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

March 27, 2025

Abstract

This paper 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 and t 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 instrumental variable estimation.

JEL Codes: C12, C13, C22

1 Introduction

Phillips and Kheifets (2024, hereafter, PK(2024)) recently developed a trend instrumental variable approach for estimating cointegrated systems, allowing both singular cointegration (multicointegration (Granger and Lee, 1990)) and nonsingular cointegration. That research considered both large-K asymptotics where the number of instruments K grows to infinity with the sample

*All authors contributed equally to this work. For helpful comments and suggestions, we thank the editor, Brendan Kline, and the two anonymous referees. Email: yisun@ucsd.edu; peter.phillips@yale.edu; igor.kheifets@charlotte.edu. Phillips acknowledges research support from the Kelly Fund at the University of Auckland and a KLC Fellowship at Singapore Management University. Address formal correspondence to: Yixiao Sun, Department of Economics, UC San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0508.

size n, as in earlier work on cointegration (Phillips, 2014), and fixed-K asymptotics where K is held fixed. PK (2024) showed that, in the case of multicointegration and under fixed-K asymptotics, the asymptotic mixed normality of the trend instrumental variable estimator holds if there is no long-run correlation between the multicointegrated equilibrium error and the integrated regressor process. In this paper, as a first contribution, we show that asymptotic mixed normality still holds under fixed-K asymptotics even if there is long-run endogeneity in a multicointegrated system. Hence, asymptotic pivotal inference under fixed-K asymptotics is possible regardless of the presence of long-run endogeneity.

As a second contribution, we develop asymptotic F and t limit theory under fixed-K asymptotics for a multicointegrated system. Our theory complements the results in Hwang and Sun (2017) where similar F and t limit theory was established for the cointegration regression in the absence of multicointegration. The asymptotic F and t limit theory is very convenient to use, as critical values are readily available from statistical software packages and do not require simulations. It can also deliver more reliable inference in practical applications, as shown in other related work (Hwang and Sun, 2017; Müller and Watson, 2018).

The asymptotic mixed normality of the trend IV estimator under both fixed-K asymptotics and large-K asymptotics simplifies the study of efficiency issues. As the third contribution, we compare the conditional asymptotic variances for different values of K for a cointegration system with or without multicointegration. We find that the efficiency gain from increasing K beyond a certain threshold becomes less significant. This, together with the fact that a finite number of instruments is always employed in practical work, provides justification for using fixed-K asymptotics. On the other hand, the fixed-K asymptotic distribution approaches the large-K asymptotic distribution as K increases to infinity. Consequently, fixed-K asymptotic critical values are also asymptotically valid under large-K asymptotics. In this sense, critical values from the F or t distribution are robust with respect to the value of K.

The remainder of this paper is organized as follows. Section 2 establishes the asymptotic mixed normality under fixed-K asymptotics for a cointegration regression, with or without multicointegration. Section 3 evaluates the asymptotic relative efficiency of the trend IV estimator as K varies. Section 4 presents concluding remarks. Proofs are given in the appendix.

2 Asymptotic mixed normality & asymptotic F and t tests

As is common in much of the applied work, we consider the following standard cointegration model:

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

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:

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

In this paper, we study the estimation of the cointegrating vector a_0 in the presence of endogeneity, that is, $\{\Delta x_t\}$ may be correlated with $\{u_{0t}\}$. 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 that the initial values are zero: $x_0 = 0$ and $e_0 = 0$. In practice, nonzero initial values can be accommodated by including an intercept in the regression, as discussed at the end of this section.

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

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 $\Omega_{00\cdot x} = 0$,

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

where $\mathbf{0}$ is a $d_x \times 1$ 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 integrated and augmented model[1]

$$Y = [X, C]\,\gamma_0 + e^{+}, \text{ for } \gamma_0 := (a_0', b_0')', \tag{4}$$

where $b_0 = (f_0', g_0')'$,

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

$$C := \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ \Delta x_1 & \Delta x_2 & \cdots & \Delta x_n \end{bmatrix}' = \begin{bmatrix} x' \\ \Delta x' \end{bmatrix}',$$

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 $P_{\Phi_K} = \Phi_K (\Phi_K'\Phi_K)^{-1}\Phi_K'$.

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

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

where

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

[1]The OLS estimator of the parameters in the regression: $Y_t = X_t'a_0 + x_t'f_0 + \tilde{e}_t$, is referred to as the Integrated Modified OLS (IM-OLS) estimator by Vogelsang and Wagner (2014), who considered only the nonsingular case when $\Omega_{00\cdot x} > 0$ and showed that this estimator is asymptotically mixed normal. Phillips and Kheifets (2021) established the asymptotic distribution of the IM-OLS estimator when $\Omega_{00\cdot x} = 0$ and demonstrated that singularity introduces nonstandard asymptotics. As we show below, the application of the trend IV method to the model in equation (4) restores asymptotic mixed normality. Interestingly, applying the trend IV method to the regression $Y_t = X_t'a_0 + x_t'f_0 + \tilde{e}_t$ does not yield an asymptotically mixed normal estimator either, as shown in Kheifets and Phillips (2024). Therefore, to achieve asymptotic mixed normality in the singular case, we have to include both x_t and Δx_t in the regression and apply the trend IV method.

Alternatively, and equivalently,

$$(\hat{a}_{\text{TIV}}, \hat{b}_{\text{TIV}}) = \arg\min_{(a,b)} (Y - Xa - Cb)' \, P_{\Phi_K} (Y - Xa - Cb) \, .$$

For orthonormal basis functions $\{\varphi_j(r)\}_{j=1}^{\infty}$, we have $\left\|n^{-1}\Phi_K'\Phi_K - I_K\right\|_2 = o(1)$ 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

$$V_Y = V_X a_0 + V_x f_0 + V_{\Delta x} g_0 + V_{e^+} = V_X a_0 + V_C b_0 + V_{e^+}, \tag{7}$$

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 $\hat{b}_{\text{fTIV}}$ of b_0 as

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

The estimators $\hat{a}_{\text{fTIV}}$ and $\hat{b}_{\text{fTIV}}$ 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 paper, 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 paper $\hat{a}_{\text{fTIV}}$ is the transformed and augmented OLS estimator 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 size n 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 and orthonormal basis functions on $L_2[0,1]$.*

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

$$\begin{aligned} B_X(r) &= \int_0^r B_x(s)\,ds, \;\; \mu_K = \int_0^1 \tilde{\varphi}_K(r)\, B_X(r)'\,dr, \;\; \eta_K = \int_0^1 \tilde{\varphi}_K(r)\, B_x(r)'\,dr, \\ \xi_K &= \int_0^1 \tilde{\varphi}_K(r)\, dB_x(r)', \;\; \psi_{0\cdot x,K} = \int_0^1 \tilde{\varphi}_K(r)\, B_{0\cdot x}(r)\,dr, \end{aligned}$$

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

Note that μ_K, η_K, ξ_K and S_K are $K \times d_x$ matrices, $\psi_{0\cdot x,K}$ and $\psi_{e\cdot x,K}$ are $K \times 1$ vectors, and $J_K = Q_{[\xi_K,\eta_K]}$ is a $K \times K$ projection matrix, which projects onto the orthogonal complement of the space spanned by the columns of $[\xi_K, \eta_K]$.

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 $n \to \infty$,*

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

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

$$
n^2\left(\hat{a}_{\mathrm{fTIV}} - a_0\right) \rightsquigarrow S_K'\psi_{e\cdot x,K} \equiv \mathcal{MN}\left(0, \omega_{ee\cdot x}\left(\mu_K' J_K \mu_K\right)^{-1}\right). \tag{10}
$$

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 Theorem 1(ii) of 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, serving as an addendum to Theorem 1(ii) of PK (2024).

The zero-mean mixed normal asymptotic distribution enables the construction of asymptotically pivotal tests about cointegration parameters. Consider, for example, the case of multicointegration.[2] 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:

$$
\hat{V}_{e^+} = V_Y - V_X\hat{a}_{\mathrm{fTIV}} - V_C\hat{b}_{\mathrm{fTIV}},
$$

and then compute the average of the squared residuals:

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

[2] Asymptotic F and t tests can also be developed for the case of nonsingular cointegration within the present framework; but we focus on multicointegration here.

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

$$\mathbb{W}_{\text{fTIV}} = \frac{1}{\hat{\omega}_{ee\cdot x}} \left[H\hat{a}_{\text{fTIV}} - h\right]' \left[H\left(V_X' Q_{V_C} V_X\right)^{-1} H'\right]^{-1} \left[H\hat{a}_{\text{fTIV}} - h\right]/p. \tag{11}$$

When $p = 1$, we may calculate the t statistic

$$\mathbb{T}_{\text{fTIV}} = \frac{H\hat{a}_{\text{fTIV}} - h}{\sqrt{\hat{\omega}_{ee\cdot x} H\left(V_X' Q_{V_C} V_X\right)^{-1} H'}}. \tag{12}$$

Theorem 2 ***(Asymptotic F and t tests with fTIV)*** *Let Assumptions 1 and 2 hold. Consider the case of multicointegration with $\Omega_{00\cdot x} = 0$. If $K > 3d_x$ and H has full row rank p, then*

$$\mathbb{W}^*_{\text{fTIV}} \ := \frac{K - 3d_x}{K}\mathbb{W}_{\text{fTIV}} \rightsquigarrow F_{p,K-3d_x}, \tag{13}$$

$$\mathbb{T}^*_{\text{fTIV}} \ := \sqrt{\frac{K - 3d_x}{K}}\mathbb{T}_{\text{fTIV}} \rightsquigarrow t_{K-3d_x}, \tag{14}$$

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

Theorem 2 is new and extends the corresponding result in Hwang and Sun (2018), which rules out multicointegration, to the case of multicointegration. 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 b_0 and $\omega_{ee\cdot x}$, which are often ignored in the use of fully modified methodology. For the corresponding 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 an intercept λ_0 is included in (1), we have the model

$$y_t = x_t' a_0 + \lambda_0 + u_{0t}.$$

The integrated and augmented model is then given by

$$Y_t = X_t' a_0 + x_t' b_0 + \Delta x_t' g_0 + \mu_0 + \lambda_0 t + e^+_{0\cdot x,t},$$

where μ_0 is a constant that captures the effects of the initial values x_0 and e_0. In matrix form the equations corresponding to (4) and (7) take the same forms, viz.,

$$Y = [X, C]\gamma_0 + e^+ = Xa_0 + Cb_0 + e^+, \tag{15}$$

$$V_Y = [X, C]\gamma_0 + V_{e^+} = V_X a_0 + V_C b_0 + V_{e^+}, \tag{16}$$

but now $\gamma_0 = (a_0', b_0')'$ for $b_0 = (f_0', g_0', \mu_0, \lambda_0)'$ and

$$C = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ \Delta x_1 & \Delta x_2 & \cdots & \Delta x_n \\ 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & n \end{bmatrix}'.$$

With these obvious modifications, the fTIV estimator $\hat{a}_{\text{fTIV}}$ of a_0 and the Wald and t statistics $\mathbb{W}_{\text{fTIV}}$ and $\mathbb{T}_{\text{fTIV}}$ can be computed in the same way as before.

The asymptotic theory for $\hat{a}_{\text{fTIV}}$ in Theorem 1 continues to hold if we redefine ξ_K as

$$\xi_K = \left[\int_0^1 \tilde{\varphi}_K(r)\, dB_x(r)', \int_0^1 \tilde{\varphi}_K(r)\, dr, \int_0^1 \tilde{\varphi}_K(r)\, r dr\right] \in \mathbb{R}^{K\times(d_x+2)},$$

which contains two additional columns, reflecting the inclusion of an intercept and a linear trend in (15) and (16). Further, the asymptotic F and t test theory in Theorem 2 remains valid after a simple adjustment. We have, in place of (13) and (14),

$$\mathbb{W}^*_{\text{fTIV}} := \frac{K-(3d_x+2)}{K}\mathbb{W}_{\text{fTIV}} \rightsquigarrow F_{p,K-(3d_x+2)}, \tag{17}$$

$$\mathbb{T}^*_{\text{fTIV}} := \sqrt{\frac{K-(3d_x+2)}{K}}\mathbb{T}_{\text{fTIV}} \rightsquigarrow t_{K-(3d_x+2)}, \tag{18}$$

provided that $K > 3d_x + 2$ and $[\mu_K, \eta_K, \xi_K]$ has full column rank of $3d_x + 2$. This adjustment occurs because there are two additional regressors, resulting in a loss of two degrees of freedom. If the basis functions integrate to zero (i.e., $\int_0^1 \tilde{\varphi}_K(r)\, dr = 0$) and $[\mu_K, \eta_K, \xi_K]$ has a column rank of $3d_x + 1$, then the results in (17) and (18) hold with $3d_x + 2$ replaced by $3d_x + 1$. Deterministic trends can also be included in (1) or (4), and the asymptotic F and t limit theory remains valid, albeit with a different multiplicative adjustment factor and different degrees of freedom for the F and t distributions (see, e.g., Section 4 in Hwang and Sun (2018)).

3 Asymptotic relative efficiency

We now compare the asymptotic distributions of the fTIV and TIV estimators under Assumption 1. For the large-K asymptotic results in this section, we follow PK (2024) and consider the basis functions $\varphi_j(r) = \sqrt{2}\sin((j-1/2)\pi r)$ for $j = 1, 2, \ldots, K$.[3] Theorem 2 of PK (2024) showed that, 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 $K = o\left(n^{4/5-\delta}\right)$ for some $\delta > 0$, the following holds:

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

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

$$n(\hat{a}_{\text{TIV}} - a_0) \quad \rightsquigarrow \quad \mathcal{A}^{-1}_{X\cdot x}\left(\int_0^1 B_{X\cdot x}(r)\, B_{0\cdot x}(r)\, dr\right)$$

[3]The large-K asymptotic results can be extended to allow for any complete set of orthonormal basis functions on $L_2[0,1]$ that are twice continuously differentiable, cf., Phillips (2005, pp.126).

$$\equiv \quad \mathcal{MN}\left(0, \Omega_{00\cdot x}\mathcal{A}_{X\cdot x}^{-1}\int_0^1\int_0^1 (r\wedge s)\, B_{X\cdot x}(r)\, B_{X\cdot x}(s)'\, drds\mathcal{A}_{X\cdot x}^{-1}\right). \quad (19)$$

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 $\Omega_{00\cdot x}=0$,

$$n^2\left(\hat{a}_{\mathrm{TIV}}-a_0\right) \rightsquigarrow \mathcal{A}_{X\cdot x}^{-1}\left(\int_0^1 B_{X\cdot x}(r)\, dB_{e\cdot x}(r)\right) \equiv \mathcal{MN}\left(0, \omega_{ee\cdot x}\mathcal{A}_{X\cdot x}^{-1}\right) \quad (20)$$

under the joint large-K asymptotics specified above.

Following arguments similar to those in PK (2024), we can show that $\hat{a}_{\mathrm{fTIV}}$ and $\hat{a}_{\mathrm{TIV}}$ share the same joint large-K asymptotic distributions, as given in (19) and (20), respectively, for the cointegration and multicointegration cases. 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 (19) and (20), respectively. So the critical values obtained from fixed-K approximations are asymptotically valid under large-K asymptotics.

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 impact of 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.

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 $K\to\infty$ but $K=o\left(n^{4/5-\delta}\right)$). Because the asymptotic distributions are all mixed normal, it is simplest to compare conditional 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), (19) or (20), 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 ($K \to \infty$).

	Cointegration		Multicointegration	
	$K = 7$	$K = 12$	$K = 7$	$K = 12$
mean	1.171	1.071	1.282	1.129
std	0.360	0.134	0.330	0.119
min	0.584	0.682	1.004	1.004
25%	1.029	1.014	1.097	1.053
50%	1.074	1.036	1.182	1.094
75%	1.184	1.083	1.344	1.164
max	9.038	3.875	6.453	3.050

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 ($K \to \infty$).

	Cointegration		Multicointegration	
Coverage	$K = 7$	$K = 12$	$K = 7$	$K = 12$
0.99	1.274	1.123	1.493	1.174
0.95	1.244	1.129	1.348	1.134
0.90	1.187	1.094	1.327	1.145

4 Final remarks

The primary advantage of the high-dimensional trend IV approach to estimation and inference 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 specific asymptotics that hold the number of instruments fixed may provide more reliable distributional approximations. Under this approach to the limit theory, fTIV delivers an asymptotically valid, easy-to-use, and more accurate F and t tests while retaining the nonparametric advantage of TIV.

Appendix of proofs

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),

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

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 $\mathcal{N}\left(0, \Omega_{00\cdot x}\left(\int_0^1\int_0^1 (r \wedge s)\, \tilde{\varphi}_K(r)\, \tilde{\varphi}_K(s)'\, dr ds\right)\right)$. 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 $\mathcal{N}(0,\omega_{ee\cdot x}\int_0^1\tilde{\varphi}_K(r)\,\tilde{\varphi}_K'(r)\,dr)$. Due to the orthonormality of the basis functions ($\int_0^1\tilde{\varphi}_K(r)\,\tilde{\varphi}_K'(r)\,dr=I_K$), the limit distribution is mixed normal with the conditional variance matrix simplified from a sandwich form to a single matrix component:

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

■

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:

$$\begin{aligned} \mathbb{W}_{\text{fTIV}} &= \frac{\left[H\hat{a}_{\text{fTIV}}-h\right]'\left[H\left(V_X'Q_{V_C}V_X\right)^{-1}H'\right]^{-1}\left[H\hat{a}_{\text{fTIV}}-h\right]/p}{V_{e^+}'Q_{[V_X,V_C]}V_{e^+}/K} \\ &= \frac{\left[\hat{a}_{\text{fTIV}}-a_0\right]'H'\left[H\left(V_X'Q_{V_C}V_X\right)^{-1}H'\right]^{-1}H\left[\hat{a}_{\text{fTIV}}-a_0\right]/p}{V_{e^+}'Q_{[V_X,V_C]}V_{e^+}/K} \\ &\rightsquigarrow \frac{\psi_{e\cdot x,K}'S_KH'\left[H\left(\mu_K'J_K\mu_K\right)^{-1}H'\right]^{-1}HS_K'\psi_{e\cdot x,K}/p}{\psi_{e\cdot x,K}'Q_{[\mu_K,\eta_K,\xi_K]}\psi_{e\cdot x,K}/K} \\ &= \frac{\left\|P_{[S_KH']}\psi_{e\cdot x,K}\right\|^2/p}{\left\|Q_{[\mu_K,\eta_K,\xi_K]}\psi_{e\cdot x,K}\right\|^2/K}. \end{aligned} \tag{21}$$

Under the assumption that $\left\{\varphi_j(\cdot)\right\}_{j=1}^K$ 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),

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

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 $HS_K'\psi_{e\cdot x,K}$. The conditional independence between these two normals holds because, conditional on (μ_K,η_K,ξ_K), we have

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

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

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

and

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

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:

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

The result for the t-statistic can be proved similarly, and the details are omitted. ■

References

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.

Hwang, J. and Sun, Y. (2017). Asymptotic F and t tests in an efficient gmm setting. *Journal of Econometrics*, 198(2):277–295.

Hwang, J. and Sun, Y. (2018). Simple, robust, and accurate F and t tests in cointegrated systems. *Econometric Theory*, 34:949–984.

Kheifets, I. L. and Phillips, P. C. B. (2024). High-dimensional IV multicointegration estimation and inference. *Working Paper, Yale University.*

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

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

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

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

Phillips, P. C. B. and Kheifets, I. L. (2021). On multicointegration. *Cowles Foundation Discussion Paper, Yale University.*

Phillips, P. C. B. and Kheifets, I. L. (2024). High-dimensional IV cointegration estimation and inference. *Journal of Econometrics*, 238(2):105622.

Sun, Y. (2023). Some extensions of asymptotic F and t theory in nonstationary regressions. *Advances in Econometrics*, 45A:319–347.

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.