On Multicointegration in the I(1) Model*

Peter C. B. Philllips[a,b,c,d]

[a]Yale University, [b]University of Auckland,

[c]University of Southampton & [d]Singapore Management University

February 10, 2019

Abstract

A semiparametric triangular systems approach is used to show how multicointegration can occur naturally in an $I(1)$ cointegrated regression model. The framework reveals the source of multicointegration as singularity of the long run error covariance matrix in an $I(1)$ system, a feature noted but not emphasized or explored in earlier work. Under such singularity, cointegrated $I(1)$ systems embody a multicointegrated structure and may be analyzed and estimated without appealing to the associated $I(2)$ system but with consequential asymptotic properties that introduce asymptotic bias into most conventional methods of cointegrating regression. Under certain conditions, trend IV estimation with deterministic orthonormal instruments (Phillips, 2014) is shown to provide efficient estimation with mixed normal limit theory and pivotal inference in singular multicointegrated systems in addition to standard cointegrated $I(1)$ systems. Some implications for robust inference are discussed.

Keywords: Cointegration, Integration, Long run variance matrix, Multicointegration, Singularity, Trend IV estimation.

JEL Codes: C12,C13,C22

1 Introduction

Economic identities that link some variables to partial sums of constituent variables arise frequently in economic data. Examples include common relations between stock and flow versions of variables such as the capital stock and fixed investment, inventory investment and inventory stock, housing construction completions and housing units under construction. Many of these variables

*This paper has its origins in a 2011 Yale Take Home Examination. Phillips acknowledges research support from a Kelly Fellowship at the University of Auckland.

have nonstationary characteristics and cointegration models have proved a frequently used framework for empirical work investigating such time series.

The concept of multicointegration was introduced by Granger and Lee (1989, 1990) to allow explicitly for linkages among stock and flow forms of integrated order one ($I(1)$) variables. In particular, the orginal concept of multicointegration was suggested to capture the notion that equilibrium errors in an $I(1)$ cointegrating relation may accumulate so that they cointegrate with the original variables. Engsted and Haldrup (1999) remark that this phenomenon is likely to occur in practice when characterizing the dynamic interactions of stock and flow variables. Granger and Lee (1990) and Lee (1996) showed how multicointegration can arise in the context of optimum control problems and infinite horizon quadratic adjustment cost models.

In these models the equilibrium errors (or residuals in the $I(1)$ cointegrating relation) are considered $I(0)$ or stationary, so that upon cumulation these errors become I(1), and then subsequent cointegration may occur with the original variables or partial sums of them. Somewhat naturally it has therefore been posited in the multicointegration literature that the following statements hold:

1. "Engsted and Johansen (1999) show that when variables are multicointegrated the requirements for the system to be an $I(1)$ system will fail; in fact, an $I(1)$ specification will be misspecifed even though the main interest lies in the analysis of the $I(1)$ series. Instead the system should be formulated as an $I(2)$ model where multicointegration can be shown to result in cointegration amongst generated $I(2)$ variables and their first differences" (Engsted and Haldrup, 1999, p.237)

2. "If the process is given by the cointegrated VAR model for $I(1)$ variables, then multicointegration cannot occur" (Engsted and Johansen, 1999).

These ideas seem natural in the stock and flow framework and appear to have been universally accepted in the literature. But they were developed in a VAR framework and do not necessarily hold in more general models, including semiparametric $I(1)$ cointegrating settings such as the triangular system of Phillips (1991). In fact, as demonstrated in the present paper, multicointegration occurs naturally in a cointegrated $I(1)$ model whenever there is a rank deficiency in the long run error covariance matrix. The phenomenon is a general one and rank deficiency turns out to be the determining factor of multicointegration in an $I(1)$ system. Multicointegration arises because singularity in the long run error covariance matrix induces a further long run cointegrating relation simply because the singularity implies a moving average $I(-1)$ component in a certain direction in the system error, which leads directly to cointegration upon accumulation. The phenomenon has an analogue reduced rank structure in the parametric VAR model context and was noted but not further analyzed by Engsted and Haldrup (1999, p.241).

The masterful treatment of reduced rank VAR systems by Johansen (1992, 1995) provides explicit representations of the reduced rank structures which ensure the existence of cointegrated $I(1)$ and $I(2)$ VAR systems. The implications of these conditions for characterizing systems with multicointegration are employed in Engsted and Johansen (1999), which demonstrates the relevance of the $I(2)$ system for embodying multicointegrated structures in VAR systems. Outside the VAR setting, multicointegration can exist in an $I(1)$ reduced rank VARMA setting or in $I(1)$ cointegrated systems with infinite order bidirectional lags. These models and approaches to multicointegration are reconciled in what follows.

The present paper makes two main contributions. First, a general analysis of multicointegration is provided within an $I(1)$ cointegrated system using the semiparametric triangular model framework. Multicointegration in such systems depends on singularity in the long run error covariance matrix, which in turn is shown to affect the asymptotic behavior of most cointegrated system estimation procedures by introducing bias and non-pivotal inference. Second, it is shown that under certain conditions trend IV estimation with deterministic orthonormal instruments (Phillips, 2014) provides efficient estimation with mixed normal limit theory and pivotal inference in singular multicointegrated systems, as well as standard cointegrated $I(1)$ systems, thereby providing an approach to estimation and inference in $I(1)$ cointegrated systems that is robust to multicointegration under these conditions.

2 Multicointegration in the $I(1)$ framework

The starting point in developing a framework for the source of multicointegation is the following $I(1)$ triangular matrix system of cointegration (Phillips, 1991)

$$y_t = Ax_t + u_{0t} \tag{1}$$
$$x_t = x_{t-1} + u_{xt}, \quad t = 1, \ldots, T. \tag{2}$$

Here A is an $m_0 \times m_x$ cointegrating coefficient matrix, the $I(1)$ m_x-vector x_t is initialized at $t = 0$ by $x_0 = O_p(1)$, and the composite error vector $u_t = (u_{0t}', u_{0xt}')'$ is assumed to follow the linear process

$$u_t = D(L)\eta_t = \sum_{j=0}^{\infty} D_j \eta_{t-j}, \quad \text{with } \sum_{j=0}^{\infty} j||D_j|| < \infty, \quad \eta_t \sim iid(0, I_m), \tag{3}$$

where $m = m_0 + m_x$. Let $\Gamma_h = \mathbb{E}u_t u_{t+h}'$ and $\mathbb{V}^{\mathrm{LR}}(u_t) = \sum_{h=-\infty}^{\infty} \Gamma_h$ denote the long run variance matrix of u_t. The linear operator $D(L)$ and long run variance matrix $\Omega = \sum_{h=-\infty}^{\infty} \Gamma_h = D(1)D(1)' = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} D_j D_k'$ and one sided long run covariance matrix $\Delta = \sum_{h=-\infty}^{0} \mathbb{E}(u_h u_0')$ of u_t are partitioned conformably with u_t as

$$D(L) = \begin{bmatrix} D_{00}(L) & D_{0x}(L) \\ D_{x0}(L) & D_{xx}(L) \end{bmatrix}, \ \Omega = \begin{bmatrix} \Omega_{00} & \Omega_{0x} \\ \Omega_{x0} & \Omega_{xx} \end{bmatrix}, \ \Delta = \begin{bmatrix} \Delta_{00} & \Delta_{0x} \\ \Delta_{x0} & \Delta_{xx} \end{bmatrix} \tag{4}$$

where $\Omega_{xx} > 0$ is positive definite, ensuring that x_t is a full rank unit root vector process which delivers m_x common stochastic trends to the $I(1)$ system

(1)-(2). The conditional long run variance matrix $\Omega_{00,x} = \Omega_{00} - \Omega_{0x}\Omega_{xx}^{-1}\Omega_{x0}$ is the Schur complement of the block Ω_{xx} in Ω and this matrix is positive (semi-) definite if and only if Ω is positive (semi-) definite by virtue of the Guttman rank additivity formula $\operatorname{rank}(\Omega) = \operatorname{rank}(\Omega_{xx}) + \operatorname{rank}(\Omega_{00,x})$).

The case of nonsingular Ω is well studied. The situation where Ω is singular and, in particular, where the conditional long run variance matrix $\Omega_{00,x}$ is singular seems largely to have been ignored[1] in the now vast literature on cointegration and none of the implications of singularity for estimation and inference have been explored in the (1)-(2) setting. This neglect is partly because, as we will show, singularity in the long run error covariance matrix leads to an $I(1)$ reduced rank VARMA representation rather than a reduced rank $I(1)$ VAR representation. So while such systems fall naturally within the semiparametric framework above, they do not fall so neatly within the VAR framework, at least without raising the order of the system to $I(2)$. Nonetheless, the singular long run variance matrix case is especially interesting because it leads directly to a situation where partial sums of the observed variables y_t and x_t (which then become $I(2)$ variables) are cointegrated with x_t in some unknown direction - see (9) below. The importance of this situation is that it provides a primitive (that is, within the $I(1)$ system) link to the phenomenon of multicointegration as envisaged in special cases by Granger and Lee (1989). But the source of the multicointegration is now firmly evident in the $I(1)$ framework (1)-(2).

The multicointegration model is well known to be empirically important in cases involving variables such as production, sales and inventories (Granger and Lee, 1990) or housing completions, starts, and construction (Lee, 1992), where aggregation plays a critical role in relating key variables of economic interest. More recent applications of multicointegration involve issues of fiscal sustainability (Berenguer-Rico and Carrion-i-Silvestre, 2011; Escario et al., 2012). The present formulation is a general semiparametric one in the sense that the short run dynamics are left completely unspecified beyond the linear process framework (3) and both cointegrating and multicointegrating relations are parameterized with unknown coefficients rather than through the special case of identities, stock-flow relationships, or posited behavioral relations with known coefficients.

The time series $u_{0t-1} = y_{t-1} - Ax_{t-1}$ is the lagged equilibrium error and the system (1)-(2) may therefore be written in the following error correction model (ECM) form (Phillips, 1991)

$$\begin{aligned}
\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} &= \begin{bmatrix} -u_{0t-1} \\ 0 \end{bmatrix} + \begin{bmatrix} A\Delta x_t + u_{0t} \\ u_{xt} \end{bmatrix} \\
&= \begin{bmatrix} -I_{m_0} \\ 0 \end{bmatrix} [I_{m_0}, -A] \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} A\Delta x_t + u_{0t} \\ u_{xt} \end{bmatrix} \\
&=: \alpha\beta' \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} A\Delta x_t + u_{0t} \\ u_{xt} \end{bmatrix} \qquad (5)
\end{aligned}$$

[1] The possibility is mentioned by Engsted and Haldrup (1999, p.241) but is not analyzed.

or, setting $z_t = (y_t', x_t')'$, as

$$\Delta z_t = \alpha\beta' z_{t-1} + u_{zt}, \tag{6}$$

where

$$\alpha = \begin{bmatrix} -I_{m_0} \\ 0 \end{bmatrix}, \quad \beta' = [I_{m_0}, -A], \quad u_{zt} = \begin{bmatrix} A\Delta x_t + u_{0t} \\ u_{xt} \end{bmatrix}, \tag{7}$$

with the $m \times m_0$ loading coefficient matrix α and the $(m_0 \times m)$ cointegrating matrix β'. The vector $\beta' z_{t-1} = y_{t-1} - Ax_{t-1} = u_{0t-1}$ is just the lagged equilibrium error term from (1). The ECM error vector u_{zt} in (6) is serially dependent and follows a general linear process induced by u_t and the mechanism (3).

An alternate representation of (1) which is useful in the development of efficient estimation methods of $I(1)$ cointegrated systems by fully modified least squares (Phillips and Hansen, 1990) or trend IV regression (Phillips, 2014) is the augmented regression equation

$$\begin{aligned} y_t &= Ax_t + \Omega_{0x}\Omega_{xx}^{-1}\Delta x_t + u_{0.xt}, \quad u_{0.xt} := u_{0t} - \Omega_{0x}\Omega_{xx}^{-1}u_{xt} \\ &=: Ax_t + F\Delta x_t + u_{0.xt}, \quad \Delta x_t = u_{xt}, \end{aligned} \tag{8}$$

where both the cointegrating coefficient matrix A and the long run regression coefficient $F = \Omega_{0x}\Omega_{xx}^{-1}$ are treated as unknown. Importantly, the long run regression coefficient matrix F is nonparametric. Applying partial sum operations to (8) gives

$$Y_t = AX_t + F(x_t - x_0) + U_{0.xt}, \tag{9}$$

with $Y_t = \sum_{s=1}^t y_s, X_t = \sum_{s=1}^t x_s$, and $U_{0.xt} = \sum_{s=1}^t u_{0.xs}$. Now suppose that the long run (conditional) variance matrix $\Omega_{00.x}$ of $u_{0.xt}$ is singular of rank $0 < p < m_0$ and H is an $m \times p$ matrix of full rank p spanning the null space of $\Omega_{00.x}$, so that

$$H'\Omega_{00.x}H = 0. \tag{10}$$

Then in this direction the transformed error $H'u_{0.xt}$ has zero long run variance matrix and zero spectral density matrix at the origin. There therefore exists some p dimensional $I(0)$ process ε_{Ht} for which $H'u_{0.xt} = \Delta\varepsilon_{Ht}$ *a.s.*, leading to the representation

$$H'y_t = H'Ax_t + H'F\Delta x_t + \Delta\varepsilon_{Ht},$$

and by partial summation to

$$H'Y_t = H'AX_t + H'F(x_t - x_0) + (\varepsilon_{Ht} - \varepsilon_{H0}).$$

It follows that

$$H'Y_t = H'AX_t + H'Fx_t + (\varepsilon_{Ht} - \varepsilon_{H0} - H'Fx_0) =: H'AX_t + H'Fx_t + \eta_{Ht}, \tag{11}$$

where $\eta_{Ht} = \varepsilon_{Ht} - \varepsilon_{H0} - H'Fx_0$ is $I(0)$ up to (and conditional on) the initial condition $x_0 = O_p(1)$. From (11) it follows that the variables (Y_t, X_t, x_t) are cointegrated, involving both the $I(2)$ time series (Y_t, X_t) and the $I(1)$ time series x_t. This accords with the conventional definition of multicointegration.

Now define the partial sum process $Z_t = \sum_{s=1}^{t} z_s = (Y_t', X_t')'$ and note that $z_t = (y_t', x_t')'$ is an $I(1)$ process whose common stochastic trends are embodied in x_t. In the notation of Engsted and Johansen (1999), the linear combination $\tau' z_t := [I_{m_0}, -A] z_t = u_{0t}$ is $I(0)$ and the cumulated process $\tau' Z_t = \sum_{s=1}^{t} \tau' z_s$ cointegrates with x_t in the sense that there exist matrices $\rho' = H_0'$ and $\psi' = -H_0' F$ (again using the notation of Engsted and Johansen) such that

$$\rho' \sum_{s=1}^{t} \tau' z_s + \psi' x_t = H_0' [I_{m_0}, -A] Z_t - H_0' F x_t = H_0' Y_t - H_0' A X_t - H_0' F x_t = \eta_{Ht} \equiv I(0). \tag{12}$$

The m dimensional $I(1)$ process z_t is therefore multicointegrated in the sense of Granger and Lee (1989) and Engsted and Johansen (1999). This general case of multicointegration therefore appears to contradict the claims made in #1 and #2 above that "multicointegration cannot take place in the error correction model for $I(1)$ variables." Of course, neither the $I(1)$ ECM (6) nor the $I(1)$ augmented regression model (8) is specified in a VAR form. It turns out that requiring an $I(1)$ VAR specification is a binding restriction that eliminates multicointegrated $I(1)$ systems in VAR format. In this sense the semiparametric setting is materially more general because it admits multicointegrated $I(1)$ versions simply as a property of the error process in the formulation.

The reason is straightforward and is explained by writing the ECM (6) as follows

$$\begin{aligned}
\Delta z_t &= \alpha\beta' z_{t-1} + \begin{bmatrix} A\Delta x_t + u_{0t} \\ u_{xt} \end{bmatrix} = \alpha\beta' z_{t-1} + \begin{bmatrix} A\Delta x_t + F\Delta x_t + u_{0.xt} \\ u_{xt} \end{bmatrix} \\
&= \alpha\beta' z_{t-1} + \begin{bmatrix} A+F \\ 0 \end{bmatrix} \Delta x_t + \begin{bmatrix} u_{0.xt} \\ u_{xt} \end{bmatrix} \\
&= \alpha\beta' z_{t-1} + \begin{bmatrix} A+F \\ 0 \end{bmatrix} \Delta x_{t-1} + \begin{bmatrix} u_{0.xt} \\ u_{xt} \end{bmatrix} + \begin{bmatrix} A+F \\ 0 \end{bmatrix} \Delta u_{xt} \\
&= \alpha\beta' z_{t-1} + \begin{bmatrix} 0 & A+F \\ 0 & 0 \end{bmatrix} \Delta z_{t-1} + \begin{bmatrix} u_{0.xt} \\ u_{xt} \end{bmatrix} + \begin{bmatrix} 0 & A+F \\ 0 & 0 \end{bmatrix} \Delta u_t.
\end{aligned} \tag{13}$$

Now the long run variance matrix of the error vector in (13)

$$\begin{bmatrix} u_{0.xt} \\ u_{xt} \end{bmatrix} + \begin{bmatrix} 0 & A+F \\ 0 & 0 \end{bmatrix} \Delta u_t \tag{14}$$

is the same as the long run variance matrix of the first member of (14) and is therefore

$$\begin{bmatrix} \Omega_{00.x} & 0 \\ 0 & \Omega_{xx} \end{bmatrix},$$

which is singular. So the system (13) is a reduced rank regression but has non-invertible moving average error components ($H' u_{0.xt} = \Delta\varepsilon_{Ht}$ and Δu_t)

and cannot therefore be written in standard reduced rank $I(1)$ VAR form with lagged regressors and martingale difference errors. However, it can be viewed as a reduced rank $I(1)$ VARMA model with MA unit roots; and taking partial sums of (13), subject to initial conditions, leads to a reduced rank $I(2)$ system

$$\Delta Z_t = \alpha\beta' Z_{t-1} + \begin{bmatrix} 0 & A+F \\ 0 & 0 \end{bmatrix} z_{t-1} + \begin{bmatrix} U_{0.xt} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} A+F \\ I \end{bmatrix} u_{xt}, \qquad (15)$$

which provides a reduced rank linear combination of the $I\,(2)$ vector Z_t, the $I\,(1)$ vectors $z_t, x_t, U_{0.xt} = \sum_{s=1}^{t} u_{0.xs}$, and the stationary vector u_{xt}. Thus, there is an $I\,(2)$ multicointegrated system with weakly dependent errors corresponding to the $I\,(1)$ multicointegrated system (6), matching the reasoning that leads to the $I\,(2)$ system in Engsted and Johansen (1999). Note that the lower block of (15) is an identity and the error vector u_{xt} in (15) therefore has lower dimension but has nonsingular long run variance matrix Ω_{xx}. The process $U_{0.xt} = \beta' Z_t - F x_t$, on the other hand, is not full rank $I\,(1)$ and therefore can be expected to affect inference, just as it does in the $I\,(1)$ system (8).

3 Reconciliation with the VAR

3.1 Cointegrating relations and the moving average representation

It is helpful to reconcile the above discussion with the analysis of multicointegration given in Engsted and Johansen (1999). Start by writing the triangular system (1)-(2) as

$$\begin{bmatrix} I_{m_0} & -A \\ 0 & \Delta I_{m_x} \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} = u_t = D(L)\eta_t.$$

Formally solving this system yields the 'reduced form'

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} I_{m_0} & -A \\ 0 & \Delta I_{m_x} \end{bmatrix}^{-1} D(L)\eta_t = \begin{bmatrix} I_{m_0} & \Delta^{-1} A \\ 0 & \Delta^{-1} I_{m_x} \end{bmatrix} D(L)\eta_t,$$

and factoring Δ^{-1} leads to

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} \Delta I_{m_0} & A \\ 0 & I_{m_x} \end{bmatrix} D(L)\eta_t := (1-L)^{-1} C(L)\eta_t \qquad (16)$$

where

$$C(L) = \begin{bmatrix} \Delta I_{m_0} & A \\ 0 & I_{m_x} \end{bmatrix} \begin{bmatrix} D_{00}(L) & D_{0x}(L) \\ D_{x0}(L) & D_{xx}(L) \end{bmatrix}.$$

System (16) may be interpreted as the usual moving average Wold representation $\Delta z_t = C(L)\eta_t$. In this system Engsted and Johansen (1999) assume that the roots of $|C(z)| = 0$ are either bounded away from unity or $z = 1$. Observe that the matrix

$$C(1) = \begin{bmatrix} 0 & A \\ 0 & I_{m_x} \end{bmatrix} \begin{bmatrix} D_{00}(1) & D_{0x}(1) \\ D_{x0}(1) & D_{xx}(1) \end{bmatrix} = \begin{bmatrix} A \\ I \end{bmatrix} \begin{bmatrix} D_{x0}(1) & D_{xx}(1) \end{bmatrix} =: \xi\epsilon'$$

has reduced rank, as expected in a standard cointegrated $I(1)$system, with cointegrating matrix given by the orthogonal complement $\xi'_\perp = \begin{bmatrix} I_{m_0} & -A \end{bmatrix}$ (in the more common reduced rank notation (6) we have $\xi'_\perp = \beta'$) so that $\xi'_\perp C(1) = 0$. The matrix $\epsilon' = \begin{bmatrix} D_{x0}(1) & D_{xx}(1) \end{bmatrix}$ has full rank m_x and, by reordering of coordinates as may be needed, we can assume $D_{xx}(1)$ to be nonsingular[2]. An orthogonal complement matrix of ϵ may then be constructed as

$$\epsilon_\perp = \begin{bmatrix} D_{x0}(1)' \\ D_{xx}(1)' \end{bmatrix}_\perp = \begin{bmatrix} I_{m_0} \\ -D_{xx}(1)^{-1}D_{x0}(1) \end{bmatrix}.$$

Following the algebraic approach in Engsted and Johansen (1999), the derivative matrix of $C(z)$ is

$$\dot{C}(z) = \begin{bmatrix} -I_{m_0} & 0 \\ 0 & 0 \end{bmatrix} D(z) + \begin{bmatrix} (1-z)\, I_{m_0} & A \\ 0 & I_{m_x} \end{bmatrix} \dot{D}(z),$$

so that

$$\begin{aligned} \xi'_\perp \dot{C}(1)\epsilon_\perp = -\xi'_\perp & \begin{bmatrix} D_{00}(1) & D_{0x}(1) \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_{m_0} \\ -D_{xx}(1)^{-1}D_{x0}(1) \end{bmatrix} \\ &= -\begin{bmatrix} D_{00}(1) & D_{0x}(1) \end{bmatrix} \begin{bmatrix} I_{m_0} \\ -D_{xx}(1)^{-1}D_{x0}(1) \end{bmatrix} \\ &= -\left\{ D_{00}(1) - D_{0x}(1)\, D_{xx}(1)^{-1}D_{x0}(1) \right\}. \end{aligned}$$

The matrix $D_{00}(1) - D_{0x}(1)\, D_{xx}(1)^{-1}D_{x0}(1)$ is the Shur complement of the block $D_{xx}(1)$ in $D(1)$ and is singular if and only if the matrix

$$D(1) = \begin{bmatrix} D_{00}(1) & D_{0x}(1) \\ D_{x0}(1) & D_{xx}(1) \end{bmatrix}$$

is singular since by the Schur determinantal formula

$$|D(1)| = \left| D_{00}(1) - D_{0x}(1)\, D_{xx}(1)^{-1}D_{x0}(1) \right| |D_{xx}(1)| = 0$$

if and only if $\left| D_{00}(1) - D_{0x}(1)\, D_{xx}(1)^{-1}D_{x0}(1) \right| = 0$ because $|D_{xx}(1)| \neq 0$ by construction. But the long run error variance matrix in (1)-(2) is $\Omega = D(1)\, D(1)'$. It follows that the matrix $\xi'_\perp \dot{C}(1)\epsilon_\perp$ has reduced rank if and only if the long run error variance matrix Ω is singular. Hence, the criterion given in Engsted and Johansen (1999) for multicointegration in an $I(2)$ system (that

[2] Since $\Omega_{xx} = D_{x0}(1)\, D_{x0}(1)' + D_{xx}(1)\, D_{xx}(1)'$ is positive definite, the matrix $[D_{x0}(1), D_{xx}(1)]$ has full row rank m_x and the columns (coordinates) may be rearranged as needed to ensure that the $m_x \times m_x$ matrix $D_{xx}(1)$ is nonsingular.

$\xi_{\perp}'\dot{C}(1)\epsilon_{\perp}$ has reduced rank) reduces to the multicointegration criterion in an $I(1)$ system given here – namely that the long run error covariance matrix in that system (here the triangular system given by (1)-(2)) is singular. Importantly, however, the algebraic analysis in Engsted and Johansen (1999) restricts attention to autoregressive formulations of cointegrated I(1) systems and in doing so eliminates cointegrated I(1) models such as (1)-(2) with singular long run error variance matrices[3].

3.2 Parametric augmented regression

The augmented regression (8) provides another mechanism for reconciling the existence of multicointegration without specifying an $I(2)$ system. In fact, (8) may be converted into an equivalent augmented parametric system of distributed lags as follows. We begin by noting that

$$y_t = Ax_t + u_{0t} = Ax_t + \sum_{k=-\infty}^{\infty} G_k \Delta x_{t+k} + u_{0.xt} \tag{17}$$

The latter equation arises from the well known relation (Saikkonen, 1991)

$$u_{0t} = \sum_{k=-\infty}^{\infty} G_k u_{xt+k} + u_{0.xt} = \sum_{k=-\infty}^{\infty} G_k \Delta x_{t+k} + u_{0.xt}$$

which explicitly relates the regression errors u_{0t} and $u_{0.xt}$ in terms of leads and lags of the errors u_{xt} so that the orthogonality

$$\mathbb{E}\left(u_{xt+k}u_{0.xt}'\right) = 0, \quad k = 0, \pm 1, \pm 2, ...$$

holds and the long run variance matrix of $(u_{0.xt}, u_{xt})$ is the block diagonal matrix $\mathrm{diag}\left[\Omega_{00.x}, \Omega_{xx}\right]$. In this formulation we have the long run regression coefficient equivalence $\sum_{k=-\infty}^{\infty} G_k = \Omega_{0x}\Omega_{xx}^{-1}$, leading to (8) as is now demonstrated.

In particular, using the BN decomposition under the summability conditions in Phillips and Solo (1992) that are satisfied by (3) we have

$$u_t = D(L)\eta_t = D\left(1\right)\eta_t + \tilde{\eta}_{t-1} - \tilde{\eta}_t = D\left(1\right)\eta_t - \Delta\tilde{\eta}_t,$$

where $\tilde{\eta}_t = \sum_{j=0}^{\infty} \tilde{D}_j \eta_{t-j}$ and $\tilde{D}_j = \sum_{k=j+1}^{\infty} D_k$. The long run variance matrix Ω is partitioned conformably with $D\left(1\right) = \left[D_0\left(1\right)' \vdots D_x\left(1\right)'\right]'$ as

$$\Omega = D\left(1\right)D\left(1\right)' = \begin{bmatrix} D_0\left(1\right)D_0\left(1\right)' & D_0\left(1\right)D_x\left(1\right)' \\ D_x\left(1\right)D_0\left(1\right)' & D_x\left(1\right)D_x\left(1\right)' \end{bmatrix} = \begin{bmatrix} \Omega_{00} & \Omega_{0x} \\ \Omega_{x0} & \Omega_{xx} \end{bmatrix},$$

[3] In particular, Engsted and Johansen (1999) show that multicointegration cannot appear in a cointegrated $I(1)$ autoregressive model because a requisite condition for the autoregressive representation is that $\left|\xi_{\perp}'\dot{C}(1)\epsilon_{\perp}\right| \neq 0$ or $\xi_{\perp}'\dot{C}(1)\epsilon_{\perp}$ to have full rank – see their conditions (1) and (4).

and then

$$F = \Omega_{0x}\Omega_{xx}^{-1} = D_0(1)\,D_x(1)'\left(D_x(1)\,D_x(1)'\right)^{-1} = \sum_{j=-\infty}^{\infty} G_j = G(1)$$

because we have

$$\sum_{j=-\infty}^{\infty} G_j \Delta x_{t+j} = \sum_{j=-\infty}^{\infty} G_j u_{xt+j} = \left(\sum_{k=-\infty}^{\infty} G_j\right) u_{xt} + \tilde{u}_{xt-1} - \tilde{u}_{xt} = G(1)\,u_{xt} - \Delta\tilde{u}_{xt},$$

where $\tilde{u}_{xt} = \tilde{G}(L)\,u_{xt}$ with $\tilde{G}(L) = \sum_{k=j+1}^{\infty} G_k \mathbf{1}\{j \geq 0\} - \sum_{k=-\infty}^{j} G_k \mathbf{1}\{j < 0\}$ using the two-sided version of the Phillips and Solo (1992) BN decomposition (see Corbae et al, 2002; Lemma D).

It now follows that

$$\begin{aligned} y_t &= Ax_t + \sum_{k=-\infty}^{\infty} G_k \Delta x_{t+k} + u_{0.xt} & (18) \\ &= Ax_t + G(1)\,\Delta x_t + u_{0.xt} - \Delta\tilde{u}_{xt}, & \\ &= Ax_t + G(1)\,\Delta x_t + u_{0.xt}^{+}, \quad \text{with } u_{0.xt}^{+} = u_{0.xt} - \Delta\tilde{u}_{xt} & (19) \end{aligned}$$

for which we have the long run variance matrix equivalence

$$\mathbb{V}^{\mathrm{LR}}\left(u_{0.xt}^{+}\right) = \mathbb{V}^{\mathrm{LR}}\left(u_{0.xt}\right) = \Omega_{00} - \Omega_{0x}\Omega_{xx}^{-1}\Omega_{x0},$$

because $\Delta\tilde{u}_{xt}$ has zero long run variance matrix and zero long run covariance with $u_{0.xt}$. This equivalence confirms that the models (18) and (19) are long-run equivalent in the sense that the difference between them has zero long run covariance matrix. Thus, the multicointegrated augmented regression system (8) has an analogue (18) in the parametric model context but requires modeling with infinite order bidirectional lags. In both cases, asymptotic theory and inferential methods need to take account of the singularity of $\Omega_{00.x}$.

4 Estimation and Inference

With the exception of certain specialized models involving known relationships between variables such as stocks and flows, the existence of multicointegration will often not be anticipated in practical applied work on estimation and inference in $I(1)$ cointegrated systems. Tests for the presence of multicointegration have been developed (Engsted et al., 1997) but multicointegration may not be suspected and pre-test analyses may therefore not be conducted or lead to misleading outcomes. In the absence of such tests it is obviously useful to have methods of estimating $I(1)$ cointegrated systems that are robust to the presence of multicointegration.

Since semiparametric formulations of cointegrated I(1) systems may be conducted in the presence of multicointegration, standard efficient methods of estimating of such systems such as FM-OLS (Phillips and Hansen, 1990; Phillips, 1995), trend IV regression (Phillips, 2014), or dynamic OLS (Saikkonen, 1991; Phillips and Loretan, 1991; Stock and Watson, 1993) may continue to be employed in practical work. However, the properties of such regressions are influenced by the singularity of the long run error covariance matrix. The typical impact of singularity is to raise the rate of convergence in the direction of singularity, thereby producing a degenerate limit theory for the estimate of the full cointegrating matrix. Moreover, common semiparametric methods of estimation such as FM-OLS involve the use of nonparametric kernel estimates of the long run variance and covariance matrices for bias correction and inference. In consequence, the accelerated rate of convergence in FM-OLS estimation is affected by the asymptotic behavior of these kernel estimates under rank degeneracy, as in the analysis of Phillips (1995). Inference is correspondingly affected with further nonstandard limit distribution complications and non-pivotal limit theory in test statistics. These consequences are analyzed in an Online Supplement.

In the present paper, we proceed to examine two straightforward approaches to estimation and inference. To keep the analysis brief we confine attention to a scalar cointegrating relationship, which enables a convenient introduction of the basic ideas, highlights the main implications, and covers one of the most common cases arising in practice.

4.1 Estimation Approaches

To fix ideas, consider the following scalar version of the augmented $I(1)$ cointegrating equation (8)

$$y_t = a'x_t + f'\Delta x_t + u_{0.xt}, \quad \Delta x_t = u_{xt}, \; u_{0.xt} = u_{0t} - \Omega_{0x}\Omega_{xx}^{-1}u_{xt} \tag{20}$$

where the conditional long run variance is now zero, viz., $\Omega_{00.x} = \Omega_{00} - \Omega_{0x}\Omega_{xx}^{-1}\Omega_{x0} = 0$. In this event, write $u_{0.xt} = \Delta e_t$ where e_t has variance σ_e^2 and long run variance $\omega_e^2 > 0$. The latter positivity condition is not necessary but its relaxation leads to further interesting complications on which we will comment later. In what follows, we consider three methods of estimation of the parameters in (20).

The analysis that follows considers both singular and non-singular $\Omega_{00.x}$ cases. We start by requiring the following high-level conditions which hold under well-known conditions (e.g., Phillips and Solo, 1992):

$$(a) \quad \frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor n\cdot \rfloor} e_t \quad \rightsquigarrow \quad B_e(\cdot) \equiv BM\left(\omega_e^2\right), \quad \text{when } \Omega_{00.x} = 0 \tag{21}$$

$$(b) \quad \frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor n\cdot \rfloor} u_{0.xt} \quad \rightsquigarrow \quad B_{0.x}(\cdot) = BM\left(\Omega_{00.x}\right), \quad \text{when } \Omega_{00.x} > 0. \tag{22}$$

In case (a) we further assume the joint functional law

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor n\cdot\rfloor}\left(e_t, u_{xt}'\right)' \rightsquigarrow \left(B_e\left(\cdot\right), B_x\left(\cdot\right)'\right)' \equiv BM\left(\left[\begin{array}{cc}\omega_e^2 & \omega_{ex}\\ \omega_{xe} & \Omega_{xx}\end{array}\right]\right) \tag{23}$$

$$\text{with } \left[\begin{array}{cc}\omega_e^2 & \omega_{ex}\\ \omega_{xe} & \Omega_{xx}\end{array}\right] > 0 \text{ and } \Delta_{xe} = \sum_{h=0}^{\infty} E\left(u_{x0}e_h\right). \tag{24}$$

The functional law (22) already holds under (3), and (23) similarly holds under analogous linear process conditions, as in Phillips and Solo (1992). Although $u_{0.xt} = \Delta e_t$ has zero long run covariance with u_{xt} in case (a) the same is not necessarily so of e_t. For instance, if $e_t = \alpha' u_{xt} + \varepsilon_t$ where $\varepsilon_t \sim iid\left(0, \sigma_\varepsilon^2\right)$ and independent of u_{xt}, then $u_{0.xt} = \alpha'\Delta u_{xt} + \Delta\varepsilon_t$ has zero long run covariance with u_{xt} but the long run covariance of u_{xt} and e_t is $\mathbb{CV}^{\mathrm{LR}}\left(u_{xt}, e_t\right) = \omega_{xe} = \Omega_{xx}\alpha \neq 0$. Condition (23) allows for both possibilities.

The first method of estimation that we consider raises the integration order of the system by partial summation of (20), a process that can be performed whether or not $\Omega_{00.x} = 0$. But singularity obviously affects limit behavior, as demonstrated below. This method is always available and has been considered in other work, including predictive regression cases (Phillips and Lee, 2013), and, of course, aggregated VAR representations and ECM systems such as (15) above. Vogelsang and Wagner (2014) recently proposed a version of this procedure for estimating $I(1)$ systems under the condition $\Omega_{00.x} > 0$ and called the method IM-OLS (integrated modified least squares). The estimation method has an advantage over FM-OLS in that it does not require estimation of long run variance matrices and avoids use of kernels and bandwidth choices, but it is asymptotically inefficient relative to FM-OLS and other efficient methods of $I(1)$ system estimation.

Using capitals as before to denote partial summation (e.g., $Y_t = \sum_{s=1}^{t} y_s$), write the transformed system (20), up to initial conditions, as

$$Y_t = a'X_t + f'x_t + e_t^+ \tag{25}$$

$$e_t^+ = e_t\mathbf{1}\left\{\Omega_{00.x} = 0\right\} + U_{0.xt}\mathbf{1}\left\{\Omega_{00.x} > 0\right\} \tag{26}$$

covering the two cases. Applying least squares regression to (25) gives

$$\hat{a} - a = \left(X'Q_xX\right)^{-1} X'Q_xe^+, \;\; \hat{f} - f = \left(x'Q_Xx\right)^{-1} x'Q_Xe^+ \tag{27}$$

in usual partitioned matrix regression notation. Standard weak convergence methods for nonstationary regression (Phillips, 1986, 1988) lead to the following asymptotics as $n \to \infty$, where we focus on estimation of the cointegrating vector a. The result given in (ii) for the case $\Omega_{00.x} > 0$ corresponds to the finding in Vogelsang and Wagner (2014, theorem 2).

Theorem 1

(a) *When* $\Omega_{00.x} = 0$ *and (23) holds*

(i) $n^2(\hat{a}-a) \rightsquigarrow \left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1} \left\{\int_0^1 \tilde{B}_X dB_e - \left(\int_0^1 \breve{B}_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1} \Delta_{xe}\right\}$,

where $\tilde{B}_X(r) = \breve{B}_X(r) - \int_0^1 \breve{B}_X B_x \left(\int_0^1 B_x B_x'\right)^{-1} B_x(r)$ *and* $\breve{B}_X(r) = \int_0^r B_x$.

(b) *When* $\Omega_{00.x} > 0$ *and (22) holds*

(ii) $n(\hat{a}-a) \rightsquigarrow \left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1}\left(\int_0^1 \tilde{B}_X B_{0.x}\right) = \left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1} \int_0^1 \overrightarrow{\tilde{B}_X} dB_{0.x}$

$$\equiv \mathcal{MN}\left(0, \Omega_{00.x}\left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1} \int_0^1 \overrightarrow{\tilde{B}_X}\overrightarrow{\tilde{B}_X}' \left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1}\right),$$

where $\overrightarrow{\tilde{B}_X}(r) = \int_r^1 \tilde{B}_X$.

Importantly and as expected, the limit distributions (i) and (ii) are very different for the two cases $\Omega_{00.x} = 0$ and $\Omega_{00.x} > 0$. Singularity raises the rate of convergence in (i) to $O(n^2)$ but introduces nonstandard asymptotics with second order bias effects from endogeneity and serial dependence (Δ_{xe}). Thus, raising the integration order of the system fails to resolve these problems in the least squares asymptotic theory. When $\Omega_{00.x} > 0$, mixed normal asymptotic theory applies as for other procedures like FM-OLS but with some efficiency loss, as discussed in Vogelsang and Wagner (2014). Thus, while this approach of raising the integration order leads to a viable estimation and testing methodology in nonsingular systems, it does not seem to provide a useful methodology for singular, multicointegrated models, at least without the introduction of further modifications to address endogeneity and serial dependence.

The second approach is based on the trend IV (TIV) method of Phillips (2014)[4] or a version of this method based on a fixed number of instruments that is used in recent work by Hwang and Sun (2018). According to this method deterministic instrumental variables $\left\{\varphi_k\left(\frac{t}{n}\right)\right\}_{k=1}^K$ are used to estimate the equation (20) where $\{\varphi_k(r)\}_{k=1}^\infty$ are orthonormal basis functions in $L_2[0,1]$. In what follows, we let $\varphi_k(r) = \sqrt{2}\sin(2\pi k r)$ and $\Phi_K = [\varphi_1, ..., \varphi_K]$, where $\varphi_k' = \left(\varphi_k\left(\frac{1}{n}\right), ..., \varphi_k\left(\frac{n}{n}\right)\right)$. As shown in Phillips (2005, Lemma A) $n^{-1}\Phi_K'\Phi_K = I_K + O\left(\frac{1}{n}\right)$ and then TIV estimation of (20) is asymptotically equivalent to simple least squares regression on the linearly transformed K dimensional system

$$V_y = V_x a + V_{\Delta x} f + V_{u_{0.x}} =: V_c \theta + V_{u_{0.x}}, \quad \theta = (a', f')',$$

using the notation $V_c = \Phi_K' c = \sum_{t=1}^n \tilde{\varphi}_K\left(\frac{t}{n}\right) c_t'$ where $\tilde{\varphi}_K(r) = (\varphi_1(r), ..., \varphi_K(r))'$, $\Phi_K' = \left[\tilde{\varphi}_K\left(\frac{1}{n}\right), ..., \tilde{\varphi}_K\left(\frac{n}{n}\right)\right]$, and $c_t' = (x_t', u_{xt}')$ with similar definitions for $V_y, V_x, V_{\Delta x}, V_{u_{0.x}}$. In standard partitioned regression form, these estimates are

$$\hat{a}_{TIV} - a = \left(V_x' Q_{V_{\Delta x}} V_x\right)^{-1} V_x' Q_{V_{\Delta x}} V_{u_{0.x}}, \tag{28}$$

$$\hat{f}_{TIV} - f = \left(V_{\Delta x}' Q_{V_x} V_{\Delta x}\right)^{-1} V_{\Delta x}' Q_{V_x} V_{u_{0.x}}. \tag{29}$$

[4] This TIV approach was originally proposed by the author in a York University Workshop conference presentation given in 2003 and the full paper was presented in the Faro Time Series Econometrics Conference 2005.

This procedure is called transformed augmented least squares (TA-OLS) in Hwang and Sun (2018), who investigate its asymptotic properties when $\Omega_{00.x} > 0$ and K is fixed as $n \to \infty$.

The third approach is to apply TIV regression to the following augmented regression form of (25).

$$Y_t = a'X_t + f'x_t + g'\Delta x_t + e_t^+ = a'X_t + f'x_t + g'u_{xt} + e_t^+, \tag{30}$$

where $e_t^+ = e_t\mathbf{1}\{\Omega_{00.x} = 0\} + U_{0.xt}\mathbf{1}\{\Omega_{00.x} > 0\}$, as before. For fixed K, this approach is equivalent to least squares regression on the transformed system

$$V_Y = V_X a + V_x f + V_{\Delta x} g + V_{e^+} =: V_X a + V_c \ell + V_{e^+},$$

using $V_C = \Phi_K' C$ with $C_t' = (x_t', u_{xt}')$ and $\ell' = (f', g')$. This procedure utilizes the deterministic transform of the time aggregated version of the model but augments the model further by the inclusion of the regressor Δx_t. Standard partitioned regression on () leads to the following aggregated TIV (ATIV) estimator of a

$$\hat{a}_{ATIV} - a = \left(V_X' Q_{V_C} V_x\right)^{-1} V_x' Q_{V_C} V_{e^+}.$$

The following result shows the asymptotic properties of these various methods as $n \to \infty$, and as $(K, n) \to \infty$, when $\Omega_{00.x} = 0$ in addition to the nonsingular case $\Omega_{00.x} > 0$.

Theorem 2

(a) *When $\Omega_{00.x} = 0$, (23) holds, K is fixed and $n \to \infty$*

(i) $n^2\left(\hat{a}_{TIV} - a\right) \rightsquigarrow \left(\eta_K' Q_{\xi_K} \eta_K\right)^{-1}\left(\eta_K' Q_{\xi_K} \int_0^1 \tilde{\varphi}_K^{(1)} dB_e\right),$

where $\eta_K = \int_0^1 \tilde{\varphi}_K B_x'$, $\xi_K = \int_0^1 \tilde{\varphi}_K dB_x'$, and $Q_{\xi_K} = I_K - \xi_K\left(\xi_K'\xi_K\right)^{-1}\xi_K'$.

(b) *When $\Omega_{00.x} > 0$, (22) holds, K* is fixed *and $n \to \infty$*

(ii) $n\left(\hat{a}_{TIV} - a\right) \rightsquigarrow \left(\eta_K' Q_{\xi_K} \eta_K\right)^{-1}\left(\eta_K' Q_{\xi_K} \int_0^1 \tilde{\varphi}_K dB_{0.x}\right) \equiv \mathcal{MN}\left(0, \Omega_{00.x}\left(\eta_K' Q_{\xi_K} \eta_K\right)^{-1}\right),$

where $B_{0.x}(r) = B_0(r) - \Omega_{0x}\Omega_{xx}^{-1} B_x(r) \equiv BM(\Omega_{00.x})$ and is independent of the Brownian motion B_x.

(c) *When $\Omega_{00.x} > 0$, (22) holds, and $(K, n) \to \infty$* with $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$

(iii) $n\left(\hat{a}_{TIV} - a\right) \rightsquigarrow \left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x dB_{0.x}\right) \equiv \mathcal{MN}\left(0, \Omega_{00.x}\left(\int_0^1 B_x B_x'\right)^{-1}\right).$

(d) *When $\Omega_{00.x} = 0, \omega_{ex} = 0$, (23) holds, and $(K, n) \to \infty$* with $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$

(iv) $n^2\left(\hat{a}_{TIV} - a\right) \rightsquigarrow \left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x dB_{e.x}\right) \equiv \mathcal{MN}\left(0, \omega_e^2\left(\int_0^1 B_x B_x'\right)^{-1}\right).$

(e) *When* $\Omega_{00.x} > 0$, *(22) holds, and* $(K,n) \to \infty$ with $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$

$$\textbf{(v)}\ \ n\left(\hat{a}_{ATIV} - a\right) \rightsquigarrow \left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1} \left(\int_0^1 \tilde{B}_X B_{0.x}\right) = \left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1} \int_0^1 \overrightarrow{\tilde{B}_X} dB_{0.x}$$

$$\equiv \mathcal{MN}\left(0, \Omega_{00.x}\left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1} \int_0^1 \overrightarrow{\tilde{B}_X}\overrightarrow{\tilde{B}_X}' \left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1}\right),$$

where $\overrightarrow{\tilde{B}_X}(r) = \int_r^1 \tilde{B}_X$.

(f) *When* $\Omega_{00.x} = 0$, *(23) holds, and* $(K,n) \to \infty$ with $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$

$$\textbf{(vi)}\ \ n^2\left(\hat{a}_{ATIV} - a\right) \rightsquigarrow \left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1} \left(\int_0^1 \tilde{B}_X dB_{e.x}\right) \equiv \mathcal{MN}\left(0, \omega_{e.x}^2 \left(\int_0^1 \tilde{B}_X \tilde{B}_X'\right)^{-1}\right),$$

where $B_{e.x}(r) = B_e(r) - \omega_{ex}\Omega_{xx}^{-1} B_x(r) \equiv BM\left(\omega_{e.x}^2\right)$, with $\omega_{e.x}^2 = \omega_e^2 - \omega_{ex}\Omega_{xx}^{-1}\omega_{xe}$, *and* $B_{e.x}$ *independent of the Brownian motion* B_x.

Again, the limit distributions (i) and (ii) differ for the two cases $\Omega_{00.x} = 0$ and $\Omega_{00.x} > 0$. The non-singular TIV regression case (ii) has the usual $O(n)$ convergence rate and leads to mixed normal limit theory as in other methods such as FM-OLS regression but without asymptotic efficiency. As shown in (iii) $\hat{a}_{TIV}$ achieves full asymptotic efficiency when $K \to \infty$ as $n \to \infty$ when $K/n \to \infty$, a result proved in Phillips (2014).

The singular case (i) is particularly intriguing because it is so similar in form to case (ii) while raising the rate of convergence to $O\left(n^2\right)$. Importantly, there is no serial correlation bias in contrast to the limit theory for the IM-OLS estimator $\hat{a}$ because in the TIV limit theory the component $\int_0^1 \tilde{\varphi}_K^{(1)} dB_e$ is centred on zero, so the deterministic regressors and their derivatives are valid instruments that necessarily satisfy the orthogonality condition. These features and the focus on long run properties that is brought out by the transform are the strengths of the deterministic IV regression approach developed in Phillips (2005, 2014). The advantages of using such deterministic trend regressors, and a finite number of them in asymptotic calculations, have proved useful in much recent work on the long run properties of economic time series (see, e.g., Müller, 2007; Müller and Watson, 2016; Sun, 2018) and in establishing F and t distribution limit theory in cointegrating regression tests (Hwang and Sun, 2018). See also Lazarus et al. (2018) for a recent overview of some aspects of these methods.

When the long run covariance $\omega_{ex} = 0$, the TIV estimator $\hat{a}_{TIV}$ has mixed normal limit theory

$$n^2\left(\hat{a}_{TIV} - a\right) \rightsquigarrow \left(\eta_K' Q_{\xi_K} \eta_K\right)^{-1} \left(\eta_K' Q_{\xi_K} \int_0^1 \tilde{\varphi}_K^{(1)} dB_e\right) \equiv \mathcal{MN}\left(0, \omega_e^2 \left(\eta_K' Q_{\xi_K} \eta_K\right)^{-1}\right), \tag{31}$$

analogous to that of the non-singular case $\Omega_{00.x} > 0$ in (ii), the only modification being the scale long run variance ω_e^2 in place of $\Omega_{00.x}$.

In the general case where the long run covariance $\mathbb{CV}^{\text{LR}}(e_t, u_{xt}) = \omega_{ex} \neq 0$ and there is long run endogeneity in the singular multicointegrated model, the limit distribution in (i) is no longer mixed normal. But the limit distribution is pivotal up to scale and the long run covariance ω_{ex}. Write $B_x = \Omega_{xx}^{1/2} W_x$ and $B_e = \omega_e W_e$ in terms of standard vector Brownian motion W_x and standard scalar Brownian motion W_e, and define the components $\eta_K^W = \int_0^1 \tilde{\varphi}_K W_x'$ and $\xi_K^W = \int_0^1 \tilde{\varphi}_K dW_x'$. Then from (42) in the proof of the Theorem 2 we have the following alternate representation of the limit theory in terms of the standardized components $\left(\eta_K^W, \xi_K^W, W_e\right)$

$$n^2 (\hat{a}_{TIV} - a) \rightsquigarrow \omega_e \Omega_{xx}^{-1/2} \left(\eta_K^{W\prime} Q_{\xi_K^W} \eta_K^W\right)^{-1} \left(\eta_K^{W\prime} Q_{\xi_K^W} \int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right), \quad (32)$$

which depends on scale multiplication by $\omega_e \Omega_{xx}^{-1/2}$ and possible correlation between W_e and W_x, notably through potential correlation between $\eta_K = \int_0^1 \tilde{\varphi}_K B_x'$ and $\int_0^1 \tilde{\varphi}_K^{(1)} dW_e$ which reflects underlying endogeneity in the singular system. In particular, $\mathbb{E}\left(\int_0^1 \tilde{\varphi}_K^{(1)} dB_e vec(\eta_K)'\right) = I_k \otimes \omega_{ex}$, as shown in Lemma A in the Appendix. But when $\mathbb{CV}^{\text{LR}}(e_t, u_{xt}) = \omega_{ex} = 0$, the limit result is mixed normal with standardized form

$$\mathcal{MN}\left(0, \omega_e^2 \Omega_{xx}^{-1/2} \left(\eta_K^{W_x\prime} Q_{\xi_K^{W_x}} \eta_K^{W_x}\right)^{-1} \Omega_{xx}^{-1/2}\right). \quad (33)$$

The singular case (iv) shows that under singularity ($\Omega_{00.x} = 0$) and no endogeneity ($\omega_{ex} = 0$) the usual TIV estimator $\hat{a}_{TIV}$ is asymptotically mixed normal just as in the nonsingular $\Omega_{00.x} > 0$ case (iii) and with a limit distribution entirely analogous to case (iii) but with higher convergence rate and scale coefficient limit variance ω_e^2 in place of the usual $\Omega_{00.x}$ that applies in the nonsingular case. Thus, use of a full set of deterministic instrumental variables in the limit as $n \to \infty$ delivers mixed normal asymptotics and efficiency, analogous to the $\Omega_{00.x} > 0$ case, when $\Omega_{00.x} = 0$ and $\omega_{ex} = 0$.

Limit theory for the endogenous case where $\omega_{ex} \neq 0$ is given in the singular case (vi) and uses ATIV estimation, or trend IV regression on the augmented aggregated system. (More detail will be provided in the next draft of the paper.)

4.2 Inference

Theorems 1 and 2 show that both TIV and IM-OLS provide consistent and asymptotically mixed normal estimation procedures which are convenient for inference in the standard $I(1)$ cointegrating regression model with nonsingular $\Omega_{00.x} > 0$. But when $\Omega_{00.x} = 0$, IM-OLS suffers from asymptotic second order bias and limit theory that is unsuited to pivotal inference. In what follows we therefore concentrate on the TIV approach to testing.

We start by considering the Wald statistic for testing the full dimension hypothesis $\mathcal{H}_0 : a = a^0$ about the cointegrating vector a. (General linear restrictions will be provided in a later draft.) The usual regression form of this statistic based on the TIV estimate $\hat{a}_{TIV}$ is

$$\text{Wald}_{TIV} = \left(\hat{a}_{TIV} - a^0\right)' \left(V_x' Q_{V_{\Delta x}} V_x\right) \left(\hat{a}_{TIV} - a^0\right) / \hat{\omega}^2 \tag{34}$$

allowing for an estimated long run equation error variance $\hat{\omega}^2$, as usual in cointegrating regression testing. Here we use the simple variance estimate $\hat{\omega}^2 = K^{-1}\hat{V}_{0.x}'\hat{V}_{0.x}$ suggested in Phillips (2005), where the residual vector is $\hat{V}_{0.x} = V_y - V_x\hat{a}_{TIV} - V_{\Delta x}\hat{f}_{TIV}$ and normalization is by K^{-1} in view of the dimension of the system after transformation by the instruments. The following result gives the limit theory of the test statistic Wald_{TIV}.

Theorem 3 *As* $n \to \infty$ with K *fixed the following hold.*

(a) *When* $\Omega_{00.x} = 0$ *and (23) holds*

(i) $\text{Wald}_{TIV} \rightsquigarrow K \frac{\left(\int_0^1 dW_e \tilde{\varphi}_K^{(1)\prime}\right) P_{G_x} \left(\int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right)}{\left(\int_0^1 dW_e \tilde{\varphi}_K^{(1)\prime}\right)\left(I_K - P_{\zeta_c}\right)\left(\int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right)}$

where $G_x = Q_{\xi_K^W}\eta_K^W$, $\eta_K = \int_0^1 \tilde{\varphi}_K B_x = \eta_K^W \Omega_{xx}^{1/2}$, $\xi_K = \int_0^1 \tilde{\varphi}_K dB_x = \xi_K^W \Omega_{xx}^{1/2}$, *and* $Q_{\xi_K} = I_K - \xi_K\left(\xi_K'\xi_K\right)^{-1}\xi_K' = Q_{\xi_K^{W_x}}$.

(i*) $\text{Wald}_{TIV} \rightsquigarrow \frac{Km_x}{K-2m_x} \times F_{m_x, K-2m_x}$ *when* $\omega_{ex} = 0$.

(b) *When* $\Omega_{00.x} > 0$ *and (22) holds*

(ii) $\text{Wald}_{TIV} \rightsquigarrow \frac{Km_x}{K-2m_x} \times F_{m_x, K-2m_x}$,

where $F_{p,q}$ *denotes an* F *distribution with degrees of freedom* p *and* q.

These findings for the Wald test extend in a straightforward way to t ratio statistics which have asymptotic t distributions when $n \to \infty$ and K is fixed, even with $\Omega_{00.x} = 0$ when $\omega_{ex} = 0$. Theorem 3(ii) corresponds to the result[5] in Hwang and Sun (2018) when $\Omega_{00.x} > 0$. Theorem 3(i*) shows that under long run orthogonality of e_t and u_x, precisely the same limit theory holds under singularity when $\Omega_{00.x} = 0$ as when $\Omega_{00.x} > 0$. This feature of TIV inference is remarkable and shows that TIV estimation and associated inferential methods can deliver 'finite sample style F and t distribution asymptotics' in cointegrated systems which are robust to multicointegration provided the long run orthogonality condition $\omega_{ex} = 0$ holds.

5 Conclusion

The above analysis deals with estimation and inference in a scalar cointegrating relationship. But the main ideas extend in a straightforward way to systems

[5] The Wald statistic in Hwang and Sun (2018) is standardized by the numerator degrees of freedom (here m_x), so their limit result in this case is $F_{m_x, K-2m_x}$.

estimation. A general analysis will follow in a later draft that will include various extensions of the present framework. Some of these are as follows.

1. Discussion of the local to singular long run variance matrix $\Omega_{00.x}$ case, which may be useful for inference and robustification of inference.

2. The case of double singularity where $u_{0.xt} = \Delta^2 e_t$ with consequent changes to limit theory and rates of convergence.

3. General Wald statistics to test $Ra = r$ or nonlinear analytic restrictions.

4. Overview on robust inference in cointegrating regression, allowing for singularity.

6 Appendix

This Appendix provides proofs of some subsidiary results and the theorems in the paper. We start with the following.

Lemma A:

(i) $\mathbb{E}\left(\int_0^1 \tilde{\varphi}_K^{(1)} dB_e vec\left(\eta_K\right)'\right) = I_K \otimes \omega_{ex}$, $\mathbb{E}\left(\int_0^1 B_x \tilde{\varphi}_K\left(r\right)' \int_0^1 \tilde{\varphi}_K^{(1)} dB_e\right) = K\omega_{xe}$

(ii) $\mathbb{E}\left(\int_0^1 \tilde{\varphi}_K^{(1)} dB_e vec\left(\xi_K\right)'\right) = 0.$

(iii) $\mathbb{E}\left(\eta_K \xi_K'\right) = 0.$

where $\eta_K = \int_{r=0}^1 \tilde{\varphi}_K\left(r\right) B_x\left(r\right)' dr$ and $\xi_K = \int_{r=0}^1 \tilde{\varphi}_K\left(r\right) dB_x\left(r\right)'$.

Proof of Lemma A

Part (i). Using $vec\left(\cdot\right)$ to vectorize a matrix by rows, we have

$$\begin{aligned}
&\mathbb{E}\left(\int_0^1 \tilde{\varphi}_K^{(1)} dB_e vec\left(\eta_K\right)'\right) = \mathbb{E}\left\{\int_{s=0}^1 \tilde{\varphi}_K^{(1)}\left(s\right) dB_e\left(s\right) vec\left\{\int_{r=0}^1 \tilde{\varphi}_K\left(r\right) B_x\left(r\right)' dr\right\}'\right\} \\
=\;& \mathbb{E}\left\{\int_{s=0}^1 \tilde{\varphi}_K^{(1)}\left(s\right) dB_e\left(s\right) \int_{r=0}^1 \tilde{\varphi}_K\left(r\right)' \otimes B_x\left(r\right)' dr\right\} \\
=\;& \int_{r=0}^1 \int_{s=0}^1 \tilde{\varphi}_K^{(1)}\left(s\right)\left[\tilde{\varphi}_K\left(r\right)' \otimes \mathbb{E}\left\{dB_e\left(s\right) \int_0^r dB_x\left(p\right)'\right\}\right] dr = \int_{r=0}^1 \int_{s=0}^r \tilde{\varphi}_K^{(1)}\left(s\right)\left[\tilde{\varphi}_K\left(r\right)' \otimes \omega_{ex} ds\right] dr \\
=\;& \int_{r=0}^1 \int_{s=0}^r \tilde{\varphi}_K^{(1)}\left(s\right) ds \tilde{\varphi}_K\left(r\right)' dr \otimes \omega_{ex} = \int_{r=0}^1 \tilde{\varphi}_K\left(r\right) \tilde{\varphi}_K\left(r\right)' dr \otimes \omega_{ex} = I_K \otimes \omega_{ex},
\end{aligned}$$

by orthonormality of the $\{\varphi_k\}$ sequence. The second result follows immediately.

Part (ii). By direct calculation, we have

$$\begin{aligned}
&\mathbb{E}\left(\int_0^1 \tilde{\varphi}_K^{(1)} dB_e vec\left(\xi_K\right)'\right) = \mathbb{E}\left\{\int_{s=0}^1 \tilde{\varphi}_K^{(1)}\left(s\right) dB_e\left(s\right) vec\left\{\int_{r=0}^1 \tilde{\varphi}_K\left(r\right) dB_x\left(r\right)'\right\}'\right\} \\
= \;& \mathbb{E}\left\{\int_{s=0}^1 \tilde{\varphi}_K^{(1)}\left(s\right) dB_e\left(s\right) \int_{r=0}^1 \tilde{\varphi}_K\left(r\right)' \otimes dB_x\left(r\right)'\right\} \\
= \;& \int_{r=0}^1 \int_{s=0}^1 \tilde{\varphi}_K^{(1)}\left(s\right)\left[\tilde{\varphi}_K\left(r\right)' \otimes \mathbb{E}\left\{dB_e\left(s\right) dB_x\left(r\right)'\right\}\right] dr \\
= \;& \int_0^1 \tilde{\varphi}_K^{(1)}\left(r\right) \tilde{\varphi}_K\left(r\right)' dr \otimes \omega_{ex} = 0,
\end{aligned}$$

since

$$\begin{aligned}
&\int_0^1 \varphi_j^{(1)}\left(r\right) \varphi_k\left(r\right) dr = 4\pi j \int_0^1 \cos\left\{2\pi j r\right\} \sin\left\{2\pi k r\right\} dr \\
= \;& 2\pi j \int_0^1 \left[\sin\left\{2\pi\left(j+k\right) r\right\} + \sin\left\{2\pi\left(j-k\right) r\right\}\right] dr \\
= \;& 2\pi j \left\{\left[-\cos\left\{2\pi\left(j+k\right) r\right\}\right]_0^1 + \left[-\cos\left\{2\pi\left(j-k\right) r\right\}\right]_0^1\right\} \\
= \;& 2\pi j \left\{\left[-1+1\right] + \left[-1+1\right]\right\} = 0.
\end{aligned}$$

Part (iii). By direct calculation again, we have

$$\begin{aligned}
&\mathbb{E}\left(\eta_K \xi_K'\right) = \mathbb{E}\left\{\int_{r=0}^1 \tilde{\varphi}_K\left(r\right) B_x\left(r\right)' dr \int_{s=0}^1 dB_x\left(s\right) \tilde{\varphi}_K\left(s\right)''\right\} \\
= \;& \int_{r=0}^1 \int_{s=0}^1 \tilde{\varphi}_K\left(r\right) \int_0^r \mathbb{E}\left(dB_x\left(p\right)' dB_x\left(s\right)\right) \tilde{\varphi}_K\left(s\right)' dr \\
= \;& \int_{r=0}^1 \tilde{\varphi}_K\left(r\right) \int_0^r \tilde{\varphi}_K\left(p\right)' dp dr \times \text{trace}\left[\Omega_{xx}\right] = 0,
\end{aligned}$$

because the (j,k) element is

$$\begin{aligned}
&\int_{r=0}^1 \tilde{\varphi}_j\left(r\right) \int_0^r \tilde{\varphi}_k\left(p\right)' dp dr = 2\int_{r=0}^1 \sin\left(2\pi j r\right) \int_0^r \sin\left(2\pi k p\right) dp dr \\
= \;& 2\int_{r=0}^1 \sin\left(2\pi j r\right)\left[\frac{-\cos\left(2k\pi p\right)}{2\pi k}\right]_0^r dr = \frac{1}{\pi k}\int_{r=0}^1 \sin\left(2\pi j r\right)\left[1-\cos\left(2k\pi r\right)\right] dr \\
= \;& \frac{1}{2\pi k}\int_{r=0}^1 \sin\left(4\pi j r\right) dr = 0, \text{ for all } \left(j,k\right).
\end{aligned}$$

Proof of Theorem 1

Part (a) In this case $u_{0.xt} = \Delta e_t$, $\sum_{t=1}^{\lfloor n\cdot\rfloor} u_{0.xt} = e_{\lfloor n\cdot\rfloor} - e_0 \rightsquigarrow e_\infty - e_0$ as $n \to \infty$ and no invariance principle holds for $\sum_{t=1}^{\lfloor n\cdot\rfloor} u_{0.xt}$. Further, by standard weak convergence methods (Phillips, 1986, 1988) we have the following limit behavior for the component sample moments:

(a-1) $n^{-1}\sum_{t=1}^{n}x_te_t \rightsquigarrow \int_0^1 B_x dB_e + \Delta_{xe}$, where $\Delta_{xe} = \sum_{h=0}^{\infty}\mathbb{E}\left(u_{x0}e_h\right)$.

(a-2) $n^{-3}\sum_{t=1}^{n}X_tx_t' = n^{-1}\sum_{t=1}^{n}\frac{X_t}{n^{3/2}}\frac{x_t'}{n^{1/2}} \rightsquigarrow \int_0^1 \breve{B}_X B_x'$, where $B_x(r) = BM(\Omega_{xx})$ and $\breve{B}_X(r) = \int_0^r B_x$.

(a-3) $n^{-4}X'Q_xX = n^{-1}\sum_{t=1}^{n}\frac{X_t}{n^{3/2}}\frac{X_t'}{n^{3/2}} - \left(n^{-3}\sum_{t=1}^{n}X_tx_t'\right)\left(n^{-2}\sum_{t=1}^{n}x_tx_t'\right)^{-1}\left(n^{-3}\sum_{t=1}^{n}x_tX_t'\right) \rightsquigarrow \int_0^1 \tilde{B}_X\tilde{B}_X'$ where $\tilde{B}_X(r) = \breve{B}_X(r) - \left(\int_0^1 \breve{B}_X B_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}B_x(r)$ is the L_2 projection residual of B_X on B_x.

(a-4) $n^{-2}X'Q_xe = \sum_{t=1}^{n}\frac{X_t}{n^{3/2}}\frac{e_t}{n^{1/2}} - \left(n^{-3}\sum_{t=1}^{n}X_tx_t'\right)\left(n^{-2}\sum_{t=1}^{n}x_tx_t'\right)^{-1}\left(n^{-1}\sum_{t=1}^{n}x_te_t\right) \rightsquigarrow \int_0^1 \tilde{B}_X dB_e$.

Results (a-1)-(a-3) follow immediately by standard manipulations (Phillips, 1986, 1988). In particular, note that $n^{-3/2}X_{\lfloor nr\rfloor} = n^{-3/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}x_t \rightsquigarrow \breve{B}_X(r)$ and setting $E_t = \sum_{s=1}^{t}e_s$, $E_0 = 0$, we have $n^{-1/2}E_{\lfloor nr\rfloor} \rightsquigarrow B_e(r)$. To confirm (a-4), use partial summation to write

$$\begin{aligned}\sum_{t=1}^{n}X_te_t &= \sum_{t=1}^{n}X_t\Delta E_t = \Delta\left(\sum_{t=1}^{n}X_tE_t\right) - \sum_{t=1}^{n}\Delta X_tE_{t-1}\\ &= X_nE_n - \sum_{t=1}^{n}x_tE_{t-1}.\end{aligned}$$

Then

$$\begin{aligned}\frac{1}{n^2}\sum_{t=1}^{n}X_te_t &= \frac{X_n}{n^{3/2}}\frac{E_n}{n^{1/2}} - \frac{1}{n}\sum_{t=1}^{n}\frac{x_t}{n^{1/2}}\frac{E_{t-1}}{n^{1/2}} \rightsquigarrow \breve{B}_X(1)B_e(1) - \int_0^1 B_xB_e\\ &= \int_0^1 \breve{B}_X dB_e, \qquad (35)\end{aligned}$$

by integration by parts since $\breve{B}_X(r) = \int_0^r B_x$ is of bounded variation and so $\int_0^1 \breve{B}_X dB_e = \left[\breve{B}_X B_e\right]_0^1 - \int_0^1 B_xB_e = \breve{B}_X(1)B_e(1) - \int_0^1 B_xB_e$. Next by (a-1) and (35) we have

$$\begin{aligned}n^{-2}X'Q_xe &\rightsquigarrow \int_0^1 \breve{B}_X dB_e - \left(\int_0^1 \breve{B}_X B_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\left(\int_0^1 B_x dB_e + \Delta_{xe}\right)\\ &= \int_0^1 \tilde{B}_X dB_e - \left(\int_0^1 \breve{B}_X B_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\Delta_{xe},\end{aligned}$$

with $\tilde{B}_X(r) = \breve{B}_X(r) - \left(\int_0^1 \breve{B}_X B_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}B_x(r)$, giving result (a-4). Combining (a-3) and (a-4) and using continuous mapping leads to the stated limit result

$$n^2(\hat{a} - a) \rightsquigarrow \left(\int_0^1 \tilde{B}_X\tilde{B}_X'\right)^{-1}\left\{\int_0^1 \tilde{B}_X dB_e - \left(\int_0^1 \breve{B}_X B_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\Delta_{xe}\right\}$$

given in (i) for $n^2(\hat{a}-a)$.

Part (b) When $\Omega_{00.x}>0$, the system (25) is $Y_t=a'X_t+f'x_t+U_{0.xt}$ and then

$$\hat{a}-a=\left(X'Q_xX\right)^{-1}X'Q_xU_{0.x},\ \hat{f}-f=\left(x'Q_Xx\right)^{-1}x'Q_XU_{0.x}$$

Standard methods lead to the following component limits:

(b-1) $n^{-1}\sum_{t=1}^{n}x_tu_{0.xt}\rightsquigarrow\int_0^1 B_xdB_{0.x}+\Delta_{x0}^{+}$,
where $\Delta_{x0}^{+}=\sum_{h=-\infty}^{0}\mathbb{E}\left(u_{xh}u_{0.x,0}\right)=\Delta_{x0}-\Delta_{xx}\Omega_{xx}^{-1}\Omega_{x0}$ as $u_{0.xt}=u_{0t}-\Omega_{0x}\Omega_{xx}^{-1}u_{xt}=u_{0t}-u_{xt}'\Omega_{xx}^{-1}\Omega_{x0}$.

(b-2) $n^{-1}\sum_{t=1}^{n}\frac{x_t}{n^{1/2}}\frac{U_{0.xt}}{n^{1/2}}\rightsquigarrow\int_0^1 B_xB_{0.x}$, $\frac{1}{n}\sum_{t=1}^{n}\frac{X_t}{n^{3/2}}\frac{U_{0.xt}}{n^{1/2}}\rightsquigarrow\int_0^1\breve{B}_XB_{0.x}$.

(b-3) $n^{-3}X'Q_xU_{0.x}=\frac{1}{n}\sum_{t=1}^{n}\frac{X_t}{n^{3/2}}\frac{U_{0.xt}}{n^{1/2}}-\left(n^{-3}\sum_{t=1}^{n}X_tx_t'\right)\left(n^{-2}\sum_{t=1}^{n}x_tx_t'\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}\frac{x_t}{n^{1/2}}\frac{U_{0.xt}}{n^{1/2}}\right)$
$\rightsquigarrow\int_0^1\breve{B}_XB_{0.x}-\left(\int_0^1\breve{B}_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\left(\int_0^1 B_xB_{0.x}\right)=\int_0^1\tilde{B}_XB_{0.x}$.

It then follows that

$$n\left(\hat{a}-a\right)=\left(n^{-4}X'Q_xX\right)^{-1}\left(n^{-3}X'Q_xU_{0.x}\right)\rightsquigarrow\left(\int_0^1\tilde{B}_X\tilde{B}_X'\right)^{-1}\left(\int_0^1\tilde{B}_XB_{0.x}\right),$$

as stated in (i). Note that

$$\int_0^1\tilde{B}_XB_{0.x}=\left[\left(\int_0^r\tilde{B}_X\right)B_{0.x}(r)\right]_0^1-\int_0^1\left(\int_0^r\tilde{B}_X\right)dB_{0.x}(r)$$
$$=\ \widetilde{\tilde{B}_X}(1)\,B_{0.x}(1)-\int_0^1\widetilde{\tilde{B}_X}(r)\,dB_{0.x}(r)=\widetilde{\tilde{B}_X}(1)\int_0^1 dB_{0.x}(r)-\int_0^1\widetilde{\tilde{B}_X}(r)\,dB_{0.x}(r)$$
$$=\ \int_0^1\left[\widetilde{\tilde{B}_X}(1)-\widetilde{\tilde{B}_X}(r)\right]dB_{0.x}(r)=\int_0^1\overrightarrow{\tilde{B}_X}(r)\,dB_{0.x}(r)$$

where $\widetilde{\tilde{B}_X}(r):=\int_0^r\tilde{B}_X$ and $\overrightarrow{\tilde{B}_X}(r):=\int_r^1\tilde{B}_X$. The alternative version of the limit result now follows immediately.

Proof of Theorem 2

Part (a) We proceed as in the proof of Theorem 1 by examining the limit behavior of the various components of $\hat{a}_{TIV}$. These arise from the transform limit theory, as in Phillips (2005, 2014) and Hwang and Sun (2018). But here we take into account the degeneracy $\Omega_{00.x}=0$.

We first consider $V_{u_{0.x}}$. Using partial summation and the bounded continuous differentiability of $\varphi_k(r) = \sqrt{2}\sin(2\pi kr)$, we have

$$\begin{aligned} V_{u_{0.x}} &= \sum_{t=1}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right) u_{0.xt} = \sum_{t=1}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right) \Delta e_t = \Delta\left\{\sum_{t=1}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right) e_t\right\} - \sum_{t=1}^{n}\left\{\Delta\tilde{\varphi}_K\left(\frac{t}{n}\right)\right\} e_{t-1} \\ &= \tilde{\varphi}_K(1) e_n - \tilde{\varphi}_K(0) e_0 - \frac{1}{n}\sum_{t=1}^{n}\left\{\tilde{\varphi}_K^{(1)}\left(\frac{t}{n}\right) + O\left(\frac{1}{n}\right)\right\} e_{t-1} \\ &= -\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \tilde{\varphi}_K^{(1)}\left(\frac{t}{n}\right)\frac{e_{t-1}}{\sqrt{n}}\left\{1 + o_p(1)\right\}, \end{aligned}$$

so that

$$n^{1/2} V_{u_{0.x}} \rightsquigarrow \int_0^1 \tilde{\varphi}_K^{(1)} dB_e, \tag{36}$$

where $\tilde{\varphi}_K^{(1)}(r) = \left(\varphi_1^{(1)}(r), ..., \varphi_K^{(1)}(r)\right)'$ and $\varphi_k^{(1)}(r) = 2^{3/2}\pi k\cos(2\pi kr)$. Next note that

$$\frac{1}{n^{3/2}} V_x = \frac{1}{n}\sum_{t=1}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right)\frac{x_t'}{\sqrt{n}} \rightsquigarrow \int_0^1 \tilde{\varphi}_K B_x' =: \eta_K, \tag{37}$$

$$\frac{1}{\sqrt{n}} V_{\Delta x} = \sum_{t=1}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right)\frac{u_{xt}'}{\sqrt{n}} \rightsquigarrow \int_0^1 \tilde{\varphi}_K dB_x' =: \xi_K, \tag{38}$$

$$\begin{aligned} \frac{1}{n^3} V_x' Q_{V_{\Delta x}} V_x &= \frac{1}{n^3} V_x' V_x - \frac{V_x'}{n^{3/2}}\frac{V_{\Delta x}}{\sqrt{n}}\left(\frac{V_{\Delta x}'}{\sqrt{n}}\frac{V_{\Delta x}}{\sqrt{n}}\right)^{-1}\frac{V_{\Delta x}'}{\sqrt{n}}\frac{V_x}{n^{3/2}} \\ &\rightsquigarrow \eta_K' Q_{\xi_K}\eta_K, \end{aligned} \tag{39}$$

$$\begin{aligned} \frac{1}{n} V_x' Q_{V_{\Delta x}} V_{u_{0.x}} &= \frac{V_x'}{n^{3/2}}\left(n^{1/2} V_{u_{0.x}}\right) - \frac{V_x'}{n^{3/2}}\frac{V_{\Delta x}}{\sqrt{n}}\left(\frac{V_{\Delta x}'}{\sqrt{n}}\frac{V_{\Delta x}}{\sqrt{n}}\right)^{-1}\frac{V_{\Delta x}'}{\sqrt{n}}\left(n^{1/2} V_{u_{0.x}}\right) \\ &\rightsquigarrow \eta_K' Q_{\xi_K}\int_0^1 \tilde{\varphi}_K^{(1)} dB_e. \end{aligned} \tag{40}$$

Then, as $n \to \infty$

$$\begin{aligned} n^2\left(\hat{a}_{TIV} - a\right) &= \left(\frac{1}{n^3} V_x' Q_{V_{\Delta x}} V_x\right)^{-1}\left(\frac{1}{n} V_x' Q_{V_{\Delta x}} V_{u_{0.x}}\right) \\ &\rightsquigarrow \left(\eta_K' Q_{\xi_K}\eta_K\right)^{-1}\left(\eta_K' Q_{\xi_K}\int_0^1 \tilde{\varphi}_K^{(1)} dB_e\right), \end{aligned} \tag{41}$$

giving the stated result. Note that $u_{0.xt} = \Delta e_t$ and although the long run covariance of u_{xt} and $u_{0.xt}$ is zero this does not mean that the long run covariance of u_{xt} and e_t is zero. For instance, if $e_t = \alpha' u_{xt}$, then $\mathbb{CV}^{\mathrm{LR}}(u_{xt}, u_{0.xt}) = \mathbb{CV}^{\mathrm{LR}}(u_{xt}, \Delta u_{xt}'\alpha) = 0$ but $\mathbb{CV}^{\mathrm{LR}}(u_{xt}, e_t) = \Omega_{xx}\alpha \neq 0$.

We now use the representations $B_x = \Omega_{xx}^{1/2} W_x$ and $B_e = \omega_{ee}^{1/2} W_e$ in terms of standard vector Brownian motion W_x and standard Brownian motion W_e, and define $\eta_K^W = \int_0^1 \tilde{\varphi}_K W_x$ and $\xi_K^W = \int_0^1 \tilde{\varphi}_K dW_x$. Then

$$\eta_K' Q_{\xi_K} \eta_K = \Omega_{xx}^{1/2} \left(\eta_K^{W\prime} Q_{\xi_K^W} \eta_K^W \right) \Omega_{xx}^{1/2} = \Omega_{xx}^{1/2} \left(\eta_K^{W\prime} Q_{\xi_K} \eta_K^W \right) \Omega_{xx}^{1/2},$$

since $Q_{\xi_K^W} = \xi_K^W \Omega_{xx}^{1/2} \left(\Omega_{xx}^{1/2} \xi_K^{W\prime} \xi_K^W \Omega_{xx}^{1/2} \right)^{-1} \Omega_{xx}^{1/2} \xi_K^{W\prime} = \xi_K \left(\xi_K' \xi_K \right)^{-1} \xi_K' = Q_{\xi_K}$. Hence, the limit theory can be written in the following standardized form

$$n^2 \left(\hat{a}_{TIV} - a \right) \rightsquigarrow \omega_e \Omega_{xx}^{-1/2} \left(\eta_K^{W\prime} Q_{\xi_K^W} \eta_K^W \right)^{-1} \left(\eta_K^{W\prime} Q_{\xi_K^W} \int_0^1 \tilde{\varphi}_K^{(1)} dW_e \right), \quad (42)$$

which is important for inference.

Parts (b) & (c) Next consider the case where $\Omega_{00.x} > 0$. Here we have the standard limit theory

$$V_{u_{0.x}} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \tilde{\varphi}_K \left(\frac{t}{n} \right) u_{0.xt} \rightsquigarrow \int_0^1 \tilde{\varphi}_K dB_{0.x} = \Omega_{00.x}^{1/2} \int_0^1 \tilde{\varphi}_K dW_{0.x}.$$

Combining this result with (37)-(39) we obtain the following mixed normal limit theory as $n \to \infty$

$$\begin{aligned} n \left(\hat{a}_{TIV} - a \right) &\rightsquigarrow \Omega_{00.x}^{1/2} \Omega_{xx}^{-1/2} \left(\eta_K^{W\prime} Q_{\xi_K} \eta_K^W \right)^{-1} \left(\eta_K^{W\prime} Q_{\xi_K} \int_0^1 \tilde{\varphi}_K dW_{0.x} \right) \\ &= \Omega_{00.x}^{1/2} \left(\eta_K' Q_{\xi_K} \eta_K \right)^{-1} \left(\eta_K' Q_{\xi_K} \int_0^1 \tilde{\varphi}_K dW_{0.x} \right) \\ &\equiv \mathcal{MN} \left(0, \Omega_{00.x} \left(\eta_K' Q_{\xi_K} \eta_K \right)^{-1} \right), \end{aligned}$$

as in Hwang and Sun (2018), giving part (b). When $K \to \infty$ as $n \to \infty$ under the rate conditions given in Phillips (2014) we have

$$n \left(\hat{a}_{TIV} - a \right) \rightsquigarrow \left(\int_0^1 B_x B_x' \right)^{-1} \left(\int_0^1 B_x dB_{0.x} \right) \equiv \mathcal{MN} \left(0, \Omega_{00.x} \left(\int_0^1 B_x B_x' \right)^{-1} \right),$$

where the limit variance achieves the semiparametric efficiency bound (Phillips, 1991), giving part (c).

Part (d) The proof of this part follows the general line of argument that was developed in the proof of the main theorem of Phillips (2014). From (41) we already have the large $n \to \infty$, fixed K limit theory

$$n^2 \left(\hat{a}_{TIV} - a \right) = \left(\frac{1}{n^3} V_x' Q_{V_{\Delta x}} V_x \right)^{-1} \left(\frac{1}{n} V_x' Q_{V_{\Delta x}} V_{u_{0.x}} \right) \rightsquigarrow \left(\eta_K' Q_{\xi_K} \eta_K \right)^{-1} \left(\eta_K' Q_{\xi_K} \int_0^1 \tilde{\varphi}_K^{(1)} dB_e \right).$$

We now examiine the behavior as $(K,n)\to\infty$ of these two components

$$\frac{1}{n^3}V_x'Q_{V_{\Delta x}}V_x \rightsquigarrow \int_0^1 B_xB_x' \text{ and } \frac{1}{n}V_x'Q_{V_{\Delta x}}V_{u_{0.x}} \rightsquigarrow \int_0^1 B_x dB_{e.x}. \tag{43}$$

In the full TIV approach the system is projected on the range space of the instruments Φ_K using the projector $P_{\Phi_K}=\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'$. It is useful to write out the full expression in this case as follows

$$n^2\left(\hat{a}_{TIV}-a\right)=\left\{\frac{1}{n^2}X'P_{\Phi_K}X-\frac{1}{K}\left(\frac{1}{n}X'P_{\Phi_K}U_x\right)\left(\frac{1}{K}U_x'P_{\Phi_K}U_x\right)^{-1}\left(\frac{1}{n}U_x'P_{\Phi_K}X\right)\right\}$$
$$\times\left\{X'P_{\Phi_K}U_{0.x}-\left(X'P_{\Phi_K}U_x\right)\left(\frac{1}{K}U_x'P_{\Phi_K}U_x\right)^{-1}\left(\frac{1}{K}U_x'P_{\Phi_K}U_{0.x}\right)\right\}. \tag{44}$$

To develop joint $(K,n)\to\infty$ asymptotics for these components we use an expansion of the probability space that includes the limit processes (B_e,B_x) and within that space use an 'in probability' version of the convergence (23) to the limiting Brownian motions (B_e,B_x), as in Lemma A of Phillips (2014) or Lemma C of Phillips (2007). This device leads to the establishment of weak convergence in the original space. Starting with the first member of (43) we have the following decomposition

$$\frac{1}{n^3}V_x'Q_{V_{\Delta x}}V_x=\frac{1}{n^3}V_x'V_x-\frac{V_x'}{n^{3/2}}\frac{V_{\Delta x}}{\sqrt{n}}\left(\frac{V_{\Delta x}'}{\sqrt{n}}\frac{V_{\Delta x}}{\sqrt{n}}\right)^{-1}\frac{V_{\Delta x}'}{\sqrt{n}}\frac{V_x}{n^{3/2}}. \tag{45}$$

Looking at the lead term on the right side of (45) and the corresponding term $n^{-2}X'P_{\Phi_K}X$ in the first member of (44) we have

$$\frac{1}{n^2}X'P_{\Phi_K}X=\frac{1}{n^2}X'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'X$$
$$=\left(\frac{1}{n}\sum_{t=1}^n\frac{x_t}{\sqrt{n}}\tilde{\varphi}_{Kt}'\right)\left(\frac{1}{n}\sum_{t=1}^n\tilde{\varphi}_{Kt}\tilde{\varphi}_{Kt}'\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^n\tilde{\varphi}_{Kt}\frac{x_t'}{\sqrt{n}}\right)$$
$$=\left(\frac{1}{n}\sum_{t=1}^n\frac{x_t}{\sqrt{n}}\tilde{\varphi}_{Kt}'\right)\left(I_K+O\left(\frac{1}{n}\right)\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^n\tilde{\varphi}_{Kt}\frac{x_t'}{\sqrt{n}}\right)=\frac{1}{n^3}V_x'V_x\left\{1+O\left(\frac{1}{n}\right)\right\},$$

where we use the notation $\tilde{\varphi}_{Kt}=\tilde{\varphi}_K\left(\frac{t}{n}\right)$; and, as shown in equation (34) of Phillips (2014)

$$\frac{1}{n^2}X'P_{\Phi_K}X\rightsquigarrow\int_0^1 B_xB_x'. \tag{46}$$

Next, in place of $\frac{V_{\Delta x}'}{\sqrt{n}}\frac{V_{\Delta x}}{\sqrt{n}}$ in (45) we have

$$\frac{1}{K}U_x'P_{\Phi_K}U_x={U_x}'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'U_x=\frac{1}{K}\left(\sum_{t=1}^n\frac{u_{xt}}{\sqrt{n}}\tilde{\varphi}_{Kt}'\right)\left(\frac{1}{n}\sum_{t=1}^n\tilde{\varphi}_{Kt}\tilde{\varphi}_{Kt}'\right)^{-1}\left(\sum_{t=1}^n\tilde{\varphi}_{Kt}\frac{u_{xt}'}{\sqrt{n}}\right)$$
$$=\frac{1}{K}\left(\sum_{t=1}^n\frac{u_{xt}}{\sqrt{n}}\tilde{\varphi}_{Kt}'\right)\left(\sum_{t=1}^n\tilde{\varphi}_{Kt}\frac{u_{xt}'}{\sqrt{n}}\right)\left\{1+O\left(\frac{1}{n}\right)\right\}=\frac{1}{K}\frac{V_{\Delta x}'}{\sqrt{n}}\frac{V_{\Delta x}}{\sqrt{n}}\left\{1+O\left(\frac{1}{n}\right)\right\}, \tag{47}$$

and so

$$K^{-1}U_x'P_{\Phi_K}U_x \to_p \Omega_{xx} \tag{48}$$

as $(K,n) \to \infty$ by Phillips (2005; 2014, Lemma C). Further

$$\begin{aligned}
& \frac{1}{n}X'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'U_x = \left(\frac{1}{n}\frac{X}{\sqrt{n}}'\Phi_K\right)\left(\frac{1}{n}\Phi_K'\Phi_K\right)^{-1}\left(\frac{1}{\sqrt{n}}\Phi_K'U_x\right) \\
= & \left(\frac{1}{n}\sum_{t=1}^{n}\frac{x_t}{\sqrt{n}}\tilde{\varphi}_{Kt}'\right)\left(\sum_{t=1}^{n}\tilde{\varphi}_{Kt}\frac{u_{xt}'}{\sqrt{n}}\right)\left\{1+O\left(\frac{1}{n}\right)\right\} \\
= & \frac{V_x'}{n^{3/2}}\frac{V_{\Delta x}}{\sqrt{n}}\left\{1+O\left(\frac{1}{n}\right)\right\} = O(1). \qquad (49)
\end{aligned}$$

Combining (46) - (49), we deduce that as $(K,n) \to \infty$

$$\begin{aligned}
& \frac{1}{n^2}X'P_{\Phi_K}X - \frac{1}{K}\left(\frac{1}{n}X'P_{\Phi_K}U_x\right)\left(\frac{1}{K}U_x'P_{\Phi_K}U_x\right)^{-1}\left(\frac{1}{n}U_x'P_{\Phi_K}X\right) \\
= & \frac{1}{n^2}X'P_{\Phi_K}X + O\left(\frac{1}{K}\right) \rightsquigarrow \int_0^1 B_xB_x', \qquad (50)
\end{aligned}$$

as in equation (46) of Phillips (2014). Equivalently, in the notation of (45) as $(K,n) \to \infty$

$$\begin{aligned}
\frac{1}{n^3}V_x'Q_{V_{\Delta x}}V_x &= \left\{\frac{1}{n^2}X'P_{\Phi_K}X\right\}\left\{1+O\left(\frac{1}{n}\right)\right\} - \frac{1}{K}\frac{V_x'}{n^{3/2}}\frac{V_{\Delta x}}{\sqrt{n}}\left(\frac{1}{K}\frac{V_{\Delta x}'}{\sqrt{n}}\frac{V_{\Delta x}}{\sqrt{n}}\right)^{-1}\frac{V_{\Delta x}'}{\sqrt{n}}\frac{V_x}{n^{3/2}} \\
&= \frac{1}{n^2}X'P_{\Phi_K}X + O\left(\frac{1}{K}\right) \rightsquigarrow \int_0^1 B_xB_x'.
\end{aligned}$$

We next consider the second factor in braces of (44). In this factor there are two terms

$$X'P_{\Phi_K}U_{0.x} \text{ and } \left(X'P_{\Phi_K}U_x\right)\left(\frac{1}{K}U_x'P_{\Phi_K}U_x\right)^{-1}\left(\frac{1}{K}U_x'P_{\Phi_K}U_{0.x}\right), \tag{51}$$

which we consider in turn. First, using the notation $\tilde{\varphi}_{Kt}^{(1)} = \tilde{\varphi}_K^{(1)}\left(\frac{t}{n}\right)$ we have from earlier analysis that

$$\Phi_K'U_{0.x} = \sum_{t=1}^{n}\tilde{\varphi}_{Kt}u_{0.xt} = \sum_{t=1}^{n}\tilde{\varphi}_{Kt}\Delta e_t = -\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\tilde{\varphi}_{Kt}^{(1)}\frac{e_{t-1}}{\sqrt{n}}\left\{1+o_p(1)\right\},$$

so that

$$\begin{aligned}
X'P_{\Phi_K}U_{0.x} &= X'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'U_{0.x} = \left(\frac{1}{n}X'\Phi_K\right)\left(\frac{1}{n}\Phi_K'\Phi_K\right)^{-1}\left\{-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\tilde{\varphi}_{Kt}^{(1)}\frac{e_{t-1}}{\sqrt{n}}\left\{1+o_p(1)\right\}\right\} \\
&= -\left(\frac{1}{n^{3/2}}X'\Phi_K\right)\left(I_K+O\left(\frac{1}{n}\right)\right)^{-1}\left(\sum_{t=1}^{n}\tilde{\varphi}_{Kt}^{(1)}\frac{e_{t-1}}{\sqrt{n}}\right)\left\{1+o_p(1)\right\} \\
&= -\left(\frac{1}{n^{3/2}}X'\Phi_K\right)\left(\frac{1}{n^{1/2}}\Phi_K^{(1)\prime}E_{-1}\right)\left\{1+o_p(1)\right\}, \qquad (52)
\end{aligned}$$

where $E = [e_1, ..., e_n]$ and $\Phi_K^{(1)} = \left[\tilde{\varphi}_{K1}^{(1)}, ..., \tilde{\varphi}_{Kn}^{(1)}\right]$.

As already noted, when $n \to \infty$ for all fixed K

$$\frac{1}{\sqrt{n}} U_x' \Phi_K \rightsquigarrow \int_0^1 dB_x(r) \tilde{\varphi}_K(r)', \quad \frac{1}{\sqrt{n}} \Phi_K^{(1)\prime} E_{-1} \rightsquigarrow \int_0^1 \tilde{\varphi}_K^{(1)}(r) dB_e(r),$$

and

$$\begin{aligned} \mathbb{E}\left(\frac{1}{\sqrt{n}} U_x' \Phi_K\right)\left(\frac{1}{\sqrt{n}} \Phi_K^{(1)} E_{-1}\right) &\to \mathbb{E}\left\{\int_0^1 dB_x(r) \tilde{\varphi}_K(r)'\right\} \int_0^1 \tilde{\varphi}_K^{(1)}(r) dB_e(r) \\ &= \omega_{xe}\left\{\int_0^1 \tilde{\varphi}_K(r)' \tilde{\varphi}_K^{(1)}(r) d(r)\right\} = 0, \end{aligned}$$

as in the proof of Lemma A (ii) since

$$\begin{aligned} &\int_0^1 \tilde{\varphi}_K(r)' \tilde{\varphi}_K^{(1)}(r) dr = 2\sum_{k=1}^K 2\pi k \int_0^1 (\sin(2\pi kr)\cos(2\pi kr)) dr \\ =\ & 2\sum_{k=1}^K \pi k \int_0^1 \sin(4\pi kr) dr = 2\sum_{k=1}^K \pi k \left[-\frac{\cos(4\pi kr)}{4\pi k}\right]_0^1 \\ =\ & \frac{1}{2}\left[\sum_{k=1}^K [1 - \cos(4\pi k)]\right] = 0. \end{aligned}$$

It follows that the components $\frac{1}{\sqrt{n}} U_x' \Phi_K$ and $\frac{1}{\sqrt{n}} \Phi_K^{(1)} E_{-1}$ are asymptotically uncorrelated as $n \to \infty$ and therefore independent since both are asymptotically normal. We deduce that as $n, K \to \infty$

$$\frac{1}{K}\left(\frac{1}{n^{3/2}} X' \Phi_K\right)\left(\frac{1}{\sqrt{n}} \Phi_K' U_x\right) \Omega_{xx}^{-1}\left(\frac{1}{\sqrt{n}} U_x' \Phi_K\right)\left(\frac{1}{\sqrt{n}} \Phi_K^{(1)\prime} E_{-1}\right) = O_p\left(\frac{1}{K}\right)$$

$$\begin{aligned} & X' P_{\Phi_K} U_{0.x} - (X' P_{\Phi_K} U_x)\left(\frac{1}{K} U_x' P_{\Phi_K} U_x\right)^{-1}\left(\frac{1}{K} U_x' P_{\Phi_K} U_{0.x}\right) \\ =\ & -\left(\frac{1}{n^{3/2}} X' \Phi_K\right)\left\{\left(\frac{1}{n^{1/2}} \Phi_K^{(1)} E_{-1}\right) + O_p\left(\frac{1}{K}\right)\right\} \\ \sim\ & -\int_0^1 B_x dB_e, \end{aligned}$$

using the same line of argument as that of equations (48), (49) and ultimately (57) in the proof of the main theorem in Phillips (2014).

It follows that when $\Omega_{00.x} = 0$ and $(K, n) \to \infty$ with $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$ and $\omega_{xe} = 0$, we have

$$n^2(\hat{a}_{TIV} - a) \rightsquigarrow \left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x dB_e\right) \equiv \mathcal{MN}\left(0, \omega_e^2\left(\int_0^1 B_x B_x'\right)^{-1}\right),$$

giving the stated result.

Parts (e) and (f) will be proved in a later draft.

Proof of Theorem 3

Part (a) First consider the limit behavior of the residual long run variance estimate $\hat{\omega}^2$ when $\Omega_{00.x}=0$. The residuals take the form

$$\begin{aligned}\hat{V}_{0.x} &= V_y - V_x\hat{a}_{TIV} - V_{\Delta x}\hat{f}_{TIV} = V_y - V_c\hat{\theta}_{TIV} \\ &= V_{u_{0.x}} - V_c\left(\hat{\theta}_{TIV}-\theta\right) = V_{u_{0.x}} - V_c\left(V_c'V_c\right)^{-1}V_c'V_{u_{0.x}} \\ &= \left(I_K - P_{V_c}\right)V_{u_{0.x}} = \left(I_K - P_{V_c}\right)\Phi_K u_{0.x},\end{aligned}$$

with projection matrix $P_{V_c}=V_c\left(V_c'V_c\right)^{-1}V_c'$ of rank m_x with $V_c=\left[V_x,V_{u_x}\right]$. Hence

$$\begin{aligned}\hat{\omega}^2 &= K^{-1}\hat{V}_{0.x}'\hat{V}_{0.x} = K^{-1}u_{0.x}'\Phi_K'\left(I_K-P_{V_c}\right)\Phi_K u_{0.x} \\ &= K^{-1}V_{u_{0.x}}'\left(I_K-P_{V_c}\right)V_{u_{0.x}}.\end{aligned}$$

As shown in (36) $n^{1/2}V_{u_{0.x}} \rightsquigarrow \int_0^1\tilde{\varphi}_K^{(1)}dB_e$ and setting $D_n^{1/2}=\operatorname{diag}\left[n^{3/2}I_K, n^{1/2}I_K\right]$ and $\Omega_{cc}^{1/2}=\operatorname{diag}\left[\Omega_{xx}^{1/2},\Omega_{xx}^{1/2}\right]$ we find that

$$\begin{aligned}V_cD_n^{-1/2} &= \begin{bmatrix} V_x & V_{u_x}\end{bmatrix}D_n^{-1/2} = \begin{bmatrix} n^{-3/2}\sum_{t=1}^n\tilde{\varphi}_K\left(\frac{t}{n}\right)x_t' & n^{-1/2}\sum_{t=1}^n\tilde{\varphi}_K\left(\frac{t}{n}\right)u_{xt}'\end{bmatrix} \\ &\rightsquigarrow \begin{bmatrix}\int_0^1\tilde{\varphi}_KB_x' & \int_0^1\tilde{\varphi}_KdB_x'\end{bmatrix} = \begin{bmatrix}\int_0^1\tilde{\varphi}_KW_x' & \int_0^1\tilde{\varphi}_KdW_x'\end{bmatrix}\Omega_{cc}^{1/2} =: \zeta_c\Omega_{cc}^{1/2},\end{aligned}$$

where $\zeta_c'=\left[\int_0^1\tilde{\varphi}_KW_x', \int_0^1\tilde{\varphi}_KdW_x\right]$. Hence,

$$\begin{aligned}& nV_{u_{0.x}}'\left(I_K-P_{V_c}\right)V_{u_{0.x}} \\ =\ & \left(\sqrt{n}V_{u_{0.x}}\right)'\left(\sqrt{n}V_{u_{0.x}}\right) - \left(\sqrt{n}V_{u_{0.x}}\right)'V_cD_n^{-1/2}\left(D_n^{-1/2}V_c'V_cD_n^{-1/2}\right)^{-1}D_n^{-1/2}V_c'\left(\sqrt{n}V_{u_{0.x}}\right) \\ \rightsquigarrow\ & \left(\int_0^1 dB_e\tilde{\varphi}_K^{(1)\prime}\right)\left(I_K-\zeta_c\Omega_{cc}^{1/2}\left(\Omega_{cc}^{1/2}\zeta_c'\zeta_c\Omega_{cc}^{1/2}\right)^{-1}\Omega_{cc}^{1/2}\zeta_c'\right)\left(\int_0^1 dB_e\tilde{\varphi}_K^{(1)\prime}\right) \\ =\ & \omega_{ee}^2\left(\int_0^1 dW_e\tilde{\varphi}_K^{(1)\prime}\right)\left(I_K-\zeta_c\left(\zeta_c'\zeta_c\right)^{-1}\zeta_c'\right)\left(\int_0^1\tilde{\varphi}_K^{(1)}dW_e\right).\end{aligned}$$

It follows that

$$n\hat{\omega}^2 \rightsquigarrow \frac{\omega_e^2}{K}\left(\int_0^1 dW_e\tilde{\varphi}_K^{(1)\prime}\right)\left(I_K-P_{\zeta_c}\right)\left(\int_0^1\tilde{\varphi}_K^{(1)}dW_e\right).$$

Observe that if W_x and W_e are independent (i.e., $\omega_{xe}=0$) then

$$\left(I_K-P_{\zeta_c}\right)\left(\int_0^1\tilde{\varphi}_K^{(1)}dW_e\right) \equiv \mathcal{MN}\left(0,\left(I_K-P_{\zeta_c}\right)\left(\int_0^1\tilde{\varphi}_K^{(1)}\tilde{\varphi}_K^{(1)\prime}\right)\left(I_K-P_{\zeta_c}\right)\right).$$

As is easily demonstrated $\int_0^1 \tilde{\varphi}_K^{(1)} \tilde{\varphi}_K^{(1)\prime} = I_K$ and $rank\left(I_K - P_{\zeta_c}\right) = K - 2m_x$, so that conditional on $\sigma\left\{W_x\left(r\right) : r \in [0,1]\right\}$

$$\left(\int_0^1 dW_e \tilde{\varphi}_K^{(1)\prime}\right)\left(I_K - P_{\zeta_c}\right)\left(\int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right)|_{W_x} \equiv \varkappa^2_{K-2m_x}$$

which is also the unconditional distribution, since it is independent of W_x It follows that in this case $n\hat{\omega}^2 \rightsquigarrow \frac{\omega_{ee}^2}{K} \times \varkappa^2_{K-2m_x}$.

Next, write the Wald statistic as

$$\begin{aligned} W_{TIV} &= \frac{n}{n\hat{\omega}^2}\left(\hat{a}_{TIV} - a^0\right)'\left(V_x' Q_{V_{\Delta x}} V_x\right)\left(\hat{a}_{TIV} - a^0\right) \\ &= \frac{1}{n\hat{\omega}^2}\left\{n^2\left(\hat{a}_{TIV} - a^0\right)'\right\}\left(n^{-3} V_x' Q_{V_{\Delta x}} V_x\right)\left[n^2\left(\hat{a}_{TIV} - a^0\right)\right] \end{aligned}$$

We know from Theorem 2(i), (32) and (33) that

$$\begin{aligned} n^2\left(\hat{a}_{TIV} - a\right) &\rightsquigarrow \omega_e \Omega_{xx}^{-1/2}\left(\eta_K^{W_x\prime} Q_{\xi_K^{W_x}} \eta_K^{W_x}\right)^{-1}\left(\eta_K^{W_x\prime} Q_{\xi_K^{W_x}} \int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right) \\ &= \omega_e \Omega_{xx}^{-1/2}\left(G_x' G_x\right)^{-1} G_x'\left(\int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right) \quad \text{with } G_x = Q_{\xi_K^{W_x}} \eta_K^{W_x} \\ &\equiv \mathcal{MN}\left(0, \omega_e^2 \Omega_{xx}^{-1/2}\left(G_x' G_x\right)^{-1} \Omega_{xx}^{-1/2}\right) \text{ when } \omega_{ex} = 0, \end{aligned}$$

and

$$n^{-3} V_x' Q_{V_{\Delta x}} V_x \rightsquigarrow \Omega_{xx}^{1/2} \eta_K^{W_x\prime} Q_{\xi_K^{W_x}} \eta_K^{W_x} \Omega_{xx}^{1/2} = \Omega_{xx}^{1/2} G_x' G_x \Omega_{xx}^{1/2}$$

Hence

$$\begin{aligned} &\text{Wald}_{TIV} = \frac{1}{n\hat{\omega}^2}\left\{n^2\left(\hat{a}_{TIV} - a^0\right)'\right\}\left(n^{-3} V_x' Q_{V_{\Delta x}} V_x\right)\left[n^2\left(\hat{a}_{TIV} - a^0\right)\right] \\ \rightsquigarrow\ & \omega_e^2\left[\left(\int_0^1 dW_e \tilde{\varphi}_K^{(1)\prime}\right) G_x\left(G_x' G_x\right)^{-1} \Omega_{xx}^{-1/2}\left[\Omega_{xx}^{1/2}\left(G_x' G_x\right)\Omega_{xx}^{1/2}\right]\Omega_{xx}^{-1/2}\left(G_x' G_x\right)^{-1} G_x'\left(\int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right)\right] \\ & \div\left[\frac{\omega_e^2}{K}\left(\int_0^1 dW_e \tilde{\varphi}_K^{(1)\prime}\right)\left(I_K - P_{\zeta_c}\right)\left(\int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right)\right] \\ =\ & K\frac{\left(\int_0^1 dW_e \tilde{\varphi}_K^{(1)\prime}\right) G_x\left(G_x' G_x\right)^{-1} G_x'\left(\int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right)}{\left(\int_0^1 dW_e \tilde{\varphi}_K^{(1)\prime}\right)\left(I_K - P_{\zeta_c}\right)\left(\int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right)} \\ =\ & K\frac{\left(\int_0^1 dW_e \tilde{\varphi}_K^{(1)\prime}\right) P_{G_x}\left(\int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right)}{\left(\int_0^1 dW_e \tilde{\varphi}_K^{(1)\prime}\right)\left(I_K - P_{\zeta_c}\right)\left(\int_0^1 \tilde{\varphi}_K^{(1)} dW_e\right)} \\ \equiv\ & K \times \frac{\varkappa^2_{m_x}}{\varkappa^2_{K-2m_x}} \equiv \frac{Km_x}{K-2m_x} \times \frac{\varkappa^2_{m_x}/m_x}{\varkappa^2_{K-2m_x}/\left(K-2m_x\right)} \equiv \frac{Km_x}{K-2m_x} \times F_{m_x, K-2m_x}, \text{ when } \omega_{xe} = 0, \end{aligned}$$

giving the stated results for $\omega_{xe} \neq 0$ and $\omega_{xe} = 0$.

Part (b) This part follows by existing theory (Phillips, 2014; Hwang and Sun, 2018).

7 References

Berenguer-Rico, V. and Carrion-i-Silvestre, J.L., (2011). "Regime shifts in stock–flow $I(2)$–$I(1)$ systems: the case of US fiscal sustainability," *Journal of Applied Econometrics*, 26, 298-321.

Corbae, D., Ouliaris, S. and Phillips, P. C. B. (2002). "Band spectral regression with trending data". *Econometrica,* 70, 1067 - 1110.

Engsted, T., Gonzalo, J. and Haldrup, N., 1997. Testing for multicointegration. *Economics Letters*, 56, 259-266.

Engsted, T. and Haldrup, N. (1999). "Multicointegration in stock-flow models," *Oxford Bulletin of Economics and Statistics*, 61, 237-254.

Engsted, T. and S. Johansen (1999). "Granger's representation theorem and multicointegration," pp. 200-2012 in R. Engle and H. White (eds.), *Cointegration, Causality and Forecasting : Festschrift in Honour of Clive Granger*, Oxford University Press

Escario, R., Gadea, M.D. and Sabate, M. (2012). "Multicointegration, seigniorage and fiscal sustainability. Spain 1857–2000," *Journal of Policy Modeling*, 34, 270-283.

Granger, C. W. J., and Lee, T. (1989). Investigation of production, sales and inventory relations using multicointegration and non-symmetric error correction models. *Journal of Applied Econometrics*, 4 (supplement), pp S145-S159.

Granger, C. W. J., and Lee, T. (1990). "Multicointegration". in Rhodes G. F. and Fomby T. B., *Advances in Econometrics*, 8, Greenwich CT JAI Press, pp. 71-84.

Hwang, J. and Sun, Y. (2018). "Simple, robust, and accurate F and t tests in cointegrated systems," *Econometric Theor*y, 34, 949-984.

Johansen, S. (1992). "A representation of vector autoregressive processes integrated of order 2," *Econometric Theory*, 8, 188-202.

Johansen, S. (1995). "A statistical analysis of cointegration for I (2) variables," *Econometric Theory*, 11, 25-59.

Lazarus, E., Lewis, D.J., Stock, J.H. and Watson, M.W. (2018). "HAR Inference: Recommendations for Practice," *Journal of Business & Economic Statistics*, 36, 541-559.

Lee, T.-H. (1992). "Stock-flow relationships in US housing construction," *Oxford Bulletin of Economics and Statistics*, 54, 419-430.

Lee, T.-H. (1996). Stock adjustment for multicointegrated series. *Empirical Economics*, 21, pp 633-639.

Müller, U. K. (2007), "A Theory of Robust Long-run Variance Estimation," *Journal of Econometrics*, 141, 1331–1352. [

Müller, U. K., and Watson, M. W. (2016), "Low Frequency Econometrics," Working Paper, Department of Economics, Princeton University. [567]

Phillips, P. C. B. (1986). "Understanding spurious regressions in econometrics," *Journal of Econometrics* 33, 311–340.

Phillips, P. C. B. (1988). "Multiple regression with integrated processes." In N. U. Prabhu, (ed.), *Statistical Inference from Stochastic Processes, Contemporary Mathematics* 80, 79–106.

Phillips, P. C. B. (1991). "Optimal inference in cointegrated systems," *Econometrica*, 59, 283-306.

Phillips, P. C. B. (1995). "Fully modified least squares and vector autoregression," *Econometrica,* 63, 1023-1078.

Phillips, P. C. B. (2005). "HAC Estimation by Automated Regression." *Econometric Theory,* 21, 116-142.

Phillips, P.C. B. (2007). "Unit Root Log Periodogram Regression," *Journal of Econometrics,* 138, 104-124.

Phillips, P. C. B. (2014). "Optimal estimation of cointegrated systems with irrelevant instruments," *Journal of Econometrics,* 178, 210-224.

Phillips, P. C. B. and B. E. Hansen (1990). "Statistical inference in instrumental variables regression with I(1) processes," *Review of Economic Studies* 57, 99–125.

Phillips, P.C. and Lee, J.H., 2013. Predictive regression under various degrees of persistence and robust long-horizon regression. *Journal of Econometrics*, 177, 250-264.

Phillips, P. C. B. and M. Loretan (1991). "Estimating Long-Run Economic Equilibria," *Review of Economic Studies* 59, 407–436.

Phillips, P. C. B. and V. Solo (1992). "Asymptotics for Linear Processes," *Annals of Statistics* 20, 971–1001.

Saikkonen, P. (1991). "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory* 7, 1–21.

Stock, J. H. and M. W. Watson (1993). "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica* 61, 783–821.

Sun, Y. (2018). "Comment on 'HAR Inference: Recommendations for Practice'," *Journal of Business & Economic Statistics*, 36, 565-568.

Vogelsang, T.J. and Wagner, M. (2014). "Integrated modified OLS estimation and fixed-b inference for cointegrating regressions," *Journal of Econometrics*, 178, 741-760.