ESTIMATION AND INFERENCE IN A POSSIBLY MULTI-COINTEGRATED SYSTEM WITH A FIXED NUMBER OF INSTRUMENTS

By

Yixiao Sun, Peter C. B. Phillips and Igor L. Kheifets

October 2024

COWLES FOUNDATION DISCUSSION PAPER NO. 2410

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Estimation and Inference in a Possibly Multicointegrated System with a Fixed Number of Instruments*

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

[a]University of California, San Diego
[b]Yale University, [c]University of Auckland, [d]Singapore Management University
[e]University of North Carolina at Charlotte

October 10, 2024

Abstract

This note shows that the mixed normal asymptotic limit of the trend IV estimator with a fixed number of deterministic instruments (fTIV) holds in both singular (multicointegrated) and nonsingular cointegration systems, thereby relaxing the exogeneity condition in (Phillips and Kheifets, 2024, Theorem 1(ii)). The mixed normality of the limiting distribution of fTIV allows for asymptotically pivotal F tests about the cointegration parameters and for simple efficiency comparisons of the estimators for different numbers K of instruments, as well as comparisons with the trend IV estimator when $K \to \infty$ with the sample size.

Keywords: Asymptotic F test, Cointegration, Fixed-K asymptotics, Long-run variance, Multicointegration, Singularity, Trend IV estimation.

JEL Codes: C12, C13, C22

1 Asymptotic Mixed Normality and Asymptotic F Test

Phillips and Kheifets (2024, hereafter, PK(2024)) recently developed a high-dimensional approach to estimating cointegrated systems in the presence of multicointegration. That approach allowed the number of deterministic instruments K to pass to infinity with the sample size n, as in earlier work on cointegration (Phillips, 2014). We complement that study by establishing similar asymptotics for estimation and inference with a fixed number of instruments, which can deliver more reliable inference in practical applications, as shown in other related work (Hwang and Sun, 2017; Müller and Watson, 2018).

*Email: yisun@ucsd.edu; peter.phillips@yale.edu; igor.kheifets@charlotte.edu

Applied work often focuses on estimating cointegration coefficients through the following standard cointegrating equation:

$$y_t = x_t' a_0 + u_{0t},\ \Delta x_t = u_{xt},\ \text{for } t = 1, \ldots, n, \tag{1}$$

where $x_t \in \mathbb{R}^{d_x}$ for some $d_x \in \mathbb{Z}^+$ and $u_t := (u_{0t}, u_{xt}')'$ are weakly dependent with long-run variance matrix

$$\Omega = \begin{bmatrix} \Omega_{00} & \Omega_{0x} \\ \Omega_{x0} & \Omega_{xx} \end{bmatrix} \in \mathbb{R}^{(d_x+1)\times(d_x+1)}. \tag{2}$$

This note is concerned with estimating the cointegrating vector a_0 in the presence of endogeneity. For this purpose, an augmented form of (1) is useful, as in equation (9) of PK(2024),

$$y_t = x_t' a_0 + \Delta x_t' f_0 + u_{0\cdot x,t},\ u_{0\cdot x,t} = u_{0t} - \Omega_{0x}\Omega_{xx}^{-1} u_{xt}, \tag{3}$$

with $f_0 = \Omega_{xx}^{-1}\Omega_{x0}$ and conditional long-run variance $\Omega_{00\cdot x} = \Omega_{00} - \Omega_{0x}\Omega_{xx}^{-1}\Omega_{x0} \geq 0$. We consider both the standard cointegration case where $\Omega_{00\cdot x} > 0$ and the multicointegration case where $\Omega_{00\cdot x} = 0$; however, our primary focus is on the latter. In this case, we write $u_{0\cdot x,t} = \Delta e_t$, thereby assuring the presence of multicointegration, where e_t has both positive variance and long-run variance. To simplify the exposition, we assume the initial condition $e_0 = 0$. In practice, the initial condition can be accommodated by including an intercept in the regression, as discussed in PK(2024).

Denote Brownian motion with variance (matrix) $\mathcal{V}$ by $\mathrm{BM}(\mathcal{V})$, let "$\rightsquigarrow$" signify weak convergence in the relevant probability space, and make the following assumption.

Assumption 1 ***(Functional Central Limit Theorem (FCLT))***

(a) *For the nonsingular case with $\Omega_{00\cdot x} > 0$, the following joint FCLT holds*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor n\cdot \rfloor} \begin{pmatrix} u_{0t} \\ u_{xt} \end{pmatrix} \rightsquigarrow \begin{pmatrix} B_0(\cdot) \\ B_x(\cdot) \end{pmatrix} \equiv \mathrm{BM}(\Omega),$$

where Ω, given in (2), is positive definite.

(b) *For the singular case with $\Omega_{00\cdot x} = 0$, the following joint FCLT holds*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor n\cdot \rfloor} \begin{pmatrix} e_t \\ u_{xt} \end{pmatrix} \rightsquigarrow \begin{pmatrix} B_e(\cdot) \\ B_x(\cdot) \end{pmatrix} \equiv \mathrm{BM}\left(\begin{bmatrix} \omega_{ee} & \omega_{ex} \\ \omega_{xe} & \Omega_{xx} \end{bmatrix} \right),$$

where the variance matrix of the above Brownian motion is positive definite.

Define

$$e_{0\cdot x,t} = e_t - \omega_{ex}\Omega_{xx}^{-1} u_{xt} \text{ and } \omega_{ee\cdot x} = \omega_{ee} - \omega_{ex}\Omega_{xx}^{-1}\omega_{xe}.$$

Under Assumption 1, for the case $\Omega_{00\cdot x} > 0$, we have

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor n\cdot \rfloor} \begin{pmatrix} u_{0\cdot x,t} \\ u_{xt} \end{pmatrix} \rightsquigarrow \begin{pmatrix} B_{0\cdot x}(\cdot) \\ B_x(\cdot) \end{pmatrix} \equiv \mathrm{BM}\left(\begin{bmatrix} \Omega_{00\cdot x} & \mathbf{0}' \\ \mathbf{0} & \Omega_{xx} \end{bmatrix} \right);$$

and for the case $\Omega_{00\cdot x} = 0$,

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor n\cdot \rfloor} \left(\begin{array}{c} e_{0\cdot x,t} \\ u_{xt} \end{array} \right) \rightsquigarrow \left(\begin{array}{c} B_{e\cdot x}(\cdot) \\ B_x(\cdot) \end{array} \right) \equiv \mathrm{BM}\left(\left[\begin{array}{cc} \omega_{ee\cdot x} & \mathbf{0}' \\ \mathbf{0} & \Omega_{xx} \end{array} \right] \right),$$

where $\mathbf{0}$ is a d_x-vector of zeros, with $B_{0\cdot x} := B_0 - \Omega_{0x}\Omega_{xx}^{-1}B_x$ and $B_{e\cdot x} := B_e - \omega_{ex}\Omega_{xx}^{-1}B_x$ independent of B_x.

Using capitals to signify partial summation, we write $Y_t = \sum_{s=1}^{t} y_s$, $X_t = \sum_{s=1}^{t} x_s$, $U_{0\cdot x,t} = \sum_{s=1}^{t} u_{0\cdot x,s}$. Define

$$\begin{aligned} e_{0\cdot x,t}^{+} &= U_{0\cdot x,t} \cdot \mathbf{1}\{\Omega_{00\cdot x} > 0\} + e_{0\cdot x,t} \cdot \mathbf{1}\{\Omega_{00\cdot x} = 0\}, \\ g_0 &= \Omega_{xx}^{-1}\omega_{xe} \cdot \mathbf{1}\{\Omega_{00\cdot x} = 0\}, \end{aligned}$$

where $\mathbf{1}\{\cdot\}$ is the indicator function. In matrix form we have the doubly augmented model

$$Y = [X, C]\,\gamma_0 + e^{+}, \text{ for } \gamma_0' := (a_0', f_0', g_0') := (a_0', \ell_0'), \tag{4}$$

where

$$Y = [Y_1, \ldots, Y_n]', X = [X_1, \ldots, X_n]',$$

$$C := \left[\begin{array}{ccc} x_1 & \cdots & x_n \\ \Delta x_1 & \cdots & \Delta x_n \end{array} \right]' = \left[\begin{array}{c} x' \\ \Delta x' \end{array} \right]',$$

and

$$e^{+} = \left[e_{0\cdot x,1}^{+}, \ldots, e_{0\cdot x,n}^{+}\right]'.$$

PK(2024) estimated the above cointegrating model using deterministic instrumental variables $\{\varphi_j(t/n)\}_{j=1}^{K}$, where $\{\varphi_j(r)\}_{j=1}^{\infty}$ is a complete set of basis functions of $L_2[0,1]$. In what follows, we let

$$\tilde{\varphi}_K(r) = (\varphi_1(r), \ldots, \varphi_K(r))', \ \tilde{\varphi}_{K,t} = \tilde{\varphi}_K\left(\frac{t}{n}\right) = \left[\varphi_1\left(\frac{t}{n}\right), \ldots, \varphi_K\left(\frac{t}{n}\right)\right]',$$

and so

$$\Phi_K = [\tilde{\varphi}_{K,1}, \ldots, \tilde{\varphi}_{K,n}]'$$

is the observation matrix of the instruments. The projection matrix that projects onto the column space of Φ_K is given by $P_{\Phi_K} = \Phi_K\left(\Phi_K'\Phi_K\right)^{-1}\Phi_K'$.

Based on the K instrumental variables $\tilde{\varphi}_{K,t}$, the trend IV (TIV) estimator of a_0 is defined as:

$$\hat{a}_{\mathrm{TIV}} = \arg\min_a (Y - Xa)' R_K (Y - Xa) = \left(X'R_KX\right)^{-1}(X'R_KY), \tag{5}$$

where

$$R_K = P_{\Phi_K} - P_{\Phi_K}C\left(C'P_{\Phi_K}C\right)^{-1}C'P_{\Phi_K}. \tag{6}$$

Alternatively, and equivalently,

$$(\hat{a}_{\mathrm{TIV}}, \hat{\ell}_{\mathrm{TIV}}) = \arg\min_{(a,\ell)} (Y - Xa - C\ell)' P_{\Phi_K} (Y - Xa - C\ell).$$

If $\left\|n^{-1}\Phi_K'\Phi_K - I_K\right\|_2 = o(1)$ so that $\left\|P_{\Phi_K} - n^{-1}\Phi_K\Phi_K'\right\|_2 = o(1)$ under an asymptotic specification of K (either fixed or growing with n), then the TIV is asymptotically equivalent to OLS applied to the transformed and augmented system

$$V_Y = V_X a_0 + V_x f_0 + V_{\Delta x} g_0 + V_{e^+} = V_X a_0 + V_C \ell_0 + V_{e^+}, \tag{7}$$

where we employ the notation $V_Z = \Phi_K' Z$ for an observation matrix Z. Transformations to V_Z were used, for example, in Hwang and Sun (2018). Standard partitioned least squares regression on (7) leads to the following estimator of a_0 :

$$\hat{a}_{\text{fTIV}} = \left(V_X' Q_{V_C} V_X\right)^{-1} V_X' Q_{V_C} V_Y,$$

where, for an observation matrix Z with d_Z rows, $Q_Z = I_{d_Z} - P_Z$ for $P_Z = Z\left(Z'Z\right)^{-1} Z'$. The estimator $\hat{a}_{\text{fTIV}}$ is the same as $\hat{a}_{\text{TIV}}$ but with P_{Φ_K} replaced by $n^{-1}\Phi_K\Phi_K'$ in the definitions of $\hat{a}_{\text{TIV}}$ and R_K in (5) and (6). A similar construction gives estimators $\hat{f}_{\text{fTIV}}$ and $\hat{g}_{\text{fTIV}}$ of f_0 and g_0 as

$$(\hat{f}_{\text{fTIV}}', \hat{g}_{\text{fTIV}}')' = \left(V_C' Q_{V_X} V_C\right)^{-1} V_C' Q_{V_X} V_Y.$$

The estimators $\hat{a}_{\text{fTIV}}$, $\hat{f}_{\text{fTIV}}$, and $\hat{g}_{\text{fTIV}}$ are the fixed-K Trend IV (fTIV) estimators in PK(2024), which may also be referred to as the transformed and augmented OLS (TA-OLS), following Hwang and Sun (2018). In this note, we use the same notation and terminology as in PK(2024). We focus on $\hat{a}_{\text{fTIV}}$, for which the estimation error is given by

$$\hat{a}_{\text{fTIV}} - a_0 = \left(V_X' Q_{V_C} V_X\right)^{-1} V_X' Q_{V_C} V_{e^+}. \tag{8}$$

Unless stated otherwise, throughout this note $\hat{a}_{\text{fTIV}}$ is the transformed and augmented OLS based on a fixed number of basis functions (i.e., K is fixed), while $\hat{a}_{\text{TIV}}$ is the trend IV estimator based on an increasing number of instruments (i.e., the high-dimensional trend IV estimator that lets K approach infinity as the sample n size grows). Nevertheless, both estimators can be analyzed under both types of asymptotics.

To establish fixed-K asymptotics we make the following assumption about the basis functions:

Assumption 2 $\{\varphi_j(\cdot)\}_{j=1}^K$ *are continuously differentiable basis functions on* $L_2[0,1]$.

For ease of comparison, we use the same definitions as in PK(2024) given below:

$$\begin{aligned}
B_X(r) &= \int_0^r B_x(s)\,ds,\ \mu_K = \int_0^1 \tilde{\varphi}_K(r)\, B_X'(r)\,dr,\ \eta_K = \int_0^1 \tilde{\varphi}_K(r)\, B_x(r)'\,dr, \\
\xi_K &= \int_0^1 \tilde{\varphi}_K(r)\, dB_x(r)'\,;\ \psi_{0\cdot x,K} = \int_0^1 \tilde{\varphi}_K(r)\, B_{0\cdot x}(r)\,dr; \\
\psi_{e\cdot x,K} &= \int_0^1 \tilde{\varphi}_K(r)\, dB_{e\cdot x}(r) = \int_0^1 \tilde{\varphi}_K(r)\, dB_e(r) - \xi_K \Omega_{xx}^{-1}\omega_{xe}; \\
Q_{\xi_K} &= I_K - \xi_K(\xi_K'\xi_K)^{-1}\xi_K'; \\
J_K &= Q_{\xi_K} - Q_{\xi_K}\eta_K\left(\eta_K' Q_{\xi_K}\eta_K\right)^{-1}\eta_K' Q_{\xi_K};\ \text{and } S_K = J_K\mu_K\left(\mu_K' J_K \mu_K\right)^{-1}.
\end{aligned}$$

Note that μ_K, η_K and ξ_K are $K \times d_x$ matrices, $\psi_{0\cdot x,K}$ and $\psi_{e\cdot x,K}$ are K-vectors, and $J_K = Q_{[\xi_K,\eta_K]}$, which projects onto the orthogonal complement of the space spanned by $[\xi_K, \eta_K]$.

Theorem 1 ***(Asymptotic Mixed Normality of fTIV)*** *Let Assumptions 1 and 2 hold.*

(a) *When $\Omega_{00\cdot x} > 0$ we have, for fixed K as $n \to \infty$,*

$$n\left(\hat{a}_{\text{fTIV}} - a_0\right) \rightsquigarrow S_K' \psi_{0\cdot x,K} \equiv \mathcal{MN}\left(0, \Omega_{00\cdot x} S_K' \left(\int_0^1 \int_0^1 (r \wedge s)\, \tilde{\varphi}_K(r)\, \tilde{\varphi}_K(s)'\, drds\right) S_K\right). \tag{9}$$

(b) *When $\Omega_{00\cdot x} = 0$ we have, for fixed K as $n \to \infty$,*

$$n^2\left(\hat{a}_{\text{fTIV}} - a_0\right) \rightsquigarrow S_K' \psi_{e\cdot x,K} \equiv \mathcal{MN}\left(0, \omega_{ee\cdot x} S_K' \left(\int_0^1 \tilde{\varphi}_K(r)\, \tilde{\varphi}_K'(r)\, dr\right) S_K\right). \tag{10}$$

Proof of Theorem 1. By virtue of summation by parts, integration by parts, and the continuous mapping theorem, the following weak convergence results hold:

(a) $n^{-1/2} V_{e^+} \rightsquigarrow \psi_{e\cdot x,K}$ when $\Omega_{00\cdot x} = 0$ and $n^{-3/2} V_{e^+} \rightsquigarrow \psi_{0\cdot x,K}$ when $\Omega_{00\cdot x} > 0$;

(b) $n^{-1/2} V_{\Delta x} \rightsquigarrow \xi_K$;

(c) $n^{-3/2} V_x \rightsquigarrow \eta_K$;

(d) $n^{-5/2} V_X \rightsquigarrow \mu_K$.

Then, for Part (a),

$$n\left(\hat{a}_{\text{fTIV}} - a_0\right) = \left(n^{-5} V_X' Q_{V_C} V_X\right)^{-1} n^{-4} V_X' Q_{V_C} V_{e^+} \rightsquigarrow \left(\mu_K' J_K \mu_K\right)^{-1}\left(\mu_K' J_K \psi_{0\cdot x,K}\right) = S_K' \psi_{0\cdot x,K}.$$

Since the randomness of (μ_K, η_K, ξ_K) is fully driven by $B_x(\cdot)$, which is uncorrelated with and hence independent of $B_{0\cdot x}(\cdot)$, it follows that $\psi_{0\cdot x,K} = \int_0^1 \tilde{\varphi}_K(r)\, B_{0\cdot x}(r)\, dr$ is independent of (μ_K, η_K, ξ_K). Therefore, conditional on (μ_K, η_K, ξ_K), $\psi_{0\cdot x,K}$ follows the normal distribution $\mathcal{N}\left(0, \Omega_{00\cdot x}\left(\int_0^1 \int_0^1 (r \wedge s)\, \tilde{\varphi}_K(r)\, \tilde{\varphi}_K(s)'\, drds\right)\right)$. Consequently, the limit distribution is mixed normal:

$$n\left(\hat{a}_{\text{fTIV}} - a_0\right) \rightsquigarrow S_K' \psi_{0\cdot x,K} \equiv \mathcal{MN}\left(0, \Omega_{00\cdot x} S_K' \left(\int_0^1 \int_0^1 (r \wedge s)\, \tilde{\varphi}_K(r)\, \tilde{\varphi}_K(s)'\, drds\right) S_K\right).$$

For Part (b),

$$\begin{aligned} n^2\left(\hat{a}_{\text{fTIV}} - a_0\right) &= \left(n^{-5} V_X' Q_{V_C} V_X\right)^{-1} n^{-3} V_X' Q_{V_C} V_{e^+} \\ &\rightsquigarrow \left(\mu_K' J_K \mu_K\right)^{-1}\left(\mu_K' J_K \psi_{e\cdot x,K}\right) = S_K' \psi_{e\cdot x,K}. \end{aligned}$$

Since (μ_K, η_K, ξ_K) depends only on $B_x(\cdot)$, which is uncorrelated with and hence independent of $B_{e\cdot x}(\cdot)$, $\psi_{e\cdot x,K} = \int_0^1 \tilde{\varphi}_K(r)\, dB_{e\cdot x}(r)$ is independent of (μ_K, η_K, ξ_K). Conditional on

(μ_K, η_K, ξ_K), $\psi_{e\cdot x,K}$ follows the normal distribution $\mathcal{N}(0, \omega_{ee\cdot x} \int_0^1 \tilde{\varphi}_K(r) \tilde{\varphi}_K'(r) dr)$. Therefore, the limit distribution is mixed normal:

$$n^2 (\hat{a}_{\text{fTIV}} - a_0) \rightsquigarrow S_K' \psi_{e\cdot x,K} \equiv \mathcal{MN}\left(0, \omega_{ee\cdot x} S_K' \left(\int_0^1 \tilde{\varphi}_K(r) \tilde{\varphi}_K'(r) dr\right) S_K\right).$$

■

In both the singular and nonsingular cases, the limiting distribution is mixed normal with a zero mean. Unlike the use of OLS applied directly to (3), the fTIV estimator has no second-order endogeneity bias. This bias is removed simply by using an IV approach that involves deterministic instruments.

The asymptotic mixed normality in the multicointegration case was shown in PK(2024) under the exogeneity assumption that $\omega_{xe} = 0$ (i.e., $g_0 = 0$), indicating that the multicointegration error $\{e_t\}$ has no long-run correlation with the integrated process $\{x_t\}$. Theorem 1(b) establishes the asymptotic mixed normality without the exogeneity assumption. Also, unlike the result in PK(2024), Theorem 1 provides the asymptotic distribution of the fTIV estimator for any continuously differentiable basis functions. Flexibility in the choice of the basis gives an additional advantage as we discuss next.

The conditional asymptotic variance of the fTIV depends on the basis functions used. When $\Omega_{00\cdot x} > 0$, we may choose $\{\varphi_j(\cdot)\}_{j=1}^K$ such that

$$\int_0^1 \int_0^1 (r \wedge s) \tilde{\varphi}_K(r) \tilde{\varphi}_K(s)' drds = I_K, \tag{11}$$

in which case we have

$$n(\hat{a}_{\text{fTIV}} - a_0) \rightsquigarrow \mathcal{MN}(0, \Omega_{00\cdot x} S_K' S_K) = \mathcal{MN}(0, \Omega_{00\cdot x} (\mu_K' J_K \mu_K)^{-1}).$$

Similarly, when $\Omega_{00\cdot x} = 0$, we may choose $\{\varphi_j(\cdot)\}_{j=1}^K$ such that

$$\int_0^1 \tilde{\varphi}_K(r) \tilde{\varphi}_K'(r) dr = I_K,$$

that is, $\{\varphi_j(\cdot)\}_{j=1}^K$ are orthonormal basis functions of $L_2[0,1]$. Then

$$n^2 (\hat{a}_{\text{fTIV}} - a_0) \rightsquigarrow \mathcal{MN}(0, \omega_{ee\cdot x} S_K' S_K) = \mathcal{MN}(0, \omega_{ee\cdot x} (\mu_K' J_K \mu_K)^{-1}). \tag{12}$$

In both cases, the conditional variance matrix in the mixed normal distribution, which takes a sandwich product form in general, collapses to a single matrix component.

The zero-mean mixed normal asymptotic distribution enables the construction of asymptotically pivotal tests about cointegration parameters. Consider, for example, the case of multicointegration using the fTIV based on orthonormal basis functions. To test $H_0 : Ha = h$ against $H_1 : Ha \neq h$ for some restriction matrix $H \in \mathbb{R}^{p \times d_x}$ and vector $h \in \mathbb{R}^{p \times 1}$, we first obtain the OLS residual vector:

$$\hat{V}_{e^+} = V_Y - V_X \hat{a}_{\text{fTIV}} - V_x \hat{f}_{\text{fTIV}} - V_{\Delta x} \hat{g}_{\text{fTIV}},$$

and then construct the long-run variance estimator:

$$\hat{\omega}_{ee\cdot x} = \frac{1}{K}\left\|\hat{V}_{e^+}\right\|^2,$$

where $\|\cdot\|$ denotes the Euclidean norm. Based on $\hat{\omega}_{ee\cdot x}$, we calculate the Wald statistic in the usual way:

$$\mathbb{W}_{\text{fTIV}} = \frac{1}{\hat{\omega}_{ee\cdot x}}\left[H\hat{a}_{\text{fTIV}} - h\right]'\left[H\left(V_X' Q_{V_C} V_X\right)^{-1} H'\right]^{-1}\left[H\hat{a}_{\text{fTIV}} - h\right]/p. \tag{13}$$

Theorem 2 ***(Asymptotic F test with fTIV)*** *Let Assumptions 1 and 2 hold. In the case of multicointegration with $\Omega_{00\cdot x} = 0$ and $K > 3d_x$, if $\{\varphi_j(\cdot)\}_{j=1}^K$ are orthonormal basis functions of $L_2[0,1]$, and H has full row rank p, then*

$$\mathbb{W}^*_{\text{fTIV}} := \frac{K - 3d_x}{K}\mathbb{W}_{\text{fTIV}} \rightsquigarrow F_{p,K-3d_x}, \tag{14}$$

for fixed K as $n \to \infty$, where $F_{p,K-3d_x}$ is the standard F distribution with degrees of freedom p and $K - 3d_x$.

Proof of Theorem 2. Since $\hat{V}_{e^+} = Q_{[V_X,V_C]}V_{e^+}$, we have

$$\hat{\omega}_{ee\cdot x} = V_{e^+}' Q_{[V_X,V_C]} V_{e^+}/K.$$

Using this and Theorem 1, we obtain:

$$\begin{aligned}
\mathbb{W}_{\text{fTIV}} &= \frac{\left[H\hat{a}_{\text{fTIV}} - h\right]'\left[H\left(V_X' Q_{V_C} V_X\right)^{-1} H'\right]^{-1}\left[H\hat{a}_{\text{fTIV}} - h\right]/p}{V_{e^+}' Q_{[V_X,V_C]} V_{e^+}/K} \\
&= \frac{\left[\hat{a}_{\text{fTIV}} - a_0\right]' H'\left[H\left(V_X' Q_{V_C} V_X\right)^{-1} H'\right]^{-1} H\left[\hat{a}_{\text{fTIV}} - a_0\right]/p}{V_{e^+}' Q_{[V_X,V_C]} V_{e^+}/K} \\
&\rightsquigarrow \frac{\psi_{e\cdot x,K}' S_K H'\left[H\left(\mu_K' J_K \mu_K\right)^{-1} H'\right]^{-1} H S_K' \psi_{e\cdot x,K}/p}{\psi_{e\cdot x,K}' Q_{[\mu_K,\eta_K,\xi_K]}\psi_{e\cdot x,K}/K} \\
&= \frac{\left\|P_{[S_K H']}\psi_{e\cdot x,K}\right\|^2/p}{\left\|Q_{[\mu_K,\eta_K,\xi_K]}\psi_{e\cdot x,K}\right\|^2/K}.
\end{aligned} \tag{15}$$

Under the assumption that $\{\varphi_j(\cdot)\}_{j=1}^K$ are orthonormal, $\psi_{e\cdot x,K}$ follows the normal distribution $\mathcal{N}(0, \omega_{ee\cdot x} I_K)$. Hence, conditional on (μ_K, η_K, ξ_K),

$$\left\|P_{[S_K H']}\psi_{e\cdot x,K}\right\|^2/\omega_{ee\cdot x} =^d \chi^2_p,$$
$$\left\|Q_{[\mu_K,\eta_K,\xi_K]}\psi_{e\cdot x,K}\right\|^2/\omega_{ee\cdot x} =^d \chi^2_{K-3d_x},$$

where $=^d$ denotes distributional equivalence. The two chi-square variates above are conditionally independent, as they are based on two conditionally independent normals, namely

$Q_{[\mu_K,\xi_K,\eta_K]}\psi_{e\cdot x,K}$ and $S_K'\psi_{e\cdot x,K}$. The conditional independence between these two normals holds because, conditional on (μ_K,η_K,ξ_K), we have

$$\begin{aligned}
&cov(Q_{[\mu_K,\eta_K,\xi_K]}\psi_{e\cdot x,K}, S_K'\psi_{e\cdot x,K}) \\
&= \omega_{ee\cdot x}Q_{[\mu_K,\eta_K,\xi_K]}S_K = \omega_{ee\cdot x}Q_{[\mu_K,\eta_K,\xi_K]}Q_{[\xi_K,\eta_K]}\mu_K\left(\mu_K'J_K\mu_K\right)^{-1} \\
&= \omega_{ee\cdot x}\left\{Q_{[\xi_K,\eta_K]} - P_{\left[Q_{[\xi_K,\eta_K]}\mu_K\right]}\right\}Q_{[\xi_K,\eta_K]}\mu_K\left(\mu_K'J_K\mu_K\right)^{-1} \\
&= \omega_{ee\cdot x}Q_{[\xi_K,\eta_K]}\mu_K\left(\mu_K'J_K\mu_K\right)^{-1} - \omega_{ee\cdot x}Q_{[\xi_K,\eta_K]}\mu_K\left(\mu_K'J_K\mu_K\right)^{-1} \\
&= 0.
\end{aligned}$$

Therefore, conditional on (μ_K,η_K,ξ_K) ,

$$\frac{\left\|P_{[S_KH']}\psi_{e\cdot x,K}\right\|^2/p}{\left\|Q_{[\mu_K,\eta_K,\xi_K]}\psi_{e\cdot x,K}\right\|^2/K} =^d \frac{\chi_p^2/p}{\chi_{K-3d_x}^2/K},$$

and

$$\frac{K-3d_x}{K}\frac{\left\|P_{[S_KH']}\psi_{e\cdot x,K}\right\|^2/p}{\left\|Q_{[\mu_K,\eta_K,\xi_K]}\psi_{e\cdot x,K}\right\|^2/K} =^d \frac{\chi_p^2/p}{\chi_{K-3d_x}^2/\left(K-3d_x\right)} =^d F_{p,K-3d_x}.$$

The conditional distribution does not depend on the conditioning variables (μ_K,η_K,ξ_K), and hence it is also the unconditional distribution. We have therefore shown that:

$$\mathbb{W}_{\text{fTIV}}^* = \frac{K-3d_x}{K}\mathbb{W}_{\text{fTIV}} \rightsquigarrow \frac{\chi_p^2/p}{\chi_{K-3d_x}^2/\left(K-3d_x\right)} =^d F_{p,K-3d_x}.$$

■

Theorem 2 is new and extends the corresponding result in Hwang and Sun (2018), which rules out multicointegration, to the case of multicointegration. An asymptotic F test based on $\hat{a}_{\text{fTIV}}$ can also be developed for the usual nonsingular cointegration case if the basis functions satisfy (11). For example, we can use the eigenfunctions of the covariance kernel operator $r \wedge s$ as the basis functions. We omit the details here. If the null hypothesis involves a single restriction (i.e., $p=1$), then it can be shown that the t-statistic based on OLS applied to (7) is asymptotically t-distributed. The asymptotic F and t approximations are not only convenient to use but also more accurate, as the F and t distributions capture the estimation errors in estimating f_0, g_0, and $\omega_{ee\cdot x}$ (or $\Omega_{00\cdot x}$), which are often ignored by the fully modified methodology. For the F and t asymptotic theory in other nonstationary and stationary settings, the reader is referred to Sun (2023), Hwang and Sun (2017), and the references therein.

In practical work when the doubly augmented model (4) is modified to include an intercept and the basis functions do not integrate to zero, we have, in place of (14),

$$\mathbb{W}_{\text{fTIV}}^* := \frac{K-3d_x-1}{K}\mathbb{W}_{\text{fTIV}} \rightsquigarrow F_{p,K-3d_x-1},$$

provided that $K > 3d_x+1$. This adjustment occurs because there is an additional regressor in the transformed and augmented model in (7), resulting in a loss of one degree of freedom. No adjustment is needed so that (14) still holds if the basis functions integrate to zero ($\int_0^1 \varphi_j(r)\,dr = 0$ for $j = 1, ..., K$). Deterministic trends can also be included in (4), and the asymptotic F limit theory remains valid, albeit with a different multiplicative adjustment factor and different degrees of freedom for the F distribution.

2 Asymptotic Relative Efficiency

We now compare the asymptotic distributions of the fTIV and TIV estimators under Assumptions 1 and 2. For the large-K asymptotic results in this section, we impose the stronger assumption that $\{\varphi_j(\cdot)\}_{j=1}^{\infty}$ is a complete set of twice continuously differentiable and orthonormal basis functions of $L_2[0,1]$. For example, we can take $\varphi_j(r) = \sqrt{2}\sin((j-1/2)\pi r)$ for $j = 1, 2, \ldots, K$. Theorem 2 of PK(2024) showed in the cointegration case with $\Omega_{00\cdot x} > 0$ under joint large-K asymptotics where both $K \to \infty$ and $n \to \infty$, but $K = o\left(n^{4/5-\delta}\right)$, that for some $\delta > 0$

$$n\left(\hat{a}_{\text{TIV}} - a_0\right) \rightsquigarrow \mathcal{A}_{X\cdot x}^{-1} \int_0^1 \overrightarrow{B_{X\cdot x}}(r)\, dB_{0\cdot x}(r) \equiv \mathcal{MN}\left(0, \Omega_{00\cdot x}\mathcal{A}_{X\cdot x}^{-1} \int_0^1 \overrightarrow{B_{X\cdot x}}(r)\, \overrightarrow{B_{X\cdot x}}(r)'\, dr \mathcal{A}_{X\cdot x}^{-1}\right),$$

where $\mathcal{A}_{X\cdot x} = \int_0^1 B_{X\cdot x}(r)\, B'_{X\cdot x}(r)\, dr$, $\overrightarrow{B_{X\cdot x}}(r) = \int_r^1 B_{X\cdot x}(s)\, ds$, and

$$B_{X\cdot x}(r) = B_X(r) - \left(\int_0^1 B_X(s)\, B'_x(s)\, ds\right)\left(\int_0^1 B_x(s)\, B'_x(s)\, ds\right)^{-1} B_x(r).$$

The asymptotic result can be equivalently presented as

$$\begin{aligned} n\left(\hat{a}_{\text{TIV}} - a_0\right) &\rightsquigarrow \mathcal{A}_{X\cdot x}^{-1}\left(\int_0^1 B_{X\cdot x}(r)\, B_{0\cdot x}(r)\, dr\right) \\ &\equiv \mathcal{MN}\left(0, \Omega_{00\cdot x}\mathcal{A}_{X\cdot x}^{-1} \int_0^1 \int_0^1 (r \wedge s)\, B_{X\cdot x}(r)\, B_{X\cdot x}(s)'\, dr ds \mathcal{A}_{X\cdot x}^{-1}\right). \quad (16) \end{aligned}$$

The above representation takes a form similar to that given in Theorem 1(a), as both contain the covariance kernel (i.e., $(r \wedge s)$) of standard Brownian motion. PK(2024) also showed in their Theorem 2 that in the multicointegration case with $\Omega_{00\cdot x} = 0$,

$$n^2\left(\hat{a}_{\text{TIV}} - a_0\right) \rightsquigarrow \mathcal{A}_{X\cdot x}^{-1}\left(\int_0^1 B_{X\cdot x}(r)\, dB_{e\cdot x}(r)\right) \equiv \mathcal{MN}\left(0, \omega_{ee\cdot x}\mathcal{A}_{X\cdot x}^{-1}\right) \quad (17)$$

under the joint large-K asymptotics specified above.

Following arguments similar to those in PK(2024), we can show that if trigonometric orthonormal polynomials are used as the basis functions, $\hat{a}_{\text{fTIV}}$ and $\hat{a}_{\text{TIV}}$ share the same large-K asymptotic distributions, as given in (16) and (17), respectively, for the cointegration and multicointegration cases.

Under multicointegration, it turns out that when K grows with n to infinity, the trend IV method provides asymptotically jointly efficient estimators of the cointegrating coefficient a_0 and the multicointegrating coefficient f_0. The joint mixed normal limit distribution of a trend IV estimator corresponds to that in a multicointegrated, correctly specified parametric VAR model with iid Gaussian innovations, as shown in Kheifets and Phillips (2024). The reason is that the trend IV method not only fully removes endogeneity effects by reducing the error process asymptotically to $B_{e\cdot x}$ but also fully captures the path of the regressors, reproducing B_X, B_x, and $B_{X\cdot x}$ in the limit as $K \to \infty$. This asymptotic efficiency result extends that in Phillips (1991), which considered only the nonsingular cointegration case and dealt only with optimal estimation of cointegrating coefficients. Moreover, the result shows that precise VAR specification (with assumed iid errors) is unnecessary for optimal estimation provided that efficient methods like the trend IV method with a growing number of instruments are employed. It is particularly noteworthy that this observation applies to the regression coefficient that is effectively nonparametric, as the multicointegrating coefficient f_0 is a nonparametric long-run regression coefficient.

It can be shown that when we let $K \to \infty$, the fixed-K asymptotic distributions in (9) and (10) of Theorem 1 converge to the joint large-K asymptotic distributions in (16) and (17), respectively. From a theoretical perspective, the large-K asymptotic distributions of the fTIV estimator can be obtained using a two-step sequential limit, where we first hold K fixed and let $n \to \infty$, followed by letting $K \to \infty$. Given this, it is of interest to compare the asymptotic distributions of fTIV for different numbers of instruments, including those of TIV (where $K \to \infty$ at a certain rate). Because the asymptotic distributions are all mixed normal, it is simplest to compare variances or standard deviations. To this end, we compute the ratio of the (asymptotic and random) standard derivation of the fTIV to that of the TIV. For the cointegration case with $\Omega_{00\cdot x} > 0$, the ratio is

$$\sqrt{\frac{S_K' \left(\int_0^1 \int_0^1 (r \wedge s)\, \tilde{\varphi}_K(r)\, \tilde{\varphi}_K(s)'\, drds\right) S_K}{\mathcal{A}_{X\cdot x}^{-1} \left(\int_0^1 \int_0^1 (r \wedge s)\, B_{X\cdot x}(r)\, B_{X\cdot x}(s)'\, drds\right) \mathcal{A}_{X\cdot x}^{-1}}}.$$

For the multicointegration case with $\Omega_{00\cdot x} = 0$, the ratio is

$$\sqrt{\frac{\mathcal{A}_{X\cdot x}}{\mu_K' J_K \mu_K.}}$$

We simulate these ratios for the case when $d_x = 1$, so that $B_x(\cdot)$ is one-dimensional Brownian motion. In this case, the ratios do not depend on the variance of $B_x(\cdot)$, so it can be replaced by standard Brownian motion. We simulate standard Brownian motion by $\{\sum_{t=1}^{[nr]} u_{x,t}/\sqrt{n} : r \in [0,1]\}$, where $u_{x,t} \sim iid\mathcal{N}(0,1)$. We set $n = 10,000$, and the number of simulation replications is also 10,000. The results are presented in Table 1. Under cointegration with $\Omega_{00\cdot x} > 0$, in 75% of the cases, the standard deviation of the fTIV is no more than 1.18 times as large as that of the TIV for $K = 7$ and 1.08 times as large for $K = 12$. Under multicointegration with $\Omega_{00\cdot x} = 0$, in 75% of the cases, the standard deviation of fTIV is no more than 1.34 times as large as that of TIV for $K = 7$ and 1.16 times as large for $K = 12$. This shows that the fTIV with a moderately large K becomes nearly as efficient as the TIV.

Table 2 reports the ratios of the confidence interval lengths based on the fTIV and the TIV estimators. Confidence intervals for the cointegration parameter a_0 are defined as $[\hat{a} -$

$q/n^\iota, \hat{a} + q/n^\iota]$, where $\hat{a}$ is either the fTIV or TIV estimator, q is the quantile of the asymptotic distribution given in (9), (10), (16) or (17), and $\iota = 2$ in the case of multicointegration and 1 otherwise. These confidence intervals are infeasible because they depend on unknown conditional variances. However, the ratio of the lengths of the confidence intervals is the ratio of the quantiles, which is nuisance-parameter free and can be easily simulated. For example, in the case of multicointegration, the ratio of the lengths of the 95% confidence intervals is $q_{\text{fTIV}}/q_{\text{TIV}}$ where q_{fTIV} and q_{TIV} are the 95% quantiles of $\mathcal{MN}(0, (\mu_K' J_K \mu_K)^{-1})$ and $\mathcal{MN}\left(0, \mathcal{A}_{X\cdot x}^{-1}\right)$, respectively. Simulations show that 95% confidence intervals based on the fTIV are 25% and 35% larger for $K = 7$ than those based on the TIV in the cointegration and multicointegration cases. For $K = 12$, they are 13% larger in both cases. Note that for the construction of feasible confidence intervals, the length comparisons will depend on the efficiency of estimating both a_0 and the quantiles of the asymptotic distributions.

Table 1: Descriptive statistics of the ratio of the (asymptotic) standard deviations of the fTIV with $K = 7$ and $K = 12$ to that of the TIV ($K \to \infty$).

Model	Cointegration		Multicointegration	
K	7	12	7	12
mean	1.171463	1.070536	1.282318	1.129075
std	0.359790	0.134466	0.329768	0.118773
min	0.584240	0.681827	1.004457	1.003903
25%	1.028656	1.014433	1.097037	1.052504
50%	1.073846	1.035775	1.182339	1.094366
75%	1.183864	1.083237	1.344061	1.163552
max	9.037757	3.874641	6.453239	3.049757

Table 2: Descriptive statistics of ratios of the confidence interval lengths based on the fTIV with $K = 7$ and $K = 12$ to that based on the TIV ($K \to \infty$).

Model	Cointegration		Multicointegration	
K	7	12	7	12
Coverage				
0.99	1.273595	1.122713	1.493161	1.174461
0.95	1.243763	1.128659	1.347902	1.134481
0.90	1.186814	1.093507	1.326981	1.145040

3 Final Remarks

The primary advantage of the high-dimensional trend IV method is its joint asymptotic efficiency in estimating the cointegrating and multicointegrating parameters, while treating the system innovations nonparametrically. The gains in efficiency and confidence region precision from high-dimensional trend IV are evident in simulations but are by no means excessive compared to fTIV with a moderate number of instruments. In practical work, a finite number of instruments is always employed, and the asymptotics that hold the number of instruments fixed may provide more reliable distributional approximations. Under this type of asymptotics, fTIV delivers an asymptotically valid, easy-to-use, and more accurate F test that avoids the use of any additional tuning parameters while retaining the nonparametric advantage of TIV.

References

Hwang, J. and Y. Sun (2017): "Asymptotic F and t tests in an efficient GMM setting," *Journal of Econometrics*, 198, 277–295.

——— (2018): "Simple, robust, and accurate F and t tests in cointegrated systems," *Econometric Theory*, 34, 949–984.

Kheifets, I. and P. C. B. Phillips (2024): "High-Dimensional IV Multicointegration Estimation and Inference," *Working Paper, Yale University.*

Müller, U. K. and M. W. Watson (2018): "Long-Run Covariability," *Econometrica*, 86, 775–804.

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

——— (2014): "Optimal estimation of cointegrated systems with irrelevant instruments," *Journal of Econometrics*, 178, 210–224.

Phillips, P. C. B. and I. L. Kheifets (2024): "High-dimensional IV cointegration estimation and inference," *Journal of Econometrics*, 238, 105622.

Sun, Y. (2023): "Some Extensions of Asymptotic F and t Theory in Nonstationary Regressions," *Advances in Econometrics*, 45A, 319–347.