Journal of Economic Theory and Econometrics, Vol. 33, No. 4, Dec. 2022, 54–76

Note on Testing for Linear Trends in Cointegrating Regressions

Cheol-Keun Cho*

Abstract In this study, I address the testing problem on the regression trend slope in cointegrating regressions when the stochastic regressors have nonzero drifts. A test statistic constructed using demeaned integrated modified ordinary least squares (IMOLS) residuals is considered. Asymptotic theory for the test is developed under the standard small-*b* framework, resorting to the consistency of heteroskedasticity and autocorrelation consistent (HAC) estimator. The simulation experiment shows the proposed test performs reasonably compared to the existing fully modified OLS-based test in Hansen (1992b).

Keywords Cointegration, Drift, HAC, IMOLS

JEL Classification C12, C22

*Department of Economics, University of Ulsan, E-mail: chkcho@gmail.com. The author gratefully acknowledges financial support from the University of Ulsan under Grant No. 2020-0945. I thank the editor, and the two anonymous referees for helpful and constructive comments.

Submitted September 5, 2022, Accepted November 30, 2022

1. INTRODUCTION

In this study, I address the trend coefficient testing problem in cointegrating regressions when $I(1)$ regressors have nonzero drifts (i.e. contain deterministic linear trends). Testing for the regression trend slope is important to learn about the nature of the cointegrating relationships. If the regression trend coefficient is zero, the $I(1)$ series are 'deterministically cointegrated', in which the linear combination with a cointegrating vector eliminates the deterministic trends as well as the stochastic trends. Otherwise, the $I(1)$ series are only stochastically cointegrated and the cointegrating relationship involves a deterministic trend.[1] See Ogaki, 1992; Ogaki and Park, 1997; Han and Ogaki, 1997; Kakkar and Ogaki, 1999; Wagner, 2015; and Mikayilov *et al.*, 2018 for related empirical applications.

I develop a test on the trend slope in the integrated modified ordinary least squares (IMOLS) estimation (Vogelsang and Wagner, 2014) framework, using an IMOLS residual that is obtained by plugging in the IMOLS estimator in the original regression equation. However, I find that this residual must be demeaned to apply it to construct a consistent long-run variance estimator. Without demeaning, the resulting heteroskedasticity and autocorrelation consistent (HAC) estimator is not consistent and the associated test is not appropriate for statistical inference unless the regressors have no drift. The performance of the proposed IMOLS-based test is evaluated via a simulation experiment. The simulation result shows that the IMOLS-based test has smaller size distortions with a moderate loss of power, compared to the fully modified OLS-based test.

The remainder of the paper is organized as follows. Section 2 presents the model and provides a review of the IMOLS estimator. Section 3 develops the asymptotic theory for the test. Section 4 evaluates the performance of the test in finite samples via a Monte Carlo simulation, and Section 5 provides an empirical application. Section 6 concludes the paper.

[1] See Ogaki and Park, 1997; Perron and Campbell, 1993; Campbell and Perron, 1991; and Perron and Rodríguez, 2016 for further discussion.

2. MODEL SETUP AND PRELIMINARY RESULTS

2.1. MODEL AND ASSUMPTION

Suppose that y_t (scalar valued) and x_t (k dimensional) are $I(1)$ series with drifts:

$$y_t = \alpha_y + \delta_y t + y_t^0, \qquad (1)$$
$$x_t = \alpha_x + \delta_x t + x_t^0,$$

where $\delta_x \neq 0$, $y_t^0 = y_{t-1}^0 + v_t^y$, and $x_t^0 = x_{t-1}^0 + v_t$ with v_t^y and v_t being $I(0)$. I consider the unrestricted regression equation:

$$y_t = \delta_0 + \delta_1 t + x_t'\beta + u_t, \qquad (2)$$

where $\delta_0 = \alpha_y - \beta'\alpha_x$, $\delta_1 = \delta_y - \beta'\delta_x$, and $u_t = y_t^0 - \beta' x_t^0 \sim I(0)$. The null hypothesis of interest is $H_0 : \delta_1 = 0$ (deterministic cointegration) and the alternative hypothesis is $H_1 : \delta_1 \neq 0$ (stochastic cointegration).

Assumption 1.
Let $\eta_t = (u_t, v_t')'$, and assume a functional central limit theorem of the form:

$$T^{-1/2}\sum_{t=1}^{[rT]} \eta_t \Rightarrow B(r) = \begin{pmatrix} B_u(r) \\ B_v(r) \end{pmatrix} = \Omega^{1/2} W(r), \quad r \in [0,1],$$

where $W(r) = (w_{uv}(r), w_v'(r))'$ is $(k+1)$-dimensional standard Brownian motion with

$$\Omega = \lim_{T\to\infty}\left(T^{-1} var\left(\sum_{t=1}^{T} \eta_t\right)\right) = \begin{bmatrix} \Omega_{uu} & \Omega_{uv} \\ \Omega_{vu} & \Omega_{vv} \end{bmatrix} > 0. \qquad (3)$$

Under Assumption 1, y_t and x_t are cointegrated up to the deterministic trend. Following the literature, the Cholesky form of $\Omega^{1/2}$ is used:

$$\Omega^{1/2} = \begin{bmatrix} \sigma_{uv} & \lambda_{uv} \\ \mathbf{0} & \Omega_{vv}^{1/2} \end{bmatrix}, \text{ where } \sigma_{uv}^2 = \Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}. \qquad (4)$$

Note that $w_{uv}(\cdot)$ represents the part of $B_u(\cdot)$ that is independent of $B_v(\cdot) = \Omega_{vv}^{1/2} w_v(\cdot)$.

Define the one-sided long-run covariance matrix

$$\Lambda = \sum_{j=1}^{\infty} E\left(\eta_{t-j}\eta_t'\right) = \begin{bmatrix} \Lambda_{uu} & \Lambda_{uv} \\ \Lambda_{vu} & \Lambda_{vv} \end{bmatrix} \text{ and } \Sigma = E\left(\eta_t \eta_t'\right) = \begin{bmatrix} \Sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{bmatrix}. \qquad (5)$$

Note that $\Omega = \Sigma + \Lambda + \Lambda'$. Further, define $\Delta \equiv \Sigma + \Lambda$ and denote $\Delta = \begin{bmatrix} \Delta_{uu} & \Delta_{uv} \\ \Delta_{vu} & \Delta_{vv} \end{bmatrix}$.

2.2. IMOLS ESTIMATOR

Vogelsang and Wagner (2014) (hereafter, VW (2014)) proposed the IMOLS estimator, which is the OLS estimator in the following partial-summed and augmented regression:

$$S_t^y = S_t^{f\prime}\delta + S_t^{x\prime}\beta + x_t^\prime\gamma + S_t^u, \tag{6}$$

with $\delta = (\delta_0, \delta_1)^\prime$, $S_t^y = \sum_{j=1}^t y_j$, $S_t^f = \sum_{j=1}^t f_j = \sum_{j=1}^t (1, j)^\prime$, $S_t^x = \sum_{j=1}^t x_j$, and $S_t^u = \sum_{j=1}^t u_j$. In matrix form,

$$S^y = S^{\tilde{x}}\theta + S^u, \tag{7}$$

where $S^{\tilde{x}} = \left(S^f \vdots S^x \vdots X\right)$ and $\theta = (\delta^\prime, \beta^\prime, \gamma^\prime)^\prime$ with $S^f = \left(S_1^f, \ldots S_T^f\right)^\prime$, $S^x = (S_1^x, \ldots S_T^x)^\prime$, and $X = (x_1, \ldots x_T)^\prime$. The IMOLS estimator of θ is given by

$$\widehat{\theta} = \left(\widehat{\delta}^\prime, \widehat{\beta}^\prime, \widehat{\gamma}^\prime\right)^\prime = \left(S^{\tilde{x}\prime}S^{\tilde{x}}\right)^{-1} S^{\tilde{x}\prime}S^y. \tag{8}$$

VW (2014) derive the following limit result for the case of $\delta_x = \mathbf{0}$:

$$A_{IM}^{-1}\begin{pmatrix} \hat{\delta}-\delta \\ \hat{\beta}-\beta \\ \hat{\gamma}-\Omega_{vv}^{-1}\Omega_{vu} \end{pmatrix} = \begin{pmatrix} T^{1/2}\tau_F(\hat{\delta}-\delta) \\ T\left(\hat{\beta}-\beta\right) \\ \left(\hat{\gamma}-\Omega_{vv}^{-1}\Omega_{vu}\right) \end{pmatrix} \xrightarrow{d} \sigma_{uv}\left(\Pi\int g(s)g(s)^\prime ds\Pi^\prime\right)^{-1}\Pi\int g(s)w_{uv}(s)ds, \tag{9}$$

where

$$\Pi = \begin{pmatrix} I_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Omega_{vv}^{1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Omega_{vv}^{1/2} \end{pmatrix},\ g(r) = \begin{pmatrix} \int_0^r f(s)ds \\ \int_0^r w_v(s)ds \\ w_v(r) \end{pmatrix},\ A_{IM} = \begin{pmatrix} T^{-1/2}\tau_F^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & T^{-1}I_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_k \end{pmatrix}, \tag{10}$$

and τ_F is a diagonal matrix with diagonal elements 1, T satisfying

$$T^{-1}\tau_F^{-1}\sum_{t=1}^{[rT]} f_t \to \int_0^r f(s)ds,\ r\in[0,1] \text{ as } T \text{ grows with } f(s) = (1,s)^\prime. \tag{11}$$

When $\delta_x \neq \mathbf{0}$, the limit result in (9) is still valid for $\hat{\beta}$ and $\hat{\gamma}$ but not for $\hat{\delta}$. The fully valid limit result for the present case of nonzero drifts (δ_x) in the regressors is derived in Cho (2022a). Define $\theta_* = \left(\delta^\prime, \beta^\prime, \Omega_{vu}^\prime\Omega_{vv}^{-1}\right)^\prime$ and $G(r) \equiv \int_0^r g(s)ds$. Cho (2022a) shows that

$$\left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta}-\theta_{*}\right)+\begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}=\left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta}-\theta_{**}\right) \quad (12)$$

$$\xrightarrow{d}\sigma_{uv}\Pi^{-1}\left(\int_0^1 g(s)g(s)'ds\right)^{-1}\int_0^1 (G(1)-G(s))\,dw_{uv}(s):=A^{\infty}=\Big(\underset{1\times 2}{A_{\delta}^{\infty\prime}},\ \underset{1\times k}{A_{\beta}^{\infty\prime}},\underset{1\times k}{A_{\gamma}^{\infty\prime}}\Big)',$$

where

$$\theta_{**}\equiv\theta_{*}-A_{IM}^{D\prime}\begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}=\begin{pmatrix} \delta-D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \beta \\ \Omega_{vv}^{-1}\Omega_{vu} \end{pmatrix},$$

and

$$A_{IM}^{D}=\begin{pmatrix} T^{-1/2}\tau_F^{-1} & \mathbf{0} & \mathbf{0} \\ -T^{-1}D_x & T^{-1}I_k & \mathbf{0} \\ -D_x^F & \mathbf{0} & I_k \end{pmatrix}, \quad (13)$$

with $D_x\equiv\left(\mathbf{0}\,\vdots\,\delta_x\right)$ and $D_x^F\equiv\left(\delta_x\,\vdots\,\mathbf{0}\right)$, both being $k\times 2$ matrices.[2]

Conditional on $w_v(\cdot)$, the variance of A^{∞} in (12) is given by $\sigma_{uv}^2\Pi^{-1}V_{IM}^{\Xi}\Pi^{-1}:=\sigma_{uv}^2V_{IM}^{o}$, with

$$V_{IM}^{\Xi}\equiv\left(\int_0^1 g(s)g(s)'ds\right)^{-1}\int_0^1 (G(1)-G(s))\,(G(1)-G(s))'\,ds\left(\int_0^1 g(s)g(s)'ds\right)^{-1}, \quad (14)$$

and also conditional on $w_v(\cdot)$, $var\left(A_{\beta}^{\infty}\right)=\sigma_{uv}^2\left(E\cdot V_{IM}^{o}\cdot E'\right):=V_{IM}^{\beta}$, with $E\equiv\left[\mathbf{0}_{k\times 2}\ I_k\ \mathbf{0}_{k\times k}\right]$.

Building on the result in (12), the asymptotic result that can be directly applicable for the inference on δ_1 can be derived (see Cho (2022b)):

$$T\left(\widehat{\delta}_1-\delta_1\right)\xrightarrow{d}-\delta_x'A_{\beta}^{\infty}, \quad (15)$$

where A_{β}^{∞} is the segment of A^{∞} that corresponds to $\widehat{\beta}$ defined in (12).

---

[2] $\left(A_{IM}^{D\prime}\right)^{-1}=\begin{pmatrix} T^{1/2}\tau_F & T^{1/2}\tau_F D_x' & T^{1/2}\tau_F D_x^{F\prime} \\ \mathbf{0} & TI_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_k \end{pmatrix}$

3. INFERENCE FOR THE TREND COEFFICIENT

3.1. IMOLS-BASED TEST

The limit distribution in (15) provides a basis for constructing a test statistic, which is given by

$$t_{\delta_1}^{IM} = \frac{T\widehat{\delta}_1}{\sqrt{\widehat{\delta}_x' \widehat{V}_{IM}^{\beta} \widehat{\delta}_x}}, \tag{16}$$

where $\widehat{\delta}_x = \frac{1}{T-1}\sum_{t=2}^{T} \triangle x_t$ and $\widehat{V}_{IM}^{\beta}$ denotes a consistent estimator of $V_{IM}^{\beta} = \sigma_{uv}^2 \left(E \cdot V_{IM}^{o} \cdot E'\right)$ with $E \equiv [\mathbf{0}_{k\times 2}\ I_k\ \mathbf{0}_{k\times k}]$. Thus, an estimator $\widehat{V}_{IM}^{\beta}$ is composed of two parts: estimators of σ_{uv}^2 and $E \cdot V_{IM}^{o} \cdot E'$.

As shown in VW (2014) $\check{V}_{IM}^{o}$ is consistent for V_{IM}^{o}, and Cho (2022a) shows that $E \cdot \check{V}_{IM}^{o} \cdot E'$ is a consistent estimator for $E \cdot V_{IM}^{o} \cdot E'$ even in the present case $\delta_x \neq \mathbf{0}$:

$$\check{V}_{IM}^{o} = A_{IM}^{-1} \left(S^{\widetilde{x}\prime} S^{\widetilde{x}}\right)^{-1} \left(C'C\right) \left(S^{\widetilde{x}\prime} S^{\widetilde{x}}\right)^{-1} A_{IM}^{-1},$$

where $C = (c_1, c_2, \ldots, c_T)'$, with $c_t = \sum_{j=1}^{T} S_j^{\widetilde{x}} - \sum_{j=1}^{t-1} S_j^{\widetilde{x}}$ and $S_j^{\widetilde{x}\prime} =$ the j^{th} row in $S^{\widetilde{x}}$.

To estimate σ_{uv}^2, one might consider using $\widehat{u}_t \equiv y_t - f_t'\widehat{\delta} - x_t'\widehat{\beta}$, which is obtained by plugging in the IMOLS estimator in the original regression equation (2). It will be shown that this residual contains terms involving the nonzero drift (δ_x), and it does not deliver a consistent estimator of the long-run variance (Corollary 1). However, the demeaned residual leads to a consistent long-run variance estimator and a valid statistical inference with the standard normal approximation. The demeaned residual is given by

$$\widehat{u}_t^d \equiv \widehat{u}_t - \frac{1}{T}\sum_{j=1}^{T} \widehat{u}_j,$$

and $\triangle \ddot{x}_t \equiv \triangle x_t - \overline{\triangle x} = (v_t + \delta_x) - \frac{1}{T-1}\sum_{t=2}^{T}(v_t + \delta_x) = v_t - \frac{1}{T-1}\sum_{t=2}^{T} v_t = v_t^d$ with $\overline{\triangle x} = \frac{1}{T-1}\sum_{t=2}^{T} \triangle x_t$. Define $\widehat{\eta}_t \equiv \left(\widehat{u}_t^d,\ \triangle \ddot{x}_t'\right)'$. The nonparametric kernel-based HAC estimator for the long-run variance Ω is given by

$$\widehat{\Omega} = \frac{1}{T}\sum_{i=1}^{T}\sum_{j=1}^{T} k\left(\frac{|i-j|}{M}\right)\widehat{\eta}_i \widehat{\eta}_j' := \begin{bmatrix} \widehat{\Omega}_{uu} & \widehat{\Omega}_{uv} \\ \widehat{\Omega}_{vu} & \widehat{\Omega}_{vv} \end{bmatrix},$$

where $k(x)$ is a kernel function, for example, the Bartlett kernel: $k(x) = (1-|x|) \cdot \mathbf{1}(|x| \leq 1)$.

Provided that $\widehat{\Omega}$ is consistent for Ω, $\widehat{\sigma}_{uv}^2 \equiv \widehat{\Omega}_{uu} - \widehat{\Omega}_{uv}\widehat{\Omega}_{vv}^{-1}\widehat{\Omega}_{vu}$ and $\widehat{V}_{IM}^{\beta} = \widehat{\sigma}_{uv}^2 \left(E \cdot \check{V}_{IM}^{o} \cdot E'\right)$ are consistent for σ_{uv}^2 and V_{IM}^{β}, respectively. Upon this consistency property, by the standard conditioning argument, $t_{\delta_1}^{IM}$ converges in distribution to $N(0,1)$ under H_0. For the consistency of the long-run variance estimator, some regularity conditions for $\eta_t \equiv (u_t,\, v_t')'$ are required. In addition, for the case where $(u_t,\, v_t')'$ needs to be estimated, another set of sufficient conditions such as those found in Hansen (1992a) may be invoked.

First, different sets of regularity conditions for consistency of HAC estimators are provided in the literature (see Newey and West, 1987; Andrews, 1991; Hansen, 1992a; and Jansson, 2002 among others). The unobserved series $\{\eta_t\}$ should be assumed to meet at least one of those conditions. In the present paper, I assume that the conditions (K), (S), and (V1) in Hansen (1992a) hold for $\{\eta_t\}$:

Let $\{\alpha_m\}_{m=1}^{\infty}$ denote the α-mixing coefficient for $\{\eta_t\}$.[3]

(K) For all $x \in \mathbb{R}$, $|k(x)| \leq 1$ with $k(x) = k(-x)$ and $k(0) = 1$; $k(x)$ is continuous at zero and for almost all x; $\int_{\mathbb{R}} |k(x)|\, dx < \infty$.

(S) $M \to \infty$ as T grows, and for some $q \in (1/2, \infty)$, $M^{1+2q}/T = O_p(1)$.

(V1) For some $r \in (2,4]$ such that $r > 2 + 1/q$, and some $p > r$,

$$\text{(V1-i) } 12\textstyle\sum_{m=1}^{\infty} \alpha_m^{2(1/r-1/p)} < \infty; \quad \text{(V1-ii) } \sup_{t\geq 1} \|\eta_t\|_p < \infty.$$

The condition (K) is satisfied by most popular kernels including the Bartlett, and QS kernels. As documented in Hansen (1992a), the condition (S) is also satisfied for the data-driven bandwidth proposed in Andrews (1991), with different values of q depending on the kernel. For the Bartlett and QS kernel, the condition holds with $q = 1$ and $q = 2$, respectively. Note that this condition implies that $M = O_p(T^{1/3})$ for the Bartlett kernel, and $M = O_p(T^{1/5})$ for the QS kernel. (V1-i) imposes a restriction on the degree of temporal dependence for $\{\eta_t\}$, and (V1-ii) states that the L^p norm of the series is uniformly bounded, putting a restriction on the tail-thickness of the distribution.

Second, for the case where $(u_t,\, v_t')'$ needs to be estimated, Hansen (1992a) provides a set of sufficient condition (condition (V3) in his paper) that guarantees the consistency of a HAC estimator. Suppose that we construct a HAC estimator using the residuals $\mathbf{V}_t(\theta) = \mathbf{V}_t - \mathbf{X}_t'\left(\widehat{\theta} - \theta_0\right)$ to estimate the long-run variance of $\mathbf{V}_t$. Here, θ_0 represents a model parameter, $\widehat{\theta}$ is an estimator for θ_0, and $\mathbf{X}_t$ denotes a set of regressors. Let $\{\delta_T\}$ be an appropriate sequence of deterministic

[3] The condition (V1-i) may be stated in terms of the φ mixing coefficients as in Hansen (1992a).

nonsingular matrices. The set of sufficient condition in Hansen (1992a) is as follows:

$$\text{(V3-ii)} \sup_{t\leq T}\left\|\mathbf{X}_t'\delta_T'\right\| = O_p(1)\text{, and}$$

$$\text{(V3-iii)} \sqrt{T}\left(\delta_T'\right)^{-1}\left(\widehat{\theta}-\theta_0\right) = O_p(1)\,.$$

As the HAC estimator is constructed with $\widehat{\eta}_t \equiv \left(\widehat{u}_t^d,\, \triangle\ddot{x}_t'\right)'$, the sufficient condition in Hansen (1992a) should be checked for $\widehat{\eta}_t$.

Note that

$$\widehat{u}_t = y_t - f_t'\widehat{\delta} - x_t'\widehat{\beta} = u_t - \left(f_t',\, x_t',\, \mathbf{0}'\right)\left(\widehat{\theta}-\theta_*\right)$$

and

$$\widehat{u}_t^d = u_t^d - \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)\left(\widehat{\theta}-\theta_*\right) = u_t^d - \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)A_{IM}^{D\prime}\left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta}-\theta_*\right), \tag{17}$$

where $u_t^d \equiv u_t - \frac{1}{T}\sum_{j=1}^{T}u_j$, $f_t^d = f_t - \frac{1}{T}\sum_{j=1}^{T}f_j$, and $x_t^d = x_t - \frac{1}{T}\sum_{j=1}^{T}x_j$.

The proof for Theorem 1 in the Appendix shows

$$\widehat{u}_t^d = u_t^d - \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)A_{IM}^{D\prime}\left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta}-\theta_{**}\right),$$

where

$$\theta_{**} \equiv \theta_* - A_{IM}^{D\prime}\begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0}\end{pmatrix} = \begin{pmatrix} \delta - D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \beta \\ \Omega_{vv}^{-1}\Omega_{vu}\end{pmatrix}.$$

Now, rewrite $\widehat{\eta}_t$ as

$$\begin{aligned}\widehat{\eta}_t &\equiv \begin{pmatrix}\widehat{u}_t^d \\ \triangle\ddot{x}_t\end{pmatrix} = \begin{pmatrix}u_t^d \\ v_t^d\end{pmatrix} - \begin{pmatrix}\left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)\left(\widehat{\theta}-\theta_{**}\right) \\ \mathbf{0}_{k\times 1}\end{pmatrix} \\ &= \begin{pmatrix}u_t^d \\ v_t^d\end{pmatrix} - \begin{pmatrix}\left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)A_{IM}^{D\prime} \\ \mathbf{0}_{k\times(2k+2)}\end{pmatrix}\left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta}-\theta_{**}\right).\end{aligned}$$

This equation shows that $\widehat{\eta}_t$ satisfies the conditions (V3-ii) and (V3-iii) if $\widehat{u}_t^d$ does. In the proof for Theorem 1 it is shown that

$$\left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta}-\theta_{**}\right) = O_p(1)\,,$$

and

$$\sup_{1\le t\le T}\left\|\sqrt{T}\left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)A_{IM}^{D\prime}\right\| = O_p(1),$$

which verifies that $\widehat{u}_t^d$ meets the sufficient condition with $\delta_T = A_{IM}^D$, $\mathbf{X}_t = \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)'$, $\widehat{\theta} = \widehat{\theta}$, and $\theta_0 = \theta_{**}$.

Theorem 1 states that the HAC estimator $\widehat{\Omega}$ is consistent, which is established by showing that the conditions (V3-ii) and (V3-iii) hold.

Theorem 1. *Under Assumption 1, (K), (S), and (V1), as T and M grow, it holds that*

$$\widehat{\Omega} \xrightarrow{p} \Omega, \text{ and } \widehat{\sigma}_{uv}^2 \xrightarrow{p} \sigma_{uv}^2. \tag{18}$$

Proof: See the Appendix.

Remark. Unlike $\widehat{u}_t^d$, the undemeaned residual $\widehat{u}_t$ does not satisfy the Hansen's sufficient condition. To see this, rewrite the residual as

$$\begin{aligned}\widehat{u}_t &= u_t - \left(f_t',\, x_t',\, \mathbf{0}'\right)\left(\widehat{\theta} - \theta_*\right) \qquad (19)\\ &= u_t - \left(f_t',\, x_t',\, \mathbf{0}'\right)\left(\widehat{\theta} - \theta_{**}\right) - \left(f_t',\, x_t',\, \mathbf{0}'\right)A_{IM}^{D\prime}\begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0}\end{pmatrix}\\ &= u_t - \left(f_t',\, x_t',\, \mathbf{0}'\right)\left(\widehat{\theta} - \theta_{**}\right) - \delta_x'\Omega_{vv}^{-1}\Omega_{vu},\end{aligned}$$

and observe that the last term $\delta_x'\Omega_{vv}^{-1}\Omega_{vu}$ is generally not zero if the regressors have nonzero drifts ($\delta_x \neq \mathbf{0}$). Also, with a nonzero δ_x,

$$\sqrt{T}\left(f_t',\, x_t',\, \mathbf{0}'\right)A_{IM}^{D\prime} = \left(1,\, t/T,\, T^{-1/2}\alpha_x' + T^{-1/2}x_t^{0\prime},\, -\sqrt{T}\delta_x'\right)$$

explodes as T grows, violating the condition (V3-ii). Next Corollary directly shows that the HAC estimator constructed with $\widehat{u}_t$ is not consistent for Ω_{uu}. Define

$$\widetilde{\Omega}_{uu} = \widetilde{\Gamma}_0 + 2\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\widetilde{\Gamma}_j, \tag{20}$$

where $\widetilde{\Gamma}_j = \frac{1}{T}\sum_{t=j+1}^{T}\widehat{u}_t\widehat{u}_{t-j}$.

Corollary 1. *Under Assumption 1, (K), (S), and (V1), as T and M grow, it holds that*

$$\frac{1}{M}\widetilde{\Omega}_{uu} \xrightarrow{d} 2B_\infty^2\int_0^1 k(x)dx, \tag{21}$$

where $B_\infty = \delta_x'\left(A_\gamma^\infty - \Omega_{vv}^{-1}\Omega_{vu}\right)$.

Proof: See the Appendix.

Remark. Define $\widetilde{\Omega} = \frac{1}{T}\sum_{i=1}^{T}\sum_{j=1}^{T} k\left(\frac{|i-j|}{M}\right)\widetilde{\eta}_i\widetilde{\eta}_j' = \begin{bmatrix} \widetilde{\Omega}_{uu} & \widetilde{\Omega}_{uv} \\ \widetilde{\Omega}_{vu} & \widetilde{\Omega}_{vv} \end{bmatrix}$ with $\widetilde{\eta}_t \equiv (\widehat{u}_t,\ \triangle \ddot{x}_t')'$. Note that $\widetilde{\Omega}_{vv} = \widehat{\Omega}_{vv} \xrightarrow{p} \Omega_{vv}$. In addition, it can be shown that $\widetilde{\Omega}_{uv}$ (and $\widetilde{\Omega}_{vu}$) is $O_p\left(\frac{M^{3/2}}{T}\right)$ as $\widetilde{\Omega}_{uu}$ is, by using the same arguments in the proof of Corollary 1. Thus, $\widetilde{\sigma}_{uv}^2 \equiv \widetilde{\Omega}_{uu} - \widetilde{\Omega}_{uv}\widetilde{\Omega}_{vv}^{-1}\widetilde{\Omega}_{vu} = O_p(M) - O_p\left(\frac{M^{3/2}}{T}\right)O_p(1)O_p\left(\frac{M^{3/2}}{T}\right) = O_p(M) - O_p\left(\left(\frac{M}{T}\right)^2 M\right) = O_p(M)$. Hence the associated t statistic converges to zero under H_0. Therefore, the standard normal approximation is invalid.

3.2. FMOLS-BASED TEST

For the fully modified OLS (FMOLS) estimator, Hansen (1992a) shows

$$T\left(\widehat{\beta}_{FM} - \beta\right) \xrightarrow{d} \left(\mathbf{0},\ \Omega_{vv}^{-1/2}\right) \times \sigma_{uv}\left(\int_0^1 J_u(r)J_u(r)'dr\right)^{-1}\int_0^1 J_u(r)w_{uv}(r)dr$$
$$= \sigma_{uv}E^{FM}\Pi_{FM}^{-1}\left(\int_0^1 J_u(r)J_u(r)'dr\right)^{-1}\int_0^1 J_u(r)w_{uv}(r)dr := A_{\beta_{FM}}^{\infty},$$

with $J_u(r) = (f(r)', w_v(r)')'$, $E^{FM} = (\mathbf{0}_{k\times 2}, I_k)$, and $\Pi_{FM} = diag\left(I_2, \Omega_{vv}^{1/2}\right)$. The asymptotic distribution for the FMOLS estimator of δ_1 can be similarly obtained by following the steps applied for deriving (15) (see Cho (2022a) and Hansen (1992a)).

$$T\left(\widehat{\delta}_1^{FM} - \delta_1\right) \xrightarrow{d} -\delta_x' A_{\beta_{FM}}^{\infty}. \tag{22}$$

The variance of the limit in (22) conditional on $w_v(\cdot)$ is given by

$$\delta_x' V_{FM}^{\beta}\delta_x = \delta_x'\left(\sigma_{uv}^2 E^{FM}\Pi_{FM}^{-1}\left(\int_0^1 J_u(r)J_u(r)'dr\right)^{-1}\Pi_{FM}^{-1}E^{FM\prime}\right)\delta_x \tag{23}$$
$$:= \delta_x'\left(\sigma_{uv}^2 E^{FM}V_{FM}^{o}E^{FM\prime}\right)\delta_x,$$

where $V_{FM}^{o} \equiv \Pi_{FM}^{-1}\left(\int_0^1 J_u(r)J_u(r)'dr\right)^{-1}\Pi_{FM}^{-1}$. The inference for δ_1 can be conducted with

$$t_{\delta_1}^{FM} = \frac{T\widehat{\delta}_1^{FM}}{\sqrt{\widehat{\delta}_x'\widetilde{V}_{FM}^{\beta}\widehat{\delta}_x}}, \tag{24}$$

where $\widehat{\delta}_x = \frac{1}{T}\sum_{t=2}^{T} \triangle x_t$ and $\widetilde{V}^{\beta}_{FM} = \overline{\sigma}^2_{uv}\left(E^{FM}\cdot \widetilde{V}^{o}_{FM}\cdot E^{FM\prime}\right)$. Here,

$$\widetilde{V}^{o}_{FM} = \left(D^{FM}\frac{1}{T}\sum_{t=1}^{T} x_t^{FM} x_t^{FM\prime} D^{FM\prime}\right)^{-1}$$

with $x_t^{FM} = (f_t', x_t')'$ and $D^{FM} = diag\left(\tau_F^{-1}, T^{-1/2} I_k\right)$, and $\overline{\sigma}^2_{uv}$ is a nonparametric kernel-based HAC estimator constructed using the FMOLS residuals (see Section 4 in Hansen (1992a)). The standard conditioning argument yields $t^{FM}_{\delta_1} \xrightarrow{d} N(0,1)$ under H_0 as T grows.

3.3. **COMPARISON OF LOCAL POWER**

To derive the local power of the tests, consider the local alternative $H_A : \delta_1 = \frac{c}{T}$. Under H_A, as the sample size increases

$$t^{IM}_{\delta_1} = \frac{T\widehat{\delta}_1}{\sqrt{\widehat{\delta}_x'\widehat{V}^{\beta}_{IM}\widehat{\delta}_x}} = \frac{T\left(\widehat{\delta}_1 - \frac{c}{T}\right)+c}{\sqrt{\widehat{\delta}_x'\widehat{V}^{\beta}_{IM}\widehat{\delta}_x}} \xrightarrow{d} Z + \frac{c}{\sqrt{\sigma^2_{uv}\delta_x'(E\cdot V^{o}_{IM}\cdot E')\delta_x}};$$

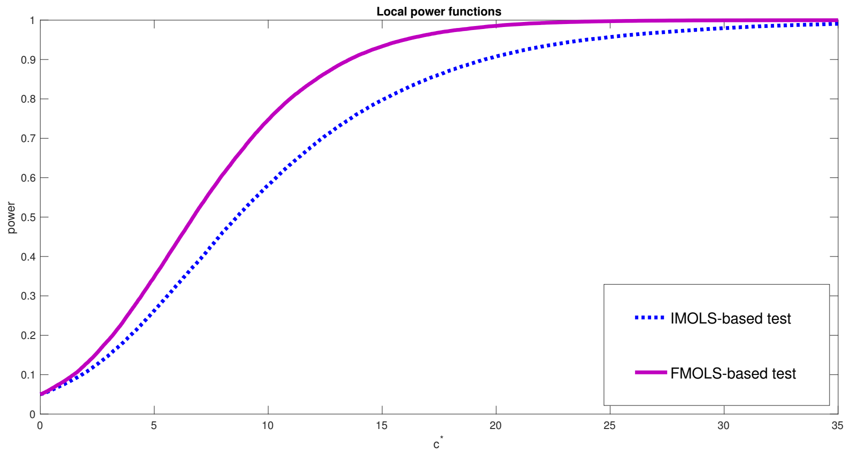
$$t^{FM}_{\delta_1} = \frac{T\widehat{\delta}_1^{FM}}{\sqrt{\widehat{\delta}_x'\widehat{V}^{\beta}_{FM}\widehat{\delta}_x}} = \frac{T\left(\widehat{\delta}_1^{FM} - \frac{c}{T}\right)+c}{\sqrt{\widehat{\delta}_x'\widehat{V}^{\beta}_{FM}\widehat{\delta}_x}} \xrightarrow{d} Z + \frac{c}{\sqrt{\sigma^2_{uv}\delta_x'(E^{FM}\cdot V^{o}_{FM}\cdot E^{FM\prime})\delta_x}} \text{ with } Z \sim N(0,1).$$

Consider the local power in the right-tail test for the case $k = 1$. To this end, define $c^* \equiv \frac{c}{\sigma_{uv}|\delta_x|\Omega_{vv}^{-1/2}}$. Under the local alternative: $\delta_1 = \frac{c}{T}$, $c > 0$, it holds that

$$t^{IM}_{\delta_1} \xrightarrow{d} Z + \frac{c^*}{\sqrt{\mathbf{e}_3' V^{\Xi}_{IM}\mathbf{e}_3}}; \quad t^{FM}_{\delta_1} \xrightarrow{d} Z + \frac{c^*}{\sqrt{\mathbf{e}_{FM}' V^{\Xi}_{FM}\mathbf{e}_{FM}}}, \text{ where } V^{\Xi}_{FM} = \left(\int_0^1 J_u(r)J_u(r)'dr\right)^{-1},$$

$\mathbf{e}_3 = (0,0,1,0)'$, $\mathbf{e}_{FM} = (0,0,1)'$, and V^{Ξ}_{IM} is defined in (14). Figure 1 depicts the local power curves of the two tests over $c^* \in [0,35]$. I used i.i.d N(0,1) pseudo random variables with $T = 1{,}000$ and the number of replications is 50,000. The nominal size is 5%.

Figure 1. Local Power Curve



The FMOLS-based test exhibits a higher local power than the IMOLS-based test, reflecting the asymptotic efficiency of the FMOLS estimator. However, the higher local power may mask the serious size distortion of the tests in small samples. The next section examines the finite-sample size property of the tests.

4. SIMULATION STUDY

In this section, the finite-sample properties of the tests are examined via a Monte Carlo simulation. The data generation process is given by

$$y_t = \delta_0 + \delta_1 t + \beta x_t + u_t \text{ with } \delta_0 = 0,\ \beta = 1 \tag{25}$$

and

$$x_t = \delta_x t + x_t^0 \text{ with } \delta_x = 1 \text{ and } x_t^0 = x_{t-1}^0 + v_t. \tag{26}$$

The error terms u_t and v_t are generated from AR(1) processes

$$u_t = \alpha u_{t-1} + \varepsilon_t^u \text{ and } v_t = \theta v_{t-1} + \varepsilon_t^v, \tag{27}$$

where

$$\begin{pmatrix} \varepsilon_t^u \\ \varepsilon_t^v \end{pmatrix} \sim \text{i.i.d}\ \mathbf{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).$$

The considered sample sizes are 100, 200, and 500, and the number of simulation replications is 5,000. The parameters α and θ control the persistence of u_t and v_t, respectively, and ρ captures the strength of endogeneity. The set of parameter values considered for this experiment is as follows: $\alpha \in \{0, 0.8, 0.9\}$, $\theta \in \{0, 0.2, 0.5, 0.8\}$, $\rho \in \{0, 0.2, 0.5, 0.8\}$, and $\delta_1 \in \{0, 0.1, 0.5\}$. The null and alternative hypotheses of interest are $H_0 : \delta_1 = 0$ and $H_1 : \delta_1 \neq 0$. Both tests are two-sided with the nominal size of 5%. For constructing HAC estimators, the Andrews' AR(1) plug-in data-dependent bandwidth (denoted by M^*) scheme (Andrews, 1991) is applied and the Bartlett kernel is used.

Table 1 presents the empirical rejection frequencies for the selected parameter values. The size-adjusted powers are reported in the parentheses. The FMOLS-based test ($t_{\delta_1}^{FM}$) exhibits a slightly higher power than the IMOLS-based test ($t_{\delta_1}^{IM}$), whereas $t_{\delta_1}^{IM}$ test is less size distorted. When the regression error does not exhibit autocorrelation (i.e. $\alpha = 0$), both tests can control the size very well, with the empirical size being close to the nominal size of 5%. However, as α increases, both tests suffer from size distortion (over-rejections). However, the IMOLS-based test is less size distorted and its empirical size is less sensitive to the strength of endogeneity, compared to $t_{\delta_1}^{FM}$ test. For the DGPs with no endogeneity ($\rho = 0$), the overall size property of the IMOLS-based test is comparable to or slightly better than that of the FMOLS-based test. For example, for DGPs with $\rho = 0$ and $\alpha = 0.9$, the FMOLS-based test shows rejection frequencies of around 18% even in the largest sample (T=500), whereas the empirical size of the IMOLS-based test ($t_{\delta_1}^{IM}$) is about 15%. However, for the DGPs with endogeneity, the size distortion of the FMOLS-based test is substantially more severe than that of the IMOLS-based test. For the last DGP in Table 1, the rejection frequency of $t_{\delta_1}^{FM}$ is almost twice higher than that of $t_{\delta_1}^{IM}$ when the sample size is 500.

By contrast, the empirical power of $t_{\delta_1}^{FM}$ is slightly better than that of $t_{\delta_1}^{IM}$, particularly when the true value of δ_1 is close to the null value, which is in line with the local power result in the previous section. However, when δ_1 is 0.5, the difference in the power becomes smaller. In addition, the size adjusted power of the $t_{\delta_1}^{IM}$ test is higher than that of $t_{\delta_1}^{FM}$ for the DGPs with endogeneity and highly persistent error term.

In Section 3 it was shown that the undemeaned IMOLS residuals lead to an invalid test. The null rejection frequencies of this invalid test are reported in Table 2. The test exhibits under-rejections when the null is true, , which corroborates the theoretical results in Section 3. The test power is extremely low when δ_1 is relatively small and the power seems to be overly sensitive to the strength of endogeneity and the degree of persistence.

Table 1. Empirical size and power

		T=100		T=200		T=500	
(ρ,α,θ)	δ_1	$t_{\delta_1}^{FM}$	$t_{\delta_1}^{IM}$	$t_{\delta_1}^{FM}$	$t_{\delta_1}^{IM}$	$t_{\delta_1}^{FM}$	$t_{\delta_1}^{IM}$
(0,0,0)	0	.0784	.0542	.0628	.0542	.0520	.0492
	0.1	.6740	.5114	.9736	.8704	1	.9974
	0.5	1	.9974	1	1	1	1
(0.2,0,0)	0	.0788	.0566	.0638	.0538	.0544	.0506
	0.1	.6726	.5194	.9748	.8758	1	.9984
	0.5	1	.9984	1	1	1	1
(0.8,0,0)	0	.0704	.0628	.0620	.0542	.0554	.0532
	0.1	.8496	.7972	.9968	.9810	1	.9998
	0.5	.9998	1	1	1	1	1
(0,0.8,0.2)	0	.2740	.2562	.1904	.1588	.1372	.1018
	0.1	.3634	.3238	.4780	.3828	.8652	.7264
		(.1056)	(.0912)	(.2356)	(.2016)	(.7588)	(.6490)
	0.5	.9366	.8442	.9980	.9690	1	.9996
		(.7712)	(.6548)	(.9860)	(.9342)	(1)	(.9988)
(0,0.8,0.5)	0	.2778	.2652	.1898	.1630	.1374	.1046
	0.1	.4746	.4280	.7094	.5828	.9812	.9114
		(.1892)	(.1582)	(.4590)	(.3836)	(.9516)	(.8676)
	0.5	.9862	.9504	1	.9942	1	1
		(.9326)	(.8614)	(.9994)	(.9862)	(1)	(1)
(0,0.8,0.8)	0	.3008	.3446	.1984	.1976	.1368	.1136
	0.1	.7764	.7424	.9582	.9146	1	.9972
		(.5116)	(.3480)	(.8844)	(.7986)	(.9998)	(.9942)
	0.5	.9992	.9962	1	1	1	1
		(.9902)	(.9430)	(1)	(.9994)	(1)	(1)
(0,0.9,0.2)	0	.3516	.3654	.2510	.2430	.1776	.1500
	0.1	.3702	.3850	.3674	.3326	.5706	.4450
		(.0774)	(.0682)	(.1120)	(.0926)	(.3378)	(.2744)
	0.5	.8082	.7206	.9510	.8630	.9996	.9868
		(.4750)	(.3820)	(.8074)	(.6710)	(.9982)	(.9680)
(0,0.9,0.5)	0	.3636	.3830	.2538	.2514	.1798	.1526
	0.1	.4362	.4382	.4972	.4400	.7984	.6700
		(.1068)	(.0976)	(.1960)	(.1658)	(.6214)	(.5090)
	0.5	.9324	.8724	.9930	.9606	1	.9976
		(.7182)	(.6150)	(.9570)	(.8802)	(1)	(.9958)
(0,0.9,0.8)	0	.3900	.4538	.2668	.2780	.1798	.1532
	0.1	.6714	.6740	.8390	.7796	.9926	.9654
		(.2902)	(.1930)	(.5936)	(.5006)	(.9754)	(.9264)
	0.5	.9900	.9776	.9998	.9980	1	1
		(.9242)	(.8120)	(.9986)	(.9854)	(1)	(1)
(0.5,0.9,0.5)	0	.4556	.4304	.3378	.2774	.2226	.1578
	0.1	.3476	.3644	.3530	.3670	.7356	.6894
		(.0202)	(.0282)	(.0542)	(.0852)	(.4746)	(.5100)
	0.5	.8954	.8582	.9892	.9654	1	.9992
		(.5002)	(.4756)	(.8998)	(.8656)	(1)	(.9982)
(0.8,0.9,0.5)	0	.6692	.5538	.5158	.3516	.3252	.1636
	0.1	.3928	.3580	.3020	.3246	.7454	.7866
		(.0080)	(.0084)	(.0098)	(.0356)	(.4210)	(.6160)
	0.5	.8972	.8902	.9950	.9872	1	.9998
		(.3310)	(.3752)	(.8672)	(.8902)	(1)	(.9998)

Note: In the parentheses are the size-adjusted powers. The simulation DGP is described in (25), (26), and (27). $H_0 : \delta_1 = 0$. $t_{\delta_1}^{IM}$ and $t_{\delta_1}^{FM}$ are defined in (16) and (24), respectively. The Bartlett kernel and the Andrews' AR(1) plug-in data-driven bandwidth are used. Both tests are two-sided with the nominal size of 5%.

Table 2. Empirical size and power, IMOLS-based test with undemeaned residuals, $T = 200$

		$\delta_x = 1$			$\delta_x = 2$		
(ρ, α, θ)		$(0.5, 0.9, 0.5)$	$(0.8, 0.9, 0.5)$	$(0.8, 0.9, 0.2)$	$(0.5, 0.9, 0.5)$	$(0.8, 0.9, 0.5)$	$(0.8, 0.9, 0.2)$
δ_1	0	.0752	.0222	.0142	.0326	.0084	.0066
	0.1	.1130	.0212	.0090	.0302	.0054	.0048
	0.5	.7050	.5010	.1028	.1828	.0354	.0090

Note: The simulation DGP is described in (25), (26), and (27). $H_0 : \delta_1 = 0$. The Bartlett kernel and the Andrews' AR(1) plug-in bandwidth are used. The test is two-sided and the reference distribution is the standard normal.

5. EMPIRICAL APPLICATION

To illustrate the use of the proposed test, I consider the cointegrating relationship between U.S. aggregate consumption (c_t), household wealth (a_t), and labor income (y_t) developed by Lettau and OgLudvigson (2001):

$$c_t = \delta_0 + \beta_a a_t + \beta_y y_t + u_t. \tag{28}$$

The CAY data is available for download at Martin Lettau's website.[4] The data covers the period 1952Q1-2019Q3 and the sample size is 271.

First, to test for the presence of nonzero drifts in the $I(1)$ regressors, I carried out t tests separately for household wealth and labor income series in the following intercept-only equations.

$$\triangle a_t = \delta_a + v_t^a, \text{ and } \triangle y_t = \delta_y + v_t^y.$$

The OLS estimates for the drifts are given by $\widehat{\delta}_a = \frac{1}{T-1}\sum_{t=2}^{T} \triangle a_t = 0.0063$, and $\widehat{\delta}_y = \frac{1}{T-1}\sum_{t=2}^{T} \triangle y_t = 0.0053$. The Newey-West standard errors are 0.0012 and 5.3/10^4, respectively. To calculate the standard errors, the Andrews' AR(1) plug-in data-driven bandwidth rule (Andrews, 1991) was applied. The t values are 5.2091, and 10.0047, being supportive for nonzero drifts in the regressors.

Next, to test for the regression trend coefficient, I considered the following regression equation.

$$c_t = \delta_0 + \delta_1 t + \beta_a a_t + \beta_y y_t + u_t. \tag{29}$$

The estimated equations are:

$$\widehat{c}_t^{FM} = 0.4120 + 0.00042t + 0.1560a_t + 0.7822y_t$$

[4]https://sites.google.com/view/martinlettau/data

in the FMOLS estimation, and

$$\widehat{c}_t^{IM} = -0.5028 - 0.00002t + 0.2127a_t + 0.8140y_t$$

in the IMOLS estimation. To obtain the FMOLS estimate, 42 lags were used to calculate the HAC estimator, following the Andrew's data-driven bandwidth scheme. The same bandwidth value was used to calculate $t_{\delta_1}^{FM}$. In the IMOLS inference, σ_{uv}^2 was estimated with 40 lags, as also suggested by the data-dependent bandwidth rule. The t values for the regression trend coefficient (δ_1) were given by

$$t_{\delta_1}^{FM} = 0.5968, \text{ and } t_{\delta_1}^{IM} = -0.0241,$$

whereas the 95% standard normal critical value is 1.96 in the two-tail tests. Thus, both tests fail to reject $H_0 : \delta_1 = 0$, which implies that the original regression specification in (28) is correct and the three series are deterministically cointegrated.

6. CONCLUSION

In this study, I proposed an IMOLS-based test on the regression trend slope coefficient. To construct the test statistic, I considered the IMOLS residuals obtained by plugging in the IMOLS estimator of the regression parameter in the original regression equation. This residual, after demeaning, can be used for constructing a consistent estimator of the long-run variance. In the simulation experiment, the proposed IMOLS-based test displayed better size property compared to the FMOLS-based test.

APPENDIX. Mathematical Proof

Proof of Theorem 1

Rearranging (17) by adding and subtracting $A_{IM}^{D\prime}\left(\left(T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu}\right)', \mathbf{0}', \mathbf{0}'\right)'$ yields

$$\begin{aligned}\widehat{u}_t^d = u_t^d &- \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right) A_{IM}^{D\prime} \\ &\times \left[\left(A_{IM}^{D\prime}\right)^{-1} \left\{ \left(\widehat{\theta} - \theta_*\right) + A_{IM}^{D\prime} \begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \right\} \right] \\ &- \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right) A_{IM}^{D\prime} \begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.\end{aligned} \tag{30}$$

The last term of the right handside in (30) is zero:

$$
\begin{aligned}
&\left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right) A_{IM}^{D\prime} \begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \qquad (31)\\
&= \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right) \begin{pmatrix} T^{-1/2}\tau_F^{-1} & -T^{-1}D_x' & -D_x^{F\prime} \\ \mathbf{0} & T^{-1}I_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_k \end{pmatrix} \begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}\\
&= \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right) \begin{pmatrix} D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} = f_t^{d\prime} D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} = 0,
\end{aligned}
$$

since $f_t^{d\prime} D_x^{F\prime} = \left(0,\, t - \frac{T+1}{2}\right) \begin{pmatrix} \delta_x' \\ \mathbf{0}' \end{pmatrix} = \mathbf{0}$.

This yields

$$
\begin{aligned}
\widehat{u}_t^d &= u_t^d - \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right) A_{IM}^{D\prime} \qquad (32)\\
&\quad \times \left[\left(A_{IM}^{D\prime}\right)^{-1} \left\{ \left(\widehat{\theta} - \theta_*\right) + A_{IM}^{D\prime} \begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \right\} \right]\\
&= u_t^d - \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right) \left\{ \left(\widehat{\theta} - \theta_*\right) + A_{IM}^{D\prime} \begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \right\}\\
&= u_t^d - \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right) \left(\widehat{\theta} - \theta_{**}\right),
\end{aligned}
$$

where $\theta_{**} = \theta_* - A_{IM}^{D\prime} \begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \delta - D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \beta \\ \Omega_{vv}^{-1}\Omega_{vu} \end{pmatrix}$.

Now, to check on the condition (V3-ii), consider

$$\sqrt{T}\left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)A_{IM}^{D\prime} = \sqrt{T}\left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)\begin{pmatrix} T^{-1/2}\tau_F^{-1} & -T^{-1}D_x' & -D_x^{F\prime} \\ \mathbf{0} & T^{-1}I_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_k \end{pmatrix} \tag{33}$$

$$= \left(f_t^{d\prime}\tau_F^{-1},\, \sqrt{T}\left(-T^{-1}f_t^{d\prime}D_x' + T^{-1}x_t^{d\prime}\right),\, -\sqrt{T}f_t^{d\prime}D_x^{F\prime}\right)$$

$$= \left(0,\, \frac{1}{T}\left(t - \frac{T+1}{2}\right),\, \frac{1}{\sqrt{T}}\left(x_t^{0\prime} - \frac{1}{T}\sum_{j=1}^{T} x_j^{0\prime}\right),\, \mathbf{0}'\right),$$

from which, by standard arguments, one can show

$$\sqrt{T}\left(f_{[rT]}^{d\prime},\, x_{[rT]}^{d\prime},\, \mathbf{0}'\right)A_{IM}^{D\prime} \Rightarrow \left(0,\, r - \frac{1}{2},\, \left(w_v(r) - \int_0^1 w_v(s)ds\right)'\Omega_{vv}^{1/2},\, \mathbf{0}'\right). \tag{34}$$

Thus,

$$\sup_{1\le t\le T}\left\|\sqrt{T}\left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)A_{IM}^{D\prime}\right\| = O_p(1), \tag{35}$$

which verifies that the condition (V3-ii) holds.

Next, by (12)

$$\left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta} - \theta_{**}\right) = \left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta} - \theta_{*}\right) + \begin{pmatrix} T^{1/2}\tau_F D_x^{F\prime}\Omega_{vv}^{-1}\Omega_{vu} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} = O_p(1), \tag{36}$$

which verifies the condition (V3-iii).

Proof of Corollary 1

Rewrite the last equation in (19) using $u_t = u_t^d + \frac{1}{T}\sum_{j=1}^{T} u_j$ and

$$\left(f_t',\, x_t',\, \mathbf{0}'\right)\left(\widehat{\theta} - \theta_{**}\right) = \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)\left(\widehat{\theta} - \theta_{**}\right) + \left(\frac{1}{T}\sum_{i=1}^{T} f_i',\, \frac{1}{T}\sum_{i=1}^{T} x_i',\, \mathbf{0}'\right)\left(\widehat{\theta} - \theta_{**}\right)$$

to obtain

$$\widehat{u}_t = u_t^d - \left(f_t^{d\prime},\, x_t^{d\prime},\, \mathbf{0}'\right)\left(\widehat{\theta} - \theta_{**}\right) + B_{1T} + B_{2T} + B_{3T} \tag{37}$$

$$= \widehat{u}_t^d + B_{1T} + B_{2T} + B_{3T} := \widehat{u}_t^d + B_T$$

with $B_{1T} = \frac{1}{T}\sum_{j=1}^{T} u_j$, $B_{2T} = -\left(\frac{1}{T}\sum_{i=1}^{T} f_i',\ \frac{1}{T}\sum_{i=1}^{T} x_i',\ \mathbf{0}'\right)\left(\widehat{\theta} - \theta_{**}\right)$ and $B_{3T} = -\delta_x'\Omega_{vv}^{-1}\Omega_{vu}$.

Note that $B_{1T} \xrightarrow{p} 0$, $\frac{1}{T^2}\sum_{i=1}^{T} f_i' D_x' \to \frac{1}{2}\delta_x'$, $\frac{1}{T^2}\sum_{i=1}^{T} x_i' = \frac{1}{T^2}\sum_{i=1}^{T}\left(\alpha_x + \delta_x i + x_i^0\right)' = \frac{1}{2}\delta_x' + o_p(1)$, and $\frac{1}{T}\sum_{i=1}^{T} f_i' D_x^{F\prime} = \delta_x'$. Thus,

$$
\begin{aligned}
B_{2T} &= -\left(\frac{1}{T}\sum_{i=1}^{T} f_i',\ \frac{1}{T}\sum_{i=1}^{T} x_i',\ \mathbf{0}'\right) A_{IM}^{D\prime}\left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta} - \theta_{**}\right) \qquad (38)\\
&= -\left(T^{-1/2}\frac{1}{T}\sum_{i=1}^{T} f_i'\tau_F^{-1},\ -\frac{1}{T^2}\sum_{i=1}^{T} f_i' D_x' + \frac{1}{T^2}\sum_{i=1}^{T} x_i',\ -\frac{1}{T}\sum_{i=1}^{T} f_i' D_x^{F\prime}\right)\left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta} - \theta_{**}\right)\\
&\Rightarrow -\left(\mathbf{0}',\ \mathbf{0}',\ -\delta_x'\right) A^{\infty} = \delta_x' A_\gamma^{\infty}
\end{aligned}
$$

by applying (11) and (12). Combining these results yields $B_T \xrightarrow{d} \delta_x' A_\gamma^{\infty} - \delta_x'\Omega_{vv}^{-1}\Omega_{vu} := B_\infty$.

Next, substitute $\widehat{u}_s^d + B_T$ for $\widehat{u}_s$ in $\widetilde{\Gamma}_j$ in (20) to rewrite $\widetilde{\Gamma}_j$ as

$$
\widetilde{\Gamma}_j = \widetilde{\Gamma}_j^I + \widetilde{\Gamma}_j^{II} + \widetilde{\Gamma}_j^{III} + \widetilde{\Gamma}_j^{IV},
$$

where $\widetilde{\Gamma}_j^I = \frac{1}{T}\sum_{t=j+1}^{T}\widehat{u}_t^d\widehat{u}_{t-j}^d$, $\widetilde{\Gamma}_j^{II} = \frac{1}{T}\sum_{t=j+1}^{T}\widehat{u}_t^d B_T$, $\widetilde{\Gamma}_j^{III} = \frac{1}{T}\sum_{t=j+1}^{T}\widehat{u}_{t-j}^d B_T$, and $\widetilde{\Gamma}_j^{IV} = \frac{T-j}{T}B_T^2$. Hence

$$
\begin{aligned}
\widetilde{\Omega}_{uu} &= \left(\widetilde{\Gamma}_0^I + 2\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\widetilde{\Gamma}_j^I\right) + \left(\widetilde{\Gamma}_0^{II} + 2\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\widetilde{\Gamma}_j^{II}\right) \qquad (39)\\
&\quad + \left(\widetilde{\Gamma}_0^{III} + 2\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\widetilde{\Gamma}_j^{III}\right) + \left(\widetilde{\Gamma}_0^{IV} + 2\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\widetilde{\Gamma}_j^{IV}\right)\\
&:= \widetilde{\Omega}_{uu}^I + \widetilde{\Omega}_{uu}^{II} + \widetilde{\Omega}_{uu}^{III} + \widetilde{\Omega}_{uu}^{IV}.
\end{aligned}
$$

Note that $\widetilde{\Gamma}_0^{II} = \widetilde{\Gamma}_0^{III} = 0$ because $\sum_{t=1}^{T}\widehat{u}_t^d = 0$.

The first term $\widetilde{\Omega}_{uu}^I$ in above equation converges to Ω_{uu} by Theorem 1. The second term, by recalling that $\sum_{t=1}^{T}\widehat{u}_t^d = 0$, can be rewritten as

$$
\widetilde{\Omega}_{uu}^{II} = 2\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\widetilde{\Gamma}_j^{II} = 2\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\frac{1}{T}\sum_{t=j+1}^{T}\widehat{u}_t^d B_T = \frac{2}{T}B_T\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\left(-\sum_{i=1}^{j}\widehat{u}_i^d\right). \qquad (40)
$$

To derive a limit result for $\sum_{i=1}^{j}\widehat{u}_i^d$, note that

$$
\sum_{t=1}^{[rT]}\widehat{u}_t^d = \sum_{t=1}^{[rT]} u_t^d - \sum_{t=1}^{[rT]}\left(f_t^{d\prime},\ x_t^{d\prime},\ \mathbf{0}'\right) A_{IM}^{D\prime}\left(A_{IM}^{D\prime}\right)^{-1}\left(\widehat{\theta} - \theta_{**}\right)
$$

by (32) and that Equation (33) implies

$$\sum_{t=1}^{[rT]} \widehat{u}_t^d = \sum_{t=1}^{[rT]} u_t^d - \sum_{t=1}^{[rT]} \left(0, \frac{1}{T^{3/2}}\left(t - \frac{T+1}{2}\right), \frac{1}{\sqrt{T}}\left(x_t^{0\prime} - \frac{1}{T}\sum_{j=1}^{T} x_j^{0\prime}\right),\ \mathbf{0}'\right) \left(A_{IM}^{D\prime}\right)^{-1} \left(\widehat{\theta} - \theta_{**}\right). \tag{41}$$

Then, by Assumption 1, Equation (12) and by applying the continuous mapping theorem (CMT),

$$T^{-1/2}\sum_{t=1}^{[rT]} \widehat{u}_t^d \Rightarrow B_u(r) - rB_u(1) - \left(0, \frac{r(r-1)}{2}, \int_0^r B_v'(s)ds - r\int_0^1 B_v'(s)ds, \mathbf{0}'\right) \times A^{\infty} := \Xi(r). \tag{42}$$

Therefore,

$$\begin{aligned} \widetilde{\Omega}_{uu}^{II} &= \frac{2}{T} B_T \sum_{j=1}^{M} k\left(\frac{j}{M}\right)\left(-\sum_{i=1}^{j} \widehat{u}_i^d\right) \qquad (43) \\ &= -\frac{M^{3/2}}{T} 2B_T \frac{1}{M}\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\left(M^{-1/2}\sum_{i=1}^{j}\widehat{u}_i^d\right) = O_p\left(\frac{M^{3/2}}{T}\right) \end{aligned}$$

because

$$2B_T\frac{1}{M}\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\left(M^{-1/2}\sum_{i=1}^{j}\widehat{u}_i^d\right) \xrightarrow{d} 2B_{\infty}\int_0^1 k(u)\Xi(u)\,du \tag{44}$$

as M and T grow. This implies $\frac{1}{M}\widetilde{\Omega}_{uu}^{II}$ converges in probability to zero as T and M grow with $\frac{M}{T}$ shrinking to zero.

The last term $\widetilde{\Omega}_{uu}^{IV}$ in (39) can be rewritten as

$$B_T^2\left(1 + 2\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\frac{T-j}{T}\right). \tag{45}$$

The summation in this expression can be further rewritten by using $\frac{T-j}{T} = -\frac{M}{T}\left(\frac{j}{M}\right) + 1$ as

$$\begin{aligned} \sum_{j=1}^{M} k\left(\frac{j}{M}\right)\frac{T-j}{T} &= \sum_{j=1}^{M} k\left(\frac{j}{M}\right)\left[-\frac{M}{T}\left(\frac{j}{M}\right)+1\right] \qquad (46) \\ &= -\frac{M^2}{T}\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\left(\frac{j}{M}\right)\frac{1}{M} + M\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\frac{1}{M}. \end{aligned}$$

Note that $\frac{1}{M}\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\left(1-\frac{j}{M}\right) \to \int_0^1 k(x)\,(1-x)\,dx < \infty$, and $\frac{1}{M}\sum_{j=1}^{M} k\left(\frac{j}{M}\right) \to \int_0^1 k(x)dx < \infty$ as M increases, by the integrability condition in (K). Thus, as T and M grow with $\frac{M}{T}$ shrinking to zero,

$$\frac{1}{M}\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\frac{T-j}{T} = -\frac{M}{T}\sum_{j=1}^{M} k\left(\frac{j}{M}\right)\left(\frac{j}{M}\right)\frac{1}{M} + \sum_{j=1}^{M} k\left(\frac{j}{M}\right)\frac{1}{M} \to \int_0^1 k(x)dx, \tag{47}$$

from which one can deduce $\widetilde{\Omega}_{uu}^{IV} = O_p(M)$ by recalling $B_T^2 = O_p(1)$.

Finally, by combining these results one can deduce that $\widetilde{\Omega}_{uu} = O_p(M)$ and

$$\frac{1}{M}\widetilde{\Omega}_{uu} \xrightarrow{d} 2B_\infty^2 \int_0^1 k(x)dx. \tag{48}$$

REFERENCES

Andrews, D. W. K. (1991). "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, 59, 817-854.

Campbell, J. Y. and Perron, P. (1991). "Pitfalls and opportunities: What macroeconomists should know about unit roots," *NBER Macroeconomic Annual*, 6, 141-201.

Cho, C.K. (2022a). "Tests of the null of cointegration using Integrated and Modified OLS residuals," *Available at SSRN 4112080*.

Cho, C.K. (2022b). "Self-normalization inference for linear trends in cointegrating regressions," *Available at SSRN 4197849*.

Han, H.L. and Ogaki, M. (1997). "Consumption, income and cointegration," *International Review of Economics & Finance*, 6(2), 107-117.

Hansen, B. E. (1992a). "Consistent covariance matrix estimation for dependent heterogenous processes," *Econometrica*, 60, 967–972.

Hansen, B. E. (1992b). "Efficient estimation and testing of cointegrating vectors in the presence of deterministic trends," *Journal of Econometrics*, 53, 87–121.

Jansson, M. (2002). "Consistent covariance estimation for linear processes," *Econometric Theory*, 18, 1449–1459.

Kakkar, V. and Ogaki, M. (1999). "(1999), Real exchange rates and nontradables: A relative price approach," *Journal of Empirical Finance*, 6(2), 193–215.

Lettau, M. and Ludvigson, S. (2001). "Consumption, aggregate wealth, and expected stock returns," *the Journal of Finance*, 56(3), 815–849.

Mikayilov, J. I., Galeotti, M. and Hasanov, F. J. (2018). "The impact of economic growth on CO2 emissions in Azerbaijan," *Journal of cleaner production*, 197, 1558–1572.

Newey, W. K. and West, K. D. (1987). "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55, 703–708.

Ogaki, M. (1992). "Engel's law and cointegration," *Journal of Political Economy*, 100(5), 1027–1046.

Ogaki, M. and Park, J. Y. (1997). “A cointegration approach to estimating preference parameters,” *Journal of Econometrics*, 82, 107–134.

Perron, P. and Campbell, J. Y. (1993). “A note on Johansen’s cointegration procedure when trends are present,” *Empirical Economics*, 18, 777–790.

Perron, P. and Rodríguez, G. (2016). “Residuals-based tests for cointegration with generalized least-squares detrended data,” *The Econometrics Journal*, 19, 84–111.

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

Wagner, M. (2015). “The environmental Kuznets curve, cointegration and nonlinearity,” *Journal of Applied Econometrics*, 30(6), 948–967.