University of Connecticut

Department of Economics Working Paper Series

Low Frequency Robust Cointegrated Regression in the Presence of a Near-Unity Regressor

by

Jungbin Hwang
University of Connecticut

Gonzalo Valdés
Universidad de Tarapacá

Working Paper 2020-03
January 2020

365 Fairfield Way, Unit 1063
Storrs, CT 06269-1063
Phone: (860) 486-3022
Fax: (860) 486-4463
http://www.econ.uconn.edu/

Low Frequency Robust Cointegrated Regression in the Presence of a Near-Unity Regressor*

Jungbin Hwang
Department of Economics,
University of Connecticut

Gonzalo Valdés
Departamento de Ingeniería Industrial y de Sistemas,
Universidad de Tarapacá

January, 2020

Abstract

This paper develops a robust t and F inferences on a triangular cointegrated system when one may not be sure the economic variables are exact unit root processes. We show that the low frequency transformed augmented (TA) OLS method possesses an asymptotic bias term in the limiting distribution, and corresponding t and F inferences in Hwang and Sun (2017) are asymptotically invalid. As a result, the size of the cointegration vector can be extremely large for even very small deviations from the unit root regressors. We develop a method to correct the asymptotic bias of the TA-OLS test statistics for the cointegration vector. Our modified statistics not only adjusts the locational bias but also reflects the estimation uncertainty of the long-run endogenity parameter in the bias correction term and has asymptotic t and F limits. Based on the modified TA-OLS test statistics, the paper provides a simple Bonferroni method to test for the cointegration parameter.

JEL Classification: C12, C13, C32

Keywords: Cointegration, Local to Unity, t and F tests, Alternative Asymptotics, Low Frequency Econometrics, Transformed and Augmented OLS

*Email: jungbin.hwang@uconn.edu and gvaldes@uta.cl. Correspondence to: Jungbin Hwang, 365 Fairfield Way, U-1063, Storrs, CT 06269-1063.

1 Introduction

In economic theory, we often try to conclude long run structural relationships among economic variables over periods of time. When economic time series possess an exact unit-root, the structural relationships between the non-stationary I(1) variables are captured by the concept of cointegration in Engle and Granger (1987). In most of macroeconomic applications, however, it is arguable that fundamental economic variables follow the exact unit root process, e.g., Christiano and Eichenbaum (1990). In fact, modeling key variables in the cointegration system using unit roots usually come in practice through failure to reject the unit root hypothesis with a limited span of time series data (Elliott, 1998). Thus, the assumption of unit root in cointegration model may represent a lack of 'knowledge' about economic interactions behind the common stochastic trends. See, for more details, Christiano and Eichenbaum (1990), Elliott (1998), and Müller and Watson (2008, 2013).

In this paper, we develop robust t and F inferences on the triangular cointegrated system when we are not sure whether the economic variables exhibit the exact unit root or are close to being local to unity. The long run relationship in the cointegration regression can be equivalently understood as low-frequency behaviors of time series. By transforming time series from the original time domain, the analysis is carried on the domain of frequencies, such as short-run or long-run business cycles. There is a growing focus in recent time series literature that projects the time series on the domain of low frequencies and make inference about long-run variability using the transformed data. See Müller and Watson (2017, 2018) for more discussion on the low-frequency transformations and their applications in econometrics. The low frequency approach also includes the analysis of cointegrated relation, as in Bierens (1997), Müller and Watson (2013), Phillips (2014), and Hwang and Sun (2017).

A recent study by Hwang and Sun (2017, HS hereafter) develops convenient t and F tests for the cointegration vector. The triangular cointegrated system is characterized by I(1) regressors, which are endogenous within the structural relation. To keep it general, the short-run dynamics deviated from the cointegrated relationship are allowed to have serial dependence of unknown forms. After transforming the original non-stationary time series data and its first differences into a K number of low-frequency transformations, HS (2017) runs a transformed and augmented ordinary least square (TA-OLS) with the K number of observations. Then, practitioners can easily perform robust t and F inferences in cointegration regression using any canned statistical program that can compute the t and F statistics in the classical Gaussian linear regression. Small sample simulation evidence in HS (2017) is favorably compared to existing tests such as the fully-modified OLS estimator of Phillips and Hansen (1990) and the trend IV estimator of Phillips (2014).

One of the key assumptions in HS (2017) is that I(1) regressors in the cointegrated system are under the exact unit root process. Since the asymptotic inference of TA-OLS method in HS (2017) crucially relies on the exact unit root assumption, it is questionable whether the t and F tests of TA-OLS are still (asymptotically) valid once the cointegration system departs from the unit root assumption. In fact, the inference about cointegration vectors in the time domain framework can lead to flawed inference once we fail to account for the exact order of integration of the data, e.g. Elliott (1998). To account for the order of integration, we adopt a local to unity approximation of cointegration regressors and investigate the asymptotic behaviors of the TA-OLS estimator and corresponding test statistics. The local to unity assumption has gained an attractive feature of the modeling devices for the nearly integrated regressor, as in Bobkoski (1983), Cavanagh (1985), and Phillips (1987). Instead of maintaining a strict dichotomy between integrated and non-integrated regressor, the assumption of the local to unity regressor allows for a smoother transition between two processes and thus can provide a more reasonable approximation to the TA-OLS methods, especially when the length of time series is small.

We first derive the fixed-K limiting distributions of TA-OLS and show that the TA-OLS is still consistent and share a common mixture of the normal distribution as found in HS (2017). However, due to the local to unity regressor, the limits of the TA-OLS estimator have an asymptotic bias term. The asymptotic bias is a product of the two important characteristics in our cointegration model: the deviation from the exact unit root and the degree of long run endogeneity within the cointegration system. It is analytically shown that the limiting distributions of TA-OLS statistics are mixtures of non-central t and F distributions where the random non-centrality parameter depends on the asymptotic bias from the local to unity regressors. The presence of the random non-centrality parameter indicates that the convenient t and F approximations in HS (2017) are no longer valid asymptotically. This result is consistent with Elliott (1998) whose approximation of cointegration model is based on the time-domain. Our numerical results also show that the empirical size of TA-OLS method to test the cointegration vector can be extremely large for even very small deviations from a unit root regressor. On the other hand, we find that the TA-OLS estimator of the long run endogeneity coefficient in the augmented cointegrated system is still asymptotically centered toward its true value. Thus, even if there is a source of asymptotic bias in the cointegration system by the mistakenly first differenced I(1) regressors, one can still precisely perform the long-run endogeneity test using the t and F tests with the TA-OLS framework.

Since the goal of an empirical researcher is making a valid inference for the cointegration vector, our next analysis is to provide modified TA-OLS statistics that correct the asymptotic bias. The modified statistics not only adjust the locational bias but also correct the estimation uncertainty of the long run endogeneity parameter in the bias correction term. After we fully account for both effects on the plugged-in bias correction formula, we show that the modified statistics have the asymptotic t and F limits. Thus, using our modified TA-OLS statistics, practitioners can conveniently implement robust t and F tests for the cointegration vector.

The modified test statistics in this paper require the knowledge of the local to unity parameter which is not consistently estimable in general. However, there are several ways developed in the time series literature to measure the uncertainty of the local to unity parameter in the context of unit -root testing problem. For example, Elliott and Stock (2001) constructs a nontrivial confidence interval (CI) for the unknown local to unity parameter by inverting a sequence of optimal tests in Gaussian autoregressions. Using the CI of local to unity parameter in Elliott and Stock (2001), we provide a simple Bonferroni method to the modified TA-OLS in the second stage. The idea of the Bonferroni confidence interval in the presence of unidentified nuisance parameters has been widely used in various context in statistics and econometrics. See, for example, Cavanagh, Elliott, and Stock (1995), Cambell and Yogo (2006), and McCloskey (2017). From Bonferroni's inequality, our Bonferroni CI for the cointegration parameter yields an asymptotically correct inference at least nominal coverage rate.

Our Monte Carlo results show that under the local to unity regressor, the unmodified TA-OLS methods in HS (2017) suffer from severe size distortions, especially when there exists a moderate amount of long-run endogeneity. Our finite sample studies further show the modified TA-OLS statistic plugged by the true local to unity parameter methods successfully controls the size distortions. The feasible version of the modified TA-OLS statistics using the Bonferroni method also has asymptotically correct sizes but is expected to result in some power loss because the Bonferroni correction yields to more conservative tests. We show that the power loss is increasing in the squared long-run correlation and the local to unity parameter.

Our paper contributes to recent literature in low frequency econometrics which first transforms time series onto a space of low frequencies and employs the fixed-K asymptotics for the finite number of transformed data (Müller and Watson; 2008, 2017). The low frequency transformation of time series has a long history in statistics and engineering; see, for example, the contribution of Thomson (1982) and Chapter 5 of Stoica and Moses (2005) for a textbook treatment. In the context of the cointegrated time series, Phillips (1991) estimates the cointegration parameter using frequency domain techniques, and Bierens (1997) proposes a nonparametric tests for the number of cointegrations using a transformed time series. More recently, Phillips (2014) develops an optimal estimation of cointegration using trend instrumental variables, and Müller and Watson (2013) use the Neyman-Pearson decision-theoretic framework to design robust and nearly optimal tests about the cointegration vectors using a fixed number of transformed data.

While this paper employs basis function transformation as a tool to estimate the main parameters of interest, the approach has also been used in the recent heteroskedasticity and autocorrelation robust inference (HAR) literature for time series models. The recent research along this line was inspired by Phillips (2005), Müller (2007), and Sun et al. (2008). See also Hwang and Sun (2017), Lazarus et al. (2018), and Lazarus et al. (2019).

The rest of the paper is organized as follows. Section 2 introduces an idea of low-frequency transformed regression analysis of cointegration and the fixed-K asymptotics limits of the TA-OLS estimator and the corresponding t and F tests. Section 3 extends the low frequency transformed cointegration system in the presence of a local to unity regressor. The next sections provide a method to correct the asymptotic bias of TA-OLS test statistics and suggest a feasible Bonferroni approach. Section 6 presents simulation evidence. The last section concludes. Proofs are given in the Appendix.

2 Low Frequency Transformation of Time Series

To illustrate the idea of the low-frequency transformed regression analysis of cointegration, we start by considering

$$y_t = \alpha_0 + x_t'\beta_0 + u_{0t} \quad \text{for } t = 1, \ldots, T, \tag{1}$$

where y_t is a scalar time series and x_t is a $d \times 1$ vector of time series. The main focus of interest is a parameter vector $\beta_0 \in R^d$. There are two-key features for the cointegration regression equation in (1). First, the regressor x_t has a unit root with a stationary innovation u_{xt} as

$$x_t = x_{t-1} + u_{xt} \quad \text{for } t = 1, \ldots, T, \tag{2}$$

and $x_0 = O_p(1)$.To maintain the generality, we allow the $I(0)$ errors $u_t \equiv (u_{0t}, u_{xt}')' \in R^m$ with $m = d + 1$ to be weakly stationary with serial dependence of unknown forms with the following long run variance (LRV) matrix Ω:

$$\underset{m\times m}{\Omega} = \sum_{j=-\infty}^{\infty} E u_t u_{t-j}' = \begin{pmatrix} \underset{1\times 1}{\sigma_0^2} & \underset{1\times d}{\sigma_{0x}} \\ \underset{d\times 1}{\sigma_{x0}} & \underset{d\times d}{\Omega_{xx}} \end{pmatrix}.$$

We assume that Ω_{xx} is positive definite, and hence x_t is a full-rank integrated process. Letting $u_{0\cdot xt} = u_{0t} - \delta_0' u_{xt}$ for $\delta_0 = \Omega_{xx}^{-1}\sigma_{x0}$, a long-run projection of u_{0t} onto u_{xt}, we can re-write the cointegrated regression equation in (1) in the following augmented form

$$y_t = \alpha_0 + x_t'\beta_0 + \delta_0' \Delta x_t + u_{0\cdot xt} \text{ for } t = 1, ..., T, \tag{3}$$

where $\Delta x_t = x_t - x_{t-1} = u_{xt}$.

The low frequency transformation of the cointegration system starts by projecting the original time series data $\{y_t, x_t', \Delta x_t'\}_{t=1}^T$ onto a space spanned by K number of basis functions $\{\phi_i\}_{i=1}^K$, which leads the following set of transformed data: For $i = 1, ..., K$,

$$\mathbb{W}_{y,i} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_t \phi_i(\frac{t}{T}), \; \mathbb{W}_{x,i} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_t \phi_i(\frac{t}{T}), \text{ and } \mathbb{W}_{\Delta x,i} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \Delta x_t \phi_i(\frac{t}{T}). \tag{4}$$

With these transformed data, the long run behaviors of the original time series are captured by choosing a proper set of basis functions $\{\phi_i(\cdot)\}_{i=1}^K$ which can concentrate on the low frequency components of time series. Examples include Fourier basis functions considered in Sun (2013, 2014) and HS (2017):

$$\left\{\phi_{2j-1}(r) = \sqrt{2}\cos(2j\pi r), \; \phi_{2j} = \sqrt{2}\sin(2j\pi r), \; j = 1, \ldots, K/2\right\}, \tag{5}$$

and cosine basis functions suggested in Müller and Watson (2008, 2013):

$$\left\{\phi_j(r) = \sqrt{2}\cos(j\pi r), \; j = 1, \ldots, K\right\}. \tag{6}$$

The low-frequency transformations enjoy several advantages for estimating and making an inference about the parameter of the long-run relationship β_0. Letting

$$\Phi_i = [\phi_i(1/T), \ldots, \phi_i((T-1)/T), \phi_i(1)]' \in \mathbb{R}^T$$

as a basis vector corresponding to the basis functions in (5)–(6), one can easily show that a matrix of K basis vectors $\Phi = [l_T, \Phi_1, \ldots, \Phi_K] \in \mathbb{R}^{T\times(K+1)}$ including the column of ones $l_T = (1, ..., 1)' \in \mathbb{R}^T$ satisfy $(\Phi'\Phi)^{-1} = T^{-1}I_{K+1}$. Therefore, the transformed data becomes a scale of the OLS regression coefficient of the original time series data on the space basis functions. For example, with $X = (x_1, x_2, ..., x_T)'$, the vector of (scaled) transformed data $\tilde{\mathbb{W}}_x = \{\tilde{\mathbb{W}}_{x,i}\}_{i=1}^K$ with $\tilde{\mathbb{W}}_{x,i} = \mathbb{W}_{x,i}/\sqrt{T}$ and the sample mean $\bar{x}_T = T^{-1}\sum_{j=1}^T x_t$ is equal to the OLS coefficient of $(\Phi'\Phi)^{-1}\Phi'X = \Phi'X$. Then, the low-frequency movement of time series can be captured by using the non-stochastic trend predictor Φ multiplied by the OLS coefficient $(\bar{x}_T, \tilde{\mathbb{W}}_x')'$ as

$$x_t = \bar{x} + \underbrace{\phi_1(\frac{t}{T})\tilde{\mathbb{W}}_{x,1} + \ldots + \phi_K(\frac{t}{T})\tilde{\mathbb{W}}_{x,K}}_{\text{Low Frequency Components}} + \hat{u}_{xt}.$$

The low frequency component captures the long run movements of the original data with periodicity longer than $2T/j$ for $j = 1, ..., K$ years of cycles. A useful rule of thumb introduced in Müller (2014) and Müller and Watson (2017) suggests a choice of $K = 16$ to capture the low-frequency movements of $T = 65$ years of Post World War II macro data with periodicity higher than the commonly accepted business cycle period of $T/(K/2) \simeq 8$ years. The low-frequency transformation also has substantive empirical content in the context of the cointegration regression system in (1)–(2), as the cointegration model itself seeks a long run relation among economic time series.

Using linearity of low frequency transformations in (4), we can translate the augmented cointegration regression in (3) into the following form of transformed and augmented (TA) regression:

$$\mathbb{W}_{y,i} = \mathbb{W}_{x,i}'\beta_0 + \mathbb{W}_{\Delta x,i}'\delta_0 + \mathbb{W}_{0\cdot x,i} \text{ for } i = 1, ..., K, \tag{7}$$

where $\mathbb{W}_{0\cdot x,i} := T^{-1//2}\sum_{t=1}^T \phi_i(\frac{t}{T})u_{0\cdot xt}$. Throughout the paper, we maintain functional central limit theorem (FCLT) for $\{u_t\}$

$$T^{-1/2}\sum_{t=1}^{[T\cdot]} u_t \Rightarrow B(\cdot) := \Omega^{1/2}W(\cdot) = \begin{pmatrix} \sigma_{0\cdot x}w_0(\cdot) + \sigma_{0x}\Omega_{xx}^{-1/2}W_x(\cdot) \\ \Omega_{xx}^{1/2}W_x(\cdot) \end{pmatrix}, \tag{8}$$

where $W(\cdot) := (w_0(\cdot), W_x'(\cdot))'$ is an m-dimensional standard Brownian process, $\sigma_{0\cdot x}^2 = \sigma_0^2 - \sigma_{0x}\Omega_{xx}^{-1}\sigma_{x0}$, and $\Omega^{1/2}$ is a Cholesky decomposition of the LRV Ω. One can find primitive conditions to hold the FCLT assumption in Durlauf (1986), Phillips and Solo (1992), Davidson (1994), among others. With the FCLT assumption in (8), we can use summation by parts, continuous mapping theorem, and integration by parts to get

$$\mathbb{W}_{\Delta x,i} \Rightarrow \Omega_{xx}^{1/2}\int_0^1 \phi_i(r)\,dW_x(r) \sim N(0, \Omega_{xx}), \tag{9}$$

$$\mathbb{W}_{0\cdot x,i} \Rightarrow \sigma_{0\cdot x}\int_0^1 \phi_i(r)\,dw_0(r) \sim N(0, \sigma_{0\cdot x}^2) \tag{10}$$

for $i=1,...,K$. Also, invoking the continuous mapping theorem together with (8), we have

$$\frac{\mathbb{W}_{x,i}}{T}=\frac{1}{T^{3/2}}\sum_{s=1}^{T}\phi_i\left(\frac{s}{T}\right)x_s\Rightarrow\Omega_{xx}^{1/2}\int_0^1\phi_i\left(r\right)W_x(r)dr\sim N(0,\Omega_{xx}^{1/2}\Sigma\Omega_{xx}^{1/2}),\tag{11}$$

where $\Sigma=\int_0^1\int_0^1\phi_i(r)\phi_i(s)\min(r,s)drds\cdot I_d$, for $i=1,...,K$. Since the weak convergences in (9)–(11) hold jointly, the TA regression in (7) naturally leads us to consider a small sample approximately Gaussian linear regression model

$$\mathbb{W}_{y,i}\simeq(\mathbb{S}_{x,i}')\beta_{T,0}+\mathbb{S}_{\Delta x,i}'\delta_0+\mathbb{S}_{0\cdot x,i}\text{ for }i=1,...,K,\tag{12}$$

where $\beta_{T,0}=T\beta_0$, $\mathbb{S}_{\Delta x,i}$, $\mathbb{S}_{0\cdot x,i}$, and $\mathbb{S}_{x,i}$ are the Gaussian weak convergence limits of $\mathbb{W}_{\Delta x,i}$, $\mathbb{W}_{0\cdot x,i}$, and $\mathbb{W}_{x,i}/T$, respectively, which are specified in (9), (10), and (11), respectively. Since $W_x(\cdot)$ and $w_0(\cdot)$ are independent, $\{\mathbb{S}_{x,i},\mathbb{S}_{\Delta x,i}\}_{i=1}^K$,the functional of $W_x(\cdot)$, and $\{\mathbb{S}_{0\cdot x,i}\}_{i=1}^K$, the functional of $w_0(\cdot)$, are independent. Also, the orthonormal property of the basis functions $\{\phi_i(\cdot)\}_{i=1}^K$ ensures the errors of regression $\{\mathbb{S}_{0\cdot x,i}\}_{i=1}^K$ are i.i.d normal with zero mean and variance $\sigma_{0\cdot x}^2$. Therefore, standard OLS framework of the sample Gaussian linear regression model can be applied to estimate the parameters $\beta_{T,0}$ and δ_0. HS (2017) runs the OLS estimator for $\gamma_0=(\beta_0',\delta_0')'$ based on (7) and defines TA-OLS estimator of γ_0 as

$$\hat{\gamma}=(\hat{\beta}',\hat{\delta}')'=(\mathbb{W}_X'\mathbb{W}_X)^{-1}\mathbb{W}_X'\mathbb{W}_y,$$

where $\mathbb{W}_X=(\mathbb{W}_x,\mathbb{W}_{\Delta x})$. HS (2017) shows

$$\hat{\beta}\overset{A}{\sim}N\left[\beta_0,\sigma_{0\cdot x}^2(\mathbb{W}_x'M_{\Delta x}\mathbb{W}_x)^{-1}\right],\tag{13}$$

and

$$\hat{\delta}\overset{A}{\sim}N\left[\delta_0,\sigma_{0\cdot x}^2(\mathbb{W}_{\Delta x}'M_x\mathbb{W}_{\Delta x})^{-1}\right],\tag{14}$$

where $M_{\Delta x}=I_K-\mathbb{W}_{\Delta x}\left(\mathbb{W}_{\Delta x}'\mathbb{W}_{\Delta x}\right)^{-1}\mathbb{W}_{\Delta x}'$ and $M_x=I_K-\mathbb{W}_x\left(\mathbb{W}_x'\mathbb{W}_x\right)^{-1}\mathbb{W}_x'$. To test a hypothesis of

$$H_0^\beta: R_\beta\beta_0=r_\beta\text{ vs. }H_1:\ R_\beta\beta_0\neq r_\beta,\tag{15}$$

where R is a $p_\beta\times d$ matrix, HS (2017) constructs the following Wald statistic and derives its limiting distribution by

$$\begin{aligned}F(\hat{\beta})&=\frac{1}{\hat{\sigma}_{0\cdot x}^2}(R_\beta\hat{\beta}-r_\beta)'\left[R_\beta(\mathbb{W}_x'M_{\Delta x}\mathbb{W}_x)^{-1}R_\beta'\right]^{-1}(R_\beta\hat{\beta}-r_\beta)/p_\beta\\&\Rightarrow\frac{K}{K-2d}\cdot F_{p_\beta,K-2d},\end{aligned}\tag{16}$$

where $F_{p_\beta,K-2d}$ is the F distribution with degrees of freedom p_β and $K-2d$. When $p=1$, the t-statistic can be constructed in a similar manner. Here, $\hat{\sigma}_{0\cdot x}^2=K^{-1}\sum_{i=1}^K\hat{\mathbb{W}}_{0\cdot x,i}^2$ is a natural variance estimate of the regression error, where $\hat{\mathbb{W}}_{0\cdot x,i}=\mathbb{W}_{y,i}-\mathbb{W}_{x,i}'\hat{\beta}-\mathbb{W}_{\Delta x,i}'\hat{\delta}$ is a residual of the small sample regression in (12).

It is important to note that the asymptotic variances in (13)–(14) are different with convergence orders, $(\mathbb{W}_x'M_{\Delta x}\mathbb{W}_x)^{-1}=O_p(T^{-2})$ while $(\mathbb{W}_{\Delta x}'M_x\mathbb{W}_{\Delta x})^{-1}=O_p(1)$. The different convergence rates imply different orders of convergence for estimators $\hat{\beta}$ and $\hat{\delta}$ with $T(\hat{\beta}-\beta_0)=O_p(1)$

and $(\hat{\delta}-\delta_0)=O_p(1)$, respectively. The latter estimator $\hat{\delta}$ for the long-run endogeneity parameter is inconsistent but yields to asymptotically valid t and F tests for $H_0:\delta=\delta_0$ as in (16). The testing hypothesis is

$$H_0^{\delta}: R_{\delta}\delta_0 = r_{\delta} \text{ vs. } H_1^{\delta}: \ R_{\delta}\delta_0 \neq r_{\delta}, \tag{17}$$

where R is a $p_{\delta}\times d$ matrix, one can construct Wald statistic and obtain its limiting distribution as

$$\begin{aligned} F(\hat{\delta}) &= \frac{1}{\hat{\sigma}^2_{0\cdot x}}(R_{\delta}\hat{\delta}-r_{\delta})'\left[R_{\delta}(\mathbb{W}'_{\Delta x}M_x\mathbb{W}_{\Delta x})^{-1}R_{\delta}\right]^{-1}(R_{\delta}\hat{\delta}-r_{\delta})/p_{\delta} \\ &\Rightarrow \frac{K}{K-2d}\cdot F_{p_{\delta},K-2d}. \end{aligned} \tag{18}$$

3 Asymptotic Behavior of TA-OLS with a Near Unity Regressor

Although the TA-OLS method is very convenient for practitioners with standard t and F tests, it crucially relies on the exact unit root assumption on the cointegration regressor x_t. The exact unit root assumption in x_t makes the low-frequency transformations of the first difference $\{\mathbb{W}_{\Delta x,i}\}_{i=1}^K$ to be same as those of $\{u_{xt}\}$, i.e. $\mathbb{W}_{\Delta x,i}=\mathbb{W}_{u_x,i}=T^{-1/2}\sum_{t=1}^T u_{xt}\phi_i(\frac{t}{T})$. As a result, the low-frequency transformations of the projected errors $\mathbb{W}_{0\cdot x,i}=T^{-1/2}\sum_{t=1}^T\phi_i(\frac{t}{T})u_{0\cdot xt}$, are asymptotically independent of the regressors $\{\mathbb{W}_{x,i}\}_{i=1}^K$ and $\{\mathbb{W}_{\Delta x,i}\}_{i=1}^K$ which govern long-run and short-run dynamics of the TA regression system, respectively. However, once the cointegration system departs from the unit root assumption, it is questionable whether the Gaussian approximation of the TA cointegration system is still valid. To answer this, we adopt a local to unity approximation of the cointegration regressor

$$x_t = \rho_T x_{t-1} + u_{xt} \text{ where } \rho_T = 1-\frac{c}{T} \tag{19}$$

for $c\geq 0$. When $c=0$, the regressor x_t has an exact unit root the I(0) errors. Modeling the cointegration regressor x_t as in (19) allows for a smooth transition between stationary but highly persistent and the "exact" I(1) non-stationary series and provides a more reasonable approximation to the TA cointegration system in (7). This is especially when the length of time series is not enough to identify the exact nature of the auto-regressive root of x_t.

With the local to unity approximation of regressor x_t in (19), the first difference process Δx_t becomes

$$\Delta x_t = -\frac{cx_{t-1}}{T}+u_{x,t} \text{ for } t=1,...,T.$$

Thus, the low frequency transformation $\{\mathbb{W}_{\Delta x,i}\}_{i=1}^K$ is no longer the same as $\{\mathbb{W}_{u_x,i}\}_{i=1}^K$ but is now a combination of two transformed data

$$\mathbb{W}_{\Delta x,i}=\frac{1}{\sqrt{T}}\sum_{t=1}^T u_{xt}\phi_i(\frac{t}{T})-c\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[\frac{x_{t-1}}{T}\right]\phi_i(\frac{t}{T}) \tag{20}$$

for $i=1,...,K$. The augmented cointegration regression in (3) is modified into

$$y_t=\alpha_0+x_t'\beta_0+\delta_0'\Delta x_t+\tilde{u}_{0\cdot xt} \text{ for } t=1,...,T,$$

where

$$\tilde{u}_{0\cdot xt} := u_{0\cdot xt} + c\left[\frac{\delta' x_{t-1}}{T}\right],$$

and thus the transformed regression model in (7) changes to

$$\mathbb{W}_{y,i} = \mathbb{W}'_{x,i}\beta_0 + \mathbb{W}'_{\Delta x,i}\delta_0 + \tilde{\mathbb{W}}_{0\cdot x,i} \text{ for } i = 1, ..., K, \tag{21}$$

where

$$\tilde{\mathbb{W}}_{0\cdot x,i} := \mathbb{W}_{0\cdot x,i} + \frac{c}{T^{3/2}}\left[\sum_{t=1}^{T}\delta_0' x_{t-1}\phi_i(\frac{t}{T})\right].$$

Since the null distributions of parameters in small sample Gaussian regression are invariant to the (asymptotic) variance of the regressors, we expect the two sets of transformed regressors $\{\mathbb{W}_{x,i}\}_{i=1}^{K}$ and $\{\mathbb{W}_{\Delta x,i}\}_{i=1}^{K}$in (21) to have the same role as what we obtained under the exact unit root regressor in (7). However, the regression equation in (21) now involves an additional error of low frequency transformation inside $\tilde{\mathbb{W}}_{0\cdot x,i}$, and it is now questionable whether the convenient features of the asymptotic t and F tests in (16)–(18) can still be maintained by the structure of small sample Gaussian regression model in (21). To answer this, we first make the following assumptions to formally establish the asymptotic properties of the TA-OLS estimator $\hat{\gamma} = (\hat{\beta}', \hat{\delta}')'$.

Assumption 1 *The vector process $\{u_t = (u_{0t}, u'_{xt})'\}_{t=1}^{T}$ satisfies the FCLT in (8).*

Assumption 2 *(i) For $i = 1, \ldots, K$, each function $\phi_i(\cdot)$ is continuously differentiable; (ii) For $i = 1, \ldots, K$, each function $\phi_i(\cdot)$ satisfies $\int_0^1 \phi_i(x)\,dx = 0$; (iii) The functions $\{\phi_i(\cdot)\}_{i=1}^{K}$ are orthonormal in $L^2[0,1]$.*

Together with the local to unity regressors in (19), Assumption 1 of FCLT enables us to invoke the result in Phillips (1987) and get

$$\frac{1}{\sqrt{T}}x_{[Tr]} \Rightarrow \Omega_{xx}^{1/2}J_c(r), \tag{22}$$

where the Ornstein-Uhlenbeck (OU) process is defined by $J_c(r) = \int_0^r \exp(-c(r-s))dW_x(s)$. Since Assumption 2 holds in both (5) and (6), we can repeat the weak convergence approximations in (11) allowing the local to unity assumption in (19)

$$\frac{\mathbb{W}_{x,i}}{T} \Rightarrow \Omega_{xx}^{1/2}\int_0^1 \phi_i(r)\,J_c(r)dr \sim N(0, \Omega_{xx}^{1/2}\Sigma_c\Omega_{xx}^{1/2}), \tag{23}$$

where $\Sigma_c = \frac{1}{2c}\int_0^1\int_0^1 \phi_i(r)\phi_i(s)\{\exp[-c|r-s|] - \exp[-c(r+s)]drds \cdot I_d$, for $i = 1, ..., K$. The above weak convergence shows that the local to unity assumption does not change the Gaussian limits but is has different asymptotic variance from (10). In the proof of Proposition 1, we show that

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{t-1}\phi_i(\frac{t}{T}) = \frac{\mathbb{W}_{x,i}}{T} + O_p(\frac{1}{T}).$$

Thus, the transformed first difference $\mathbb{W}_{\Delta x,i}$ and the regression error $\tilde{\mathbb{W}}_{0\cdot x,i}$ have the following weak convergence limits of

$$\begin{aligned}\mathbb{W}_{\Delta x,i} &\Rightarrow \Omega_{xx}^{1/2}\left[\int_0^1 \phi_i(r)dW_x(r) - c\cdot\int_0^1 \phi_i(r)J_c(r)dr\right], \\ \tilde{\mathbb{W}}_{0\cdot x,i} &\Rightarrow \sigma_{0\cdot x}\int_0^1 \phi_i(r)dw_0(r) + c\cdot\left[\Omega_{xx}^{1/2}\delta_0\right]'\int_0^1 \phi_i(r)J_c(r)dr\end{aligned} \tag{24}$$

for $i = 1, ...K$, respectively. Combining these results, the TA regression in (21) is now asymptotically equivalent to:

$$\mathbb{W}_{y,i} \simeq \mathbb{S}'_{x,i}\beta_{T,0} + \mathbb{S}'_{\Delta x,i}\delta_0 + \left[\mathbb{S}_{0\cdot x,i} + c\delta_0'\mathbb{S}_{x,i}\right] \text{ for } i = 1, ..., K,$$

where $\mathbb{S}_{x,i}$, $\mathbb{S}_{\Delta x,i}$, and $\mathbb{S}_{0\cdot x,i}$ are the Gaussian random limits of $\mathbb{W}_{x,i}/T$, $\mathbb{W}_{\Delta x,i}$, and $\mathbb{W}_{0\cdot x,i}$, respectively, which are specified in (23), (24), and (10), respectively. Then, the asymptotic behavior of TA-OLS estimator is captured by

$$\begin{aligned}T(\hat{\beta} - \beta_0) &= \left[\frac{\mathbb{W}_x'}{T}(I_K - P_{\Delta x})\frac{\mathbb{W}_x}{T}\right]^{-1}\left[\frac{\mathbb{W}_x'}{T}(I_K - P_{\Delta x})\tilde{\mathbb{W}}_{0\cdot x}\right] \\ &\Rightarrow \left[\mathbb{S}_x'(I_K - P_{\mathbb{S}_{\Delta x}})\mathbb{S}_x\right]^{-1}\mathbb{S}_x'(I_K - P_{\mathbb{S}_{\Delta x}})\mathbb{S}_{0\cdot x} + c\delta_0,\end{aligned}$$

where $P_{\mathbb{S}_{\Delta x}} = \mathbb{S}_{\Delta x}(\mathbb{S}_{\Delta x}'\mathbb{S}_{\Delta x})^{-1}\mathbb{S}_{\Delta x}'$. Conditioning on $\mathbb{S}_x$ and $\mathbb{S}_{\Delta x}$, the first majorant term characterizes the weak Gaussian limit of TA-OLS estimator under the unit root regressors which is centered toward the true parameter β_0. This limit is the same as what is derived under the exact unit root regressor in HS (2017), except for the covariance structure of the conditioning random variables $\mathbb{S}_x$ and $\mathbb{S}_{\Delta x}$. The second term $c\delta_0$ indicates the asymptotic distribution of $\hat{\beta}$ possesses a bias term $c\delta_0$. When $c = 0$, the results are the same as the previous I(0) cointegrated regression. We formally state the weak convergences result of TA-OLS estimator including $\hat{\delta}$ in the following Proposition. Define

$$\Upsilon_T = \begin{pmatrix} T\cdot I_d & \underset{d\times d}{0} \\ \underset{d\times d}{0} & I_d \end{pmatrix}.$$

Proposition 1 *Let $\mathbb{S}_X = [\mathbb{S}_x', \mathbb{S}_{\Delta x}']'$. Under Assumptions 1–2, and the local to unity regressors in (19), as $T \to \infty$ but holding K fixed,*

$$\Upsilon_T(\hat{\gamma} - \gamma_0) = \begin{pmatrix} T(\hat{\beta} - \beta_0) \\ \hat{\delta} - \delta_0 \end{pmatrix} \Rightarrow \begin{bmatrix} c\delta_0 \\ 0 \end{bmatrix} + MN(0, \sigma_{0\cdot x}^2(\mathbb{S}_X'\mathbb{S}_X)^{-1}).$$

The result of Proposition 1 can be summarized by

$$\begin{aligned}T(\hat{\beta} - \beta_0) &\Rightarrow MN\left[c\delta_0, \sigma_{0\cdot x}^2(\mathbb{S}_x'M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x)^{-1}\right], \\ \hat{\delta} - \delta_0 &\Rightarrow MN\left[0, \sigma_{0\cdot x}^2(\mathbb{S}_{\Delta x}'M_{\mathbb{S}_x}\mathbb{S}_{\Delta x})^{-1}\right],\end{aligned}$$

where the convergences hold jointly. As we expected, the local to unity regressor affects the limit behavior of $\hat{\beta}$ by shifting the center of the weak limit $T(\hat{\beta} - \beta_0)$ from zero to the asymptotic bias term $c\delta_0$. This implies the TA-OLS estimator $\hat{\beta}$ is asymptotically unbiased i) if the regressors have the exact unit root processes, i.e. $c = 0$ or ii) there is no long-run simultaneity between u_t and u_{xt}, i.e. $\delta_0 = 0$. Both of these cases, however, are unlikely to show up in practice. The results are similar with Elliott (1998) which finds the fragility of time-domain cointegration inference in the presence of local to unity regressors. Our work also shows that the same asymptotic bias terms appear in the domain of low frequencies.

Although the limiting distribution of cointegration vector is affected by the local to unity regressor, the result in Proposition 1 also indicates that $\hat{\delta}$ is still asymptotically centered toward δ_0 and has the exact same asymptotic behavior as the case of exact unit root regressors. Given that the source of asymptotic bias $\hat{\beta}$ is originated by the mistakenly first differenced data Δx_t in (21), it is very interesting to observe that the TA-OLS estimator $\hat{\delta}$ still yields an asymptotically unbiased estimation of the long-run endogeneity parameter δ_0. The TA-OLS estimator $\hat{\delta}$ of δ_0 is not consistent in our framework, but weakly converges to a random limit centered toward the true parameter β_0. This is because the underlying approximation scheme of our low-frequency transformed regression is based on "fixed-K" asymptotics which let the sample size T grow to infinity but holding K fixed. If one considers a different limiting experiment of approximating $\hat{\gamma}$ where K increases with T but at a slower rate, e.g. Phillips (2005, 2014), we expect $\hat{\delta}$ becomes a consistent estimator for δ_0. Searching for more accurate approximations of finite sample estimator $\hat{\gamma}$ (and thus $\hat{\delta}$), however, the results of our "fixed-K" asymptotics about $\hat{\delta}$ can provide a robust way of making inference for δ_0. Formally, under the null hypotheses in (15) and (17), the results in Proposition 1 gives

$$\begin{aligned} T(R_\beta\hat{\beta} - r_\beta) &\Rightarrow MN(cR_\beta\delta_0, \sigma^2_{0\cdot x}\left[R_\beta(\mathbb{S}_x' M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x)^{-1}R_\beta'\right]), \\ (R_\delta\hat{\delta} - r_\delta) &\Rightarrow MN(0, \sigma^2_{0\cdot x}\left[R_\delta(\mathbb{S}_{\Delta x}' M_{\mathbb{S}_x}\mathbb{S}_{\Delta x})^{-1}R_\delta'\right]), \end{aligned} \tag{25}$$

respectively. In view of the joint weak convergence results in (23)–(24), it is easy to check

$$\begin{aligned} R_\beta\left[(\mathbb{W}_x'/T)M_{\Delta x}(\mathbb{W}_x'/T)\right]^{-1}R_\beta' &\Rightarrow R_\beta\left[\mathbb{S}_x' M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\right]^{-1}R_\beta', \\ R_\delta(\mathbb{W}_{\Delta x}' M_x \mathbb{W}_{\Delta x})^{-1}R_\delta' &\Rightarrow R_\delta\left[\mathbb{S}_{\Delta x}' M_{\mathbb{S}_x}\mathbb{S}_{\Delta x}\right]^{-1}R_\delta'. \end{aligned} \tag{26}$$

Thus, if one finds an asymptotic behavior of $\hat{\sigma}^2_{0\cdot x}$ under the near-unity regressor in (19), we are able to find a weak limit of Wald and t statistics for the parameters $\gamma = (\beta_0', \delta_0')$. The results are summarized in the following Proposition.

Proposition 2 *Let Assumptions 1 and 2, and the null hypotheses in (15)-(17) hold . Under the fixed-K asymptotics, we have*

(a) $F(\hat{\beta}) \Rightarrow \frac{K}{K-2d} \cdot F_{p_\beta, K-2d}(\| \theta \|^2)$;

(b) $t(\hat{\beta}) \Rightarrow \sqrt{\frac{K}{K-2d}} \cdot t_{K-2d}(\theta)$, *where*

$$\theta = \left[R_\beta\left[\mathbb{S}_x' M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\right]^{-1}R_\beta'\right]^{-1/2} \cdot \left[\frac{cR_\beta\delta_0}{\sigma_{0\cdot x}}\right].$$

(c) $F(\hat{\delta}) \Rightarrow \frac{K}{K-2d} \cdot F_{p_\delta, K-2d}$;

(d) $t(\hat{\delta}) \Rightarrow \sqrt{\frac{K}{K-2d}} \cdot t_{K-2d}$.

In the proof of Proposition 2, we show the asymptotic variance estimate $\hat{\sigma}^2_{0\cdot x}$ for the long-run projected variance $\sigma^2_{0\cdot x}$ weakly converges to χ^2_{K-2d} limiting distribution. Since all other components of test statistics except the bias term $cR_\beta\delta_0$ behave the same way as in the case of the exact unit root regressors, we can capture the effect of the local to unity regressors on the hypothesis tests of β_0 only by looking at the random non-centrality parameter $\| \theta \|^2$in the limiting F and t distributions.

To get some intuition on the random noncentrality parameter, suppose $p = 1$ and testing for a single hypothesis about the cointegration parameter $H_0 : R\beta = r$. Then, the non-centrality parameter $\| \theta \|^2$ becomes

$$\| \theta \|^2 = \frac{c^2}{\sigma^2_{0\cdot x}} \cdot \frac{\| R_\beta \delta_0 \|^2}{\left[R_\beta \Omega_{xx}^{-1/2} \right] \left[\eta_x' M_\xi \eta_x \right]^{-1} \left[\Omega_{xx}^{-1/2} R_\beta' \right]},$$

for a random variable $\eta_x' M_\xi \eta_x$ where $\eta_x = (\eta_{x,1}, ..., \eta_{x,K})'$, $M_\xi = I_K - \xi(\xi'\xi)^{-1}\xi'$, and $\xi = (\xi_1, ..., \xi_K)'$ with

$$\eta_{x,i} := \int_0^1 \phi_i(r) J_c(r) dr, \ \xi := \int_0^1 \phi_i(r) dW(r),$$

for $i = 1, ..., K$. Choose $H = ((R_\beta \Omega_{xx}^{-1/2})' / \left\| R_\beta \Omega_{xx}^{-1/2} \right\|, \tilde{H})'$ for some $\tilde{H}$ such that H is orthogonal, then we can express the denominator by

$$\begin{aligned}
\left[R_\beta \Omega_{xx}^{-1/2} \right] \left[\eta_x' M_\xi \eta_x \right]^{-1} \left[\Omega_{xx}^{-1/2} R_\beta' \right] &= R_\beta \Omega_{xx}^{-1/2} H' (H \left[\eta_x' M_\xi \eta_x \right]^{-1} H') H \Omega_{xx}^{-1/2} R_\beta' \\
&= \left\| R_\beta \Omega_{xx}^{-1/2} \right\|^2 \left[e_d' \left[H \eta_x' M_\xi \eta_x H' \right]^{-1} e_d \right] \\
&\overset{d}{=} R_\beta \Omega_{xx}^{-1} R_\beta' \cdot e_d' \left[\eta_x' M_\xi \eta_x \right]^{-1} e_d,
\end{aligned}$$

where $e_d = (1, 0, ..., 0)' \in \mathbb{R}^d$ and the last equality comes from a rotational invariance property of random vector η_x. With this result and some additional algebra, we can show that the random non-centrality parameter $\| \theta \|^2$ is equivalent in distribution to

$$\|\theta\|^2 \overset{d}{=} c^2 \cdot \left[\frac{\sigma_{0x} \Omega_{xx}^{-1/2}}{\sigma_{0\cdot x}} \right] P_{\Omega_{xx}^{-1/2} R_\beta'} \left[\frac{\Omega_{xx}^{-1/2} \sigma_{x0}}{\sigma_{0\cdot x}} \right] \left[\frac{1}{e_d' \left[\eta_x' M_\xi \eta_x \right]^{-1} e_d} \right],$$

where $P_{\Omega_{xx}^{-1/2} R_\beta'}$ is a projection matrix onto a space spanned by $\Omega_{xx}^{-1/2} R_\beta'$. Since

$$\frac{\sigma_{0x} \Omega_{xx}^{-1} \sigma_{x0}}{\sigma^2_{0\cdot x}} = \frac{r^2}{1 - r^2} \text{ and } r^2 = \frac{\sigma_{0x} \Omega_{xx}^{-1} \sigma_{x0}}{\sigma_0},$$

the random variable $\| \theta \|^2$ is proportional to a long-run correlation vector between $\{u_{0t}\}$ and $\{u_{xt}\}$ projected on to $\Omega_{xx}^{-1/2} R_\beta'$. When $d = p = 1$, the non-random part of $\|\theta\|^2$ is equal to $c^2 \cdot r^2/(1 - r^2)$ so that the degree of overrejection approaches to one when the squared long-run correlation r^2 gets close to one. The quantitative message is consistent with Elliott (1998) whose approximation of cointegration model is based on the time-domain. The presence of non-zero $\| \theta \|^2$ implies that the hypothesis test using the Wald statistics in (16) will tend to over-reject. However, the results in Proposition 2 (c)–(d) indicate we can still perform asymptotically valid Wald and t tests about the long run endogeneity coefficient δ_0. This is expected from our previous investigation on the limit behavior of $\hat{\delta}$ which is not affected by the local to unity regressors.

Although the TA-OLS tests for $\hat{\delta}$ lead asymptotically valid F and t-tests regardless of the dependency properties of the cointegration regressors, the goal of an empirical researcher is making a valid inference for the cointegration vector β_0. Since the presence of the local to unity regressors has an impact on the limiting distributions of $\hat{\beta}$, the result of Proposition 2 (a)–(b) indicates that the corresponding testing procedures no longer have t and F limits and the inferences on β_0 are in danger of severe size distortions. The resulting mixed noncentral F and t limiting distributions in Proposition 2 (a)–(b) shows that the random non-centrality parameter $\| \theta \|^2$ (and θ) depend on the local to unity parameter c, the basis functions, and function of LRV matrix Ω. Given c and TA-OLS estimators $\hat{\delta}$, $\hat{\sigma}^2_{0\cdot x}$, and HAR estimator of LRV $\hat{\Omega}_{xx}$, one can consider the following plug-in estimation of random non-central parameter θ:

$$\hat{\theta} = c\left[R_\beta \hat{\Omega}_{xx}^{-1/2}(\eta_x' M_\xi \eta_x)^{-1}\hat{\Omega}_{xx}^{-1/2} R_{\beta'}\right]^{-1/2}\left[\frac{R_\beta\hat{\delta}}{\hat{\sigma}_{0\cdot x}}\right].$$

With this random variable $\hat{\theta}$, one can simulate the critical values of mixed non-central $F_{p_\beta,K-2d}(\| \hat{\theta} \|^2)$ and $t_{K-2d}(\hat{\theta})$ random variables. These critical values are data-dependent and rely on the knowledge of local to unity parameter c, but similar ideas of plugged-in critical values have been suggested by Cavanagh, Elliott, and Stock (1995) and Campbell and Yogo (2006) in the context of time series regression with local to unity regressor. However, we note that this method of corrected critical value still has some difficulties for practitioners to apply in practice. The issue involves estimation uncertainty of plugged-in parameters $\hat{\delta}$ and $\hat{\sigma}_{0\cdot x}$. The results in this section show that these two parameters weakly converge to random limit and thus inconsistent under the fixed-K asymptotics. Thus, plugging these parameters to critical value simulations without considering these random limits might lead a poor finite sample approximation of underlying test statistics. A similar issue involves with a large estimation uncertainty in $\hat{\Omega}_{xx}$ which is typically estimated nonparametrically, e.g. Newey and West (1987) and Andrews (1991). The nonparametric estimation of $\hat{\Omega}_{xx}$ requires a user choice of weighting (kernel) function and the smoothing parameter. In finite samples, both the kernel function and the bandwidth, especially the latter, do affect the sampling distribution of $\hat{\Omega}_{xx}$ and largely affects the preciseness of associated critical values. See, for example, Kiefer and Vogelsang (2005), Sun, Phillips, and Jin (2008), and Hwang and Sun (2018).

4 Bias Corrected Inferences for β_0 under a near unity regressor

In this section, we provide a method to correct the asymptotic bias of TAOLS test statistics for β_0. Let Γ_c be $p \times 2d$ matrix formed by the hypothesis matrix R_β and the local to unity parameter c.

$$\Gamma_c := \begin{pmatrix} R_\beta & -cR_\beta \end{pmatrix}.$$

Then, under $H_0^\beta : R_\beta\beta_0 = r_\beta$,

$$\begin{aligned}\Gamma_c \Upsilon_T \left[\hat{\gamma} - \gamma_0\right] &= \begin{pmatrix} R_\beta & -cR_\beta \end{pmatrix}\begin{pmatrix} T(\hat{\beta}-\beta_0) \\ \hat{\delta}-\delta_0 \end{pmatrix} \\ &= T\left[R_\beta(\hat{\beta} - c\cdot\hat{\delta}/T) - r\right] + cR_\beta\delta_0.\end{aligned} \tag{27}$$

Using the joint convergence result in Proposition 1 and continuous mapping theorem, we have

$$\begin{aligned}\Gamma_c \Upsilon_T [\hat{\gamma} - \gamma_0] &= T(R_\beta(\hat{\beta} - c \cdot \hat{\delta}/T) - r_\beta) + cR_\beta \delta_0 \\ &\Rightarrow \Gamma_c \begin{bmatrix} c\delta_0 \\ 0 \end{bmatrix} + \Gamma_c (\mathbb{S}_X' \mathbb{S}_X)^{-1} \mathbb{S}_X' \mathbb{S}_{0\cdot x} \\ &\overset{d}{=} MN(cR_\beta \delta_0, \sigma_{0\cdot x}^2 \Gamma_c (\mathbb{S}_X' \mathbb{S}_X)^{-1} \Gamma_c').\end{aligned} \tag{28}$$

Therefore, the plug in estimator of $\hat{\beta}+c\cdot\frac{\hat{\delta}}{T}$ can correct the bias of $c\cdot\frac{\delta_0}{T}$ in the limiting distribution of $T(\hat{\beta}-\beta_0)$, because (28) implies that the limiting distribution of $\hat{\beta}+c\cdot\frac{\hat{\delta}}{T}$ is

$$T(R_\beta(\hat{\beta} - c \cdot \hat{\delta}/T) - r_\beta) \Rightarrow MN(0, \sigma_{0\cdot x}^2 \Gamma_c (\mathbb{S}_X' \mathbb{S}_X)^{-1} \Gamma_c'), \tag{29}$$

which is centered toward zero. It is important to point out the asymptotic variance of the plug-in estimator $\hat{\beta} - c\cdot\hat{\delta}/T$ is no longer the same as that of $T(\hat{\beta}-\beta_0)$. This is because the asymptotic variance of the plug-in estimator has to reflect the estimation uncertainty of $\hat{\delta}$ in its limiting distribution. This motivates us to construct the following modified Wald statistic:

$$\begin{aligned}F(\hat{\beta}; c) &= \frac{T^2}{\hat{\sigma}_{0\cdot x}^2} (R_\beta[\hat{\beta} - c\cdot\hat{\delta}/T] - r_\beta)' \left[\Gamma_c (\Upsilon_T^{-1} \mathbb{W}_X' \mathbb{W}_X \Upsilon_T^{-1})^{-1} \Gamma_c'\right]^{-1} \\ &\quad \times (R_\beta[\hat{\beta} - c\cdot\hat{\delta}/T] - r_\beta)/p.\end{aligned} \tag{30}$$

When $p=1$ and for a one-sided alternative hypothesis, we would construct the modified t statistic as below

$$t(\hat{\beta}; c) = \frac{T(R_\beta[\hat{\beta} - c\cdot\hat{\delta}/T] - r_\beta)}{\sqrt{\hat{\sigma}_{0\cdot x}^2 \Gamma_c (\Upsilon_T^{-1} \mathbb{W}_X' \mathbb{W}_X \Upsilon_T^{-1})^{-1} \Gamma_c'}}. \tag{31}$$

The theorem below establishes the limiting null distributions of $F(\hat{\beta};c)$ and $t(\hat{\beta};c)$ under the fixed-K asymptotics.

Theorem 3 *Under Assumptions 1–2, as $T\to\infty$ but holding K fixed,*

$$F(\hat{\beta}; c) \Rightarrow \frac{K}{K-2d} \cdot F_{p,K-2d},$$

and

$$t(\hat{\beta}; c) \Rightarrow \sqrt{\frac{K}{K-2d}} \cdot t_{K-2d}.$$

The results of theorem indicate one can construct valid t and F tests using the modified t and Wald statistics. The modified statistics not only adjust the locational bias but also reflect the estimation uncertainty of the $\hat{\delta}$ in the bias correction term. After we fully account the effect of the plugged-in bias correction $c\cdot\hat{\delta}/T$ on the modified statistics, we obtain the exact same asymptotic F and t limits. This means one can conveniently implement the modified test statistics using the standard t and F testing methods.

When $p_\beta = 1$, Theorem 3 shows a valid $(1-\alpha)\cdot 100\%$ confidence interval (CI) for the testing parameter $R\beta_0$ can be constructed as

$$CI_{R\beta_0}(c; 1-\alpha) = [\beta_{R,l}^{1-\alpha}(c), \beta_{R,h}^{1-\alpha}(c)], \tag{32}$$

where

$$\beta_{R,l}^{1-\alpha}(c) = R_\beta\left[\hat\beta - \frac{c\hat\delta}{T}\right] - \frac{1}{T}\sqrt{\hat\sigma_{0\cdot x}^2\Gamma_c\left[\Upsilon_T^{-1}\mathbb{W}_X'\mathbb{W}_X\Upsilon_T^{-1}\right]^{-1}\Gamma_c'}\cdot\sqrt{\frac{K}{K-2d}}\cdot t_{K-2d}^{1-\alpha/2},$$
$$\beta_{R,h}^{1-\alpha}(c) = R_\beta\left[\hat\beta - \frac{c\hat\delta}{T}\right] + \frac{1}{T}\sqrt{\hat\sigma_{0\cdot x}^2\Gamma_c\left[\Upsilon_T^{-1}\mathbb{W}_X'\mathbb{W}_X\Upsilon_T^{-1}\right]^{-1}\Gamma_c'}\cdot\sqrt{\frac{K}{K-2d}}\cdot t_{K-2d}^{1-\alpha/2},$$

and t_{K-2d}^{α} is the $(1-\alpha)$ quantile from the $t_{p,K-2d}$ distribution. When the regressors are nearly integrated with $c>0$, the modified confidence interval in (32) shifts the location of the interval up to $-cR\hat\delta/T$. With the location adjustment $-cR\hat\delta/T$, one may come up with the following CI

$$R_\beta\left[\hat\beta - \frac{c\hat\delta}{T}\right] \pm \frac{1}{T}\sqrt{\hat\sigma_{0\cdot x}^2 R_\beta(\mathbb{W}_X'M_{\Delta x}\mathbb{W}_X)^{-1}R_\beta'}\cdot\sqrt{\frac{K}{K-2d}}\cdot t_{K-2d}^{1-\alpha/2}. \tag{33}$$

The common critical value $t_{K-2d}^{1-\alpha/2}$ and estimated variance terms $\hat\sigma_{0\cdot x}^2$ reflect the uncertainty of time series in the (un)modified confidence intervals, but there is notable difference in the margin of errors of two confidence intervals in (32) and (33). With some additional algebra, we can express the term in (32) by

$$\begin{aligned}&\Gamma_c\left[\Upsilon_T^{-1}\mathbb{W}_X'\mathbb{W}_X\Upsilon_T^{-1}\right]^{-1}\Gamma_c'\\ =\;& R_\beta\left[\Lambda_1(c)\left(\mathbb{W}_x'M_{\Delta x}\mathbb{W}_x\right)^{-1} + \Lambda_2(c)\left(\mathbb{W}_{\Delta x}'M_x\mathbb{W}_{\Delta x}\right)^{-1}\right]R_\beta',\end{aligned}$$

where

$$\begin{aligned}\Lambda_1(c) &= T^2(I_d + cT^{-1}\left[\mathbb{W}_{\Delta x}'\mathbb{W}_{\Delta x}\right]^{-1}\mathbb{W}_{\Delta x}'\mathbb{W}_x),\\ \Lambda_2(c) &= c^2 I_d + cT\left[\mathbb{W}_x'\mathbb{W}_x\right]^{-1}\mathbb{W}_x'\mathbb{W}_{\Delta x}.\end{aligned}$$

That is, the measure of uncertainty in the confidence interval (32) is a weight average of standard error terms for $\hat\beta$ and $\hat\delta$ weighted by $\Lambda_1(c) = O_p(T^2)$ and $\Lambda_2(c) = O_p(1)$, respectively. The relative difference in the order of magnitude between these weights is based on the different convergence rates of the variance estimates $(\mathbb{W}_x'M_{\Delta x}\mathbb{W}_x)^{-1} = O_p(T^{-2})$ and $(\mathbb{W}_{\Delta x}'M_x\mathbb{W}_{\Delta x})^{-1} = O_p(1)$ for the estimators $\hat\beta$ and $\hat\delta$, respectively. Interestingly, the weights are functions of the OLS coefficients from the two transformed regressors $\mathbb{W}_x$ and $\mathbb{W}_{\Delta x}$ and the local to unity parameter c.

Lastly, when $c = 0$, i.e. the regressor x_t has an exact unit root, it is easy to check that the above confidence interval of β_0 reduces to the standard form of symmetric confidence interval,

$$R_\beta\hat\beta \pm \sqrt{\hat\sigma_{0\cdot x}^2 R_\beta(\mathbb{W}_X'M_{\Delta x}\mathbb{W}_X)^{-1}R_\beta'}\cdot\sqrt{\frac{K}{K-2d}}\cdot t_{K-2d}^{1-\alpha/2},$$

which is same as the TA-OLS tests in HS (2017).

5 Implementation: Bonferroni Confidence Interval

The near-unity approximation of the modified test statistics $F(\hat{\beta}; c)$ and $t(\hat{\beta}; c)$ requires the knowledge of the local to unity parameter c which is not consistently estimable in general. However, There are several ways developed in the literature on time series to measure the uncertainty of c in the context of unit root testing problem, andone can still construct a nontrivial and informative confidence interval (CI) for the unknown parameter c. See, for example, Stock (1991), Andrews (1993), Elliott and Stock (2001), Mikusheva (2007), and Phillips (2014) for constructing CI of the local to unity parameter c. In this section, we follow Elliott and Stock (1991)'s approach to construct CI $[c_l, c_h]$ for c with $100(1-\varepsilon)\%$ coverage rate where $\varepsilon \leq \alpha$. The detailed procedures are summarized in the Appendix. When $p_\beta = 1$, the Bonferroni CI for $R\beta_0$ can be simply constructed as

$$\begin{aligned} CI^{\mathrm{B}}_{R\beta_0}(c; 1-\alpha) &= \bigcup_{c\in[c_l,c_h]} CI_{R\beta_0}(c; 1-\alpha+\varepsilon) \\ &= \left[\min_{c\in[c_l,c_h]} \beta^{\alpha-\epsilon}_{R,l}(c), \max_{c\in[c_l,c_h]} \beta^{\alpha-\epsilon}_{R,h}(c)\right], \end{aligned}$$

where $\beta^{\alpha-\varepsilon}_{R,l}(c)$ and $\beta^{\alpha-\varepsilon}_{R,h}(c)$ are defined in (32). The idea of constructing the robust Bonferroni confidence interval where there exists unidentified nuisance parameters has been used in various contexts in statistics and econometrics. See, for example, McCloskey (2012) and references therein. By Bonferroni's inequality, the above Bonferroni CI yields a confidence level of at least $100(1-\alpha)\%$. Since the infeasible confidence interval $[\beta^{\alpha-\varepsilon}_{R,l}(c), \beta^{\alpha-\varepsilon}_{R,h}(c)]$ depends on c only through $T^{-1}c\hat{\delta}$ and $\Gamma(c)$, the computational cost of finding is low when one searches the maximum (and minimum) of $\beta^{\alpha-\varepsilon}_{R,u}(c)$ (and $\beta^{\alpha-\varepsilon}_{R,l}(c)$) over $[c_l, c_h]$.

One potential limitation to the Bonferroni based method is that the CI is often too wide, and the resulting coverage rate is usually higher than the nominal one. One convenient way is to consider the union of $CI_{R\beta_0}(c; 1-\alpha)$ for nominal α-test, instead of $CI_{R\beta_0}(c; 1-\alpha+\epsilon)$, over $[c_l, c_h]$ which is $100(1-\epsilon)\%$ coverage rate. For the choice of ϵ, we follow Sun (2014b) and set $\epsilon = 0.10$ in our Monte Carlo simulations. There can be several ways to improve the performance of the Bonferroni based method. See Cavanagh, Elliott, and Stock (1995) and McCloskey (2017) for more details.

6 Monte Carlo Evidences

We evaluate the performance of our modified TA-OLS method, presented in the previous section, in a finite sample. We compare it with several other methods, including the unmodified TA-OLS approach in HS (2017). For cointegration models, we consider the following bivariate cointegration model as in Phillips (2014) and HS (2017)

$$\begin{matrix} y_t = \alpha_0 + x_t\beta_0 + u_{0t} \\ x_t = \rho_T x_{t-1} + u_{xt} \end{matrix}, u_t = \begin{pmatrix} u_{0t} \\ u_{xt} \end{pmatrix} = \Theta u_{t-1} + \epsilon_t, \quad (34)$$

with a local to unity coefficient $\rho_T = 1 - \frac{c}{T}$ and

$$\epsilon_t = \left(\begin{array}{c} \epsilon_{0t} \\ \epsilon_{xt} \end{array} \right) \sim \text{i.i.d } N\left(0, \Sigma\right), \ \Theta = \psi \cdot I_2, \ \Sigma = J_{2,2} \cdot \phi + I_2 \cdot (1 - \phi),$$

and $J_{2,2}$ is the 2×2 matrix of ones. The true cointegration regression coefficients are set to be $\alpha_0 = 1$ and $\beta_0 = 1$. The initial value of the error process u_t drawn from standard normal distribution. To minimize the initialization effect, we generate a time series of length $2T$ and drop the first T observations. The parameter ψ controls the persistence of individual components in $u_t = (u_{0t}, u'_{xt})' \in \mathbb{R}^{d+1}$. We set the values of ψ as $\{0.50, 0.75\}$, so the stationary cointegration error u_t reasonably persistent. The parameter ϕ is a pairwise correlation coefficient between the elements of u_t and characterizes the degree of endogenity. With some algebraic manipulations, it is straightforward to obtain the LRV Ω of u_t as

$$\Omega = \left(I_{d+1} - \Theta\right)^{-1} \Sigma \left(I_{d+1} - \Theta\right)^{-1\prime} = \left(\frac{1}{1-\psi}\right)^2 \cdot \left(\begin{array}{cc} 1 & \phi \cdot J_{1,d} \\ \phi \cdot J_{d,1} & J_{d,d} \cdot \phi + I_d \cdot (1-\phi) \end{array} \right).$$

It then follows that the squared long run correlation is expressed by

$$r^2 = \frac{\sigma_{0x}\Omega_{xx}^{-1}\sigma_{x0}}{\sigma_0^2} = \frac{d\phi^2}{(1-\phi) + d\phi}.$$

Using the formula above, we set ϕ to satisfy $r^2 \in \{0, 0.15,\ 0.25,\ 0.35,\ 0.50,\ 0.75\}$. For the autoregressive coefficient of the cointegration regressor x_t, we take $\rho_T \in \{1,\ 0.975,\ 0.95,\ 0.925\}$ with sample size $T = 200$, so the corresponding local to unity parameters are $c \in \{0, 5, 10, 15\}$.

The null hypotheses of interest for the true parameter is

$$H_0 : \beta_0 = 1 \text{ and } H_0 : \beta_0 \neq 1.$$

Also, in the case of $r^2 = 0$, we test the long run endogenity parameter with the following null hypothesis of

$$H_0 : \delta_0 = 0 \text{ and } H_0 : \delta_0 \neq 0.$$

We consider Fourier basis functions given in (5) for our TAOLS framework, as the same numerical evidences hold for the cosine transformation in (6). See Section 6 of HS (2017) for the detail. For fixed values of K, we set $K = 8$ and $K = 16$ for the AR(1) parameters $\psi = 0.75$ and $\psi = 0.50$, respectively. These choices of K are shown to have good finite sample performances in various literature of fixed smoothing asymptotics with extensive numerical experiments. See, for example, Müller and Watson (2013, 2017), HS (2017), and Lazarus, Lewis, Stock, and Watson (2018).

We first examine the empirical size of four different type of TAOLS t-tests at nominal size $\alpha = 5\%$. The first test is the unmodified TA-OLS test considered in HS (2017) and Phillips (2014). As a second group of tests, we consider two infeasible TA-OLS t-tests which treat the true local to unity parameter c as known: the first one is the plugged-in type modification of TA-OLS in (33), and the second one is the modified TA-OLS t-test in (31) and (32). It is imporant to point out that the plugged-in TA-OLS test only shifts the location of the confidence interval by $c\hat{\delta}/T$, whereas the modified TA-OLS fully accounts for the asymptotic uncertainty of plugged-in estimator $\hat{\delta}$ as well as the bias correction term $c\hat{\delta}/T$. All of these three tests employ the same

t_{K-2} critical values. The last test we consider is a feasible version of the modified TA-OLS test which is based on the Bonferroni correction (Bonferroni TA-OLS, hereafter) in Section 5.

Figures 1–2 and Tables 1–2 report the empirical size of the four different TAOLS tests. The number of simulation replications is $10,000$. The results are summarized below.

From Table 1, in the unit root case, that is $c = 0$, the first three TA-OLS tests have empirical sizes very close to the nominal size of 5%. This is not surprising given that the unmodified TA-OLS asymptotically valid under the exact unit-root assumption. The plugged-in TAOLS and the modified TAOLS are numerically equivalent to the unmodified TAOLS when $c = 0$. The Bonferroni TA-OLS yields to a correct empirical size although it is mildly undersized varying from 3.2% to 4.2%. While the feasible Bonferroni TA-OLS test is conservative, the conservatism only comes from the Bonferroni step as we check the infeasible modified TA-OLS provides very accurate size control.

Second, as c deviates from zero, the unmodified TA-OLS suffers from severe size distortions, especially, when the squared long-run correlation, r^2, and the local to unity parameter, c, grow. This confirms our theoretical results in Proposition 2. The plugged-in TA-OLS test improves the size distortion of the unmodified TA-OLS, but still has empirical rejection rates greater than the nominal size. For example, when $\psi = 0.75$ in Table 1, the empirical rejection rate of the plugged-in TAOLS is between $8\% \sim 17\%$ for $c \in \{5, 10, 15\}$. This is because there is a large amount of finite sample noise in $\hat{\delta}$ which is $O_p(1)$ in our fixed-K asymptotics. We find that our modified TA-OLS test, equipped with both the bias correction and the variance adjustment, has accurate finite sample size for all values of r^2 and c considered. Lastly, the Bonferroni TA-OLS test has correct size, varying from 1.8% to 6.2%, and also it is conservative in finite sample as expected.

To sum up, first there is a large amount of size distortions for the unmodified TA-OLS in the local to unity case when r^2 and c deviate from zero. Second, treating c as known, our modified TA-OLS successfully corrects the size distortions of the unmodified TA-OLS. When c is unknown, the feasible version of our modified TAOLS with Bonferroni correction is also size-corrected but is mildly undersized. The conservatism is expected to result in power loss compared to the modified TA-OLS test which is infeasible in practice. To investigate the power loss of the feasible Bonferroni TAOLS procedure, we compare its finite sample power with the modified TA-OLS. The true parameter of cointegration is now from the local alternative hypothesis $\beta = \beta_0 + b/T$, where $b \in [-25, 25]$ measures the magnitude of the local departure. All other DGPs are the same as before.

Figures 3–5 present the finite sample power curve of each procedure for $r^2 \in \{0.15, 0.35, 0.50, 0.75\}$ and $c \in \{5, 10, 15\}$ when $\psi = 0.75$. In all cases, the power of the infeasible modified TA-OLS test outperforms the feasible Bonferroni TA-OLS. However, the modified TA-OLS test is infeasible when c is unknown. The cost of lack of knowledge of c is reflected on the relative power loss of the Bonferroni TA-OLS test. Figures 3–5 indicate that the power loss is increasing in the squared long-run correlation r^2, but we also show that the power of the Bonferroni TA-OLS test increases along with $|b|$.

Lastly, our results also indicate that we can precisely perform the endogenity test, i.e., a test of whether $\delta_0 = 0$, regardless of the local to -unity parameters c. This is consistent with our fixed-K asymptotic results in Proposition 2 (c)-(d) which indicate that, in the presence of the local to unity regressor, $\hat{\delta}$ is still asymptotically centered toward its true value and yields a robust test for the long-run endogenity parameter δ_0.

7 Conclusion

In this paper, we develop a theory that adopts a local to unity approximations to a triangular cointegrated system. Our analysis is carried on the domain of low frequencies by transforming data from the original time domain. Instead of maintaining a strict dichotomy between integrated and non-integrated regressors, our assumption of the local to unity regressor allows for a smoother transition between the two processes. It thus can provide a more reasonable approximation to the low frequency transformed methods.

We show that the unmodified TA-OLS in HS (2017) approach is still consistent and share a common mixture of normal distribution in the limit. However, the unmodified TA-OLS possesses an asymptotic bias term in the limiting distribution. As a result, the unmodified TA-OLS suffers from severe size distortions, especially, when the degree of long run endogeneity grows, or the cointegration regressor deviates from the exact unit root.

We develop a modified TA-OLS test statistics, which yields to a convenient t and F inferences for the cointegrating vector and long run endogeneity parameter. The modified TA-OLS not only adjusts for the asymptotic bias arising from the local to unity regressor but also corrects for the uncertainty of the plugged-in bias correction term. When the local to unity parameter is unknown, we also provide a feasible version of our modified TA-OLS, which considers a Bonferroni correction. Our Monte Carlo analysis shows that the Bonferroni TA-OLS test is size-corrected and mildly undersized. Also, there is an increasing power gain of the Bonferroni TA-OLS test when the true parameter departs from the value of the parameter under the null hypothesis.

References

[1] Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–854.

[2] Andrews, D.W.K. (1993). Exactly median-unbiased estimation of first order autoregressive/unit root models. *Econometrica*, 61(1), 139-165. doi:10.2307/2951781

[3] Bobkoski, M. J. (1983). Hypothesis testing in nonstationary time series, Unpublished Ph.D. Thesis, Department of Statistics, University of Wisconsin.

[4] Cavanagh, C. (1985). Roots local to unity, Manuscript, Department of Economics, Harvard University.

[5] Cavanagh, C. L., Elliott, G., Stock, J. H. (1995). Inference in models with nearly integrated regressors. *Econometric theory*, 11(5), 1131-1147.

[6] Campbell, J. Y., Yogo, M. (2006). Efficient tests of stock return predictability. *Journal of Financial Economics*, 81(1), 27-60.

[7] Bierens, H. J. (1997). Nonparametric cointegration analysis. *Journal of Econometrics* 77(2), 379–404.

[8] Christiano, L. J., Eichenbaum, M. (1990). Unit roots in real GNP: Do we know, and do we care?. In Carnegie-Rochester Conference Series on Public Policy (Vol. 32, pp. 7-61). North-Holland.

[9] Elliott, G. (1998). On the robustness of cointegration methods when regressors almost have unit roots. *Econometrica*, 66(1), 149-158. doi:10.2307/2998544

[10] Elliott, G., Stock, J. H. (2001). Confidence intervals for autoregressive coefficients near one. *Journal of Econometrics*, 103(1-2), 155-181.

[11] Engle, R. F., Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, 251-276.

[12] Hwang, J., Sun, Y. (2017). Simple, robust, and accurate F and t tests in cointegrated systems. *Econometric Theory*, 1-36.

[13] Hwang, J. and Sun, Y. (2018). Should we go one step further? An accurate comparison of one-step and two-step procedures in a generalized method of moments framework. Journal of econometrics, 207(2), pp.381-405.

[14] Kiefer, N.M., Vogelsang, T.J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory* 21, 1130–1164.

[15] Lazarus, E., Lewis, D.J. and Stock, J.H., (2019). The size-power tradeoff in HAR inference. Available at SSRN 3436372.

[16] Lazarus, E., Lewis, D.J., Stock, J.H. and Watson, M.W., (2018). HAR Inference: Recommendations for Practice. *Journal of Business & Economic Statistics*, 36(4), pp.541-559.

[17] McCloskey, A. (2017). Bonferroni-based size-correction for nonstandard testing problems. *Journal of Econometrics*, 200(1), pp.17-35.

[18] Mikusheva, A. (2007). Uniform inference in autoregressive models. *Econometrica*, 75(5), 1411-1452.

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

[20] Müller, U. K. (2014). HAC corrections for strongly autocorrelated time series. *Journal of Business & Economic Statistics* 32(3), 311–322.

[21] Müller, U.K. and Watson, M.W., (2008). Testing models of low-frequency variability. *Econometrica*, 76(5), pp.979-1016.

[22] Müller, U. K., Watson, M.W. (2013). Low-frequency robust cointegration testing. *Journal of Econometrics* 174(2), 66–81.

[23] Müller, U., & Watson, M. (2017). Low-Frequency Econometrics. In B. Honoré, A. Pakes, M. Piazzesi, & L. Samuelson (Eds.), Advances in Economics and Econometrics: Eleventh World Congress (Econometric Society Monographs, pp. 53-94). Cambridge: Cambridge University Press. doi:10.1017/9781108227223.003

[24] Müller, U. K., Watson, M. W. (2018). Long-Run Covariability. *Econometrica*, 86(3), 775-804.

[25] Newey, W., West, K. (1987). A simple positive definite, heteroscedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, 55, 103080.

[26] Phillips, P.C., (1987). Towards a unified asymptotic theory for autoregression. *Biometrika*, 74(3), pp.535-547.

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

[28] Phillips, P.C.B. (2014). Optimal estimation of cointegrated systems with irrelevant instruments, *Journal of Econometrics* 178(2), 210–224.

[29] Phillips, P.C.B., Durlauf, S.N. (1986). Multiple regression with integrated processes. *Review of Economic Studies* 53, 473–496.

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

[31] Stock, J. (1991). Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series, *Journal of Monetary Economics*, 28, issue 3, p. 435-459.

[32] Stoica, P., Moses, R. (2005). *Spectral Analysis of Signals.* Pearson: New Jersey.

[33] Sun, Y. (2013). A heteroskedasticity and autocorrelation robust F test using an orthonormal series variance estimator. The Econometrics Journal, 16(1), 1-26.

[34] Sun, Y., (2014). HAC Corrections for Strongly Autocorrelated Time Series Comment. *Journal of Business & Economic Statistics*, 32:3, 311-322.

[35] Sun, Y., Phillips, P.C.B., Jin, S. (2008). Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing. *Econometrica* 76(1), 175–94.

[36] Thomson, D.J. (1982). Spectrum estimation and harmonic analysis. *IEEE Proceedings* 70, 1055–1096.

Table 1: Empirical size of 5% various TA-OLS methods with T=200, K=8 and AR(1) error with $\psi = 0.75$.

$\psi = 0.75$ and $K = 8$

$c = 0$	Unmodified	Plugged-in (Infeasible)	Modified (Infeasible)	Bonferroni
0	0.055	0.055	0.055	0.040
0.15	0.058	0.058	0.058	0.042
0.25	0.056	0.056	0.056	0.039
0.35	0.053	0.053	0.053	0.037
0.50	0.054	0.054	0.054	0.036
0.75	0.050	0.050	0.050	0.032
$c = 5$	Unmodified	Plugged-in (Infeasible)	Modified (Infeasible)	Bonferroni
0	0.052	0.079	0.054	0.029
0.15	0.075	0.074	0.051	0.027
0.25	0.090	0.076	0.050	0.021
0.35	0.117	0.078	0.053	0.021
0.50	0.166	0.080	0.055	0.019
0.75	0.367	0.081	0.054	0.018
$c = 10$	Unmodified	Plugged-in (Infeasible)	Modified (Infeasible)	Bonferroni
0	0.052	0.112	0.053	0.026
0.15	0.097	0.115	0.050	0.025
0.25	0.142	0.125	0.056	0.026
0.35	0.190	0.114	0.049	0.022
0.50	0.299	0.120	0.055	0.027
0.75	0.624	0.120	0.049	0.032
$c = 15$	Unmodified	Plugged-in (Infeasible)	Modified (Infeasible)	Bonferroni
0	0.048	0.158	0.047	0.022
0.15	0.114	0.166	0.049	0.029
0.25	0.176	0.173	0.054	0.032
0.35	0.251	0.171	0.053	0.034
0.50	0.393	0.170	0.057	0.041
0.75	0.757	0.170	0.054	0.062

Table 2: Empirical size of 5% various TA-OLS methods with T=200, K=16 and AR(1) error with $\psi = 0.50$.

$\psi = 0.50$ and $K = 16$

$c = 0$	Unmodified	Plugged-in (Infeasible)	Modified (Infeasible)	Bonferroni
0	0.053	0.053	0.053	0.042
0.15	0.052	0.052	0.052	0.038
0.25	0.056	0.056	0.056	0.038
0.35	0.053	0.053	0.053	0.036
0.50	0.052	0.052	0.052	0.033
0.75	0.053	0.053	0.053	0.033
$c = 5$	Unmodified	Plugged-in (Infeasible)	Modified (Infeasible)	Bonferroni
0	0.058	0.072	0.056	0.032
0.15	0.079	0.068	0.053	0.025
0.25	0.107	0.071	0.054	0.017
0.35	0.133	0.063	0.049	0.014
0.50	0.206	0.071	0.053	0.012
0.75	0.458	0.072	0.056	0.014
$c = 10$	Unmodified	Plugged-in (Infeasible)	Modified (Infeasible)	Bonferroni
0	0.054	0.092	0.052	0.026
0.15	0.118	0.091	0.051	0.022
0.25	0.174	0.093	0.052	0.018
0.35	0.250	0.092	0.054	0.017
0.50	0.399	0.095	0.053	0.017
0.75	0.783	0.096	0.055	0.023
$c = 15$	Unmodified	Plugged-in (Infeasible)	Modified (Infeasible)	Bonferroni
0	0.054	0.115	0.053	0.029
0.15	0.152	0.117	0.052	0.024
0.25	0.235	0.122	0.054	0.025
0.35	0.355	0.125	0.057	0.026
0.50	0.538	0.117	0.056	0.024
0.75	0.898	0.125	0.055	0.035

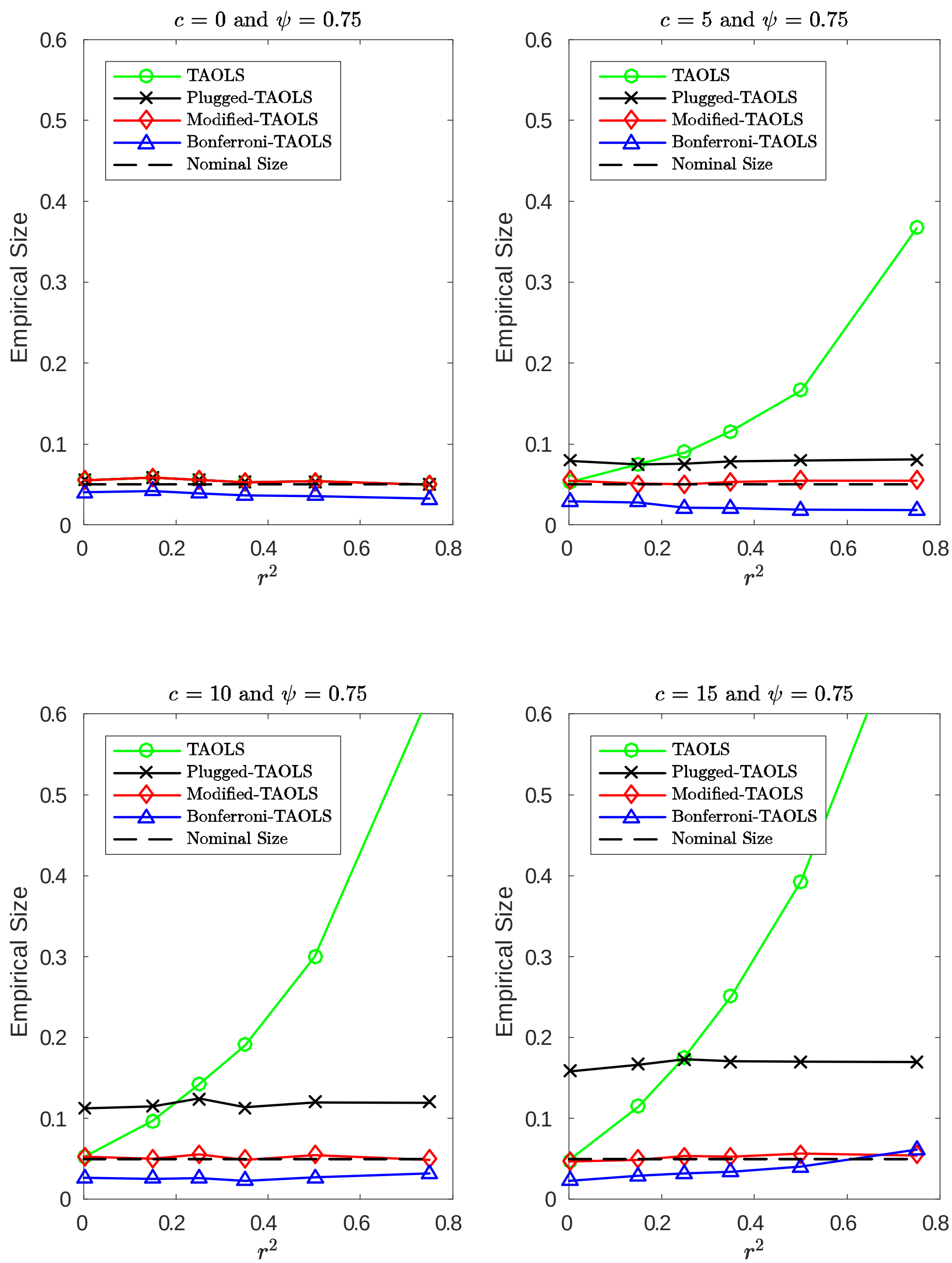

Figure 1: Empirical size of 5% fixed-smoothing tests (TAOLS, Infeasible Plugged-in TAOLS, Infeasible Modified TAOLS, and Feasible Bonferrnoi Modified TAOLS) with $K = 8$ and AR(1) error with $\psi = 0.75$.

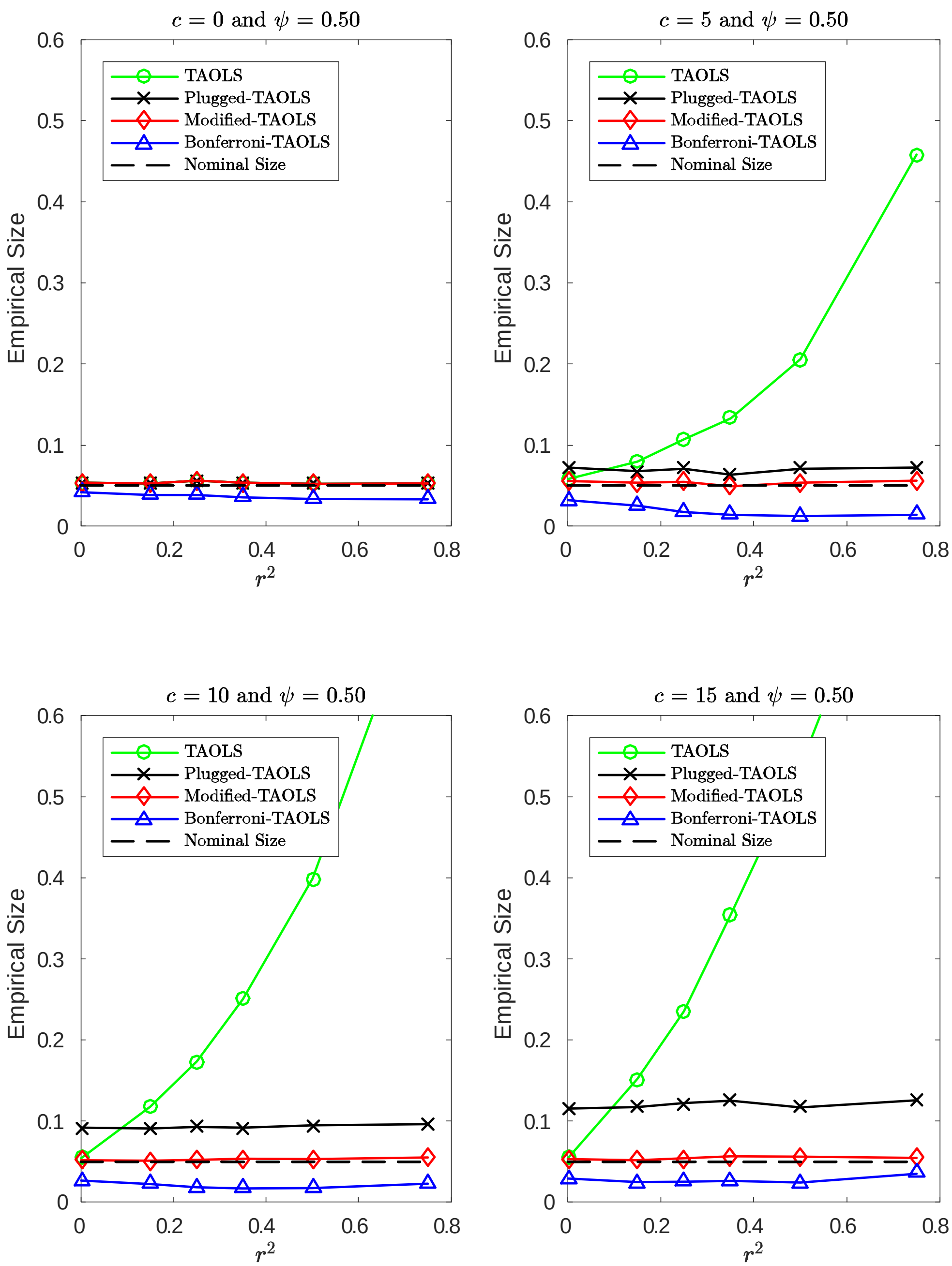

Figure 2: Empirical size of 5% fixed-smoothing tests (TAOLS, Infeasible Plugged-in TAOLS, Infeasible Modified TAOLS, and Feasible Bonfernoi Modified TAOLS) with $K = 16$ and AR(1) error with $\psi = 0.50$.

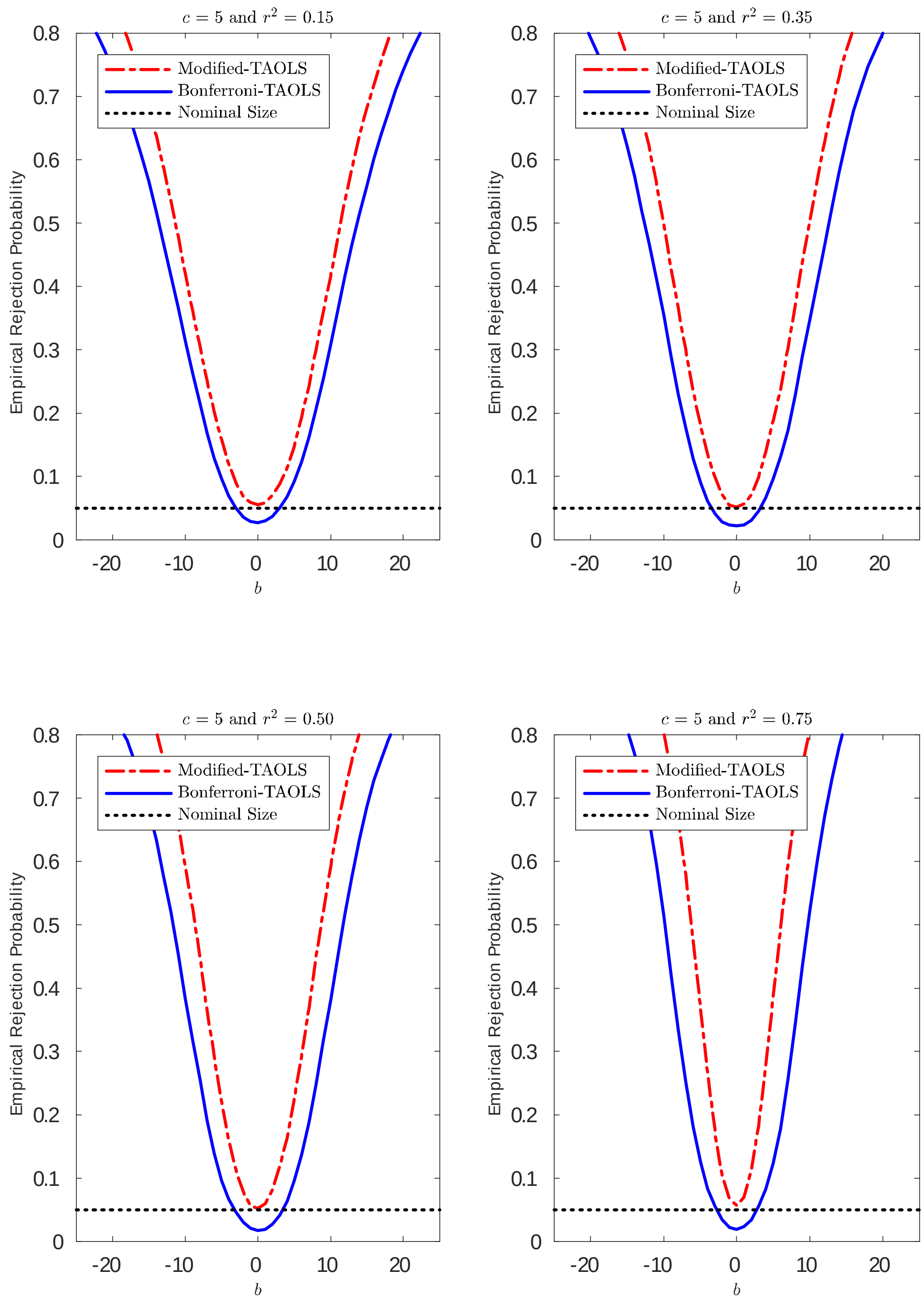

Figure 3: Finite sample power of 5% fixed-smoothing tests (Modified TAOLS and Feasible Bonferrnoi Modified TAOLS) with $K = 8$, $c = 5$, and AR(1) error with $\psi = 0.75$.

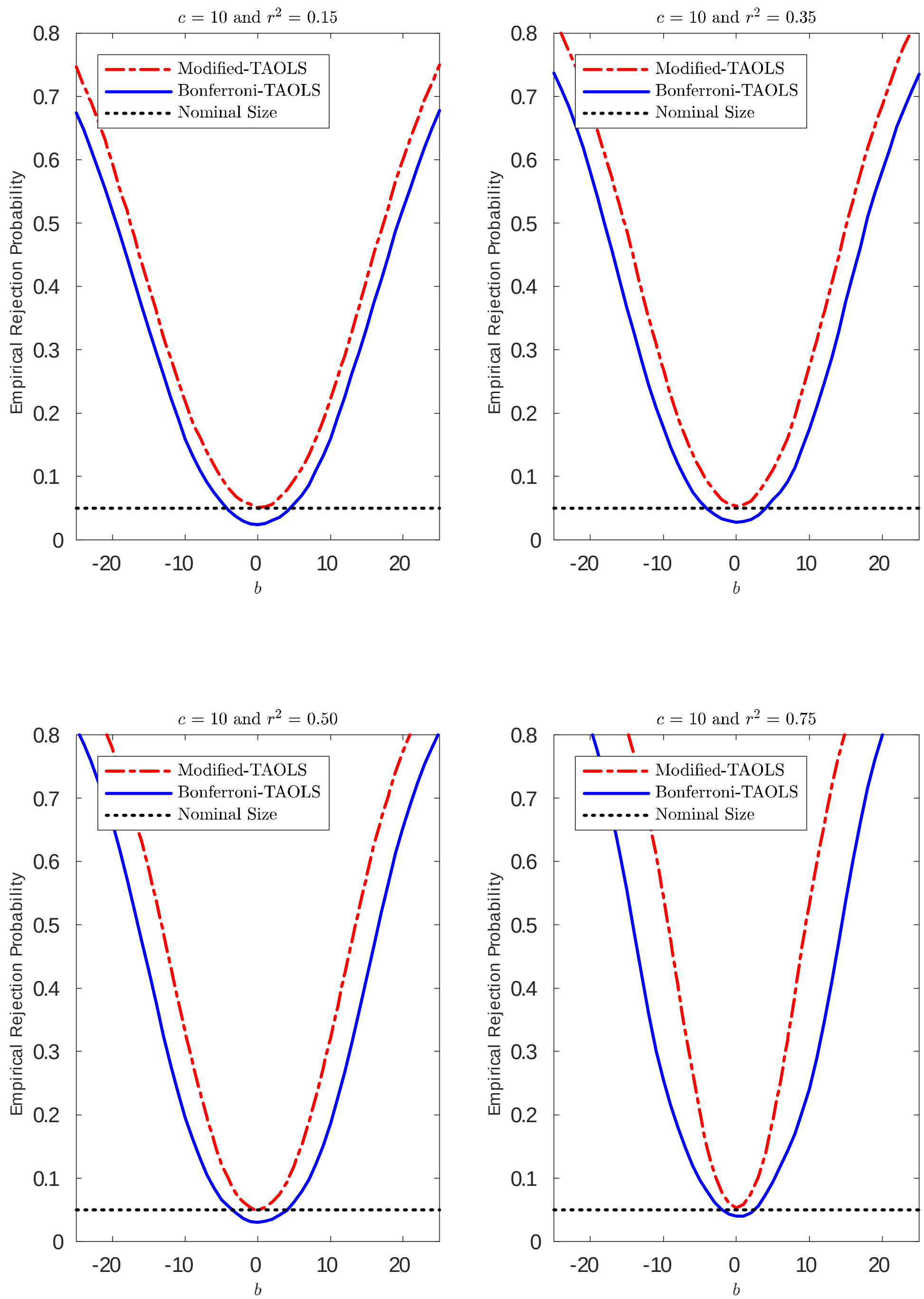

Figure 4: Finite sample power of 5% fixed-smoothing tests (Modified TAOLS and Feasible Bonferrnoi Modified TAOLS) with $K = 8$, $c = 10$, and AR(1) error with $\psi = 0.75$.

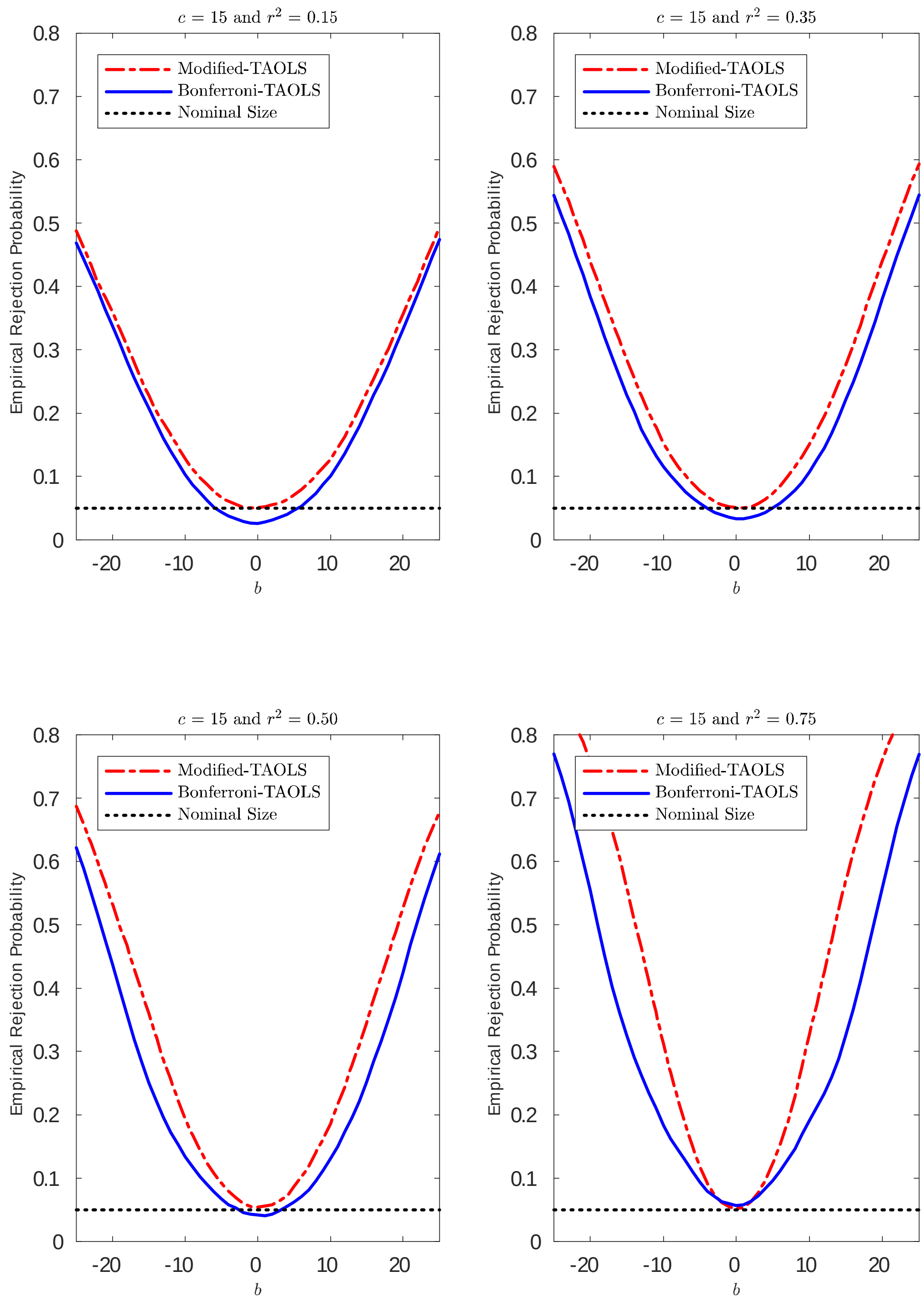

Figure 5: Finite sample power of 5% fixed-smoothing tests (Modified TAOLS and Feasible Bonferrnoi Modified TAOLS) with $K = 8$, $c = 15$, and AR(1) error with $\psi = 0.75$.

8 Appendix of Proofs

Proof of Proposition 1. We begin by showing the asymptotic equivalence between $\frac{1}{T^{3/2}}\sum_{t=1}^{T} x_{t-1}\phi_i(\frac{t}{T})$ and the transformed regressor $\mathbb{W}_x/T$ in (24), that is,

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T} x_{t-1}\phi_i\left(\frac{t}{T}\right) = \frac{1}{T^{3/2}}\sum_{t=1}^{T} x_t\phi_i\left(\frac{t}{T}\right) + O_p\left(\frac{1}{T}\right).$$

The left side of equation is

$$\frac{1}{T}\sum_{t=1}^{T}\frac{x_{t-1}}{\sqrt{T}}\phi_i\left(\frac{t}{T}\right) = \frac{1}{T}\sum_{s=0}^{T-1}\frac{x_s}{\sqrt{T}}\phi_i\left(\frac{s}{T}\right) + \frac{1}{T}\sum_{t=1}^{T}\frac{x_{t-1}}{\sqrt{T}}\left[\phi_i\left(\frac{t}{T}\right) - \phi_i\left(\frac{t-1}{T}\right)\right]. \tag{35}$$

By mean value theorem,

$$\phi_i\left(\frac{t}{T}\right) = \phi_i\left(\frac{t-1}{T}\right) + \phi'(r_t^*)\left(\frac{1}{T}\right) \text{ for some } r_t^* \in \left[\frac{t-1}{T}, \frac{t}{T}\right],$$

and Assumption 2 yields

$$\phi_i\left(\frac{t}{T}\right) - \phi_i\left(\frac{t-1}{T}\right) = \frac{\phi'(r_t^*)}{T} \leq \frac{M}{T}$$

for some $M > 0$ uniformly over t. Therefore, the second term in (35) satisfies

$$\frac{1}{T}\sum_{t=1}^{T}\frac{x_{t-1}}{\sqrt{T}}\left[\phi_i\left(\frac{t}{T}\right) - \phi_i\left(\frac{t-1}{T}\right)\right] \leq \left(\frac{M}{T}\right)\left[\frac{1}{T}\sum_{t=0}^{T-1}\frac{x_t}{\sqrt{T}}\right] = O_p\left(\frac{1}{T}\right).$$

For the first term in (35),

$$\begin{aligned}\frac{1}{T}\sum_{s=0}^{T-1}\frac{x_s}{\sqrt{T}}\phi_i\left(\frac{s}{T}\right) &= \frac{1}{T}\sum_{s=1}^{T}\frac{x_s}{\sqrt{T}}\phi_i\left(\frac{s}{T}\right) + \frac{x_0}{T^{3/2}}\phi_i(0) - \frac{x_T}{T^{3/2}}\phi_i(1) \\ &= \frac{1}{T}\sum_{s=1}^{T}\frac{x_s}{\sqrt{T}}\phi_i\left(\frac{s}{T}\right) + O_p\left(\frac{1}{T}\right),\end{aligned} \tag{36}$$

where the second equality follows from $x_0 = O_p(1)$ and the equation (22). With this result and the weak convergences in (9), (10), and (11), we get

$$\begin{aligned}\Upsilon_T^{-1}\mathbb{W}_X &= (\mathbb{W}_x/T, \mathbb{W}_{\Delta x}) \Rightarrow \mathbb{S}_X = (\mathbb{S}_x, \mathbb{S}_{\Delta x}), \\ \tilde{\mathbb{W}}_x &\Rightarrow \mathbb{S}_{0\cdot x} + c\cdot\mathbb{S}_x\delta_0,\end{aligned} \tag{37}$$

where $\tilde{\mathbb{W}}_x = (\tilde{\mathbb{W}}_{x,1}, ..., \tilde{\mathbb{W}}_{x,K})'$.Then, by the definition of $\hat{\gamma}$ and Υ_T, we have

$$\begin{aligned}\Upsilon_T(\hat{\gamma} - \gamma_0) &= (\Upsilon_T^{-1}\mathbb{W}_X'\mathbb{W}_X\Upsilon_T^{-1})^{-1}\Upsilon_T^{-1}\mathbb{W}_X'\tilde{\mathbb{W}}_{0\cdot x} \\ &\Rightarrow \left(\mathbb{S}_X'\mathbb{S}_X\right)^{-1}\mathbb{S}_X'\left[\mathbb{S}_{0\cdot x} + c\cdot\mathbb{S}_x\delta_0\right] \\ &= \left(\mathbb{S}_X'\mathbb{S}_X\right)^{-1}\mathbb{S}_X'\mathbb{S}_{0\cdot x} + c\cdot\left(\mathbb{S}_X'\mathbb{S}_X\right)^{-1}\mathbb{S}_X'\mathbb{S}_x\delta_0.\end{aligned}$$

Since $\mathbb{S}_{0\cdot x,i}$s are i.i.d normal random variable with variance $\sigma^2_{0\cdot x}$ over $i = 1, ..., K$ and independent with $\mathbb{S}_X = (\mathbb{S}_x, \mathbb{S}_{\Delta x})$, the latter component can be expressed by a mixture of normal distribution

$$MN(0, \sigma^2_{0\cdot x}\left(\mathbb{S}'_X\mathbb{S}_X\right)^{-1}).$$

The second component can be written more explicitly as

$$\begin{aligned} c\cdot\left(\mathbb{S}'_X\mathbb{S}_X\right)^{-1}\mathbb{S}'_X\mathbb{S}_x\delta_0 &= c\cdot\begin{pmatrix} \mathbb{S}'_x\mathbb{S}_x & \mathbb{S}'_x\mathbb{S}_{\Delta x} \\ \mathbb{S}'_{\Delta x}\mathbb{S}_x & \mathbb{S}'_{\Delta x}\mathbb{S}_{\Delta x}\end{pmatrix}^{-1}\begin{pmatrix}\mathbb{S}'_x\mathbb{S}_x\delta_0 \\ \mathbb{S}_{\Delta x}\mathbb{S}_x\delta_0\end{pmatrix} \\ = \begin{pmatrix} c\cdot\left(\mathbb{S}'_x M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\right)^{-1}\mathbb{S}'_x M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\delta_0 \\ c\cdot\left(\mathbb{S}'_{\Delta x} M_{\mathbb{S}_x}\mathbb{S}_{\Delta x}\right)^{-1}\mathbb{S}'_{\Delta x} M_{\mathbb{S}_x}\mathbb{S}_x\delta_0\end{pmatrix} &= \begin{pmatrix} c\delta_0 \\ 0\end{pmatrix},\end{aligned}$$

which finishes the proof. ■

Proof of Proposition 2. We prove the result for the F statistic only, as the result for t statistic can be proved in a similar manner. Note that

$$\begin{aligned}\hat{\sigma}^2_{0\cdot x} &= \frac{1}{K}\sum_{i=1}^{K}\hat{\omega}^2_{0\cdot x,i} = \frac{1}{K}\mathbb{W}'_Y\left[I_K - \mathbb{W}_X(\mathbb{W}'_X\mathbb{W}_X)^{-1}\mathbb{W}'_X\right]\mathbb{W}_Y \qquad (38) \\ &= \frac{1}{K}\tilde{\mathbb{W}}'_{0\cdot x}\left[I_K - \mathbb{W}_X(\mathbb{W}'_X\mathbb{W}_X)^{-1}\mathbb{W}'_X\right]\tilde{\mathbb{W}}_{0\cdot x} \\ &\Rightarrow \frac{1}{K}\left[\mathbb{S}_{0\cdot x} + c\cdot\mathbb{S}_x\delta_0\right]'\left[I_K - \mathbb{S}_X(\mathbb{S}'_X\mathbb{S}_X)^{-1}\mathbb{S}'_X\right]\left[\mathbb{S}_{0\cdot x} + c\cdot\mathbb{S}_x\delta_0\right].\end{aligned}$$

Since $P_{\mathbb{S}_X} = \mathbb{S}_X(\mathbb{S}'_X\mathbb{S}_X)^{-1}\mathbb{S}'_X$ is a projection matrix onto a space generated by $[\mathbb{S}_x, \mathbb{S}_{\Delta x}]$, it is easy to check

$$\left[I_K - \mathbb{S}_X(\mathbb{S}'_X\mathbb{S}_X)^{-1}\mathbb{S}'_X\right]\left[c\cdot\mathbb{S}_x\delta_0\right] = 0.$$

Therefore, the weak convergence limit of the estimator $\hat{\sigma}^2_{0\cdot x}$ simplifies to

$$\hat{\sigma}^2_{0\cdot x} \Rightarrow \frac{1}{K}\mathbb{S}'_{0\cdot x}M_{\mathbb{S}_X}\mathbb{S}_{0\cdot x} \sim \frac{\sigma^2_{0\cdot x}}{K}\chi^2_{K-2d},$$

where $M_{\mathbb{S}_X} := I_K - \mathbb{S}_X(\mathbb{S}'_X\mathbb{S}_X)^{-1}\mathbb{S}'_X$. Combining this result with

$$T(R_\beta\hat{\beta} - r_\beta) \Rightarrow R_\beta\left[\mathbb{S}'_x M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\right]^{-1}\mathbb{S}'_x M_{\mathbb{S}_{\Delta x}}\mathbb{S}_{0\cdot x} + cR_\beta\delta_0$$

and

$$R_\beta\left[(\mathbb{W}'_x/T)M_{\Delta x}(\mathbb{W}'_x/T)\right]^{-1}R'_\beta \Rightarrow R_\beta\left[\mathbb{S}'_x M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\right]^{-1}R'_\beta,$$

we get

$$F(\hat{\beta}) \Rightarrow \frac{K}{p_\beta}\frac{\left\|\frac{Z}{\sigma_{0\cdot x}} + \left[R_\beta\left[\mathbb{S}'_x M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\right]^{-1}R'_\beta\right]^{-1/2}\cdot\left[\frac{cR_\beta\delta_0}{\sigma_{0\cdot x}}\right]\right\|^2}{\left[\frac{\mathbb{S}'_{0\cdot x}M_{\mathbb{S}_X}\mathbb{S}_{0\cdot x}}{\sigma^2_{0\cdot X}}\right]}, \qquad (39)$$

where

$$Z = \left[R_\beta\left[\mathbb{S}'_x M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\right]^{-1}R_\beta\right]^{-1/2}R_\beta\left[\mathbb{S}'_x M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\right]^{-1}\mathbb{S}'_x M_{\mathbb{S}_{\Delta x}}\mathbb{S}_{0\cdot x} \sim N(0, \sigma^2_{0\cdot x}\cdot I_K).$$

Conditional on $\mathbb{S}_X = (\mathbb{S}_x, \mathbb{S}_{\Delta x})$, $M_{\mathbb{S}_X}\mathbb{S}_{0\cdot x}$ and $\mathbb{S}_x M_{\mathbb{S}_{\Delta x}}\mathbb{S}_{0\cdot x}$ are independent, as both $M_{\mathbb{S}_X}\mathbb{S}_{0\cdot x}$ and $\mathbb{S}_x' M_{\mathbb{S}_{\Delta x}}\mathbb{S}_{0\cdot x}$ are normal and its conditional covariance is

$$cov\left(M_{\mathbb{S}_X}\mathbb{S}_{0\cdot x}, \mathbb{S}_x' M_{\mathbb{S}_{\Delta x}}\mathbb{S}_{0\cdot x}\right) = \sigma_{0\cdot x}^2\left[I_K - \mathbb{S}_X(\mathbb{S}_X'\mathbb{S}_X)^{-1}\mathbb{S}_X'\right]M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x = 0.$$

This implies that Z is independent of $\mathbb{S}_{0\cdot x}' M_{\mathbb{S}_X}\mathbb{S}_{0\cdot x}$ conditional on $\mathbb{S}_X = (\mathbb{S}_x, \mathbb{S}_{\Delta x})$, and hence

$$\begin{aligned}
&\frac{K}{p_\beta}\frac{\left\|\frac{Z}{\sigma_{0\cdot x}} + \left[R_\beta\left[\mathbb{S}_x' M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\right]^{-1}R_\beta'\right]^{-1/2}\cdot\left[\frac{cR_\beta\delta_0}{\sigma_{0\cdot x}}\right]\right\|^2}{\left[\frac{\mathbb{S}_{0\cdot x}' M_{\mathbb{S}_X}\mathbb{S}_{0\cdot x}}{\sigma_{0\cdot X}^2}\right]}\\
&\overset{d}{=} \frac{K}{K-2d}F_{p_\beta, K-2d}\left(\|\theta\|^2\right),
\end{aligned}$$

where

$$\theta = \left[R_\beta\left[\mathbb{S}_x' M_{\mathbb{S}_{\Delta x}}\mathbb{S}_x\right]^{-1}R_\beta'\right]^{-1/2}\cdot\left[\frac{cR_\beta\delta_0}{\sigma_{0\cdot x}}\right].$$

Similarly, with $Z = \left[R_\delta\left[\mathbb{S}_{\Delta x}' M_{\mathbb{S}_X}\mathbb{S}_{\Delta x}\right]^{-1}R_\delta\right]^{-1/2} R_\delta\left[\mathbb{S}_{\Delta x}' M_{\mathbb{S}_X}\mathbb{S}_{\Delta x}\right]^{-1}\mathbb{S}_{\Delta x}' M_{\mathbb{S}_X}\mathbb{S}_{0\cdot x}$, we obtain

$$F(\hat{\delta}) \Rightarrow \frac{K}{p_\delta}\frac{\left\|\frac{Z}{\sigma_{0\cdot x}}\right\|^2}{\left[\frac{\mathbb{S}_{0\cdot x}' M_{\mathbb{S}_X}\mathbb{S}_{0\cdot x}}{\sigma_{0\cdot X}^2}\right]} \overset{d}{=} \frac{K}{K-2d}F_{p_\delta, K-2d}.$$

■

Proof of Theorem 3.

We prove the result for the Wald statistic only as the same proof goes through for the t statistic with obvious modifications. From (29) and (29), we have

$$\begin{aligned}
T\left(R_\beta\left[\hat{\beta} - c\cdot\frac{\hat{\delta}}{T}\right] - r_\beta\right) &\Rightarrow \Gamma_c(\mathbb{S}_X'\mathbb{S}_X)^{-1}\mathbb{S}_X'\mathbb{S}_{0\cdot x},\\
\Gamma_c(\Upsilon_T^{-1}\mathbb{W}_X'\mathbb{W}_X\Upsilon_T^{-1})^{-1}\Gamma_c' &\Rightarrow \Gamma_c(\mathbb{S}_X'\mathbb{S}_X)^{-1}\Gamma_c'.
\end{aligned}$$

Combining these results with (38), we have

$$\begin{aligned}
F(\hat{\beta}; c) &= \frac{T^2}{\hat{\sigma}_{0\cdot x}^2}(R_\beta[\hat{\beta} - c\cdot\hat{\delta}/T] - r_\beta)'\left[\Gamma_c(\Upsilon_T^{-1}\mathbb{W}_X'\mathbb{W}_X\Upsilon_T^{-1})^{-1}\Gamma_c'\right]^{-1}\\
&\quad\times(R_\beta[\hat{\beta} - c\cdot\hat{\delta}/T] - r_\beta)/p.\\
&\Rightarrow \left[\frac{K}{p_\beta}\right]\frac{\left[\Gamma_c(\mathbb{S}_X'\mathbb{S}_X)^{-1}\mathbb{S}_X'\mathbb{S}_{0\cdot x}\right]'\left[\Gamma_c(\mathbb{S}_X'\mathbb{S}_X)^{-1}\Gamma_c'\right]^{-1}\left[\Gamma_c(\mathbb{S}_X'\mathbb{S}_X)^{-1}\mathbb{S}_X'\mathbb{S}_{0\cdot x}\right]}{\mathbb{S}_{0\cdot x}' M_{\mathbb{S}_X}\mathbb{S}_{0\cdot x}}
\end{aligned}$$

Using a similar argument in the proof of Proposition 2, the conditional limit of Wald statistics

$F(\hat{\beta}; c)$ can be expressed as

$$\left[\frac{K}{p_\beta}\right] \frac{\left[\Gamma_c(\mathbb{S}_X'\mathbb{S}_X)^{-1}\mathbb{S}_X'\mathbb{S}_{0\cdot x}\right]' \left[\Gamma_c(\mathbb{S}_X'\mathbb{S}_X)^{-1}\Gamma_c'\right]^{-1} \left[\Gamma_c(\mathbb{S}_X'\mathbb{S}_X)^{-1}\mathbb{S}_X'\mathbb{S}_{0\cdot x}\right]}{\mathbb{S}_{0\cdot x}' M_{\mathbb{S}_X} \mathbb{S}_{0\cdot x}}$$
$$\stackrel{d}{=} \frac{K}{p} \frac{\chi^2_{p_\beta}}{\chi^2_{K-2d}}, \;\; \chi^2_p \perp \chi^2_{K-2d}$$
$$\stackrel{d}{=} \frac{K}{K-2d} F_{p,K-2d},$$

which is invariant to the conditioning variable $\mathbb{S}_X$. Thus, it is also the unconditional distribution which proves

$$F(\hat{\beta}; c) \Rightarrow \frac{K}{K-2d} F_{p,K-2d}.$$

■

9 Appendix: Construction of $[c_l, c_h]$ in Elliott and Stock (2001)

Consider

$$x_t = \mu + \rho x_{t-1} + u_{xt} \text{ where } \rho = 1 - \frac{c}{T},$$

where u_{xt} has a serial dependence of unknown forms with $\Omega_{xx} = \sum_{j=-\infty}^{\infty} E[u_{xt}u'_{xt-j}]$. For simplicity, we only discuss a scalar regressor x_t with $d = 1$.

Step 0: Obtain a heteroskedasticity autocorrelation robust (HAR) estimator $\hat{\Omega}_{xx}$ by

$$\hat{\Omega}_{xx} = \sum_{j=-(T-1)}^{T-1} k\left(\frac{j}{bT}\right)\hat{\Gamma}_j \text{ where } \hat{\Gamma}_j = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} \hat{u}_t^* \hat{u}_{t-j}^{*\prime} \text{ for } j \geqq 0; \\ \frac{1}{T}\sum_{t=-j+1}^{T} \hat{u}_{t+j}^* \hat{u}_t^{*\prime} \text{ for } j < 0, \end{cases}$$

and $\hat{u}_t^* = \hat{u}_t - T^{-1}\sum_{t=1}^{T}\hat{u}_t$ with $\hat{u}_t = x_t - \hat{\rho}_{\text{OLS}} x_{t-1}$ for $t = 1, .., T$, and $b \in (0, 1]$ is the bandwidth parameter with a kernel function $k(\cdot)$. Here, we use Bartlett-kernel function in Newey and West (1987) and the optimal bandwidth rule suggested by Andrews (1991).

Step 1: Given the choice of the number of grids m, say $m = 200$, make a fine grid to get $\mathcal{C} = [0, c_1^*, c_2^*, c_T^*]$ with $c_T^* = 0.2T$.

Step 2: Following Elliott and Stock (2001, pp161), we choose $\bar{c} = 7$ with $\bar{\rho} = 1 - \frac{\bar{c}}{T}$ and construct the following test statistics:

$$P_T(0, \bar{c}) := \frac{1}{\hat{\Omega}_{xx}}\left[\sum_{t=1}^{T}(u_{\text{GLS},t}(\bar{\rho}))^2 - \bar{\rho}\sum_{t=1}^{T}(u_{\text{GLS},t}(1))^2\right],$$

where

$$Z(\rho) = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_T \end{bmatrix} = \begin{bmatrix} 1 \\ 1-\rho \\ \vdots \\ 1-\rho \end{bmatrix}, \ x(\rho) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ \vdots \\ x_T - x_{T-1} \end{bmatrix},$$

and

$$\begin{aligned} u_{\text{GLS},t}(\rho) &= x_t(\rho) - z_t(\rho)'\beta(\rho) \text{ for } t = 1, ..., T, \\ \beta(\rho) &= (Z'(\rho)Z(\rho))^{-1}Z'(\rho)x(\rho). \end{aligned}$$

Step 3: Let $W(\cdot)$ be a standard Wiener process and $J_c(\cdot)$ be OU-process $J_c(r) = \int_0^r \exp(-c(r-s))dW(s)$. Given $c^* \in \mathcal{C}$, simulate the following two quantities $(p_1(c^*), p_2(c^*))$

$$p_1(c^*) = 100 \cdot \epsilon_1\% \text{ percentile of } P(c^*, \bar{c});$$
$$p_2(c^*) = 100 \cdot (1-\epsilon_2)\% \text{ percentile of } P(c^*, \bar{c}),$$

where

$$P(c^*, \bar{c}) = \bar{c}^2 \int_0^1 (J_{c^*}(s))^2\, ds + \bar{c}J_{c^*}^2(1).$$

For $\epsilon = 0.10$, we choose $\epsilon_1 = 0.06$ and $\epsilon_2 = 0.04$ by following Eliiott and Stock (2001).

Here, with a very large number B, say $B = 5,000$, the random variable $P(c^*; 0, \bar{c})$ can be simulated by

$$\hat{p}_B(c^*; 0, \bar{c}) := \bar{c}^2 \cdot \frac{1}{B} \sum_{b=1}^{B} \left(\hat{J}_{c^*}(\frac{b}{B}) \right)^2 + \bar{c} \left(\hat{J}_{c^*}(1) \right)^2,$$

$$\hat{J}_{c^*}(\frac{s}{B}) := \frac{1}{\sqrt{B}} \sum_{b=1}^{s} \exp\left(c^*(\frac{s-b}{B}) \right) e_b,$$

where e_b's are $\overset{\text{i.i.d}}{\sim} N(0,1)$.

Step 4 Keep $c^* \in [c_l, c_u]$ if

$$p_{\alpha_l}(c^*, \bar{c}) \leq P_T(0, \bar{c}) \leq p_{1-\alpha_u}(c^*, \bar{c}).$$