ON MULTICOINTEGRATION

By

Peter C. B. Phillips and Igor Kheifets

October 2021

COWLES FOUNDATION DISCUSSION PAPER NO. 2306

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

On Multicointegration*

Peter C. B. Phillips[a,b,c,d] and Igor L. Kheifets[e]

[a]Yale University, [b]University of Auckland,
[c]University of Southampton & [d]Singapore Management University
[e]HSE University

August 29, 2021

Abstract

A semiparametric triangular systems approach shows 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 little 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 can introduce asymptotic bias into conventional methods of cointegrating regression. The present paper shows how estimation of such systems may be accomplished under multicointegration without losing the nice properties that hold under simple cointegration, including mixed normality and pivotal inference. The approach uses an extended version of high-dimensional trend IV (Phillips, 2006, 2014) estimation with deterministic orthonormal instruments that leads to mixed normal limit theory and pivotal inference in singular multicointegrated systems in addition to standard cointegrated $I(1)$ systems. Wald tests of general linear restrictions are constructed using a fixed-b long run variance estimator that leads to robust pivotal HAR inference in both cointegrated and multicointegrated cases. Simulations show the properties of the estimation and inferential procedures in finite samples, contrasting the cointegration and multicointegration cases. An empirical illustration to housing stocks, starts and completions is provided.

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

JEL Codes: C12, C13, C22

*Phillips acknowledges research support from the NSF under Grant No. SES 18-50860, a Kelly Fellowship at the University of Auckland, and an LKC Fellowship at Singapore Management University.

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 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, 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. One of the latest empirical applications of multicointegration has been to global climate change modeling, where the effects of accumulating cointegration disequilibria in temperature and surface radiation raise oceanic heat storage which leads to a multicointegrated linkage influencing global temperature (Bruns et al., 2020).

In these models the equilibrium errors (or residuals in an $I(1)$ 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 well 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 here and in related work on fully modified least squares (Kheifets and Phillips, 2021, hereafter, KP), multicointegration occurs naturally in a cointegrated $I(1)$ model whenever there is a rank deficiency in the long run conditional covariance matrix of the cointegrating equation error. 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 conditional covariance matrix induces a further long run cointegrating relation simply because the singularity implies a moving average $I(-1)$ (or higher level) component in that direction in the equation error, which leads directly to cointegration upon accumulation. This formulation of multicointegration in terms of rank deficiency in the long run conditional error covariance matrix is intuitive because it points directly to latent higher order relations in the $I(1)$ system and indicates their direction without further complications or the use of additional notation. The phenomenon has an analogue reduced rank structure in the parametric VAR model context and was noted but not 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, hereafter, EJ), 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 three 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 standard cointegrated system estimation procedures by introducing bias and non-pivotal inference. These findings are illustrated here in the case of the integrated modified least squares (IM-OLS) approach (Vogelsang and Wagner, 2014, hereafter, VW). Similarly, KP(2021) recently developed asymptotics for the fully modified least squares (FM-OLS) cointegration coefficient estimator under multicointegration, showing degenerate limit theory in general but accelerated convergence over the usual $O(n)$ rate in the direction of multicointegration, accompanied by second order bias in the limit theory. In a second contribution, it is shown that an extended version of high-dimensional trend IV (TIV) estimation with deterministic orthonormal instruments (Phillips, 2006, 2014) provides a robust approach to estimation with mixed normal limit theory and pivotal inference in singular multicointegrated systems as well as standard cointegrated $I(1)$ systems. This TIV method therefore provides a convenient IV approach to estimation and inference in $I(1)$ cointegrated systems that is robust to multicointegration without the need for pretesting. The system TIV estimator has the further advantage of a higher convergence rate under multicointegration than the FM-OLS estimator studied in KP(2021) and this estimator enables robust inference using standard Wald statistics formulated in the same way with a HAR variance matrix under both cointegration and multicointegration.

A further contribution of the paper is technical, with a group of new findings concerning the limit theory of functionals of trend transformed stationary and nonstationary variables in the case of asymptotically infinite instrument numbers. This contribution includes some new methods of developing limit theory for estimators and Wald statistics in highly complex cases involving singularities in signal matrices and partitioned regression asymptotics that require component-wise analysis or matrix normalization rather than diagonal matrix normalization to extract the correct asymptotics. These methods and results are of independent interest given recent research on large instrument numbers and deterministic transforms of variables that enable empirical investigations to focus on long run behavior.

The paper is organized as follows. The next section explains the source of multicointegration in the standard semiparametric triangular cointegrated system of $I(1)$ variables. Section 3 reconciles these origins with VAR and augmented regression representations that are commonly used in practical work. Section 4 presents and analyzes IM-OLS and TIV approaches to the estimation of cointegrated systems under conditions of multicointegration. Limit theory for both approaches is provided. Section 5 develops methods of inference using HAR methods that lead to pivotal asymptotics suited for inference in practical work. Section 6 reports some simulation results and an empirical illustration is given in Section 7. Section 8 concludes and proofs are given in the Appendix in Section 9. As noted above some proofs involve complex methods and derivations. As an aid to readers in following the derivations, a glossary of notation 9.3 for the most common functionals that appear in formulae is given for convenient reference at the end of the paper. Proofs and some additional technical results of interest, including a reverse partial summation formula, are given in the Online Supplement that accompanies the paper.

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 throughout the paper 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 $\mathbb{V}^{\text{LR}}(u_t) = \sum_{h=-\infty}^{\infty} \Gamma_h$ denote the long run variance matrix of u_t. The linear operator $D(L)$ and long run variance matrix $\mathbb{V}^{LR}(u_t) = \Omega = \sum_{h=-\infty}^{\infty} \Gamma_h = D(1)D(1)' = \sum_{j,k=0}^{\infty} D_j D_k'$ and one sided long run covariance matrix $\Delta = \sum_{h=0}^{\infty} \mathbb{E}(u_{t-h}u_t')$ 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). This full rank condition is maintained throughout the paper. 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 $\text{rank}(\Omega) = \text{rank}(\Omega_{xx}) + \text{rank}(\Omega_{00.x})$.

The case of nonsingular Ω is well studied. The case where Ω_{xx} may be singular and the regressors x_t not full rank $I(1)$ processes was studied in Phillips (1995). But the situation 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, with the exception of KP(2021), 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). Moreover, the condition for multicointegration is straightforwardly expressed in terms of the existing parameters of the $I(1)$ system without further notation or complications.

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, 1996), 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; Kheifets and Phillips, 2021) and climate change (Bruns et al., 2020). The present formulation is a general semiparametric one in the tradition of Phillips and Hansen (1990), so 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)

[1] The possibility of full system singularity with rank failure in Ω is mentioned by Engsted and Haldrup (1999, p.241) but is not analyzed. Singularity in the long run conditional variance matrix and the implications of this singularity on estimation procedures such as fully modified least squares (FM-OLS) were the subject of a Yale Take Home Examination in 2011 (http://korora.econ.yale.edu/phillips/teach/ex/553a-ex11a.pdf). That approach was analyzed in KP(2021).

may therefore be written in the following error correction model (ECM) form (Phillips, 1991)

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

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 FM-OLS or trend IV regression (Phillips, 2014) is the augmented regression

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

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.*, in the absence of fractional antipersistence[2] which is ruled out by the absolute 1-summability condition (3), leading to the representation

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

[2]Under fractional antipersistence in which $\varepsilon_{Ht} = H'u_{0.xt} = \Delta^d\varepsilon_{Ht} = (1-L)^d\varepsilon_{Ht}$ for some $d \in (0,1)$, the system (11) would be replaced by the equation $H'(1-L)^{-d}y_t = H'A(1-L)^{-d}x_t + H'F(1-L)^{-d}u_{xt} + \varepsilon_{Ht}$. Both the matrix transform H and the antipersistence parameter d would be unknown in this case, leading to further complications that are left for future research.

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 $x_0 = O_p(1)$, and provided no further level of long run degeneracy (or higher order multicointegration) is present for which $\mathbb{V}^{LR}(\varepsilon_{Ht}) = 0$. 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. Importantly, in this general framework the multicointegration parameters, notably H and $H'F = H'\Omega_{0x}\Omega_{xx}^{-1}$, are nonparametric.

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 EJ(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

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

The m dimensional $I(1)$ process z_t is therefore multicointegrated, which 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 of the $I(1)$ system.

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), viz.,

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

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

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

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 EJ(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 EJ(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

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

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 EJ(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 $\xi_\perp' = \left[\begin{array}{cc} I_{m_0} & -A \end{array}\right]$ (in the more common reduced rank notation (6) we would have $\xi_\perp' = \beta'$) so that $\xi_\perp' C(1) = 0$. The matrix $\epsilon' = \left[\begin{array}{cc} D_{x0}(1) & D_{xx}(1) \end{array}\right]$ has full rank m_x and, by reordering of coordinates as may be needed, we can assume $D_{xx}(1)$ to be nonsingular[3]. An orthogonal complement matrix of ϵ may then be constructed as

$$\epsilon_\perp = \left[\begin{array}{c} D_{x0}(1)' \\ D_{xx}(1)' \end{array}\right]_\perp = \left[\begin{array}{c} I_{m_0} \\ -D_{xx}(1)^{-1}D_{x0}(1) \end{array}\right].$$

Following the algebraic approach in EJ(1999) the derivative matrix of $C(z)$ is

$$\dot{C}(z) = \left[\begin{array}{cc} -I_{m_0} & 0 \\ 0 & 0 \end{array}\right] D(z) + \left[\begin{array}{cc} (1-z)\, I_{m_0} & A \\ 0 & I_{m_x} \end{array}\right] \dot{D}(z),$$

so that

$$\begin{aligned} \xi_\perp' \dot{C}(1)\epsilon_\perp &= -\xi_\perp' \left[\begin{array}{cc} D_{00}(1) & D_{0x}(1) \\ 0 & 0 \end{array}\right] \left[\begin{array}{c} I_{m_0} \\ -D_{xx}(1)^{-1}D_{x0}(1) \end{array}\right] \\ &= -\left[\begin{array}{cc} D_{00}(1) & D_{0x}(1) \end{array}\right] \left[\begin{array}{c} I_{m_0} \\ -D_{xx}(1)^{-1}D_{x0}(1) \end{array}\right] \\ &= -\left\{ D_{00}(1) - D_{0x}(1)\, D_{xx}(1)^{-1} D_{x0}(1) \right\}. \end{aligned}$$

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

$$D(1) = \left[\begin{array}{cc} D_{00}(1) & D_{0x}(1) \\ D_{x0}(1) & D_{xx}(1) \end{array}\right]$$

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 EJ(1999) for multicointegration in an $I(2)$ system (that $\xi_\perp' \dot{C}(1)\epsilon_\perp$ 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 EJ(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[4].

[3] 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.

[4] In particular, EJ(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 $\xi_\perp' \dot{C}(1)\epsilon_\perp$ has full rank – see their conditions (1) and (4).

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 diag $[\Omega_{00.x}, \Omega_{xx}]$. 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 validating summability conditions of Phillips and Solo (1992) that are satisfied by (3) we have

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

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 $D(1) = \left[D_0(1)' \vdots D_x(1)'\right]'$ as

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

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 BN decomposition (Corbae et al., 2002, Lemma D).

It now follows that

$$y_t \quad = \quad Ax_t + \sum_{k=-\infty}^{\infty} G_k \Delta x_{t+k} + u_{0.xt} \tag{18}$$

$$
\begin{aligned}
&= Ax_t + G(1)\Delta x_t + u_{0.xt} - \Delta \tilde{u}_{xt}, \\
&= Ax_t + G(1)\Delta x_t + u_{0.xt}^{+}, \quad \text{with } u_{0.xt}^{+} = u_{0.xt} - \Delta \tilde{u}_{xt} \qquad (19)
\end{aligned}
$$

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 $\Omega_{00.x}$. The formulation (18) extends the earlier bidirectional (leads and lags dynamic OLS regression) model of cointegration (Saikkonen, 1991; Phillips and Loretan, 1991; Stock and Watson, 1993) to the multicointegration case.

4 Estimation

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 for VAR systems Engsted et al. (1997) but multicointegration may not be suspected, pre-test analyses may not be conducted or they may lead to pre-test bias and misleading outcomes; and empirical research may be conducted using triangular cointegrated systems rather than VAR specifications. 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 these systems such as FM-OLS or dynamic OLS may continue to be employed in practical work. But 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 regression with cointegrated regressors and unrestricted VAR regression in the presence of cointegration (Phillips, 1995). Inference is correspondingly affected with further nonstandard limit distribution complications and non-pivotal limit theory in test statistics. These consequences may be analyzed[5] but are not pursued here. Instead, for reasons explained below the present paper explores the implications for integrated modified least squares (IM-OLS) estimation of such systems and develops new methods of estimation based on trend instrumental variable (TIV) methods that have clear advantages for efficient estimation and robust inference.

Another potentially interesting option that is not pursued here is the use of IVX estimation (Phillips and Magdalinos, 2009; Kostakis et al, 2014), which is known to provide robust estimation and inference in cointegrating regression and predictive regression models with many integrated or near-integrated regressors. In the present context, it can be shown that these robustness characteristics continue to hold. In particular, the same mixed normal estimation limit theory and chi-squared Wald statistic inferential limit theory applies even when $\Omega_{00.x} = 0$, provided that $\Omega_{00} > 0$, which is explained by the fact that the IVX limit theory depends only on Ω_{00}. This favorable robust outcome holds because IVX avoids endogeneity corrections by the use of mildly integrated instrument regressors that are endogenously constructed from the regressors themselves but with persistence outside the local to unity vicinity. However, the same rate of convergence is maintained in the limit theory for both singular and nonsingular cases, so that this procedure is asymptotically inefficient in both cases. The case where $\Omega_{00} = 0$ is more complex and it turns out that again IVX estimation has a mixed normal limit theory and the rate of convergence is faster than when $\Omega_{00} > 0$. But the IVX estimation method is still inefficient in this case. While this approach is certainly a worthwhile option to explore in view of its generality and robustness to departures from integration, it is not pursued here but will be examined in later work.

The present paper examines two approaches to estimation and inference which offer a rate efficient method of estimation in both singular and nonsingular cases: TIV regression (Phillips, 2005b, 2014) and IM-OLS regression (VW, 2014). To keep the analysis as brief as possible 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. Extension to the multivariate model follows usual lines but inferential analysis using Wald statistics is further complicated[6] by the need for matrix normalization to take account of differing rates of convergence in differing directions and arbitrary linear combinations of the matrix coefficients under test.

[5] KP(2021) developed the limit theory of FM-OLS estimation under rank failure of the long run conditional variance matrix of the error in the augmented regression equation.

[6] This complication is by no means trivial. Often in such cases, simplifying assumptions are made to assure no loss of rank or degrees of freedom in the limit, e.g., Andrews and Cheng (2012, 2014), and VW(2014). A general analysis of Wald statistic testing under matrix normalization without such prior requirements is examined in Magdalinos and Phillips (2019) and under ongoing development.

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 $f' = \Omega_{0x}\Omega_{xx}^{-1}$ and the conditional long run variance $\Omega_{00.x} = \Omega_{00} - \Omega_{0x}\Omega_{xx}^{-1}\Omega_{x0} \geq 0$. We will consider both the standard form of the equation where $\Omega_{00.x} > 0$ and the singular form where $\Omega_{00.x} = \Omega_{00} - \Omega_{0x}\Omega_{xx}^{-1}\Omega_{x0} = 0$. In that event, we write $u_{0.xt} = \Delta e_t$ where e_t has variance σ_e^2 and long run variance $\omega_{ee} > 0$. The latter positivity condition is not necessary but its relaxation leads to further complications involving higher order singularity (with consequential effects on multicointegration) that may be dealt with using similar methods to those developed here but these complications do not appear to be empirically relevant and are not pursued in the present work. In what follows, we consider two methods of estimation of the parameters in (20).

We start by requiring the following high-level conditions, which hold under well-known conditions (e.g., Phillips and Solo (1992)). Here and in what follows we use $\rightsquigarrow$ to signify weak convergence in the relevant probability space.

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

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

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

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor n\cdot \rfloor} \left(e_t, u_{xt}'\right)' \rightsquigarrow \left(B_e(\cdot), B_x(\cdot)'\right)' \equiv BM\left(\begin{bmatrix} \omega_{ee} & \omega_{ex} \\ \omega_{xe} & \Omega_{xx} \end{bmatrix}\right), \text{ with} \tag{23}$$

$$\begin{bmatrix} \omega_{ee} & \omega_{ex} \\ \omega_{xe} & \Omega_{xx} \end{bmatrix} > 0, \; \Delta_{xx} = \sum_{h=0}^{\infty} \mathbb{E}\left(u_{x0}u_{xh}'\right), \; \Delta_{xe} = \sum_{h=0}^{\infty} \mathbb{E}\left(u_{x0}e_h\right), \tag{24}$$

$$\Delta_{x0} = \sum_{h=0}^{\infty} \mathbb{E}\left(u_{xh}u_{0h}\right), \; \Delta_{x0}^{+} = \Delta_{x0} - \Omega_{0x}\Omega_{xx}^{-1}\Delta_{xx} = \Delta_{x0} - \Delta_{xx}f. \tag{25}$$

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 $\mathbb{CV}^{\text{LR}}\left(u_{xt}, e_t\right) = \omega_{xe} = \Omega_{xx}\alpha \neq 0$. Condition (23) allows for both $\omega_{xe} = 0$ and $\omega_{xe} \neq 0$ possibilities.

4.1.1 Integrated modified least squares (IM-OLS)

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. The method of raising the order of system integration 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. VW(2014) recently proposed an important new version of this procedure for estimating $I(1)$ systems under the standard condition $\Omega_{00.x} > 0$ and called the method integrated modified least squares (IM-OLS). The IM-OLS method has an appealing practical advantage over FM-OLS in that it involves simple least squares regression and does not require estimation of long run one-sided covariance matrices and avoids use of kernels and bandwidth choices. The method does not lead to consistent estimation of all the coefficients in the system because the equation necessarily has spurious regression components in which the error is $I(1)$ just as some of the regressors. Moreover, the approach is asymptotically inefficient relative to FM-OLS and other efficient methods of $I(1)$ system estimation. Procedures of inference are also considerably more complex than usual because a simple consistent estimator of $\Omega_{00.x}$ is not readily available, due to the partially spurious regression feature of the fitted equation and the inconsistent estimates of some of the coefficients.[7] The present paper makes a separate contribution to IM-OLS inference by providing a new pivotal approach to inference in the cointegration case. The method developed here makes use of a sandwich-form asymptotic covariance variance matrix estimator that can be constructed in the usual way and which applies in the same form for both cointegrated and multicointegrated systems.

Using capitals as before to denote partial summation, write $Y_t = \sum_{s=1}^{t} y_s$, $X_t = \sum_{s=1}^{t} x_s$, and $U_{0.xt} = \sum_{s=1}^{t} u_{0.xs}$. The transformed system (20), up to initial conditions (in particular, taking $e_0 = 0$), is then

$$\begin{aligned} Y_t &= a'X_t + f'x_t + e_t^+ && (26) \\ e_t^+ &= e_t \mathbf{1}\{\Omega_{00.x} = 0\} + U_{0.xt}\mathbf{1}\{\Omega_{00.x} > 0\}, && (27) \end{aligned}$$

a formulation that covers both singular and non-singular cases. Applying least squares regression to (26) gives the IM-OLS error of estimation of the cointegrating coefficients

$$\hat{a} - a = (X'Q_x X)^{-1} X'Q_x e^+, \ \hat{f} - f = (x'Q_X x)^{-1} x'Q_X e^+ \tag{28}$$

in standard partitioned matrix regression notation with orthogonal projector $Q_x = x(x'x)^{-1}x'$, where $X' = [X_1, ..., X_n]$, $x' = [x_1, ..., x_n]$, and $e^{+\prime} = \left[e_1^+, ..., e_n^+\right]$. There are no modifications in the OLS estimation procedure, just the use of least squares on the partial summed augmented system (26). In practice it is useful to include a fitted intercept in regressions on (26), which

[7] VW(2014) examine three methods as a basis for inference with IM-OLS: (i) using an FM-OLS regression estimate of the long run error variance from the original $I(1)$ cointegrating equation, thereby avoiding use of residuals from the partially spurious IM-OLS regression, but introducing another estimation procedure that is based on the original model; (ii) differencing the IM-OLS residuals and using these to construct a long run variance estimate; (iii) bias-adjusting the residuals in (ii) by means of a further regression and using these adjusted residuals for long run variance estimation. Each of these methods relies on information that there is no multicointegration in the system and the properties of these methods under multicointegration are unexplored. The present work provides a new approach to inference in IM-OLS regression that uses a standard form of HAR variance matrix estimation that leads to pivotal inference in a cointegrated system.

takes care of any initialization effects if $e_0 \neq 0$ in the singular case where $e_t^+ = e_t \mathbf{1}\{\Omega_{00.x} = 0\}$. For such regressions, all the following results are modified by demeaning the limit processes in the usual manner (Park and Phillips, 1988, 1989). To avoid notational clutter we do not make this modification in the limit theory or derivations that follow so that the results apply when $e_0 = 0$.

Standard weak convergence methods for nonstationary regression lead to the following asymptotics as $n \to \infty$, where we focus on estimation of the cointegrating vector a. The result given in Theorem 1(i) below for the case $\Omega_{00.x} > 0$ corresponds to the finding in VW(2014, Theorem 2), with a somewhat simpler form of the limit theory. The result given in (ii) shows the effects of the presence of multicointegration on the IM-OLS estimator. These results are not directly needed in our subsequent development. But they are useful for comparative purposes and in detailing some of the challenges involved in robust estimation and inference in cointegrated/multicointegrated systems.

Theorem 1 (IM-OLS Estimation)

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

(i) $n(\hat{a} - a) \rightsquigarrow \mathcal{A}_{X.x}^{-1} \int_0^1 B_{X.x} B_{0.x} = \mathcal{A}_{X.x}^{-1} \int_0^1 \overrightarrow{B_{X.x}} dB_{0.x}$

$$\equiv \mathcal{MN}\left(0, \Omega_{00.x} \mathcal{A}_{X.x}^{-1} \left(\int_0^1 \overrightarrow{B_{X.x}}\,\overrightarrow{B_{X.x}}'\right) \mathcal{A}_{X.x}^{-1}\right),$$

where $\mathcal{A}_{X.x} = \int_0^1 B_{X.x} B_{X.x}'$, $\overrightarrow{B_{X.x}}(r) = \int_r^1 B_{X.x}$, $B_{X.x}(r) = B_X(r) - \int_0^1 B_X B_x' \left(\int_0^1 B_x B_x'\right)^{-1} B_x(r)$ *and* $B_X(r) = \int_0^r B_x$.

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

(ii) $n^2(\hat{a} - a) \rightsquigarrow \mathcal{A}_{X.x}^{-1} \left\{\int_0^1 B_{X.x} dB_e - \left(\int_0^1 B_X B_x'\right) \left(\int_0^1 B_x B_x'\right)^{-1} \Delta_{xe}\right\}.$

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 (ii) to $O(n^2)$ but introduces nonstandard asymptotics with second order bias effects from endogeneity (correlation between the Brownian motions B_x and B_e when $\omega_{xe} \neq 0$) and serial dependence (Δ_{xe}). Thus, raising the integration order of the system fails to resolve these standard problems of least squares asymptotic theory. Even when $\omega_{xe} = 0$, bias remains whenever $\Delta_{xe} \neq 0$. When $\Omega_{00.x} > 0$, mixed normal asymptotic theory applies as for other procedures like FM-OLS but with some efficiency loss and some complexities in inference, as discussed in VW(2014) and further complexities in the estimation of $\Omega_{00.x}$ due to the partial spurious regression feature of (26).

Thus, while this approach of raising the integration order leads to a viable estimation and testing methodology in nonsingular systems, it does not provide a robust methodology for singular, multicointegrated models, at least without the introduction of new modifications to address endogeneity and serial dependence. Akin to other methods like FM-OLS, IM-OLS does not provide a generally robust estimation methodology for cointegrated systems that encompasses multicointegration. A new approach is needed to accomplish this.

4.1.2 Trend Instrumental Variable Estimation

The approach we develop here is based on the trend IV (TIV) method of Phillips (2014)[8] which employs orthonormal (ON) deterministic trend functions as instrumental variables for the regressors in (20). These ON instruments are designed to transform the system so that its long run properties are brought into primary focus both for regression estimation (Phillips, 1998) and for long run variance matrix estimation Phillips (2005b); Müller (2007). These methods have recently become popular in examining various properties of long run relations among time series variables (e.g., Phillips (2005a); Müller and Watson (2018); Hwang and Sun (2018)) and have numerous empirical applications as revealed in these studies.

In the TIV method of estimating cointegrating equations such as (20), deterministic instrumental variables $\left\{\varphi_k\left(\frac{t}{n}\right)\right\}_{k=1}^{K}$ are employed, where $\{\varphi_k(r)\}_{k=1}^{\infty}$ are orthonormal basis functions of $L_2[0,1]$ and K is allowed to pass to infinity as $n \to \infty$. This approach is high-dimensional TIV estimation. An alternate version of this method is based on a fixed number K of orthonormal trend instruments and is used in recent work by Hwang and Sun (2018). We call this method the fixed-K approach of TIV regression. Various classes of orthonormal functions may be used in these regressions without materially affecting the limit theory or finite sample performance, as demonstrated in Phillips (2014) and Hwang and Sun (2018). The latter paper shows a particular advantage in terms of F and t distribution limit theory for conventional test statistics of coefficient restrictions, which can enhance inference in finite samples in standard cointegrated systems. This advantage has received wider attention recently (Lazarus et al., 2018). But as shown later in the current work, the fixed-K approach does not deal as effectively with multicointegrated systems.

In what follows, we let $\tilde{\varphi}_K(r) = (\varphi_1(r), ..., \varphi_K(r))'$, and $\Phi_K' = [\tilde{\varphi}_{K1}, ..., \tilde{\varphi}_{Kn}]$ where $\tilde{\varphi}_{Kt} = \tilde{\varphi}_K\left(\frac{t}{n}\right) = \left[\varphi_1\left(\frac{t}{n}\right), ..., \varphi_K\left(\frac{t}{n}\right)\right]'$. The projector matrix onto the space of the instruments is $P_{\Phi_K} = \Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'$. For trigonometric orthonormal polynomials we have $n^{-1}\Phi_K'\Phi_K = I_K + O\left(\frac{1}{n}\right)$, as shown in Phillips (2005b, Lemma A) when $\varphi_k(r) = \sqrt{2}\sin\left\{\left(k-\frac{1}{2}\right)\pi r\right\}$, so that $P_{\Phi_K} \sim n^{-1}\Phi_K\Phi_K'$. TIV estimation of (20) is then asymptotically equivalent to simple least squares regression on the linearly transformed K-dimensional system

$$V_y = V_x a + V_{\Delta x} f + V_{u_{0.x}}, \tag{29}$$

where we use the general notation $V_c = \Phi_K' c = \sum_{t=1}^{n} \tilde{\varphi}_{Kt} c_t'$ for the trigonometric transform of a time series c_t. The resulting coefficient estimates of (29) have the following form in standard

[8] The TIV approach was originally proposed by the author in a York University Workshop conference presentation given in 2003. The same paper was presented in the Faro Time Series Econometrics Conference 2005 and distributed as a Cowles Foundation Discussion Paper (Phillips, 2006). That paper also introduced the concept of a trend likelihood associated with the low frequency components of a time series obtained by fitted regression on a number of deterministic orthonormal regressors. Phillips (2005b) introduced the related idea of trend coordinates based on these fitted regression components to study long run covariability among trending time series, a subject that has been extensively investigated recently by Müller and Watson (2018). The approach has earlier origins in band spectral regression (Hannan, 1963; Engle, 1974; Corbae et al., 2002) in the frequency domain.

partitioned regression notation

$$\hat{a}_{TIV} - a = \left(V_x' Q_{V_{\Delta x}} V_x\right)^{-1} V_x' Q_{V_{\Delta x}} V_{u_{0.x}}, \tag{30}$$

$$\hat{f}_{TIV} - f = \left(V_{\Delta x}' Q_{V_x} V_{\Delta x}\right)^{-1} V_{\Delta x}' Q_{V_x} V_{u_{0.x}}. \tag{31}$$

This least squares 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 $n \to \infty$. It is asymptotically equivalent to fixed-K TIV.

The approach we suggest here is designed to robustify the TIV procedure to the presence of multicointegration and singularity. The idea is to apply TIV regression to the following augmented regression form of (26)

$$Y_t = a' X_t + f' x_t + g' \Delta x_t + e_t^+ = a' X_t + f' x_t + g' u_{xt} + e_t^+, \tag{32}$$

where the additional (redundant) regressor Δx_t is included with coefficient $g = 0$ and the regression error is $e_t^+ = e_t \mathbf{1}\{\Omega_{00.x} = 0\} + U_{0.xt} \mathbf{1}\{\Omega_{00.x} > 0\}$ as before. Thus, this time aggregated version of the model is augmented by the inclusion of the additional regressor Δx_t, analogous to the original system (8). As before, in practical work it is useful to include a fitted intercept in (32), which is innocuous but takes care of initialization effects in the singular case where $e_t^+ = e_t \mathbf{1}\{\Omega_{00.x} = 0\}$ and $e_0 \neq 0$. Again, the limit theory is simply adjusted to employ deviations from means for the relevant stochastic processes, which for ease of notation is not done here.

In observation form, we write (32) as

$$Y = [X, C_x]\,\gamma + e^+, \text{ with } \gamma' = (a', f', g') =: (a', \ell') \tag{33}$$

and

$$C_x' = [c_{x1}, ..., c_{xn}] = \begin{bmatrix} x_1 & \cdots & x_n \\ u_{x1} & \cdots & u_{xn} \end{bmatrix} =: \begin{bmatrix} x' \\ u_x' \end{bmatrix}.$$

Equation (32) may, of course, also be estimated by direct application of least squares, leading to a form of augmented IM-OLS regression with

$$\hat{a} = \left(X' Q_W X\right)^{-1} \left(X' Q_W Y\right),$$

where $Q_W = Q_x - Q_x u_x \left(u_x' Q_x u_x\right)^{-1} u_x' Q_x$ in standard notation. The asymptotic theory for such direct least squares estimation is derived in the Appendix. For the cointegration case with $\Omega_{00.x} > 0$ we find that

$$n\left(\hat{a} - a\right) \rightsquigarrow \mathcal{A}_{X.x}^{-1} \int_0^1 B_{X.x} B_{0.x} \equiv \mathcal{MN}\left(0, \Omega_{00.x} \times \Omega_{XX}\right) \tag{34}$$

with

$$\Omega_{XX} = \mathcal{A}_{X.x}^{-1} \left(\int_0^1 \overrightarrow{B_{X.x}}(r)\, \overrightarrow{B_{X.x}}(r)' \, dr \right) \mathcal{A}_{X.x}^{-1}, \tag{35}$$

which is identical to the limit distribution of the IM-OLS estimator. Thus, inclusion of the surplus and irrelevant regressor u_{xt} in the fitted model (32) has no effect on the limit theory under cointegrating regression and the same limit theory as in Theorem 1(i) applies. In the multicointegrated case, we find that

$$n^2(\hat{a}-a) \rightsquigarrow \mathcal{A}_{X.x}^{-1}\left\{\int_0^1 B_{X.x}dB_e - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\Delta_{xe} \right.$$
$$\left. - \left[\int_0^1 B_{X.x}dB_x' - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\Delta_{xx}\right]\Sigma_{xx}\sigma_{xe}\right\}, \tag{36}$$

where $\Sigma_{xx} = \mathbb{E}(u_{xt}u_{xt}')$, $\sigma_{xe} = \mathbb{E}(u_{xt}e_t)$ and the one-sided long run covariances are given in (24). But inference in this system is further complicated by the fact that in the partially spurious regression (32) the coefficients of the additional regressors $\Delta x_t = u_{xt}$ are inconsistently estimated in both cointegration and multicointegration cases. So, the use of direct least squares on the augmented system does not resolve the endogeneity and serial dependence issues of IM-OLS and nuisance parameter dependencies in the limit. Some form of fully modified version of least squares regression on (32) might be employed to improve asymptotic properties but this avenue leads to further difficulties and will not be pursued in what follows.

Instead, we proceed with the analysis of TIV estimation of the augmented system. The TIV estimator of a in (32) has the form

$$\hat{a}_{TIV} = \arg\min_a (Y-Xa)' R_K (Y-Xa) = (X'R_KX)^{-1}(X'R_KY) \tag{37}$$

where the projector matrix is $R_K = P_{\Phi_K} - P_{\Phi_K}C_x(C_x'P_{\Phi_K}C_x)^{-1}C_x'P_{\Phi_K}$ and $Y' = [Y_1, ...Y_n]$. The TIV estimation procedure projects all the variables in the augmented system (32) onto the deterministic instruments using the projector P_{Φ_K}. For fixed K, this approach is, as above in (29), asymptotically equivalent to least squares regression on the transformed system

$$V_Y = V_X a + V_x f + V_{\Delta x} g + V_{e^+} =: V_X a + V_C \ell + V_{e^+}, \tag{38}$$

where we employ the notation $V_Z = \Phi_K' Z$ for an observation matrix Z. Standard partitioned least squares regression on (38) leads to the following estimator of a

$$\hat{a}_{fTIV} - a = (V_X' Q_{V_C} V_X)^{-1} V_X' Q_{V_C} V_{e^+}, \tag{39}$$

giving the fixed-K trend IV (fTIV) estimate.

The results that follow provide the asymptotic theory for TIV estimation with fixed-K and as $K \to \infty$ in both $\Omega_{00.x} > 0$ and $\Omega_{00.x} = 0$ cases. The proofs involve new complications due to the presence of the redundant regressor Δx_t in the fitted equation, the partially spurious nature of the regression equation when $\Omega_{00.x} > 0$, and the impact of singularity when $\Omega_{00.x} = 0$.

New asymptotic theory is provided to address these complications. The analysis is particularly difficult when $K \to \infty$ as $n \to \infty$ and the development of the asymptotic theory of inference in the following section involves new methods and results. But the final limit theory is satisfyingly simple for both the singular and nonsingular $\Omega_{00.x}$ cases. The result for fixed-K TIV estimation is given in Theorem 2. The main result is given in Theorem 3 for TIV estimation when $K \to \infty$.

This approach leads to mixed normal limit distribution theory in $\Omega_{00.x} > 0$ and $\Omega_{00.x} = 0$ cases, therefore providing a basis for robust estimation and inference in cointegrated/multicointegrated systems even when the presence of multicointegration is unknown a priori.

Theorem 2 (TIV estimation with fixed K)

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

(iii) $n\left(\hat{a}_{fTIV} - a\right) \rightsquigarrow S_K' \Psi_{0.xK} \equiv \mathcal{MN}\left(0, \Omega_{00.x} S_K' \left(\int_0^1 \int_0^1 (r \wedge s)\, \tilde{\varphi}_K(r)\, \tilde{\varphi}_K(s)' \, drds\right) S_K\right),$

where $S_K = J_K \mu_K \left(\mu_K' J_K \mu_K\right)^{-1}$, $\mu_K = \int_0^1 \tilde{\varphi}_K B_X'$, $\Psi_{0.xK} = \int_0^1 \tilde{\varphi}_K B_{0.x}$, $J_K = Q_{\xi_K} - Q_{\xi_K} \eta_K \left(\eta_K' Q_{\xi_K} \eta_K\right)^{-1} \eta_K' Q_{\xi_K}$, $\eta_K = \int_0^1 \tilde{\varphi}_K(r)\, B_x(r)'\, dr$, $\xi_K = \int_0^1 \tilde{\varphi}_K(r)\, dB_x(r)'$, *and* $Q_{\xi_K} = I_K - \xi_K \left(\xi_K' \xi_K\right)^{-1} \xi_K'$.

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

(iv) $n^2\left(\hat{a}_{fTIV} - a\right) \rightsquigarrow S_K' \psi_{eK},$

where $\psi_{eK} = \int_0^1 \tilde{\varphi}_K dB_e$. When $\omega_{ex} = 0$, the limit distribution is mixed normal and

(iv)* $n^2\left(\hat{a}_{fTIV} - a\right) \rightsquigarrow \mathcal{MN}\left(0, \omega_{ee}\left(\mu_K' R_K \mu_K\right)^{-1}\right).$

Theorem 3 (TIV estimation with $K \to \infty$)

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

(v) $n\left(\hat{a}_{TIV} - a\right) \rightsquigarrow \mathcal{A}_{X.x}^{-1}\left(\int_0^1 B_{X.x} B_{0.x}\right) = \mathcal{A}_{X.x}^{-1} \int_0^1 \overrightarrow{B_{X.x}} dB_{0.x}$

$$\equiv \mathcal{MN}\left(0, \Omega_{00.x} \mathcal{A}_{X.x}^{-1} \int_0^1 \overrightarrow{B_{X.x}}\, \overrightarrow{B_{X.x}}' \mathcal{A}_{X.x}^{-1}\right).$$

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

(vi) $n^2\left(\hat{a}_{TIV} - a\right) \rightsquigarrow \mathcal{A}_{X.x}^{-1}\left(\int_0^1 B_{X.x} dB_{e.x}\right) \equiv \mathcal{MN}\left(0, \omega_{ee.x} \mathcal{A}_{X.x}^{-1}\right),$

where $B_{e.x}(r) = B_e(r) - \omega_{ex} \Omega_{xx}^{-1} B_x(r) \equiv BM(\omega_{ee.x})$ where $B_{e.x}$ is independent of the Brownian motion B_x and $\omega_{ee.x} = \omega_{ee} - \omega_{ex} \Omega_{xx}^{-1} \omega_{xe}$.

As expected, in both Theorems 2 and 3 the limit distributions differ for the two cases $\Omega_{00.x} = 0$ and $\Omega_{00.x} > 0$. For Theorem 3 we employ the expansion rate condition on the instrument number $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$. The same condition was used in Phillips (2014) and facilitates the joint limit theory $(K, n) \to \infty$.

The non-singular TIV regression has the usual $O(n)$ convergence rate for cointegrating regressions when $\Omega_{00.x} > 0$ in both fixed K and $K \to \infty$ cases. Just as in the standard cointegrating regression theory with $\Omega_{00.x} > 0$ mixed normal limit theory applies, as it does for other methods of estimation such as FM-OLS regression. Noticeably, when $K \to \infty$ as $n \to \infty$, Theorem 3 (v) shows that TIV reproduces the limit theory of the IM-OLS estimator given in Theorem 1 (i). As remarked above in connection with (35), IM-OLS may also be applied directly to the augmented model (32) with the redundant regressor Δx_t but without the long run transforms and again the same limit theory applies as in Theorem 1 (i) and Theorem 2 (iv). So the presence of the redundant regressor Δx_t in the fitted regression model (32) has no asymptotic effects, at least when $\Omega_{00.x} > 0$. There are, however, non-trivial effects on the estimated residuals from the use of IM-OLS on the augmented system (32) that make inference difficult.

The singular case with $\Omega_{00.x} = 0$ is much more intriguing. First, rates of convergence rise to $O\left(n^2\right)$ as they do for IM-OLS. But the limit theory for TIV is much simpler because the long run transforms are effective in focusing attention on long run properties. Second, the TIV regression is successful in removing both endogeneity and serial correlation biases in both singular and nonsingular cases under joint convergence when $K \to \infty$ as $n \to \infty$. Third, the limit theory is mixed normal and conducive to pivotal inference in both cases, even though the rates of convergence are different for singular and nonsingular systems. Fourth, the mixed normal limit theory in (vi) may be written in standardized form as

$$\mathcal{MN}\left(0, \omega_{ee.x}\Omega_{xx}^{-1/2}\mathcal{A}_{W,X.x}^{-1}\Omega_{xx}^{-1/2}\right) \equiv \omega_{ee.x}^{-1/2}\Omega_{xx}^{-1/2} \times \mathcal{MN}\left(0, \mathcal{A}_{W,X.x}^{-1}\right), \tag{40}$$

with $\mathcal{A}_{W,X.x} = \int_0^1 W_{X.x}W_{X.x}'$, since by simple matrix scale manipulations we have the representation

$$\begin{aligned} B_{X.x}\left(r\right) &= B_X\left(r\right) - \int_0^1 B_X B_x' \left(\int_0^1 B_x B_x'\right)^{-1} B_x\left(r\right) \\ &= \Omega_{xx}^{1/2}\left\{W_X\left(r\right) - \int_0^1 W_X W_x' \left(\int_0^1 W_x W_x'\right)^{-1} W_x\left(r\right)\right\} =: \Omega_{xx}^{1/2} W_{X.x}, \end{aligned} \tag{41}$$

where $B_x = \Omega_{xx}^{1/2}W_x$, $B_X\left(r\right) = \Omega_{xx}^{1/2}\int_0^r W_x$, and $W_x \equiv BM\left(I_{m_x}\right)$. The limit distribution (40) is then a matrix scaled form of a mixed normal distribution that depends only on functionals of standard Brownian motion. Importantly, the convergence rate of TIV regression is faster than that of FM-OLS in the multicointegrated case where the rate does not achieve $O(n^2)$ – see KP(2021).

Theorems 2 and 3 highlight differences in TIV estimation between the fixed K and high-dimensional $K \to \infty$ cases. For the fixed K case. TIV does not fully remove endogeneity bias as the limiting error transform $\psi_{eK} = \int_0^1 \tilde{\varphi}_K dB_e$ in the limit distribution (iv) remains correlated with the regressor variable limiting transforms $(\mu_K, \eta_K, \xi_K) = \left(\int_0^1 \tilde{\varphi}_K B_X', \int_{r=0}^1 \tilde{\varphi}_K B_x' dr, \int_0^1 \tilde{\varphi}_K dB_x\right)$ when the long run covariance $\omega_{ex} \neq 0$. But when $\omega_{ex} = 0$ and K is fixed the TIV estimator $\hat{a}_{TIV}$ does have mixed normal limit theory, given by

$$n^2\left(\hat{a}_{TIV} - a\right) \rightsquigarrow \mathcal{MN}\left(0, \omega_{ee}\left(\mu_K' R_K \mu_K\right)^{-1}\right), \tag{42}$$

which may be written in standardized Brownian motion form, analogous to (40) in this case. So under the long run orthogonality condition $\omega_{ex} = 0$ TIV estimation with fixed K instruments provides robust estimation and is effective in pivotal inference. But in the general case where the long run covariance $\mathbb{CV}^{\text{LR}}\left(e_t, u_{xt}\right) = \omega_{ex} \neq 0$ and there is long run endogeneity in the singular multicointegrated model, the limit distribution in (v) for the fixed K case is no longer mixed normal.

These results show the key advantage of high-dimensional trend IV regression on the augmented aggregated system (32). The limit theory of TIV regression is mixed normal in *both* non-singular and singular cases when $K \to \infty$ as $n \to \infty$. The method therefore provides a useful foundation for a robust approach to estimation and inference about cointegrating coefficients in both cointegrated and multicointegrated systems.

Our primary focus in this paper is on the estimation of the cointegrating vector a, the key linkage parameter in an $I(1)$ cointegrated system and to develop a new procedure that is robust to the possible presence of multicointegration. In cases where multicointegration is known to be present, or at least strongly suspected, the methods developed in this paper also provide for estimation of the multicointegration vector f.

For completeness but to keep the present paper within manageable length we give only a brief outline here of TIV estimation of f. For this purpose, it is convenient to use a different partitioned model representation than (33). In observation form, we write (32) as

$$Y = [x, C_X]\,\gamma_x + e^+, \text{ with } \gamma_x' = (f', a', g') =: (f', h') \tag{43}$$

and

$$C_X' = [c_{X1}, ..., c_{Xn}] = \begin{bmatrix} X_1 & \cdots & X_n \\ u_{x1} & \cdots & u_{xn} \end{bmatrix} =: \begin{bmatrix} X' \\ u_x' \end{bmatrix}.$$

The TIV estimator of the multicointegration parameter f is then

$$\hat{f}_{TIV} = \arg\min_f (Y - xf)' S_K (Y - xf) = (x' S_K x)^{-1} (x' S_K Y),$$

where the projection matrix is now $S_K = P_{\Phi_K} - P_{\Phi_K} C_X (C_X' P_{\Phi_K} C_X)^{-1} C_X' P_{\Phi_K}$. The following limit theory for $\hat{f}_{TIV}$ extends Theorem 3 to the multicointegration parameter.

Theorem 4 (TIV multicointegration parameter estimation with $K \to \infty$)
When $\Omega_{00.x} = 0$, (23) holds, and $(K, n) \to \infty$ with $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$, $\hat{f}_{TIV} \to_p f$ and

$$n\left(\hat{f}_{TIV} - f\right) \rightsquigarrow \mathcal{A}_{x.X}^{-1}\left(\int_0^1 B_{x.X} dB_{e.x}\right) \equiv \mathcal{MN}\left(0, \omega_{ee.x}\mathcal{A}_{x.X.}^{-1}\right), \tag{44}$$

in which $\mathcal{A}_{x.X} = \int_0^1 B_{x.X} B_{x.X}'$, $B_{x.X}(r) = B_x(r) - \int_0^1 B_x B_X' \left(\int_0^1 B_X B_X'\right)^{-1} B_X(r)$, *and* $B_{e.x}(r) = B_e(r) - \omega_{ex}\Omega_{xx}^{-1} B_x(r) \equiv BM(\omega_{ee.x})$, *where* $B_{e.x}$ *and* $\omega_{ee.x}$ *are defined in Theorem 3.*

The convenient mixed normal limit theory (44) enables pivotal inference in a similar way to that for the cointegration estimator $\hat{a}$. The $O(n)$ convergence rate matches that of simple cointegration estimation without multicointegration. Moreover, the high-dimensional TIV estimator $\hat{f}_{TIV}$ has analogous optimal estimation properties for the multicointegration parameter f as those of the TIV cointegration estimator in a semiparametric cointegrated system without multicointegration (Phillips, 1991, 2014). These properties will be analyzed in full in later work, as will the proof of Theorem 4 which is lengthy and complex.

5 Inference

Theorems 1-3 show that both TIV and IM-OLS methods provide consistent and asymptotically mixed normal estimation procedures which might be expected to form a basis for inference in the standard $I(1)$ cointegrating regression model with nonsingular $\Omega_{00.x} > 0$. But when $\Omega_{00.x} > 0$ the augmented system (32) is a partially spurious regression, just like the original aggregated system (26) with $I(1)$ regressors and an $I(1)$ error. The spurious nature of this regression complicates inference and requires special methods to estimate the long run variance $\Omega_{00.x}$ in constructing Wald tests. Moreover, when $\Omega_{00.x} = 0$, IM-OLS suffers from asymptotic second order bias and limit theory that is unsuited to pivotal inference, thereby failing to resolve endogeneity and serial correlation bias problems in the limit. In what follows we therefore concentrate on the TIV approach to testing.

More specifically, consider a Wald test of the linear hypothesis $\mathcal{H}_0 : Ha = h$ about the cointegrating vector a where H is $q \times m_x$ of rank q and h is a q-vector. Just as in estimation, the problem of inference is complicated by the fact that it is unknown a priori whether the system is singular or not in the absence of prior information or pre-testing. Robust inference therefore requires that the same approach be employed in both cases since $\Omega_{00.x}$ is, of course, unknown. For this purpose it is convenient to employ a sandwich form in estimating the covariance matrix metric for the Wald statistic in order to deal in a comprehensive way with the different types of temporal dependencies that arise in the nonsingular $\Omega_{00.x} > 0$ and singular $\Omega_{00.x} = 0$ cases. This matrix can be constructed in a general way by using the form of the TIV estimate $\hat{a}_{TIV}$. In view of (33) and (37), $\hat{a}_{TIV}$ satisfies

$$\hat{a}_{TIV} - a = (X'R_K X)^{-1} (X'R_K e^+) = G_K \Phi_K' e^+ = G_K \sum_{t=1}^{n} \tilde{\varphi}_K \left(\frac{t}{n}\right) e_t^+ \tag{45}$$

where $R_K = P_{\Phi_K} - P_{\Phi_K} C_x (C_x' P_{\Phi_K} C_x)^{-1} C_x' P_{\Phi_K}$, so that

$$H(\hat{a}_{TIV} - a) = HG_K \Phi_K' e^+ = HG_K \left(\sum_{t=1}^{n} \tilde{\varphi}_K \left(\frac{t}{n}\right) e_t^+ \right),$$

in which the coefficient matrix G_K has the form

$$G_K = (X'R_K X)^{-1} \left\{ X'\Phi_K (\Phi_K'\Phi_K)^{-1} - X'P_{\Phi_K} C_x (C_x' P_{\Phi_K} C_x)^{-1} C_x' \Phi_K (\Phi_K'\Phi_K)^{-1} \right\}, \tag{46}$$

and $\Phi_K' e^+ = \sum_{t=1}^{n} \tilde{\varphi}_K \left(\frac{t}{n}\right) e_t^+$ is the transformed error vector in the model after projection on the instruments Φ_K. We may estimate the residuals e_t^+ from the fitted TIV regression giving

$$\begin{aligned} \hat{e}_t^+ &= Y_t - \hat{a}_{TIV}' X_t - \hat{f}_{TIV}' x_t - \hat{g}_{TIV}' \Delta x_t \\ &= e_t^+ - (\hat{a}_{TIV} - a)' X_t - \left(\hat{f}_{TIV} - f\right)' x_t - (\hat{g}_{TIV} - g)' u_{xt}. \end{aligned}$$

and construct the kernel estimates

$$\hat{V}_{Kn} = \sum_{j=-M+1}^{M} k\left(\frac{j}{M}\right) \frac{1}{n} \sum_{1 \le t, t+j \le n} \tilde{\varphi}_K \left(\frac{t}{n}\right) \tilde{\varphi}_K \left(\frac{t+j}{n}\right)' \hat{e}_t^+ \hat{e}_{t+j}^+, \tag{47}$$

$$\hat{\omega}^2_{e^+} = \sum_{j=-M+1}^{M} k\left(\frac{j}{M}\right)\frac{1}{n}\sum_{1\le t,t+j\le n}\hat{e}^+_t\hat{e}^+_{t+j}, \tag{48}$$

as if we were estimating a long run variance matrix of $\tilde{\varphi}_K\left(\frac{t}{n}\right)e^+_t$ and long run variance of e^+_t, thereby ignoring the spurious nature of the regression when $\Omega_{00.x} > 0$.

The lag kernel function $k(\cdot): \mathbb{R} \to [0,1]$ used in (47) and (48) is assumed to be a symmetric, piecewise smooth density with $k(x) = 0$ for $|x| > 1$, and $\int_{-1}^{1} k(x)\,dx = 1$. In the case of standard HAC estimation, the lag truncation parameter M is assumed to satisfy $\frac{1}{M} + \frac{M}{n} \to 0$ as $n \to \infty$. In the case of HAR inference with a fixed-b setting leading to $M = bn$, we use the notation $k_b(x) = k(\frac{x}{b})$ and correspondingly define the HAR kernel estimator as

$$\hat{V}_{bKn} = \sum_{j=-n+1}^{n-1} k_b\left(\frac{j}{n}\right)\frac{1}{n}\sum_{1\le t,t+j\le n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t+j}{n}\right)'\hat{e}^+_t\hat{e}^+_{t+j}. \tag{49}$$

With these components we can construct the following HAC and HAR Wald statistics in conventional form as follows

$$\text{Wald}_{TIV} = (H\hat{a}_{TIV} - h)'\left[HG_K\left(n\hat{V}_{Kn}\right)G'_K H'\right]^{-1}(H\hat{a}_{TIV} - h), \tag{50}$$

$$\text{Wald}_{TIV,b} = (H\hat{a}_{TIV} - h)'\left[HG_K\left(n\hat{V}_{bKn}\right)G'_K H'\right]^{-1}(H\hat{a}_{TIV} - h). \tag{51}$$

The regression error is $e^+_t = e_t\mathbf{1}\{\Omega_{00.x} = 0\} + U_{0.xt}\mathbf{1}\{\Omega_{00.x} > 0\}$. So the asymptotic properties of (47), (48) and therefore both Wald test statistics Wald_{TIV} and $\text{Wald}_{TIV,b}$ depend on the asymptotic behavior of the residuals $\hat{e}^+_t$, the long run error variance estimate $\hat{\omega}^2_{e^+}$, and the long run variance matrix estimates $\hat{V}_{Kn}$ and $\hat{V}_{bKn}$ associated with the transformed error components $\tilde{\varphi}_K\left(\frac{t}{n}\right)e^+_t$.

Two forms of TIV inference can be considered, corresponding to fixed-K and $K \to \infty$ cases, just as in estimation. A disadvantage of the the fixed-K approach is that the partially spurious nature of the fitted regression carries the inconsistencies of the estimates $\left(\hat{f}_{TIV}, \hat{g}_{TIV}\right)$ into the regression residuals $\hat{e}^+_t$ and their $I(1)$ character in the usual $\Omega_{00.x} > 0$ case. This leads to divergence of statistical tests as $n \to \infty$ under the null hypothesis, just as in standard spurious regression limit theory (Phillips, 1986). Even with the use of sandwich formulae and HAC estimators such as (47) the divergence rate of the Wald test for fixed K is $O_p\left(\frac{n}{M}\right)$, as shown in the proof of Theorem 5 below[9]. This divergence rate for the Wald test with a HAC covariance matrix estimate is the same as that obtained in Phillips (1998) for standard spurious regression inference with HAC error variance matrix estimators. Hence, use of fixed K inference with HAC variance estimation is not readily compatible with both singular and nonsingular cases and encounters difficulties similar to those arising in the use of IM-OLS and FM-OLS. In view of these drawbacks, we do not pursue the fixed-K TIV approach further in this context of potential singularity and multicointegration in $I(1)$ systems.

[9]See equations (111) and (112) in the proof of Theorem 5 for the residual inconsistency and (120) for the divergence rate of the Wald test of $O_p\left(\frac{n}{M}\right)$ when K is fixed.

The use of HAR inference leads to very different limit theory that is much more useful in practical work. Importantly, fixed-b settings for the bandwidth parameter as in (49) with $M = bn$ and $b \in (0, 1]$ control divergence, just as in other work on spurious regressions with HAR inference methods (Sun, 2004; Phillips et al., 2019). As usual, the HAR approach introduces nonstandard limit theory. But, as we see below, the limit theory is pivotal even for quite general linear hypothesis tests such as $\mathcal{H}_0 : Ha = h$, so that simulation based techniques and bootstrap methods are available for inference.

Under HAR inference, a substantial degree of robustness in terms of asymptotic size control in testing is achieved. Importantly, this robustness covers both cointegration and multicointegration cases. The following results give the limit theory of the two test statistics Wald_{TIV} and $\text{Wald}_{TIV,b}$ when $(K, n) \to \infty$ when $\Omega_{00.x} > 0$ and $\Omega_{00.x} = 0$.

Theorem 5 (TIV inference with $K \to \infty$) *Under the assumptions of Theorem 3 and under the null hypothesis $\mathcal{H}_0 : Ha = h$, the following hold as $(K, n) \to \infty$ with $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$:*

When $\Omega_{00.x} > 0$:

(vii) $\text{Wald}_{TIV} = O_p\left(\frac{n}{M}\right), \quad \text{Wald}_{TIV,b} \rightsquigarrow \eta'_{\mathcal{E}_W} L \left\{LL'\right\}^{-1} L' \eta_{\mathcal{E}_W},$

where $L = \mathcal{E}_W^{1/2} \Omega_{xx}^{-1/2} H$, *and setting* $k_b(\cdot) = k\left(\frac{\cdot}{b}\right)$,

$$\mathcal{E}_W := \mathcal{A}_{W,X.x}^{-1} \left(\int_0^1 \int_0^1 k_b(r-p) \left(W_{X.x}(r) W_{X.x}(p)'\right) \widetilde{W_{0.x}}(r) \widetilde{W_{0.x}}(p)\, dr dp \right) \mathcal{A}_{W,X.x}^{-1},$$

$$\eta_{\mathcal{E}_W} := \mathcal{E}_W^{-1/2} \left(\int_0^1 W_{X.x} W'_{X.x} \right)^{-1} \int_0^1 \overrightarrow{W_{X.x}} dW_{0.x}.$$

where

$$W_{0.x}(r) = W_0(r) - \Omega_{0x}\Omega_{xx}^{-1} W_x(r),$$

$$\widetilde{W_{0.x}}(r) = W_{0.x}(r) - \int W_{0.x} W'_{X.x} \left(\int W_{X.x} W'_{X.x} \right)^{-1} W_X(r) - \int W_{0.x} W'_{x.X} \left(\int W_{x.X} W'_{x.X} \right)^{-1} W_x(r),$$

$$W_{x.X}(r) = W_x(r) - \int W_x W'_X \left(\int W_X W'_X \right)^{-1} W_X(r),$$

$$W_{X.x}(r) = W_X(r) - \int W_X W'_x \left(\int W_x W'_x \right)^{-1} W_x(r)$$

When $\Omega_{00.x} = 0$:

(viii) $\text{Wald}_{TIV} \rightsquigarrow \chi_q^2, \quad \text{Wald}_{TIV,b} \rightsquigarrow \eta'_{e.x} \mathcal{J}'_q \left\{\mathcal{J}_q \mathcal{F}_W \mathcal{J}'_q\right\}^{-1} \mathcal{J}_q \eta_{e.x},$

where

$$\mathcal{F}_W := \mathcal{A}_{W,X.x}^{-1} \left(\int_0^1 \int_0^1 k_b(r-p) W_{X.x}(p) W_{X.x}(r)' \, d\mathcal{Q}_W(p)\, d\mathcal{Q}_W(r) \right) \mathcal{A}_{W,X.x}^{-1},$$

$$\eta_{e.x} := \mathcal{A}_{W,X.x}^{-1}\left(\int_0^1 W_{X.x} dW_{e.x}\right), \; W_{e.x}(r) = \omega_{ee.x}^{-1/2} B_{e.x} = \omega_{ee.x}^{-1/2}\left(B_e - \omega_{ex}\Omega_{xx}^{-1} B_x\right),$$

$$\mathcal{A}_{W,X.x} = \int_0^1 W_{X.x} W_{X.x}', \text{ and } \mathcal{J}_q = [I_q, 0].$$

$\mathcal{F}_W$ and $\eta_{e.x}$ depend only on the vector standard Brownian motions (W_x, W_X) and the standard Brownian motion $W_{e.x}$ which is independent of (W_x, W_X). The stochastic process $\mathcal{Q}_W(\cdot)$ is also a functional of these standard Brownian motions and is defined in (136).

Remarks

(a) The first result of **(vii)** shows that the HAC-based Wald statistic Wald_{TIV} diverges at rate $O_p\left(\frac{n}{M}\right)$, just as the squared t-statistic in Phillips (1998). So HAC variance matrices in the construction of the Wald statistic fail to resolve the partially spurious nature of the regression (32) and are therefore not recommended in the present context where there is potential multicointegration.

(b) On the other hand, the second result of **(vii)** shows that the fixed-b HAR variance matrix estimator leads to the modified Wald statistic $\text{Wald}_{TIV,b}$ whose limit distribution can be represented by the pivotal quadratic form quantity $\eta_{\mathcal{E}_W}' L \{LL'\}^{-1} L' \eta_{\mathcal{E}_W}$. Importantly, the random projection matrix $P_L = L\{LL'\}^{-1} L'$ has rank $q = \text{rank}(L) = \text{rank}(H)$ *a.s.* and is diagonalizable by an orthogonal matrix. Since the distribution of the random vector $\eta_{\mathcal{E}_W} = \mathcal{E}_W^{1/2} \eta_W$ is invariant to orthogonal transformation in the same way as the vector standard Brownian motions (W_x, W_X), the random quadratic form $\eta_{\mathcal{E}_W}' L\{LL'\}^{-1} L' \eta_{\mathcal{E}_W}$, which is a nonlinear functional of these standard Brownian motions and $W_{0.x}$, depends only on the rank of the matrix L, viz. the number of restrictions q. This pivotal limit theory for the HAR statistic $\text{Wald}_{TIV,b}$ makes valid asymptotic inference possible by direct simulation or by use of the bootstrap. The HAR statistic $\text{Wald}_{TIV,b}$ is constructed in the usual manner for trend IV inference and in the cointegration case with $\Omega_{00.x} > 0$ provides a simple alternative to the procedures suggested in Vogelsang and Wagner (2014).[10]

(c) Analysis under the local alternative hypothesis $\mathcal{H}_A : Ha = h + \frac{d(a)}{n}$ shows that the Wald test based on the $\text{Wald}_{TIV,b}$ statistic has non-trivial asymptotic power under cointegration, with strength that depends on a random noncentrality parameter involving the quadratic form $\theta_d = d(a)' \mathcal{E}_W^{1/2} L\{LL'\}^{-1} L' \mathcal{E}_W^{1/2} d(a)$.

(d) In **(viii)** under multicointegration, the HAC-based Wald statistic $\text{Wald}_{TIV} \rightsquigarrow \chi_q^2$ and the HAR-based statistic

$$\text{Wald}_{TIV,b} \rightsquigarrow \eta_{e.x}' \mathcal{J}_q' \left\{\mathcal{J}_q \mathcal{F}_W \mathcal{J}_q'\right\}^{-1} \mathcal{J}_q \eta_{e.x}, \quad \text{with } \mathcal{J}_q = [I_q, 0].$$

Both test statistics have nontrivial asymptotic power under multicointegration and local alternative hypotheses of the form $\mathcal{H}_A : Ha = h + \frac{d(a)}{n^2}$. The statistic Wald_{TIV} has a

[10] The procedures suggested in Vogelsang and Wagner (2014) are designed only for the cointegration case with $\Omega_{00.x} > 0$ and do not apply under multicointegration.

noncentral χ_q^2 limit distribution with noncentrality parameter $d(a)'d(a)$; and the $\text{Wald}_{TIV,b}$ statistic has a noncentral limit distribution involving the random noncentrality parameter $\vartheta_d = d(a)'\mathcal{J}_q'\left\{\mathcal{J}_q\mathcal{F}_W\mathcal{J}_q'\right\}^{-1}\mathcal{J}_q d(a)$.

(e) Theorem 5 **(vii)** and **(viii)** show that the same HAR Wald statistic $\text{Wald}_{TIV,b}$ is asymptotically valid and pivotal for both cointegrated and multicointegrated systems, therefore providing a robust approach to inference concerning the cointegrating coefficients even under singularity.

(f) These findings for the Wald test $\text{Wald}_{TIV,b}$ extend in a straightforward way to HAR-based t ratio statistics which produce asymptotically pivotal tests for both $\Omega_{00.x} = 0$ and $\Omega_{00.x} > 0$ cases.

In nonsingular systems with $\Omega_{00.x} > 0$ we can expect some loss of cointegration estimation efficiency and test power when using TIV estimation on the extended system (26) and the associated robust $\text{Wald}_{TIV,b}$ test rather than TIV estimation of (8) and associated Wald tests that rely on correct prior knowledge that the conditional error variance $\Omega_{00.x} > 0$. But when $\Omega_{00.x} = 0$, the faster $O(n^2)$ convergence rate of the estimator sharpens estimation efficiency and improves the discriminatory power of both the Wald_{TIV} test and the $\text{Wald}_{TIV,b}$ test.

We close this section by mentioning that the inferential apparatus above that leads to the high-dimensional TIV Wald statistic $\text{Wald}_{TIV,b}$ for inference about the cointegration vector a may be applied to construct similar high-dimensional TIV Wald statistics for testing hypotheses about the multicointegration vector f. The associated limit theory is chi-squared for a HAC based Wald statistic which uses a consistent estimator of $\omega_{ee.x}$ and nonstandard but still pivotal when a fixed-b estimator of $\omega_{ee.x}$ is used. These results align with those given in Theorem 5 **(viii)** for the two Wald statistics Wald_{TIV} and $\text{Wald}_{TIV,b}$ for testing hypotheses about a. Both results rely on the mixed normal limit theory given in Theorem 4 for the estimator $\hat{f}_{TIV}$. Details of these results will be reported in later work.

6 Simulations

This Section reports the finite sample performance of TIV estimation of cointegrating relationships and compares TIV performance with IM-OLS estimation for various model specifications that include time series with and without multicointegration. Finite sample properties of the TIV Wald statistics are also studied in cases of cointegration and multicointegration. As a baseline for cointegrated series without multicointegration, simulations in past work (Phillips, 2014) showed good performance characteristics for TIV estimation in relation to other standard procedures such as FM-OLS and Dynamic Least Squares in triangular systems as well as reduced rank regression (RRR) in VAR system formulations with cointegration but not multicointegration. Those findings are now extended to include comparisons with IM-OLS in the present case.

Several experimental designs were employed based on the data generating process

$$y_t = ax_t + u_{0t}$$
$$x_t = x_{t-1} + u_{xt}, \quad t = 1, \ldots, n,$$

where $u_t = \eta_t + D_1\eta_{t-1}$, $\eta_t \sim iidN(0, \Sigma)$, $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, the cointegrating coefficient $a = 2$, and the initialization of x_t is $x_0 = 0$. Both cointegrated and multicointegrated systems are considered and these are determined by the parameter settings of the (endogeneity) correlation coefficient ρ and the moving average coefficient matrix D_1. Various sample sizes are used and the number of replications in each experiment is $10,000$. The following models were used.

Cointegrated models

Model 10: $D_1 = 0_{2\times 2}, \ \rho = 0$

Model 11: $D_1 = 0_{2\times 2}, \ \rho = 0.5$

Model 12: $D_1 = \begin{bmatrix} 0.3 & 0.4 \\ 0.8 & 0.6 \end{bmatrix}, \ \rho = 0.5$

Multicointegrated models

Model 20: $D_1 = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \ \rho = 0$

Model 21: $D_1 = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \ \rho = 0.5$

Model 22: $D_1 = \begin{bmatrix} 0.3 & 0.4 \\ 5.2 & 0.6 \end{bmatrix}, \ \rho = 0.5$

Model 23: $D_1 = \begin{bmatrix} -0.3 & 0.4 \\ 0.7 & -0.6 \end{bmatrix}, \ \rho = 0.5$

The models with $\rho = 0$ and zero diagonal elements in D_1 do not generate endogeneity or serial cross-correlation. So those models are pure cointegrated systems with exogenous regressors and *iid* innovations. Model 12 has been used in the cointegration literature in earlier work (Phillips and Loretan, 1991), and Model 22 modifies model 12 by introducing multicointegration into the system. Model 23 also generates a multicointegrated system, but with less variability in u_x compared to Model 22.

For TIV estimation the orthonormal trigonometric polynomials $\varphi_k(r) = \sqrt{2}\sin\{(k - 1/2)\pi r\}$ were used as instrumental variabes and the number of instruments was based on the setting $K = n^{3.8/5}$ in accord with the requirement in Theorems 3 and 4 that $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$. Following the recommendation in the paper the model was estimated by TIV with a fitted intercept. The asymptotic distributions in Theorems 3 and 4 were obtained by numerical computation from simulations with time series of length $n = 1,000$ using $1,000$ replications.

6.1 Finite sample distributions of the estimators

This subsection compares finite sample performance and convergence rates of TIV, RRR and OLS estimators of the cointegrating parameter a. Empirical densities of the centred and scaled TIV and IM-OLS estimators are calibrated against the asymptotic distributions given in Theorem 3.

The centred densities of the TIV, RRR and OLS estimators are shown in Figure 1 for $n = 50$ for three models. In the pure cointegration model 10, the three estimators show similar behavior although TIV, which is not needed in this pure cointegration case, shows somewhat greater dispersion than OLS and RRR. In models 22 and 23 under multicointegration the TIV estimator shows much greater concentration and little bias compared with OLS and RRR which are biased and skewed with greater dispersion. These results corroborate the limit theory in which TIV has an $O(n^2)$ convergence rate in multicointegrated models instead of the usual $O(n)$ rate for cointegrated systems.

We now compare the performance characteristics of TIV and IM-OLS in finite samples. Figure 2 plots the densities of the centred TIV and IM-OLS estimators scaled by the appropriate convergence rate for each model against the mixed-normal asymptotic distribution. For the cointegration models 10-12, Figure 2 plots the densities of the standardized TIV estimator $n(\hat{a}_{TIV} - a)$ based on the sample sizes $n = 50$ and $n = 100$ together with the asymptotic mixed normal density given in Theorem 3(v). For the three models, the mixed-normal approximation works well as an approximation to the finite sample distributions of TIV, even for $n = 50$. The same is true for the densities of the standardized IM-OLS estimators, confirming the result in Theorem 1(i) and earlier results in VW(2014).

For the multicointegrated models 20-23, the densities of the standardized TIV estimator $n^2(\hat{a}_{TIV} - a)$, based on sample sizes $n = 50$ and $n = 100$ and the simulated asymptotic mixed normal density, based on Theorem 3(vi), are plotted in Figure 3. For all these models and cases the mixed-normal approximation to the distribution of the TIV estimator works well, again even for $n = 50$, whereas the IM-OLS estimator shows clear evidence of bias, skewness and greater dispersion for models 21-23. For model 20, where no endogeneity or serial correlation is present, which is the perfect set of conditions for the IM-OLS estimator, the densities of both estimators are approximated well by the mixed normal density, as predicted by Theorem 1(ii) and Theorem 3(vi) with some finite sample advantage in terms of reduced dispersion to the IM-OLS estimator in this case.

6.2 Size and power properties of the Wald test

Finite sample performance of Wald test statistics for testing the null hypothesis $\mathcal{H}_0 : a = 2$ were explored next. The empirical rejection rates under the null for the Wald statistics using the HAR variance estimate and the fixed-b asymptotic distribution given in Theorem 5 were calculated with the setting $b = 1$ and are reported in Table 1 for levels 10%, 5% and 1%. The results show excellent size control in all cases even for $n = 50$ in both the cointegrated and multicointegrated models.

For the Wald statistics using the HAC variance estimate calculated with the setting $M = 3n^{1/5}$ and using a χ^2 asymptotic distribution are presented in Table 2. For the cointegration models size is not controlled and the statistics diverge with the sample size. For the multicointegration models the rejection rates are 2–3 times larger than the nominal ones. Both cases show the importance of the HAR specification and appropriate limit theory for controlling size in Wald statistic testing.

Two control parameters – the number of instruments K and the bandwidth M (or b, the sample fraction) – are used in variance estimation. These parameters need to be set by the user. We analyzed the sensitivity of the Wald test to these parameter settings for models 12 and 22. Empirical rejection rates of the Wald test at the 5% nominal level were studied, varying K as fractions $\{0.2, 0.4, 0.6, 0.8\}$ of the sample size n and M as fractions $\{0.2, 0.4, 0.6, 0.8, 1\}$ of the sample size n. The rates under the null in Table 3 show: (i) that size is stable across a wide range of values of K and b in the cointegrated case; and (ii) that the size is stable across a wide range of values of K, when $b > 0.5$ in the multicointegrated case.

Size-adjusted power calculations under the alternative $H_1 : a = 2.1$ are reported in Table 4. The results show that power is stable across all K values with a minor drop for larger bandwidths in the cointegration case. The size-adjusted power results in the multicointegration case under the alternative $H_1 : a = 2.001$ in Table 4 show that the power is high and increases with the sample size but with a minor drop for larger K and bandwidth size. In view of the faster convergence rate in the multicointegration case, local power in this case is evident for the much smaller departure $H_1 : a = 2.001$ from the null compared with the cointegration case where results for $H_1 : a = 2.1$ are reported.

Finally, in Table 5 we calculate the empirical rejection rates of Wald test statistics at the 5% level varying K as fractions $\{0.2, 0.4, 0.6, 0.8\}$ of the sample size n and (small) bandwidth as fractions $\{0.02, 0.04, 0.06, 0.08, 0.1\}$ of the sample size n using a χ^2 approximation instead of the correct limit theory. The test statistic diverges for all values of K and bandwidths in the cointegration case, while the size in the multicointergation case is sensitive to both number of instruments and bandwidth size.

7 Empirical Illustration

Lee (1996) considered a model of the housing market that implies a long run equilibrium relationship between time series of housing starts and housing completions. If these series are multicointegrated then a parametric VAR $I(1)$ model will be misspecified. Engsted and Haldrup (1999) therefore analyzed the time series within an $I(2)$ framework allowing for multicointegration. In this section, we analyze the long run relationship between housing starts and completions over the five decade period $1970 - 2020$ in an $I(1)$ semiparametric triangular model using the new TIV estimator and associated Wald tests.

The data are provided by the U.S. Census Bureau and the U.S. Department of Housing and Urban Development. They were obtained from FRED, the Federal Reserve Bank of St. Louis on March 16, 2021. We consider two series: *starts* = Housing Starts, which comprise Total New Privately-Owned Housing Units Started [HOUST]; and *completions* = Total New Privately-Owned Housing Units Completed [COMPUTSA]. Both series are reported in thousands of units and are seasonally adjusted. Our empirical analysis considers the following five decadal periods: (1) 1970-01-01 — 1979-12-31, (2) 1980-01-01 — 1989-12-31, (3) 1990-01-01 — 1999-12-31, (4) 2000-01-01 — 2009-12-31, (5) 2010-01-01 — 2019-12-31.

The cointegration relationship between *completions* and *starts* is estimated in each of these decades. In estimation no *a priori* assumption is made about the existence or non-existence of multicointegration. The results are given in Table 6. Over decades (1) and (2) to 1990, the estimate 0.98 is basically the same as that found in Lee (1996). The estimate then declines to 0.96 in 1990-2000 and to 0.95 in recent years. A possible interpretation is that 5% of houses under construction were never completed in those decades. A practical question is whether this fraction of uncompleted houses is significant, which can be formalized as a test of the null hypothesis $\mathcal{H}_0 : a = 1$ against the alternative $\mathcal{H}_1 : a < 1$.

The equilibrium errors from the cointegrated relationship between *completions* and *starts* accumulate into a stock variable of incomplete constructions. In each period, the inventory stock variable is measured as

$$Stock_t = \sum_{j=1}^{t} \left(\hat{a}_{TIV} * start_j - completed_j \right), \tag{52}$$

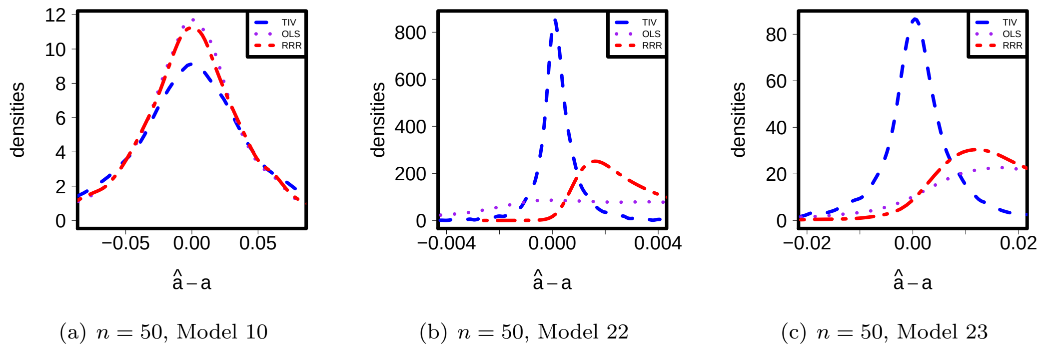
and is plotted together with the flow variables *starts* and *completions* in Figure 4. The figure reveals that these variables are again cointegrated, revealing a multicointegrated relationship between *completions* and *starts*. To conduct a test of the null $\mathcal{H}_0$, the asymptotic distributions of the Wald test statistic given in Theorem 4 are approximated by Monte Carlo simulations with 1000 replications for a sample size of 1000, and p-values for the two distributions (under cointegration and multicointegration) are calculated for each period.

The empirical findings for these tests are shown in Table 6. Allowing for multicointegration in the relationship we conclude that the null hypothesis $\mathcal{H}_0 : a = 1$ is rejected for periods (2), (3), and (4) (and nearly rejected for period (5)) at the 5% level as indicated by the p-values shown in the column 'pvalue-M'. If the multicointegrated relationship is ignored, the null hypothesis would not be rejected for any period, except for period (4), as indicated by the p-values given in the column 'pvalue-C'. Allowing for the presence of a multicointegrated relationship among starts, completions, and the housing stock therefore has a material impact on the empirical (cointegrating) relationship between starts and completions that suggests a significant shift in the relationship that raises the fraction of uncompleted houses.

8 Conclusion

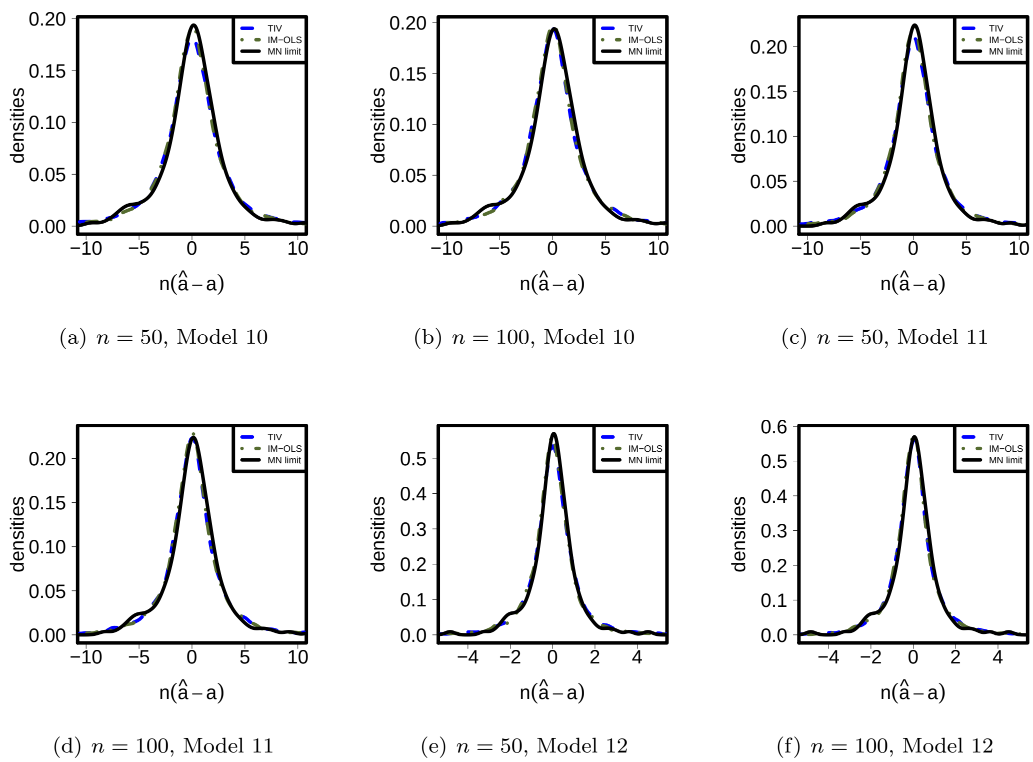
This paper has studied the effects of singularities in long run conditional covariance matrices on estimation and inference in cointegrating regression models. Such singularities are shown to be present whenever a cointegrated $I(1)$ system happens to involve multicointegrated time series. Singularities complicate estimation and inference by leading to non-pivotal, nuisance parameter dependencies in all existing methods of estimating nonstationary time series regressions. But in view of their natural focus on the analysis of long run properties, instrumental variable regression with deterministic trend regressors or similar trend transforms have appealing properties even under singularities. The results of the present analysis show that, in spite of the complications introduced by long run variance matrix singularities, certain key advantages of the trend IV regression approach continue to apply. Notably, the limit theory of trend IV regression is mixed normal and Wald tests based on traditional sandwich formulae may be conducted under pivotal asymptotics without knowledge of potential singularities or the presence of multicointegration in the time series. Use of fixed-b methods in conjunction with trend IV regression are particularly helpful in achieving pivotal limit theory when the regression equation is partially spurious with nonstationary errors and usual HAC-based test statistics are divergent under the null.

The analysis in this paper deals with estimation and inference in a scalar cointegrating relationship. The main ideas and methods of estimation and inference extend to systems estimation. In such cases, the higher convergence rate $O\left(n^2\right)$ applies in the (possibly matrix) direction L_1 of singularity of $\Omega_{00.x}$ for which $L_1'\Omega_{00.x}L_1 = 0$ and the slower $O\left(n\right)$ rate applies in the orthogonal direction L_2. The full matrix of cointegrating coefficients then converges to a mixed normal limit distribution which is a matrix transform of the slower rate limit distribution, just as in usual cointegration limit theory (Park and Phillips, 1988, 1989; Phillips, 1988, 1989). The analysis and algebra in this general case follows the same approach as that in cointegrated regression systems with cointegrated regressors and unrestricted VAR estimation with cointegrated variates, as detailed in Phillips (1995). But inferential limit theory is more subtle in this case of singularity in the matrix $\Omega_{00.x}$ because of interaction between the restriction matrix H, the rotation matrix $L = [L_1, L_2]$ isolating the two directions of convergence, and the matrix normalization involved in standardizing the TIV estimation errors. A full analysis of this case requires the use of methods and limit theory for Wald tests under general conditions of matrix normalization as recently developed in Magdalinos and Phillips (2019). The application of those methods in the present context is left for subsequent work.



(a) $n = 50$, Model 10 (b) $n = 50$, Model 22 (c) $n = 50$, Model 23

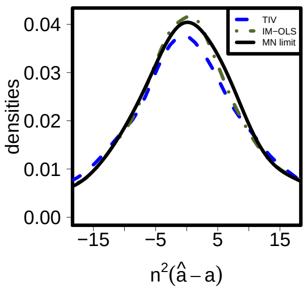
Figure 1: Kernel estimates of the density functions of the estimation errors $\widehat{a} - a$ for the TIV, RRR and LS estimators for sample size $n = 50$ in the pure cointegration model 10 and the multicointegration models 22 and 23.



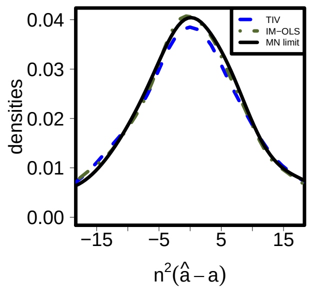
(a) $n = 50$, Model 10 (b) $n = 100$, Model 10 (c) $n = 50$, Model 11

(d) $n = 100$, Model 11 (e) $n = 50$, Model 12 (f) $n = 100$, Model 12

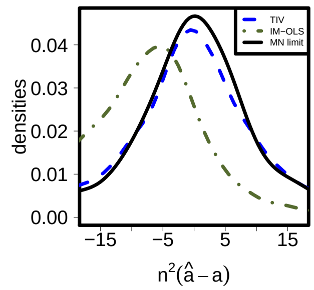
Figure 2: Kernel density estimates of the density functions of the estimation error $n(\widehat{a} - a)$ for the TIV and the IM-OLS estimators and the density of the mixed-normal limit of the TIV estimator for sample sizes $n = 50$ and $n = 100$ and cointegration models 10, 11, and 12.



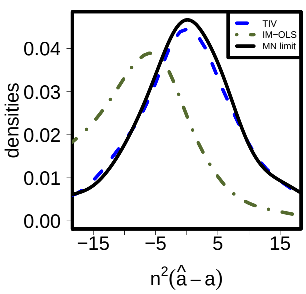
(a) $n = 50$, Model 20



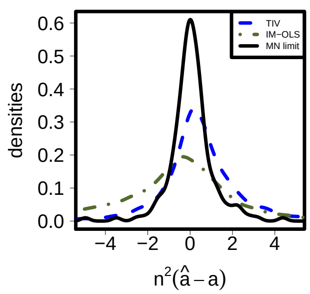
(b) $n = 100$, Model 20



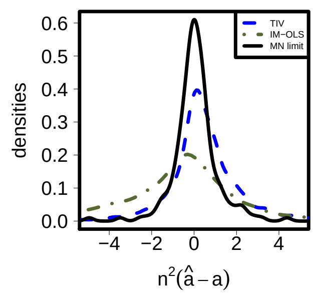
(c) $n = 50$, Model 21



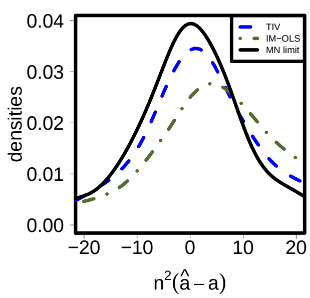
(d) $n = 100$, Model 21



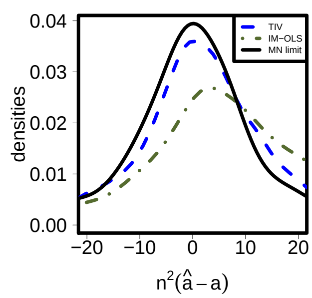
(e) $n = 50$, Model 22



(f) $n = 100$, Model 22



(g) $n = 50$, Model 23



(h) $n = 100$, Model 23

Figure 3: Kernel density estimates of the density functions of the estimation error $n^2(\widehat{a} - a)$ for the TIV and the IM-OLS estimators and the density of the mixed-normal limit of the TIV estimator for sample sizes $n = 50$ and $n = 100$ and multicointegration models 20, 21, 22 and 23.

Table 1: Test size using HAR variance estimates. Empirical rejection rates are shown at nominal $10\%, 5\%$ and 1% levels Wald test using the fixed-b asymptotic approximation, calculated for different models and sample sizes.

Model	n	10%	5%	1%
10	50	0.1178	0.0612	0.0163
10	100	0.1100	0.0591	0.0139
11	50	0.1178	0.0612	0.0163
11	100	0.1100	0.0591	0.0139
12	50	0.1139	0.0585	0.0154
12	100	0.1141	0.0581	0.0163
20	50	0.1070	0.0552	0.0130
20	100	0.0958	0.0479	0.0095
21	50	0.1070	0.0552	0.0130
21	100	0.0958	0.0479	0.0095
22	50	0.1242	0.0653	0.0139
22	100	0.1161	0.0623	0.0138
23	50	0.1201	0.0613	0.0135
23	100	0.0950	0.0513	0.0121

Table 2: Test size using HAC variance estimates. Empirical rejection rates are shown at nominal $10\%, 5\%$ and 1% levels for the Wald test using χ^2 critical values as approximations, calculated for different models and sample sizes.

Model	n	10%	5%	1%
10	50	0.6981	0.6442	0.5494
10	100	0.7302	0.6852	0.5949
11	50	0.6981	0.6442	0.5494
11	100	0.7302	0.6852	0.5949
12	50	0.6893	0.6348	0.5391
12	100	0.7359	0.6833	0.5907
20	50	0.2212	0.1488	0.0661
20	100	0.1722	0.1056	0.0396
21	50	0.2212	0.1488	0.0661
21	100	0.1722	0.1056	0.0396
22	50	0.2492	0.1730	0.0788
22	100	0.2318	0.1560	0.0638
23	50	0.2601	0.1871	0.0932
23	100	0.1983	0.1312	0.0536

Table 3: Test size across K and b. Empirical rejection rates at nominal 5% level Wald test using the fixed-b asymptotic approximation, calculated for different models and sample sizes, for a range of instrument numbers K (in rows) and a range of bandwidths used in the kernel estimation of the variance determined by b (in columns).

Model	n	K\b	0.2	0.4	0.6	0.8	1
12	50	10	0.0533	0.0550	0.0576	0.0622	0.0633
12	50	20	0.0582	0.0541	0.0553	0.0593	0.0577
12	50	30	0.0580	0.0549	0.0540	0.0582	0.0581
12	50	40	0.0577	0.0543	0.0543	0.0587	0.0585
12	100	20	0.0612	0.0594	0.0606	0.0646	0.0625
12	100	40	0.0607	0.0557	0.0555	0.0588	0.0583
12	100	60	0.0599	0.0546	0.0544	0.0579	0.0567
12	100	80	0.0594	0.0548	0.0542	0.0576	0.0554
22	50	10	0.0345	0.0604	0.0660	0.0682	0.0662
22	50	20	0.0744	0.0766	0.0753	0.0716	0.0670
22	50	30	0.0982	0.0856	0.0803	0.0781	0.0713
22	50	40	0.1070	0.0935	0.0889	0.0839	0.0766
22	100	20	0.0622	0.0634	0.0627	0.0631	0.0599
22	100	40	0.0886	0.0723	0.0715	0.0672	0.0612
22	100	60	0.1020	0.0798	0.0716	0.0710	0.0629
22	100	80	0.1011	0.0855	0.0823	0.0754	0.0672

Table 4: Size-adjusted power across K and b. Empirical rejection rates at nominal 5% level Wald test using fixed-b approximation, calculated for different models and sample sizes, for a range of number of instruments, K (shown in rows), and a range of bandwidths used in the kernel estimation of the variance determined by b (shown in columns).

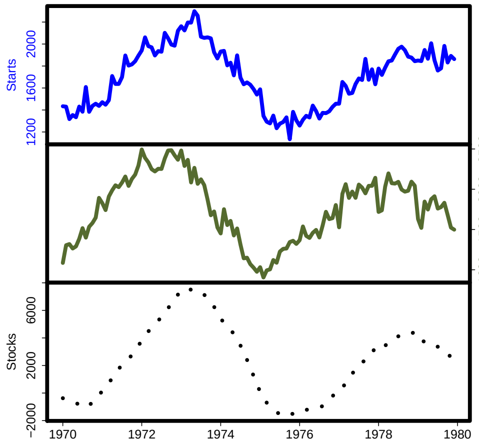
Model	n	K\b	0.2	0.4	0.6	0.8	1
12	50	10	0.8011	0.7657	0.7312	0.7036	0.6773
12	50	20	0.8143	0.7883	0.7593	0.7278	0.7057
12	50	30	0.8177	0.7883	0.7644	0.7320	0.7058
12	50	40	0.8208	0.7884	0.7624	0.7311	0.7019
12	100	20	0.9440	0.9233	0.9044	0.8819	0.8650
12	100	40	0.9484	0.9276	0.9095	0.8905	0.8747
12	100	60	0.9493	0.9289	0.9116	0.8921	0.8778
12	100	80	0.9494	0.9281	0.9126	0.8926	0.8786
22	50	10	0.5742	0.5313	0.4955	0.4627	0.4291
22	50	20	0.5222	0.4793	0.4369	0.3966	0.3621
22	50	30	0.4229	0.3749	0.3280	0.2976	0.2782
22	50	40	0.3129	0.2693	0.2423	0.2148	0.2049
22	100	20	0.9671	0.9577	0.9453	0.9282	0.9108
22	100	40	0.9396	0.9246	0.9084	0.8891	0.8652
22	100	60	0.8941	0.8722	0.8533	0.8258	0.8006
22	100	80	0.8274	0.8013	0.7787	0.7488	0.7238

Table 5: Test size across K and small b. Empirical rejection rates at nominal 5% level for the Wald test using χ^2 critical values, calculated for different models and sample sizes, for a range instrument numbers K (shown in rows), and a range of bandwidths used in the kernel estimation of the variance determined by b (shown in columns).

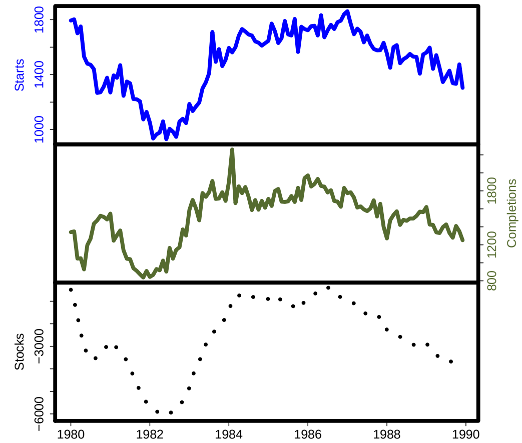
Model	n	K\b	0.02	0.04	0.06	0.08	0.1
12	50	10	0.7086	0.7086	0.6790	0.6450	0.6281
12	50	20	0.7460	0.7460	0.7174	0.6838	0.6614
12	50	30	0.7497	0.7497	0.7222	0.6891	0.6660
12	50	40	0.7518	0.7518	0.7222	0.6905	0.6658
12	100	20	0.8140	0.7525	0.7081	0.6773	0.6565
12	100	40	0.8238	0.7617	0.7182	0.6853	0.6615
12	100	60	0.8250	0.7635	0.7189	0.6865	0.6612
12	100	80	0.8256	0.7631	0.7183	0.6863	0.6607
22	50	10	0.0009	0.0009	0.0025	0.0129	0.0340
22	50	20	0.0172	0.0172	0.0301	0.0694	0.1172
22	50	30	0.0686	0.0686	0.0953	0.1524	0.2047
22	50	40	0.1481	0.1481	0.1757	0.2214	0.2549
22	100	20	0.0008	0.0062	0.0401	0.0863	0.1238
22	100	40	0.0104	0.0544	0.1367	0.1946	0.2285
22	100	60	0.0483	0.1284	0.2100	0.2511	0.2702
22	100	80	0.1172	0.1890	0.2378	0.2617	0.2761

Table 6: US housing construction data. Wald test statistics and p-values for the null hypothesis $\mathcal{H}_0 : a = 1$ under cointegration and multicointegration for successive decades over 1970-2020.

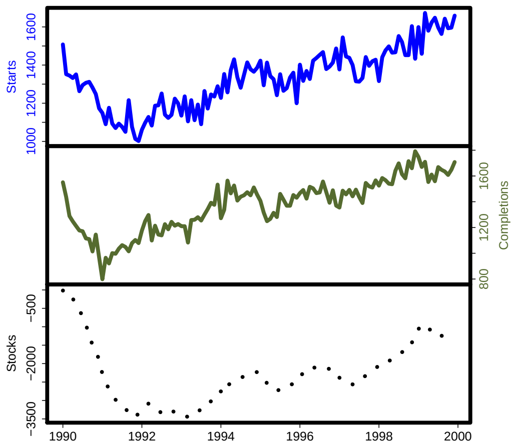
Period begins	Period ends	TIV	pvalue-M	pvalue-C
1970-01-01	1979-12-31	0.9784476	0.06218905	0.18656716
1980-01-01	1989-12-31	0.9735254	0.01492537	0.07462687
1990-01-01	1999-12-31	0.9606591	0.02736318	0.11691542
2000-01-01	2009-12-31	0.9709445	0.00497512	0.04477612
2010-01-01	2019-12-31	0.9454967	0.05223881	0.16666667



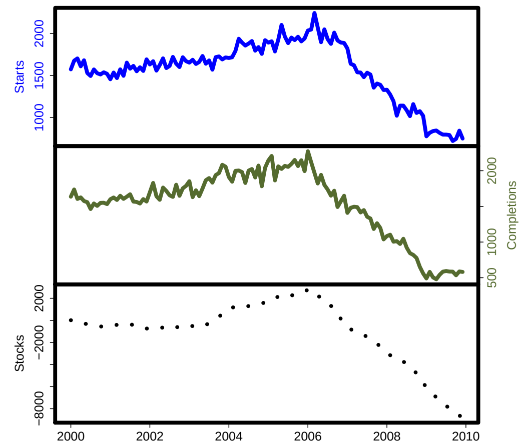
(a) 1970s



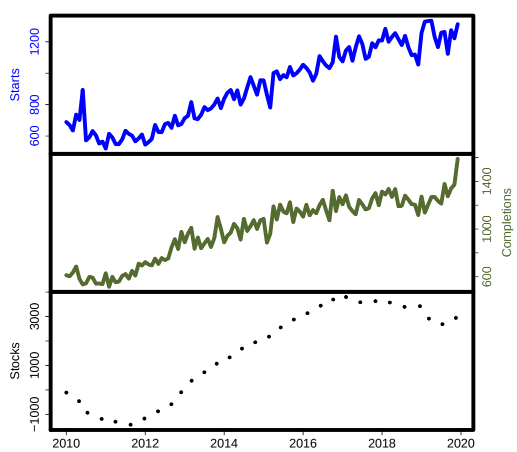
(b) 1980s



(c) 1990s



(d) 2000s



(e) 2010s

Figure 4: Housing starts (Starts), completions (Completions) and accumulated difference (Stock) data for successive decades (a) 1970s, (b) 1980s, (c) 1990s (d) 2000s and (e) 2010s.

9 Appendix

This Appendix provides proofs of subsidiary results and all the main theorems in the paper. The following glossary of notation that is used in the paper is provided for convenient reference.

9.1 Subsidiary Results

Lemma A (Reverse partial summation) : $\sum_{s=1}^{n} a_s b_s = \sum_{t=1}^{n} \left(\sum_{s=t}^{n} a_s\right) \Delta b_t$ *if* $b_0 = 0$.

Proof of Lemma A

By partial summation $f_n g_n - f_0 g_0 = \sum_{t=1}^{n} (\Delta f_t) g_t + \sum_{t=1}^{n} f_{t-1} \Delta g_t$. Setting $f_t = \sum_{s=1}^{t} a_s$ so that $\Delta f_t = a_t$ and $b_t = g_t$ we have, with $b_0 = 0$,

$$\sum_{s=1}^{n} a_s b_s = \left(\sum_{s=1}^{n} a_s\right) b_n - \sum_{t=1}^{n} \left(\sum_{s=1}^{t-1} a_s\right) \Delta b_t = \left(\sum_{s=1}^{n} a_s\right) b_n - \sum_{t=1}^{n} \left(\sum_{s=1}^{n} a_s - \sum_{s=t}^{n} a_s\right) \Delta b_t$$
$$= \left(\sum_{s=1}^{n} a_s\right) b_n - \left(\sum_{s=1}^{n} a_s\right) (b_n - b_0) + \sum_{t=1}^{n} \left(\sum_{s=t}^{n} a_s\right) \Delta b_t = \sum_{t=1}^{n} \left(\sum_{s=t}^{n} a_s\right) \Delta b_t,$$

giving a reverse form of the usual partial summation formula which involves only a single term and is useful in simplifying finite sample expressions and limit formulae.

Lemma B:

(i) $a_n V_{e^+} := \begin{cases} n^{1/2} \sum_{t=1}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right) e_t & \Omega_{00.x} = 0 \\ n^{-3/2} \sum_{t=1}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right) U_{0.xt} & \Omega_{00.x} > 0 \end{cases}$

$$\rightsquigarrow \begin{cases} \int_0^1 \tilde{\varphi}_K dB_e & \Omega_{00.x} = 0 \\ \int_0^1 \tilde{\varphi}_K B_{0.x} & \Omega_{00.x} > 0 \end{cases} =: \begin{cases} \psi_{eK} & \Omega_{00.x} = 0 \\ \Psi_{0.xK} & \Omega_{00.x} > 0 \end{cases};$$

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

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

(iv) $\frac{1}{n^{5/2}} V_X := \frac{1}{n} \sum_{t=1}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right) \frac{X_t'}{n^{3/2}} \rightsquigarrow \int_0^1 \tilde{\varphi}_K B_X' = \mu_K$;

(v) $\frac{1}{n} \sum_{t=1}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right) \frac{x_t'}{n^{3/2}} \rightsquigarrow \int_0^1 \tilde{\varphi}_K B_x' = \int_0^1 \left(\int_r^1 \tilde{\varphi}_K\right) dB_x'$

(vi) $\mathbb{E}\left(\eta_K \xi_K'\right) = \int_{r=0}^{1} \tilde{\varphi}_K(r) \int_0^r \tilde{\varphi}_K(p)' \, dp dr \times \operatorname{trace}\left[\Omega_{xx}\right]$.

where $\tilde{\varphi}_K(r) = (\varphi_1(r), ..., \varphi_K(r))'$, *and* $B_X(r) = \int_0^r B_x$.

Proof of Lemma B

Parts (i)-(iv). These results follow by standard weak convergence methods for these orthonormal linear transform functionals (Phillips (2005a, 2014).

Part (v) Convergence to $\int_0^1 \tilde{\varphi}_K B_x'$ is immediate. for the second representation, we use a version of partial integration, analogous to the version of partial summation given in Lemma A, viz.,

$$\int_0^1 \tilde{\varphi}_K B_x' \quad = \quad \left[-\left(\int_r^1 \tilde{\varphi}_K(s) ds\right) B_x'(r)\right]_0^1 + \int_0^1 \left(\int_r^1 \tilde{\varphi}_K(s) ds\right) dB_x'(r)$$

$$= \int_0^1 \left(\int_r^1 \tilde{\varphi}_K(s) ds \right) dB_x'(r),$$

using the fact that $B_x(0) = 0$.

Part (vi) Direct calculation gives

$$\mathbb{E}\left(\eta_K \xi_K'\right) = \mathbb{E}\left\{ \int_{r=0}^1 \tilde{\varphi}_K(r) B_x(r)' dr \int_{s=0}^1 dB_x(s) \tilde{\varphi}_K(s)' \right\}$$

$$= \int_{r=0}^1 \int_{s=0}^1 \tilde{\varphi}_K(r) \int_0^r \mathbb{E}\left(dB_x(p)' dB_x(s)\right) \tilde{\varphi}_K(s)' dr$$

$$= \int_{r=0}^1 \tilde{\varphi}_K(r) \int_0^r \tilde{\varphi}_K(p)' dp dr \times \text{trace}\left[\Omega_{xx}\right].$$

The following results provide limit theory for certain quadratic forms of $I(2)$, $I(1)$, and $I(0)$ time series where the quadratic forms involve projection matrices onto the space of orthonormal polynomials in which the dimension of the space $K \to \infty$ as $n \to \infty$. The resutls are stated here for convenient reference and proofs are available elsewhere. In particular, results (54) and (55) are proved in the proof of Theorem 3, (58) is proved in Phillips (2005a), and (53) and (57) are proved in Phillips (2014).

Lemma C: As $(K, n) \to \infty$ with $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$,

$$\frac{1}{n^2} x' P_{\Phi_K} x = \left(\frac{1}{n} \frac{x'\Phi_K}{\sqrt{n}}\right)\left(I_K + O\left(\frac{1}{n}\right)\right)\left(\frac{1}{n}\frac{\Phi_K' x}{\sqrt{n}}\right) \rightsquigarrow \int_0^1 B_x B_x' \tag{53}$$

$$\frac{1}{n^4} X' P_{\Phi_K} X = \left(\frac{1}{n} \frac{X'\Phi_K}{n^{3/2}}\right)\left(I_K + O\left(\frac{1}{n}\right)\right)\left(\frac{1}{n}\frac{\Phi_K' X}{n^{3/2}}\right) \rightsquigarrow \int_0^1 B_X B_X' \tag{54}$$

$$\frac{1}{n^3} X' P_{\Phi_K} x = \left(\frac{1}{n} \frac{X'\Phi_K}{n^{3/2}}\right)\left(I_K + O\left(\frac{1}{n}\right)\right)\left(\frac{1}{n}\frac{\Phi_K' x}{n^{1/2}}\right) \rightsquigarrow \int_0^1 B_X B_x' \tag{55}$$

$$\frac{1}{n^3} X' P_{\Phi_K} U_{0.x} = \left(\frac{1}{n} \frac{X'\Phi_K}{n^{3/2}}\right)\left\{I_K + O\left(\frac{1}{n}\right)\right\}\left(\frac{1}{n}\Phi_K' \frac{U_{0.x}}{n^{1/2}}\right) \rightsquigarrow \int_0^1 B_X B_{0.x} \tag{56}$$

$$\frac{1}{n} x' P_{\Phi_K} u_x \frac{1}{K^{1/2}} = O_p\left(\frac{1}{K^{1/2}}\right) = o_p(1), \tag{57}$$

$$\frac{1}{K} u_x' P_{\Phi_K} u_x = \frac{1}{K}\frac{u_x'\Phi_K}{\sqrt{n}}\left(I_K + O\left(\frac{1}{n}\right)\right)\frac{\Phi_K' u_x}{\sqrt{n}} \to_p \Omega_{xx} \tag{58}$$

where $P_{\Phi_K} = \Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'$ and $\Phi_K' = \left[\tilde{\varphi}_{K1}, ..., \tilde{\varphi}_{Kn}\right]$, where $\tilde{\varphi}_{Kt} = \tilde{\varphi}_K\left(\frac{t}{n}\right) = \left[\varphi_1\left(\frac{t}{n}\right), ..., \varphi_K\left(\frac{t}{n}\right)\right]'$.

9.2 Proofs of the Main Theorems

Proof of Theorem 1

Part (i) When $\Omega_{00.x} > 0$, this result follows as in Vogelsang and Wagner (2014) with only minor modification. The system (26) is $Y_t = a'X_t + f'x_t + U_{0.xt}$ and then

$$\hat{a} - a = \left(X'Q_x X\right)^{-1} X'Q_x U_{0.x}, \ \hat{f} - f = \left(x'Q_X x\right)^{-1} x'Q_X U_{0.x}$$

Standard weak convergence methods (Phillips, 1986, 1988) give the following component limits:

(a-1) $n^{-2}\sum_{t=1}^n x_t U_{0.xt} \rightsquigarrow \int_0^1 B_x B_{0.x}$, $n^{-3}\sum_{t=1}^n X_t U_{0.xt} \rightsquigarrow \int_0^1 B_X B_{0.x}$, where $n^{-3/2}X_{\lfloor nr\rfloor} = n^{-3/2}\sum_{t=1}^{\lfloor n\cdot\rfloor} x_t \rightsquigarrow B_X(r) = \int_0^r B_x,\ B_x(r) = BM(\Omega_{xx}),$ and $B_{0.x}(r) = BM(\Omega_{00.x})$, where $\Omega_{00.x} = \Omega_{00} - \Omega_{0x}\Omega_{xx}^{-1}\Omega_{x0}$.

(a-2) $n^{-4}X'Q_xX = n^{-4}\sum_{t=1}^{n} X_tX_t' - \left(n^{-3}\sum_{t=1}^{n} X_tx_t'\right)\left(n^{-2}\sum_{t=1}^{n} x_tx_t'\right)^{-1}\left(n^{-3}\sum_{t=1}^{n} x_tX_t'\right)$

$$\rightsquigarrow \int_0^1 B_XB_X' - \left(\int_0^1 B_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\left(\int_0^1 B_xB_X'\right) = \int_0^1 B_{X.x}B_{X.x}',$$

(a-3) $n^{-3}X'Q_xU_{0.x} = n^{-3}\sum_{t=1}^{n} X_tU_{0.xt} - \left(n^{-3}\sum_{t=1}^{n} X_tx_t'\right)\left(n^{-2}\sum_{t=1}^{n} x_tx_t'\right)^{-1}\left(n^{-2}\sum_{t=1}^{n} x_tU_{0.xt}\right)$

$$\rightsquigarrow \int_0^1 B_XB_{0.x} - \left(\int_0^1 B_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\left(\int_0^1 B_xB_{0.x}\right) = \int_0^1 B_{X.x}B_{0.x},$$

where $B_{X.x}(r) = B_X(r) - \left(\int_0^1 B_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1} B_x(r)$, the L_2 projection residual of B_X on B_x. It follows that

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

as stated in (i). Note that

$$\int_0^1 B_{X.x}B_{0.x} = \int_0^1 \overrightarrow{B_{X.x}}(r)\,dB_{0.x}(r) \tag{59}$$

where $\overrightarrow{B_{X.x}}(r) := \int_r^1 B_{X.x}$. The representation (59) follows as in Lemma B (v) or by using reverse partial summation as in Lemma A. ■

Part (ii) 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 $\sum_{t=1}^{\lfloor n\cdot\rfloor} u_{0.xt}$. Instead, the following limits hold for the component sample moments:

(b-1) $n^{-1}\sum_{t=1}^{n} x_te_t \rightsquigarrow \int_0^1 B_xdB_e + \Delta_{xe}$, where $\Delta_{xe} = \sum_{h=0}^{\infty}\mathbb{E}(u_{x0}e_h)$;

(b-2) $n^{-3}\sum_{t=1}^{n} X_tx_t' \rightsquigarrow \int_0^1 B_XB_x'$;

(b-3) $n^{-4}X'Q_xX \rightsquigarrow \int_0^1 B_{X.x}B_{X.x}'$ as in (a-2);

(b-4) $n^{-2}X'Q_xe = \sum_{t=1}^{n}\frac{X_t}{n^{3/2}}\frac{e_t}{n^{1/2}} - \left(n^{-3}\sum_{t=1}^{n} X_tx_t'\right)\left(n^{-2}\sum_{t=1}^{n} x_tx_t'\right)^{-1}\left(n^{-1}\sum_{t=1}^{n} x_te_t\right) \rightsquigarrow \int_0^1 B_{X.x}dB_e - \left(\int_0^1 B_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\Delta_{xe},.$

Results (b-1)-(b-3) follow by standard manipulations as in Part (a). Setting $E_t = \sum_{s=1}^{t} e_s$, $E_0 = 0$, we have $n^{-1/2}E_{\lfloor nr\rfloor} \rightsquigarrow B_e(r)$, and to confirm (b-4) use partial summation to write

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

Then

$$\frac{1}{n^2}\sum_{t=1}^{n} X_te_t = \frac{X_n}{n^{3/2}}\frac{E_n}{n^{1/2}} - \frac{1}{n}\sum_{t=1}^{n}\frac{x_t}{n^{1/2}}\frac{E_{t-1}}{n^{1/2}} \tag{60}$$

$$\rightsquigarrow B_X(1)\,B_e(1) - \int_0^1 B_xB_e = \int_0^1 B_XdB_e, \tag{61}$$

by integration by parts since $B_X(r) = \int_0^r B_x$ is of bounded variation and so $\int_0^1 B_X dB_e = [B_X B_e]_0^1 - \int_0^1 B_x B_e = B_X(1) B_e(1) - \int_0^1 B_x B_e$. Using (b-1) and (61) we have

$$\begin{aligned} n^{-2} X' Q_x e &\rightsquigarrow \int_0^1 B_X dB_e - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x dB_e + \Delta_{xe}\right) \\ &= \int_0^1 B_{X.x} dB_e - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1} \Delta_{xe}, \end{aligned}$$

giving result (b-4). Combining (b-3) and (b-4) and using continuous mapping leads to the stated limit result

$$n^2(\hat{a} - a) \rightsquigarrow \left(\int_0^1 B_{X.x} B_{X.x}'\right)^{-1}\left\{\int_0^1 B_{X.x} dB_e - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1} \Delta_{xe}\right\}$$

giving (ii) for the limit distribution of $n^2(\hat{a} - a)$. ■

The following proofs refer to the augmented model given by (32) in the text, which is repeated here for convenience

$$Y_t = a' X_t + f' x_t + g' \Delta x_t + e_t^+ = a' X_t + f' x_t + g' u_{xt} + e_t^+, \tag{62}$$

where the regression error is $e_t^+ = e_t \mathbf{1}\{\Omega_{00.x} = 0\} + U_{0.xt} \mathbf{1}\{\Omega_{00.x} > 0\}$ and $U_{0.xt} = \sum_{s=1}^t u_{0.xs}$. We first derive limit results for the application of IM-OLS in this augmented model and then consider the use of fixed-K TIV regression and TIV regression with $K \to \infty$.

Proofs of (35) and (36)

When $\Omega_{00.x} > 0$, the system (62) is

$$Y_t = a' X_t + f' x_t + g' u_{xt} + U_{0.xt} =: a' X_t + d' w_t + U_{0.xt}, \tag{63}$$

with $d' = (f', g')$ and $w_t' = (x_t, u_{xt}')$. Least squares estimation of (63) gives

$$\hat{a} - a = (X' Q_W X)^{-1} X' Q_W U_{0.x}$$

where $Q_W = Q_x - Q_x u_x (u_x' Q_x u_x)^{-1} u_x' Q_x$ in standard notation with $Q_x = I_n - x(x'x)^{-1} x'$. Then $n(\hat{a} - a) = (n^{-4} X' Q_W X)^{-1} (n^{-3} X' Q_W U_{0.x})$ and the component limits follow by standard methods. In particular

$$\begin{aligned} \frac{1}{n} u_x' Q_x u_x &= \frac{1}{n} u_x' u_x - \frac{1}{n}\left(\frac{u_x' x}{n}\right)\left(\frac{x'x}{n^2}\right)^{-1}\left(\frac{x' u_x}{n}\right) \to_p E(u_{xt} u_{xt}'), \\ \frac{1}{n} u_x' Q_x U_{0.x} &= \frac{u_x'}{\sqrt{n}} \frac{U_{0.x}}{\sqrt{n}} - \left(\frac{u_x' x}{n}\right)\left(\frac{x'x}{n^2}\right)^{-1}\left(\frac{x' U_{0.x}}{n^2}\right) \\ &\rightsquigarrow \int_0^1 dB_x B_{0.x} + \Delta_{0x}^+ - \left(\int_0^1 dB_x B_x' + \Delta_{xx}\right)\left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x B_{0.x}'\right), \end{aligned}$$

$$\frac{1}{n^2} X' Q_x u_x = \frac{X'}{n^{3/2}} \frac{u_x}{\sqrt{n}} - \left(\frac{1}{n} \frac{X'x}{n^2}\right)\left(\frac{x'x}{n^2}\right)^{-1}\left(\frac{x'}{\sqrt{n}} \frac{u_x}{\sqrt{n}}\right)$$

$$\begin{aligned}
&\rightsquigarrow \left(\int_0^1 B_X dB_x'\right) - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x dB_x' + \Delta_{xx}'\right) \\
&= \int_0^1 B_{X.x} dB_x' - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\Delta_{xx}', \\
\frac{1}{n^3}X'Q_xU_{0.x} &= \frac{1}{n}\frac{X'}{n^{3/2}}\frac{U_{0.x}}{n^{1/2}} - \left(\frac{1}{n}\frac{X'x}{n^2}\right)\left(\frac{x'x}{n^2}\right)^{-1}\left(\frac{1}{n}\frac{x'}{\sqrt{n}}\frac{U_{0.x}}{\sqrt{n}}\right) \\
&\rightsquigarrow \left(\int_0^1 B_X B_{0.x}'\right) - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x B_{0.x}'\right) \\
&= \int_0^1 B_{X.x} B_{0.x}',
\end{aligned}$$

where $\Delta_{0x}^+ = \Delta_{0x} - \Omega_{0x}\Omega_{xx}^{-1}\Delta_{xx} = \Delta_{0x} - f'\Delta_{xx}$. Using these results and (b-3) above we obtain

$$n^{-4}X'Q_WX = n^{-4}X'Q_xX - \frac{1}{n}\left(\frac{1}{n^2}X'Q_xu_x\right)\left(\frac{1}{n}u_x'Q_xu_x\right)^{-1}\left(\frac{1}{n^2}u_x'Q_xX\right) \rightsquigarrow \int_0^1 B_{X.x}B_{X.x}', \tag{64}$$

and

$$\begin{aligned}
n^{-3}X'Q_WU_{0.x} &= n^{-3}X'Q_xU_{0.x} - n^{-3}X'Q_xU_x\left(U_x'Q_xU_x\right)^{-1}U_x'Q_xU_{0.x} \\
&= n^{-3}X'Q_xU_{0.x} - \frac{1}{n}\left(\frac{1}{n^2}X'Q_xU_x\right)\left(\frac{1}{n}U_x'Q_xU_x\right)^{-1}\left(\frac{1}{n}U_x'Q_xU_{0.x}\right) \\
&\rightsquigarrow \int_0^1 B_{X.x}B_{0.x}'.
\end{aligned}$$

It follows that

$$\begin{aligned}
n(\hat{a}-a) &= \left(n^{-4}X'Q_xX\right)^{-1}\left(n^{-3}X'Q_xU_{0.x}\right) \rightsquigarrow \left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1 B_{X.x}B_{0.x}\right) \\
&\equiv \mathcal{MN}\left(0, \Omega_{00.x}\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1 \overrightarrow{B_{X.x}}(r)\,\overrightarrow{B_{X.x}}(r)'\,dr\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\right),
\end{aligned}$$

which is identical to the limit result for the IM-OLS estimator in the $\Omega_{00.x} > 0$ case. Thus, inclusion of the surplus and irrelevant regressor u_{xt} in the fitted model (62) has no effect on the limit theory of IM-OLS in the base case of cointegrating regression. This proves (35).

To prove (36), consider the singular case where $\Omega_{00.x} = 0$. The system (62) is now

$$Y_t = a'X_t + f'x_t + g'u_{xt} + U_{0.xt} =: a'X_t + d'w_t + e_t, \tag{65}$$

with $d' = (f', g')$ and $w_t' = (x_t, u_{xt}')$. Least squares estimation of (63) now gives $\hat{a} - a = (X'Q_WX)^{-1}X'Q_We$. Then $n^2(\hat{a}-a) = \left(n^{-4}X'Q_WX\right)^{-1}\left(n^{-2}X'Q_We\right)$ whose component limits are as follows

$$\begin{aligned}
\frac{1}{n}u_x'Q_xu_x &= \frac{1}{n}u_x'u_x - \frac{1}{n}\left(\frac{u_x'x}{n}\right)\left(\frac{x'x}{n^2}\right)^{-1}\left(\frac{x'u_x}{n}\right) \to_p \mathbb{E}\left(u_{xt}u_{xt}'\right) = \Sigma_{xx}, \\
\frac{1}{n}u_x'Q_xe &= \frac{1}{n}u_x'e - \frac{1}{n}\left(\frac{u_x'x}{n}\right)\left(\frac{x'x}{n^2}\right)^{-1}\left(\frac{x'e}{n}\right) \to_p \mathbb{E}\left(u_{xt}e_t\right) = \sigma_{xe},
\end{aligned}$$

$$\frac{1}{n^2}X'Q_x u_x = \frac{X'}{n^{3/2}}\frac{u_x}{\sqrt{n}} - \left(\frac{1}{n}\frac{X'x}{n^2}\right)\left(\frac{x'x}{n^2}\right)^{-1}\left(\frac{x'}{\sqrt{n}}\frac{u_x}{\sqrt{n}}\right)$$
$$\rightsquigarrow \left(\int_0^1 B_X dB_x'\right) - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x dB_x' + \Delta_{xx}\right)$$
$$= \int_0^1 B_{X.x} dB_x' - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\Delta_{xx},$$
$$\frac{1}{n^2}X'Q_x e = \frac{X'}{n^{3/2}}\frac{e}{n^{1/2}} - \left(\frac{1}{n}\frac{X'x}{n^2}\right)\left(\frac{x'x}{n^2}\right)^{-1}\left(\frac{x'}{\sqrt{n}}\frac{e}{\sqrt{n}}\right)$$
$$\rightsquigarrow \left(\int_0^1 B_X dB_e\right) - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x dB_e + \Delta_{xe}\right)$$
$$= \int_0^1 B_{X.x} dB_e - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\Delta_{xe},$$

where $\Delta_{xx} = \sum_{h=0}^{\infty}\mathbb{E}\left(u_{x0}u_{xh}'\right)$, and $\Delta_{xe} = \sum_{h=0}^{\infty}\mathbb{E}\left(u_{x0}e_h\right)$. Using these results, (b-3), and $n^{-4}X'Q_W X \rightsquigarrow \int_0^1 B_{X.x}B_{X.x}'$ from (64), we have

$$\begin{aligned} & n^{-2}X'Q_W e = n^{-2}X'Q_x e - n^{-2}X'Q_x u_x\left(u_x'Q_x u_x\right)^{-1}u_x'Q_x e \\ = \quad & n^{-2}X'Q_x e - \left(\frac{1}{n^2}X'Q_x u_x\right)\left(\frac{1}{n}u_x'Q_x u_x\right)^{-1}\left(\frac{1}{n}u_x'Q_x e\right) \\ \rightsquigarrow \quad & \int_0^1 B_{X.x}dB_e - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\Delta_{xe} \\ & -\left\{\int_0^1 B_{X.x}dB_x' - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\Delta_{xx}'\right\}\Sigma_{xx}\sigma_{xe}. \end{aligned} \tag{66}$$

It follows from (64) and (66) that

$$\begin{aligned} & n^2\left(\hat{a}-a\right) = \left(n^{-4}X'Q_W X\right)^{-1}\left(n^{-2}X'Q_W e\right) \\ \rightsquigarrow \quad & \left(\int_0^1 B_{X.x}B_{X.x}'\right)\left\{\int_0^1 B_{X.x}dB_e - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\Delta_{xe}\right. \\ & \left. -\left\{\int_0^1 B_{X.x}dB_x' - \left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\Delta_{xx}\right\}\Sigma_{xx}\sigma_{xe}\right\}, \end{aligned}$$

as stated in (36). ■

Proof of Theorem 2

Part (iii) The proof proceeds as follows. In this case $\Omega_{00.x} > 0$ and $e_t^+ = U_{0.xt}$ so we have

$$n\left(\hat{a}_{fTIV} - a\right) = \left(\frac{1}{n^5}V_X'Q_{V_C}V_X\right)^{-1}\left(\frac{1}{n^4}V_X'Q_{V_C}V_{U_{0.x}}\right).$$

By standard partitioned regression, $Q_{V_C} = I - V_C\left(V_C'V_C\right)^{-1}V_C' = Q_{u_x} - Q_{u_x}V_x\left(V_x'Q_{u_x}V_x\right)^{-1}V_x'Q_{u_x}$, so that by the results in Lemma B for the component factors we have

$$n^{-5}V_X'Q_{V_C}V_X = n^{-5}V_X'\left\{Q_{u_x} - Q_{u_x}V_x\left(V_x'Q_{u_x}V_x\right)^{-1}V_x'Q_{u_x}\right\}V_X$$

$$
\begin{aligned}
&= \frac{V_X'}{n^{5/2}}\frac{V_X}{n^{5/2}} - \frac{V_X'}{n^{5/2}}\frac{V_{\Delta x}}{n^{1/2}}\left(\frac{V_{\Delta x}'}{n^{1/2}}\frac{V_{\Delta x}}{n^{1/2}}\right)^{-1}\frac{V_{\Delta x}'}{n^{1/2}}\frac{V_X}{n^{5/2}} - \frac{1}{n^5}V_X'Q_{u_x}V_x\left(V_x'Q_{u_x}V_x\right)^{-1}V_x'Q_{u_x}V_X' \\
&\rightsquigarrow \mu_K'\mu_K - \mu_K'\xi_K\left(\xi_K'\xi_K\right)^{-1}\xi_K'\mu_K - \mu_K'Q_{\xi_K}\eta_K\left(\eta_K'Q_{\xi_K}\eta_K\right)^{-1}\eta_K'Q_{\xi_K}\mu_K \\
&= \mu_K'J_K\mu_K, \qquad (67)
\end{aligned}
$$

where $\mu_K = \int_0^1 \tilde{\varphi}_K B_X'$, $\eta_K = \int_0^1 \tilde{\varphi}_K B_x'$, and

$$
J_K := Q_{\xi_K} - Q_{\xi_K}\eta_K\left(\eta_K'Q_{\xi_K}\eta_K\right)^{-1}\eta_K'Q_{\xi_K},
$$

with $Q_{\xi_K} = I - \xi_K\left(\xi_K'\xi_K\right)^{-1}\xi_K'$ and $\xi_K = \int_0^1 \tilde{\varphi}_K dB_x'$. In a similar way, using Lemma B we have

$$
\begin{aligned}
&\frac{1}{n^4}V_X'\left\{Q_{u_x} - Q_{u_x}V_x\left(V_x'Q_{u_x}V_x\right)^{-1}V_x'Q_{u_x}\right\}V_{U_{0.x}} \\
&= \frac{V_X'}{n^{5/2}}Q_{u_x}\frac{V_{U_{0.x}}}{n^{3/2}} - \frac{V_X'}{n^{5/2}}Q_{u_x}\frac{V_x}{n^{3/2}}\left(\frac{V_x'}{n^{3/2}}Q_{u_x}\frac{V_x}{n^{3/2}}\right)^{-1}\frac{V_x'}{n^{3/2}}Q_{u_x}\frac{V_{U_{0.x}}}{n^{3/2}} \\
&\rightsquigarrow \mu_K'Q_{\xi_K}\Psi_{0.xK} - \mu_K'Q_{\xi_K}\eta_K\left(\eta_K'Q_{\xi_K}\eta_K\right)^{-1}\eta_K'Q_{\xi_K}\Psi_{0.xK} \\
&= \mu_K J_K \Psi_{0.xK}.
\end{aligned}
$$

Now,

$$
\Psi_{0.xK} = \int_0^1 \tilde{\varphi}_K B_{0.x} \equiv \mathcal{N}\left(0, \Omega_{00.x}\left(\int_0^1\int_0^1 (r\wedge s)\,\tilde{\varphi}_K(r)\,\tilde{\varphi}_K(s)'\,drds\right)\right),
$$

as

$$
\mathbb{E}\left(\Psi_{0.xK}\Psi_{0.xK}'\right) = \int_0^1\int_0^1 \tilde{\varphi}_K \mathbb{E}\left\{B_{0.x}(r)\,B_{0.x}(s)\right\}\tilde{\varphi}_K' drds = \Omega_{00.x}\int_0^1\int_0^1 (r\wedge s)\,\tilde{\varphi}_K(r)\,\tilde{\varphi}_K(s)'\,drds.
$$

It follows that when $\Omega_{00.x} > 0$

$$
\begin{aligned}
& n\left(\hat{a}_{fTIV} - a\right) \rightsquigarrow \left(\mu_K'J_K\mu_K\right)^{-1}\left(\mu_K'J_K\Psi_{0.xK}\right) = S_K'\Psi_{0.xK} \\
\equiv\ & \mathcal{MN}\left(0, \Omega_{00.x}S_K'\left(\int_0^1\int_0^1 (r\wedge s)\,\tilde{\varphi}_K(r)\,\tilde{\varphi}_K(s)'\,drds\right)S_K\right), \qquad (68)
\end{aligned}
$$

where $S_K = J_K\mu_K\left(\mu_K'J_K\mu_K\right)^{-1}$. Mixed normality follows by virtue of the asymptotic independence of $\Psi_{0.xK}$ and (μ_K, ξ_K, η_K). ■

Part (iv). In this case $\Omega_{00.x} = 0$ and $e_t^+ = e_t$ so we have

$$
n^2\left(\hat{a}_{fTIV} - a\right) = \left(\frac{1}{n^5}V_X'Q_{V_C}V_X\right)^{-1}\left(\frac{1}{n^3}V_X'Q_{V_C}V_e\right). \qquad (69)
$$

As in (67) of Part (iii), $n^{-5}V_X'Q_{V_C}V_X \rightsquigarrow \mu_K'J_K\mu_K$. The second component of (69) is

$$
\begin{aligned}
\frac{1}{n^3}V_X'Q_{V_C}V_e &= \frac{V_X'}{n^{5/2}}\frac{V_e}{n^{1/2}} - \frac{V_X'}{n^{5/2}}\frac{U_x}{n^{1/2}}\left(\frac{U_x'}{n^{1/2}}\frac{U_x}{n^{1/2}}\right)^{-1}\frac{U_x'}{n^{1/2}}\frac{V_e}{n^{1/2}} \\
&- \frac{V_X'}{n^{5/2}}Q_{u_x}\frac{V_x}{n^{3/2}}\left(\frac{V_x'}{n^{3/2}}Q_{u_x}\frac{V_x}{n^{3/2}}\right)^{-1}\frac{V_x'}{n^{3/2}}Q_{u_x}\frac{V_e}{n^{1/2}} \\
&\rightsquigarrow \mu_K'J_K\psi_{eK}.
\end{aligned}
$$

Combining these factors and using continuous mapping we deduce that

$$n^2 (\hat{a}_{fTIV} - a) \rightsquigarrow (\mu_K' J_K \mu_K)^{-1} (\mu_K' J_K \psi_{eK}) = S_K' \psi_{eK}, \tag{70}$$

as stated in (iv). Note that this limit distribution is not mixed normal because when $\omega_{ex} \neq 0$, the component $\psi_{eK} = \int_0^1 \tilde{\varphi}_K dB_e$ is not independent (ξ_K, μ_K, η_K), all of which depend on B_x which is correlated with B_e. However, when $\omega_{ex} = 0$, the component ψ_{eK} is independent of (ξ_K, μ_K, η_K) and mixed normality holds, giving part (iv)*. In particular

$$\begin{aligned} n^2 (\hat{a}_{fTIV} - a) \quad &\rightsquigarrow \quad S_K' \psi_{eK} \equiv \mathcal{MN}\left(0, \omega_{ee} S_K' \left(\int_0^1 \tilde{\varphi}_K \tilde{\varphi}_K'\right) S_K\right) \\ &\equiv \quad \mathcal{MN}\left(0, \omega_{ee} (\mu_K' J_K \mu_K)^{-1}\right), \end{aligned}$$

since $S_K = J_K \mu_K (\mu_K' J_K \mu_K)^{-1}$ and $\left(\int_0^1 \tilde{\varphi}_K \tilde{\varphi}_K'\right) = I_K$. ■

Proof of Theorem 3

Part (v)

The proof follows a general line of argument that was developed in the proof of the main theorem of Phillips (2014) but with considerable additional complications in the present case arising from the more complex augmented model and the singularity in the conditional long run variance matrix. To facilitate the development of joint $(K, n) \to \infty$ asymptotics we use an expansion of the probability space that includes the limit processes $(B_e, B_x, B_{0.x})$ and within that space use an 'in probability' version of weak convergence to the limit Brownian motions $(B_e, B_x, B_{0.x})$, as in Lemma A of Phillips (2014) or Lemma C of Phillips (2007). This device leads in the usual manner to the establishment of weak convergence in the original space.

In the present case the full TIV approach projects the entire aggregated system

$$Y_t = a' X_t + f' x_t + g' \Delta x_t + e_t^+ = a' X_t + f' x_t + g' u_{xt} + e_t^+, \tag{71}$$

onto the range space of the instruments Φ_K using $P_{\Phi_K} = \Phi_K (\Phi_K' \Phi_K)^{-1} \Phi_K'$. When $\Omega_{00.x} > 0$ the regression error is $e_t^+ = U_{0.xt} = \sum_{s=1}^t u_{0.xs}$ and then (71) is

$$Y_t = a' X_t + f' x_t + g' u_{xt} + U_{0.xt} =: a' X_t + \ell' c_{xt} + U_{0.xt}, \tag{72}$$

where $c_t' = (x_t', u_{xt}')$ and $\ell' = (f', g') = (f', 0)$ since the true value of the coefficient of u_{xt} is zero. Write (72) in observation matrix form as

$$Y = [X, C_x] \gamma + U_{0.x}, \text{ with } \gamma' = (a', \ell'),$$

where $Y' = [Y_1, ... Y_n]$, $X' = [X_1, ..., X_n]$, $U_{0.x}' = (U_{0.x,1}, ..., U_{0.x,n})$, and

$$C_x' = [c_{x1}, ..., c_{xn}] = \begin{bmatrix} x_1 & \cdots & x_n \\ u_{x1} & \cdots & u_{xn} \end{bmatrix} =: \begin{bmatrix} x' \\ u_x' \end{bmatrix},$$

noting that u_x is the matrix of observations of $u_{xt} = \Delta x_t$, in contrast to the vector of partial sums $U_{0.x}$. The centred and suitably scaled TIV estimator of a then has the following form

$$
\begin{aligned}
n\left(\hat{a}_{TIV} - a\right) &= \left\{\frac{1}{n^4} X' R_K X\right\}^{-1} \left\{\frac{1}{n^3} X' R_K U_{0.x}\right\} \\
&= \left\{\frac{1}{n^4} X' P_{\Phi_K} X - \left(\frac{1}{n^2} X' P_{\Phi_K} C_x F_n\right) \left(F_n C_x' P_{\Phi_K} C_x F_n\right)^{-1} \left(F_n C_x' P_{\Phi_K} X \frac{1}{n^2}\right)\right\}^{-1} \times \\
&\quad \left\{\frac{1}{n^3} X' P_{\Phi_K} U_{0.x} - \frac{1}{n^3} \left(X' P_{\Phi_K} C_x F_n\right) \left(F_n C_x' P_{\Phi_K} C_x F_n\right)^{-1} F_n C_x' P_{\Phi_K} U_{0.x}\right\}, \qquad (73)
\end{aligned}
$$

where $F_n = \operatorname{diag}\left[n^{-1} I_{m_x}, K^{-1/2} I_{m_x}\right]$. We now proceed to derive the limit theory for the two major factors in this matrix quotient.

The first factor in braces in (73) is

$$
\frac{1}{n^4} X' P_{\Phi_K} X - \left(\frac{1}{n^2} X' P_{\Phi_K} \left[\begin{array}{cc} x & u_x \end{array}\right] F_n\right) \left(F_n \left[\begin{array}{cc} x' P_{\Phi_K} x & x' P_{\Phi_K} u_x \\ u_x' P_{\Phi_K} x & u_x' P_{\Phi_K} u_x \end{array}\right] F_n\right)^{-1} \left(F_n \left[\begin{array}{c} x' P_{\Phi_K} X \\ u_x' P_{\Phi_K} X \end{array}\right] \frac{1}{n^2}\right), \qquad (74)
$$

and the components of (74) are now considered in turn. Proceeding as in the proof of equation (34) of the main theorem and Lemmas B and D of Phillips (2014), we find that as $(K, n) \to \infty$ with $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$

$$
\begin{aligned}
\frac{1}{n^4} X' P_{\Phi_K} X &= \frac{1}{n} \frac{X' \Phi_K}{n^2} \left\{I_K + O\left(\frac{1}{n}\right)\right\} \frac{\Phi_K' X}{n^2} \\
&= \frac{X' \Phi_K}{n^{5/2}} \frac{\Phi_K' X}{n^{5/2}} \left\{1 + O_p\left(\frac{1}{n}\right)\right\} \rightsquigarrow \int_0^1 B_X B_X'. \qquad (75)
\end{aligned}
$$

To show (75) we use the $a.s.$ convergent series representation $B_X(r) = \sum_{m=1}^{\infty} \varphi_m(r) \nu_m$ of the continuous stochastic process $\breve{B}_x$ in terms of the orthonormal sequence $\{\varphi_m\}_{m=1}^{\infty}$ over $[0,1]$. This series can be constructed by integrating the uniformly and almost surely convergent Karhunen-Loève series $B_x(r) = \sum_{k=1}^{\infty} \lambda_k \varphi_k(r) \xi_{xk}$, giving

$$
\begin{aligned}
B_X(r) &= \int_0^r B_x(s)\, ds = \sum_{k=1}^{\infty} \lambda_k \left(\int_0^r \varphi_k(s)\, ds\right) \xi_{xk} = \sum_{k=1}^{\infty} \lambda_k \psi_k(r) \xi_{xk} \\
&= \sum_{k=1}^{\infty} \lambda_k \left(\sum_{m=1}^{\infty} \delta_{km} \varphi_m(r)\right) \xi_{xk} = \sum_{m=1}^{\infty} \varphi_m(r) \left(\sum_{k=1}^{\infty} \delta_{km} \lambda_k \xi_{xk}\right) \\
&= \sum_{m=1}^{\infty} \varphi_m(r) \nu_m, \qquad (76)
\end{aligned}
$$

with $\psi_k(r) = \int_0^r \varphi_k(r) = \sum_{m=1}^{\infty} \delta_{km} \varphi_m(r)$ in which each $\psi_k(r)$ is represented by its expansion in terms of the ON functions $\{\varphi_j\}$, and the random sequence ν_m is defined by $\nu_m = \sum_{k=1}^{\infty} \delta_{km} \lambda_k \xi_{xk}$. Using this representation of the process $B_X(r)$ and the expanded probability space, we have

$$
\frac{1}{n} \sum_{t=1}^{n} \frac{X_t}{n^{3/2}} \tilde{\varphi}_K \left(\frac{t}{n}\right)' = \int_0^1 B_X(r) \tilde{\varphi}_K(r)'\, dr \left\{1 + o_p(1)\right\}
$$

$$
\begin{aligned}
&= \int_0^1 \left\{ \sum_{m=1}^{\infty} \varphi_m(r)\, \nu_m \right\} \tilde{\varphi}_K(r)'\, dr \{1 + o_p(1)\} \\
&= \sum_{m=1}^{K} \nu_m \int_0^1 \varphi_m(r)\, \tilde{\varphi}_K(r)'\, dr \{1 + o_p(1)\} + \sum_{m=K+1}^{\infty} \nu_m \int_0^1 \varphi_m(r)\, \tilde{\varphi}_K(r)'\, dr \{1 + o_p(1)\} \\
&= \sum_{m=1}^{K} \nu_m \int_0^1 \varphi_m(r)\, \tilde{\varphi}_K(r)'\, dr \{1 + o_p(1)\} = \mathcal{V}_K \{1 + o_p(1)\},
\end{aligned}
$$

where $\mathcal{V}_K = [\nu_1, ..., \nu_K]$. Standardizing the matrix quadratic form $X' P_{\Phi_K} X$ and allowing $(K, n) \to \infty$ we have

$$
\begin{aligned}
\frac{1}{n^4} X' P_{\Phi_K} X = \frac{X' \Phi_K}{n^{5/2}} \frac{\Phi_K' X}{n^{5/2}} \left\{ 1 + O_p\left(\frac{1}{n}\right) \right\} &= \mathcal{V}_K \mathcal{V}_K' \{1 + o_p(1)\} \\
= \sum_{m=1}^{K} \nu_m \nu_m' \{1 + o_p(1)\} = \int_0^1 B_X B_X' \{1 + o_p(1)\} \qquad (77)
\end{aligned}
$$

because

$$
\int_0^1 B_X B_X' = \int_0^1 \left\{ \sum_{m=1}^{\infty} \varphi_m(r)\, \nu_m \right\} \left\{ \sum_{\ell=1}^{\infty} \varphi_\ell(r)\, \nu_\ell' \right\} dr = \sum_{m,\ell=1}^{\infty} \nu_m \nu_\ell' \int_0^1 \varphi_m(r)\, \varphi_\ell(r)\, dr = \sum_{m=1}^{\infty} \nu_m \nu_m',
$$

which is a series representation of $\int_0^1 B_X B_X'$ in terms of the component Gaussian vector variates $\{\nu_m\}_{m=1}^{\infty}$.

Moving to the second term of (74), consider the central matrix factor

$$
F_n \begin{bmatrix} x' P_{\Phi_K} x & x' P_{\Phi_K} u_x \\ u_x' P_{\Phi_K} X & u_x' P_{\Phi_K} u_x \end{bmatrix} F_n = \begin{bmatrix} \frac{1}{n^2} x' P_{\Phi_K} x & \frac{1}{n} x' P_{\Phi_K} u_x \frac{1}{K^{1/2}} \\ \frac{1}{K^{1/2}} u_x' P_{\Phi_K} x \frac{1}{n} & \frac{1}{K} u_x' P_{\Phi_K} u_x \end{bmatrix}. \qquad (78)
$$

First, as $(K, n) \to \infty$

$$
K^{-1} u_x' P_{\Phi_K} u_x \to_p \Omega_{xx}, \qquad (79)
$$

by Phillips (2005a; 2014, Lemma C). Further

$$
\begin{aligned}
\frac{1}{n} x' \Phi_K \left(\Phi_K' \Phi_K\right)^{-1} \Phi_K' u_x &= \left(\frac{1}{n} \frac{x'}{\sqrt{n}} \Phi_K \right) \left(\frac{1}{n} \Phi_K' \Phi_K \right)^{-1} \left(\frac{1}{\sqrt{n}} \Phi_K' u_x \right) \\
= \left(\frac{1}{n} \sum_{t=1}^{n} \frac{x_t}{\sqrt{n}} \tilde{\varphi}_{Kt}' \right) &\left\{ I_K + O\left(\frac{1}{n}\right) \right\} \left(\sum_{t=1}^{n} \tilde{\varphi}_{Kt} \frac{u_{xt}'}{\sqrt{n}} \right) = O_p(1), \qquad (80)
\end{aligned}
$$

and

$$
\frac{1}{n^2} x' P_{\Phi_K} x \rightsquigarrow \int_0^1 B_x B_x', \qquad (81)
$$

as in equation (46) of Phillips (2014). Thus, as $(K, n) \to \infty$

$$
\begin{aligned}
\frac{1}{K} u_x' P_{\Phi_K} u_x &= \frac{1}{K} \frac{u_x' \Phi_K}{\sqrt{n}} \frac{\Phi_K' u_x}{\sqrt{n}} \to_p \Omega_{xx} \qquad (82) \\
\frac{1}{n} x' P_{\Phi_K} u_x \frac{1}{K^{1/2}} &= O_p\left(\frac{1}{K^{1/2}} \right) = o_p(1), \qquad (83)
\end{aligned}
$$

$$\frac{1}{n^2}x'P_{\Phi_K}x \quad \rightsquigarrow \quad \int_0^1 B_xB_x', \tag{84}$$

so that, at least to first order, we have

$$F_n\begin{bmatrix} x'P_{\Phi_K}x & x'P_{\Phi_K}u_x \\ u_x'P_{\Phi_K}x & u_x'P_{\Phi_K}u_x \end{bmatrix}F_n \rightsquigarrow \begin{bmatrix} \int_0^1 B_xB_x' & 0 \\ 0 & \Omega_{xx} \end{bmatrix}. \tag{85}$$

Higher order terms in the off diagonal elements of this matrix will be needed and constructed later in analyzing the second major factor of (73).

Next consider the first matrix factor in the matrix quadratic form in the second term of (74), viz.,

$$\begin{aligned} &\frac{1}{n^2}X'P_{\Phi_K}\begin{bmatrix} x & u_x \end{bmatrix}F_n = \begin{bmatrix} \frac{1}{n^3}X'P_{\Phi_K}x & \frac{1}{n^2}X'P_{\Phi_K}u_x\frac{1}{K^{1/2}} \end{bmatrix} \\ = & \begin{bmatrix} \left(\frac{1}{n}\frac{X'}{n^{3/2}}\Phi_K\right)\left(\frac{1}{n}\Phi_K'\Phi_K\right)^{-1}\left(\frac{1}{n}\Phi_K'\frac{x}{\sqrt{n}}\right) & \frac{1}{n^2}X'P_{\Phi_K}u_x\frac{1}{K^{1/2}} \end{bmatrix} \\ = & \begin{bmatrix} \left(\frac{1}{n}\frac{X'}{n^{3/2}}\Phi_K\right)\left\{I_K+O\left(\frac{1}{n}\right)\right\}\left(\frac{1}{n}\Phi_K'\frac{x}{\sqrt{n}}\right) & \frac{1}{n^2}X'P_{\Phi_K}u_x\frac{1}{K^{1/2}} \end{bmatrix} \\ \rightsquigarrow & \begin{bmatrix} \int_0^1 B_XB_x' & 0 \end{bmatrix}, \end{aligned} \tag{86}$$

where series arguments similar to those yielding (77) and (81) above are used to show that as $(K,n)\to\infty$

$$\frac{1}{n^3}X'P_{\Phi_K}x = \left(\frac{1}{n}\frac{X'}{n^{3/2}}\Phi_K\right)\left\{I_K+O\left(\frac{1}{n}\right)\right\}\left(\frac{1}{n}\Phi_K'\frac{x}{\sqrt{n}}\right) \rightsquigarrow \sum_{m=1}^{\infty}\lambda_m\nu_m\xi_m' = \int_0^1 B_XB_x', \tag{87}$$

Combining (85) and (86) we have

$$\begin{aligned} &\left(\frac{1}{n^2}X'P_{\Phi_K}\begin{bmatrix} x & u_x \end{bmatrix}F_n\right)\left(F_n\begin{bmatrix} x'P_{\Phi_K}x & x'P_{\Phi_K}u_x \\ u_x'P_{\Phi_K}x & u_x'P_{\Phi_K}u_x \end{bmatrix}F_n\right)^{-1}\left(F_n\begin{bmatrix} x'P_{\Phi_K}X \\ u_x'P_{\Phi_K}X \end{bmatrix}\frac{1}{n^2}\right) \\ \rightsquigarrow & \begin{bmatrix} \int_0^1 B_XB_x' & 0 \end{bmatrix}\begin{bmatrix} \int_0^1 B_xB_x' & 0 \\ 0 & \Omega_{xx} \end{bmatrix}^{-1}\begin{bmatrix} \int_0^1 B_xB_X' \\ 0 \end{bmatrix} \\ = & \left(\int_0^1 B_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\left(\int_0^1 B_xB_X'\right). \end{aligned} \tag{88}$$

Using (77) and (88) we then obtain

$$\int_0^1 B_XB_X' - \left(\int_0^1 B_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\left(\int_0^1 B_xB_X'\right) = \int_0^1 B_{X.x}B_{X.x}',$$

where $B_{X.x}(r) = B_X(r) - \left(\int_0^1 B_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}B_x(r)$ is the orthogonal projection residual of B_X on B_x. Thus,

$$\frac{1}{n^4}X'P_{\Phi_K}X - \left(\frac{1}{n^2}X'P_{\Phi_K}C_xF_n\right)\left(F_nC_x'P_{\Phi_K}C_xF_n\right)^{-1}\left(F_nC_x'P_{\Phi_K}X\frac{1}{n^2}\right) \rightsquigarrow \int_0^1 B_{X.x}B_{X.x}'. \tag{89}$$

Next we move to the second major factor in braces in equation (73), which we write as

$$\frac{1}{n^{3/2}}X'P_{\Phi_K}\frac{U_{0.x}}{n^{3/2}}-\left(\frac{1}{n^{3/2}}X'P_{\Phi_K}C_xF_n\right)\left(F_nC_x'P_{\Phi_K}C_xF_n\right)^{-1}\left(F_nC_x'P_{\Phi_K}\frac{U_{0.x}}{n^{3/2}}\right). \tag{90}$$

The first term of (90) is

$$\begin{aligned}\frac{1}{n^{3/2}}X'P_{\Phi_K}\frac{U_{0.x}}{n^{3/2}} &= \left(\frac{1}{n}\frac{X'}{n^{3/2}}\Phi_K\right)\left(\frac{1}{n}\Phi_K'\Phi_K\right)^{-1}\left(\frac{1}{n}\Phi_K'\frac{U_{0.x}}{\sqrt{n}}\right)\\ &= \left(\frac{1}{n}\frac{X'}{n^{3/2}}\Phi_K\right)\left\{I_K+O\left(\frac{1}{n}\right)\right\}\left(\frac{1}{n}\Phi_K'\frac{U_{0.x}}{\sqrt{n}}\right)\rightsquigarrow\int_0^1 B_XB_{0.x}',\end{aligned}$$

just as in (86). The second term of (90), ignoring the sign, reduces as follows

$$\begin{aligned}&\left(\frac{1}{n^{3/2}}X'P_{\Phi_K}C_xF_n\right)\left(F_nC_x'P_{\Phi_K}C_xF_n\right)^{-1}\left(F_nC_x'P_{\Phi_K}\frac{U_{0.x}}{n^{3/2}}\right)\\ =\;&\left(\frac{1}{n^2}X'P_{\Phi_K}\begin{bmatrix} x & u_x\end{bmatrix}F_n\right)\left(F_n\begin{bmatrix} x'P_{\Phi_K}x & x'P_{\Phi_K}u_x\\ u_x'P_{\Phi_K}X & u_x'P_{\Phi_K}u_x\end{bmatrix}F_n\right)^{-1}\left(F_n\begin{bmatrix} x'P_{\Phi_K}U_{0.x}\\ u_x'P_{\Phi_K}U_{0.x}\end{bmatrix}\frac{1}{n}\right)\\ \rightsquigarrow\;&\begin{bmatrix}\int_0^1 B_XB_x' & 0\end{bmatrix}\begin{bmatrix}\int_0^1 B_xB_x' & 0\\ 0 & \Omega_{xx}\end{bmatrix}^{-1}\begin{bmatrix}\int_0^1 B_xB_{0.x}\\ 0\end{bmatrix}\\ =\;&\left(\int_0^1 B_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\left(\int_0^1 B_xB_{0.x}\right),\end{aligned}$$

since

$$F_n\begin{bmatrix}\frac{1}{n^2}x'P_{\Phi_K}U_{0.x}\\ u_x'P_{\Phi_K}U_{0.x}\end{bmatrix}\frac{1}{n}=\begin{bmatrix}\frac{1}{\eta^2}x'P_{\Phi_K}U_{0.x}\\ \frac{1}{K^{1/2}n}u_x'P_{\Phi_K}U_{0.x}\end{bmatrix}=\begin{bmatrix}\frac{1}{\eta^2}x'P_{\Phi_K}U_{0.x}\\ \frac{1}{K^{1/2}n}u_x'P_{\Phi_K}U_{0.x}\end{bmatrix}$$

and

$$\begin{aligned}\frac{1}{n^2}x'P_{\Phi_K}U_{0.x} &= \frac{1}{n^2}x'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'U_{0.x}\\ &= \frac{x'}{n^{1/2}}\Phi_K\left\{I_K+O\left(\frac{1}{n}\right)\right\}\Phi_K'\frac{U_{0.x}}{n^{3/2}}\rightsquigarrow\int_0^1 B_xB_{0.x},\end{aligned}$$

analogous to (81), whereas

$$\begin{aligned}&\frac{1}{n}U_{0.x}'P_{\Phi_K}u_x\frac{1}{K^{1/2}}=\frac{1}{n}U_{0.x}'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'u_x\frac{1}{K^{1/2}}\\ =\;&\frac{1}{n^{3/2}}U_{0.x}'\Phi_K\left\{I_K+O\left(\frac{1}{n}\right)\right\}\Phi_K'\frac{u_x}{\sqrt{n}}\frac{1}{K^{1/2}}=O_p\left(\frac{1}{K^{1/2}}\right)=o_p(1)\end{aligned}$$

analogous to (80). We deduce that

$$\begin{aligned}&\frac{1}{n^{3/2}}X'P_{\Phi_K}\frac{U_{0.x}}{n^{3/2}}-\left(\frac{1}{n^{3/2}}X'P_{\Phi_K}C_xF_n\right)\left(F_nC_x'P_{\Phi_K}C_xF_n\right)^{-1}\left(F_nC_x'P_{\Phi_K}\frac{U_{0.x}}{n^{3/2}}\right)\\ \rightsquigarrow\;&\int_0^1 B_XB_{0.x}'-\left(\int_0^1 B_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\left(\int_0^1 B_xB_{0.x}\right)=\int_0^1 B_{X.x}B_{0.x}'\end{aligned} \tag{91}$$

It follows directly from (89) and (91) that

$$n\left(\hat{a}_{TIV}-a\right)\rightsquigarrow\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1 B_{X.x}B_{0.x}'\right)=\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\int_0^1\overrightarrow{B_{X.x}}dB_{0.x}$$

$$\equiv \quad \mathcal{MN}\left(0, \Omega_{00.x}\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\int_0^1 \overrightarrow{B_{X.x}}\,\overrightarrow{B_{X.x}}'\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\right), \tag{92}$$

where $\overrightarrow{B_{X.x}}(r) = \int_r^1 B_{X.x}$, using the same argument as in Lemma B(v) and (59). The limit distribution is therefore identical to that of IM-OLS in the case where $\Omega_{00.x} > 0$. Nothing is lost in the asymptotic theory in this case by working with the additional augmentation of the model to include the regressor $\Delta x_t = u_{xt}$. ■

Part (vi)

In this case the model is written in observation matrix form as

$$Y = [X, C_x]\,\gamma + e, \text{ with } \gamma' = (a', \ell'),$$

where $Y' = [Y_1, ...Y_n]$, $X' = [X_1, ..., X_n]$, $e' = (e_1, ..., e_n)$, $\ell = (f', g')'$ and

$$C_x' = [c_{x1}, ..., c_{xn}] = \begin{bmatrix} x_1 & \cdots & x_n \\ u_{x1} & \cdots & u_{xn} \end{bmatrix} =: \begin{bmatrix} x' \\ u_x' \end{bmatrix}.$$

The centred and scaled TIV estimator of a then has the form

$$\begin{aligned} & n^2\left(\hat{a}_{TIV} - a\right) \\ = \quad & \left\{\frac{1}{n^4}X'P_{\Phi_K}X - \left(\frac{1}{n^2}X'P_{\Phi_K}C_x\right)\left(C_x'P_{\Phi_K}C_x\right)^{-1}\left(C_x'P_{\Phi_K}X\frac{1}{n^2}\right)\right\}^{-1} \times \\ & \left\{\frac{1}{n^{3/2}}X'P_{\Phi_K}\frac{e}{n^{1/2}} - \left(\frac{1}{n^{3/2}}X'P_{\Phi_K}C_x\right)\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}\frac{e}{n^{1/2}}\right\}. \end{aligned} \tag{93}$$

We derive the limit theory for the two factors in this matrix quotient. The first factor in braces in (93) is identical to (74), and its limit behavior is therefore determined as in (89), giving

$$\begin{aligned} & \left\{\frac{1}{n^4}X'P_{\Phi_K}X - \left(\frac{1}{n^2}X'P_{\Phi_K}C_xF_n\right)\left(F_nC_x'P_{\Phi_K}C_xF_n\right)^{-1}\left(F_nC_x'P_{\Phi_K}X\frac{1}{n^2}\right)\right\} \\ & \rightsquigarrow \int_0^1 B_XB_X' - \left(\int_0^1 B_XB_x'\right)\left(\int_0^1 B_xB_x'\right)^{-1}\left(\int_0^1 B_xB_X'\right) = \int_0^1 B_{X.x}B_{X.x}', \end{aligned} \tag{94}$$

where $B_{X.x}(r) = B_X(r) - \int_0^1 B_XB_x\left(\int_0^1 B_xB_x'\right)^{-1}B_x(r)$, as before.

The second factor in braces in (93) is

$$\frac{1}{n^{3/2}}X'P_{\Phi_K}\frac{e}{n^{1/2}} - \left(\frac{1}{n^{3/2}}X'P_{\Phi_K}C_xF_n\right)\left(F_nC_x'P_{\Phi_K}C_xF_n\right)^{-1}F_nC_x'P_{\Phi_K}\frac{e}{n^{1/2}}. \tag{95}$$

For the first component of (95) note that

$$\frac{X'}{n^{3/2}}P_{\Phi_K}\frac{e}{n^{1/2}} = \frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'e}{n^{1/2}}\left\{1 + o_p(1)\right\}. \tag{96}$$

For the second component of (95) a more complex calculation is required. It turns out that because of the relative orders of magnitude of the submatrix elements in the matrix multiplication involved in this second term we need to include higher order terms in the inverse of the matrix $F_n C_x' P_{\Phi_K} C_x F_n$. To do so we employ the standard block inverse formula

$$\left[\begin{array}{cc} A_1'A_1 & A_1'A_2 \\ A_2'A_1 & A_2'A_2 \end{array}\right]^{-1} = \left[\begin{array}{cc} A_{11.2}^{-1} & -\left(A_1'A_1\right)^{-1} A_1'A_2 A_{11.2}^{-1} \\ -A_{22.1}^{-1} A_2'A_1 \left(A_1'A_1\right)^{-1} & \left(A_2'A_2\right)^{-1} + \left(A_2'A_2\right)^{-1} A_2'A_1 A_{11.2}^{-1} A_1'A_2 \left(A_2'A_2\right)^{-1} \end{array}\right]$$

which gives, using the fact that $C_x = [x \ , \ u_x]$,

$$\left(F_n \left[\begin{array}{cc} x'P_{\Phi_K}x & x'P_{\Phi_K}u_x \\ u_x'P_{\Phi_K}x & u_x'P_{\Phi_K}u_x \end{array}\right] F_n\right)^{-1} = \left[\begin{array}{cc} \frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'x}{n^{3/2}} & \frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'u_x}{n^{1/2}}\frac{1}{K^{1/2}} \\ \frac{1}{K^{1/2}}\frac{u_x'\Phi_K}{n^{1/2}}\frac{\Phi_K'x}{n^{3/2}} & \frac{1}{K}u_x'P_{\Phi_K}u_x \end{array}\right]^{-1}$$

$$=: \left[\begin{array}{cc} A_1'A_1 & A_1'A_2 \\ A_2'A_1 & A_2'A_2 \end{array}\right]^{-1} = \left[\begin{array}{cc} A_{11.2}^{-1} & -\left(A_1'A_1\right)^{-1} A_1'A_2 A_{11.2}^{-1} \\ -A_{22.1}^{-1} A_2'A_1 \left(A_1'A_1\right)^{-1} & A_{22.1}^{-1} \end{array}\right] \tag{97}$$

with $A_{11.2} = A_1'A_1 - A_1'A_2\left(A_2'A_2\right)^{-1} A_2'A_1$, and $A_{22.1} = A_2'A_2 - A_2'A_1\left(A_1'A_1\right)^{-1} A_1'A_2$. Using the results above, direct calculation of the block entries in the matrix (97) leads to

$$\begin{aligned} A_{11.2} &= \frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'x}{n^{3/2}} - \frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'u_x}{n^{1/2}}\frac{1}{K^{1/2}}\left(\frac{1}{K}u_x'P_{\Phi_K}u_x\right)^{-1}\frac{1}{K^{1/2}}\frac{u_x'\Phi_K}{n^{1/2}}\frac{\Phi_K'x}{n^{3/2}} \\ &= \frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'x}{n^{3/2}} - \frac{1}{K^{1/2}}\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'u_x}{n^{1/2}}\left(\frac{1}{K}u_x'P_{\Phi_K}u_x\right)^{-1}\left(\frac{1}{K^{1/2}}\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'u_x}{n^{1/2}}\right)' \\ &= \frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'x}{n^{3/2}} + O_p\left(\frac{1}{K}\right) \rightsquigarrow \int_0^1 B_x B_x', \end{aligned}$$

$$\begin{aligned} A_{22.1} &= \frac{1}{K}u_x'P_{\Phi_K}u_x - \frac{1}{K^{1/2}}\frac{u_x'\Phi_K}{n^{1/2}}\frac{\Phi_K'x}{n^{3/2}}\left(\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'x}{n^{3/2}}\right)^{-1}\left(\frac{1}{K^{1/2}}\frac{X'\Phi_K}{n^{3/2}}\frac{\Phi_K'u_x}{n^{1/2}}\right) \\ &= \frac{1}{K}u_x'P_{\Phi_K}u_x + O_p\left(\frac{1}{K}\right) \to_p \Omega_{xx}, \end{aligned}$$

and the off diagonal block entry

$$\begin{aligned} & -A_{22.1}^{-1}A_2'A_1\left(A_1'A_1\right)^{-1} \\ &= -\left\{\left(\frac{1}{K}u_x'P_{\Phi_K}u_x\right)^{-1} + O_p\left(\frac{1}{K}\right)\right\}\left(\frac{1}{K^{1/2}}\frac{u_x'\Phi_K}{n^{1/2}}\frac{\Phi_K'x}{n^{3/2}}\right)\left\{\left(\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'x}{n^{3/2}}\right)^{-1} + O_p\left(\frac{1}{K}\right)\right\} \\ &= -\frac{1}{K^{1/2}}\Omega_{xx}^{-1}\left(\frac{u_x'\Phi_K}{n^{1/2}}\frac{\Phi_K'x}{n^{3/2}}\right)\left(\int_0^1 B_x B_x'\right)^{-1}\left\{1 + o_p(1)\right\}. \end{aligned}$$

Then

$$\begin{aligned} & \left(F_n \left[\begin{array}{cc} x'P_{\Phi_K}x & x'P_{\Phi_K}u_x \\ u_x'P_{\Phi_K}x & u_x'P_{\Phi_K}u_x \end{array}\right] F_n\right)^{-1} \\ & \sim_a \left[\begin{array}{cc} \left(\int_0^1 B_x B_x'\right)^{-1} & -\frac{1}{K^{1/2}}\left(\int_0^1 B_x B_x'\right)^{-1}\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'U_x}{n^{1/2}}\Omega_{xx}^{-1} \\ -\frac{1}{K^{1/2}}\Omega_{xx}^{-1}\left(\frac{U_x'\Phi_K}{n^{1/2}}\frac{\Phi_K'x}{n^{3/2}}\right)\left(\int_0^1 B_x B_x'\right)^{-1} & \Omega_{xx}^{-1} \end{array}\right]. \end{aligned} \tag{98}$$

Retention of the $O_p\left(K^{-1/2}\right)$ off-diagonal blocks in (98) is particularly important, as will now become apparent. In particular, the second component of (95) is

$$-\left(\frac{1}{n^{3/2}}X'P_{\Phi_K}C_xF_n\right)\left(F_nC_x'P_{\Phi_K}C_xF_n\right)^{-1}\left(F_nC_x'P_{\Phi_K}\frac{e}{n^{1/2}}\right).$$

The component matrices involved in this block multiplication are

$$\frac{1}{n^{3/2}}X'P_{\Phi_K}C_xF_n=\left[\begin{array}{cc}\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'x}{n^{3/2}} & \frac{X'\Phi_K}{n^{5/2}}\Phi_K'u_x\frac{1}{K^{1/2}}\end{array}\right]+o_p(1),$$

$$(F_nC_x'P_{\Phi_K}C_xF_n)^{-1}=\left[\begin{array}{cc}\left(\int_0^1 B_xB_x'\right)^{-1}+o_p(1) & -\left(\int_0^1 B_xB_x'\right)^{-1}\frac{1}{K^{1/2}}\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'u_x}{n^{1/2}}\Omega_{xx}^{-1}\\ -\Omega_{xx}^{-1}\left(\frac{1}{K^{1/2}}\frac{u_x'\Phi_K}{n^{1/2}}\frac{\Phi_K'x}{n^{3/2}}\right)\left(\int_0^1 B_xB_x'\right)^{-1} & \Omega_{xx}^{-1}+o_p(1)\end{array}\right],$$

and

$$F_nC_x'P_{\Phi_K}\frac{e}{n^{1/2}}=\left[\begin{array}{c}\frac{1}{n}\frac{x'\Phi_K}{n^{1/2}}\frac{\Phi_K'e}{n^{1/2}}\\ \frac{1}{K^{1/2}}u_x'P_{\Phi_K}\frac{e}{n^{1/2}}\end{array}\right]+o_p(1).$$

Then

$$\begin{aligned}
&\left(\frac{1}{n^{3/2}}X'P_{\Phi_K}C_xF_n\right)(F_nC_x'P_{\Phi_K}C_xF_n)^{-1}F_nC_x'P_{\Phi_K}\frac{e}{n^{1/2}}\\
&=\left[\begin{array}{c}\left(\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'x}{n^{3/2}}\right)\left(\int_0^1 B_xB_x'\right)^{-1}-\frac{1}{K}\frac{X'\Phi_K}{n^{5/2}}\Phi_K'u_x\Omega_{xx}^{-1}\left(\frac{u_x'\Phi_K}{n^{1/2}}\frac{\Phi_K'x}{n^{3/2}}\right)\left(\int_0^1 B_xB_x'\right)^{-1}\\ -\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'x}{n^{3/2}}\left(\int_0^1 B_xB_x'\right)^{-1}\frac{1}{K^{1/2}}\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'u_x}{n^{1/2}}\Omega_{xx}^{-1}+\frac{X'\Phi_K}{n^{5/2}}\Phi_K'u_x\frac{1}{K^{1/2}}\Omega_{xx}^{-1}\end{array}\right]\\
&\times\left[\begin{array}{c}\frac{1}{n}\frac{x'\Phi_K}{n^{1/2}}\frac{\Phi_K'e}{n^{1/2}}\\ \frac{1}{K^{1/2}}u_x'P_{\Phi_K}\frac{e}{n^{1/2}}\end{array}\right]+o_p(1)\\
&=\left(\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'x}{n^{3/2}}\right)\left(\int_0^1 B_xB_x'\right)^{-1}\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'e}{n^{1/2}}\\
&-\frac{1}{K}\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'u_x}{n^{1/2}}\Omega_{xx}^{-1}\left(\frac{u_x'\Phi_K}{n^{1/2}}\frac{\Phi_K'x}{n^{3/2}}\right)\left(\int_0^1 B_xB_x'\right)^{-1}\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'e}{n^{1/2}}\\
&-\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'X}{n^{3/2}}\left(\int_0^1 B_xB_x'\right)^{-1}\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'u_x}{n^{1/2}}\Omega_{xx}^{-1}\frac{1}{K}u_x'P_{\Phi_K}e+\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'u_x}{n^{1/2}}\Omega_{xx}^{-1}\frac{1}{K}u_x'P_{\Phi_K}e+o_p(1)\\
&=\int_0^1 B_XB_x\left(\int_0^1 B_xB_x'\right)^{-1}\frac{X'\Phi_K}{n^{3/2}}\frac{\Phi_K'e}{n^{1/2}}+O_p\left(\frac{1}{K}\right)-\int_0^1 B_XB_x\left(\int_0^1 B_xB_x'\right)^{-1}\frac{X'\Phi_K}{n^{3/2}}\frac{\Phi_K'u_x}{n^{1/2}}\Omega_{xx}^{-1}\omega_{xe}\\
&+\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'u_x}{n^{1/2}}\Omega_{xx}^{-1}\omega_{xe}+o_p(1)\\
&=\left(\int_0^1 B_XB_x\right)\left(\int_0^1 B_xB_x'\right)^{-1}\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'\left(e-u_x'\Omega_{xx}^{-1}\omega_{xe}\right)}{n^{1/2}}+\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'u_x}{n^{1/2}}\Omega_{xx}^{-1}\omega_{xe}+o_p(1)\\
&=\left(\int_0^1 B_XB_x\right)\left(\int_0^1 B_xB_x'\right)^{-1}\left(\int_0^1 B_xdB_{e.x}\right)+\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'u_x}{n^{1/2}}\Omega_{xx}^{-1}\omega_{xe}+o_p(1), \qquad (99)
\end{aligned}$$

since

$$\frac{1}{K}u_x'P_{\Phi_K}e\to_p\omega_{xe},$$

just as in (79). Now combine (99) with (96) in (95) and we have

$$\begin{aligned}
&\frac{1}{n^{3/2}}X'P_{\Phi_K}\frac{e}{n^{1/2}}-\left(\frac{1}{n^{3/2}}X'P_{\Phi_K}C_xF_n\right)(F_nC_x'P_{\Phi_K}C_xF_n)^{-1}F_nC_x'P_{\Phi_K}\frac{e}{n^{1/2}}\\
&=\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'e}{\sqrt{n}}-\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'u_x}{\sqrt{n}}\Omega_{xx}^{-1}\omega_{xe}-\int_0^1 B_XB_x\left(\int_0^1 B_xB_x'\right)^{-1}\int_0^1 B_xdB_{e.x}+o_p(1)\\
&=\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K'\left(e-u_x\Omega_{xx}^{-1}\omega_{xe}\right)}{\sqrt{n}}-\int_0^1 B_XB_x\left(\int_0^1 B_xB_x'\right)^{-1}\int_0^1 B_xdB_{e.x}+o_p(1)\\
&\rightsquigarrow\int_0^1 B_XdB_{e.x}-\int_0^1 B_XB_x\left(\int_0^1 B_xB_x'\right)^{-1}\int_0^1 B_xdB_{e.x}=\int_0^1 B_{X.x}dB_{e.x}. \qquad (100)
\end{aligned}$$

Finally, combining the results for the two factors (94) and (100) and using continuous mapping we obtain the stated result that

$$n^2\left(\hat{a}_{TIV}-a\right) \rightsquigarrow \left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\int_0^1 B_{X.x}dB_{e.x} \equiv \mathcal{MN}\left(0,\omega_{ee.x}\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\right), \tag{101}$$

with $\omega_{ee.x}=\omega_{ee}-\omega_{ex}\Omega_{xx}\omega_{xe}$. ■

Proof of Theorem 5: Construction of the Wald statistic

We start the proof of Theorem 5 with some preliminary exposition of the two forms of the Wald statistic. The HAC Wald statistic $\text{Wald}_{TIV}=\left(H\hat{a}_{TIV}-h\right)'\left[HG_K\left(n\hat{V}_{Kn}\right)G_k'H'\right]^{-1}\left(H\hat{a}_{TIV}-h\right)$ uses an implicit sandwich form and relies on the kernel estimate

$$\hat{V}_{Kn}=\sum_{j=-M}^{M}k\left(\frac{j}{M}\right)\frac{1}{n}\sum_{1\le t,t+j\le n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t+j}{n}\right)'\hat{e}_t^+\hat{e}_{t+j}^+, \tag{102}$$

with TIV regression residuals

$$\hat{e}_t^+=e_t^+-\left(\hat{a}_{TIV}-a\right)'X_t-\left(\hat{f}_{TIV}-f\right)'x_t-\left(\hat{g}_{TIV}-g\right)'u_{xt}, \tag{103}$$

where $f=\Omega_{0x}\Omega_{xx}$, $g=0$ and the true regression error is $e_t^+=e_t\mathbf{1}\left\{\Omega_{00.x}=0\right\}+U_{0.xt}\mathbf{1}\left\{\Omega_{00.x}>0\right\}$. The limit behavior of the statistic Wald_{TIV} depends on that of the estimate $\hat{a}_{TIV}$, viz.,

$$\hat{a}_{TIV}-a=\left(X'R_KX\right)^{-1}\left(X'R_Ke^+\right)=G_K\Phi_K'e^+=G_K\sum_{t=1}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)e_t^+,$$

where $R_K=P_{\Phi_K}-P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}$ and

$$\begin{aligned} G_K &= \left(X'R_KX\right)^{-1}\left\{X'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}-X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\right\} \\ &= \left(X'R_KX\right)^{-1}\left\{X'\Phi_K-X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'\Phi_K\right\}\left(\Phi_K'\Phi_K\right)^{-1} \end{aligned}$$

as well as the estimation error effects of $\hat{f}_{TIV}-f$ and $\hat{g}_{TIV}-g=\hat{g}_{TIV}$ on the fitted residuals $\hat{e}_t^+$.

In the nonsingular case where $\Omega_{00.x}>0$, we have $e_t^+=U_{0.xt}$ so the true regression error is $I(1)$ and the regression equation is a partially spurious regression, as discussed in the text of the paper. Usual long run variance estimates of the equation error are therefore no longer consistent but tend to a random variable after suitable renormalization. The same is true for IM-OLS regression, a fact that substantially complicates inference in IM-OLS regression, as recognized in Vogelsang and Wagner (2014) and discussed in the main text – see footnote 7.

In the singular case where $\Omega_{00.x}=0$, we have $e_t^+=e_t$ and the regression equation is an augmented cointegrating regression with an $I(0)$ error. This equation no longer involves spurious elements and conventional methods of long run variance estimation work as usual. The limit behavior of the residuals and the kernel estimate $\hat{V}_{Kn}$ are therefore very different in these two cases, as is to be expected. They are examined separately below in **Part (vii)** and **Part (viii)** of the proof corresponding to the results given in the statement of Theorem 5. Similar considerations apply to the long run error variance estimator

$$\hat{\omega}_{e^+}^2 = \sum_{j=-M}^{M} k\left(\frac{j}{M}\right)\frac{1}{n}\sum_{1\le t,t+j\le n}^{n} \hat{e}_t^+\hat{e}_{t+j}^+, \tag{104}$$

based directly on the residuals (103).

The second form of the Wald statistic uses the HAR long run variance estimate leading to

$$\text{Wald}_{TIV,b} = \left(H\hat{a}_{TIV} - h\right)'\left[HG_K\left(n\hat{V}_{bKn}\right)G_k'H'\right]^{-1}\left(H\hat{a}_{TIV} - h\right),$$

which uses an implicit sandwich form with the fixed-b kernel estimate in (49)

$$\hat{V}_{bKn} = \sum_{j=-n+1}^{n-1} k_b(j)\frac{1}{n}\sum_{1\le t,t+j\le n}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t+j}{n}\right)'\hat{e}_t^+\hat{e}_{t+j}^+,$$

where $k_b(j) = k\left(\frac{j}{bn}\right)$. With these preliminaries in hand we now proceed with the proof of Theorem 5

Part (vii)

The HAC Case In this case $\Omega_{00.x} > 0$, $e_t^+ = U_{0.xt}$ and the fitted equation is a partially spurious regression because of the presence of the $I(1)$ regressor x_t and the $I(1)$ error $U_{0.xt}$ in the transformed model

$$Y_t = a'X_t + f'x_t + g'\Delta x_t + e_t^+ = a'X_t + f'x_t + g'u_{xt} + U_{0.xt}, \tag{105}$$

where $f = \Omega_{xx}^{-1}\Omega_{x0}$ and $g = 0$ by construction. The TIV regression produces consistent estimates of the cointegrating vector a, as shown in Theorem 3 (v), where

$$n\left(\hat{a}_{TIV} - a\right) = \left\{\frac{1}{n^4}X'R_KX\right\}^{-1}\left\{\frac{1}{n^3}X'R_KU_{0.x}\right\} \rightsquigarrow \left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\int_0^1 \overrightarrow{B_{X.x}}dB_{0.x}$$

$$\equiv \mathcal{MN}\left(0, \Omega_{00.x}\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\int_0^1 \overrightarrow{B_{X.x}}\,\overrightarrow{B_{X.x}}'\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\right). \tag{106}$$

But due to the spurious regression feature of (105), the estimates $\hat{f}_{TIV}$ and $\hat{g}_{TIV}$ of f and g are not consistent. In particular,

$$\hat{f}_{TIV} - f = \left(\frac{1}{n^2}x'R_{fK}x\right)^{-1}\left(\frac{1}{n^2}x'R_{fK}U_{0.x}\right),$$

where

$$\begin{aligned} R_{fK} &= P_{\Phi_K} - P_{\Phi_K}C_f\left(C_f'P_{\Phi_K}C_f\right)^{-1}C_f'P_{\Phi_K}, \\ C_f' &= \begin{bmatrix} X_1 & \cdots & X_n \\ u_{x1} & \cdots & u_{xn} \end{bmatrix} = \begin{bmatrix} X' \\ u_x' \end{bmatrix}. \end{aligned}$$

After some calculations, using $P_{\Phi_K} = \Phi_K (\Phi_K'\Phi_K)^{-1} \Phi_K' = \frac{1}{n}\Phi_K \left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1} \Phi_K' = \frac{1}{n}\Phi_K \left(I_K + O\left(\frac{1}{n}\right)\right) \Phi_K'$, we obtain

$$\frac{1}{n^2}x'R_{fK}x \rightsquigarrow \int_0^1 B_x B_x' - \int_0^1 B_x B_X' \left(\int_0^1 B_X B_X'\right)^{-1} \int_0^1 B_X B_x' = \int_0^1 B_{x.X} B_{x.X}', \qquad (107)$$

where $B_{x.X}(r) = B_x(r) - \int_0^1 B_x B_X' \left(\int_0^1 B_X B_X'\right)^{-1} B_X(r)$. Using the normalization matrix $L_n = diag\left[n^{-2} I_{m_x}, K^{-1/2} I_{m_x}\right]$, we have

$$\begin{aligned}
\frac{1}{n^2}x'R_{xK}U_{0.x} &= \left(\frac{1}{n^2}x'P_{\Phi_K}U_{0.x} - \frac{1}{n^2}x'P_{\Phi_K}C_f L_n \left(L_n C_f' P_{\Phi_K} C_f L_n\right)^{-1} L_n C_f' P_{\Phi_K} U_{0.x}\right) \\
&= \frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K' U_{0.x}}{n^{3/2}} - \frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K' C_f}{\sqrt{n}} L_n \left(L_n C_f' P_{\Phi_K} C_f L_n\right)^{-1} L_n \frac{C_f'\Phi_K}{\sqrt{n}}\frac{\Phi_K' U_{0.x}}{n^{3/2}} + o_p(1) \\
&\rightsquigarrow \int_0^1 B_x B_{0.x}' - \int_0^1 B_x B_X' \left(\int_0^1 B_X B_X'\right)^{-1} \int_0^1 B_X B_{0.x}' = \int_0^1 B_{x.X} B_{0.x}', \qquad (108)
\end{aligned}$$

since

$$L_n C_f' P_{\Phi_K} C_f L_n = \begin{bmatrix} \frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K' X}{n^{5/2}} & \frac{1}{K^{1/2}}\frac{X'\Phi_K}{n^{5/2}}\frac{\Phi_K' u_x}{n^{1/2}} \\ \frac{1}{K^{1/2}}\frac{u_x'\Phi_K}{n^{1/2}}\frac{\Phi_K' X}{n^{5/2}} & \frac{1}{K}\frac{u_x'\Phi_K}{n^{1/2}}\frac{\Phi_K' u_x}{n^{1/2}} \end{bmatrix} \rightsquigarrow \begin{bmatrix} \int_0^1 B_X B_X' & 0 \\ 0 & \Omega_{xx} \end{bmatrix}, \qquad (109)$$

and the terms involving u_x are $o_p(1)$ because $\frac{1}{K^{1/2}}\frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K' u_x}{\sqrt{n}} = O\left(K^{-1/2}\right)$. Then

$$\hat{f}_{TIV} - f = \left(\frac{1}{n^2}x'R_{fK}x\right)^{-1} \left(\frac{1}{n^2}x'R_{fK}U_{0.x}\right) \rightsquigarrow \left(\int_0^1 B_{x.X} B_{x.X}'\right)^{-1} \int_0^1 B_{x.X} B_{0.x}',$$

and $\hat{f}_{TIV}$ is inconsistent. Next consider $\hat{g}_{TIV}$. Since $g = 0$ by construction, we have

$$\hat{g}_{TIV} = \left(\frac{1}{K}u_x' R_{gK} u_x\right)^{-1} \left(\frac{1}{K}u_x' R_{gK} e^+\right) = \left(\frac{1}{K}u_x' R_{gK} u_x\right)^{-1} \left(\frac{1}{K}u_x' R_{gK} U_{0.x}\right), \qquad (110)$$

where

$$\begin{aligned}
R_{gK} &= P_{\Phi_K} - P_{\Phi_K} C_g \left(C_g' P_{\Phi_K} C_g\right)^{-1} C_g' P_{\Phi_K}, \\
C_g' &= \begin{bmatrix} X_1 & \cdots & X_n \\ x_1 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} X' \\ x' \end{bmatrix}.
\end{aligned}$$

Using the normalization matrix $D_n = diag\left[n^{-2} I_{m_x}, n^{-1} I_{m_x}\right]$ we have

$$\begin{aligned}
&\frac{1}{K}u_x' R_{gK} u_x = \frac{1}{K}u_x' P_{\Phi_K} u_x - \frac{1}{K}u_x' P_{\Phi_K} C_g D_n \left(D_n C_g' P_{\Phi_K} C_g D_n\right)^{-1} D_n C_g' P_{\Phi_K} u_x \\
= \;& \frac{1}{K}u_x' P_{\Phi_K} u_x - \left(\frac{1}{\sqrt{K}}\frac{u_x'\Phi_K}{\sqrt{n}}\right) \begin{bmatrix} \frac{1}{n}\frac{\Phi_K' X}{n^{3/2}} \\ \frac{1}{n}\frac{\Phi_K' x}{\sqrt{n}} \end{bmatrix} \left(D_n C_g' P_{\Phi_K} C_g D_n\right)^{-1} \begin{bmatrix} \frac{1}{n}\frac{X'\Phi_K}{n^{3/2}} \\ \frac{1}{n}\frac{x'\Phi_K}{\sqrt{n}} \end{bmatrix} \left(\frac{\Phi_K' u_x}{\sqrt{n}}\frac{1}{\sqrt{K}}\right) \\
= \;& \frac{1}{K}u_x' P_{\Phi_K} u_x + O_p\left(\frac{1}{K}\right) \to_p \Omega_{xx}.
\end{aligned}$$

Turning to the second factor in (110) and using the normalization matrix $D_n = diag\left[n^{-2} I_{m_x}, n^{-1} I_{m_x}\right]$ we have

$$\frac{1}{n}u_x' R_{gK} U_{0.x} = u_x' P_{\Phi_K} U_{0.x} - \frac{1}{n}u_x' P_{\Phi_K} C_g D_n \left(D_n C_g' P_{\Phi_K} C_g D_n\right)^{-1} D_n C_g' P_{\Phi_K} U_{0.x}$$

$$
\begin{aligned}
&= \frac{u_x'\Phi_K}{\sqrt{n}}\left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1}\frac{\Phi_K U_{0.x}}{n^{3/2}} - \left(\frac{u_x'\Phi_K}{\sqrt{n}}\right)\left[\begin{array}{c}\frac{1}{n}\frac{\Phi_K' X}{n^{3/2}} \\ \frac{1}{n}\frac{\Phi_K' x}{\sqrt{n}}\end{array}\right]\left(D_n C_g' P_{\Phi_K} C_g D_n\right)^{-1}\left[\begin{array}{c}\frac{1}{n}\frac{X'\Phi_K}{n^{3/2}} \\ \frac{1}{n}\frac{x'\Phi_K}{\sqrt{n}}\end{array}\right]\frac{\Phi_K U_{0.x}}{n^{3/2}} \\
&\rightsquigarrow \int_0^1 dB_x B_{0.x} - \int_0^1 dB_x B_{[X,x]}'\left(\int_0^1 B_{[X,x]}B_{[X,x]}'\right)^{-1}\int_0^1 B_{[X,x]}B_{0.x} = \int_0^1 dB_x B_{0.x}^{\#},
\end{aligned}
$$

where $B_{0.x}^{\#}(r) = B_{0.x}(r) - \int_0^1 B_{0.x}B_{[X,x]}'\left(\int_0^1 B_{[X,x]}B_{[X,x]}'\right)^{-1} B_{[X,x]}(r)$ and $B_{[X,x]} = [B_X, B_x]$. Then

$$
\hat{g}_{TIV} = \left(\frac{1}{K}u_x' R_{gK} u_x\right)^{-1}\left(\frac{1}{K}u_x' R_{gK} U_{0.x}\right) \sim_a \Omega_{xx}^{-1} \times O_p\left(\frac{n}{K}\right)
$$

so that the TIV regression residuals are

$$
\begin{aligned}
\hat{e}_t^+ &= \hat{U}_{0.xt} = U_{0.xt} - \left(\hat{a}_{TIV} - a\right)' X_t - \left(\hat{f}_{TIV} - f\right)' x_t - \hat{g}_{TIV}' u_{xt} \\
&= U_{0.xt} - n\left(\hat{a}_{TIV} - a\right)'\frac{X_t}{n} - \left(\hat{f}_{TIV} - f\right)' x_t - \hat{g}_{TIV}' u_{xt},
\end{aligned}
$$

and, standardizing, we have

$$
\begin{aligned}
\frac{\hat{e}_{t=\lfloor nr\rfloor}^+}{\sqrt{n}} &= \frac{U_{0.x\lfloor nr\rfloor}}{\sqrt{n}} - n\left(\hat{a}_{TIV} - a\right)'\frac{X_{\lfloor nr\rfloor}}{n^{3/2}} - \left(\hat{f}_{TIV} - f\right)'\frac{x_{\lfloor nr\rfloor}}{\sqrt{n}} - \hat{g}_{TIV}'\frac{u_{xt}}{\sqrt{n}} \\
&= \frac{U_{0.x\lfloor nr\rfloor}}{\sqrt{n}} - n\left(\hat{a}_{TIV} - a\right)'\frac{X_{\lfloor nr\rfloor}}{n^{3/2}} - \left(\hat{f}_{TIV} - f\right)'\frac{x_{\lfloor nr\rfloor}}{\sqrt{n}} + O_p\left(\frac{\sqrt{n}}{K}\right) \\
&= \frac{U_{0.x\lfloor nr\rfloor}}{\sqrt{n}} - n\left(\hat{a}_{TIV} - a\right)'\frac{X_{\lfloor nr\rfloor}}{n^{3/2}} - \left(\hat{f}_{TIV} - f\right)'\frac{x_{\lfloor nr\rfloor}}{\sqrt{n}} + o_p(1) \qquad (111)
\end{aligned}
$$

if $n = o\left(K^2\right)$ as $n \to \infty$. Then

$$
\begin{aligned}
&\frac{\hat{e}_{t=\lfloor nr\rfloor}^+}{\sqrt{n}} = \frac{U_{0.x\lfloor nr\rfloor}}{\sqrt{n}} - n\left(\hat{a}_{TIV} - a\right)'\frac{X_{\lfloor nr\rfloor}}{n^{3/2}} - \left(\hat{f}_{TIV} - f\right)'\frac{x_{\lfloor nr\rfloor}}{\sqrt{n}} + o_p(1) \\
&\rightsquigarrow B_{0.x}(r) - \left(\int_0^1 dB_{0.x}\overrightarrow{B_{X.x}}'\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1} B_X(r) - \left(\int_0^1 B_{0.x}B_{x.X}'\right)\left(\int_0^1 B_{x.X}B_{x.X}'\right)^{-1} B_x(r) \\
&= B_{0.x}(r) - \left(\int_0^1 B_{0.x}B_{X.x}'\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1} B_X(r) - \left(\int_0^1 B_{0.x}B_{x.X}'\right)\left(\int_0^1 B_{x.X}B_{x.X}'\right)^{-1} B_x(r) \\
&=: \widetilde{B_{0.x}}(r). \qquad (112)
\end{aligned}
$$

Using these results, the HAC kernel estimate of the transformed residuals and standardized residuals based on $\hat{e}_t^+ = \hat{U}_{0.xt}$ is

$$
\hat{V}_{Kn} = \sum_{j=-M}^{M} k\left(\frac{j}{M}\right)\frac{1}{n}\sum_{1\le t,t+j\le n}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t+j}{n}\right)'\hat{e}_t^+\hat{e}_{t+j}^+,
$$

and using (112)

$$
\frac{1}{nM}\hat{V}_{Kn} = \frac{1}{M}\sum_{j=-M}^{M} k\left(\frac{j}{M}\right)\frac{1}{n}\sum_{1\le t,t+j\le n}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t+j}{n}\right)'\frac{\hat{U}_{0.xt}}{\sqrt{n}}\frac{\hat{U}_{0.xt+j}}{\sqrt{n}}
$$

$$
\begin{aligned}
&= \left(\frac{1}{M}\sum_{j=-M}^{M} k\left(\frac{j}{M}\right)\right)\frac{1}{n}\sum_{t=1}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t}{n}\right)'\left(\frac{\hat{U}_{0.xt}}{\sqrt{n}}\right)^2 + o_p(1)\\
&= \left(\int_{-1}^{1} k(r)\,dr\right)\frac{1}{n}\Phi_K'\Lambda_{U_{0.x},n}\Phi_K\\
&\rightsquigarrow \left(\int_{-1}^{1} k(r)\,dr\right)\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\,\widetilde{B_{0.x}}(r)^2\,dr, \qquad (113)
\end{aligned}
$$

where $\Lambda_{U_{0.x},n} = diag\left[\left(\hat{U}_{0.x1}/\sqrt{n}\right)^2, ..., \left(\hat{U}_{0.xn}/\sqrt{n}\right)^2\right]$. Hence,

$$
n\hat{V}_{Kn} \sim_a n^2 M\left(\int_{-1}^{1} k(r)\,dr\right)\frac{1}{n}\Phi_K'\Lambda_{U_{0.x},n}\Phi_K \sim_a n^2 M\left(\int_{-1}^{1} k(r)\,dr\right)\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\,\widetilde{B_{0.x}}(r)^2\,dr. \qquad (114)
$$

Next observe that

$$
\begin{aligned}
&G_K\left(n\hat{V}_{Kn}\right)G_k' = nM\left(\int_{-1}^{1} k(r)\,dr\right)G_K\Phi_K'\Lambda_{U_{0.x},n}\Phi_K G_k' + o_p(1)\\
&= nM\left(\int_{-1}^{1} k(r)\,dr\right)(X'R_KX)^{-1}\left\{X'\Phi_K - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'\Phi_K\right\}\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'\Lambda_{U_{0.x},n}\\
&\times \Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\left\{\Phi_K'X - \Phi_K'C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}X\right\}(X'R_KX)^{-1} + o_p(1) \qquad (115)\\
&= nM\left(\int_{-1}^{1} k(r)\,dr\right)(X'R_KX)^{-1}\left(X'P_{\Phi_K} - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}\right)\Lambda_{U_{0.x},n}\\
&\times\left(P_{\Phi_K}X - P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}X\right)(X'R_KX)^{-1} + o_p(1)\\
&= nM\left(\int_{-1}^{1} k(r)\,dr\right)\left(\frac{X'R_KX}{n^4}\right)^{-1}\left(\frac{X'P_{\Phi_K} - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}}{n^4}\right)\Lambda_{U_{0.x},n}\\
&\times\left(\frac{P_{\Phi_K}X - P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}X}{n^4}\right)\left(\frac{X'R_KX}{n^4}\right)^{-1} + o_p(1). \qquad (116)
\end{aligned}
$$

Consider the first term in the expanded form of the central three factors in (116), viz.,

$$
\begin{aligned}
&E_{n0} := \frac{1}{n^4}X'P_{\Phi_K}\Lambda_{U_{0.x},n}P_{\Phi_K}X = \frac{1}{n^5}X'P_{\Phi_K}\Lambda_{U_{0.x},n}P_{\Phi_K}X\\
&= \frac{X'\Phi_K}{n^{5/2}}\left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1}\left(\frac{1}{n}\Phi_K'\Lambda_{U_{0.x},n}\Phi_K\right)\left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1}\frac{\Phi_K'X}{n^{5/2}}\\
&= \frac{X'\Phi_K}{n^{5/2}}\left(\frac{1}{n}\Phi_K'\Lambda_{U_{0.x},n}\Phi_K\right)\frac{\Phi_K'X}{n^{5/2}} + o_p(1)\\
&= \frac{X'\Phi_K}{n^{5/2}}\left(\frac{1}{n}\sum_{t=1}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t}{n}\right)'\left(\frac{\hat{U}_{0.xt}}{\sqrt{n}}\right)^2\right)\frac{\Phi_K'X}{n^{5/2}} + o_p(1)\\
&\sim_a \left(\int_0^1 B_X(r)\tilde{\varphi}_K(r)'\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\,\widetilde{B_{0.x}}(r)^2\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)B_X(r)'\,dr\right).
\end{aligned}
$$

Using the normalizing matrix $F_n = \text{diag}\left[n^{-1}I_{m_x}, K^{-1/2}I_{m_x}\right]$ as earlier, the cross product terms in the expanded form of the central three factors in braces in (116) are (ignoring the negative sign): first

$$
\begin{aligned}
E_{n1} &:= \frac{1}{n^4} X' P_{\Phi_K} C_x \left(C_x' P_{\Phi_K} C_x\right)^{-1} C_x' P_{\Phi_K} \Lambda_{U_{0.x},n} P_{\Phi_K} X \\
&= \left(\frac{1}{n^2} X' P_{\Phi_K} C_x F_n\right) \left(F_n C_x' P_{\Phi_K} C_x F_n\right)^{-1} \left(F_n C_x' P_{\Phi_K} \Lambda_{U_{0.x},n} P_{\Phi_K} X \frac{1}{n^2}\right) \\
&= \frac{X'\Phi_K}{n^{5/2}} \left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1} \left(\frac{\Phi_K' C_x F_n}{\sqrt{n}}\right) \left(F_n C_x' P_{\Phi_K} C_x F_n\right)^{-1} \\
&\quad \times \left(\frac{F_n C_x' \Phi_K}{\sqrt{n}}\right) \left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1} \left(\frac{1}{n} \Phi_K' \Lambda_{U_{0.x},n} \Phi_K\right) \left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1} \frac{\Phi_K' X}{n^{5/2}} \\
&\sim_a \left(\int_0^1 B_X B_x'\right) \left(\int_0^1 B_x B_x'\right)^{-1} \left(\int_0^1 B_x(r)\, \tilde{\varphi}_K(r)'\, dr\right) \left(\int_0^1 \tilde{\varphi}_K(r)\, \tilde{\varphi}_K(r)'\, \widetilde{B_{0.x}}(r)^2\, dr\right) \\
&\quad \times \left(\int_0^1 \tilde{\varphi}_K(r)\, \breve{B}_X(r)'\, dr\right),
\end{aligned}
$$

and second (again ignoring the negative sign)

$$
\begin{aligned}
E_{n2} &= \frac{1}{n^4} X' P_{\Phi_K} \Lambda_{U_{0.x},n} P_{\Phi_K} C_x \left(C_x' P_{\Phi_K} C_x\right)^{-1} C_x' P_{\Phi_K} X \\
&= \left(\frac{1}{n^2} X' P_{\Phi_K} \Lambda_{U_{0.x},n} P_{\Phi_K} C_x F_n\right) \left(F_n C_x' P_{\Phi_K} C_x F_n\right)^{-1} \left(F_n C_x' P_{\Phi_K} X \frac{1}{n^2}\right) \\
&= \frac{X'\Phi_K}{n^{5/2}} \left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1} \left(\frac{1}{n} \Phi_K' \Lambda_{U_{0.x},n} \Phi_K\right) \left(\frac{\Phi_K' C_x F_n}{\sqrt{n}}\right) \left(F_n C_x' P_{\Phi_K} C_x F_n\right)^{-1} \\
&\times \left(\frac{F_n C_x' \Phi_K}{\sqrt{n}}\right) \left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1} \frac{\Phi_K' X}{n^{5/2}} \\
&\sim_a \left(\int_0^1 B_X(r)\, \tilde{\varphi}_K(r)'\, dr\right) \left(\int_0^1 \tilde{\varphi}_K(r)\, \tilde{\varphi}_K(r)'\, \widetilde{B_{0.x}}(r)^2\, dr\right) \left(\int_0^1 \tilde{\varphi}_K(r)\, B_x(r)'\, dr\right) \\
&\times \left(\int_0^1 B_x B_x'\right)^{-1} \left(\int_0^1 B_x B_X'\right).
\end{aligned}
$$

The third term in the expanded form of the central three factors in braces in (116) is

$$
\begin{aligned}
E_{n3} &:= \frac{1}{n^4} X' P_{\Phi_K} C_x \left(C_x' P_{\Phi_K} C_x\right)^{-1} \left(C_x' P_{\Phi_K} \Lambda_{U_{0.x},n} P_{\Phi_K} C_x\right) \left(C_x' P_{\Phi_K} C_x\right)^{-1} C_x' P_{\Phi_K} X \\
&= \left(\frac{1}{n^2} X' P_{\Phi_K} C_x F_n\right) \left(F_n C_x' P_{\Phi_K} C_x F_n\right)^{-1} \left(F_n C_x' P_{\Phi_K} \Lambda_{U_{0.x},n} P_{\Phi_K} C_x F_n\right) \\
&\times \left(F_n C_x' P_{\Phi_K} C_x F_n\right)^{-1} \left(\frac{1}{n^2} F_n C_x' P_{\Phi_K} X\right) \\
&= \frac{X'\Phi_K}{n^{5/2}} \left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1} \left(\frac{\Phi_K' C_x F_n}{\sqrt{n}}\right) \left(F_n C_x' P_{\Phi_K} C_x F_n\right)^{-1} \left(\frac{F_n C_x' \Phi_K}{\sqrt{n}}\right) \left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1} \left(\frac{1}{n} \Phi_K' \Lambda_{U_{0.x},n} \Phi_K\right) \\
&\times \left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1} \left(\frac{\Phi_K' C_x F_n}{\sqrt{n}}\right) \left(F_n C_x' P_{\Phi_K} C_x F_n\right)^{-1} \left(\frac{F_n C_x' \Phi_K}{\sqrt{n}}\right) \left(\frac{\Phi_K'\Phi_K}{n}\right)^{-1} \frac{\Phi_K' X}{n^{5/2}} \\
&\sim_a \left(\int_0^1 B_X B_x'\right) \left(\int_0^1 B_x B_x'\right)^{-1} \left(\int_0^1 B_x(r)\, \tilde{\varphi}_K(r)'\, dr\right) \left(\int_0^1 \tilde{\varphi}_K(r)\, \tilde{\varphi}_K(r)'\, \widetilde{B_{0.x}}(r)^2\, dr\right) \\
&\times \left(\int_0^1 \tilde{\varphi}_K(r)\, B_x(r)'\, dr\right) \left(\int_0^1 B_x B_x'\right)^{-1} \left(\int_0^1 B_x B_X'\right).
\end{aligned}
$$

Combining these three terms, adjusting for sign, and including the leading term we obtain

$$
\begin{aligned}
&\frac{1}{n^4}\left(X'P_{\Phi_K} - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}\right)\Lambda_{U_{0.x},n}\left(P_{\Phi_K}X - P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}X\right)\\
&= E_{n0} - E_{n1} - E_{n2} + E_{n3}\\
&\sim_a \left(\int_0^1 B_X(r)\tilde{\varphi}_K(r)'\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\widetilde{B_{0.x}}(r)^2\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)B_X(r)'\,dr\right)\\
&-\left\{\left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x(r)\tilde{\varphi}_K(r)'\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\widetilde{B_{0.x}}(r)^2\,dr\right)\right.\\
&\left.\times\left(\int_0^1 \tilde{\varphi}_K(r)B_X(r)'\,dr\right)\right\}\\
&-\left\{\left(\int_0^1 B_X(r)\tilde{\varphi}_K(r)'\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\widetilde{B_{0.x}}(r)^2\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)B_x(r)'\,dr\right)\right.\\
&\left.\times\left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x B_X'\right)\right\}\\
&+\left\{\left(\int_0^1 B_X B_x'\right)\left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x(r)\tilde{\varphi}_K(r)'\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\widetilde{B_{0.x}}(r)^2\,dr\right)\right.\\
&\left.\times\left(\int_0^1 \tilde{\varphi}_K(r)B_x(r)'\,dr\right)\left(\int_0^1 B_x B_x'\right)^{-1}\left(\int_0^1 B_x \breve{B}_X'\right)\right\}\\
&= \left(\int_0^1 B_{X.x}(r)\tilde{\varphi}_K(r)'\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\widetilde{B_{0.x}}(r)^2\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)B_{X.x}(r)'\,dr\right).
\end{aligned}
\tag{117}
$$

The full expression for $G_K\left(n\hat{V}_{Kn}\right)G_K'$ is then

$$
\begin{aligned}
G_K\left(n\hat{V}_{Kn}\right)G_K' &= nM\left(\int_{-1}^1 k(r)\,dr\right)G_K\Phi_K'\Lambda_{U_{0.x},n}\Phi_K G_K' + o_p(1)\\
&= \frac{nM}{n^4}\left(\int_{-1}^1 k(r)\,dr\right)\left(\frac{X'R_KX}{n^4}\right)^{-1}\left(E_{n0} - E_{n1} - E_{n2} + E_{n3}\right)\left(\frac{X'R_KX}{n^4}\right)^{-1} + o_p(1)\\
&\sim_a \frac{nM}{n^4}\left(\int_{-1}^1 k(r)\,dr\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1 B_{X.x}(r)\tilde{\varphi}_K(r)'\,dr\right)\\
&\quad\times\left(\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\widetilde{B_{0.x}}(r)^2\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)B_{X.x}(r)'\,dr\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\\
&=: \frac{nM}{n^4}\left(\int_{-1}^1 k(r)\,dr\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}E_{nK}\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1},
\end{aligned}
\tag{118}
$$

where

$$
E_{nK} = \left(\int_0^1 B_{X.x}(r)\tilde{\varphi}_K(r)'\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\widetilde{B_{0.x}}(r)^2\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)B_{X.x}(r)'\,dr\right).
\tag{119}
$$

It follows that the Wald statistic is

$$\begin{aligned}
&\text{Wald}_{TIV} = (H\hat{a}_{TIV} - h)' \left[HG_K \left(n\hat{V}_{Kn}\right) G_K' H'\right]^{-1} (H\hat{a}_{TIV} - h) \\
= \quad & \{n(\hat{a}_{TIV} - a)\}' H' \left[n^2 HG_K \left(n\hat{V}_{Kn}\right) G_K' H'\right]^{-1} H\{n(\hat{a}_{TIV} - a)\} \\
\sim_a \quad & \{n(\hat{a}_{TIV} - a)\}' H' \left[n^2 H \left\{\frac{nM}{n^4}\left(\int_{-1}^1 k(r)\,dr\right)\left(\int_0^1 B_X B_X'\right)^{-1} E_{nK}\left(\int_0^1 B_X B_X'\right)^{-1}\right\} H'\right]^{-1} \\
& \times H\{n(\hat{a}_{TIV} - a)\} \\
= \quad & \left(\frac{n}{M}\right)\{n(\hat{a}_{TIV} - a)\}' H' \left[H\left\{\left(\int_{-1}^1 k(r)\,dr\right)\left(\int_0^1 B_X B_X'\right)^{-1} E_{nK}\left(\int_0^1 B_X B_X'\right)^{-1}\right\} H'\right]^{-1} \\
& \times H\{n(\hat{a}_{TIV} - a)\} \\
= \quad & O_p\left(\frac{n}{M}\right) \text{ for fixed } K. \qquad (120)
\end{aligned}$$

Thus, for fixed K, the HAC Wald statistic is $O_p\left(\frac{n}{M}\right)$ and diverges whenever the lag trunction parameter $M = o(n)$. This outcome is analogous to the result in Phillips (1998) corresponding to the behavior of the coefficient t statistics constructed with HAC standard errors in a spurious regression of an integrated time series on deterministic orthonormal regressors $\{\varphi_k\}_{k=1}^K$. It is, like that result and as subsequent research (Sun, 2004; Phillips et al., 2019) has confirmed, indicative of the important property that when $M = bn$ for some fixed $b \in (0,1]$ we would expect to have test statistic asymptotic behavior of the form $\text{Wald}_{TIV} = O_p(1)$. In the HAR section of the proof given below we show that this is indeed so and establish the limit theory in this case of fixed-b long run variance matrix estimation.

First we complete the development of the limit behavior of the HAC Wald statistic when $K \to \infty$. To do so, working from (120) we need to evaluate the asymptotic behavior of E_{nK}. Observe that

$$\begin{aligned}
&E_{nK} = \left(\int_0^1 B_{X.x}(r)\tilde{\varphi}_K(r)'\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)\tilde{\varphi}_K(r)'\widetilde{B_{0.x}}(r)^2\,dr\right)\left(\int_0^1 \tilde{\varphi}_K(r)B_{X.x}(r)'\,dr\right) \\
&= \int_{r=0}^1\int_{s=0}^1\int_{p=0}^1 B_{X.x}(r)B_{X.x}(p)'\tilde{\varphi}_K(r)'\tilde{\varphi}_K(s)\tilde{\varphi}_K(s)'\tilde{\varphi}_K(p)\widetilde{B_{0.x}}(s)^2\,drdsdp \\
&\sim_a \int_0^1 \widetilde{B_{0.x}}(r)^2 B_{X.x}(r)B_{X.x}(r)'\,dr, \qquad (121)
\end{aligned}$$

To verify (121) we proceed as follows. First, as in the representation (76) in which $B_X(r) = \sum_{m=1}^\infty \varphi_m(r)\nu_m$, we note that $B_{X.x}(r)$ can be written in orthormal expansion form as $B_{X.x}(r) = \sum_{j=1}^\infty \varphi_j(r)\zeta_j$ for suitably chosen m_x-vector variates $\{\zeta_j\}_{j=1}^\infty$. Then, $\int_0^1 B_{X.x}(r)\tilde{\varphi}_K(r)'\,dr = (\zeta_1, ..., \zeta_K) =: \vartheta_K$ and

$$\begin{aligned}
E_{nK} &= \int_0^1 [\vartheta_K\tilde{\varphi}_K(r)]\left[\tilde{\varphi}_K(r)'\vartheta_K'\right]\widetilde{B_{0.x}}(r)^2\,dr \\
&= \int_0^1 \left[\sum_{j=1}^K \varphi_j(r)\zeta_j\right]\left[\sum_{j=1}^K \varphi_j(r)\zeta_j\right]'\widetilde{B_{0.x}}(r)^2\,dr
\end{aligned}$$

$$\rightarrow_{a.s.} \int_0^1 B_{X.x}(r) B_{X.x}(r)' \widetilde{B_{0.x}}(r)^2 \, dr, \text{ as } K \rightarrow \infty, \tag{122}$$

giving (121). Using the fact that $n(\hat{a}_{TIV} - a) = O_p(1)$ with limit distribution given by (106) as $(K, n) \rightarrow \infty$, we deduce that the Wald statistic Wald_{TIV} has the following asymptotic behavior

$$\begin{aligned}
&\text{Wald}_{TIV} = (H\hat{a}_{TIV} - h)' \left[HG_K \left(n\hat{V}_{Kn} \right) G_k' H' \right]^{-1} (H\hat{a}_{TIV} - h) \\
&\sim_a \left(\frac{n}{M} \right) \{ n(\hat{a}_{TIV} - a) \}' H' \times \\
&\left[H \left\{ \left(\int_{-1}^1 k(r) \, dr \right) \left(\int_0^1 B_X B_X' \right)^{-1} \left(\int_0^1 \widetilde{B_{0.x}}(r)^2 B_{X.x}(r) B_{X.x}(r)' \, dr \right) \left(\int_0^1 B_X B_X' \right)^{-1} \right\} H' \right]^{-1} \\
&\times H \{ n(\hat{a}_{TIV} - a) \} = O_p\left(\frac{n}{M} \right),
\end{aligned}$$

so that the Wald statistic Wald_{TIV} diverges at rate $O_p\left(\frac{n}{M}\right)$ even as $K \rightarrow \infty$ when $M = o(n)$ as $n \rightarrow \infty$. The same divergence result therefore holds for this case as for the regression with fixed K.

The HAR Case When $M = bn$ and a fixed-b kernel approach is employed, we find that $\text{Wald}_{TIV} = O_p(1)$. To find the limit distribution of Wald_{TIV} in this case we proceed as follows. Define $k_b\left(\frac{j}{n}\right) = k\left(\frac{j}{bn}\right)$, where $k(\cdot)$ is the lag kernel function as before, and set $M = bn$ for some $b \in 0, 1]$. First, consider the limit behavior of the fixed b kernel long run variance estimator $\hat{V}_{bKn}$, defined as

$$n\hat{V}_{bKn} = \sum_{j=-n+1}^{n-1} k\left(\frac{j}{bn}\right) \sum_{1 \le t, t+j \le n}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right) \tilde{\varphi}_K\left(\frac{t+j}{n}\right)' \hat{U}_{0.xt} \hat{U}_{0.xt+j}.$$

Then, in place of (113) and using (112), we have as $n \rightarrow \infty$

$$\begin{aligned}
\frac{1}{n^2} \hat{V}_{bKn} &= \frac{1}{n^2} \sum_{j=-n+1}^{n-1} k\left(\frac{j}{bn}\right) \sum_{1 \le t, t+j \le n}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right) \tilde{\varphi}_K\left(\frac{t+j}{n}\right)' \frac{\hat{U}_{0.xt}}{\sqrt{n}} \frac{\hat{U}_{0.xt+j}}{\sqrt{n}} \\
&= \frac{1}{n^2} \sum_{1 \le t, s \le n}^{n} k_b\left(\frac{s-t}{n}\right) \tilde{\varphi}_K\left(\frac{t}{n}\right) \tilde{\varphi}_K\left(\frac{s}{n}\right)' \frac{\hat{U}_{0.xt}}{\sqrt{n}} \frac{\hat{U}_{0.xs}}{\sqrt{n}} \\
&\rightsquigarrow \int_0^1 \int_0^1 k_b(r-p) \tilde{\varphi}_K(r) \tilde{\varphi}_K(p)' \widetilde{B_{0.x}}(r) \widetilde{B_{0.x}}(p) \, drdp.
\end{aligned}$$

Hence, $n\hat{V}_{bKn} \sim_a n^3 \int_0^1 \int_0^1 k_b(r-p) \tilde{\varphi}_K(r) \tilde{\varphi}_K(p)' \widetilde{B_{0.x}}(r) \widetilde{B_{0.x}}(p) \, drdp$ and so, in place of (118), the expression for $G_K\left(n\hat{V}_{bKn}\right) G_K'$ is now

$$\begin{aligned}
&G_K\left(n\hat{V}_{bKn}\right) G_K' \\
&\sim_a \frac{n^2}{n^4} \left(\int_0^1 B_{X.x} B_{X.x}' \right)^{-1} \left(\int_0^1 B_{X.x}(r) \tilde{\varphi}_K(r)' \, dr \right) \\
&\times \left(\int_0^1 \int_0^1 k_b(r-p) \tilde{\varphi}_K(r) \tilde{\varphi}_K(p)' \widetilde{B_{0.x}}(r) \widetilde{B_{0.x}}(p) \, drdp \right) \left(\int_0^1 \tilde{\varphi}_K(r) B_{X.x}(r)' \, dr \right) \left(\int_0^1 B_{X.x} B_{X.x}' \right)^{-1}
\end{aligned}$$

$$=: \frac{1}{n^2}\left(\int_0^1 B_X B_X'\right)^{-1} E_{bnK}\left(\int_0^1 B_X B_X'\right)^{-1},$$

and where, with $\mathcal{C}_K := \int_0^1 B_{X.x}(r)\,\tilde{\varphi}_K(r)'\,dr$,

$$E_{bnK} = \mathcal{C}_K \int_0^1\int_0^1 k_b(r-p)\,\tilde{\varphi}_K(r)\,\tilde{\varphi}_K(p)'\,\widetilde{B_{0.x}}(r)\,\widetilde{B_{0.x}}(p)\,drdp\,\mathcal{C}_K'.$$

Hence

$$\begin{aligned}
& n^2 H G_K\left(n\hat{V}_{bKn}\right)G_K'H' \sim_a H\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1} E_{bnK}\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1} H' \\
= \;& H\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\mathcal{C}_K\left(\int_0^1\int_0^1 k_b(r-p)\,\tilde{\varphi}_K(r)\,\tilde{\varphi}_K(p)'\,\widetilde{B_{0.x}}(r)\,\widetilde{B_{0.x}}(p)\,drdp\right) \\
& \times \mathcal{C}_K'\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1} H'.
\end{aligned}$$

This expression simplifies in a similar manner to (121). Indeed, in place of (122) and using the representation $\mathcal{C}_K = \int_0^1 B_{X.x}(r)\,\tilde{\varphi}_K(r)'\,dr = (\zeta_1, ..., \zeta_K) =: \vartheta_K$ as before we now have

$$\begin{aligned}
E_{bnK} &= \int_0^1\int_0^1 k_b(r-p)\left[\vartheta_K\tilde{\varphi}_K(r)\right]\left[\tilde{\varphi}_K(p)'\vartheta_K'\right]\widetilde{B_{0.x}}(r)\,\widetilde{B_{0.x}}(p)\,drdp \\
&\to_{a.s.} \int_0^1\int_0^1 k_b(r-p)\left(B_{X.x}(r)\,B_{X.x}(p)'\right)\widetilde{B_{0.x}}(r)\,\widetilde{B_{0.x}}(p)\,drdp,
\end{aligned}$$

as $K\to\infty$ since $\vartheta_K\tilde{\varphi}_K(r) = \sum_{j=1}^K \varphi_j(r)\,\zeta_j \to_{a.s.} B_{X.x}(r) = \sum_{j=1}^\infty \varphi_j(r)\,\zeta_j$. Thus, as $(K,n)\to\infty$ we have

$$\begin{aligned}
& n^2 H G_K\left(n\hat{V}_{bKn}\right)G_K'H' \\
\sim_a\; & H\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1\int_0^1 k_b(r-p)\left(B_{X.x}(r)\,B_{X.x}(p)'\right)\widetilde{B_{0.x}}(r)\,\widetilde{B_{0.x}}(p)\,drdp\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1} H'.
\end{aligned}$$

We deduce that the modified Wald statistic with fixed b HAR kernel construction is

$$\begin{aligned}
& \text{Wald}_{TIV,b} = (H\hat{a}_{TIV}-h)'\left[HG_K\left(n\hat{V}_{bKn}\right)G_K'H'\right]^{-1}(H\hat{a}_{TIV}-h) \\
= \;& \{n(\hat{a}_{TIV}-a)\}'H'\left[n^2 HG_K\left(n\hat{V}_{bKn}\right)G_K'H'\right]^{-1} H\{n(\hat{a}_{TIV}-a)\} \\
\sim_a\; & \{n(\hat{a}_{TIV}-a)\}'H' \\
& \times\left[H\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1\int_0^1 k_b(r-p)\left(B_{X.x}(r)\,B_{X.x}(p)'\right)\widetilde{B_{0.x}}(r)\,\widetilde{B_{0.x}}(p)\,drdp\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}H'\right]^{-1} \\
& \times H\{n(\hat{a}_{TIV}-a)\} \\
= \;& O_p(1) \;\text{ as }\; (K,n)\to\infty.
\end{aligned}$$

Now

$$n(\hat{a}_{TIV}-a) \rightsquigarrow \left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\int_0^1 \overrightarrow{B_{X.x}}dB_{0.x}$$

$$=_d \ \Omega_{00.x}^{1/2} \times \Omega_{xx}^{-1/2} \left(\int_0^1 W_{X.x} W_{X.x}' \right)^{-1} \int_0^1 \overrightarrow{W_{X.x}} dW_{0.x} =: \Omega_{00.x}^{1/2} \times \Omega_{xx}^{-1/2} \times \eta_W$$

$$=_d \ \mathcal{MN} \left(0, \Omega_{00.x} \times \Omega_{xx}^{-1/2} \left(\int_0^1 W_{X.x} W_{X.x}' \right)^{-1} \int_0^1 \overrightarrow{W_{X.x}} \overrightarrow{W_{X.x}}' \left(\int_0^1 W_{X.x} W_{X.x}' \right)^{-1} \Omega_{xx}^{-1/2} \right),$$

where $\eta_W := \left(\int_0^1 W_{X.x} W_{X.x}' \right)^{-1} \int_0^1 \overrightarrow{W_{X.x}} dW_{0.x}$ and $W_{0.x} = \Omega_{00.x}^{-1/2} B_{0.x}$ is a standard Brownian motion which is independent of $W_x = \Omega_{xx}^{-1/2} B_x$ and, in consequence, all functionals of W_x, including $W_{X.x}(r) = W_X(r) - \int_0^1 W_X W_x \left(\int_0^1 W_x W_x' \right)^{-1} W_x(r)$, $W_X(r) = \int_0^r W_x$, and $\overrightarrow{W_{X.x}}(r) = \int_r^1 W_{X.x}$. The limit distribution of $\text{Wald}_{TIV,b}$ therefore takes the following form

$$\text{Wald}_{TIV,b} \rightsquigarrow \Omega_{00.x} \times \eta_W' \Omega_{xx}^{-1/2} H' \left\{ H \mathcal{E}_B H' \right\}^{-1} H \Omega_{xx}^{-1/2} \eta_W,$$

where

$$\mathcal{E}_B := \left(\int_0^1 B_{X.x} B_{X.x}' \right)^{-1} \left(\int_0^1 \int_0^1 k_b(r-p) \left(B_{X.x}(r) B_{X.x}(p)' \right) \widetilde{B_{0.x}}(r) \widetilde{B_{0.x}}(p) \, dr dp \right) \left(\int_0^1 B_{X.x} B_{X.x}' \right)^{-1}$$

$$= \Omega_{00.x} \times \Omega_{xx}^{-1/2} \left(\int_0^1 W_{X.x} W_{X.x}' \right)^{-1} \left(\int_0^1 \int_0^1 k_b(r-p) \left(W_{X.x}(r) W_{X.x}(p)' \right) \widetilde{W_{0.x}}(r) \widetilde{W_{0.x}}(p) \, dr dp \right)$$

$$\times \left(\int_0^1 W_{X.x} W_{X.x}' \right)^{-1} \Omega_{xx}^{-1/2}$$

$$=: \ \Omega_{00.x} \times \Omega_{xx}^{-1/2} \times \mathcal{E}_W \times \Omega_{xx}^{-1/2},$$

where $\widetilde{W_{0.x}}(r) = \Omega_{00.x}^{-1/2} \widetilde{B_{0.x}}(r)$ and $\widetilde{B_{0.x}}$ is defined in (112). Next observe that

$$\begin{aligned} & \Omega_{00.x} \times \eta_W' \Omega_{xx}^{-1/2} H' \left\{ H \mathcal{E}_B H' \right\}^{-1} H \Omega_{xx}^{-1/2} \eta_W \\ = \ & \eta_W' \mathcal{E}_W^{-1/2} \left(\mathcal{E}_W^{1/2} \Omega_{xx}^{-1/2} H' \right) \left\{ H \Omega_{xx}^{-1/2} \mathcal{E}_W \Omega_{xx}^{-1/2} H' \right\}^{-1} \left(H \Omega_{xx}^{-1/2} \mathcal{E}_W^{1/2} \right) \mathcal{E}_W^{-1/2} \eta_W \\ =: \ & \eta_{\mathcal{E}_W}' L \left\{ L L' \right\}^{-1} L' \eta_{\mathcal{E}_W}, \end{aligned}$$

where $\eta_{\mathcal{E}_W} = \mathcal{E}_W^{-1/2} \eta_W$ and $L := \mathcal{E}_W^{1/2} \Omega_{xx}^{-1/2} H$, which leads to the following limit representation of the Wald statistic

$$\text{Wald}_{TIV,b} \rightsquigarrow \eta_{\mathcal{E}_W}' L \left\{ L L' \right\}^{-1} L' \eta_{\mathcal{E}_W} = \eta_{\mathcal{E}_W}' P_L \eta_{\mathcal{E}_W},$$

with the random projection matrix $P_L = L \left\{ L L' \right\}^{-1} L'$. Since the distribution of the random matrix

$$\mathcal{E}_W = \left(\int_0^1 W_{X.x}(r) W_{X.x}(r)' \right)^{-1} \left(\int_0^1 \int_0^1 k_b(r-p) W_{X.x}(r) W_{X.x}(p)' \widetilde{W_{0.x}}(r) \widetilde{W_{0.x}}(p) \, dr dp \right)$$

$$\times \left(\int_0^1 W_{X.x}(r) W_{X.x}(r)' \right)^{-1}$$

depends only on the vector standard Brownian motions (W_X, W_x), the scalar Brownian motion $W_{0.x}$, and the scalar kernel function $k_b(\cdot)$ this distribution is invariant to rotations, just as are the vector Brownian motion W_x and the random vector $\eta_{\mathcal{E}_W} = \mathcal{E}_W^{-1/2}\eta_W$. It follows that the limit distribution of the Wald statistic

$$\text{Wald}_{TIV,b} \rightsquigarrow \eta_{\mathcal{E}_W}' L \{LL'\}^{-1} L' \eta_{\mathcal{E}_W}$$

is pivotal and depends only on the rank of $L := \mathcal{E}_W^{1/2}\Omega_{xx}^{-1/2}H$ or equivalently the rank of the $q \times m_x$ restriction matrix H, i.e., on the number of restrictions q.

Part (viii)

The HAC Case In this case $e_t^+ = e_t$ and the fitted equation is by its partial sum construction a cointegrating equation with $I(2)$ regressor X_t, $I(1)$ regressor x_t, augmented with the additional $I(0)$ regressor $\Delta x_t = u_{xt}$, and the $I(0)$ error e_t. The TIV regression produces consistent estimates of the cointegrating vector a, as shown in Theorem 3, and also the long run regression coefficient $f = \Omega_{xx}^{-1}\Omega_{x0}$, just as in Phillips (2014). In particular, note that

$$\hat{f}_{TIV} - f = \left(\frac{1}{n^2}x'R_{fK}x\right)^{-1}(\frac{1}{n^2}x'R_{fK}e) \rightarrow_p 0 \tag{123}$$

where

$$\begin{aligned} R_{fK} &= P_{\Phi_K} - P_{\Phi_K}C_f\left(C_f'P_{\Phi_K}C_f\right)^{-1}C_f'P_{\Phi_K}, \\ C_f' &= \begin{bmatrix} X_1 & \cdots & X_n \\ u_{x1} & \cdots & u_{xn} \end{bmatrix} = \begin{bmatrix} X' \\ u_x' \end{bmatrix}. \end{aligned}$$

After calculations using techniques similar to those employed earlier[11], we obtain

$$\frac{1}{n^2}x'R_{fK}x \rightsquigarrow \int_0^1 B_xB_x' - \int_0^1 B_xB_X'\left(\int_0^1 B_XB_X'\right)^{-1}\int_0^1 B_XB_x' = \int_0^1 B_{x.X}B_{x.X}', \tag{124}$$

where $B_{x.X}(r) = B_x(r) - \int_0^1 B_xB_X'\left(\int_0^1 B_XB_X'\right)^{-1}B_X(r)$, and

$$\frac{1}{n^2}x'R_{fK}e = \frac{1}{n}\left(\frac{1}{n}x'P_{\Phi_K}e - \frac{1}{n}x'P_{\Phi_K}C_f\left(C_f'P_{\Phi_K}C_f\right)^{-1}C_f'P_{\Phi_K}e\right) = O_p\left(\frac{1}{n}\right),$$

thereby establishing (123). More specifically, and again after considerable calculation, we obtain

$$\begin{aligned} &\frac{1}{n}x'R_{fK}e = \frac{1}{n}x'P_{\Phi_K}e - \frac{1}{n}x'P_{\Phi_K}C_f\left(C_f'P_{\Phi_K}C_f\right)^{-1}C_f'P_{\Phi_K}e \\ &= \frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'e}{\sqrt{n}} - \frac{x'\Phi_K}{n^{3/2}}\frac{\Phi_K'C_f}{\sqrt{n}}L_n\left(L_nC_f'P_{\Phi_K}C_fL_n\right)^{-1}L_n\frac{C_f'\Phi_K}{\sqrt{n}}\frac{\Phi_K'e}{\sqrt{n}} + o_p(1) \\ &\rightsquigarrow \int_0^1 B_xdB_{e.x} - \int_0^1 B_xB_X'\left(\int_0^1 B_XB_X'\right)^{-1}\int_0^1 B_XdB_{e.x} = \int_0^1 B_{x.X}dB_{e.x}. \end{aligned} \tag{125}$$

Combining (124) and (125) gives

$$n(\hat{f}_{TIV} - f) = \left(\frac{1}{n^2}x'R_{fK}x\right)^{-1}\left(\frac{1}{n}x'R_{fK}e\right) \rightsquigarrow \left(\int_0^1 B_{x.X}B_{x.X}'\right)^{-1}\int_0^1 B_{x.X}dB_{e.x}. \tag{126}$$

[11] A full development will be given elsewhere.

Further, since $g=0$ by construction, we have

$$\hat{g}_{TIV} = \left(\frac{1}{K} u_x' R_{gK} u_x\right)^{-1} \left(\frac{1}{K} u_x' R_{gK} e\right) \to_p \Omega_{xx}^{-1} \omega_{xe} \tag{127}$$

where

$$\begin{aligned} R_{gK} &= P_{\Phi_K} - P_{\Phi_K} C_g \left(C_g' P_{\Phi_K} C_g\right)^{-1} C_g' P_{\Phi_K}, \\ C_g' &= \begin{bmatrix} X_1 & \cdots & X_n \\ x_1 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} X' \\ x' \end{bmatrix}. \end{aligned}$$

In particular, note that using the normalization matrix $D_n = diag\left[n^{-2} I_{m_x}, n^{-1} I_{m_x}\right]$ we have

$$\begin{aligned} &\frac{1}{K} u_x' R_{gK} u_x = \frac{1}{K} u_x' P_{\Phi_K} u_x - \frac{1}{K} u_x' P_{\Phi_K} C_g D_n \left(D_n C_g' P_{\Phi_K} C_g D_n\right)^{-1} D_n C_g' P_{\Phi_K} u_x \\ = \; &\frac{1}{K} u_x' P_{\Phi_K} u_x - \left(\frac{1}{\sqrt{K}} \frac{u_x' \Phi_K}{\sqrt{n}}\right) \begin{bmatrix} \frac{1}{n} \frac{\Phi_K' X}{n^{3/2}} \\ \frac{1}{n} \frac{\Phi_K' x}{\sqrt{n}} \end{bmatrix} \left(D_n C_g' P_{\Phi_K} C_g D_n\right)^{-1} \begin{bmatrix} \frac{1}{n} \frac{X' \Phi_K}{n^{3/2}} \\ \frac{1}{n} \frac{x' \Phi_K}{\sqrt{n}} \end{bmatrix} \left(\frac{\Phi_K' u_x}{\sqrt{n}} \frac{1}{\sqrt{K}}\right) \\ = \; &\frac{1}{K} u_x' P_{\Phi_K} u_x + O_p\left(\frac{1}{K}\right) \to_p \Omega_{xx}, \end{aligned}$$

and, similarly, $\frac{1}{K} u_x' R_{gK} e \to_p \omega_{xe}$, giving (127). In consequence $\hat{g}_{TIV}$ provides a consistent estimate of the long run regression coefficient $\Omega_{xx}^{-1} \omega_{xe}$.

It follows that the regression residuals

$$\begin{aligned} \hat{e}_t^+ &= e_t^+ - \left(\hat{a}_{TIV} - a\right)' X_t - \left(\hat{f}_{TIV} - f\right)' x_t - \hat{g}_{TIV}' u_{xt} \\ &= e_t - n^2 \left(\hat{a}_{TIV} - a\right)' \frac{X_t}{n^{3/2}} \frac{1}{\sqrt{n}} - n \left(\hat{f}_{TIV} - f\right)' \frac{x_t}{\sqrt{n}} \frac{1}{\sqrt{n}} - \omega_{xe}' \Omega_{xx}^{-1} u_{xt} + O_p\left(\frac{1}{K}\right) \\ &= e_t - \omega_{xe}' \Omega_{xx}^{-1} u_{xt} + O_p\left(\frac{1}{\sqrt{n}} + \frac{1}{K}\right) =: e_{0.x,t} + O_p\left(\frac{1}{\sqrt{n}} + \frac{1}{K}\right), \end{aligned} \tag{128}$$

consistently estimate the $e_{0.x,t} = e_t - \omega_{xe}' \Omega_{xx}^{-1} u_{xt}$. Thus the effect of the inclusion of the regressor $\Delta x_t = u_{xt}$ in the regression is to ensure that the fitted TIV residuals estimate the equation errors adjusted for the conditional long run mean, viz., $e_{0.x,t} = e_t - \omega_{xe}' \Omega_{xx}^{-1} u_{xt}$. Correspondingly, the long run variance estimator in (104) is

$$\begin{aligned} \hat{\omega}_{e^+}^2 &= \sum_{j=-M}^{M} k\left(\frac{j}{M}\right) \frac{1}{n} \sum_{1 \le t, t+j \le n}^{n} \hat{e}_t^+ \hat{e}_{t+j}^+ \\ &= \sum_{j=-M}^{M} k\left(\frac{j}{M}\right) \left(\frac{1}{n} \sum_{1 \le t, t+j \le n}^{n} e_{0.x,t} e_{0.x,t+j}\right) + o_p(1) \\ &\to_p \; \omega_{ee.x} = \omega_{ee} - \omega_{ex} \Omega_{xx}^{-1} \omega_{xe} = \mathbb{V}^{LR}\left(e_{0.x,t}\right). \end{aligned} \tag{129}$$

In a similar way, the kernel estimate $\hat{V}_{Kn}$ in (102) of the transformed residuals has the form

$$\hat{V}_{Kn} = \sum_{j=-M}^{M} k\left(\frac{j}{M}\right) \frac{1}{n} \sum_{1 \le t, t+j \le n}^{n} \tilde{\varphi}_K\left(\frac{t}{n}\right) \tilde{\varphi}_K\left(\frac{t+j}{n}\right)' \hat{e}_t^+ \hat{e}_{t+j}^+ \tag{130}$$

$$= \sum_{j=-M}^{M} k\left(\frac{j}{M}\right)\frac{1}{n}\sum_{1\le t,t+j\le n}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t+j}{n}\right)' e_{0.x,t}e_{0.x,t+j} + o_p(1)$$
$$= \sum_{j=-M}^{M} k\left(\frac{j}{M}\right)\frac{1}{n}\sum_{1\le t,t+j\le n}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t}{n}\right)' \mathbb{E}\left(e_{0.x,t}e_{0.x,t+j}\right) + o_p(1)$$
$$= \left(\frac{1}{n}\sum_{t=1}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t}{n}\right)'\right)\sum_{j=-M}^{M} k\left(\frac{j}{M}\right)\mathbb{E}\left(e_{0.x,t}e_{0.x,t+j}\right) + o_p(1)$$
$$= \left(\frac{1}{n}\Phi_K'\Phi_K\right)\omega_{e.x}^2 + o_p(1) = \omega_{ee.x}I_K + o_p(1), \tag{131}$$

since $\frac{1}{n}\Phi_K'\Phi_K = I_K + O\left(\frac{1}{n}\right)$. Further, observe that

$$\begin{aligned} G_K &= (X'R_KX)^{-1}\left\{X'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1} - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\right\} \\ &= (X'R_KX)^{-1}\left\{X'\Phi_K - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'\Phi_K\right\}\left(\Phi_K'\Phi_K\right)^{-1} \end{aligned}$$

so that

$$\begin{aligned} & G_K\left(n\hat{V}_{Kn}\right)G_K' = \omega_{ee.x}G_K\left(\Phi_K'\Phi_K\right)G_K' + o_p(1) \\ = \;& \omega_{ee.x}(X'R_KX)^{-1}\left\{X'\Phi_K - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'\Phi_K\right\}\left(\Phi_K'\Phi_K\right)^{-1} \\ & \times\left\{\Phi_K'X - \Phi_K'C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}X\right\}(X'R_KX)^{-1} + o_p(1) \\ = \;& \omega_{ee.x}(X'R_KX)^{-1}\left\{X'P_{\Phi_K}X - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}X\right. \\ & \left.-X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}X + X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'P_{\Phi_K}X\right\} \\ & \times(X'R_KX)^{-1} + o_p(1) \\ = \;& \omega_{ee.x}(X'R_KX)^{-1} + o_p(1). \end{aligned}$$

using the fact that $nP_{\Phi_K} = n\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K' = \Phi_K\left(\frac{1}{n}\Phi_K'\Phi_K\right)^{-1}\Phi_K' = \Phi_K\Phi_K'\left\{1 + O\left(\frac{1}{n}\right)\right\}$. Then, the Wald statistic is

$$\begin{aligned} & \text{Wald}_{TIV} = \left(H\hat{a}_{TIV} - h\right)'\left[HG_K\left(n\hat{V}_{Kn}\right)G_K'H'\right]^{-1}\left(H\hat{a}_{TIV} - h\right) \\ = \;& \frac{1}{\omega_{ee.x}}\left[n^2\left(\hat{a}_{TIV} - a\right)\right]'H'\left[H\left(\frac{X'R_KX}{n^4}\right)^{-1}H'\right]^{-1}H\left[n^2\left(\hat{a}_{TIV} - a\right)\right] + o_p(1) \\ \sim_a \;& \left[n^2\left(\frac{\hat{a}_{TIV} - a}{\omega_{ee.x}^{1/2}}\right)\right]'H'\left[H\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}H'\right]^{-1}H\left[n^2\left(\frac{\hat{a}_{TIV} - a}{\omega_{ee.x}^{1/2}}\right)\right] \\ \equiv \;& \chi_q^2 \end{aligned}$$

since

$$n^2H\left(\hat{a}_{TIV} - a\right) \rightsquigarrow \mathcal{MN}\left(0, \omega_{ee.x}H\left(\int_0^1 B_{X.x}B_{X,x}'\right)^{-1}H'\right), \tag{132}$$

thereby giving the required limit theory for the Wald_{TIV} statistic.

The HAR Case In the HAR variance matrix case in view of the fixed-b kernel, we use the following full representation in place of (128)

$$\begin{aligned}\hat{e}_t^+ &= e_{0.x,t} - n^2\left(\hat{a}_{TIV}-a\right)'\frac{X_t}{n^{3/2}}\frac{1}{\sqrt{n}} - n\left(\hat{f}_{TIV}-f\right)'\frac{x_t}{\sqrt{n}}\frac{1}{\sqrt{n}} + O_p\left(\frac{1}{K}\right)\\ &= e_{0.x,t} - \frac{1}{\sqrt{n}}\mathcal{D}_n' w_{nt} + O_p\left(\frac{1}{K}\right), \qquad (133)\end{aligned}$$

where $\mathcal{D}_n' := \left(n^2\left(\hat{a}_{TIV}-a\right)', n(\hat{f}_{TIV}-f)'\right)$ and $w_{nt} = (X_t/n^{3/2}, x_t/n^{1/2})'$. In view of (126) and (132) we have

$$\begin{aligned}\mathcal{D}_n' w_{n\lfloor nr\rfloor} &= n^2\left(\hat{a}_{TIV}-a\right)'\frac{X_{\lfloor nr\rfloor}}{n^{3/2}} + n\left(\hat{f}_{TIV}-f\right)'\frac{x_{\lfloor nr\rfloor}}{\sqrt{n}}\\ &\rightsquigarrow \int_0^1 dB_{e.x}B_{X.x}'\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}B_X(r) + \int_0^1 dB_{e.x}B_{x.X}\left(\int_0^1 B_{x.X}B_{x.X}'\right)^{-1}B_x(r)\\ &= \ell_X' B_X(r) + \ell_x' B_x(r) =: \ell_+' B_+(r) \qquad (134)\end{aligned}$$

where $\ell_X' = \int_0^1 dB_{e.x}B_{X.x}'\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}$, $\ell_x' = \int_0^1 dB_{e.x}B_{x.X}\left(\int_0^1 B_{x.X}B_{x.X}'\right)^{-1}$, $\ell_+ = (\ell_X', \ell_x')'$, and $B_+(r) = (B_X(r)', B_x(r)')'$.

The following fixed-b kernel estimate replaces (130) and when $K^2/n \to \infty$ we have

$$\begin{aligned}\hat{V}_{bKn} &= \sum_{j=-n+1}^{n-1} k_b\left(\frac{j}{n}\right)\frac{1}{n}\sum_{1\le t,t+j\le n}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t+j}{n}\right)'\hat{e}_t^+\hat{e}_{t+j}^+\\ &= \sum_{j=-n+1}^{n-1} k_b\left(\frac{j}{n}\right)\sum_{1\le t,t+j\le n}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t+j}{n}\right)'\frac{e_{0.x,t}}{\sqrt{n}}\frac{e_{0.x,t+j}}{\sqrt{n}}\\ &- \frac{2}{n}\sum_{j=-n+1}^{n-1} k_b\left(\frac{j}{n}\right)\sum_{1\le t,t+j\le n}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t+j}{n}\right)'\frac{e_{0.x,t}}{\sqrt{n}}\mathcal{D}_n' w_{nt+j}\\ &+ \frac{1}{n^2}\sum_{j=-n+1}^{n-1} k_b\left(\frac{j}{n}\right)\sum_{1\le t,t+j\le n}^{n}\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{t+j}{n}\right)'\left(\mathcal{D}_n' w_{nt}\mathcal{D}_n' w_{nt+j}\right) + o_p(1)\\ &= \sum_{1\le t,s\le n}^{n} k_b\left(\frac{s-t}{n}\right)\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{s}{n}\right)'\frac{e_{0.x,t}}{\sqrt{n}}\frac{e_{0.x,s}}{\sqrt{n}}\\ &- \frac{2}{n}\sum_{1\le t,s\le n}^{n} k_b\left(\frac{s-t}{n}\right)\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{s}{n}\right)'\frac{e_{0.x,t}}{\sqrt{n}}\mathcal{D}_n' w_{ns}\\ &+ \frac{1}{n^2}\sum_{1\le t,s\le n}^{n} k_b\left(\frac{s-t}{n}\right)\tilde{\varphi}_K\left(\frac{t}{n}\right)\tilde{\varphi}_K\left(\frac{s}{n}\right)'\mathcal{D}_n' w_{nt}\mathcal{D}_n' w_{ns} + o_p(1)\\ &\sim_a \int_0^1\int_0^1 k_b(p-r)\,\tilde{\varphi}_K(p)\,\tilde{\varphi}_K(r)'\,dB_{e.x}(p)\,dB_{e.x}(r)\\ &- 2\int_0^1\int_0^1 k_b(p-r)\,\tilde{\varphi}_K(p)\,\tilde{\varphi}_K(r)'\,dB_{e.x}(p)\ell_+' B_+(r)dr\\ &+ \int_0^1\int_0^1 k_b(p-r)\,\tilde{\varphi}_K(p)\,\tilde{\varphi}_K(r)'\,\ell_+' B_+(p)\ell_+' B_+(r)dpdr\end{aligned}$$

$$= \int_0^1 \int_0^1 k_b\left(p-r\right)\tilde{\varphi}_K\left(p\right)\tilde{\varphi}_K d\mathcal{Q}_B(p)d\mathcal{Q}_B(r) \tag{135}$$

as $n\to\infty$, where the stochastic process $\mathcal{Q}_B(r)$ is defined by the stochastic differential equation $d\mathcal{Q}_B(r) = dB_{e.x}(r) - \ell_+' B_+(r)dr$. From the definition of the components of $\mathcal{Q}_B(r)$ we have the equivalent representation

$$d\mathcal{Q}_B(r) =_d \omega_{ee.x}^{1/2} d\mathcal{Q}_W(r) := \omega_{ee.x}^{1/2}\left[dW_{e.x}(r) - \ell_{W,+}' W_+(r)dr\right], \tag{136}$$

in terms of functionals of standard Brownian motion processes. This representation is obtained by noting that: (i) $\ell_+ = (\ell_X', \ell_x')' =_d \omega_{ee.x}\ell_{W,+}'\Omega_{xx}^{-1/2} = \omega_{ee.x}(\ell_{W,X}', \ell_{W,x}')'\Omega_{xx}^{-1/2}$, where $\ell_{W,X}' = \int dW_{e.x}W_{X.x}'\left(\int W_{X.x}W_{X.x}'\right)^{-1}$, $\ell_{W,x}' = \int dW_{e.x}W_{x.X}'\left(\int W_{x.X}W_{x.X}'\right)^{-1}$; and (ii) $B_+(r) = (B_X(r)', B_x(r)')' = \Omega_{xx}^{1/2}W_+(r) = \Omega_{xx}^{1/2}\left(W_X(r)', W_x(r)'\right)'$, so that we can define $d\mathcal{Q}_W(r) = dW_{e.x}(r) - \ell_{W,+}'W_+(r)dr$, which leads directly to the equivalent representation shown in (136).

In view of (135) we have the following asymptotic behavior of the scaled fixed-b kernel estimate $\hat{V}_{bKn}$

$$n\hat{V}_{bKn} \sim_a n\int_0^1 \int_0^1 k_b\left(r-p\right)\tilde{\varphi}_K\left(r\right)\tilde{\varphi}_K\left(p\right)' d\mathcal{Q}_B(p)d\mathcal{Q}_B(r).$$

Now recall that

$$\begin{aligned}
G_K &= \left(X'R_KX\right)^{-1}\left\{X'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1} - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'\Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\right\} \\
&= \left(X'R_KX\right)^{-1}\left\{X'\Phi_K - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'\Phi_K\right\}\left(\Phi_K'\Phi_K\right)^{-1} \\
&\sim_a \frac{1}{n}\left(X'R_KX\right)^{-1}\left\{X'\Phi_K - X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'\Phi_K\right\} \\
&\sim_a \frac{1}{n}\left(\frac{X'R_KX}{n^4}\right)^{-1}\left\{\frac{X'\Phi_K}{n^{5/2}} - \frac{1}{n^{5/2}}X'P_{\Phi_K}C_x\left(C_x'P_{\Phi_K}C_x\right)^{-1}C_x'\Phi_K\right\}\times\frac{n^{5/2}}{n^4} \\
&\sim_a \frac{1}{n^{5/2}}\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1 B_{X.x}\left(r\right)\tilde{\varphi}_K\left(r\right)' dr\right),
\end{aligned}$$

so that

$$\begin{aligned}
&G_K\left(n\hat{V}_{bKn}\right)G_k' \sim_a G_K\left(n\int_0^1\int_0^1 k_b\left(r-p\right)\tilde{\varphi}_K\left(r\right)\tilde{\varphi}_K\left(p\right)' d\mathcal{Q}_B(p)d\mathcal{Q}_B(r)\right)G_k' \\
&\sim_a \frac{n}{n^5}\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1 B_{X.x}\left(r\right)\tilde{\varphi}_K\left(r\right)' dr\right) \\
&\times\left(\int_0^1\int_0^1 k_b\left(r-p\right)\tilde{\varphi}_K\left(r\right)\tilde{\varphi}_K\left(p\right)' d\mathcal{Q}_B(p)d\mathcal{Q}_B(r)\right) \\
&\times\left(\int_0^1 \tilde{\varphi}_K\left(r\right)' B_{X.x}\left(r\right)' dr\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1} \\
&\sim_a \frac{1}{n^4}\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1\int_0^1 k_b\left(r-p\right)B_{X.x}\left(p\right)B_{X.x}\left(r\right)' d\mathcal{Q}_B(p)d\mathcal{Q}_B(r)\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}.
\end{aligned}$$

The final line above follows because

$$\int_0^1 B_{X.x}\left(r\right)\tilde{\varphi}_K\left(r\right)' dr\left(\int_0^1\int_0^1 k_b\left(p-r\right)\tilde{\varphi}_K\left(p\right)\tilde{\varphi}_K\left(r\right)' d\mathcal{Q}_B(p)d\mathcal{Q}_B(r)\right)\int_0^1 \tilde{\varphi}_K\left(r\right)B_{X.x}\left(r\right)' dr$$

$$\sim_a \int_0^1 \int_0^1 k_b\left(p-r\right) B_{X.x}\left(p\right) B_{X.x}\left(r\right)' d\mathcal{Q}_B(p) d\mathcal{Q}_B(r),$$

where we use precisely the same argument as earlier in (122) to show that

$$\left(\int_0^1 B_{X.x}\left(r\right)\tilde{\varphi}_K\left(r\right)' dr\right)\tilde{\varphi}_K\left(p\right) = \left(\zeta_1, ..., \zeta_K\right)\tilde{\varphi}_K\left(p\right) = \sum_{j=1}^{K}\varphi_j\left(r\right)\zeta_j \to_{a.s.} B_{X.x}\left(r\right),$$

as $K \to \infty$. The fixed-b HAR Wald statistic $\text{Wald}_{TIV,b}$ then has the following asymptotic form

$$\text{Wald}_{TIV,b} = \left(H\hat{a}_{TIV} - h\right)' \left[HG_K\left(n\hat{V}_{bKn}\right)G_K' H'\right]^{-1}\left(H\hat{a}_{TIV} - h\right)$$

$$= \left[n^2\left(\hat{a}_{TIV} - a\right)\right]' H' \left[n^4 HG_K\left(n\hat{V}_{bKn}\right)G_K' H'\right]^{-1} H\left[n^2\left(\hat{a}_{TIV} - a\right)\right]$$

$$\sim_a \left[n^2\left(\hat{a}_{TIV} - a\right)\right]' H'$$

$$\times\left[H\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1\int_0^1 k_b\left(r-p\right)B_{X.x}\left(p\right)B_{X.x}\left(r\right)' d\mathcal{Q}_B(p)d\mathcal{Q}_B(r)\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1} H'\right]^{-1}$$

$$\times H\left[n^2\left(\hat{a}_{TIV} - a\right)\right].$$

From Theorem 3, we deduce that

$$n^2 H\left(\hat{a}_{TIV} - a\right) \rightsquigarrow H\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1 B_{X.x}dB_{e.x}\right)$$

$$=_d \ \omega_{ee.x}^{1/2}\, H\,\Omega_{xx}^{-1/2}\left(\int_0^1 W_{X.x}W_{X.x}'\right)^{-1}\left(\int_0^1 W_{X.x}dW_{e.x}\right),$$

with functionals $W_{X.x}$ and $W_{e.x}$ of standard Brownian motion, corresponding to $B_{X.x}$ and $B_{e.x}$. Next, as shown above in (136) we have $d\mathcal{Q}_B(r) =_d \omega_{ee.x}^{1/2} d\mathcal{Q}_W(r)$. Then

$$H\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1}\left(\int_0^1\int_0^1 k_b\left(r-p\right)B_{X.x}\left(p\right)B_{X.x}\left(r\right)' dQ(p)dQ(r)\right)\left(\int_0^1 B_{X.x}B_{X.x}'\right)^{-1} H'$$

$$=_d \omega_{ee.x} H\Omega_{xx}^{-1/2}\left(\int_0^1 W_{X.x}W_{X.x}'\right)^{-1}\left(\int_0^1\int_0^1 k_b\left(r-p\right)W_{X.x}\left(p\right)W_{X.x}\left(r\right)' d\mathcal{Q}_W(p)d\mathcal{Q}_W(r)\right)$$

$$\times\left(\int_0^1 W_{X.x}W_{X.x}'\right)^{-1}\Omega_{xx}^{-1/2}H'.$$

It follows that the limit distribution of $\text{Wald}_{TIV,b}$ is given by

$$\text{Wald}_{TIV,b} \rightsquigarrow \left(\int_0^1 dW_{e.x}W_{X.x}'\right)\left(\int_0^1 W_{X.x}W_{X.x}'\right)^{-1}\Omega_{xx}^{-1/2}H' \times$$

$$\left[H\Omega_{xx}^{-1/2}\left(\int_0^1 W_{X.x}W_{X.x}'\right)^{-1}\left(\int_0^1\int_0^1 k_b\left(r-p\right)W_{X.x}\left(p\right)W_{X.x}\left(r\right)' d\mathcal{Q}_W(p)d\mathcal{Q}_W(r)\right)\left(\int_0^1 W_{X.x}W_{X.x}'\right)^{-1}\Omega_{xx}^{-1/2}H'\right]^{-1}$$

$$\times H\Omega_{xx}^{-1/2}\left(\int_0^1 W_{X.x}\widetilde{W}_X'\right)^{-1}\left(\int_0^1 \widetilde{W}_X dW_{e.x}\right)$$

$$=: \eta_{e.x}' L_H' \left[L_H \mathcal{F}_W L_H'\right]^{-1} L_H \eta_{e.x}$$

where $\eta_{e.x} = \left(\int_0^1 W_{X.x} W_{X.x}'\right)^{-1} \left(\int_0^1 W_{X.x} dW_{e.x}\right)$, $L_H = H\Omega_{xx}^{-1/2}$, and

$$\mathcal{F}_W = \left(\int_0^1 W_{X.x}\widetilde{W}_X'\right)^{-1} \left(\int_0^1 \int_0^1 k_b\left(r-p\right) W_{X.x}\left(p\right) W_{X.x}\left(r\right)' d\mathcal{Q}_W(p) d\mathcal{Q}_W(r)\right) \left(\int_0^1 W_{X.x} W_{X.x}'\right)^{-1}.$$

Note that H is $q \times m_x$ of rank $q \leq m_x$. Define $L = \left(L_H L_H'\right)^{-1/2} L_H$, so that $LL' = I_q$, and

$$\eta_{e.x}' L_H' \left[L_H \mathcal{F}_W L_H'\right]^{-1} L_H \eta_{Xe} = \eta_{e.x}' L' \left[L\mathcal{F}_W L'\right]^{-1} L\eta_{e.x}.$$

Construct the orthogonal $m_x \times m_x$ matrix $\mathcal{L} = \begin{bmatrix} L \\ L_\perp \end{bmatrix}$ and note that $\mathcal{L}\eta_{e.x} =_d \eta_{e.x}$. We can write

$$L\eta_{e.x} = \left[I_q, 0\right] \mathcal{L}\eta_{e.x} \quad \text{and} \quad L\mathcal{F}_W L' = \left[I_q, 0\right] \mathcal{L}\mathcal{F}_W \mathcal{L}' \begin{bmatrix} I_q \\ 0 \end{bmatrix},$$

from which it follows that

$$\begin{aligned} \eta_{e.x}' L' \left[L\mathcal{F}_W L'\right]^{-1} L\eta_{e.x} &= \eta_{e.x}' \mathcal{L}' \begin{bmatrix} I_q \\ 0 \end{bmatrix} \left\{\left[I_q, 0\right] \mathcal{L}\mathcal{F}_W \mathcal{L}' \begin{bmatrix} I_q \\ 0 \end{bmatrix}\right\}^{-1} \left[I_q, 0\right] \mathcal{L}\eta_{e.x} \\ &=_d \eta_{e.x}' \begin{bmatrix} I_q \\ 0 \end{bmatrix} \left\{\left[I_q, 0\right] \mathcal{F}_W \begin{bmatrix} I_q \\ 0 \end{bmatrix}\right\}^{-1} \left[I_q, 0\right] \eta_{e.x} \\ &= \eta_{e.x}' \mathcal{J}_q' \left\{\mathcal{J}_q \mathcal{F}_W \mathcal{J}_q'\right\}^{-1} \mathcal{J}_q \eta_{e.x}, \quad \text{with } \mathcal{J}_q = \left[I_q, 0\right]. \end{aligned}$$

Since $\mathcal{L}\eta_{e.x} =_d \eta_{e.x}$ and $\mathcal{L}\mathcal{F}_W \mathcal{L}' =_d \mathcal{F}_W$, the limit distribution of $\eta_{e.x}' L' \left[L\mathcal{F}_W L'\right]^{-1} L\eta_{e.x}$ is seen to be invariant to L and dependent only on the dimension q of the restrictions, as embodied in the matrix $\mathcal{J}_q = [I_q, 0]$. The limit distribution $\text{Wald}_{TIV,b} \rightsquigarrow \eta_{e.x}' \mathcal{J}_q' \left\{\mathcal{J}_q \mathcal{F}_W \mathcal{J}_q'\right\}^{-1} \mathcal{J}_q \eta_{e.x}$ is therefore pivotal and dependent only on the dimension parameter$q$, the constituent standard Brownian motions involved in $(\eta_{e.x}, \mathcal{F}_W)$ and the fixed-b kernel function $k_b(\cdot)$.

■

9.3 Glossary of Notation

We use the following notation for data matrices, various functionals of the Brownian motions $(B_X, B_x, B_e, B_{0.x})$, and associated stochastic processes including their standard Brownian motion analogues $(W_X, W_x, W_e, W_{0.x})$. The functionals are defined in the paper and repeated here for convenience. In the following formulae $\int$ represents $\int_0^1$ when the limits are not provided.

$$P_x = x\left(x'x\right)^{-1}x', \quad Q_x = I - x\left(x'x\right)^{-1}x'$$

$$X = [X_1, \cdots, X_n],\ x = [x_1, \cdots, x_n],\ Y = [Y_1, \cdots, Y_n],\ u_x = [u_{x1}, \cdots, u_{xn}]$$

$$B_X(r) = \int_0^r B_x = \Omega_{xx}^{1/2} W_X = \Omega_{xx}^{1/2} \int_0^r W_x$$

$$B_{X.x}(r) = B_X(r) - \int B_X B_x \left(\int B_x B_x'\right)^{-1} B_x(r) = \Omega_{xx}^{1/2}\, W_{X.x}(r)$$

$$W_{X.x}(r) = W_X(r) - \int W_X W_x' \left(\int W_x W_x'\right)^{-1} W_x(r)$$

$$\overrightarrow{B_{X.x}}(r) = \int_r^1 B_{X.x} \quad = \Omega_{xx}^{1/2}\, \overrightarrow{W_{X.x}}(r) = \Omega_{xx}^{1/2} \int_r^1 W_{X.x}$$

$$\mathcal{A}_{X.x} = \int B_{X.x} B_{X.x}' = \Omega_{xx}^{1/2} \mathcal{A}_{W,X.x} \Omega_{xx}^{1/2} = \Omega_{xx}^{1/2} \int W_{X.x} W_{X.x}'\, \Omega_{xx}^{1/2}$$

$$B_{e.x}(r) = B_e(r) - \omega_{ex}\Omega_{xx}^{-1} B_x(r) =_d \omega_{ee.x}^{1/2} W_{e.x}, \quad \omega_{ee.x} = \omega_{ee} - \omega_{ex}\Omega_{xx}^{-1}\omega_{xe},$$

$$B_{0.x}(r) = B_0(r) - \Omega_{0x}\Omega_{xx}^{-1} B_x(r) =_d \Omega_{00.x}^{1/2} W_{0.x}$$

$$B_{x.X}(r) = B_x(r) - \int B_x B_X' \left(\int B_X B_X'\right)^{-1} B_X(r) =_d \Omega_{xx}^{1/2} W_{x.X}$$

$$W_{x.X}(r) = W_x(r) - \int W_x W_X' \left(\int W_X W_X'\right)^{-1} W_X(r)$$

$$\begin{aligned}
\widetilde{B_{0.x}}(r) &= B_{0.x}(r) - \int B_{0.x} B_{X.x}' \left(\int B_{X.x} B_{X.x}'\right)^{-1} B_X(r) \\
&\quad - \int B_{0.x} B_{x.X}' \left(\int B_{x.X} B_{x.X}'\right)^{-1} B_x(r) \\
&= \Omega_{xx}^{1/2} \widetilde{W_{0.x}}(r)
\end{aligned}$$

$$\eta_K = \int_{r=0}^1 \tilde{\varphi}_K(r)\, B_x(r)'\, dr,\ \xi_K = \int_{r=0}^1 \tilde{\varphi}_K(r)\, dB_x(r)',\ \mu_K = \int_{r=0}^1 \tilde{\varphi}_K(r)\, B_X(r)'\, dr$$

$$\eta_W = \left(\int W_{X.x} W_{X.x}'\right)^{-1} \int \overrightarrow{W_{X.x}} dW_{0.x},$$

$$\eta_{e.x} = \left(\int W_{X.x} W_{X.x}'\right)^{-1} \left(\int W_{X.x} dW_{e.x}\right)$$

$$\Psi_{0.xK} = \int \tilde{\varphi}_K B_{0.x} \equiv \mathcal{N}\left(0, \Omega_{00.x}\left(\int\int (r \wedge s)\, \tilde{\varphi}_K(r)\, \tilde{\varphi}_K(s)'\, drds\right)\right)$$

$$\psi_{eK} = \int \tilde{\varphi}_K dB_e$$

$$J_K = Q_{\xi_K} - Q_{\xi_K}\eta_K\left(\eta_K' Q_{\xi_K}\eta_K\right)^{-1}\eta_K' Q_{\xi_K}$$

$$
\begin{aligned}
\mathcal{J}_q &= [I_q, 0] \\
\vartheta_K &= \int B_{X.x}(r)\,\tilde{\varphi}_K(r)'\,dr = (\zeta_1, ..., \zeta_K) \\
\ell_X' &= \int dB_{e.x}B_{X.x}'\left(\int B_{X.x}B_{X.x}'\right)^{-1} =_d \omega_{ee.x}\ell_{W,X}'\Omega_{xx}^{-1/2} \\
\ell_{W,X}' &= \int dW_{e.x}W_{X.x}'\left(\int W_{X.x}W_{X.x}'\right)^{-1} \\
\ell_x' &= \int dB_{e.x}B_{x.X}\left(\int B_{x.X}B_{x.X}'\right)^{-1} =_d \omega_{ee.x}\ell_{W,x}'\Omega_{xx}^{-1/2} \\
\ell_{W,x}' &= \int dW_{e.x}W_{x.X}'\left(\int W_{x.X}W_{x.X}'\right)^{-1} \\
\ell_+ &= (\ell_X', \ell_x')' \; =_d \omega_{ee.x}\ell_{W,+}'\Omega_{xx}^{-1/2} \\
\ell_{W,+}' &= (\ell_{W,X}', \ell_{W,x}')' \\
B_+(r) &= (B_X(r)', B_x(r)')' = \Omega_{xx}^{1/2}W_+(r) = \Omega_{xx}^{1/2}\left(W_X(r)', W_x(r)'\right)' \\
\mathcal{Q}_B &= B_{e.x}(r) + \ell_+'\int_0^r B_+ = \Omega_{ee.x}^{1/2}\mathcal{Q}_W(r) \\
\mathcal{Q}_W(r) &= W_{e.x}(r) + \ell_{W,+}'\int_0^r W_+,
\end{aligned}
$$

where $\tilde{\varphi}_K(r) = (\varphi_1(r), ..., \varphi_K(r))'$.

References

Andrews, D. W. and Cheng, X. (2012). Estimation and inference with weak, semi-strong, and strong identification. *Econometrica*, 80(5):2153–2211.

Andrews, D. W. and Cheng, X. (2014). GMM estimation and uniform subvector inference with possible identification failure. *Econometric Theory*, 30(2):287–333.

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(2):298–321.

Bruns, S. B., Csereklyei, Z., and Stern, D. I. (2020). A multicointegration model of global climate change. *Journal of Econometrics*, 214(1):175–197.

Corbae, D., Ouliaris, S., and Phillips, P. C. (2002). Band spectral regression with trending data. *Econometrica*, 70(3):1067–1109.

Engle, R. F. (1974). Band spectrum regression. *International Economic Review*, 15:1–11.

Engsted, T., Gonzalo, J., and Haldrup, N. (1997). Testing for multicointegration. *Economics Letters*, 56(3):259–266.

Engsted, T. and Haldrup, N. (1999). Multicointegration in stock-flow models. *Oxford Bulletin of Economics and Statistics*, 61(2):237–254.

Engsted, T. and Johansen, S. (1999). Granger's representation theorem and multicointegration. *Cointegration, Causality and Forecasting: Festschrift in Honour of Clive Granger*, pages 200–212.

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. and Lee, T.-H. (1989). Investigation of production, sales and inventory relationships using multicointegration and non-symmetric error correction models. *Journal of applied econometrics*, 4(S1):S145–S159.

Granger, C. W. J. and Lee, T.-H. (1990). Investigation of production, sales and inventory relationships using multicointegration and non-symmetric error correction models. *Advances in Econometrics*, 8:71–84.

Hannan, E. J. (1963). Regression for time series. In Rosenblatt, M., editor, *Time Series Analysis*, pages 17–37. Wiley.

Hwang, J. and Sun, Y. (2018). Simple, robust, and accurate F and t tests in cointegrated systems. *Econometric Theory*, 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.

Kheifets, I. and Phillips, P. C. B. (2021). Fully modified least squares for multicointegrated systems. *Journal of Econometrics (forthcoming).*

Lazarus, E., Lewis, D. J., Stock, J. H., and Watson, M. W. (2018). HAR inference: Recommendations for practice. *Journal of Business & Economic Statistics*, 36(4):541–559.

Lee, T.-H. (1996). Stock adjustment for multicointegrated series. *Empirical Economics*, 21(4):633–639.

Magdalinos, T. and Phillips, P. C. B. (2019). Wald tests under matrix normalizatio. *Working Paper, University of Southampton.*

Müller, U. K. (2007). A theory of robust long-run variance estimation. *Journal of Econometrics*, 141(2):1331–1352.

Müller, U. K. and Watson, M. W. (2018). Long-run covariability. *Econometrica*, 86(3):775–804.

Park, J. Y. and Phillips, P. C. B. (1988). Statistical inference in regressions with integrated processes: Part 1. *Econometric Theory*, 4(3):468–497.

Park, J. Y. and Phillips, P. C. B. (1989). Statistical inference in regressions with integrated processes: Part 2. *Econometric Theory*, 5(1):95–131.

Phillips, P. C. (1986). Understanding spurious regressions in econometrics. *Journal of econometrics*, 33(3):311–340.

Phillips, P. C. B. (1988). Multiple regression with integrated time series. In Prabhu, N. U., editor, *Statistical Inference from Stochastic Processes*, page 79.

Phillips, P. C. B. (1989). Partially identified econometric models. *Econometric Theory*, 5(2):181–240.

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. (1998). New tools for understanding spurious regressions. *Econometrica*, 66:1299–1325.

Phillips, P. C. B. (2005a). Challenges of trending time series econometrics. *Mathematics and Computers in Simulation*, 68(5-6):401–416.

Phillips, P. C. B. (2005b). HAC estimation by automated regression. *Econometric Theory*, 21(1):116–142.

Phillips, P. C. B. (2006). Optimal estimation of cointegrated systems with irrelevant instruments. *Cowles Foundation Discussion Paper No. 1547, Yale University.*

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

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

Phillips, P. C. B. and Loretan, M. (1991). Estimating long-run economic equilibria. *The Review of Economic Studies*, 58(3):407–436.

Phillips, P. C. B. and Solo, V. (1992). Asymptotics for linear processes. *The Annals of Statistics*, pages 971–1001.

Phillips, P. C. B., Wang, X., and Zhang, Y. (2019). HAR testing for spurious regression in trend. *Econometrics*, 7(4):50.

Saikkonen, P. (1991). Asymptotically efficient estimation of cointegration regressions. *Econometric theory*, 7(1):1–21.

Stock, J. H. and Watson, M. W. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica*, 61:783–820.

Sun, Y. (2004). A convergent t-statistic in spurious regressions. *Econometric Theory*, 20(5):943–962.

Vogelsang, T. J. and Wagner, M. (2014). Integrated modified OLS estimation and fixed-b inference for cointegrating regressions. *Journal of Econometrics*, 178(2):741–760.