Department of Economics
University of Southampton
Southampton SO17 1BJ
UK

Discussion Papers in Economics and Econometrics

OMITTED VARIABLES IN COINTEGRATION ANALYSIS

Nicoletta Pashourtidou

No. 0304

ISSN 0966-4246

DISCUSSION PAPERS

IN ECONOMICS AND ECONOMETRICS

OMITTED VARIABLES IN COINTEGRATION ANALYSIS

Nicoletta Pashourtidou[1]

School of Economic Studies, University of Manchester, Manchester M13 9PL

Discussion Paper 0304

June 2003

Abstract: This paper investigates the effects of the omission of relevant variables from the statistical model on cointegration analysis, proposed by Johansen (1988, 1991). We show that underspecification of the statistical model leads to either failure in detecting cointegration or underestimation of the cointegrating rank. Although in the underspecified statistical model the estimator of the detected cointegrating vectors is shown to be consistent, this is not the case for the estimators of the adjustment coefficient matrix and the variance of the error term. The asymptotic analysis is supplemented by a Monte Carlo experiment and an empirical example.

Key words: cointegration, omitted variables, asymptotics, Monte Carlo

JEL: C15, C32, C52

[1] I am greatly indebted to Raymond O'Brien who motivated and supervised this paper. Comments from Grayham Mizon, Bent Nielsen and Denise Osborn are gratefully acknowledged.

I. Introduction

The likelihood ratio (LR) tests for cointegration proposed by Johansen (1988, 1991) have been widely applied in empirical research and it is of interest to study their behaviour under various types of misspecification of the statistical model (SM) used for cointegration testing. The robustness of the LR tests for cointegration has been investigated using Monte Carlo simulations under omitted or irrelevant (redundant) step and impulse dummy variables (see Andrade et al., 1994), dynamic misspecification using a data generating process (DGP) with autoregressive and moving average dynamics (see Boswijk and Franses, 1992; Cheung and Lai, 1993) and non-normality assuming non-symmetric and leptokurtic innovations (see Cheung and Lai, 1993).

An interesting form of misspecification is the underspecification or overspecification of the statistical model used for cointegration testing. This means that with respect to the DGP, either some variables have been omitted from the SM or some of the variables included in the SM are irrelevant. Podivinsky (1998) investigates the performance of the LR tests for cointegration (mainly the trace statistic) when there is a mismatch between the variables used in the SM (used for the cointegration tests) and the variables entering the true cointegrating vectors. Using Monte Carlo simulations he finds that the LR tests performed on an overspecified SM detect at least the true number of cointegrating vectors. He also finds that LR tests based on only two variables: (i) have low power when there are in fact two cointegrating vectors among three variables, and (ii) may not detect a cointegrating vector if there is only one cointegrating vector among three variables. The potential importance of these results for the applied work is illustrated by DeLoach (2001) who uses LR tests for cointegration (trace and maximal eigenvalue statistics) to test the hypothesis of cointegration between the relative price of nontradables and real output (which is consistent with the productivity-bias hypothesis of Balassa and Samuelson). In a two variable model (relative price of nontradables, real output) he finds evidence of cointegration for only two out of the nine countries considered. He attributes the lack of evidence of cointegration to the fact that certain variables, which mirror long-run determinants of the relative prices, have been omitted from the SM. After having augmented the system with the variable for oil prices, he finds evidence of cointegration for four out the nine countries.

The purpose of this study is to investigate analytically the effects of $I(1)$ omitted variables from the SM, on the inference about the cointegrating rank, carried out using the LR test statistics (trace and maximal eigenvalue), proposed by Johansen (1988, 1991). The consistency of the estimators of the parameters of the error correction model under the above form of misspecification is also considered. The analytical findings are supplemented by a Monte Carlo investigation and an empirical example.

The literature concerning the effects of misspecifications on the LR tests for cointegration is limited mainly to Monte Carlo studies. The contribution of this paper is to provide an analytical (asymptotic) investigation of the robustness of LR tests for cointegration when relevant variables are omitted from the SM, and therefore an analytical formulation of earlier Monte Carlo studies.

The organisation of the paper is as follows. Section II describes the model. Sections III and IV provide some asymptotic results concerning the implications of omitted variables for the inference about the cointegrating rank and the consistency of the maximum likelihood estimators of the parameters of the error correction model. Section V provides illustrations in the form of Monte Carlo simulations and an empirical example. Section VI concludes. The proofs of all propositions are presented in the Appendix.

A word on notation. The symbols '$\xrightarrow{p}$' and '$\xrightarrow{d}$' denote convergence in probability and convergence in distribution respectively, as the sample size, T, tends to infinity. Δ is the first difference operator, $|M|$ denotes the determinant of a square matrix M, $sp(M)$ denotes the space spanned by the columns of the matrix M and I_n denotes the identity matrix of dimension n. Moreover $E(\cdot)$, $Var(\cdot)$ and $p\lim(\cdot)$ denote the expected value, variance and probability limit of the random argument respectively.

II. The model

The DGP is given by a VAR(1) model in error correction form,

$$\Delta X_t = \Pi X_{t-1} + \varepsilon_t, \ \ t = 1, 2, \ldots T \tag{1}$$

where $\varepsilon_t \sim i.i.d.(0, \Omega)$ and X_t is a $p \times 1$, $I(1)$ process. In addition X_t is cointegrated so that $\Pi = \alpha\beta^{'}$ (α and β are $p \times r$ matrices) with $r \leq p - 1$ cointegrating vectors β such that $\beta^{'} X_t \sim I(0)$.

The SM used for cointegration testing is assumed to be underspecified i.e. it includes only a subset of the variables in the DGP. More specifically, let $H = \begin{bmatrix} I_{p^*} \\ \\ \underset{k \times p^*}{0} \end{bmatrix}$ be a selection matrix, then the SM includes $p^* < p$ variables given by $X_t^* = H^{'} X_t$ so that $k \equiv (p - p^*)$ relevant variables are omitted.

The misspecified SM takes the form of a multivariate regression of $H^{'} \Delta X_t = \Delta X_t^*$ on $H^{'} X_{t-1}$. The relation between ΔX_t^* and X_{t-1}^* does not have an error correction form as the model

$$\Delta X_t^* = \Pi^* X_{t-1}^* + e_t^*, \;\; t = 1, 2, \ldots, T \tag{2}$$

is misspecified. In particular e_t^* is in general correlated with X_{t-1}^*. We also define $\beta^{(1)} = H^{'} \beta$, and $\alpha^{(1)} = H^{'} \alpha$, but $\Pi^* \neq \alpha^{(1)} \beta^{(1)'}$ and $\Pi^* \neq H^{'} \Pi H$ as $HH^{'} \neq I_p$.

Although $\beta^{'} X_t$ is $I(0)$, $\beta^{(1)'} X_t^*$ is not necessarily $I(0)$ since a linear combination of $I(1)$ variables is in general $I(1)$. The nature of $\beta^{(1)'} X_t^*$ is determined by the variables entering the cointegrating relations in the DGP. Since only the space spanned by the columns of β can be estimated, in general, $(r - k)$ cointegrating vectors (stationary relations) can be found by applying elementary row operations on $\beta^{'}$. Thus, $\beta^{'}$ can be transformed so that

$$\beta^{'} = \begin{bmatrix} \beta_{11} & \beta_{21} & \cdots & \beta_{p1} \\ \beta_{12} & \beta_{22} & \cdots & \beta_{p2} \\ \vdots & \vdots & & \vdots \\ \beta_{1r} & \beta_{2r} & \cdots & \beta_{pr} \end{bmatrix} \approx$$

$$\begin{bmatrix} \beta^{+}_{11} & \beta^{+}_{21} & \cdots & \beta^{+}_{(p-r)1} & 1 & 0 & 0 & \cdots & 0 & 0 \\ \beta^{+}_{12} & \beta^{+}_{22} & \cdots & \cdots & \beta^{+}_{(p-(r-1))2} & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta^{+}_{1r} & \beta^{+}_{2r} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \beta^{+}_{(p-1)r} & 1 \end{bmatrix} \tag{3}$$

where the symbol $\approx$ denotes the row equivalent matrix of $\beta^{'}$ given by (3) and $p-(r-i) = p^{*}-(r-k)+i$, $i = 1, 2, \ldots, r$ is the number of non-zero elements in the i-th row. Given that only p^{*} variables are included in the SM, we should be able to recover i cointegrating relations (using the underspecified SM), as long as $p^{*}-(r-k)+i \leq p^{*}$. Thus, at most , $(r-k)$ (for $i = r-k$) cointegrating relations can be estimated from the SM, by applying the same row operations on $\beta^{(1)'}$ as on $\beta^{'}$.

In what follows the analysis is based on the fact that the cointegrating vectors, β, (as well as the adjustment coefficients, α) are not identified so β (and therefore α) can be replaced by a non-singular transformation e.g. we can replace $\beta^{'}$ by a row equivalent matrix of $\beta^{'}$. To avoid complicating the notation we retain the same symbols for the parameters (and variables) and their non-singular transformations.

Below we distinguish two cases:

Case (i). $(r-k) \leq 0$, where all the cointegrating relations in the DGP involve at least one of the omitted variables, therefore $\beta^{(1)'} X_t^* \sim I(1)$.

Case (ii). $(r-k) > 0$, where there are $q < r$, $q \geq (r-k)$, cointegrating relations in the DGP which do not involve any of the k omitted variables, accounting also for the event of fortuitous zeros. Therefore, some elements of $\beta^{(1)'} X_t^*$, $\beta_{11}^{'} X_t^*$, say are stationary, where β_{11} is a submatrix of $\beta^{(1)}$ in the following partition, $\beta^{(1)} = \begin{bmatrix} \underset{p^* \times q}{\beta_{11}} & \underset{p^* \times (r-q)}{\beta_{12}} \end{bmatrix}$. Then $\beta^{(1)'} X_t^* = \begin{bmatrix} \beta_{11}^{'} X_t^* \\ \beta_{12}^{'} X_t^* \end{bmatrix}$ and $\beta_{11}^{'} X_t^* \sim I(0)$ while $\beta_{12}^{'} X_t^* \sim I(1)$. Here we assume that the actual cointegrating vectors can be found as the first q rows of $\beta^{(1)'}$. Nevertheless, if the above ordering is not satisfied, the cointegrating vectors can be isolated in the first q rows of $\beta^{(1)'}$ using elementary row operations (see above).

The eigenvalue equation that corresponds to (2) is

$$|\zeta S_{11}^* - S_{10}^* S_{00}^{*-1} S_{01}^*| = 0 \tag{4}$$

where $S_{11}^* = T^{-1} \sum_{t=1}^{T} (X_{t-1}^* - \bar{X}^*)(X_{t-1}^* - \bar{X}^*)'$, $S_{00}^* = T^{-1} \sum_{t=1}^{T} (\Delta X_t^* - \bar{\Delta} X^*)(\Delta X_t^* - \bar{\Delta} X^*)'$, $S_{10}^* = S_{01}^{*'} = T^{-1} \sum_{t=1}^{T} (X_{t-1}^* - \bar{X}^*)(\Delta X_t^* - \bar{\Delta} X^*)'$, $\bar{X}^* = T^{-1} \sum_{t=1}^{T} X_{t-1}^*$ and $\bar{\Delta} X^* = T^{-1} \sum_{t=1}^{T} \Delta X_t^*$.

The eigenvalue equation that corresponds to the DGP is

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0$$

with S_{ij}, $i,j=0,1$, defined similarly in terms of the process X_t (the DGP).

Note that we can partition the stochastic vector X_t into $X_t = \begin{bmatrix} \underset{p^* \times 1}{X_t^*} \\ \underset{k \times 1}{X_t^{(k)}} \end{bmatrix}$ where the upper $(p^* \times 1)$ block holds the variables included in the SM and the lower $(k \times 1)$ block corresponds to the omitted variables. Then, S_{ij}^*, $i,j=0,1$, is given by the top left submatrix of the corresponding S_{ij}, $i,j=0,1$.

The matrix $S_{11}^{*-1} S_{10}^* S_{00}^{*-1} S_{01}^*$ has the same eigenvalues as the roots of (4), which coincide with the non-zero eigenvalues of

$$S^* = (DS_{11}D)^+ (DS_{10}D)(DS_{00}D)^+ (DS_{01}D)$$

where $D = \begin{bmatrix} I_{p^*} & \underset{p^* \times k}{0} \\ \underset{k \times p^*}{0} & \underset{k \times k}{0} \end{bmatrix}$ and here the superscript $+$ denotes the Moore-Penrose (generalised) inverse.

Let $Q = \begin{bmatrix} S_{11}^* & 0 \\ 0 & I_k \end{bmatrix}$, $|Q| \neq 0$ then,

$$|\zeta I_p - S^*| = |Q^{-1}||Q(\zeta I_p - S^*)| = |Q^{-1}||S^*(\zeta)| = 0,$$

where $S^*(\zeta) = Q(\zeta I_p - S^*)$. Expanding the above equation,

$$|S^*(\zeta)| = \begin{vmatrix} \zeta S_{11}^* - S_{10}^* S_{00}^{*-1} S_{01}^* & 0 \\ 0 & \zeta I_k \end{vmatrix} = |\zeta I_k||\zeta S_{11}^* - S_{10}^* S_{00}^{*-1} S_{01}^*| = 0. \qquad (5)$$

As expected, there are k zero eigenvalues which correspond to the omitted variables. The second factor of (5) is the characteristic polynomial in (4) associated with the

SM. If the LR tests are to indicate the existence of cointegration in the underspecified model, the second factor of (5) must give some eigenvalues with positive probability limits.

III. Inference about the cointegrating rank

In order to investigate how the inference about the cointegrating rank is affected we need to consider the asymptotic behaviour of $S^*(\zeta)$. In particular we examine the limiting behaviour of the eigenvalue equation corresponding to the SM, in the stationary and non-stationary directions as defined by the DGP.

Define $B_T = (\beta, T^{-1/2}\bar{\beta}_\perp)$, where $\bar{\beta}_\perp = \beta_\perp(\beta_\perp^{'}\beta_\perp)^{-1}$, $\beta_\perp$ is $p \times (p-r)$ such that $\beta^{'}\beta_\perp = 0$ and $\beta = \begin{bmatrix} \underset{p^*\times r}{\beta^{(1)}} \\ \underset{k\times r}{\beta^{(2)}} \end{bmatrix}$, $\bar{\beta}_\perp = \begin{bmatrix} \underset{p^*\times(p-r)}{\bar{\beta}_\perp^{(1)}} \\ \underset{k\times(p-r)}{\bar{\beta}_\perp^{(2)}} \end{bmatrix}$ then,

$$|B_T^{'}S^*(\zeta)B_T| = \begin{vmatrix} \beta^{'}S^*(\zeta)\beta & T^{-1/2}\beta^{'}S^*(\zeta)\bar{\beta}_\perp \\ T^{-1/2}\bar{\beta}_\perp^{'}S^*(\zeta)\beta & T^{-1}\bar{\beta}_\perp^{'}S^*(\zeta)\bar{\beta}_\perp \end{vmatrix}$$

$$= \left| \begin{bmatrix} \zeta(\beta^{(1)\prime}S_{11}^*\beta^{(1)} + \beta^{(2)\prime}\beta^{(2)}) & T^{-1/2}\zeta(\beta^{(1)\prime}S_{11}^*\bar{\beta}_\perp^{(1)} + \beta^{(2)\prime}\bar{\beta}_\perp^{(2)}) \\ T^{-1/2}\zeta(\bar{\beta}_\perp^{(1)\prime}S_{11}^*\beta^{(1)} + \bar{\beta}_\perp^{(2)\prime}\beta^{(2)}) & T^{-1}\zeta(\bar{\beta}_\perp^{(1)\prime}S_{11}^*\bar{\beta}_\perp^{(1)} + \bar{\beta}_\perp^{(2)\prime}\bar{\beta}_\perp^{(2)}) \end{bmatrix} \right.$$

$$\left. - \begin{bmatrix} \beta^{(1)\prime}S_{10}^*S_{00}^{*-1}S_{01}^*\beta^{(1)} & T^{-1/2}\beta^{(1)\prime}S_{10}^*S_{00}^{*-1}S_{01}^*\bar{\beta}_\perp^{(1)} \\ T^{-1/2}\bar{\beta}_\perp^{(1)\prime}S_{10}^*S_{00}^{*-1}S_{01}^*\beta^{(1)} & T^{-1}\bar{\beta}_\perp^{(1)\prime}S_{10}^*S_{00}^{*-1}S_{01}^*\bar{\beta}_\perp^{(1)} \end{bmatrix} \right| = 0. \qquad (6)$$

In order to analyse the limiting behaviour of (6) we resort to the Granger Representation Theorem which gives the following representation for X_t in (1)

$$X_t = C\sum_{i=1}^{t}\varepsilon_i + C_1(L)\varepsilon_t \tag{7}$$

(see Johansen, 1996, Theorem 4.2). Then for the p^*-dimensional vector of variables X_t^* included in the SM we have the following representation, by using (7),

$$X_t^* = C^*\sum_{i=1}^{t}\varepsilon_i + C_1^*(L)\varepsilon_t \tag{8}$$

where $C^* = H^{'}C$, $C_1^*(L) = H^{'}C_1(L)$ both of dimensions $p^* \times p$ and $rank(C^*) = \min(p^*, p^* - (r-k))$. Thus, for case (i) $rank(C^*) = p^*$ and for case (ii) $rank(C^*) = (p^* - q)$.

Proposition 1 gives the asymptotic results for the two cases.

Proposition 1. Case (i). When $(r-k) \leq 0$, *let* $\Upsilon_T = \begin{bmatrix} T^{-1/2}I_r & 0 \\ 0 & I_{p-r} \end{bmatrix}$, *then*

$$|\Upsilon_T^{'}B_T^{'}S^*(\zeta)B_T\Upsilon_T| \xrightarrow{d} |\zeta B^{*'}C^*\int_0^1 \tilde{W}\tilde{W}^{'}C^{*'}B^*du| = 0 \tag{9}$$

where $B^* = \begin{bmatrix} \beta^{(1)} & \bar{\beta}_{\perp}^{(1)} \end{bmatrix}$, $p^* \times p$ *and* $\tilde{W} = W(u) - \int_0^1 W(u)du$ *with* $W(u)$ *being a p-dimensional Brownian motion with variance* Ω *and* $u \in [0,1]$.

Case (ii). When $(r-k) > 0$, *let* $\Upsilon_T = \begin{bmatrix} I_q & 0 & 0 \\ 0 & T^{-1/2}I_{r-q} & 0 \\ 0 & 0 & I_{p-r} \end{bmatrix}$, *then*

$$|\Upsilon_T^{'}B_T^{'}S^*(\zeta)B_T\Upsilon_T| \xrightarrow{d}$$

$$|\zeta\Sigma^*_{\beta_{11}\beta_{11}} - \Sigma^*_{\beta_{11}0}\Sigma^{*-1}_{00}\Sigma^*_{0\beta_{11}}||\zeta B^{*'}C^*\int_0^1 \tilde{W}\tilde{W}'duC^{*'}B^*| = 0 \qquad (10)$$

where $B^* = \begin{bmatrix} \beta_{12} & \bar{\beta}^{(1)}_{\perp} \end{bmatrix}$, $p^* \times (p-q)$, $Var \begin{bmatrix} \Delta X^*_t \\ \beta'_{11}X^*_t \end{bmatrix} \equiv \begin{bmatrix} \Sigma^*_{00} & \Sigma^*_{0\beta_{11}} \\ \Sigma^*_{\beta_{11}0} & \Sigma^*_{\beta_{11}\beta_{11}} \end{bmatrix}$ *and* $\tilde{W}$ *is defined as in case (i).*

(9) shows that in the limit there are p roots at zero k of which exist by construction, since the stochastic matrix $B^{*'}C^*\int_0^1 \tilde{W}\tilde{W}'C^{*'}B^*$ has rank p^* almost surely. This suggests that performing the LR tests for cointegration using the underspecified model will lead to the rejection of the hypothesis of cointegration (i.e. acceptance of $r = 0$) as the sample size becomes larger. The limit in (9) refers to case (i) where it is assumed that all the cointegrating relations in the DGP involve at least one of the omitted variables. Thus, all linear combinations of variables in the SM are $I(1)$ and therefore no cointegrating relations can be found.

(10) indicates that there are q non-zero and $(p-q)$ zero roots in the limit, which suggests that q cointegrating vectors can be detected in the underspecified model as the sample size becomes large. The first factor in (10) gives the q positive roots and the second the $(p-r)$ zero roots. This is because in (10) the stochastic matrix $B^{*'}C^*\int_0^1 \tilde{W}\tilde{W}'duC^{*'}B^*$ with dimensions $(p-q)\times(p-q)$ has rank (p^*-q) almost surely and the $k \equiv (p-p^*)$ zero roots appear in the second factor of (10) by construction. The limit in (10) refers to case (ii) where q of the cointegrating relations in the DGP do not involve any of the k omitted variables, thus some linear combinations of the

variables in the SM are $I(0)$ and therefore some cointegrating relations can be found.

IV. Consistency

The analysis of consistency is carried out only for case (ii) where some cointegrating vectors can be detected. For case (i) all the estimated eigenvalues converge in probability to zero and therefore the cointegrating space is consistently estimated by the null space.

For the analysis of consistency we use the partition of β that appears in Section II,

$$\beta = \begin{bmatrix} \underset{p^* \times q}{\beta_{11}} & \underset{p^* \times (r-q)}{\beta_{12}} \\ \underset{k \times q}{\beta_{21}} & \underset{k \times (r-q)}{\beta_{22}} \end{bmatrix}$$

where $\beta_{21} = 0$. We define $B = \begin{bmatrix} \underset{p^* \times q}{\beta_{11}} & \underset{p^* \times (p^*-q)}{\bar{\beta}_{11\perp}} \end{bmatrix}$ and $B^{-1} = \begin{bmatrix} \underset{q \times p^*}{\bar{\beta}_{11}^{'}} \\ \underset{(p^*-q) \times p^*}{\beta_{11\perp}^{'}} \end{bmatrix}$ where $\bar{\beta}_{11\perp} = \beta_{11\perp}(\beta_{11\perp}^{'}\beta_{11\perp})^{-1}$, $\bar{\beta}_{11} = \beta_{11}(\beta_{11}^{'}\beta_{11})^{-1}$ and $\beta_{11}^{'}\beta_{11\perp} = 0$. B and B^{-1} are such that the following relationship holds

$$B^{-1}B = BB^{-1} = \beta_{11}\bar{\beta}_{11}^{'} + \bar{\beta}_{11\perp}\beta_{11\perp}^{'} = I_{p^*}. \tag{11}$$

We have shown in Section III that the tests detect q cointegrating vectors, hence under the assumption of cointegration Π^* in (2) has rank q. Thus, Π^* can be expressed as

$\Pi^* = \alpha_{11}\beta_{11}^{'}$, where α_{11}[2] and β_{11} are $p^* \times q$ matrices of rank q. The SM then takes the form

$$\Delta X_t^* = \alpha_{11}\beta_{11}^{'} X_{t-1}^* + e_t^* \tag{12}$$

with $Var(e_t^*) \equiv \Lambda^*$.

Let $\hat{\beta}_{11}$, $\hat{\alpha}_{11}$ and $\hat{\Lambda}^*$ be the maximum likelihood estimators of β_{11}, α_{11} and Λ^* calculated from the SM (2) (using (4)). The parameters β_{11} and α_{11} correspond to the $p^* \times q$ submatrices of β, α in the DGP.

For the analysis of consistency we use a linear transformation of the columns of $\hat{\beta}_{11}$, which also maximises the likelihood function[3] given by

$$\begin{aligned} \tilde{\beta}_{11} &= \hat{\beta}_{11}(\bar{\beta}_{11}^{'}\hat{\beta}_{11})^{-1} \qquad (13) \\ &= \beta_{11} + \bar{\beta}_{11\perp}\beta_{11\perp}^{'}\hat{\beta}_{11}(\bar{\beta}_{11}^{'}\hat{\beta}_{11})^{-1} \\ &= \beta_{11} + \bar{\beta}_{11\perp}b_1 \end{aligned}$$

where the second equality follows by using (11) and $b_1 = \beta_{11\perp}^{'}\tilde{\beta}_{11}$.

---

[2]Partitioning α similarly to β we obtain $\alpha = \begin{bmatrix} \underset{p^*\times q}{\alpha_{11}} & \underset{p^*\times(r-q)}{\alpha_{12}} \\ \underset{k\times q}{\alpha_{21}} & \underset{k\times(r-q)}{\alpha_{22}} \end{bmatrix}$, where $H^{'}\alpha = \alpha^{(1)} = \begin{bmatrix} \underset{p^*\times q}{\alpha_{11}} & \underset{p^*\times(r-q)}{\alpha_{12}} \end{bmatrix}$ and α_{11} are the adjustment coefficients that correspond to the cointegrating vectors which can be detected in the underspecified model.

[3]In fact for any normalisation c we can define $\hat{\beta}_c = \hat{\beta}(c^{'}\hat{\beta})^{-1} = \tilde{\beta}(c^{'}\tilde{\beta})^{-1}$; expanding around β and normalising β and $\hat{\beta}$ by $c^{'}\beta = c^{'}\hat{\beta} = I_r$, we obtain $\hat{\beta} - \beta = (I_p - \beta c^{'})(\tilde{\beta} - \beta) + O_p(|\tilde{\beta} - \beta|^2)$ (see Johansen, 1996, p. 180) therefore the properties of $\hat{\beta}$ follow from those of $\tilde{\beta}$.

We also define $\tilde{\alpha}_{11} = \hat{\alpha}_{11}\hat{\beta}_{11}^{'}\bar{\beta}_{11}$ such that $\tilde{\alpha}_{11}\tilde{\beta}_{11}^{'} = \hat{\alpha}_{11}\hat{\beta}_{11}^{'}$ and

$$\begin{aligned}\tilde{\alpha}_{11} &= S_{01}^{*}\hat{\beta}_{11}(\hat{\beta}_{11}^{'}S_{11}^{*}\hat{\beta}_{11})^{-1}\hat{\beta}_{11}^{'}\bar{\beta}_{11} \\ &= S_{01}^{*}\tilde{\beta}_{11}(\tilde{\beta}_{11}^{'}S_{11}^{*}\tilde{\beta}_{11})^{-1}\end{aligned}$$

where the first equality follows from the fact that $\hat{\alpha}_{11} = S_{01}^{*}\hat{\beta}_{11}(\hat{\beta}_{11}^{'}S_{11}^{*}\hat{\beta}_{11})^{-1}$ (see equation (6.11) in Johansen, 1996) given that we can estimate β_{11} by solving (4).

In addition,

$$\begin{aligned}\hat{\Lambda}^{*} &= S_{00}^{*} - S_{01}^{*}\hat{\beta}_{11}(\hat{\beta}_{11}^{'}S_{11}^{*}\hat{\beta}_{11})^{-1}\hat{\beta}_{11}^{'}S_{10}^{*} \\ &= S_{00}^{*} - S_{01}^{*}\tilde{\beta}_{11}(\tilde{\beta}_{11}^{'}S_{11}^{*}\tilde{\beta}_{11})^{-1}\tilde{\beta}_{11}^{'}S_{10}^{*}\end{aligned}$$

where the first equality follows from the expression for the estimator of the variance-covariance matrix of the errors in the SM (see equation (6.12) in Johansen, 1996) and the second equality follows from the definition of $\tilde{\beta}_{11}$.

The proposition below establishes the consistency of the maximum likelihood estimator for the cointegrating vectors in the sense that the estimator from the underspecified SM converges in probability to a submatrix of the parameter, β, in the DGP, which is associated with the included variables.

Proposition 2. The estimator of the cointegrating vectors, $\tilde{\beta}_{11}$, associated with the underspecified model (2) *converges to vectors in $sp(\beta)$, i.e. $T^{1/2}(\tilde{\beta}_{11} - \beta_{11}) \xrightarrow{p} 0$.*

We then consider the probability limits of $\tilde{\alpha}_{11}$ and $\hat{\Lambda}^{*}$ obtained from the underspecified model. We first partition α and β conformably with $X_t = \begin{bmatrix} X_t^* \\ X_t^{(k)} \end{bmatrix}$ (see also Section II) and we use the transformed, row equivalent form of β. Then, the DGP (1) becomes[4],

$$\begin{bmatrix} \Delta X_t^* \\ \Delta X_t^{(k)} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} \beta_{11}^{'} & 0 \\ \beta_{12}^{'} & \beta_{22}^{'} \end{bmatrix} \begin{bmatrix} X_{t-1}^* \\ X_{t-1}^{(k)} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^* \\ \varepsilon_t^{(k)} \end{bmatrix}.$$

The part of the DGP that corresponds to the included variables is

$$\Delta X_t^* = \alpha_{11}\beta_{11}^{'}X_{t-1}^* + \alpha_{12}(\beta_{12}^{'}X_{t-1}^* + \beta_{22}^{'}X_{t-1}^{(k)}) + \varepsilon_t^*$$

or

$$\Delta X_t^* = \alpha_{11}\beta_{11}^{'}X_{t-1}^* + \alpha_{12}Z_{t-1} + \varepsilon_t^* \tag{14}$$

where $\varepsilon_t^* = H^{'}\varepsilon_t \sim i.i.d.(0, \Omega^*)$, $\Omega^* = H^{'}\Omega H$ and $Z_{t-1} = \beta_{12}^{'}X_{t-1}^* + \beta_{22}^{'}X_{t-1}^{(k)} \sim I(0)$, is the part of the DGP that cannot be estimated due to the omission of $X_t^{(k)}$.

The proposition below relates to the 'inconsistency' of $\tilde{\alpha}_{11}$ and $\hat{\Lambda}^*$ in the sense that their probability limits are different from the parameters, in the underspecified model, that they aim to estimate.

Proposition 3. The estimators $\tilde{\alpha}_{11}$ *and* $\hat{\Lambda}^*$ *are 'inconsistent' for the parameters* α_{11} and Ω^* *in* (14) *in the sense that they do not converge to the submatrices of* α *and* Ω *(parameters of the DGP) that correspond to the included variables i.e.* $\text{plim}\ \tilde{\alpha}_{11} \neq \alpha_{11}$ *and* $\text{plim}\ \hat{\Lambda}^* \neq \Omega^*$.

[4]Note that in the DGP, $E(\beta^{'}X_{t-1}\varepsilon_t^{'}) = 0$.

V. Illustrations

A Monte Carlo experiment

In this section we present the results of some Monte Carlo experiments in order to illustrate the asymptotic results presented in Sections III and IV and to give some idea about the consequences of possible misspecifications of the SM, in finite samples, in the case of omitted variables.

All calculations were done using Ox 3.00 (see Doornik, 1999). The number of replications is 10,000 for all experiments. We use the 95% tabulated asymptotic critical values from Osterwald-Lenum (1992, Case 0), thus the tests are carried out at 5% significance level.

We use two DGPs which are chosen on the basis of the asymptotic analysis to reflect the cases $(r-k)=0$ and $(r-k)>0$, treated in Section III. Both DGPs consist of three variables, but the first one (DGP1) has one cointegrating vector involving all three variables whereas the second one (DGP2) has two cointegrating vectors, both involving all three variables[5]. Thus,

$$\begin{bmatrix} \Delta X_{1t} \\ \Delta X_{2t} \\ \Delta X_{3t} \end{bmatrix} = \begin{bmatrix} 0.1 \\ 0.1 \\ -0.7 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} X_{1(t-1)} \\ X_{2(t-1)} \\ X_{3(t-1)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \tag{DGP1}$$

and

$$\begin{bmatrix} \Delta X_{1t} \\ \Delta X_{2t} \\ \Delta X_{3t} \end{bmatrix} = \begin{bmatrix} 0.433 & 0.233 \\ 0.5 & 0.3 \\ 0.366 & 0.366 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 1 & -0.5 & -0.5 \end{bmatrix} \begin{bmatrix} X_{1(t-1)} \\ X_{2(t-1)} \\ X_{3(t-1)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \tag{DGP2}$$

where $t=1,2,\ldots,T$, $\varepsilon_t = \begin{bmatrix} \varepsilon_{1t} & \varepsilon_{2t} & \varepsilon_{3t} \end{bmatrix}' \sim i.i.d.N_3(0,I)$ for DGP1 and DGP2.

The SMs used for the calculation of the trace and maximal eigenvalue statistics include only X_{1t} and X_{2t}.

Tables 1 and 2 show the rejection frequencies for various rank hypotheses using the trace and the maximal eigenvalue statistics, for different sample sizes.

[5] Similar DGPs were used by Podivinsky (1998).

Table 1. Rejection frequencies using the trace and

the maximal eigenvalue statistics (DGP1).

Sample size / Rank hypothesis	50	100	150	500	800
	Trace statistic				
$r = 0$	0.1363	0.1474	0.1517	0.1571	0.1606
$r \leq 1$	0.0166	0.0168	0.0178	0.0162	0.0164
	Maximal eigenvalue statistic				
$r = 0$	0.1379	0.1503	0.1563	0.1583	0.1627
$r \leq 1$	0.0166	0.0168	0.0178	0.0162	0.0164

Table 2. Rejection frequencies using the trace and

the maximal eigenvalue statistics (DGP2).

Sample size / Rank hypothesis	50	100	150	500	800
	Trace statistic				
$r = 0$	1	1	1	1	1
$r \leq 1$	0.0747	0.0686	0.0669	0.0722	0.0686
	Maximal eigenvalue statistic				
$r = 0$	1	1	1	1	1
$r \leq 1$	0.0747	0.0686	0.0669	0.0722	0.0686

$>$From Table 1 we can see that the tests might not detect any cointegrating vectors (low rejection frequencies of $r = 0$, especially for small sample sizes) which is what we expected since $(r - k) = 0$ (see Section III). From Table 2 we conclude that with DGP2 the LR tests are very likely to detect one cointegrating vector and this is in accordance with the theoretical finding which suggests that if $(r - k) > 0$ the tests detect at least $(r - k)$ (2-1=1, in this case) cointegrating vectors.

The following Monte Carlo experiments use a very large T value to evaluate the probability limits of $\tilde{\beta}$ and $\tilde{\alpha}$. We use a modified form of DGP2, in particular we use a matrix whose rows are linear transformations of the rows of β' found by adding to the first row twice the second row i.e.

$$\begin{bmatrix} 1 & -2 & 1 \\ 1 & -0.5 & -0.5 \end{bmatrix} \approx \begin{bmatrix} 3 & -3 & 0 \\ 1 & -0.5 & -0.5 \end{bmatrix},$$

where $\approx$ denotes a row equivalent matrix. Based on the asymptotic analysis of Section IV, if we omit variable X_{3t} we should expect one cointegrating vector whose estimator converges to the space spanned by β_{11} in the notation of Section IV, and in this case $\beta'_{11} = \begin{bmatrix} 3 & -3 \end{bmatrix}$. Table 3 shows the quantiles of the elements of the estimated cointegrating vector, $\tilde{\beta}_{11} = \begin{bmatrix} \tilde{\beta}_{11}^{(1)} \\ \tilde{\beta}_{11}^{(2)} \end{bmatrix}$ (associated with the largest eigenvalue) and the elements of the eigenvector corresponding to the smallest eigenvalue. In fact we use the normalised form of the estimated cointegrating vector, $\tilde{\beta}_{11}$ given in (13), in order to achieve convergence to the true (known) submatrix of the true β, β_{11}, instead of a linear combination of it. The estimation is carried out using T = 5,000 and 10,000 replications.

Table 3. Quantiles of the elements of

the estimated eigenvectors.

$\hat{v}$‡ / Quantiles	$\tilde{\beta}_{11}^{(1)}$	$\tilde{\beta}_{11}^{(2)}$	$\hat{v}_{12}$	$\hat{v}_{22}$
1%	2.9999	-3.0001	-0.0303	-0.0315
5%	3.0000	-3.0000	-0.0186	-0.0219
10%	3.0000	-3.0000	-0.0127	-0.0157
25%	3.0001	-2.9999	-0.0058	-0.0057
50%	3.0001	-2.9999	0.0001	-0.0000
75%	3.0003	-2.9997	0.0061	0.0053
90%	3.0005	-2.9995	0.0133	0.0151
95%	3.0007	-2.9993	0.0194	0.0209
99%	3.0011	-2.9989	0.0296	0.0321

‡Note. The first column of $\hat{v} = \begin{bmatrix} \tilde{\beta}_{11}^{(1)} & \hat{v}_{12} \\ \tilde{\beta}_{11}^{(2)} & \hat{v}_{22} \end{bmatrix}$ holds the eigenvector which corresponds to the largest eigenvalue, i.e. the normalised estimated cointegrating vector, $\tilde{\beta}_{11}$ whereas $(\hat{v}_{12}X_{1t} + \hat{v}_{22}X_{2t}) \sim I(1)$.

In Table 3 we can see that the elements of the estimated cointegrating vector, after normalisation converge to the appropriate elements of the submatrix of β in the DGP namely $\beta_{11}^{'} = \begin{bmatrix} 3 & -3 \end{bmatrix}$. The elements of the other estimated eigenvector, which is associated with the smallest eigenvalue seem to be sufficiently small.

Next we use DGP2 and a SM with only X_{1t} and X_{2t} to compute the quantiles of the elements of the estimated adjustment coefficient matrix. The estimator of α_{11} used in the simulations is given by $\tilde{\alpha}_{11} = \hat{\alpha}_{11}\hat{\beta}_{11}^{'}\bar{\beta}_{11}$ (Section IV) which is a transformation of $\hat{\alpha}_{11}$ such that $\tilde{\alpha}_{11}\tilde{\beta}_{11}^{'} = \hat{\alpha}_{11}\hat{\beta}_{11}^{'}$. For $T = 5{,}000$ and 10,000 replications the estimated adjustment coefficients seem to converge to the sum of the true adjustment coefficient matrix (i.e. the part of α, α_{11} say, in the DGP that corresponds to the single cointegrating vector that can be detected using the misspecified SM) and the asymptotic bias. For this case we have $\alpha_{11} = \begin{bmatrix} \alpha_{11}^{(1)} \\ \alpha_{11}^{(2)} \end{bmatrix} = \begin{bmatrix} 0.433 \\ 0.5 \end{bmatrix}$, and $\tilde{\alpha}_{11} = \begin{bmatrix} \tilde{\alpha}_{11}^{(1)} \\ \tilde{\alpha}_{11}^{(2)} \end{bmatrix}$ is the transformed estimator of α_{11}. The results appear in Table 4.

Table 4. Quantiles of the estimated

adjustment coefficients.

$\frac{\tilde{\alpha}_{11}}{\text{Quantiles}}$	$\tilde{\alpha}_{11}^{(1)}$	$\tilde{\alpha}_{11}^{(2)}$
1%	0.4879	0.5730
5%	0.4901	0.5752
10%	0.4914	0.5763
25%	0.4935	0.5783
50%	0.4957	0.5804
75%	0.4980	0.5826
90%	0.5002	0.5847
95%	0.5014	0.5859
99%	0.5036	0.5880

Table 4 provides an illustration of Proposition 3 namely that the estimator of the adjustment coefficients in an underspecified SM is inconsistent or asymptotically biased. From Table 4 we can see that the normalised estimated adjustment coefficients are biased upwards.

An empirical example

To illustrate the issue of omitted variables we use the four-equation system of narrow money (M1), prices, aggregate expenditure and interest rates for the UK. The data[6] are quarterly, seasonally adjusted, covering the period 1963Q1-1986Q2 on the following variables: nominal M1 (M), real total final expenditure at 1985 prices (I), total final expenditure deflator (P) with 1985 as the base year, three-month local authority interest rate (R_1) and learning-adjusted interest rate on checking accounts at commercial banks (R_2). In the analysis the difference $R = R_1 - R_2$ is used instead of the R_1 or R_2 individually. The logarithms of the above variables are denoted by the corresponding lower case letters. The interrelations among these variables have been investigated extensively in the literature (see inter alia, Hendry and Mizon, 1993; Hendry and Doornik, 1994; Ericsson et al., 1998; Doornik et al., 1998).

Following Doornik et al. (1998), there are two anticipated cointegrating relations

$$(m-p)_t = c_{01} + c_{11} i_t + c_{21} \Delta p_t + c_{31} R_t \tag{15}$$

$$i_t = c_{02} + c_{12} t + c_{22} \Delta p_t + c_{32} R_t \tag{16}$$

thus equation (15) imposes long-run price homogeneity and equation (16) has a linear trend (t) that captures exogenous technical progress. c_{11} is expected to be positive and it can possibly be restricted to $c_{11} = 1$, making (15) a relation in the inverse velocity of money. c_{21}, c_{31} are expected to be negative. In (16) c_{22} and c_{32} are expected to be positive and negative respectively.

[6] The data set is supplied with PcGive 10.0. The numerical results were obtained using PcGive 10.0 (see Doornik and Hendry, 2001).

For the particular sample the variables $(m-p)_t$, i_t, Δp_t and R_t were found to be $I(1)$ (the results of unit root tests are omitted for the sake of brevity).

The first SM (SM1) is a VAR(3) in all four variables, $(m-p)_t$, i_t, Δp_t and R_t, which includes also an unrestricted constant, a restricted time trend and two unrestricted dummy variables that account for shocks in output and prices. This formulation was used by Doornik et al. (1998). The second SM (SM2) is the same as the first (a VAR(3)) except that the potentially relevant variable R_t is omitted. R_t enters both anticipated cointegrating relations and if both of them exist in the DGP (and therefore can be detected by the tests with high probability) cointegration tests should detect one cointegrating relation when R_t is omitted. This follows from the asymptotic analysis and the evidence from the simulations. The third SM (SM3) is the same as SM1 (again a VAR(3)) except that the variable $(m-p)_t$ is omitted. The omitted variable in this case appears in only one of the anticipated cointegrating relations therefore its omission should not affect the detection of the cointegrating relation that does not involve $(m-p)_t$, provided that both anticipated relations are present in the DGP.

Tables 5, 6 and 7 show the statistics and p-values of the system diagnostic tests for SM1, SM2 and SM3 respectively.

Table 5. System diagnostic tests for SM1

Test	Test Statistic	p-value
Autocorrelation	F(80, 254)=1.14	0.23
Normality	$\chi^2(8)$=15.04	0.06
Heteroscedasticity	F(260, 503)=0.75	0.99

Table 6. System diagnostic tests for SM2

Test	Test Statistic	p-value
Autocorrelation	F(45, 217)=1.22	0.18
Normality	$\chi^2(6)$=7.60	0.27
Heteroscedasticity	F(120, 377)=0.78	0.94

Table 7. System diagnostic tests for SM3

Test	Test Statistic	p-value
Autocorrelation	F(45, 217)=1.15	0.25
Normality	$\chi^2(6)$=9.24	0.16
Heteroscedasticity	F(120, 377)=1.04	0.38

The first diagnostic test is a Lagrange Multiplier test for 5-th order residual vector autocorrelation, the second is a vector normality test and the third test is a vector heteroscedasticity test (see Doornik and Hendry, 2001). The results of the diagnostic tests do not indicate any source of misspecification. The omission of a potentially relevant variable does not seem to affect the statistical adequacy of SM2 and SM3.

Tables 8, 9 and 10 report the results of cointegration tests for SM1, SM2 and SM3 respectively. Rejection of the null hypothesis at 1% level of significance is indicated by **.

Table 8. Cointegration tests for SM1

Null hypothesis	Trace statistic	Maximal eigenvalue statistic
$r = 0$	76.58**	41.76**
$r \leq 1$	34.85	16.28
$r \leq 2$	18.57	12.01
$r \leq 3$	6.56	6.56

Table 9. Cointegration tests for SM2

Null hypothesis	Trace statistic	Maximal eigenvalue statistic
$r = 0$	38.37	18.18
$r \leq 1$	20.19	12.11
$r \leq 2$	8.09	8.09

Table 10. Cointegration tests for SM3

Null hypothesis	Trace statistic	Maximal eigenvalue statistic
$r = 0$	49.68**	34.57**
$r \leq 1$	15.11	11.62
$r \leq 2$	3.50	3.50

For SM1 the hypothesis that $r = 0$ is rejected by both the trace and the maximal eigenvalue tests. Therefore only one of the two anticipated cointegrating relations can be detected by the tests. Since the cointegrating vectors are not identified it cannot be determined at this stage which of the equations (15) or (16) the cointegrating vector corresponds to. When the relevant variable R_t is omitted and SM2 is used for cointegration testing neither the trace nor the maximal eigenvalue test rejects the hypothesis that $r = 0$. Hence, in the three-variable system no cointegrating relations can be detected. This finding was somehow expected given the result of cointegration tests for SM1 and given the fact that both (15) and (16) include the omitted variable R_t. However, the omission of $(m-p)_t$ does lead to rejection of the hypothesis $r = 0$ in the three-variable system. In this case the tests seem to detect the second anticipated cointegrating relation.

Carrying out restricted estimation of the cointegrating vectors, it is found that the single cointegrating vector in SM1 is identified as the first anticipated cointegrating relation given by (15) and the single cointegrating vector in SM3 is identified as the second anticipated relation given by (16). In SM1 the coefficient of the linear trend is restricted to 0 and the coefficient of i_t is restricted to -1. The test statistic for these restrictions is $\chi^2(2) = 0.617$ with p-value equal to 0.734. The results of the restricted estimation appear in Table 11. In SM3 the coefficient of the trend is restricted to -0.007 which is the negative of the mean of Δi_t and the test statistic is $\chi^2(1) = 1.112$ with p-value equal to 0.291. The results of the restricted estimation appear in Table 12. Thus, in SM3 cointegration tests detect the second anticipated cointegrating relation given by (16) which the former possibly lack power to detect in SM1. Even though the results of the diagnostic tests of Table 7 do not indicate any misspecification in SM3, the sign and the significance of Δp_t do give a hint.

Table 11. Estimates of restricted cointegrating vector and adjustment coefficients for SM1.

	$\hat{\alpha}$	$\hat{\beta}$
$(m-p)_t$	-0.103 (0.019)	1.000 (−)
i_t	-0.009 (0.012)	-1.000 (−)
Δp_t	0.004 (0.008)	6.506 (1.143)
R_t	-0.004 (0.015)	7.155 (0.553)
t	−	0.000 (−)

Table 12. Estimates of restricted cointegrating vector and adjustment coefficients for SM3.

	$\hat{\alpha}$	$\hat{\beta}$
i_t	-0.062 (0.0116)	1.000 $(-)$
Δp_t	-0.0177 (0.007)	0.564 (1.593)
R_t	-0.027 (0.014)	2.629 (0.608)
t	$-$	-0.007 $(-)$

The empirical example shows that the diagnostic tests are not always of help in pointing out misspecification due to omitted variables. This is because in an error correction model the omitted variables bias depends on submatrices of α, α_{12} (see (14) and proof of Proposition 3). Thus, if $\alpha_{12} = 0$ i.e. the variables in the DGP do not adjust to cointegrating relations that involve omitted variables, the bias is zero and therefore omission of relevant variables from the system may not be reflected in, for example autocorrelation in the residuals of the model.

VI. Concluding remarks

This paper has considered the effects of underspecifying (omission of relevant variables) the SM on the LR tests for cointegration proposed by Johansen (1988, 1996). We showed that omitting relevant variables from the SM will lead to either no detection of cointegrating relationships, if the true cointegrating rank is smaller than or equal to the number of omitted variables ($r \leq k$) or the detection of $q < r$ cointegrating relationships, if the true cointegrating rank is greater than the number of omitted variables ($r > k$). In addition, the use of an underspecified SM does not affect the consistency of the estimated cointegrating vectors since they still converge to a subspace of $sp(\beta)$ but it does affect the consistency of the estimators of the adjustment coefficient matrix and variance of the errors.

The model used to investigate the effects of omitted variables is quite simple, being a VAR(1) without deterministic terms, in order to minimise the complexity of the algebra involved. Since the effect of short-run dynamics is asymptotically negligible, their inclusion in the model would not alter the asymptotic findings. Inclusion of deterministic terms would require different scaling matrices that would take into account the deterministic direction in the p-dimensional space, however the asymptotic results would remain unchanged.

Although the analytical results are asymptotic, small sample simulations show that the theoretical findings also arise in sample sizes used in empirical work. The empirical example also illustrates this point.

The omitted variables can also be $I(0)$. Since the inclusion of a stationary variable increases the dimensions of the cointegrating space by one, omission of only $I(0)$ variables will lead to the underestimation of the cointegrating rank by the number of omitted $I(0)$ variables.

Overall we conclude that the omission of relevant variables from the SM leads to misleading inference, especially when followed by tests for linear restrictions on α and β conditional on the wrong cointegrating rank.

Appendix

Proof of Proposition 1.

Let the non-stationary direction for the process X_t^* be B^* which is $p^* \times p$ for case (i) and $p^* \times (p-q)$ for case (ii) (for the detailed form of B^* see under the relevant cases in Section III). By application of the Functional Central Limit Theorem on (8) and the Continuous Mapping Theorem (see (B.12) and Theorem B.5 in Johansen, 1996) we have

$$T^{-1/2} B^{*'} X^*_{[Tu]} = T^{-1/2} B^{*'} (C^* \sum_{i=1}^{[Tu]} \varepsilon_{[Tu]} + C_1^*(L) \varepsilon_{[Tu]}) \xrightarrow{d} B^{*'} C^* W(u)$$

$$B^{*'} \bar{X}^* \xrightarrow{d} B^{*'} C^* \int_0^1 W(u) du$$

and

$$\begin{aligned} T^{-1} B^{*'} S_{11}^* B^* &= T^{-2} B^{*'} \sum_{t=1}^{T} (X_{t-1}^* - \bar{X}^*)(X_{t-1}^* - \bar{X}^*)' B^* \\ &\xrightarrow{d} B^{*'} C^* \int_0^1 \tilde{W} \tilde{W}' C^{*'} B^* du. \end{aligned} \qquad (17)$$

Case (i)

Since $\beta^{(1)'} X_t^*$ is not $I(0)$, because of the omission of relevant variables, (6) is not appropriately scaled for convergence. Pre- and post-multiplying (6) by the scaling matrix $\Upsilon_T = \begin{bmatrix} T^{-1/2}I_r & 0 \\ 0 & I_{p-r} \end{bmatrix}$ we obtain,

$$|\Upsilon_T' B_T' S^*(\zeta) B_T \Upsilon_T| =$$

$$\left| \begin{bmatrix} T^{-1}\zeta\beta^{(1)'}S_{11}^*\beta^{(1)} + o_p(1) & T^{-1}\zeta\beta^{(1)'}S_{11}^*\bar{\beta}_\perp^{(1)} + o_p(1) \\ T^{-1}\zeta\bar{\beta}_\perp^{(1)'}S_{11}^*\beta^{(1)} + o_p(1) & T^{-1}\zeta\bar{\beta}_\perp^{(1)'}S_{11}^*\bar{\beta}_\perp^{(1)} + o_p(1) \end{bmatrix} - \begin{bmatrix} o_p(1) & o_p(1) \\ o_p(1) & o_p(1) \end{bmatrix} \right| \quad (18)$$

$$= \left| T^{-1}\zeta B^{*'} S_{11}^* B^* + o_p(1) \right|$$

where $B^* = \begin{bmatrix} \beta^{(1)} & \bar{\beta}_\perp^{(1)} \end{bmatrix}$, $p^* \times p$. The second matrix in (18) is $o_p(1)$ because its blocks are products of averages of products of either two $I(0)$ processes (S_{00}^*) or an $I(0)$ and an $I(1)$ process ($B^{*'}S_{10}^*$), which are $O_p(1)$ (see (B.12) in Johansen, 1996), thus after scaling by Υ_T they all become $o_p(1)$.

Then we have

$$|\Upsilon_T' B_T' S^*(\zeta) B_T \Upsilon_T| = \left| T^{-1} B^{*'} S_{11}^* B^* + o_p(1) \right|$$

$$\xrightarrow{d} \left| \zeta B^{*'} C^* \int_0^1 \tilde{W}\tilde{W}' C^{*'} B^* du \right| = 0 \quad (19)$$

by (17).

Case (ii)

In what follows we will use the row equivalent form of β that appears in (3). Consequently in a 2×2 block-partition of β the lower left block of β or equivalently the upper right block of $\beta^{'}$ is zero. Thus,

$$\beta = \begin{bmatrix} \underset{p^*\times q}{\beta_{11}} & \underset{p^*\times(r-q)}{\beta_{12}} \\ \underset{k\times q}{\beta_{21}} & \underset{k\times(r-q)}{\beta_{22}} \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ 0 & \beta_{22} \end{bmatrix}.$$

We then have the following partitions: $\beta^{(1)} = \begin{bmatrix} \beta_{11} & \beta_{12} \end{bmatrix}$ defined above and $\beta^{(2)} = \begin{bmatrix} \underset{k\times q}{\beta_{21}} & \underset{k\times(r-q)}{\beta_{22}} \end{bmatrix} = \begin{bmatrix} 0 & \beta_{22} \end{bmatrix}$. Note that β_{11} must satisfy the condition $\beta^{'}_{11}C^* = 0$ so that $\beta^{'}_{11}X^*_t = \beta^{'}_{11}C^*_1(L)\varepsilon_t \sim I(0)$, by (8).

Then (6) becomes

$$|B^{'}_T S^*(\zeta)B_T| =$$

$$\left| \begin{bmatrix} \zeta\beta^{'}_{11}S^*_{11}\beta_{11} & \zeta\beta^{'}_{11}S^*_{11}\beta_{12} & T^{-1/2}\zeta\beta^{'}_{11}S^*_{11}\bar{\beta}^{(1)}_{\perp} \\ \zeta\beta^{'}_{12}S^*_{11}\beta_{11} & \zeta(\beta^{'}_{12}S^*_{11}\beta_{12} + \beta^{'}_{22}\beta_{22}) & T^{-1/2}\zeta(\beta^{'}_{12}S^*_{11}\bar{\beta}^{(1)}_{\perp} + \beta^{'}_{22}\bar{\beta}^{(2)}_{\perp}) \\ T^{-1/2}\zeta\bar{\beta}^{(1)'}_{\perp}S^*_{11}\beta_{11} & T^{-1/2}\zeta(\bar{\beta}^{(1)'}_{\perp}S^*_{11}\beta_{12} + \bar{\beta}^{(2)'}_{\perp}\beta_{22}) & T^{-1}\zeta(\bar{\beta}^{(1)'}_{\perp}S^*_{11}\bar{\beta}^{(1)}_{\perp} + \bar{\beta}^{(2)'}_{\perp}\bar{\beta}^{(2)}_{\perp}) \end{bmatrix} \right.$$

$$\left. - \begin{bmatrix} \beta^{'}_{11}S^*_{10}S^{*-1}_{00}S^*_{01}\beta_{11} & \beta^{'}_{11}S^*_{10}S^{*-1}_{00}S^*_{01}\beta_{12} & \beta^{'}_{11}S^*_{10}S^{*-1}_{00}S^*_{01}\bar{\beta}^{(1)}_{\perp} \\ \beta^{'}_{12}S^*_{10}S^{*-1}_{00}S^*_{01}\beta_{11} & \beta^{'}_{12}S^*_{10}S^{*-1}_{00}S^*_{01}\beta_{12} & \beta^{'}_{12}S^*_{10}S^{*-1}_{00}S^*_{01}\bar{\beta}^{(1)}_{\perp} \\ \bar{\beta}^{(1)'}_{\perp}S^*_{10}S^{*-1}_{00}S^*_{01}\beta_{11} & \bar{\beta}^{(1)'}_{\perp}S^*_{10}S^{*-1}_{00}S^*_{01}\beta_{12} & \bar{\beta}^{(1)'}_{\perp}S^*_{10}S^{*-1}_{00}S^*_{01}\bar{\beta}^{(1)}_{\perp} \end{bmatrix} \right|. \tag{20}$$

Since $\beta^{'}_{12}X^*_t$ is assumed to be $I(1)$ the first term of (20) needs to be rescaled. Let now $\Upsilon_T = \begin{bmatrix} I_q & 0 & 0 \\ 0 & T^{-1/2}I_{r-q} & 0 \\ 0 & 0 & I_{p-r} \end{bmatrix}$ then

$$|\Upsilon^{'}_T B^{'}_T S^*(\zeta)B_T\Upsilon_T| =$$

$$\left| \begin{bmatrix} \zeta\beta_{11}^{'}S_{11}^{*}\beta_{11} & o_p(1) & o_p(1) \\ o_p(1) & \zeta T^{-1}\beta_{12}^{'}S_{11}^{*}\beta_{12}+o_p(1) & \zeta T^{-1}\beta_{12}^{'}S_{11}^{*}\bar{\beta}_{\perp}^{(1)}+o_p(1) \\ o_p(1) & \zeta T^{-1}\bar{\beta}_{\perp}^{(1)'}S_{11}^{*}\beta_{12}+o_p(1) & \zeta T^{-1}\bar{\beta}_{\perp}^{(1)'}S_{11}^{*}\bar{\beta}_{\perp}^{(1)}+o_p(1) \end{bmatrix} - \begin{bmatrix} \beta_{11}^{'}S_{10}^{*}S_{00}^{*-1}S_{01}^{*}\beta_{11} & o_p(1) & o_p(1) \\ o_p(1) & o_p(1) & o_p(1) \\ o_p(1) & o_p(1) & o_p(1) \end{bmatrix} \right|$$

$$= \begin{vmatrix} \zeta\beta_{11}^{'}S_{11}^{*}\beta_{11}-\beta_{11}^{'}S_{10}^{*}S_{00}^{*-1}S_{01}^{*}\beta_{11} & o_p(1) \\ o_p(1) & \zeta T^{-1}B^{*'}S_{11}^{*}B^{*}+o_p(1) \end{vmatrix} \tag{21}$$

where now $B^{*}=\begin{bmatrix} \beta_{12} & \bar{\beta}_{\perp}^{(1)} \end{bmatrix}$, $p^{*}\times(p-q)$.

The $o_p(1)$ blocks are blocks that were $O_p(1)$ before scaling by Υ_T because they were products of averages of products of either two $I(0)$ processes ($\beta_{11}^{'}S_{10}^{*}$, S_{00}^{*}) or an $I(0)$ and an $I(1)$ process ($B^{*'}S_{10}^{*}$, $B^{*'}S_{11}^{*}\beta_{11}$).

Next we define

$$Var\begin{bmatrix} \Delta X_t \\ \beta^{'}X_t \end{bmatrix} = \begin{bmatrix} \Sigma_{00} & \Sigma_{0\beta} \\ \Sigma_{\beta 0} & \Sigma_{\beta\beta} \end{bmatrix}.$$

In order to find the limit of (21) we need the following:

$$S_{00}^{*} \xrightarrow{p} \Sigma_{00}^{*} = H^{'}\Sigma_{00}H \tag{22}$$

$$\beta_{11}^{'}S_{10}^{*} \xrightarrow{p} \Sigma_{\beta_{11}0}^{*} = H^{'}\Sigma_{\beta 0}H \tag{23}$$

$$\beta_{11}^{'}S_{11}^{*}\beta_{11} \xrightarrow{p} \Sigma_{\beta_{11}\beta_{11}}^{*} = H^{'}\Sigma_{\beta\beta}H \tag{24}$$

and $S_{00} \xrightarrow{p} \Sigma_{00}$, $\beta^{'} S_{10} \xrightarrow{p} \Sigma_{\beta 0}$ and $\beta^{'} S_{11} \beta \xrightarrow{p} \Sigma_{\beta\beta}$ by the Weak Law of Large Numbers (see also Johansen, 1996, Lemma 10.3)). Thus,

$$|\Upsilon_T^{'} B_T^{'} S^*(\zeta) B_T \Upsilon_T| =$$

$$\begin{vmatrix} \zeta\beta_{11}^{'} S_{11}^* \beta_{11} - \beta_{11}^{'} S_{10}^* S_{00}^{*-1} S_{01}^* \beta_{11} & o_p(1) \\ o_p(1) & \zeta T^{-1} B^{*'} S_{11}^* B^* + o_p(1) \end{vmatrix} \xrightarrow{d}$$

$$= \begin{vmatrix} \zeta\Sigma_{\beta_{11}\beta_{11}}^* - \Sigma_{\beta_{11}0}^* \Sigma_{00}^{*-1} \Sigma_{0\beta_{11}}^* & 0 \\ 0 & \zeta B^{*'} C^* \int_0^1 \tilde{W}\tilde{W}' du C^{*'} B^* \end{vmatrix}$$

$$= |\zeta\Sigma_{\beta_{11}\beta_{11}}^* - \Sigma_{\beta_{11}0}^* \Sigma_{00}^{*-1} \Sigma_{0\beta_{11}}^*| |\zeta B^{*'} C^* \int_0^1 \tilde{W}\tilde{W}' du C^{*'} B^*| = 0 \qquad (25)$$

by (22)-(24) for the first factor and by (17) for the second. ■

Proof of Proposition 2.

The equations (5) and (6) have the same eigenvalues but (6) has eigenvectors $B_T^{-1}\hat{V}$ where $\hat{V} = \begin{bmatrix} \underset{p\times q}{\hat{\beta}_q} & \underset{p\times(p-q)}{\hat{V}_2} \end{bmatrix}$ is the matrix whose columns are the eigenvectors of (5) and $\hat{\beta}_q = H\hat{\beta}_{11} = \begin{bmatrix} \hat{\beta}_{11} \\ 0 \end{bmatrix}$. The eigenvalues of (6) converge to the eigenvalues of (25). Thus, the space spanned by the q first eigenvectors of (6), which correspond to the q largest eigenvalues, converges to the space spanned by vectors with zeros in the last $(p-q)$ positions. The space spanned by the first q eigenvectors of (6) is

$sp(B_T^{-1}\hat{\beta}_q) = sp(B_T^{-1}\tilde{\beta}_q)$ where $\tilde{\beta}_q = H\tilde{\beta}_{11}$ and

$$B_T^{-1}\tilde{\beta}_q = \begin{bmatrix} \bar{\beta}' \\ T^{1/2}\beta_\perp' \end{bmatrix} \tilde{\beta}_q.$$

First we analyse block (1,1). Using the formula for the partitioned inverse we have,

$$(\beta'\beta)^{-1} = \begin{bmatrix} (\beta_{11}'\beta_{11})^{-1}[I_q + \beta_{11}'\beta_{12}F\beta_{12}'\beta_{11}(\beta_{11}'\beta_{11})^{-1}] & -(\beta_{11}'\beta_{11})^{-1}\beta_{11}'\beta_{12}F \\ -F\beta_{12}'\beta_{11}(\beta_{11}'\beta_{11})^{-1} & F \end{bmatrix}$$

where

$$F = [\beta_{22}'\beta_{22} + \beta_{12}'\bar{\beta}_{11\perp}\beta_{11\perp}'\beta_{12}]^{-1}.$$

Thus,

$$(\beta'\beta)^{-1}\beta'\tilde{\beta}_q = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

where $A_1 = I_q - \bar{\beta}_{11}'\beta_{12}F\beta_{12}'\bar{\beta}_{11\perp}b_1$ and $A_2 = F\beta_{12}'\bar{\beta}_{11\perp}b_1$.

Then we analyse $\beta_\perp'\tilde{\beta}_q$ which appears in block (2,1). Partitioning $\beta_\perp'$ as in $\beta_\perp = \begin{bmatrix} \underset{(p-r)\times p^*}{\beta_\perp^{(1)'}} & \underset{(p-r)\times k}{\beta_\perp^{(2)'}} \end{bmatrix}$ we obtain

$$\beta_\perp'\tilde{\beta}_q = \begin{bmatrix} \beta_\perp^{(1)'} & \beta_\perp^{(2)'} \end{bmatrix} \begin{bmatrix} \tilde{\beta}_{11} \\ 0 \end{bmatrix} = \beta_\perp^{(1)'}\tilde{\beta}_{11} = \beta_\perp^{(1)'}\bar{\beta}_{11\perp}b_1$$

by the assumption $\beta'\beta_\perp = 0$ (or $\beta_\perp'\beta = 0$) which gives

$$\beta_\perp'\beta = \begin{bmatrix} \beta_\perp^{(1)'} & \beta_\perp^{(2)'} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} \\ 0 & \beta_{22} \end{bmatrix} = \begin{bmatrix} \beta_\perp^{(1)'}\beta_{11} & \beta_\perp^{(2)'}\beta_{12} + \beta_\perp^{(2)'}\beta_{22} \end{bmatrix} = 0$$

and therefore $\beta_{\perp}^{(1)'}\beta_{11} = 0$.

Thus,

$$B_T^{-1}\tilde{\beta}_q = \begin{bmatrix} \underset{q\times q}{I_q - \bar{\beta}_{11}^{'}\beta_{12}F\beta_{12}^{'}\bar{\beta}_{11\perp}b_1} \\ \\ \underset{(r-q)\times q}{F\beta_{12}^{'}\bar{\beta}_{11\perp}b_1} \\ \\ \underset{(p-r)\times q}{T^{1/2}\beta_{\perp}^{(1)'}\bar{\beta}_{11\perp}b_1} \end{bmatrix}. \tag{26}$$

By the form of (25) the last two blocks of (26) should converge to zero (in other words $sp(B_T^{-1}\tilde{\beta}_q)$ should converge to the space spanned by vectors with zeros in the last $(p-q)$ coordinates. A necessary condition for this is $T^{1/2}b_1 \xrightarrow{p} 0$. Then $sp(B_T^{-1}\tilde{\beta}_q) \xrightarrow{p} sp(\begin{bmatrix} I_q \\ 0 \end{bmatrix})$.

>From (13) we obtain $T^{1/2}(\tilde{\beta}_{11} - \beta_{11}) = \bar{\beta}_{11\perp}(T^{1/2}b_1) \xrightarrow{p} 0$ and that $(\tilde{\beta}_{11} - \beta_{11}) = o_p(T^{-1/2})$. ■

Proof of Proposition 3.

Using the full sample, (14) can be written as

$$\Delta X^* = \alpha_{11}\beta_{11}^{'}X_{-1}^* + \alpha_{12}Z_{-1} + \varepsilon^* \tag{27}$$

where ΔX^*, X_{-1}^*, ε^* are $p^* \times T$, Z_{-1} is $(r-q)\times T$ and they are the full sample counterparts of ΔX_t^*, X_{t-1}^*, ε_t^* and Z_{t-1} respectively.

Using the partitioned form of X_t and β,

$$\Sigma_{\beta\beta} \;=\; Var(\beta^{'}X_{t-1}) = E(\beta^{'}X_{t-1}X_{t-1}^{'}\beta) \tag{28}$$

$$= \begin{bmatrix} E(\beta_{11}^{'} X_{t-1}^{*} X_{t-1}^{*'} \beta_{11}) & E(\beta_{11}^{'} X_{t-1}^{*} Z_{t-1}^{'}) \\ E(Z_{t-1} X_{t-1}^{*'} \beta_{11}) & E(Z_{t-1} Z_{t-1}^{'}) \end{bmatrix}$$

$$\equiv \begin{bmatrix} \Sigma_{\beta_{11}\beta_{11}}^{*} & \Sigma_{\beta_{11}Z}^{*} \\ \Sigma_{Z\beta_{11}}^{*} & \Sigma_{ZZ}^{*} \end{bmatrix}$$

and the second equality follows from the fact that there are no deterministic terms in the DGP.

Since β_{11} can be estimated consistently (see Proposition 2)

$$p\text{lim}\ \tilde{\alpha}_{11} = p\text{lim}\ S_{01}^{*}\beta_{11}(\beta_{11}^{'}S_{11}^{*}\beta_{11})^{-1} = p\text{lim}\ [T^{-1}\Delta X^{*} X_{-1}^{*'}\beta_{11}(T^{-1}\beta_{11}^{'}X_{-1}^{*}X_{-1}^{*'}\beta_{11})^{-1}]$$

where the second equality is due to the absence of deterministic terms in the SM. Substituting for ΔX^{*} as it is given in (27) and using Slutsky's Theorem,

$$\begin{aligned} & p\text{lim}\ \tilde{\alpha}_{11} \\ = \ & \alpha_{11} + \alpha_{12} p\text{lim}[(T^{-1}Z_{-1}X_{-1}^{*'}\beta_{11})][p\text{lim}(T^{-1}\beta_{11}^{'}X_{-1}^{*}X_{-1}^{*'}\beta_{11})]^{-1} \\ = \ & \alpha_{11} + \alpha_{12}\Sigma_{Z\beta_{11}}^{*}\Sigma_{\beta_{11}\beta_{11}}^{*-1} \end{aligned} \tag{29}$$

and the probability limits equal the corresponding population moments since the process $\beta^{'}X_{t-1}$ (and therefore $\beta_{11}X_{t-1}^{*}$ and Z_{t-1}) is stationary and ergodic. (29) shows that $\tilde{\alpha}_{11}$ is 'inconsistent' (or asymptotically biased) unless $\alpha_{12} = 0$ or $p\text{lim}(T^{-1}Z_{-1}X_{-1}^{*'}\beta_{11}) = 0$. A stronger condition to achieve consistency is $Z_{-1}X_{-1}^{*'}\beta_{11} = 0$ i.e. Z_{-1} is orthogonal to $X_{-1}^{*'}\beta_{11}$.

For the estimator of the variance-covariance matrix of the errors (again using the consistency of $\tilde{\beta}_{11}$) we have

$$\begin{aligned} p\text{lim } \hat{\Lambda}^* &= p\text{lim } [S^*_{00} - S^*_{01}\beta_{11}(\beta'_{11}S^*_{11}\beta_{11})^{-1}\beta'_{11}S^*_{10}] \\ &= p\text{lim } (T^{-1}\Delta X^* \Delta X^{*'}) \\ &\quad - p\text{lim } [T^{-1}\Delta X^* X^{*'}_{-1}\beta_{11}(T^{-1}\beta'_{11}X^*_{-1}X^{*'}_{-1}\beta_{11})^{-1}T^{-1}\beta'_{11}X^*_{-1}\Delta X^{*'}] \\ &= p\text{lim } T^{-1}\Delta X^* M^* \Delta X^{*'} \end{aligned}$$

where $M^* = I_T - X^{*'}_{-1}\beta_{11}(\beta'_{11}X^*_{-1}X^{*'}_{-1}\beta_{11})^{-1}\beta'_{11}X^*_{-1}$. Substituting for ΔX^* using (27),

$$p\text{lim } \hat{\Lambda}^* = p\text{lim } T^{-1}[\alpha_{12}Z_{-1}M^*Z'_{-1}\alpha'_{12} + \alpha_{12}Z_{-1}M^*\varepsilon^{*'} + \varepsilon^* M^* Z'_{-1}\alpha'_{12} + \varepsilon^* M^* \varepsilon^{*'}]$$

and $M^*Z'_{-1}$ can be viewed as the residuals from the regression of Z'_{-1} on $\beta'_{11}X^*_{-1}$. By the Weak Law of Large Numbers we have

$$p\text{lim } T^{-1}Z_{-1}M^*\varepsilon^{*'} = E(Z_{-1}M^*\varepsilon^{*'}) = 0$$

since $E(Z_{-1}M^*\varepsilon^{*'}) = E[E(Z_{-1}M^*\varepsilon^{*'}|\mathcal{X}_{t-1})] = E[Z_{-1}M^*E(\varepsilon^{*'}|\mathcal{X}_{t-1})] = 0$, where $\mathcal{X}_{t-1}$ is the minimal σ-field generated by the random vector X_{t-1}. Furthermore,

$$p\text{lim } T^{-1}\beta'_{11}X^*_{-1}\varepsilon^{*'} = E(\beta'_{11}X^*_{-1}\varepsilon^{*'}) = 0$$

since $E(\beta'_{11}X^*_{-1}\varepsilon^{*'}) = E[E(\beta'_{11}X^*_{-1}\varepsilon^{*'}|\mathcal{X}_{t-1})] = E[\beta'_{11}X^*_{-1}E(\varepsilon^{*'}|\mathcal{X}_{t-1})] = 0$ (see also footnote 3). Hence,

$$\begin{aligned} p\text{lim } \hat{\Lambda}^* &= p\text{lim } (T^{-1}\varepsilon^*\varepsilon^{*'}) + p\text{lim } (T^{-1}\alpha_{12}Z_{-1}M^*Z'_{-1}\alpha'_{12}) \\ &= \Omega^* + \alpha_{12}(\Sigma^*_{ZZ} - \Sigma^*_{Z\beta_{11}}\Sigma^{*-1}_{\beta_{11}\beta_{11}}\Sigma^*_{\beta_{11}Z})\alpha'_{12} \end{aligned} \tag{30}$$

since ε^* and Z_{-1} are stationary random variables and by the Weak Law of Large Numbers the probability limits in (30) equal their corresponding population moments. Therefore, $\hat{\Lambda}^*$ is 'inconsistent' unless $\alpha_{12} = 0$. ■

References

Andrade, I. C., O'Brien, R. J. and Podivinsky, J. M. (1994). 'Cointegration tests and mean shifts', *University of Southampton Department of Economics Discussion Paper*, USDP 9405.

Boswijk, P. and Franses, P. H. (1992). 'Dynamic specification and cointegration', *Oxford Bulletin of Economics and Statistics*, Vol. 53, pp. 369-382.

Cheung, Y. W. and Lai, K. S. (1993). 'Finite-sample sizes of Johansen's likelihood ratio tests for cointegration', *Oxford Bulletin of Economics and Statistics*, Vol. 55, pp. 313-328.

DeLoach, S. B. (2001). 'More evidence in favor of the Balassa-Samuelson hypothesis', *Review of International Economics*, Vol. 9, pp. 336-342.

Doornik, J. A. (1999). *Object-Oriented Matrix Programming using Ox version 2.1*, Timberlake Consultants Ltd, London and www.nuff.ox.ac.uk/Users/Doornik.

Doornik, J. A. and Hendry, D. F. (2001). *Modelling Dynamic Systems Using PcGive*, Volume II (3rd edition), Timberlake Consultants Press, London.

Doornik, J.A., Nielsen, B. and Hendry, D.F. (1998). 'Inference in cointegrating models: UK M1 revisited', *Journal of Economic Surveys*, Vol. 12, pp. 533-572.

Ericsson, N. R., Hendry, D. F. and Mizon, G. E. (1998). 'Exogeneity, cointegration and economic policy analysis', *Journal of Business and Economic Statistics*, Vol. 16, pp. 370-387.

Hendry, D. F. and Doornik, J. A. (1994). 'Modelling linear dynamic econometric systems'. *Scottish Journal of Political Economy*, Vol. 41, pp. 1-33.

Hendry, D. F. and Mizon, G. E. (1992). 'Evaluating dynamic econometric models by encompassing the VAR'. In *Models, Methods and Applications of Econometrics: Essays in Honor of A. R. Bergstrom*, pp. 272-300, ed. P. C. B. Phillips. Blackwell, Cambridge MA.

Johansen, S. (1988). 'Statistical analysis of cointegration vectors', *Journal of Economic Dynamics and Control*, Vol. 12, pp. 231-254.

Johansen, S. (1991). 'Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models', *Econometrica,* Vol. 59, pp. 1551-1580.

Johansen, S. (1996). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, 2nd printing, Oxford University Press, Oxford.

Osterwald-Lenum, M. (1992). 'A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics', *Oxford Bulletin of Economics and Statistics*, Vol. 54, pp. 461-471.

Podivinsky, J. M. (1998). 'Testing misspecified cointegrating relationships'. *Economics Letters*, Vol 60, pp. 1-9.