hat{P}^{*\prime}_{J_{k,n}}\sqrt{T}(\hat{J}_{k,n} - J_{k,n}).
$$

Using this statistic we can test weak exogeneity of y_{nt}. We summarize this result in the following theorem.

Theorem 3.6 *For the $H_0 : J_{k,nh^*} = 0$, the Wald statistic is asymptotically equivalent to the statistic $\sqrt{T}(\hat{J}_{k,n} - J_{k,n})\hat{P}^{*}_{J_{k,n}}\hat{P}^{*\prime}_{J_{k,n}}\sqrt{T}(\hat{J}_{k,n} - J_{k,n})$; and they have asymptotic $\chi^2(h)$ distribution.*

For the general case of testing of weak exogeneity of the $(n-g) \times 1$ variable y_{2t} we have the hypothesis:

$$
H_0 : J_{2h^*} = 0 \qquad\qquad H_1 : J_{2h^*} \neq 0
$$

where J_{2h^*} is an $(n-g)\times h$ submatrix of J_2. Let $\hat{\Sigma}_{J_2}$ be a consistent estimator of $\sqrt{T}$ times the covariance matrix of the OLS estimator of J_2: $Var(\sqrt{T}\hat{J}_2)$. Let $\hat{P}_{J_2}$ be the matrix of eigenvectors such that

$$\hat{P}'_{J_2}\hat{\Sigma}_{J_2}\hat{P}_{J_2} = \hat{\Lambda}.$$

We choose the $h(n-g)$ greatest eigenvalues of $\hat{\Lambda}$ and denote it as $\hat{\Lambda}_h(n-g)$. Let $\hat{P^*_{J_2}} = \hat{P}_{J_2}\begin{pmatrix}\hat{\Lambda}^{-\frac{1}{2}}_{h(n-g)} & 0 \\ 0 & 0\end{pmatrix}$. Similar to the case of testing weak exogeneity of one variable we have the following theorem:

Theorem 3.7 *For the $H_0 : J_{2h^*} = 0$, the Wald statistic is asymptotically equivalent to the statistic$\sqrt{T}vec(\hat{J}_2 - J_2)'\hat{P}^*_{J_2}\hat{P}^{*\prime}_{J_2}\sqrt{T}vec(\hat{J}_2 - J_2)$, and they have asymptotically $\chi^2(h(n-g))$ distribution.*

4 Concluding Remarks

In this paper we present two alternative procedures to test for weak exogeneity in a cointegrated system. This procedure can be applied to test the weak exogeneity before the cointegration analysis and thus makes it possible to reduce the dimension of the problem in cointegration analysis. For future research it is planed to explore the performance of this test procedures and to study its relevance for empirical research.

References

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

I. Habro, S. Jahansen, B. N. A. R. (1998). Asymptotic Inference on Cointegrating Rank in Partial Systems. *Journal of Business&Economic Statistics*, 16:388–399.

Johansen, S. (1992). Testing Weak Exogneity and the Order of Cointegration in U.K. Money Demand Data. *Journal of Policy Modeling*, 14:313–334.

Toda, H. Y. and Phillips, P. (1993). Vector autoregression and causality. *Econometrica*, 61:1367–1393.