Testing Predictability in the Presence of Persistent Errors*

Yijie Fei
Hunan University

Yiu Lim Lui
Dongbei University of Finance and Economics

Jun Yu
University of Macau

May 12, 2024

Abstract

This paper considers testing predictability in predictive regression models with persistent errors. We derive limiting distributions of the ordinary least squares estimator and the corresponding Wald statistic under the condition of moderately integrated errors or local-to-unity errors. The asymptotic result establishes the connection between super-consistent estimation in correctly specified predictive regression and inconsistent estimation in spurious regression. To provide a robust test, a modification to the IVX-AR test of Yang et al. (2020) is proposed. The modified test is uniformly valid across different degrees of persistency in both predictors and errors. Simulation studies show that the new test enjoys satisfactory finite sample performances. Leveraging on the new test, we reexamine the predictive power of numerous economic variables in predicting the growth rate of the U.S. housing market, demonstrating the usefulness of the proposed test, particularly in the context of multivariate regression.

Keywords: Spurious regression, Predictive regression, Uniform inference; Robust test; Moderately integrated; Nearly integrated, Housing price

*Yijie Fei, College of Finance and Statistics, Hunan University, 410100, China, Email: yijiefei@hnu.edu.cn. Yiu Lim Lui, Institute for Advanced Economic Research, Dongbei University of Finance and Economics, 116000, China, Email: luiyiulim@outlook.com. Jun Yu, Faculty of Business Administration, University of Macau, Avenida da Universidade, Macao, China. Email: junyu@um.edu.mo.

1 Introduction

Predictive regression has gained widespread popularity in empirical studies. It has been used in areas such as economics, finance, many fields in social sciences and business studies. The literature typically focuses on the following linear model:

$$y_t = \alpha + \beta x_{t-1} + u_{0,t}, \tag{1}$$

where x_{t-1} is a potential predictor for variable y_t and $u_{0,t}$ is the error term, which is usually assumed to be stationary.

Stambaugh (1999) pointed out that the autoregressive property of x_{t-1} causes a finite sample bias when β is estimated by ordinary least squares (OLS). Suggestions have been made to reduce the bias in several studies, including Stambaugh (1999), Amihud and Hurvich (2004) and Lewellen (2004). Nevertheless, when x_t is stationary, the large sample properties of the OLS estimator of β are the same as in the conventional regression model and hence, the asymptotic inference is straightforward. However, since many predictors used in empirical studies have strong persistency, the assumption of stationarity for x_t is questionable.

Unsurprisingly, there has been a recent surge of interest in the literature regarding the methodology for conducting inference on predictability in the presence of nonstationary or nearly nonstationary predictors. See for example Harvey et al. (2021), Yang et al. (2021) and Gungor and Luger (2020), to name just a few. Among the existing testing procedures, arguably the most popular method in empirical research is the self-instrumentation approach called IVX, established in Phillips and Magdalinos (2009) and Kostakis et al. (2015). The essential idea of the IVX methodology is to restrict the degree of persistency of data-filtered instruments within the class of moderately integrated (MI) processes, whose asymptotic behavior has been well studied in Phillips and Magdalinos (2007) and Magdalinos and Phillips (2009). The standard instrumental variable estimation based on the constructed instruments is robust to the general time-series characteristics of regressors in the sense that the derived estimator always converges in distribution to a (mixed) normal limit. Hence, in the limit, the associated Wald test statistic follows the chi-square distribution under the null hypothesis. A great deal of research has now emerged in the literature to study the theoretical properties of the IVX method. Important contributions include, for instance, Hosseinkouchack and Demetrescu (2021), Demetrescu et al. (2022) and Demetrescu and Rodrigues (2022).

When testing for predictability based on either the OLS estimator or the IVX estimator, a commonly employed assumption is that $u_{0,t}$ in (1) follows a martingale difference sequence (mds). Using simulations, Yang et al. (2020) show that when $u_{0,t}$ is serially correlated, the IVX-Wald test suffers from a severe size distortion. To avoid this problem, they assume that $u_{0,t}$ is generated by a stationary autoregressive (AR) process with an unknown order p and proposed a novel inference procedure based on Cochrane-Orcutt-type correction. Using simulated data, they showed that their test, referred to as IVX-AR, exhibits excellent size control.

However, the dynamic of $u_{0,t}$ may not be well captured by a stationary AR process in practice. Moreover, when the AR coefficient of a time series is close to (albeit smaller than) unity, it is difficult to classify it as a stationary, or a local-to-unity (LUR), or an MI process. Furthermore, it is well-known that the finite sample distributions of many statistics are not well approximated by their asymptotic distribution derived from the stationary assumption; see, for example, Ahtola and Tiao (1984); Phillips (1987); Perron (1991). In this case, it could be better to assume $u_{0,t}$ follows an LUR or an MI process. To illustrate the relevance of our concern, consider the empirical application in Yang et al. (2020), where the predictability of the U.S. housing market growth rate by a set of economic variables is tested. If we focus on a particular predictor, namely the shares of the residential fixed investment in GDP, we will show in Section 6 that the residuals from model (1) is strongly persistent. Its persistency manifests by a large value for the KPSS statistic (0.981), greater than the 99% critical value (0.739), rejecting the null hypothesis of stationarity.

When $u_{0,t}$ follows an LUR or an MI process, another well-known problem, namely spurious regression, may arise. A leading example is when x_{t-1} and $u_{0,t}$ are two independent unit root processes. In this case, as shown by Phillips (1986), the OLS estimator for β in (1) is inconsistent and the corresponding t-statistic diverges as the sample size increases, which suggests that statistical evidence of predictability is spurious. Similarly, Lin and Tu (2020) show that, when x_{t-1} and $u_{0,t}$ are two MI processes from the stationary side and x_{t-1} and y_t are uncorrelated, the OLS estimator for β in (1) is inconsistent and the corresponding t-statistic diverges. Both sets of results are in sharp contrast to the case in which x_{t-1} follows a random walk and $u_{0,t}$ is stationary, where the OLS estimator for β is super-consistent and its corresponding t-statistic is bounded in probability.

Unfortunately, it is unknown how the limit behavior of the IVX estimator and the associated test statistic changes when the data generating process (DGP) transits from the model in Yang et al. (2020) to a spurious regression model. A natural question to ask is whether the IVX-based inferences are immune to spurious predictiveness in the presence of persistent errors. If it is not, it will then be desirable to construct a uniform predictability test that is valid regardless of the persistent level of $u_{0,t}$.

In light of the these gaps in the literature, this paper intends to make the following contributions. Firstly, we propose a multivariate predictive regression model, which allows the error term to have an LUR or MI structure. This specification bridges between the model considered in Yang et al. (2020) and spurious predictive regression. Secondly, we investigate the asymptotic behavior of various popular test statistics for predictability under this general setup. In particular, we obtain the asymptotic properties of both the OLS estimator and the IVX estimator, as well as their corresponding Wald statistics, when both the covariates and the errors are persistent. Our theoretical results provide a complete characterization of transition from super-consistent estimation in correctly specified predictive regression to inconsistent estimation in spurious regression. Furthermore, it is analytically discussed that the validity of the IVX-Wald tests proposed in Kostakis et al. (2015) and Yang et al. (2020) are critically dependent on the degree of persistency of the error term. Our third contribution is to propose a unified test for predictability, which is robust to combinations of different levels of persistency in the predictors and the errors. Via extensive simulations, we show our new test enjoys good size control and power. Lastly, we provide an empirical application of predicting the growth rate of the U.S. housing price. In a univariate regression analysis, our findings in general confirm the conclusions in Yang et al. (2020), indicating that their results are robust to the persistency in the errors and non-spurious. In a multivariate regression analysis, however, our results are significantly different from those of Yang et al. (2020). We find that almost all combinations of predictors considered in Yang et al. (2020) cannot predict the growth rate.

The paper is organized as follows. Section 2 introduces our baseline model and studies the asymptotic properties of the OLS estimator, the corresponding Wald test statistic, and the KPSS statistic when the errors are persistent. Section 3 reviews the IVX method and discusses its asymptotic behavior when the errors are persistent. Section 4 introduces the modified IVX-AR test and derives its limiting distribution. Section 5 designs several Monte Carlo experiments to check the finite sample performance of the test. The empirical study that tries to predict the growth rate of the U.S. housing market is conducted in Section 6. Section 7 concludes the paper. Proofs of the main results in the paper are given in the Online Supplement. Throughout the paper, $\xrightarrow{p}$ and $\Rightarrow$ denote convergence in probability and weak convergence, respectively. The superscript μ denotes the demeaned time series (e.g. $y_t^{\mu} = y_t - \frac{1}{T}\sum_{t=1}^{T} y_t$). $\|\cdot\|$ is used to denote the spectral norm of a matrix and $I(\cdot)$ is used to denote an indicator function.

2 Baseline Model and Asymptotics for OLS

Consider the following model:

$$\begin{cases} y_t &= \alpha + \beta' X_{t-1} + u_{0,t} \\ u_{0,t} &= \rho_T u_{0,t-1} + \varepsilon_{0,t} \\ X_t &= \Phi_T X_{t-1} + \varepsilon_{1,t} \end{cases}, \tag{2}$$

where $t = 1, 2, ..., T$, y_t and $u_{0,t}$ are scalar, X_t, β and $\varepsilon_{1,t}$ are $k \times 1$ vector, Φ_T is a $k \times k$ diagonal matrix. To avoid unnecessary complications, we assume both X_0 and $u_{0,0}$ are $O_p(1)$ so that the same initial conditions apply to LUR and MI regressors and errors. To cater for the dependence structure of the errors, we further assume

$$\varepsilon_t = \sum_{j=0}^{\infty} C_j z_{t-j}, \tag{3}$$

where $\varepsilon_t = [\varepsilon_{0,t}, \varepsilon'_{1,t}]'$, $z_t = [z_{0,t}, z'_{1,t}]'$ and $C_j = \begin{bmatrix} c_{0,j} & C_{01,j} \\ C_{10,j} & C_{1,j} \end{bmatrix}$. Throughout the paper, we impose the following condition.

Assumption 1. *For the linear process (3), we assume that* $\sum_{j=0}^{\infty} j \|C_j\| < \infty$, *and* $\{z_t\}_{t=1}^{T}$ *is a sequence of independent and identically distributed (iid) random vector, which has mean zero and positive definite variance* Σ *with* $E\|z_1\|^4 < \infty$.

The following standard notations for the spot variance (Σ) and the long run variance (Ω) are used:

$$\Sigma = E[\varepsilon_t \varepsilon_t'] = \begin{bmatrix} E\varepsilon_{0,t}^2 & E\varepsilon_{0,t}\varepsilon'_{1,t} \\ E\varepsilon_{1,t}\varepsilon_{0,t} & E\varepsilon_{1,t}\varepsilon'_{1,t} \end{bmatrix} = \begin{bmatrix} \Sigma_{00} & \Sigma_{01} \\ \Sigma_{10} & \Sigma_{11} \end{bmatrix},$$

$$\Lambda = \sum_{h=1}^{\infty} E\left[\varepsilon_t \varepsilon'_{t-h}\right] = \begin{bmatrix} \sum_{h=1}^{\infty} E\left(\varepsilon_{0,t}\varepsilon_{0,t-h}\right) & \sum_{h=1}^{\infty} E\left(\varepsilon_{0,t}\varepsilon'_{1,t-h}\right) \\ \sum_{h=1}^{\infty} E\left(\varepsilon_{1,t}\varepsilon_{0,t-h}\right) & \sum_{h=1}^{\infty} E\left(\varepsilon_{1,t}\varepsilon'_{1,t-h}\right) \end{bmatrix} = \begin{bmatrix} \Lambda_{00} & \Lambda_{01} \\ \Lambda_{10} & \Lambda_{11} \end{bmatrix},$$

$$\Omega = \Sigma + \Lambda + \Lambda' = \begin{bmatrix} \Omega_{00} & \Omega_{01} \\ \Omega_{10} & \Omega_{11} \end{bmatrix}.$$

Similarly, the variance of $u_{0,t}$ and the covariance of $[\varepsilon'_{1,t}, u_{0,t}]'$ are defined as

$$\Sigma_u = E[u_{0,t}u'_{0,t}], \Lambda_u = \sum_{k=1}^{\infty} E[u_{0,t}u_{0,t-k}], \Omega_u = \Sigma_u + \Lambda_u + \Lambda'_u,$$

$$\Sigma_{\varepsilon u} = E[\varepsilon_{1,t}u'_{0,t}], \Lambda_{\varepsilon u} = \sum_{k=1}^{\infty} E[\varepsilon_{1,t}u_{0,t-k}], \Omega_{\varepsilon u} = \Sigma_{\varepsilon u} + \Lambda_{\varepsilon u} + \Lambda'_{\varepsilon u}.$$

Suppose $\Phi_T = I_k + \frac{C_x}{T}$, where C_x is a diagonal matrix with non-positive entries. In this case, X_t is an LUR predictor and the large sample theory for various sample moments was studied in Phillips (1987) and Phillips (1988). The LUR assumption has been widely adopted in the predictive regression literature (see Cavanagh et al. (1995), Campbell and Yogo (2006), Kostakis et al. (2015) for further details). This model specification with LUR predictors has interesting interpretations under different assumptions of ρ_T. For example, if we further assume $\rho_T = \rho$ and $|\rho| < 1$, model (2) is a predictive regression model with a strictly stationary error. A closely related model was recently studied in Yang et al. (2020). When $\rho_T = 1 + \frac{c_u}{T}$, with $c_u < 0$, $u_{0,t}$ is an LUR time series. When $\rho_T = 1 + \frac{c_u}{T^{\kappa_u}}$, with $c_u < 0$ and $\kappa_u \in (0, 1)$, $u_{0,t}$ is an MI error term. The asymptotic behavior of MI processes has been well studied in Phillips and Magdalinos (2007, 2009). Model (2) also extends the spurious regression model studied in Phillips (1986) where $u_{0,t}$ and X_t are both unit root processes and $\beta = 0$. It is noteworthy to mention that when $c_u < 0$, $u_{0,t}$ is a mean reverting process under both the LUR and the MI settings. Empirically, it is difficult to distinguish between a strictly stationary process from an LUR or an MI process when the AR parameter is close to unity.

In the time series literature, it is well-known that when ρ_T is close to unity, the finite sample distribution of the OLS estimator $\hat{\rho}_T$ is poorly approximated by the Gaussian limiting distribution. In this paper, to showcase the potential size distortion induced by the persistent errors, we first investigate the limiting behavior of the OLS estimator of β and the corresponding Wald statistic, under the assumption that $u_{0,t}$ is strictly stationary or LUR or MI.

The OLS estimator of β is defined as

$$\hat{\beta} = \left(\sum_{t=1}^{T} X_{t-1}^{\mu} X_{t-1}^{\mu'} \right)^{-1} \left(\sum_{t=1}^{T} X_{t-1}^{\mu} y_t^{\mu} \right).$$

The corresponding Wald statistic W_T for the null hypothesis $H_0(R\beta = r)$ is

$$W_T = \left(R\hat{\beta} - r \right)' \left[R Q_W R' \right]^{-1} \left(R\hat{\beta} - r \right),$$

with

$$Q_W = \left[\sum_{t=1}^{T} X_{t-1}^{\mu} X_{t-1}^{\mu\prime} \right]^{-1} \hat{\Omega}_u,$$

where R is a $q \times k$ matrix with rank q and r is a $q \times 1$ vector and $\hat{\Omega}_u$ is an estimate of Ω_u using the OLS residual $e_{0,t}$ (i.e., $e_{0,t} = y_t^{\mu} - \hat{\beta}' X_{t-1}^{\mu}$).[1]

Denote $V_{xx} = \int_0^{\infty} e^{pC_x} \Omega_{11} e^{pC_x} dp$ and $\Lambda_{\tilde{z}\varepsilon} = \sum_{h=1}^{\infty} E[\tilde{z}_{0,t} \varepsilon'_{1,t-h}]$, where $\tilde{z}_{0,t}$ is defined via the Beveridge-Nelson decomposition of $u_{0,t}$, such that $u_{0,t} = \Psi_0(L) z_{0,t} = \Psi_0(1) z_{0,t} + \Delta \tilde{z}_{0,t}$, with $\Psi_0(L) z_{0,t} = \sum_{j=0}^{\infty} \psi_{0,j} z_{0,t-j} = (1 - \rho L)^{-1} \sum_{j=0}^{\infty} c_{0,j} z_{0,t-j}$, and $\tilde{z}_{0,t} = \sum_{j=0}^{\infty} a_j z_{0,t-j}$ with $a_j = -\sum_{j=0}^{\infty} \psi_{0,j+1}$. Lemma 2.1 reports the asymptotic behavior of $\hat{\beta}$ and W_T. Let C_x be a diagonal matrix with entries $\{c_{x,i}\}_{i=1}^{k}$. We define LUR, MI and ME regressor X_t as model (2) with the following AR coefficients: (LUR) $\Phi_T = I_k + C_x/T$, $C_x \leq 0$; (MI) $\Phi_T = I_k + C_x/T^{\kappa_x}$, $C_x < 0$; (ME) $\Phi_T = I_k + C_x/T^{\kappa_x}$, $C_x > 0$ with its diagonal elements $c_{x,i} \neq c_{x,j}$ for any $i \neq j$, $\kappa_x \in (1/3, 1)$.[2]

We also define the stationary, LUR and MI error $u_{0,t}$ as model (2) with the following AR coefficient: (stationary) $\rho_T = \rho$ with $|\rho| < 1$; (LUR) $\rho_T = 1 + c_u/T$ with $c_u \leq 0$; (MI) $\rho_T = 1 + c_u/T^{\kappa_u}$ with $\kappa_u \in (1/3, 1)$.

[1]The formal definition of $\hat{\Omega}_u$ is based on the Bartlett kernel function with bandwidth M_T and will be provided in the Online Supplement. As usual, we assumed that $M_T \to \infty$ and $M_T/T \to 0$ as $T \to \infty$.

[2]The assumption $c_{x,i} \neq c_{x,j}$ for any $i \neq j$ avoids the complication of a common moderately explosive behavior among predictors; see Magdalinos and Phillips (2009) for further discussion.

Lemma 2.1. *Under model (2) and Assumption 1, let $c_u < 0$, and $\kappa_x, \kappa_u \in (1/3, 1)$. Suppose that $M_T = O(T^{1/3})$.*[3] *Table 1 provides the stochastic orders of the centered OLS estimator $\hat{\beta} - \beta$ and the OLS-Wald statistic W_T under different level of persistency of X_t and $u_{0,t}$.*[4]

Table 1: Summary of results for OLS

Panel 1: Limit order of $\hat{\beta} - \beta$

		X_t		
		MI	LUR	ME
$u_{0,t}$	Stationary	$O_p(T^{-\kappa_x})$	$O_p(T^{-1})$	$O_p(T^{-\kappa_x}\Phi_T^{-T})$
	MI	$O_p(T^{-\max(\kappa_x-\kappa_u,0)})$	$O_p(T^{-(1-\kappa_u)})$	$O_p(T^{\frac{(\kappa_u-\kappa_x)}{2-I(\kappa_u\leq\kappa_x)}}\Phi_T^{-T})$
	LUR	$O_p(1)$	$O_p(1)$	$O_p(T^{\frac{(1-\kappa_x)}{2}}\Phi_T^{-T})$

Panel 2: Limit order of OLS-Wald statistic

		X_t		
		MI	LUR	ME
$u_{0,t}$	Stationary	$O_p(T^{1-\kappa_x})$	$O_p(1)$	$O_p(1)$
	MI	$O_p(\frac{T^{1-\lvert\kappa_x-\kappa_u\rvert}}{M_T})$	$O_p(\frac{T^{\kappa_u}}{M_T})$	$O_p(T^{\min(\kappa_x,\kappa_u)-\frac{1}{3}})$
	LUR	$O_p(\frac{T^{\kappa_x}}{M_T})$	$O_p(\frac{T}{M_T})$	$O_p(T^{\kappa_x-\frac{1}{3}})$

The asymptotic order of $\hat{\beta} - \beta$ is summarized in the top panel of Table 1. It is clear that the rate of convergence of $\hat{\beta}$ decreases as the degree of persistency of $u_{0,t}$ increases. Table 1 shows that, when both X_t and $u_{0,t}$ are LUR, $\hat{\beta}$ is inconsistent. Whereas, when $u_{0,t}$ is stationary or MI, $\hat{\beta}$ is consistent with different convergence rates. When X_t is MI, $\hat{\beta}$ is consistent only when the error term is stationary or X_t is more persistent than $u_{0,t}$ (i.e. $\kappa_u < \kappa_x$).

The asymptotic order of the OLS-Wald statistic is summarized in the bottom panel of Table 1. The OLS-Wald statistic diverges in probability when $u_{0,t}$ is not stationary as $T^\delta/M_T \to \infty$ for any $\delta > 1/3$. The only case where the OLS-Wald statistic has the χ^2 asymptotic distribution is when X_t is ME and $u_{0,t}$ is stationary. As the OLS-Wald statistic diverges in most cases, the

[3] The order $O(T^{1/3})$ covers popular bandwidth choices such as $M_T = o(T^{1/4})$ in KPSS and $M_T = T^{1/3}$ in Kostakis et al. (2015) and Yang et al. (2020).

[4] The detailed limiting distributions are not presented in the main body of the paper for brevity. We provide the limiting distribution of $\hat{\beta}$ in Lemma A.4 in the Online Supplement.

standard inference using the Wald statistic based on OLS in general provides a spurious conclusion of predictability.

Remark 2.1. *When $\beta = 0$, Lemma 2.1 covers the spurious regression results reported in Theorem 1 of Lin and Tu (2020). A notable difference between our study and Lin and Tu (2020) is that we allow the dependent variable and the regressors to have different deviation rates from unit root (i.e. $\kappa_x \neq \kappa_u$), whereas the rates are assumed to be the same (i.e. $\kappa_x = \kappa_u$) in Lin and Tu (2020). Moreover, Lin and Tu (2020) only consider univariate regression, while our model allows for multiple predictors.*

In our empirical example of regressing the growth rate of the housing market on a scalar predictor, we find a large value for the KPSS statistic (denoted by L_T), indicating persistent errors. The following proposition shows that when the errors is MI or LUR,[5] the statistic L_T will exceed any finite critical value when sample size goes to infinity.

Proposition 2.1. *Consider the KPSS statistic L_T*

$$L_T = \frac{\frac{1}{T^2}\sum_{t=1}^{T} S_t^2}{\hat{\Omega}_u}, S_t = \sum_{t=1}^{t} e_{0,t}. \tag{4}$$

Under the same set of assumptions as in Lemma 3.1, if X_t is LUR and $u_{0,t}$ is MI or LUR, for any $M_T = O(T^{1/3})$ and any critical value $cv \in R$, we have, as $T \to \infty$, $Pr(L_T > cv) \to 1$.

Remark 2.2. *In addition to the conventional OLS-Wald test, we have also obtained the asymptotic properties of some other popular tools for predictive regression such as the Bonferroni confidence interval of Cavanagh et al. (1995) and the Cauchy-estimator-based t-test of Choi et al. (2016). The results and the proof can be found in the Online Supplement, indicating that both methods continue to lead to spurious results if $u_{0,t}$ LUR or MI.*

[5] We only focus on the LUR case for the regressors in this proposition to keep the analysis parsimonious.

3 Limit Behavior of IVX Estimator and IVX-Wald Test

Phillips and Magdalinos (2009) introduced the IVX method that can produce a standard limiting distribution for the test statistic under the cointegrated model with LUR or MI regressor. The method is later used by Phillips and Lee (2013) and Kostakis et al. (2015) to test the predictive power of predictive regression.

To fix the idea, we define the self-instrumental variable as $Z_t = \sum_{j=1}^{t} \Upsilon_T^{t-j} \Delta X_j$ where $Z_0 = 0$, $\Upsilon_T = 1 + C_z/T^{\eta}$ and $\Delta X_t = X_t - X_{t-1}$. The tuning parameters $C_z < 0$ and $\eta \in (0,1)$ are specified by the user.[6] The IVX estimator of β is defined as

$$\hat{\beta}_{IVX} = \left[\sum_{t=1}^{T} Z_{t-1} X_{t-1}^{\mu\prime}\right]^{-1} \sum_{t=1}^{T} Z_{t-1} y_t^{\mu}$$

and the corresponding IVX-Wald statistic for $H_0(R\beta = r)$, is defined as

$$\begin{aligned} W_{\hat{\beta}_{IVX}} &\equiv \left(R\hat{\beta}_{IVX} - r\right)' Q_{IVX}^{-1} \left(R\hat{\beta}_{IVX} - r\right), \\ Q_{IVX} &\equiv R\left[\sum_{t=1}^{T} Z_{t-1} X_{t-1}^{\mu\prime}\right]^{-1} M \left[\sum_{t=1}^{T} X_{t-1}^{\mu} Z_{t-1}'\right]^{-1} R', \\ M &\equiv \left[\sum_{t=1}^{T} Z_{t-1} Z_{t-1}'\right] \hat{\Sigma}_u - T\bar{Z}_{t-1}\bar{Z}_{t-1}' \hat{\Omega}_{FM}, \end{aligned}$$

where $\hat{\Omega}_{FM} = \hat{\Sigma}_u - \hat{\Omega}_{01}\hat{\Omega}_{11}^{-1}\hat{\Omega}_{01}'$, $\bar{Z}_{t-1} = \frac{1}{T}\sum_{t=1}^{T} Z_{t-1}$.[7]

When errors are serially correlated as in (3), the IVX estimator needs to be re-centered to obtain a proper limiting distribution.[8] Hence, in addition to original IVX estimator, we also investigate the performance of the re-centered IVX estimator ($\check{\beta}_{IVX}$) and its corresponding IVX-Wald statistic ($W_{\check{\beta}_{IVX}}$) for $H_0(R\beta = r)$, defined as

$$\check{\beta}_{IVX} = \left[\sum_{t=1}^{T} Z_{t-1} X_{t-1}^{\mu\prime}\right]^{-1} \left(\sum_{t=1}^{T} Z_{t-1} y_t^{\mu} - T\hat{\Lambda}_{01}\right)$$

[6]In the simulation and the empirical studies reported later, following the suggestion of Kostakis et al. (2015) and Phillips and Lee (2016), we set $C_z = -I_k$ and $\eta = 0.95$.

[7]We define $\hat{\Sigma}_u$, $\hat{\Omega}_{01}$, and $\hat{\Omega}_{11}$ in (32), (33) and (34) in the Online Supplement.

[8]See Phillips and Magdalinos (2009) for the details except the case where the predictor is mildly explosive.

and

$$\begin{aligned}
W_{\check{\beta}_{IVX}} &\equiv \left(R\check{\beta}_{IVX} - r\right)' \check{Q}_{IVX}^{-1} \left(R\check{\beta}_{IVX} - r\right) \\
\check{Q}_{IVX} &\equiv R\left[\sum_{t=1}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1} \check{M} \left[\sum_{t=1}^{T} X_{t-1}^{\mu} Z_{t-1}'\right]^{-1} R' \\
\check{M} &\equiv \left[\sum_{t=1}^{T} Z_{t-1}Z_{t-1}'\right] \hat{\Omega}_{00} - T\bar{Z}_{t-1}\bar{Z}_{t-1}'\check{\Omega}_{FM}
\end{aligned}$$

where $\check{\Omega}_{FM} = \hat{\Omega}_{00} - \hat{\Omega}_{01}\hat{\Omega}_{11}^{-1}\hat{\Omega}_{01}'$.[9]

In the simulation studies of Yang et al. (2020), it can be observed that, when the AR coefficient is close to unity (e.g. $\rho_T = \rho = 0.9$), the $\chi^2(q)$ limiting distribution used by IVX-Wald leads to a severe size distortion. The following theorem reports the asymptotic behaviors of $\hat{\beta}_{IVX,T}$, $\check{\beta}_{IVX}$, $W_{\hat{\beta}_{IVX}}$ and $W_{\check{\beta}_{IVX}}$ under model (2), and hence, provides theoretical explanations for their finding.

Theorem 3.1. *Under the same set of assumptions as in Lemma 2.1, if* $M_T = O(T^{1/3})$, $1 > \eta > 2/3$, $\kappa_x, \kappa_u \in (1/3, 1]$, *as* $T \to \infty$, *we have following results:*[10]

(i). If X_t *is LUR or MI and* $u_{0,t}$ *is stationary, then both* $\hat{\beta}_{IVX}$ *and* $\check{\beta}_{IVX}$ *are consistent. Moreover, while* $W_{\check{\beta}_{IVX}}$ *converges in distribution to the* $\chi^2(q)$ *distribution,* $W_{\hat{\beta}_{IVX}}$ *diverges to positive infinity in probability.*

(ii). If X_t *is LUR or MI and* $u_{0,t}$ *is LUR, then both* $\hat{\beta}_{IVX}$ *and* $\check{\beta}_{IVX}$ *are inconsistent. Moreover, both* $W_{\hat{\beta}_{IVX}}$ *and* $W_{\check{\beta}_{IVX}}$ *diverge to positive infinity in probability.*

(iii). If X_t *is LUR or MI and* $u_{0,t}$ *is MI, then* $\hat{\beta}_{IVX}$ *is consistent if* $\min\{\kappa_x, \eta\} > \min\{\kappa_u, \eta\}$, *and* $\check{\beta}_{IVX}$ *is consistent if the* $\min\{\kappa_x, \eta\} > \min\{\kappa_u, \kappa_x\}$. *Otherwise, they are inconsistent. Moreover, both* $W_{\hat{\beta}_{IVX}}$ *and* $W_{\check{\beta}_{IVX}}$ *diverge to positive infinity in probability.*

[9] $\hat{\Lambda}_{0,1}$ is defined in (140) in the Online Supplement.

[10] The stochastic orders are in general quite complicated and dependent on the value of κ_x, κ_u, C_x and c_u. The stochastic orders of $\hat{\beta}_{IVX}$, $\check{\beta}_{IVX}$, $W_{\hat{\beta}_{IVX}}$ and $W_{\check{\beta}_{IVX}}$ as well as the detailed derivations are hence reported in the Online Supplement.

(iv). If X_t is ME with $\kappa_x > 1/2$ and $u_{0,t}$ is stationary, then both $\hat{\beta}_{IVX}$ and $\check{\beta}_{IVX}$ are consistent. Moreover, $W_{\hat{\beta}_{IVX}}$ converges to a distribution which is proportional to $\chi^2(q)$ and $W_{\check{\beta}_{IVX}}$ converges to $\chi^2(q)$ distribution.

(v). If X_t is ME with $\kappa_x > 1/2$ and $u_{0,t}$ is LUR, then both $\hat{\beta}_{IVX}$ and $\check{\beta}_{IVX}$ are consistent. Moreover, both $W_{\hat{\beta}_{IVX}}$ and $W_{\check{\beta}_{IVX}}$ diverge to positive infinity in probability.

(vi). If X_t is ME with $\kappa_x > 1/2$ and $u_{0,t}$ is MI, then both $\hat{\beta}_{IVX}$ and $\check{\beta}_{IVX}$ are consistent. Moreover, the asymptotic behaviours of $W_{\hat{\beta}_{IVX}}$ and $W_{\check{\beta}_{IVX}}$ depend on the values of η, κ_x, and κ_u.

Theorem 3.1 shows that the IVX estimators $\hat{\beta}_{IVX}$ and $\check{\beta}_{IVX}$ can only provide well-behaved Wald statistics $W_{\hat{\beta}_{IVX}}$ and $W_{\check{\beta}_{IVX}}$ when $u_{0,t}$ is stationary ($|\rho| < 1$). When $u_{0,t}$ is LUR or MI, the Wald statistics diverge, whether or not they are re-centered. Theorem 3.1 explains the severe size distortion in $W_{\hat{\beta}_{IVX}}$ found in Yang et al. (2020) via simulation. It also explains Remark 1 of Yang et al. (2020) about the severe size distortion in $W_{\check{\beta}_{IVX}}$.

Remark 3.1. *The condition $\kappa_x, \kappa_u \in (1/3, 1]$ suggests that X_t and $u_{0,t}$ cannot be too close to the stationary region. Phillips and Magdalinos (2009) impose a similar condition, which is sufficient to ensure that a limiting distribution can be obtained for $\check{\beta}_{IVX} - \beta$ after removing the endogeneity bias via re-centering; see Remark 3.9(iv) in Phillips and Magdalinos (2009) for further discussion.*

4 A Robust Test for Predictability

Yang et al. (2020) proposed an IVX-AR method when $u_{0,t}$ is assumed to follow a stationary $AR(p)$ process. The method first obtains the OLS residuals from the predictive regression, fits an $AR(p)$ model to the residuals, and then performs Cochrane-Orcutt correction using the estimated AR coefficients. The intuition for this method is that the serial correlation in the residuals will be asymptotically removed by the correction when $u_{0,t}$ is stationary $AR(p)$ and consequently the conventional $\chi^2(q)$ limiting distribution can be achieved for the Wald statistic.

In this paper, we adopt a similar $AR(p)$ structure. However, instead of imposing stationary errors as in Yang et al. (2020), we assume $u_{0,t}$ in (2) follows

$$\begin{cases} u_{0,t} &= \rho_{1,T}u_{0,t-1} + \rho_2\Delta u_{0,t-1} + ... + \rho_p\Delta u_{0,t-p+1} + z_{0,t}, u_{0,0} = O_p(1) \\ \rho_{1,T} &= 1 + \frac{c_u}{T^{\kappa_u}}, c_u < 0, \kappa_u \in (0,1] \end{cases} . \tag{5}$$

It is assumed that all roots of $\bar{\rho}(z) = 0$ lie outside the unit circle where $\bar{\rho}(z) = 1 - \rho_2 z - ... - \rho_p z^{p-1}$. It is also assumed that $z_{0,t}$ is a martingale difference sequence with respect to the information set at $t-1$ and has a finite second moment. If $\kappa_u = 1$, $u_{0,t}$ is LUR. If $\kappa_u \in (0,1)$, $u_{0,t}$ is MI. Following Yang et al. (2020), we propose to perform Cochrane-Orcutt-type correction before implementing IVX.

In many empirical studies that use predictive regression, the key hypothesis is the presence of predictability of a scalar predictor or a set of predictors. In this case, the null hypothesis is simply $\beta = 0$, which implies $y_t^\mu = u_{0,t}^\mu$. Therefore, we propose to conduct the Cochrane-Orcutt-type correction by performing the following p^{th} order autoregression,

$$y_t^\mu = \hat{\rho}_{1,T}y_{t-1}^\mu + \hat{\rho}_2\Delta y_{0,t-1}^\mu + ... + \hat{\rho}_p\Delta y_{0,t-p+1}^\mu + \hat{v}_t. \tag{6}$$

This procedure is different from Yang et al. (2020) where $\hat{\rho}_1, ..., \hat{\rho}_p$ are obtained from regressing $\hat{u}_{0,t}$ on its lags. Based on regression (6), we get $\tilde{y}_t \equiv y_t - \mathbf{Y}_{t-1}\hat{\rho}$ and $\tilde{X}_t \equiv X_t - \mathbf{X}_{t-1}\hat{\rho}$, where $\hat{\rho} \equiv [\hat{\rho}_{1,T}, \hat{\rho}_2, ..., \hat{\rho}_p]'$, $\mathbf{Y}_{t-1} \equiv [y_{t-1}, \Delta y_{t-1}, ..., \Delta y_{t-p+1}]$, $\mathbf{X}_{t-1} \equiv [X_{t-1}, \Delta X_{t-1}, ..., \Delta X_{t-p+1}]$, $\tilde{u}_{0,t} \equiv u_{0,t}^\mu - U_{0,t-1}\hat{\rho}$, $U_{0,t-1} = [u_{t-1}^\mu, \Delta u_{t-1}^\mu, ..., \Delta u_{t-p+1}^\mu]$.

4.1 A modified IVX-AR test

To allow for persistent errors, we estimate β by

$$\tilde{\beta}_{IVX} \equiv \left[\sum_{t=p+1}^{T} \tilde{Z}_{t-1}\tilde{X}^{\mu\prime}_{t-1}\right]^{-1} \left(\sum_{t=p+1}^{T} \tilde{Z}_{t-1}\tilde{y}^{\mu}_{t}\right),$$

where

$$\tilde{Z}_t = Z_t - \hat{\rho}_{1,T} Z_{0,t-1} - \hat{\rho}_2 \Delta Z_{0,t-1} - ... - \hat{\rho}_p \Delta Z_{0,t-p+1}.$$

Our modified IVX-AR Wald test statistic $W_{\tilde{\beta}_{IVX}}$ for $H_0(\beta = 0)$ is defined as

$$W_{\tilde{\beta}_{IVX}} \equiv \tilde{\beta}'_{IVX}\tilde{Q}^{-1}_{IVX}\tilde{\beta}_{IVX}, \tag{7}$$

where

$$\begin{aligned}
\tilde{Q}_{IVX} &\equiv \left[\sum_{t=p+1}^{T} \tilde{Z}_{t-1}\tilde{X}^{\mu\prime}_{t-1}\right]^{-1} \tilde{M} \left[\sum_{t=p+1}^{T} \tilde{X}^{\mu}_{t-1}\tilde{Z}'_{t-1}\right]^{-1}, \\
\tilde{M} &\equiv \left[\sum_{t=p+1}^{T} \tilde{Z}_{t-1}\tilde{Z}'_{t-1}\right] \tilde{\Sigma}_{00} - T\breve{Z}_{t-1}\breve{Z}'_{t-1}\tilde{\Omega}_{FM}, \\
\tilde{\Omega}_{FM} &\equiv \tilde{\Sigma}_{00} - \tilde{\Omega}_{01}\tilde{\Omega}^{-1}_{11}\tilde{\Omega}'_{01}.
\end{aligned}$$

Denote $e_{1,t} = X^{\mu}_t - \hat{\Phi}_T X^{\mu}_{t-1}$ and $\hat{\Phi}_T$ as the OLS residual and the OLS estimator from regressing X_t on X_{t-1} with an intercept, respectively, $\tilde{e}_{0,t} = \tilde{u}_{0,t}, \tilde{e}_{1,t} = e_{1,t} - [e_{1,t-1}, \Delta e_{1,t-1}, ..., \Delta e_{1,t-p+1}]\hat{\rho}$. We define

$$\begin{aligned}
\tilde{\Omega}_{ij} &= \frac{1}{T-p}\sum_{h=-M_T}^{M_T}\left(1 - \frac{h}{1+M_T}\right)\sum_{t=p+h+1}^{T} \tilde{e}_{i,t}\tilde{e}_{j,t-h},\ i,j \in \{0,1\}, \\
\tilde{\Lambda}_{01} &= \frac{1}{T-p}\sum_{h=1}^{M_T}\left(1 - \frac{h}{1+M_T}\right)\sum_{t=p+h+1}^{T} \tilde{e}_{0,t}\tilde{e}_{1,t-h}, \\
\tilde{\Sigma}_{00} &= \frac{1}{T-p}\sum_{t=p+1}^{T} \tilde{e}_{0,t}, \breve{Z}_t = \frac{1}{T}\sum_{t=p+1}^{T} \tilde{Z}_t.
\end{aligned}$$

Theorem 4.1. *Consider the error process (5). Let the same set of assumptions for ρ_T and Φ_T as in Theorem 3.1 hold. Then under the null hypothesis $\beta = 0$, as $T \to \infty$, $W_{\tilde{\beta}_{IVX}} \Rightarrow \chi^2(k)$.*

Theorem 4.1 shows that under all combinations of persistency of X_t and $u_{0,t}$, our modified test statistic converges in distribution to $\chi^2(k)$ as $T \to \infty$. Therefore, unlike $W_{\hat{\beta}_{IVX}}$ and $W_{\check{\beta}_{IVX}}$, the modified statistic does not suffer from size distortion. In Section 5, we show via Monte Carlo simulations that, while the IVX-AR method suffers from obvious size distortion when $u_{0,t}$ is highly persistent, our modified test has a well-controlled size.

Different from Yang et al. (2020), we do not use estimated residuals to perform the Cochrane-Orcutt correction, as this method is problematic when $u_{0,t}$ is persistent. The problem can be illustrated by following simple example.

Example 4.1. *Suppose that $u_{0,t}$ and X_t are generated by*

$$\begin{aligned} u_{0,t} &= \rho_{1,T} u_{0,t-1} + z_{0,t}, \rho_{1,T} = 1 + \frac{c_u}{T}, c_u < 0, \\ X_t &= \left(1 + \frac{C_x}{T}\right) X_{t-1} + \varepsilon_{1,t}, \varepsilon_{1,t} = \sum_{j=0}^{\infty} C_{1,j} z_{1,t-j}, C_x < 0, \end{aligned}$$

Further suppose that $\hat{\rho}_{1,T}$ is calculated by regressing $e_{0,t}$ on its first lag as in Yang et al. (2020). Then, according to Phillips and Ouliaris (1990), $T\left(\hat{\rho}_{1,T} - \rho_{1,T}\right) = O_p(1)$.

Note that

$$\tilde{u}_{0,t} = e_{0,t} - \hat{\rho}_{1,T} e_{0,t-1} = e_{0,t} - \left[1 + O_p(T^{-1})\right] e_{0,t-1} = \Delta e_{0,t} + o_p(1).$$

Furthermore,

$$\Delta e_{0,t} = \Delta\left(y_t^{\mu} - \hat{\beta} X_{t-1}^{\mu}\right) = \Delta y_t^{\mu} - \hat{\beta} \Delta X_{t-1}^{\mu} = z_{0,t} - \hat{\beta}\varepsilon_{1,t} + o_p(1).$$

Clearly, $\tilde{u}_{0,t}$ can have serial correlation via the linear process $\varepsilon_{1,t} = \sum_{j=0}^{\infty} C_{1,j} z_{1,t-j}$, as $\hat{\beta}$ does not converge to zero in probability as shown in Theorem 3.1. Phillips and Magdalinos (2009) showed that, when serial correlated errors are in presence, recentering the IVX estimator (as in $\check{\beta}_{IVX}$) is required for the Wald statistic to obtain a χ^2 limiting distribution. This suggests that the Wald statistic based on the corrected time series does not converge to $\chi^2(k)$ as no recentering is performed *after the Cochrane-Orcutt correction.*

Remark 4.1. *The problem persists if $u_{0,t}$ is an AR(p) process with $p \geq 2$. For example, consider $p = 2$ and we obtain $\hat{\rho}_{1,T}, \hat{\rho}_2$ from $e_{0,t} = \hat{\rho}_{1,T} e_{0,t-1} + \hat{\rho}_2 \Delta e_{0,t-1} + \hat{v}_t$. Consider the scenario (ii) in Lemma 2.1. Let $\hat{U}_{0,t-1} = [e^{\mu}_{t-1}, \Delta e^{\mu}_{t-1}, ..., \Delta e^{\mu}_{t-p+1}]$. We can express*

$$\begin{aligned}\hat{\rho} - \rho &= \left[\sum_{t=3}^{T} \hat{U}_{0,t-1}\hat{U}'_{0,t-1}\right]^{-1} \sum_{t=3}^{T} \hat{U}_{0,t-1}\left(e_{0,t} - \hat{U}'_{0,t-1}\rho\right) \\ &= \begin{bmatrix} \sum_{t=3}^{T} e^{\mu 2}_{0,t-1} & \sum_{t=3}^{T} e^{\mu}_{0,t-1}\Delta e^{\mu}_{0,t-1} \\ \sum_{t=3}^{T} e^{\mu}_{0,t-1}\Delta e^{\mu}_{0,t-1} & \sum_{t=3}^{T} \Delta e^{\mu 2}_{0,t-1} \end{bmatrix}^{-1} \begin{pmatrix} \sum_{t=3}^{T} e^{\mu}_{0,t-1}\left(e_{0,t} - \hat{U}'_{0,t-1}\rho\right) \\ \sum_{t=3}^{T} \Delta e^{\mu}_{0,t-1}\left(e_{0,t} - \hat{U}'_{0,t-1}\rho\right) \end{pmatrix}.\end{aligned}$$

Note that

$$\begin{aligned}&\frac{1}{T}\sum_{t=3}^{T} \Delta e^{\mu}_{0,t-1}\left(e_{0,t} - \hat{U}'_{0,t-1}\rho\right) \\ =\ &\frac{1}{T}\sum_{t=3}^{T} \varepsilon_{0,t-1} z_{0,t} + \left(\beta - \hat{\beta}\right)' \frac{1}{T}\sum_{t=3}^{T} \varepsilon_{0,t-1}\left(\varepsilon_{1,t} - \rho_2 \varepsilon_{1,t-1}\right) + \left(\beta - \hat{\beta}\right)' \frac{1}{T}\sum_{t=3}^{T} \varepsilon_{1,t-2} z_{0,t} \\ &+ \left(\beta - \hat{\beta}\right)' \frac{1}{T}\sum_{t=3}^{T} \varepsilon_{1,t-2}\left(\varepsilon_{1,t} - \rho_2 \varepsilon_{1,t-1}\right)\left(\beta - \hat{\beta}\right) + o_p(1).\end{aligned}$$

The above expression involves $\frac{1}{T}\sum_{t=3}^{T} \varepsilon_{0,t-1}\varepsilon_{1,t}$ and $\frac{1}{T}\sum_{t=3}^{T} \varepsilon_{0,t-1}\varepsilon_{1,t-1}$, which have non-zero probability limits as $\beta - \hat{\beta} = O_p(1)$ from Panel 1 of Lemma 2.1. Hence, $\hat{\rho}_2$ is an inconsistent estimator of ρ_2. Together with the previous example, it is clear that the Cochrane-Orcutt correction does not remove the serial correlation in the errors. As a result, using residuals to obtain the Cochrane-Orcutt corrected IVX-AR estimator cannot provide the asymptotically $\chi^2(k)$ distributed Wald statistic. The same argument applies to an AR(p) model (5) with $p > 2$.

5 Monte Carlo Simulation

In this section, we design Monte Carlo experiments to evaluate the finite sample performance of the proposed method and make comparisons with some existing methods, including the original

IVX-Wald statistic of Kostakis et al. (2015) and the IVX-AR Wald test of Yang et al. (2020). We reject $H_0(\beta = 0)$ if the Wald statistic is higher than the 95% critical value of the $\chi^2(k)$ distribution. The number of replications is set to 2500.

5.1 Data generating processes

In our simulations, we consider the following DGP:

$$y_t = \beta' X_{t-1} + u_{0,t}, t = 1, 2, ..., T, X_t = \Phi_T X_{t-1} + \varepsilon_{1,t}, X_0 = 0, \tag{8}$$

where β, X_t and $\varepsilon_{1,t}$ are $k \times 1$ vectors, y_t and $u_{0,t}$ are scalars.[11] Three empirically relevant specifications are considered.

1. The error term $u_{0,t}$ is assumed to have an AR(1) form:

   $$u_{0,t} = \rho_T u_{0,t-1} + \varepsilon_{0,t},$$

   with $\rho_T = 1+c_u/T$, $c_u \in \{-30, -10, -5, 0\}$ and $u_{0,0} = 0$. X_t is a scalar with the autoregressive parameter

   $$\Phi_T = 1 + c_x/T, c_x \in \{-20, -10, -5, -1, 0, 1, 3\}. \tag{9}$$

   The error term vector $\varepsilon_t = (\varepsilon_{0,t}, \varepsilon_{1,t})'$ has following distribution:

   $$\varepsilon_t \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix}\right). \tag{10}$$

2. The error term $u_{0,t}$ is assumed to have an AR(2) form

   $$u_{0,t} = \rho_{1,T} u_{0,t-1} + \rho_2 \Delta u_{0,t-1} + \varepsilon_{0,t}, \tag{11}$$

[11] In simulation, we either assume the AR order in the $u_{0,t}$ series is known or select it by BIC. To choose the AR order by BIC, we fit the AR(p) model to the demeaned y_t with $p = 1, ..., P_{max}$ and then select the model with the smallest BIC. We set $P_{max} = 5$ in our simulation.

with $\rho_{1,T} = 1 + c_u/T$, $c_u \in \{-30, -10, -5, 0\}$, $\rho_2 = 0.4$ and $u_{0,0} = 0$. X_t is a scalar, Φ_T and ε_t have the same forms as (9) and (10).

3. The error term $u_{0,t}$ is the same as in (11). X_t and $\varepsilon_{1,t} = (\varepsilon_{1,t}^{(1)}, \varepsilon_{1,t}^{(2)}, \varepsilon_{1,t}^{(3)}, \varepsilon_{1,t}^{(4)})'$ is a 4×1 vectors, $\Phi_T = Diag(1, 1, 0.9887, 0.9191)$, and $(\varepsilon_{0,t}, \varepsilon_{1,t}')' \sim N(\mathbf{0}, \Sigma)$ with

$$\Sigma = \begin{bmatrix} 0.0001 & -0.0016 & 0.0003 & 0 & 0.0006 \\ -0.0016 & 1.9110 & 0.1694 & -0.1893 & -0.0554 \\ 0.0003 & 0.1694 & 0.0504 & -0.0225 & 0.0043 \\ 0 & -0.1893 & -0.0225 & 0.4146 & 0.0142 \\ 0.0006 & -0.0554 & 0.0043 & 0.0142 & 0.1250 \end{bmatrix}.$$

5.2 Empirical size

Table 2 reports the empirical size of alternative methods at the 5% level of significance with $T = 200, 500$.[12] Several interesting conclusions can be drawn from Table 2. Firstly, the IVX method does not have a well-controlled size when the level of persistency of the error term is high. This result confirms the finding in Yang et al. (2020). Secondly, although the IVX-AR method has a better controlled size than the IVX method when ρ_T is near 1, its size is not close to the nominal level. This is especially true when the predictor is explosive (i.e., $c_x = 3$). Furthermore, from Panel 2 of Table 2, we observe that the empirical rejection rate is 17.8%, significantly higher than 5% when $T = 500$, $c_x = 3$ and $\rho_T = 1$. Thirdly and most importantly, both the modified IVX-AR method based on the true lag length and its BIC counterpart have a well-controlled size under various combinations of T, c_x and ρ_T.

[12] The results with $T = 100$ under DGP1 and DGP2 can be found in Section D of Online Supplement.

Table 2: Empirical sizes under DGP1

Panel 1: $T = 200$		c_x						
		-20	-10	-5	-1	0	1	3
$\rho_T = 1 - 30/T$	IVX	0.880	0.819	0.750	0.654	0.602	0.054	0.567
	IVX-AR	0.062	0.070	0.074	0.074	0.081	0.081	0.088
	Modified IVX-AR	0.042	0.043	0.040	0.042	0.046	0.051	0.051
	Modified IVX-AR (BIC)	0.041	0.043	0.040	0.043	0.047	0.052	0.051
$\rho_T = 1 - 10/T$	IVX	0.854	0.854	0.845	0.810	0.783	0.726	0.742
	IVX-AR	0.057	0.058	0.062	0.068	0.074	0.083	0.144
	Modified IVX-AR	0.050	0.049	0.048	0.047	0.047	0.050	0.055
	Modified IVX-AR (BIC)	0.048	0.049	0.047	0.047	0.047	0.050	0.054
$\rho_T = 1 - 5/T$	IVX	0.803	0.828	0.845	0.845	0.826	0.791	0.796
	IVX-AR	0.054	0.056	0.062	0.068	0.069	0.087	0.179
	Modified IVX-AR	0.053	0.050	0.051	0.052	0.050	0.052	0.066
	Modified IVX-AR (BIC)	0.051	0.049	0.051	0.052	0.050	0.052	0.065
$\rho_T = 1$	IVX	0.664	0.735	0.789	0.848	0.855	0.866	0.892
	IVX-AR	0.050	0.052	0.056	0.059	0.066	0.073	0.215
	Modified IVX-AR	0.049	0.052	0.052	0.050	0.052	0.052	0.061
	Modified IVX-AR (BIC)	0.048	0.050	0.051	0.049	0.051	0.051	0.059
Panel 2: $T = 500$		c_x						
		-20	-10	-5	-1	0	1	3
$\rho_T = 1 - 30/T$	IVX	0.944	0.913	0.877	0.811	0.771	0.729	0.714
	IVX-AR	0.056	0.058	0.061	0.060	0.064	0.074	0.101
	Modified IVX-AR	0.045	0.045	0.044	0.044	0.046	0.052	0.064
	Modified IVX-AR (BIC)	0.044	0.044	0.043	0.043	0.045	0.050	0.064
$\rho_T = 1 - 10/T$	IVX	0.914	0.912	0.896	0.868	0.861	0.847	0.824
	IVX-AR	0.052	0.056	0.054	0.060	0.060	0.067	0.135
	Modified IVX-AR	0.050	0.050	0.051	0.049	0.048	0.049	0.064
	Modified IVX-AR (BIC)	0.049	0.049	0.050	0.048	0.047	0.047	0.063
$\rho_T = 1 - 5/T$	IVX	0.878	0.900	0.894	0.893	0.882	0.867	0.876
	IVX-AR	0.050	0.054	0.059	0.054	0.058	0.068	0.150
	Modified IVX-AR	0.050	0.051	0.051	0.049	0.048	0.052	0.067
	Modified IVX-AR (BIC)	0.049	0.049	0.049	0.048	0.047	0.051	0.066
$\rho_T = 1$	IVX	0.785	0.843	0.872	0.902	0.912	0.914	0.928
	IVX-AR	0.058	0.058	0.053	0.056	0.055	0.061	0.178
	Modified IVX-AR	0.052	0.054	0.054	0.052	0.052	0.055	0.056
	Modified IVX-AR (BIC)	0.052	0.053	0.053	0.051	0.051	0.054	0.055

Notes: This table reports the empirical rejection rates of the original IVX test of Kostakis et al. (2015), the IVX-AR test of Yang et al. (2020), our modified IVX-AR test based on the true AR order and the AR order selected by BIC.

Table 3: Empirical sizes under DGP2

Panel 1: $T = 200$		c_x						
		-20	-10	-5	-1	0	1	3
$\rho_T = 1 - 30/T$	IVX	0.444	0.477	0.474	0.468	0.526	0.594	0.674
	IVX-AR	0.076	0.076	0.076	0.080	0.067	0.063	0.066
	Modified IVX-AR	0.051	0.044	0.042	0.053	0.046	0.040	0.038
	Modified IVX-AR (BIC)	0.052	0.045	0.041	0.052	0.044	0.039	0.036
$\rho_T = 1 - 10/T$	IVX	0.617	0.666	0.674	0.696	0.745	0.782	0.818
	IVX-AR	0.113	0.126	0.139	0.097	0.098	0.121	0.156
	Modified IVX-AR	0.056	0.051	0.046	0.047	0.044	0.040	0.040
	Modified IVX-AR (BIC)	0.056	0.051	0.046	0.046	0.043	0.039	0.039
$\rho_T = 1 - 5/T$	IVX	0.700	0.740	0.770	0.764	0.800	0.814	0.844
	IVX-AR	0.126	0.167	0.190	0.135	0.154	0.199	0.262
	Modified IVX-AR	0.064	0.058	0.055	0.044	0.041	0.042	0.041
	Modified IVX-AR (BIC)	0.063	0.058	0.053	0.043	0.040	0.042	0.040
$\rho_T = 1$	IVX	0.820	0.863	0.892	0.866	0.863	0.828	0.794
	IVX-AR	0.159	0.264	0.336	0.321	0.367	0.378	0.383
	Modified IVX-AR	0.066	0.075	0.067	0.044	0.042	0.040	0.040
	Modified IVX-AR (BIC)	0.066	0.074	0.065	0.043	0.041	0.039	0.039
Panel 2: $T = 500$		c_x						
		-20	-10	-5	-1	0	1	3
$\rho_T = 1 - 30/T$	IVX	0.636	0.645	0.650	0.672	0.724	0.769	0.818
	IVX-AR	0.074	0.074	0.080	0.070	0.068	0.074	0.081
	Modified IVX-AR	0.051	0.046	0.050	0.049	0.048	0.045	0.042
	Modified IVX-AR (BIC)	0.051	0.046	0.050	0.049	0.048	0.044	0.042
$\rho_T = 1 - 10/T$	IVX	0.741	0.784	0.803	0.812	0.847	0.862	0.884
	IVX-AR	0.099	0.122	0.129	0.125	0.156	0.182	0.254
	Modified IVX-AR	0.059	0.053	0.046	0.046	0.044	0.042	0.038
	Modified IVX-AR (BIC)	0.059	0.053	0.046	0.043	0.042	0.040	0.036
$\rho_T = 1 - 5/T$	IVX	0.795	0.838	0.857	0.856	0.876	0.888	0.898
	IVX-AR	0.124	0.160	0.180	0.202	0.261	0.316	0.410
	Modified IVX-AR	0.062	0.061	0.046	0.044	0.041	0.040	0.038
	Modified IVX-AR (BIC)	0.062	0.060	0.046	0.042	0.039	0.038	0.036
$\rho_T = 1$	IVX	0.893	0.923	0.930	0.919	0.898	0.892	0.870
	IVX-AR	0.147	0.424	0.321	0.476	0.526	0.544	0.527
	Modified IVX-AR	0.068	0.075	0.068	0.040	0.040	0.039	0.039
	Modified IVX-AR (BIC)	0.068	0.076	0.067	0.038	0.038	0.037	0.037

Notes: Same as Table 2.

Table 3 reports the empirical rejection rates of alternative tests under DGP2 for $T = 200, 500$. It can be seen that the IVX-AR method has a noticeable size distortion when $\rho_T \in \{1 - 10/T, 1 - 5/T, 1\}$ for various T and c_x. On the other hand, the modified IVX-AR methods enjoy the best finite sample performance, with its empirical rejection rates closest to 5% among all the methods.

Table 4: Empirical sizes under DGP3 with $T = 100, 200, 500$

		$T = 100$	$T = 200$	$T = 500$
$\rho_T = 1 - 30/T$	IVX	0.631	0.881	0.980
	IVX-AR	0.119	0.082	0.079
	Modified IVX-AR	0.044	0.044	0.049
	Modified IVX-AR (BIC)	0.044	0.044	0.049
$\rho_T = 1 - 10/T$	IVX	0.910	0.969	0.995
	IVX-AR	0.123	0.107	0.135
	Modified IVX-AR	0.038	0.048	0.043
	Modified IVX-AR (BIC)	0.039	0.048	0.042
$\rho_T = 1 - 5/T$	IVX	0.956	0.986	0.999
	IVX-AR	0.124	0.123	0.178
	Modified IVX-AR	0.035	0.038	0.043
	Modified IVX-AR (BIC)	0.036	0.038	0.043
$\rho_T = 1$	IVX	0.988	0.995	0.998
	IVX-AR	0.144	0.178	0.358
	Modified IVX-AR	0.043	0.043	0.046
	Modified IVX-AR (BIC)	0.042	0.043	0.045

Notes: Same as Table 2

Table 4 reports the empirical size under DGP3 with $T = 100, 200, 500$. It shows that when X_t is a 4×1 vector with an empirically relevant value of Φ_T, the IVX-AR method does not have a well-controlled size if ρ_T is closed to unity, with the size distortion exacerbating as ρ_T moves towards unity. On the other hand, our modified IVX-AR test and its BIC version again have the best size control.

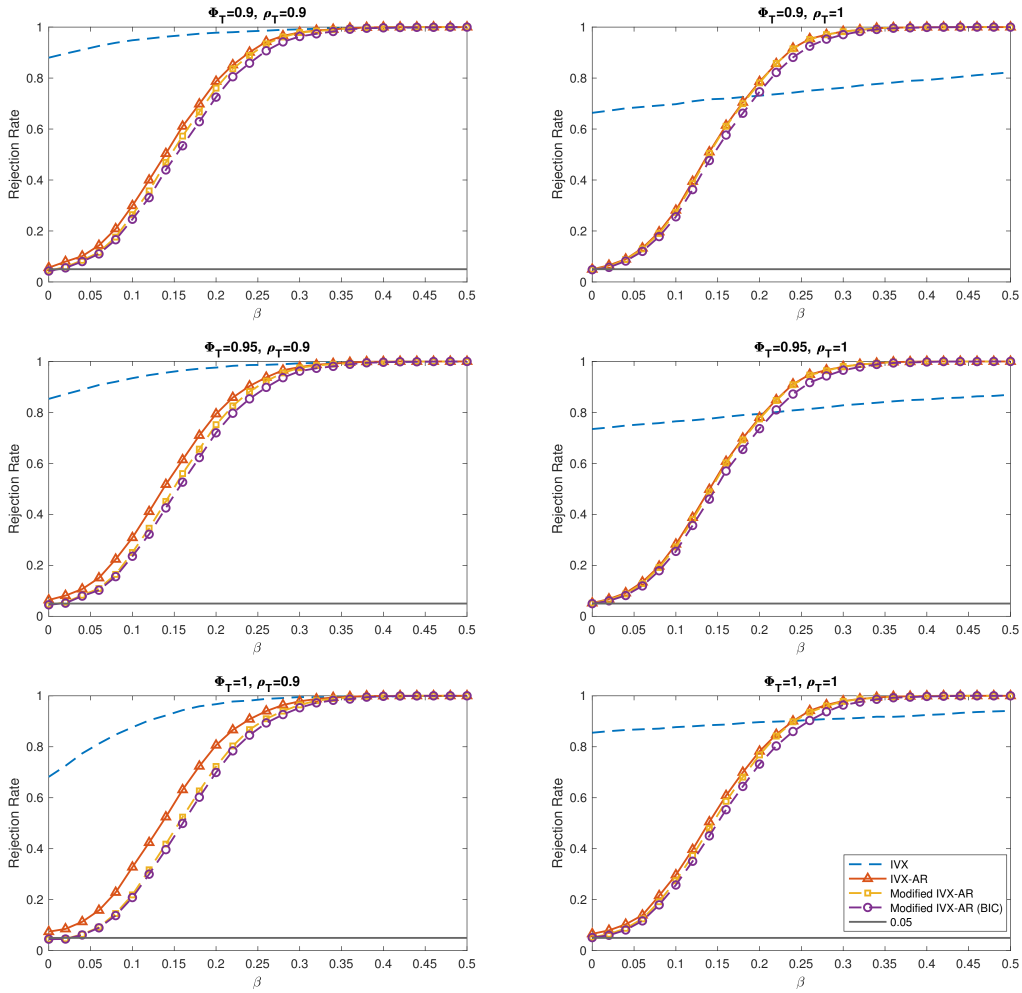

Figure 1: This figure plots the empirical rejection rates under DGP1 with $\beta \in \{0, 0.02, 0.04, ..., 0.50\}$ with $\Phi_T \in \{0.9, 0.95, 1\}$ and $\rho_T \in \{0.9, 1\}$. The dashed line represents the empirical rejection rate of IVX method. The lines with triangles, squares and circles represents the rejection rate of IVX-AR, Modified IVX-AR and Modified IVX-AR (BIC) methods, respectively. The solid line shows the nominal level.

5.3 Empirical power

To investigate the empirical power of various methods, we consider following parameter values for the first two DGPs: $T = 200, \beta \in \{0, 0.02, 0.04, ..., 0.50\}$, $\Phi_T \in \{0.9, 0.95, 1\}$,$\rho_T \in \{0.9, 1\}$ under DGP1, $\theta_{1,T} \in \{0.9, 1\}$ and $\theta_2 = 0.4$ under DGP2.

Figure 1 shows the finite sample rejection rates under DGP1. The dashed line represents the empirical rejection rate of IVX method. The lines with triangles, squares and circles represents the rejection rate of IVX-AR, Modified IVX-AR and Modified IVX-AR (BIC) methods, respectively. The solid line highlights the 0.05 nominal level. From this figure, it can be observed that the IVX method has a high rejection rate when β is near 0 under various values of Φ_T and ρ_T. However, when β deviates from 0, the IVX-AR method, the modified IVX-AR and modified IVX-AR (BIC) methods perform either similarly to or better than the IVX method. This result reinforces the findings in Yang et al. (2020). Note that, under DGP1, the IVX-AR method, the modified IVX-AR and modified IVX-AR (BIC) methods provide similar finite sample performance in rejecting the null hypothesis $\beta = 0$.

Figure 2 shows the finite sample rejection rates under DGP2 and chosen parameter values. With an AR(2) errors in DGP2, the rejection rate of IVX-AR method is higher than those of the modified methods, which should not be surprising given the size distortion shown for this method in Table 3 and rejection rate in Figure 2 at $\beta = 0$. However, the differences in empirical power diminishes as the value of β increases. When β is greater than 0.3, the IVX-AR and the modified IVX-AR methods have virtually identical rejection rates.

To investigate the finite sample power under the DGP3 with the multivariate regression, we conduct simulation and report the power curves in Figure 3. The four panels in this figure report the rejection rate of simulation under 4 different slope parameter vectors: $\beta = (j/1000, 0, 0, 0)'$ (north-west panel), $\beta = (0, j/1000, 0, 0)'$ (north-east panel), $\beta = (0, 0, j/1000, 0)'$ (south-west panel) and $\beta = (0, 0, 0, j/1000)'$ (south-east panel), with $j = 1, 2, ..., 20$, so the non-zero coefficient ranges from 0 to 0.02.

The results for multivariate regression agree with those under the scalar predictor. As shown in the north-east, south-west and south-east panels in Figure 2, the IVX method has a severe size distortion and the IVX-AR test may as well suffer from a noticeable oversize. Moreover, although the IVX-AR method enjoys a higher empirical power than the modified IVX-AR methods, the

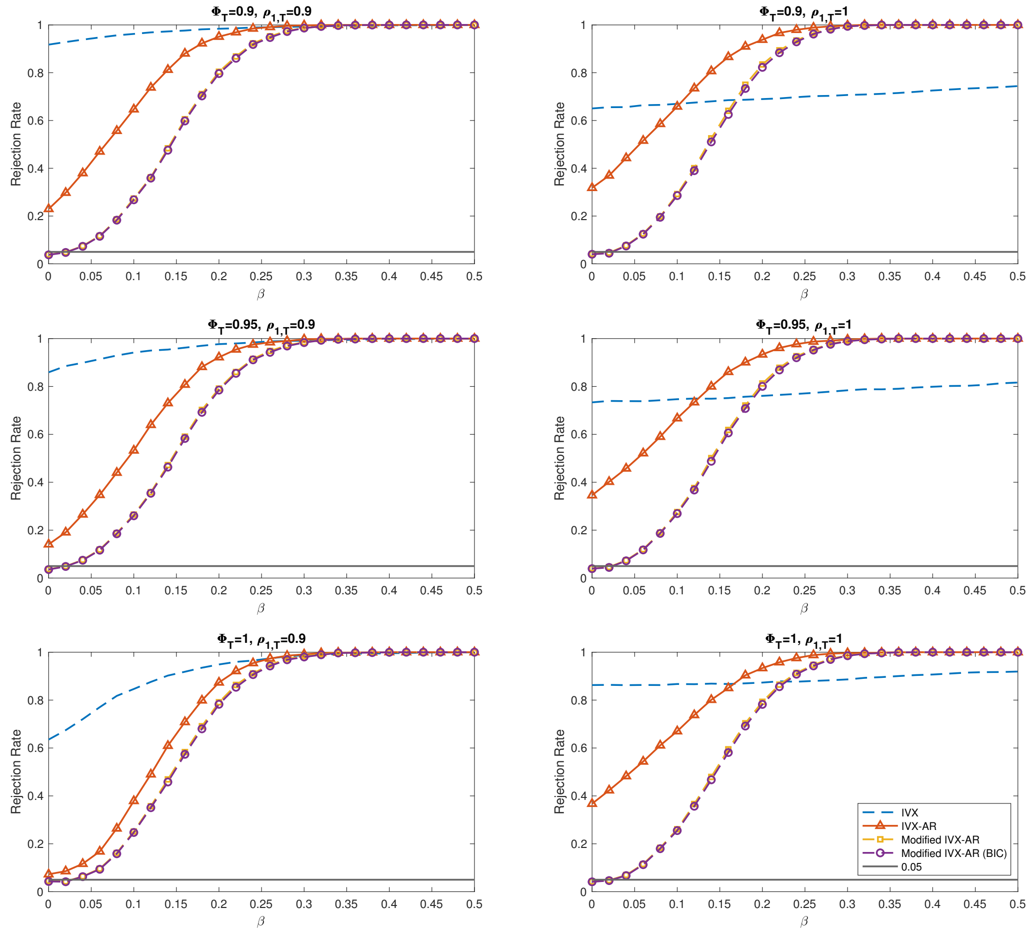

Figure 2: This figure plots the empirical rejection rates under DGP2 with $\beta \in \{0, 0.02, 0.04, ..., 0.50\}$ with $\Phi_T \in \{0.9, 0.95, 1\}$, $\rho_{1,T} \in \{0.9, 1\}$ and $\rho_2 = 0.4$. The dashed line represents the empirical rejection rate of IVX method. The lines with triangles, squares and circles represents the rejection rate of IVX-AR, Modified IVX-AR and Modified IVX-AR (BIC) methods, respectively. The solid line shows the nominal level.

difference quickly diminishes as the value of the slope coefficient increases. Finally, the modified IVX-AR method under the assumption of knowing p has a performance similar to choosing p by BIC. It can be seen from Figure 3 that the lines with squares and circles overlap with each other in all 4 panels.

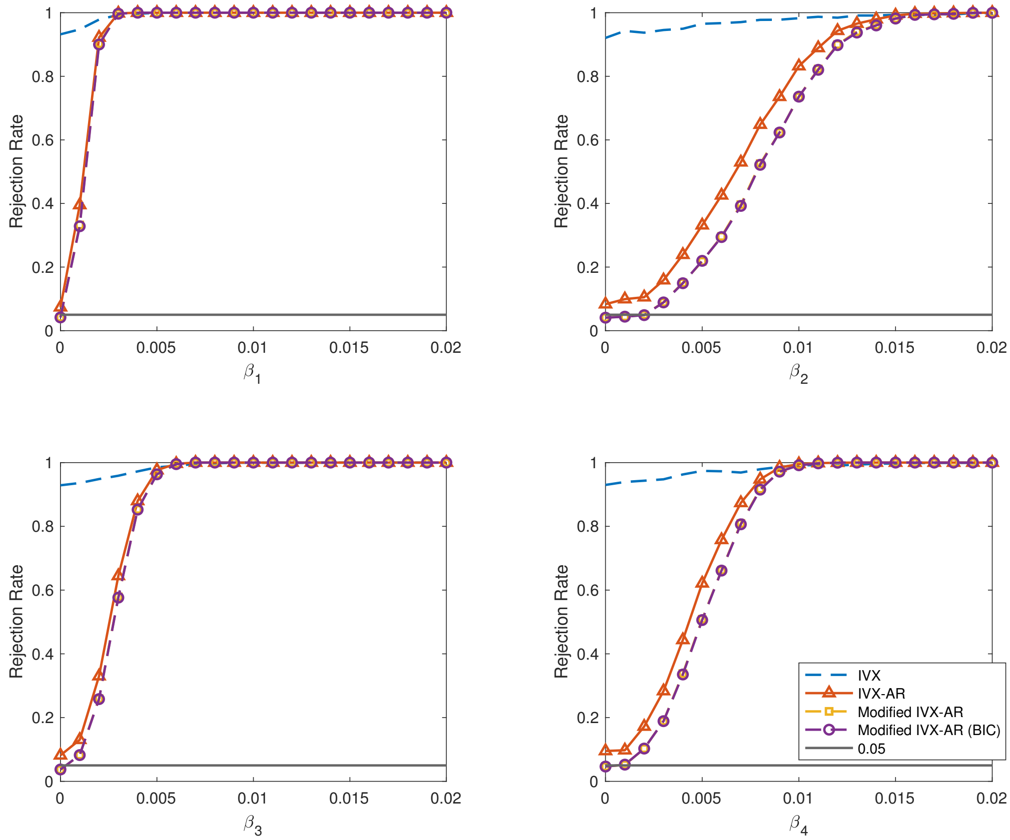

Figure 3: This figure plots the empirical rejection rates under DGP3 with 4 different slope parameter vectors, and each panel in the figure reports the rejection rate for one slope parameter vectors. The slope vector considered are: $\beta = (j/1000, 0, 0, 0)'$ (north-west panel), $\beta = (0, j/1000, 0, 0)'$ (north-east panel), $\beta = (0, 0, j/1000, 0)'$ (south-west panel) and $\beta = (0, 0, 0, j/1000)'$ (south-east panel) with $j = 1, 2, ..., 20$. The dashed line represents the empirical rejection rate of IVX method. The lines with triangles, squares and circles represents the rejection rate of IVX-AR, Modified IVX-AR and Modified IVX-AR (BIC) methods, respectively. The solid line shows the nominal level.

6 Empirical Studies

As an empirical illustration, we re-examine the predictability of quarterly growth rate of U.S. housing price index (HPI), which has recently been studied by Yang et al. (2020). Historical data of HPI is collected from the Federal Housing Finance Agency (FHFA) and the time span is from 1975:Q1 to 2023:Q3. The dependent variable, the quarterly growth rate of HPI, is computed as the percentage change of HPI level. Instead of using the full sample, we consider two sub-samples, namely sub-sample I (1975Q1-2005Q4) and sub-sample II (2006Q1-2023Q3). The reason for such a choice of sample periods is that the HPI series seems to experience a structural break in 2005 to 2006. The sample before 2006 is relatively stable with only mild fluctuations, while the period after 2006 consists of dramatic shocks like the Global Financial Crisis and the Covid19 pandemic.

As in Yang et al. (2020) we consider ten candidate predictors, including the consumer price index with all items less shelter for all urban consumers (Index 1982-1984 = 100, CPI), the implicit price deflator of gross domestic product (Index 2012=100, DEF), the percentage change of gross domestic product from last period (GDP), the percentage change of real disposable personal income from the same quarter in last year (INC), the industrial production index (Index 2012 = 100, IND), the effective Federal funds rate (IND), the shares of the residential fixed investment in GDP (INV), the 30-year mortgage rate (MOG), the total reserve balances maintained with the Federal Reserve banks (RES),[13] and the civilian unemployment rate (UNE). The time series of the predictors as well as the quarterly growth rate of HPI are plotted in Figure 4, with recessions defined by the National Bureau of Economic Research (NBER) shown as the shaded area and the ending time of first sub-sample (2005:Q4) indicated by dashed line. It can be seen that most predictors are highly persistent.

[13] The Board of Governors discontinued the release of RES on September 17, 2020. All results and plots below involving RES are obtained using data up to 2020Q3.

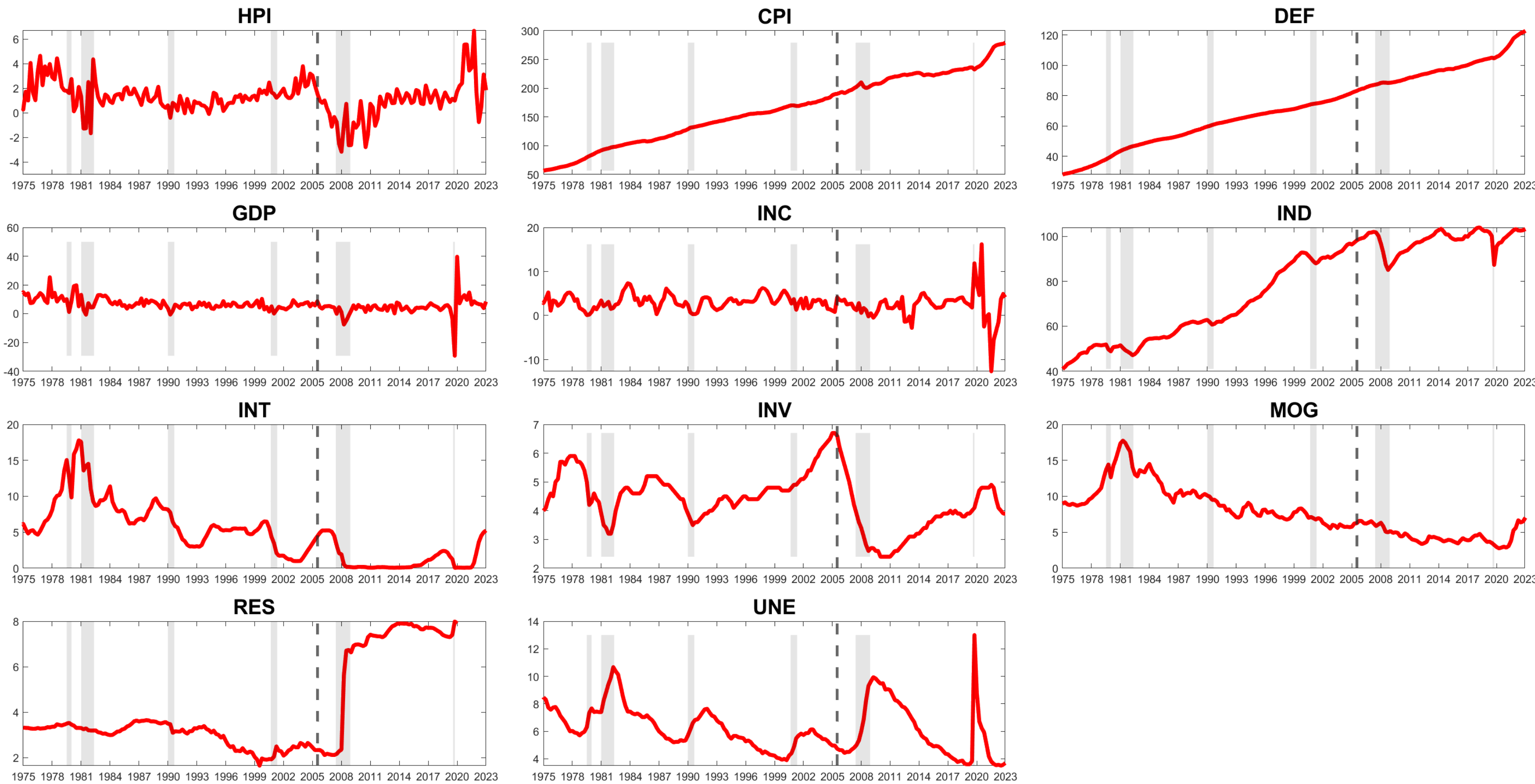

Figure 4: Time series plots of HPI growth rate and ten predictors (1975:Q1-2023:Q3).

To illustrate potential spurious regression, in Table 5 we report the largest autoregressive root of the OLS residuals from univariate predictive regression, which is high, especially in the sub-sample from 2006 to 2023. Indeed, as indicated by three common tests for unit root and stationarity, which are reported in Table 5, many residual sequences seem to exhibit high degree of persistency.[14] Although in some cases these tests yield contradictory results, it is in general plausible to conclude that most residual series are persistent. Therefore, conventional tests that are designed to deal with the mds or stationary autoregressive errors may not be applicable.

We first consider univariate predictive regressions. The results are reported in Table 6. To make a comparison, we report the results obtained from the original IVX test, the IVX-AR test, and our modified IVX-AR test.[15] In line with Yang et al. (2020), we find that among the 10 predictors they considered, the only one that is significant for both sample periods is the shares of the residential fixed investment in GDP. The difference is that, for the first sub-sample, our modified test statistic is much smaller than its IVX-AR counterpart and INV in this case is significant only at the 10% level after modification. Nevertheless, our analysis confirms that INV is in general a fairly robust predictor for the HPI growth rate. It is worthwhile to emphasize that for the post-2006 sub-sample, we find two additional variables, DEF and RES become significant at the 10% level. A notable feature is that all these indicators are closely related to monetary policy. Our findings hence suggest that in past 17 years, the monetary factor has a larger impact on the housing price movement than the macroeconomic fundamentals such as GDP and industrial production.

[14] All three tests consider in Table 5 may not be theoretically justified when the error term in the predictive regression is persistent. The results hence are only informal and serve as indicators for strong persistency.

[15] In the empirical studies we choose the AR order in the $u_{0,t}$ series by BIC. In particular, we fit the AR(p) model to the demeaned y_t with $p = 1, ..., 5$ and then select the model with the smallest BIC.

Table 5: Stationarity/Unit root test for residuals from univariate predictive regressions.

	Panel 1: Sub-sample I (1975:Q1-2005:Q4)				Panel 2: Sub-sample II (2006:Q1-2023:Q3)			
	Largest AR root	ADF	PP	KPSS	Largest AR root	ADF	PP	KPSS
CPI	0.753	-2.881***	-7.580***	0.349*	0.537	-2.269**	-4.371***	0.200*
DEF	0.751	-2.902***	-7.608***	0.348*	0.515	-2.306**	-4.502***	0.130*
GDP	0.739	-3.092***	-8.174***	0.335*	0.896	-1.347	-3.954***	1.219***
INC	0.765	-2.831***	-7.473***	0.427*	0.704	-1.485	-3.101***	1.172***
IND	0.767	-2.831***	-7.424***	0.419*	0.636	-1.842*	-3.681***	0.875***
INT	0.762	-2.902***	-7.514***	0.562**	0.697	-1.470	-3.110***	1.130***
INV	0.399	-9.553***	-9.811***	0.799***	0.893	-1.781*	-3.545***	1.182***
MOG	0.763	-2.904***	-7.617***	0.620**	0.898	-1.536	-3.724***	0.847***
RES	0.765	-2.856***	-7.471***	0.466**	0.873	-1.843*	-3.538***	0.475**
UNE	0.766	-2.844***	-7.435***	0.462*	0.649	-1.714*	-3.565***	0.906***

Notes: *,** and *** denote rejection of null hypothesis at the 10%, 5% and 1% level, respectively.

Table 6: Results of univariate predictive regressions.

	Panel 1: Sub-sample I (1975:Q1-2005:Q4)			Panel 2: Sub-sample II (2006:Q1-2023:Q3)		
	IVX	IVX-AR	modified IVX-AR	IVX	IVX-AR	modified IVX-AR
CPI	9.171*** (0.002)	2.046 (0.153)	1.919 (0.166)	13.093*** (0.000)	5.491** (0.019)	1.972 (0.160)
DEF	10.486*** (0.001)	2.404 (0.121)	2.188 (0.139)	20.080*** (0.000)	9.893*** (0.002)	3.140* (0.076)
GDP	6.874*** (0.009)	0.129 (0.719)	0.224 (0.636)	8.200*** (0.004)	1.688 (0.194)	0.028 (0.867)
INC	1.819 (0.177)	0.122 (0.727)	0.124 (0.724)	0.130 (0.718)	0.455 (0.500)	0.451 (0.502)
IND	0.018 (0.894)	0.034 (0.854)	0.029 (0.865)	9.649*** (0.002)	0.665 (0.415)	0.467 (0.494)
INT	0.481 (0.488)	0.344 (0.558)	0.345 (0.557)	1.774 (0.183)	0.355 (0.551)	0.505 (0.477)
INV	55.876*** (0.000)	15.925*** (0.000)	3.595* (0.058)	22.667*** (0.000)	18.301*** (0.000)	8.668*** (0.003)
MOG	0.859 (0.354)	1.192 (0.275)	1.116 (0.291)	1.899 (0.168)	9.388*** (0.002)	0.011 (0.917)
RES	0.054 (0.817)	0.097 (0.756)	0.092 (0.762)	0.417 (0.518)	3.958** (0.047)	2.847* (0.092)
UNE	0.473 (0.492)	0.002 (0.965)	0.003 (0.959)	11.673*** (0.001)	0.553 (0.457)	0.457 (0.499)

Notes: *,** and *** denote rejection of null hypothesis at the 10%, 5% and 1% level, respectively. Numbers in the parenthesis are p-values of the test.

Now we move to multivariate regression. In the similar spirit to Yang et al. (2020), we consider following five combinations: (1) INV + DEF + RES; (2) INV + GDP + INC + IND + UNE; (3) CPI + DEF + INT + RES; (4) INV + INT + MOG; (5) A "kitchen sink" which includes all regressors. The first combination considers the joint significance of INV, DEF and RES because Table 6 shows that all these three indicators exhibit significant predictive ability for the second sub-sample. Variables in second combination are for main macroeconomic indicators and closely tracked by investors and policymakers. The third combination mainly concentrates on the monetary policy, while the fourth combination measures the cost of housing investment, which tends to be affected by interest rates. The last combination is a "kitchen sink" which considers the joint significance of all above variables.

Table 7: Stationarity/Unit root test for residuals from multivariate predictive regressions.

	Largest AR root	ADF	PP	KPSS
Panel 1: Sub-sample I (1975:Q1-2005:Q4)				
INV+DEF+RES	0.647	-12.176***	-12.461***	0.029*
INV+GDP+INC+IND+UNE	0.454	-11.033***	-11.135***	0.184*
CPI+DEF+RES+INT	0.700	-3.621***	-8.507***	0.181*
CPI+INT+MOG	0.715	-3.581***	-8.473***	0.297*
"Kitchen Sink"	0.475	-12.489***	-12.869***	0.033*
Panel 2: Sub-sample II (2006:Q1-2023:Q3)				
INV+DEF+RES	0.839	-2.577**	-5.279***	0.185*
INV+GDP+INC+IND+UNE	0.858	-2.307**	-4.454***	1.126***
CPI+DEF+RES+INT	0.820	-2.356**	-4.535***	0.151*
CPI+INT+MOG	0.394	-3.246***	-5.446***	0.076*
"Kitchen Sink"	0.762	-9.556***	-7.225***	0.045*

Notes: This table shows the largest autoregressive (AR) root as well as the results of stationarity/unit root tests for the least-squares residuals of five multivariate predictive regressions. *,** and *** denote rejection of null hypothesis at the 10%, 5% and 1% level, respectively.

Table 7 presents the largest autoregressive root of the OLS residuals from multivariate regression, as well as the associated unit root and stationarity test results. Similar to the univariate regression, the residual sequence exhibits strong persistency in many cases, particularly for the second sub-sample. For example, when the second combination is considered in the second sub-sample, the largest AR root is 0.858. With such a large AR root, it seems plausible to assume that the error term in the predictive regression is not far from a unit root process. Indeed, we observe again that three tests lead to contradictory conclusions and it is unclear whether the residual series is stationary or not.

Table 8: Results of multivariate predictive regressions.

	IVX	IVX-AR	modified IVX-AR
Panel 1: Sub-sample I (1975:Q1-2005:Q4)			
INV+DEF+RES	87.169*** (0.000)	133.198*** (0.000)	4.742 (0.192)
INV+GDP+INC+IND+UNE	71.666*** (0.000)	30.820*** (0.000)	6.035 (0.303)
CPI+DEF+RES+INT	15.333*** (0.004)	12.572** (0.014)	4.190 (0.381)
CPI+INT+MOG	12.903*** (0.005)	6.994* (0.072)	5.087 (0.166)
"Kitchen Sink"	100.726*** (0.000)	191.418*** (0.000)	11.928 (0.290)
Panel 2: Sub-sample II (2006:Q1-2023:Q3)			
INV+DEF+RES	92.390*** (0.000)	66.137*** (0.000)	14.517*** (0.002)
INV+GDP+INC+IND+UNE	25.386*** (0.000)	24.147*** (0.000)	8.595 (0.126)
CPI+DEF+RES+INT	40.441*** (0.000)	22.321*** (0.000)	7.394 (0.116)
INV+INT+MOG	54.905*** (0.000)	38.696*** (0.000)	2.544 (0.467)
"Kitchen Sink"	131.377*** (0.000)	1039.81*** (0.000)	15.892 (0.103)

Notes: This table shows the results of multivariate predictive regressions for the first sub-sample (1975:Q1-2005:Q4) and second sub-sample (2006:Q1-2023:Q3). *,** and *** denote rejection of joint null hypothesis at the 10%, 5% and 1% level respectively. Numbers in the parenthesis are p-values of the test.

The predictability test results for multivariate regression with sub-sample I are reported in Panel 1 of Table 8 and results for sub-sample II can be found in Panel 2. We find that in multivariate cases, the conclusions of predictability obtained by the IVX-AR test and the modified IVX-AR test are quite different, with the test statistic being significantly smaller in the latter case. Using either tests, we confirm that combination 1, consisting of three individually informative predictors in sub-sample II, remains highly significant for this period when considered simultaneously. This combination, however, does not display predictive power in sub-sample I, which is in line with the results of univariate regression. Surprisingly, all other combinations, including both macroeconomic and monetary factors, are shown to be insignificant based on our modified test, while the IVX-AR test implies that most of them are significant even at the 1% level. The most striking difference between IVX-AR and our approach is for the "Kitchen Sink" regression in sub-sample II, where the IVX-AR statistic takes the value of 1039.81, which is about 65 times as large as our modified IVX-AR statistic (15.892).

7 Conclusion

In this paper, we provide limit theory for predictive regression driven by persistent errors. Various popular inference methods are shown to be invalid when both the predictors and the error term are persistent. We discuss the limiting behavior of the OLS estimator, which is known to be super-consistent when the model is correctly specified but becomes inconsistent in spurious regression. Our limit theory provides a smooth transition between these two extreme cases. In addition, it is also shown that the IVX approach proposed in Kostakis et al. (2015), with or without re-centering, fails to provide correct inference if the regression error is persistent.

To provide reliable predictability test, we propose to modify the IVX-AR statistic considered in Yang et al. (2020), which is based on Cochrane-Orcutt-type correction. The asymptotic distribution of our test statistic in the absence of predictability is χ^2, regardless of the degree of persistency of the predictors and errors. Extensive simulation studies suggest that our test has satisfactory finite sample performance under various model setups.

Using the new test, we re-examine the predictability of the quarterly growth rate of the U.S. housing price index. For univariate regression, main findings of predictability are in general consistent with those in Yang et al. (2020). In particular, we confirm that the shares of the residential fixed investment in GDP is a fairly robust predictor of the growth rate of HPI, while many other macroeconomic indicators are not. For multivariate regression, on the other hand, conclusions based on our modified test are significantly different from those using the IVX-AR test. These observations testify to the relevance of our modification for empirical research.

References

Ahtola, J. and G. C. Tiao (1984). Parameter inference for a nearly nonstationary first-order autoregressive model. *Biometrika 71*(2), 263–272.

Amihud, Y. and C. M. Hurvich (2004). Predictive regressions: A reduced-bias estimation method. *Journal of Financial and Quantitative Analysis 39*(4), 813–841.

Campbell, J. Y. and M. Yogo (2006). Efficient tests of stock return predictability. *Journal of financial economics 81*(1), 27–60.

Cavanagh, C. L., G. Elliott, and J. H. Stock (1995). Inference in models with nearly integrated regressors. *Econometric Theory 11*(5), 1131–1147.

Choi, Y., S. Jacewitz, and J. Y. Park (2016). A reexamination of stock return predictability. *Journal of Econometrics 192*(1), 168–189.

Demetrescu, M., I. Georgiev, P. M. Rodrigues, and A. R. Taylor (2022). Extensions to IVX methods of inference for return predictability. *Journal of Econometrics 237*(2), 105271.

Demetrescu, M. and P. M. Rodrigues (2022). Residual-augmented IVX predictive regression. *Journal of Econometrics 227*(2), 429–460.

Giraitis, L. and P. C. B. Phillips (2006). Uniform limit theory for stationary autoregression. *Journal of Time Series Analysis 27*(1), 51–60.

Gungor, S. and R. Luger (2020). Small-sample tests for stock return predictability with possibly non-stationary regressors and garch-type effects. *Journal of Econometrics 218*(2), 750–770.

Hansen, B. E. (1992). Convergence to stochastic integrals for dependent heterogeneous processes. *Econometric Theory 8*(4), 489–500.

Harvey, D. I., S. J. Leybourne, and A. R. Taylor (2021). Simple tests for stock return predictability with good size and power properties. *Journal of Econometrics 224*(1), 198–214.

Hosseinkouchack, M. and M. Demetrescu (2021). Finite-sample size control of IVX-based tests in predictive regressions. *Econometric Theory 37*(4), 769–793.

Kostakis, A., T. Magdalinos, and M. P. Stamatogiannis (2015). Robust econometric inference for stock return predictability. *The Review of Financial Studies 28*(5), 1506–1553.

Lewellen, J. (2004). Predicting returns with financial ratios. *Journal of Financial Economics 74*(2), 209–235.

Lin, Y. and Y. Tu (2020). Robust inference for spurious regressions and cointegrations involving processes moderately deviated from a unit root. *Journal of Econometrics 219*(1), 52–65.

Lui, Y. L., W. Xiao, and J. Yu (2021). Mildly explosive autoregression with anti-persistent errors*. *Oxford Bulletin of Economics and Statistics 83*(2), 518–539.

Magdalinos, T. and P. C. B. Phillips (2009). Limit theory for cointegrated systems with moderately integrated and moderately explosive regressors. *Econometric Theory 25*(2), 482–526.

Mikusheva, A. (2007). Uniform inference in autoregressive models. *Econometrica 75*(5), 1411–1452.

Müller, U. K. (2005). Size and power of tests of stationarity in highly autocorrelated time series. *Journal of Econometrics 128*(2), 195–213.

Perron, P. (1991). A continuous time approximation to the unstable first-order autoregressive process: The case without an intercept. *Econometrica 59*(1), 211–236.

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

Phillips, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika 74*(3), 535–547.

Phillips, P. C. B. (1988). Regression theory for near-integrated time series. *Econometrica 56*(5), 1021–1043.

Phillips, P. C. B. (2014). On confidence intervals for autoregressive roots and predictive regression. *Econometrica 82*(3), 1177–1195.

Phillips, P. C. B. and J. H. Lee (2013). Predictive regression under various degrees of persistence and robust long-horizon regression. *Journal of Econometrics 177*(2), 250–264.

Phillips, P. C. B. and J. H. Lee (2016). Robust econometric inference with mixed integrated and mildly explosive regressors. *Journal of Econometrics 192*(2), 433–450.

Phillips, P. C. B. and T. Magdalinos (2007). Limit theory for moderate deviations from a unit root. *Journal of Econometrics 136*(1), 115–130.

Phillips, P. C. B. and T. Magdalinos (2009). Econometric inference in the vicinity of unity. *Singapore Management University, CoFie Working Paper 7*.

Phillips, P. C. B. and S. Ouliaris (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica 58*(1), 165–193.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics 54*(3), 375–421.

Yang, B., X. Liu, L. Peng, and Z. Cai (2021). Unified tests for a dynamic predictive regression. *Journal of Business & Economic Statistics 39*(3), 684–699.

Yang, B., W. Long, L. Peng, and Z. Cai (2020). Testing the predictability of us housing price index returns based on an ivx-ar model. *Journal of the American Statistical Association 115*(532), 1598–1619.

Online Supplement to: 'Testing Predictability in the Presence of Persistent Errors'

(Not for Publication)[1]

Yijie Fei
Hunan University

Yiu Lim Lui
Dongbei University of Finance and Economics

Jun Yu
University of Macau

This supplementary document contains the proofs of the main results in the paper as well as additional discussions with regard to the Bonferroni confidence interval and the Cauchy estimator under persistent errors.

A. Preliminary Results and Useful Lemmas

Lemma A.1. *In model (2), assume that $u_{0,t}$ satisfies (5). The OLS estimator of ρ given by (6) is consistent and has the following stochastic orders under $H_0(\beta = 0)$.*

(i). If $\rho_{1,T} = 1 + \frac{c_u}{T}$, for $k = 1, ..., p$,

$$\begin{aligned} \hat{\rho}_{1,T} - \rho_{1,T} &= O_p(T^{-1}), \\ \hat{\rho}_k - \rho_k &= O_p(T^{-1/2}) \end{aligned}$$

(ii). If $\rho_T = 1 + \frac{c_u}{T^{\kappa_u}}$, $c_u < 0, \kappa_u \in (0, 1)$, for $k = 1, ..., p$,

$$\begin{aligned} \hat{\rho}_{1,T} - \rho_{1,T} &= O_p(T^{-\frac{1+\kappa_u}{2}}), \\ \hat{\rho}_k - \rho_k &= O_p(T^{-1/2}) \end{aligned}$$

Proof of Lemma A.1. Note that the OLS estimator of ρ can be written as

$$\begin{aligned} &\hat{\rho} - \rho \\ = &\left[\sum_{t=p+1}^{T} U_{0,t-1}U_{0,t-1}'\right]^{-1} \sum_{t=p+1}^{T} U_{0,t-1}u_{0,t} - \left[\sum_{t=p+1}^{T} U_{0,t-1}U_{0,t-1}'\right]^{-1} \sum_{t=p+1}^{T} U_{0,t-1}U_{0,t-1}'\rho \end{aligned}$$

[1]Yijie Fei, College of Finance and Statistics, Hunan University, 410100, China, Email: yijiefei@hnu.edu.cn. Yiu Lim Lui, Institute for Advanced Economic Research, Dongbei University of Finance and Economics, 116000, China, Email: luiyiulim@outlook.com. Jun Yu, Faculty of Business Administration, University of Macau, Avenida da Universidade, Macao, China. Email: junyu@um.edu.mo.

$$= \left[\sum_{t=p+1}^{T} U_{0,t-1}U_{0,t-1}'\right]^{-1} \sum_{t=p+1}^{T} U_{0,t-1}\left(u_{0,t}^{\mu} - U_{0,t-1}'\rho\right),$$

where we can express

$$\sum_{t=p+1}^{T} U_{0,t-1}U_{0,t-1}' \tag{12}$$

$$= \begin{bmatrix} \sum_{t=p+1}^{T} u_{0,t-1}^{\mu 2} & \sum_{t=p+1}^{T} u_{0,t-1}^{\mu}\Delta u_{0,t-1}^{\mu} & \cdots & \sum_{t=p+1}^{T} u_{0,t-1}^{\mu}\Delta u_{0,t-p+1}^{\mu} \\ \sum_{t=p+1}^{T} u_{0,t-1}^{\mu}\Delta u_{0,t-1}^{\mu} & \sum_{t=p+1}^{T} \Delta u_{0,t-1}^{\mu 2} & \cdots & \sum_{t=p+1}^{T} \Delta u_{0,t-1}^{\mu}\Delta u_{0,t-p+1}^{\mu} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{t=p+1}^{T} \Delta u_{0,t-p+1}^{\mu} u_{0,t-1}^{\mu} & \cdots & \cdots & \sum_{t=p+1}^{T} \Delta u_{0,t-p+1}^{\mu 2} \end{bmatrix},$$

and

$$\begin{aligned} \sum_{t=p+1}^{T} U_{0,t-1}\left(u_{0,t}^{\mu} - U_{0,t-1}'\rho\right) &= \begin{bmatrix} \sum_{t=p+1}^{T} u_{0,t-1}^{\mu}\left(u_{0,t}^{\mu} - U_{0,t-1}'\rho\right) \\ \sum_{t=p+1}^{T} \Delta u_{0,t-1}^{\mu}\left(u_{0,t}^{\mu} - U_{0,t-1}'\rho\right) \\ \cdots \\ \sum_{t=p+1}^{T} \Delta u_{0,t-p+1}^{\mu}\left(u_{0,t}^{\mu} - U_{0,t-1}'\rho\right) \end{bmatrix} \\ &= \begin{bmatrix} \sum_{t=p+1}^{T} u_{0,t-1}^{\mu} z_{0,t}^{\mu} \\ \sum_{t=p+1}^{T} \Delta u_{0,t-1}^{\mu} z_{0,t}^{\mu} \\ \cdots \\ \sum_{t=p+1}^{T} \Delta u_{0,t-p+1}^{\mu} z_{0,t}^{\mu} \end{bmatrix}. \end{aligned} \tag{13}$$

(i). Note that the off-diagonal terms in $\sum_{t=p+1}^{T} U_{0,t-1}U_{0,t-1}'$ have the order $O_p(T)$, since

$$\begin{aligned} \sum_{t=p+1}^{T} u_{0,t-1}^{\mu}\Delta u_{0,t-1}^{\mu} &= \sum_{t=p+1}^{T} u_{0,t-1}^{\mu}\left(u_{0,t-1}^{\mu} - u_{0,t-2}^{\mu}\right) \\ &= \sum_{t=p+1}^{T} u_{0,t-1}^{\mu}\left(\left(\rho_{1,T} - 1\right)u_{0,t-2}^{\mu} - \rho_2\Delta u_{0,t-2}^{\mu} - ... - \rho_P\Delta u_{0,t-p+1}^{\mu}\right) \\ &= \frac{c_u}{T}\sum_{t=p+1}^{T} u_{0,t-1}^{\mu}u_{0,t-2}^{\mu} - \rho_2\sum_{t=p+1}^{T} u_{0,t-1}^{\mu}\Delta u_{0,t-2}^{\mu} - ... \\ &\quad -\rho_p\sum_{t=p+1}^{T} u_{0,t-1}^{\mu}\Delta u_{0,t-p+1}^{\mu}. \end{aligned}$$

Applying Lemma 3.1 in Phillips (1988) several times, we can obtain

$$\sum_{t=p+1}^{T} u_{0,t-1}^{\mu} \Delta u_{0,t-1}^{\mu} = O_p(T).$$

For the second to the last diagonal terms in $\sum_{t=p+1}^{T} U_{0,t-1} U_{0,t-1}'$, note that taking first difference makes $u_{0,t}^{\mu}$ stationary. Therefore, these terms have the order $O_p(T)$ by the ergodic theorem. For $\sum_{t=p+1}^{T} (u_{0,t-1}^{\mu})^2$, applying Lemma 3.1 in Phillips (1988) gives $\sum_{t=p+1}^{T} (u_{0,t-1}^{\mu})^2 = O_p(T^2)$. The above results give

$$D_T^{-1} \sum_{t=p+1}^{T} U_{0,t-1} U_{0,t-1}' D_T^{-1} = O_p(1),$$

where

$$D_T = \begin{bmatrix} T & \cdots & \cdots & \cdots \\ \cdots & \sqrt{T} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \sqrt{T} \end{bmatrix}.$$

Similarly, $\sum_{t=p+1}^{T} u_{0,t-1}^{\mu} z_{0,t}^{\mu}$ is $O_p(T)$ by Theorem 4.4 of Hansen (1992),

$$\sum_{t=p+1}^{T} \Delta u_{0,t-j-1}^{\mu} z_{0,t}^{\mu} = O_p(\sqrt{T}) \text{ for } j = 0, ..., p$$

by the CLT of the martingale difference sequence. Therefore, we have

$$D_T^{-1} \sum_{t=p+1}^{T} U_{0,t-1} \left(u_{0,t}^{\mu} - U_{0,t-1}' \rho \right) = O_p(1).$$

Consequently, we can deduce that

$$D_T \left(\hat{\rho} - \rho \right) = \left[D_T^{-1} \sum_{t=p+1}^{T} U_{0,t-1} U_{0,t-1}' D_T^{-1} \right]^{-1} D_T^{-1} \sum_{t=p+1}^{T} U_{0,t-1} \left(u_{0,t}^{\mu} - U_{0,t-1}' \rho \right) = O_p(1).$$

(ii). Note that (5) implies $u_{0,t}^{\mu} = \rho_{1,T} u_{0,t-1}^{\mu} + \rho_2 \Delta u_{0,t-1}^{\mu} + ... + \rho_p \Delta u_{0,t-p+1}^{\mu} + z_{0,t}^{\mu}$. Note also that the above process can be written as $\alpha(L) u_{0,t}^{\mu} = z_{0,t}^{\mu}$ with

$$\alpha(L) \;=\; 1 - \rho_{1,T} L - \sum_{j=1}^{p-1} \alpha_j (1-L) L^j$$

$$= \quad (1-\mu_1 L)(1-\mu_2 L)...(1-\mu_p L),$$

where $\{\mu_j\}_{j=1}^{p}$ are the inverse characteristic roots of $\alpha(L)$ and $|\mu_1| \leq |\mu_2| \leq ... \leq |\mu_{p-1}| < \delta < 1$. Lemma 8a in the supplementary appendix of Mikusheva (2007) implies that $\Delta u_{0,t}^{\mu}$ can be represented by

$$D(L)z_{0,t}^{\mu} = \sum_{j=0}^{\infty} d_j z_{0,t-j}^{\mu}, \sum_{j=0}^{\infty} |d_j| < \infty.$$

Therefore, $\{\Delta u_{0,t-1}^{\mu}, ..., \Delta u_{0,t-p+1}^{\mu}\}$ are stationary, and we can express

$$u_{0,t}^{\mu} = \rho_{1,T} u_{0,t-1}^{\mu} + \sum_{j=0}^{\infty} \tilde{d}_j z_{0,t-j}^{\mu}$$

with $\sum_{j=0}^{\infty} \left|\tilde{d}_j\right| < \infty$. For the first term in (13), applying Lemma B4 in the online appendix of Kostakis et al. (2015) gives $\sum_{t=p+1}^{T} u_{0,t-1}^{\mu} z_{0,t}^{\mu} = O_p\left(T^{\frac{1+\kappa_u}{2}}\right)$. For the second term in (13), since $\Delta u_{0,t-1}^{\mu}$ is stationary, it can be shown that $\frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} \Delta u_{0,t-1}^{\mu} z_{0,t}^{\mu} = O_p(1)$ by the CLT of the martingale difference sequence (mds). Using the same argument, we can also show $\frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} \Delta u_{0,t-2}^{\mu} z_{0,t}^{\mu}, ..., \frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} \Delta u_{0,t-p+1}^{\mu} z_{0,t}^{\mu}$ are all $O_p(1)$. For the terms in (12), using Equation (7) and (10) in Magdalinos and Phillips (2009) and the ergodic theorem, we can establish that $\sum_{t=p+1}^{T} u_{0,t-1}^{\mu 2} = O_p(T^{1+\kappa_u})$ and all other terms in (12) are $O_p(T)$. Finally, we can show

$$G_T^{-1} \sum_{t=p+1}^{T} U_{0,t-1} U_{0,t-1}' G_T^{-1} = O_p(1),$$

where

$$G_T = \begin{bmatrix} T^{\frac{1+\kappa_u}{2}} & \cdots & \cdots & \cdots \\ \cdots & \sqrt{T} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \sqrt{T} \end{bmatrix}.$$

Note that, the off-diagonal element of $G_T^{-1} \sum_{t=p+1}^{T} U_{0,t-1} U_{0,t-1}' G_T^{-1}$ is $o_p(1)$ since their orders are $O_p(T)/T^{1+\frac{\kappa_u}{2}}$. Also note that

$$G_T^{-1} \sum_{t=p+1}^{T} U_{0,t-1} \left(u_{0,t}^{\mu} - U_{0,t-1}' \rho\right) = O_p(1).$$

Thus,

$$\begin{aligned} G_T(\hat{\rho}-\rho) &= \left[G_T^{-1}\sum_{t=p+1}^{T} U_{0,t-1}U_{0,t-1}'G_T^{-1}\right]^{-1} G_T^{-1}\sum_{t=p+1}^{T} U_{0,t-1}\left(u_{0,t}^{\mu}-U_{0,t-1}'\rho\right) \\ &= O_p(1). \end{aligned}$$

This completes the proof of Lemma A.1. ■

To prove Theorem 4.1, it is useful to introduce the following lemma.

Lemma A.2. *Under model (2), assume $u_{0,t}$ follows the AR(p) structure in (5). Let*

$$\begin{aligned} \hat{\rho}_2 &= [\hat{\rho}_2, ..., \hat{\rho}_p]', \bar{\rho}_2 = [\rho_2, ..., \rho_p]', \\ \mathbf{Z}_{\Delta,t} &= [\Delta Z_t, ..., \Delta Z_{t-p}], \\ \mathbf{X}_{\Delta,t}^{\mu} &= [\Delta X_t^{\mu}, ..., \Delta X_{t-p}^{\mu}], \\ \mathbf{u}_{\Delta,t} &= [\Delta u_{0,t}, ..., \Delta u_{0,t-p}], \\ \tilde{Z}_{t-1} &= Z_{t-1} - \hat{\rho}_{1,T} Z_{0,t-2} - \mathbf{Z}_{\Delta,t-2}\hat{\rho}_2, \\ \tilde{u}_t^{\mu} &= u_t^{\mu} - \hat{\rho}_{1,T} u_{0,t-1}^{\mu} - \mathbf{u}_{\Delta,t-1}^{\mu}\hat{\rho}_2. \end{aligned}$$

The following approximations hold:

$$\begin{aligned} \sum_{t=p+1}^{T} \tilde{Z}_{t-1}\tilde{y}_t^{\mu} &= [1+o_p(1)] \sum_{t=p+1}^{T} \tilde{Z}_{t-1,\rho} z_{0,t}^{\mu}, \\ \sum_{t=p+1}^{T} \tilde{Z}_{t-1}\tilde{X}_{t-1}^{\mu\prime} &= [1+o_p(1)] \sum_{t=p+1}^{T} \tilde{Z}_{t-1,\rho}\tilde{X}_{t-1,\rho}^{\mu\prime}. \end{aligned}$$

where

$$\begin{aligned} \tilde{Z}_{t-1,\rho} &= Z_{t-1} - \rho_{1,T} Z_{t-2} - \mathbf{Z}_{\Delta,t-2}\bar{\rho}_2, \\ \tilde{X}_{t-1,\rho}^{\mu} &= X_{t-1}^{\mu} - \rho_{1,T} X_{t-2}^{\mu} - \mathbf{X}_{\Delta,t-2}^{\mu}\bar{\rho}_2. \end{aligned}$$

Proof of Lemma A.2. The proof of Lemma A.2 is omitted because it can be proven using Lemma A.1 and the steps to prove Lemma 1 in the appendix of Yang et al. (2020). ■

Lemma A.3. *Under the same set of assumptions as in Theorem 4.1, as $T \to \infty$, the following approximations and limits hold.*

(i). Suppose that $\Phi_T = I_k + \frac{C_x}{T^{\kappa_x}}, C_x \leq 0, \kappa_x \in (0,1]$ and $\rho_{1,T} = 1 + \frac{c_u}{T^{\kappa_u}}$, $c_u < 0, \kappa_u \in (0,1]$.

If $\min\{\eta,\kappa_x\}<2\kappa_u$,

$$\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1,\rho}\tilde{X}^{\mu\prime}_{t-1,\rho}=\frac{1}{T}\sum_{t=p+1}^{T}w_{\rho,t}w'_{\rho,t}+o_p(1)\xrightarrow{p}E[w_{\rho,t}w'_{\rho,t}],$$

$$\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1,\rho}\tilde{Z}'_{t-1,\rho}=\frac{1}{T}\sum_{t=p+1}^{T}w_{\rho,t}w'_{\rho,t}+o_p(1)\xrightarrow{p}E[w_{\rho,t}w'_{\rho,t}],$$

and

$$\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\tilde{Z}_{t-1,\rho}z^{\mu}_{0,t}=\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}w_{\rho,t}z_{0,t}+o_p(1)\Rightarrow N(0,\sigma_z^2E[w_{\rho,t}w'_{\rho,t}]),$$

where $w_{\rho,t}=\varepsilon_{1,t}-\sum_{k=2}^{P}\rho_k\varepsilon_{1,t-k+1}$.

If $\min\{\eta,\kappa_x\}=2\kappa_u$,

$$\begin{aligned}&\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1,\rho}\tilde{X}^{\mu\prime}_{t-1,\rho}\\ =\;&\frac{1}{T}\sum_{t=p+1}^{T}w_{\rho,t}w'_{\rho,t}+\frac{c_u^2}{T^{2\kappa_u}}\frac{1}{T}\sum_{t=p+1}^{T}Z_{t-2}X'_{t-2}+o_p(1)\\ \Rightarrow\;&\begin{cases}E[w_{\rho,t}w'_{\rho,t}]-c_u^2{c_z}^{-1}\left(\int_0^1 J_{C_x}dB'_x+\Omega_{xx}\right)', & \text{if } \kappa_x>\eta\\ E[w_{\rho,t}w'_{\rho,t}]+c_u^2V_{xx}, & \text{if } \kappa_x<\eta\\ E[w_{\rho,t}w'_{\rho,t}]-c_u^2V_{zx}C_x, & \text{if } \kappa_x=\eta\end{cases},\end{aligned}$$

$$\begin{aligned}&\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1,\rho}\tilde{Z}'_{t-1,\rho}\\ =\;&\frac{1}{T}\sum_{t=p+1}^{T}w_{\rho,t}w'_{\rho,t}+c_u^2\frac{1}{T^{1+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}Z_{t-2}Z'_{t-2}+o_p(1)\\ \Rightarrow\;&\begin{cases}E[w_{\rho,t}w'_{\rho,t}]+c_u^2V^x_{zz}, & \text{if } \kappa_x>\eta\\ E[w_{\rho,t}w'_{\rho,t}]+c_u^2V_{xx}, & \text{if } \kappa_x<\eta\\ E[w_{\rho,t}w'_{\rho,t}]+c_u^2\int_0^\infty e^{sC_z}\left(C_xV_{xz}C_z+C_zV'_{xz}C_x\right)e^{sC_z}ds, & \text{if } \kappa_x=\eta\end{cases},\end{aligned}$$

and

$$\begin{aligned}
& \frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} \tilde{Z}_{t-1,\rho} z_{0,t}^{\mu} \\
= & \frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} w_{\rho,t} z_{0,t} - \frac{I_k c_u}{T^{\kappa_u}} \frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} Z_{t-2} z_{0,t} + o_p(1) \\
\Rightarrow & \begin{cases} B_{w_\rho}(1) - c_u U_z(1), \text{ if } \kappa_x > \eta, \\ B_{w_\rho}(1) - c_u U_x(1), \text{ if } \kappa_x < \eta \\ N\left(0, \left(E[w_{\rho,t} w_{\rho,t}'] + c_u^2 \int_0^\infty e^{sC_z}(C_x V_{xz} C_z + C_z V_{xz}' C_x) e^{sC_z} ds\right) \sigma_z^2\right), \text{ if } \kappa_x = \eta, \end{cases}
\end{aligned}$$

where $B_{w_\rho}(s)$, $U_z(s)$ and $U_x(s)$ are Brownian motions with variance $E\left[w_{\rho,t-1} w_{\rho,t-1}'\right] \sigma_z^2$, $V_{zz}^x \sigma_z^2$ and $V_{xx}\sigma_z^2$, respectively. $U_z(s)$ and $U_z(s)$ are independent of $B_{w_\rho}(s)$.

If $\min\{\eta, \kappa_x\} > 2\kappa_u$,

$$\begin{aligned}
\frac{T^{2\kappa_u}}{T^{1+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \tilde{Z}_{t-1,\rho} \tilde{X}_{t-1,\rho}^{\mu\prime} &= c_u^2 \frac{1}{T^{1+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} Z_{t-2} X_{t-2}' + o_p(1) \\
&\Rightarrow \begin{cases} -c_u^2 C_z^{-1} \left(\int_0^1 J_{C_x} dB_x' + \Omega_{xx}\right)', \text{ if } \kappa_x > \eta \\ c_u^2 V_{xx}, \text{ if } \kappa_x < \eta \\ -c_u^2 V_{zx} C_x, \text{ if } \kappa_x = \eta \end{cases},
\end{aligned}$$

$$\begin{aligned}
\frac{T^{2\kappa_u}}{T^{1+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \tilde{Z}_{t-1,\rho} \tilde{Z}_{t-1,\rho}' &= c_u^2 \frac{1}{T^{1+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} Z_{t-2} Z_{t-2}' + o_p(1) \\
&\Rightarrow \begin{cases} c_u^2 V_{zz}^x, \text{ if } \kappa_x > \eta \\ c_u^2 V_{xx}, \text{ if } \kappa_x < \eta \\ c_u^2 \int_0^\infty e^{sC_z} \left(C_x V_{xz} C_z + C_z V_{xz}' C_x\right) e^{sC_z} ds, \text{ if } \kappa_x = \eta \end{cases},
\end{aligned}$$

and

$$\begin{aligned}
& \frac{1}{T^{\frac{1+\min\{\eta,\kappa_x\}}{2} - \kappa_u}} \sum_{t=p+1}^{T} \tilde{Z}_{t-1} z_{0,t}^{\mu} \\
= & -c_u \frac{1}{T^{\frac{1+\min\{\eta,\kappa_x\}}{2}}} \sum_{t=p+1}^{T} Z_{t-2} z_{0,t} + o_p(1)
\end{aligned}$$

$$\Rightarrow \begin{cases} -c_u U_z(1), & \textit{if } \kappa_x > \eta, \\ -c_u U_z(1), & \textit{if } \kappa_x < \eta \\ N\left(0, c_u^2 \left(\int_0^\infty e^{sC_z}(C_x V_{xz} C_z + C_z V_{xz}' C_x) e^{sC_z} ds\right) \sigma_z^2\right), & \textit{if } \kappa_x = \eta. \end{cases}$$

(ii). Suppose that $\Phi_T = I_k + \frac{C_x}{T^{\kappa_x}}, C_x > 0, \kappa_x \in (0.5, 1)$ *and* $\rho_{1,T} = 1 + \frac{c_u}{T^{\kappa_u}}$, $c_u < 0, \kappa_u \in (0,1)$. *If* $\kappa_x = \kappa_u$, *we have*

$$\begin{aligned} & \frac{T^{2\kappa_x}}{T^{\kappa_x + \min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} \tilde{Z}_t \tilde{X}_t' \Phi_T^{-T} \\ = & \frac{1}{T^{\kappa_x + \min\{\eta,\kappa_x\}}} \left[\begin{array}{c} C_x \sum_{t=1}^T \Phi_T^{-T} X_{\rho,t-1} Z_{X_\rho,t-1}' \Phi_T^{-T} C_x - c_u C_x \sum_{t=1}^T \Phi_T^{-T} X_{\rho,t-1} Z_{t-1}' \Phi_T^{-T} \\ -c_u \sum_{t=1}^T \Phi_T^{-T} X_{t-1} Z_{X_\rho,t-1}' \Phi_T^{-T} C_x + c_u^2 \sum_{t=1}^T \Phi_T^{-T} X_{t-1} Z_{t-1}' \Phi_T^{-T} \end{array} \right] \\ & + o_p(1) \\ \Rightarrow & C_x C_{z,\kappa_x,\eta} W_{C_x}, \end{aligned}$$

where

$$C_{z,\kappa_x,\eta} = \begin{cases} -C_z^{-1}, & \textit{if } \kappa_x > \eta, \\ C_x^{-1}, & \textit{if } \kappa_x < \eta, \\ (C_x - C_z)^{-1}, & \textit{if } \kappa_x = \eta \end{cases}.$$

$W_{C_x} = W_{C_x}^{(1)} - W_{C_x}^{(2)} - W_{C_x}^{(3)} + W_{C_x}^{(4)}$,

$$\begin{aligned} W_{C_x}^{(1)} & \equiv C_x \int_0^\infty e^{-pC_x} Y_{C_x}^\rho Y_{C_x}^{\rho\prime} e^{-pC_x} dp C_x, W_{C_x}^{(2)} \equiv c_u \int_0^\infty e^{-pC_x} Y_{C_x} Y_{C_x}^{\rho\prime} e^{-pC_x} dp C_x, \\ W_{C_x}^{(3)} & \equiv c_u C_x \int_0^\infty e^{-pC_x} Y_{C_x}^\rho Y_{C_x}' e^{-pC_x} dp, W_{C_x}^{(4)} \equiv c_u^2 \int_0^\infty e^{-pC_x} Y_{C_x} Y_{C_x}' e^{-pC_x} dp, \end{aligned}$$

$Y_{C_x}^\rho = N\left(0, \int_0^\infty e^{-pC_x} \Omega_{w,\rho} e^{-pC_x} dp\right)$ *and* $\Omega_{w,\rho} = \sum_{j=-\infty}^\infty E[w_{\rho,t} w_{\rho,t-j}]$. *Moreover,*

$$\begin{aligned} & \frac{T^{2\kappa_x}}{T^{2\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} \tilde{Z}_t \tilde{Z}_t' \Phi_T^{-T} \\ = & \frac{\Phi_T^{-T}}{T^{\kappa_x + \min\{\eta,\kappa_x\}}} \left[\begin{array}{c} C_x \sum_{t=1}^T X_{\rho,t-1} Z_{X_\rho,t-1}' C_x - c_u C_x \sum_{t=1}^T X_{\rho,t-1} Z_{t-1}' \\ -c_u \sum_{t=1}^T X_{t-1} Z_{X_\rho,t-1}' C_x + c_u^2 \sum_{t=1}^T X_{t-1} Z_{t-1}' \end{array} \right] \Phi_T^{-T} \\ & + o_p(1) \\ \Rightarrow & C_x C_{z,\kappa_x,\eta} W_{C_x} C_{z,\kappa_x,\eta} C_x. \end{aligned}$$

$$\begin{aligned}
\frac{T^{\kappa_x}}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_{t-1}z_{0,t}^{\mu} &= C_xC_{z,\kappa_x,\eta}\frac{1}{T^{\kappa_x/2}}\sum_{t=p+1}^{T}\Phi_T^{t-T}z_{0,t}\left[\left(C_xY_{C_x}^{\rho}-c_uY_{C_x}\right)\right]+o_p(1) \\
&\Rightarrow C_xC_{z,\kappa_x,\eta}\times MN\left(0,W_{C_x}\times\sigma_z^2\right),
\end{aligned}$$

where MN stands for mixed normal distribution. If $\kappa_x > \kappa_u$, we have

$$\begin{aligned}
\frac{T^{2\kappa_u}}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_t\tilde{X}_t'\Phi_T^{-T} &= \frac{c_u^2}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}X_{t-1}Z_{t-1}'\Phi_T^{-T}+o_p(1) \\
&\Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(4)}, \\
\frac{T^{2\kappa_u}}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_t\tilde{Z}_t'\Phi_T^{-T} &= \frac{c_u^2}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{t-1}Z_{t-1}'\Phi_T^{-T}+o_p(1) \\
&\Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(4)}C_{z,\kappa_x,\eta}C_x, \\
\frac{T^{\kappa_u}}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_{t-1}z_{0,t}^{\mu} &= \frac{c_u}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{t-2}z_{0,t}+o_p(1) \\
&\Rightarrow C_xC_{z,\kappa_x,\eta}\times MN\left(0,W_{C_x}^{(4)}\times\sigma_z^2\right).
\end{aligned}$$

If $\kappa_u > \kappa_x$, we have

$$\begin{aligned}
\frac{T^{2\kappa_x}}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_t\tilde{X}_t'\Phi_T^{-T} &= \frac{C_x}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}X_{\rho,t-1}Z_{X_\rho,t-1}'\Phi_T^{-T}C_x+o_p(1), \\
&\Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(1)}, \\
\frac{T^{\kappa_x}}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_t\tilde{Z}_t'\Phi_T^{-T} &= \frac{C_x}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{X_\rho,t-1}Z_{X_\rho,t-1}'\Phi_T^{-T}C_x+o_p(1), \\
&\Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(1)}C_{z,\kappa_x,\eta}C_x, \\
\frac{T^{\kappa_x}}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_{t-1}z_{0,t}^{\mu} &= \frac{C_x}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{X_\rho,t-2}z_{0,t}+o_p(1) \\
&\Rightarrow C_xC_{z,\kappa_x,\eta}\times MN\left(0,W_{C_x}^{(1)}\times\sigma_z^2\right).
\end{aligned}$$

Proof of Lemma A.3. (i). Note that

$$\begin{aligned}
&X_{t-1}-\rho_{1,T}X_{t-2}-\mathbf{X}_{\Delta,t-2}\bar{\rho}_2 \\
=&\left(I_k+\frac{C_x}{T^{\kappa_x}}\right)X_{t-2}+\varepsilon_{1,t-1}-\rho_{1,T}X_{t-2}-\mathbf{X}_{\Delta,t-2}\bar{\rho}_2
\end{aligned}$$

$$
\begin{aligned}
&= w_{\rho,t-1} + \left(\frac{C_x}{T^{\kappa_x}} - \frac{I_k c_u}{T^{\kappa_u}}\right) X_{t-2} - \frac{C_x}{T^{\kappa_x}} \sum_{k=3}^{P} \rho_{k-1} X_{t-k} \\
&= w_{\rho,t-1} + \frac{C_x}{T^{\kappa_x}} \left(X_{t-2} - \sum_{k=3}^{p} \rho_{k-1} X_{t-k}\right) - \frac{I_k c_u}{T^{\kappa_u}} X_{t-2} \\
&= w_{\rho,t-1} + \frac{C_x}{T^{\kappa_x}} X_{\rho,t-2} - \frac{I_k c_u}{T^{\kappa_u}} X_{t-2},
\end{aligned}
$$

where $X_{\rho,t} = X_t - \sum_{k=2}^{p} \rho_k X_{t-p+1} = (1 - \rho_2 L - ... - \rho_p L^{p-1}) X_t$. And it can be seen that $X_{\rho,t}$ shares the same stochastic order as X_t. Let $Z_{X_\rho,t} = \sum_{j=1}^{t} \Upsilon_T^{t-j} \Delta X_{\rho,j}$ and $Z_{\rho,t} = \sum_{j=1}^{t} \Upsilon_T^{t-j} \Delta w_{\rho,j}$, we can express

$$
\tilde{Z}_{t-1} = Z_{\rho,t-1} + \frac{C_x}{T^{\kappa_x}} Z_{X_\rho,t-1} - \frac{I_k c_u}{T^{\kappa_u}} Z_{t-2}. \tag{14}
$$

Then

$$
\begin{aligned}
& \frac{1}{T} \sum_{t=p+1}^{T} \left(Z_{t-1} - \rho_{1,T} Z_{t-2} - \mathbf{Z}_{\Delta,t-2} \bar{\rho}_2\right) \left(X_{t-1}^{\mu} - \rho_{1,T} X_{t-2}^{\mu} - \mathbf{X}_{\Delta,t-2}^{\mu} \bar{\rho}_2\right)' \\
= & \frac{1}{T} \sum_{t=p+1}^{T} \left(Z_{\rho,t-1} + \frac{C_x}{T^{\kappa_x}} Z_{X_\rho,t-1} - \frac{I_k c_u}{T^{\kappa_u}} Z_{t-2}\right) \times \\
& (w_{\rho,t-1} + \frac{C_x}{T^{\kappa_x}} X_{\rho,t-2} - \frac{I_k c_u}{T^{\kappa_u}} X_{t-2})'.
\end{aligned} \tag{15}
$$

Suppose that $\min\{\eta, \kappa_x\} < 2\kappa_u$. It is easy to verify that all terms in (15) except $\frac{1}{T} \sum_{t=2}^{T} Z_{\rho,t-1} w'_{\rho,t-1}$ are $o_p(1)$. For example, using Proposition A2 in Phillips and Magdalinos (2009), $X_t = O_p(T^{-\kappa_x/2})$ and Lemma 3.1, 3.5, 3.6 in Phillips and Magdalinos (2009), we have

$$
\begin{aligned}
\frac{c_u^2}{T^{2\kappa_u}} \frac{1}{T} \sum_{t=p+1}^{T} Z_{t-2} X'_{t-2} &= \frac{c_u^2}{T^{2\kappa_u}} \frac{1}{T} \left(O_p(T^{1+\min\{\eta,\kappa_x\}})\right) \\
&= O_p(T^{\min\{\eta,\kappa_x\} - 2\kappa_u}) = o_p(1).
\end{aligned} \tag{16}
$$

Eventually, we have

$$
\begin{aligned}
& \frac{1}{T} \sum_{t=p+1}^{T} \left(Z_{t-1} - \rho_{1,T} Z_{t-2} - \mathbf{Z}_{\Delta,t-2} \bar{\rho}_2\right) \left(X_{t-1}^{\mu} - \rho_{1,T} X_{t-2}^{\mu} - \mathbf{X}_{\Delta,t-2}^{\mu} \bar{\rho}_2\right)' \\
= & \frac{1}{T} \sum_{t=p+1}^{T} Z_{\rho,t-1} w'_{\rho,t-1} + o_p(1)
\end{aligned}
$$

$$= \frac{1}{T}\sum_{t=p+1}^{T} w_{\rho,t-1}w'_{\rho,t-1} + o_p(1),$$

where the last equality is obtained by applying Lemma B2 in Kostakis et al. (2015). By the ergodic theorem, we have $\frac{1}{T}\sum_{t=2}^{T} w_{\rho,t-1}w'_{\rho,t-1} \overset{as}{\to} E[w_{\rho,t-1}w'_{\rho,t-1}]$.

For the term $\frac{1}{T}\sum_{t=p+1}^{T} (Z_{t-1} - \rho_{1,T}Z_{t-2} - \mathbf{Z}_{\Delta,t-2}\bar{\rho}_2)(Z_{t-1} - \rho_{1,T}Z_{t-2} - \mathbf{Z}_{\Delta,t-2}\bar{\rho}_2)'$, note that from (14), it can be rewritten as

$$\frac{1}{T}\sum_{t=p+1}^{T}\left(Z_{\rho,t-1} + \frac{C_x}{T^{\kappa_x}}Z_{X_\rho,t-1} - \frac{I_k c_u}{T^{\kappa_u}}Z_{t-2}\right)\left(Z_{\rho,t-1} + \frac{C_x}{T^{\kappa_x}}Z_{X_\rho,t-1} - \frac{I_k c_u}{T^{\kappa_u}}Z_{t-2}\right)'.$$

Similar to (16), only the first term is not $o_p(1)$. Thus,

$$\begin{aligned}
& \frac{1}{T}\sum_{t=p+1}^{T}\left(Z_{\rho,t-1} + \frac{C_x}{T^{\kappa_x}}Z_{X_\rho,t-1} - \frac{I_k c_u}{T^{\kappa_u}}Z_{t-2}\right)\left(Z_{\rho,t-1} + \frac{C_x}{T^{\kappa_x}}Z_{X_\rho,t-1} - \frac{I_k c_u}{T^{\kappa_u}}Z_{t-2}\right)' \\
= & \frac{1}{T}\sum_{t=p+1}^{T} Z_{\rho,t-1}Z'_{\rho,t-1} + o_p(1) \\
= & \frac{1}{T}\sum_{t=p+1}^{T} w_{\rho,t}w'_{\rho,t} + o_p(1) \overset{p}{\to} E[w_{\rho,t}w'_{\rho,t}],
\end{aligned}$$

where the last equality is obtained from applying Lemma B.2 in Kostakis et al. (2015).

For the term $\frac{1}{\sqrt{T}}\sum_{t=2}^{T} \tilde{Z}_{t-1}z^{\mu}_{0,t}$, we have

$$\begin{aligned}
& \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} \tilde{Z}_{t-1}z^{\mu}_{0,t} \\
= & \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\left(\left(Z_{\rho,t-1} + \frac{C_x}{T^{\kappa_x}}Z_{t-1} - \frac{C_x}{T^{\kappa_x}}\sum_{k=3}^{p}\rho_{k-1}Z_{t-k} - \frac{I_k c_u}{T^{\kappa_u}}Z_{t-2}\right)z^{\mu}_{0,t}\right) \\
= & \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} Z_{\rho,t-1}z^{\mu}_{0,t} + \frac{C_x}{T^{\kappa_x}}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} Z_{t-1}z^{\mu}_{0,t} \\
& -\frac{C_x}{T^{\kappa_x}}\sum_{k=3}^{P}\rho_{k-1}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} Z_{t-k}z^{\mu}_{0,t} - \frac{c_u}{T^{\kappa_u}}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} Z_{t-2}z^{\mu}_{0,t}.
\end{aligned}$$

Applying Lemma 3.1, 3.5 and 3.6 in Phillips and Magdalinos (2009), we can obtain the following orders

$$\begin{aligned}
\sum_{t=p+1}^{T} Z_{t-1} z_{0,t}^{\mu} &= O_p\left(T^{\frac{1+\min\{\eta,\kappa_x\}}{2}}\right), \\
\frac{C_x}{T^{\kappa_x}} \frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} Z_{t-1} z_{0,t}^{\mu} &= o_p(1), \\
\frac{C_x}{T^{\kappa_x}} \sum_{k=3}^{P} \rho_{k-1} \frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} Z_{t-k} z_{0,t}^{\mu} &= o_p(1), \\
\frac{c_u}{T^{\kappa_u}} \frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} Z_{t-2} z_{0,t}^{\mu} &= o_p(1).
\end{aligned}$$

Applying Lemma B2. in the Online appendix of Kostakis et al. (2015) and the CLT for the martingale difference sequence, we have

$$\frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} \tilde{Z}_{t-1} z_{0,t}^{\mu} = \frac{1}{\sqrt{T}} \sum_{t=p+1}^{T} w_{\rho,t-1} z_{0,t} + o_p(1) \Rightarrow N(0, \sigma_z^2 E[w_{\rho,t} w_{\rho,t}']^2).$$

Suppose that $\min\{\eta, \kappa_x\} = 2\kappa_u$, from (15) and note that $\frac{c_u^2}{T^{2\kappa_u}} \frac{1}{T} \sum_{t=2}^{T} Z_{t-2} X_{t-2}' = O_p(1)$ when $\min\{\eta, \kappa_x\} = 2\kappa_u$, we have

$$\begin{aligned}
&\frac{1}{T} \sum_{t=p+1}^{T} \left(Z_{t-1} - \rho_{1,T} Z_{t-2} - \mathbf{Z}_{\Delta,t-2} \bar{\rho}_2\right) \left(X_{t-1}^{\mu} - \rho_{1,T} X_{t-2}^{\mu} - \mathbf{X}_{\Delta,t-2}^{\mu} \bar{\rho}_2\right)' \\
= &\frac{1}{T} \sum_{t=p+1}^{T} w_{\rho,t} w_{\rho,t}' + c_u^2 \frac{T^{1+\min\{\eta,\kappa_x\}}}{T^{1+2\kappa_u}} \frac{1}{T^{1+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} Z_{t-2} X_{t-2}^{\mu\prime} + o_p(1).
\end{aligned}$$

Since $\frac{1}{T} \sum_{t=p+1}^{T} w_{\rho,t} w_{\rho,t}' \xrightarrow{p} E[w_{\rho,t} w_{\rho,t}']$, Equation (20), Lemma 3.5 and 3.6 in Phillips and Magdalinos (2009) gives

$$\frac{1}{T^{1+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} Z_{t-2} X_{t-2}^{\mu\prime} \Rightarrow -C_z^{-1} \left(\int_0^1 J_{C_x} dB_1' + \Omega_{11}\right)', \text{ if } \kappa_x > \eta.$$

$$\frac{1}{T^{1+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} Z_{t-2} X_{t-2}^{\mu\prime} \xrightarrow{p} V_{xx}, \text{ if } \kappa_x < \eta$$

$$\frac{1}{T^{1+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T} Z_{t-2}X_{t-2}^{\mu\prime} \xrightarrow{p} -V_{zx}C_x, \text{ if } \kappa_x = \eta.$$

Therefore, we have

$$\begin{aligned}
&\frac{1}{T}\sum_{t=p+1}^{T}\left(Z_{t-1}-\rho_{1,T}Z_{t-2}-\mathbf{Z}_{\Delta,t-2}\bar{\rho}_2\right)\left(X_{t-1}^{\mu}-\rho_{1,T}X_{t-2}^{\mu}-\mathbf{X}_{\Delta,t-2}^{\mu}\bar{\rho}_2\right)' \\
&\Rightarrow \begin{cases} E[w_{\rho,t}w_{\rho,t}'] - c_u^2 {C_z}^{-1}\left(\int_0^1 J_{C_x}dB_x' + \Omega_{xz}\right)', \text{ if } \kappa_x > \eta, \\ E[w_{\rho,t}w_{\rho,t}'] + c_u^2 V_{xx}, \text{ if } \kappa_x < \eta, \\ E[w_{\rho,t}w_{\rho,t}'] - c_u^2 V_{zx}C_x, \text{ if } \kappa_x = \eta. \end{cases}
\end{aligned}$$

Similarly, for $\sum_{t=p+1}^{T}\tilde{Z}_{t-1,\rho}\tilde{Z}_{t-1,\rho}'$, applying Lema 3.1, 3.5 and 3.6 in Phillips and Magdalinos (2009) gives

$$\begin{aligned}
&\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1,\rho}\tilde{Z}_{t-1,\rho}' \\
&= \frac{1}{T}\sum_{t=p+1}^{T} w_{\rho,t}w_{\rho,t}' + c_u^2\frac{T^{1+\min\{\eta,\kappa_x\}}}{T^{1+2\kappa_u}}\frac{1}{T^{1+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T} Z_{t-2}Z_{t-2}' + o_p(1) \\
&\Rightarrow \begin{cases} E[w_{\rho,t}w_{\rho,t}'] + c_u^2 V_{zz}^{x}, \text{ if } \kappa_x > \eta \\ E[w_{\rho,t}w_{\rho,t}'] + c_u^2 V_{xx}, \text{ if } \kappa_x < \eta \\ E[w_{\rho,t}w_{\rho,t}'] + c_u^2 \int_0^\infty e^{sC_z}\left(C_xV_{xz}C_z + C_zV_{xz}'C_x\right)e^{sC_z}ds, \text{ if } \kappa_x = \eta \end{cases}.
\end{aligned}$$

Similarly, for $\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}z_{0,t}^{\mu}$, since $\frac{c_u}{T^{\kappa_u}}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\left(\sum_{j=1}^{t-1}\Upsilon_T^{t-1-j}\Delta X_{j-1}\right)z_{0,t}^{\mu} = O_p(1)$ when $\min\{\eta,\kappa_x\} = 2\kappa_u$, we have

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}z_{0,t}^{\mu} \\
&= \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} Z_{\rho,t-1}z_{0,t}^{\mu} - \frac{I_kc_u}{T^{\kappa_u}}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\left(\sum_{j=1}^{t-1}\Upsilon_T^{t-1-j}\Delta X_{j-1}\right)z_{0,t}^{\mu} + o_p(1) \\
&= \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} w_{\rho,t-1}z_{0,t} - \frac{I_kc_u}{T^{\kappa_u}}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} Z_{t-1}z_{0,t} + o_p(1),
\end{aligned}$$

If $\kappa_x > \eta$, from Equation (15),(16) and Lemma 3.1 in Phillips and Magdalinos (2009), we can express

$$\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\left(w_{\rho,t-1}-\frac{I_k c_u}{T^{\min\{\eta,\kappa_x\}/2}}\tilde{Z}_{t-1}\right)z_{0,t}$$

as

$$\begin{aligned}&\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\left(w_{\rho,t-1}-\frac{I_k c_u}{T^{\eta/2}}\tilde{Z}_{t-1}\right)z_{0,t}\\ =\ &\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}w_{\rho,t-1}z_{0,t}-\frac{I_k c_u}{T^{(\eta+1)/2}}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}z_{0,t},\end{aligned}$$

where $\tilde{Z}_{t-1}$ is referring to as a mildly stationary time series.

Following the procedure to prove Proposition A1. in Phillips and Magdalinos (2009), we can show the joint convergence of

$$\left(\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}w_{\rho,t-1}z_{0,t},\frac{c_u}{T^{(1+\eta)/2}}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}z_{0,t}\right).$$

We have

$$\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\left(w_{\rho,t-1}-\frac{I_k c_u}{T^{\min\{\eta,\kappa_x\}/2}}\tilde{Z}_{t-1}\right)z_{0,t}\Rightarrow B_{w_\rho}(1)-c_u U_z(1), \tag{17}$$

If $\kappa_x<\eta$, Lemma 3.5 in Phillips and Magdalinos (2009) gives

$$\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}w_{\rho,t-1}z_{0,t}-\frac{I_k c_u}{T^{\kappa_x/2}}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}Z_{t-1}z_{0,t}=\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}w_{\rho,t-1}z_{0,t}-\frac{I_k c_u}{T^{\kappa_x/2}}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}X_{t-1}z_{0,t}.$$

Using the analogous argument to obtain (17), we can show

$$\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}w_{\rho,t-1}z_{0,t}-\frac{I_k c_u}{T^{\kappa_x/2}}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}Z_{t-1}z_{0,t}\Rightarrow B_{w_\rho}(1)-c_u U_x(1).$$

If $\kappa_x=\eta$, applying Lemma 3.6 in Phillips and Magdalinos (2009) and standard CLT for mds gives

$$\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\left(w_{\rho,t-1}-\frac{I_k c_u}{T^{\min\{\eta,\kappa_x\}/2}}\tilde{Z}_{t-1}\right)z_{0,t}$$

$$\Rightarrow \quad N\left(0, \left(p\lim_{T\to\infty}\left[\frac{1}{T}\sum_{t=p+1}^{T}\left(w_{\rho,t-1}w'_{\rho,t-1}\right) + \frac{c_u^2}{T^{1+\kappa_x}}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}\tilde{Z}'_{t-1}\right]\right)\sigma_z^2\right)$$
$$= \quad N\left(0, \left(E[w_{\rho,t}w'_{\rho,t}] + c_u^2\int_0^\infty e^{sC_z}(C_xV_{xz}C_z + C_zV'_{xz}C_x)e^{sC_z}ds\right)\sigma_z^2\right).$$

Suppose $\min\{\eta, \kappa_x\} > 2\kappa_u$. Then we have

$$\frac{T^{2\kappa_u}}{T^{1+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\left(Z_{t-1} - \rho_{1,T}Z_{t-2} - \mathbf{Z}_{\Delta,t-2}\bar{\rho}_2\right)\left(X_{t-1}^{\mu} - \rho_{1,T}X_{t-2}^{\mu} - \mathbf{X}_{\Delta,t-2}^{\mu}\bar{\rho}_2\right)'$$
$$= \quad c_u^2\frac{1}{T^{1+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}Z_{t-2}X_{t-2}^{\mu\prime} + o_p(1)$$
$$\Rightarrow \begin{cases} -c_u^2C_z{}^{-1}\left(\int_0^1 J_{C_x}dB_x' + \Omega_{xx}\right)', \text{ if } \kappa_x > \eta \\ c_u^2V_{xx}, \text{ if } \kappa_x < \eta \\ -c_u^2V_{zx}C_x, \text{ if } \kappa_x = \eta \end{cases}.$$

Similarly,

$$\frac{T^{2\kappa_u}}{T^{1+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\tilde{Z}_{t-1,\rho}\tilde{Z}'_{t-1,\rho} \quad = \quad c_u^2\frac{1}{T^{1+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}Z_{t-2}Z'_{t-2} + o_p(1)$$
$$\Rightarrow \begin{cases} c_u^2V_{zz}^x, \text{ if } \kappa_x > \eta \\ c_u^2V_{xx}, \text{ if } \kappa_x < \eta \\ c_u^2\int_0^\infty e^{sC_z}\left(C_xV_{xz}C_z + C_zV'_{xz}C_x\right)e^{sC_z}ds, \text{ if } \kappa_x = \eta \end{cases}.$$

And

$$\frac{1}{T^{\frac{1+\min\{\eta,\kappa_x\}}{2}-\kappa_u}}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}z_{0,t}^{\mu}$$
$$= \quad \frac{T^{1/2+\kappa_u}}{T^{\frac{1+\min\{\eta,\kappa_x\}}{2}}}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}w_{\rho,t-1}z_{0,t} - c_u\frac{1}{T^{\frac{1+\min\{\eta,\kappa_x\}}{2}}}\sum_{t=p+1}^{T}Z_{t-2}z_{0,t} + o_p(1)$$
$$= \quad -c_u\frac{1}{T^{\frac{1+\min\{\eta,\kappa_x\}}{2}}}\sum_{t=p+1}^{T}Z_{t-2}z_{0,t} + o_p(1)$$
$$\Rightarrow \begin{cases} -c_uU_z(1), \text{ if } \kappa_x > \eta, \\ -c_uU_z(1), \text{ if } \kappa_x < \eta \\ N\left(0, c_u^2\left(\int_0^\infty e^{sC_z}(C_xV_{xz}C_z + C_zV'_{xz}C_x)e^{sC_z}ds\right)\sigma_z^2\right), \text{ if } \kappa_x = \eta. \end{cases}$$

(ii). In the case where $C_x > 0$,

$$\begin{aligned} & X_t - \rho_{1,T} X_{t-1} - \mathbf{X}_{\Delta,t-2}\bar{\rho}_2 \\ = & \left(I_k + \frac{C_x}{T^{\kappa_x}}\right) X_{t-1} + \varepsilon_{1,t} - \left(1 + \frac{c_u}{T^{\kappa_u}}\right) X_{t-1} - \sum_{k=2}^{P} \rho_k \left(\frac{C_x}{T^{\kappa_x}} X_{t-k} + \varepsilon_{1,t-(k-1)}\right) \\ = & \frac{C_x}{T^{\kappa_x}} X_{\rho,t-1} - \frac{c_u}{T^{\kappa_u}} X_{t-1} + w_{\rho,t} \qquad (18) \end{aligned}$$

where $X_{\rho,t-1} = X_{t-1} - \sum_{k=2}^{P} \rho_k X_{t-k}$.

From (18), we can express

$$\tilde{Z}_t = \frac{C_x}{T^{\kappa_x}} Z_{X_\rho,t-1} - \frac{c_u}{T^{\kappa_u}} Z_{t-1} + Z_{\rho,t},$$

where $Z_{X_\rho,t} = \sum_{j=1}^{t} \Upsilon_T^{t-j} \Delta X_{\rho,j}$ and $Z_{\rho,t} = \sum_{j=1}^{t} \Upsilon_T^{t-j} \Delta w_{\rho,j}$. Then

$$\begin{aligned} \sum_{t=p+1}^{T} \tilde{Z}_t \tilde{X}_t' &= \sum_{t=p+1}^{T} \left(\frac{C_x}{T^{\kappa_x}} Z_{X_\rho,t-1} - \frac{c_u}{T^{\kappa_u}} Z_{t-1} + Z_{\rho,t}\right) \left(\frac{C_x}{T^{\kappa_x}} X_{\rho,t-1} - \frac{c_u}{T^{\kappa_u}} X_{t-1} + w_{\rho,t}\right)' \\ &= \frac{C_x}{T^{2\kappa_x}} \sum_{t=p+1}^{T} Z_{X_\rho,t-1} X_{\rho,t-1}' C_x - \frac{c_u C_x}{T^{\kappa_x+\kappa_u}} \sum_{t=p+1}^{T} Z_{X_\rho,t-1} X_{t-1}' \\ &\quad - \frac{c_u}{T^{\kappa_x+\kappa_u}} \sum_{t=p+1}^{T} Z_{t-1} X_{\rho,t-1}' C_x + \frac{c_u^2}{T^{2\kappa_u}} \sum_{t=p+1}^{T} Z_{t-1} X_{t-1}' \\ &\quad + \frac{C_x}{T^{\kappa_x}} \sum_{t=p+1}^{T} Z_{X_\rho,t-1} w_{\rho,t}' - \frac{c_u}{T^{\kappa_u}} \sum_{t=p+1}^{T} Z_{t-1} w_{\rho,t}' \\ &\quad + \frac{C_x}{T^{\kappa_x}} \sum_{t=p+1}^{T} Z_{\rho,t} X_{\rho,t-1}' - \frac{c_u}{T^{\kappa_u}} \sum_{t=p+1}^{T} Z_{\rho,t} X_{t-1}' + \sum_{t=p+1}^{T} Z_{\rho,t} w_{\rho,t}'. \qquad (19) \end{aligned}$$

Lemma 2.4 in Phillips and Lee (2016) and Lemma 3 in Yang et al. (2020) provide the following stochastic orders:

$$\begin{aligned} \frac{1}{T^{\kappa_x+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} X_{\rho,t-1} Z_{X_\rho,t-1}' \Phi_T^{-T} &= O_p(1), \\ \frac{1}{T^{\kappa_x+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} X_{\rho,t-1} Z_{X_\rho,t-1}' \Phi_T^{-T} &= O_p(1), \\ \frac{1}{T^{\kappa_x+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} X_{t-1} Z_{X_\rho,t-1}' \Phi_T^{-T} &= O_p(1), \end{aligned}$$

$$\frac{1}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}X_{t-1}Z_{t-1}'\Phi_T^{-T} \;=\; O_p(1).$$

Thus, the first 4 terms in (19) asymptotically dominate the other terms.

Similarly,

$$\begin{aligned}
\sum_{t=p+1}^{T}\tilde{Z}_t\tilde{X}_t' &= \sum_{t=p+1}^{T}\left(\frac{C_x}{T^{\kappa_x}}Z_{X_\rho,t-1}-\frac{c_u}{T^{\kappa_u}}Z_{t-1}+Z_{\rho,t}\right)\left(\frac{C_x}{T^{\kappa_x}}Z_{X_\rho,t-1}-\frac{c_u}{T^{\kappa_u}}Z_{t-1}+Z_{\rho,t}\right)' \\
&= \frac{C_x}{T^{2\kappa_x}}\sum_{t=p+1}^{T}Z_{X_\rho,t-1}Z_{X_\rho,t-1}'C_x-\frac{c_uC_x}{T^{\kappa_x+\kappa_u}}\sum_{t=p+1}^{T}Z_{X_\rho,t-1}Z_{t-1}' \\
&\quad -\frac{c_u}{T^{\kappa_x+\kappa_u}}\sum_{t=p+1}^{T}Z_{t-1}Z_{X_\rho,t-1}'C_x+\frac{c_u^2}{T^{2\kappa_u}}\sum_{t=p+1}^{T}Z_{t-1}Z_{t-1}' \\
&\quad +\frac{C_x}{T^{\kappa_x}}\sum_{t=p+1}^{T}Z_{X_\rho,t-1}Z_{\rho,t}'-\frac{c_u}{T^{\kappa_u}}\sum_{t=p+1}^{T}Z_{t-1}Z_{\rho,t}' \\
&\quad +\frac{1}{T^{\kappa_x}}\sum_{t=p+1}^{T}Z_{\rho,t}Z_{X_\rho,t-1}'C_x-\frac{c_u}{T^{\kappa_u}}\sum_{t=p+1}^{T}Z_{\rho,t}Z_{t-1}'+\sum_{t=p+1}^{T}Z_{\rho,t}Z_{\rho,t}'.
\end{aligned}$$

Likewise, the first 4 terms dominate the last 5 terms.

Suppose that $\kappa_x=\kappa_u$.

$$\begin{aligned}
&\frac{T^{2\kappa_x}}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_t\tilde{X}_t'\Phi_T^{-T} \qquad (20)\\
&= \frac{\Phi_T^{-T}}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\left[\begin{array}{c} C_x\sum_{t=p+1}^{T}X_{\rho,t-1}Z_{X_\rho,t-1}'C_x-c_uC_x\sum_{t=p+1}^{T}X_{\rho,t-1}Z_{t-1}' \\ -c_u\sum_{t=p+1}^{T}X_{t-1}Z_{X_\rho,t-1}'C_x+c_u^2\sum_{t=p+1}^{T}X_{t-1}Z_{t-1}'\end{array}\right]\Phi_T^{-T}+o_p(1).
\end{aligned}$$

Let k_T and k_T' be time indexes satisfying Lemma 2.3 in Phillips and Lee (2016), $\Psi_{T,t}^{\rho}=\sum_{j=1}^{t}\Upsilon_T^{t-j}X_{\rho,j-1}$, and $\check{Z}_{X_\rho,t}=Z_{X_\rho,t}-\frac{C_x}{T^{\kappa_x}}\Psi_{T,t}^{\rho}$. Following the proof of Lemma 2.4 in Phillips and Lee (2016), we can show

$$\begin{aligned}
&\frac{C_x}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{X_\rho,t-1}X_{\rho,t-1}'\Phi_T^{-T}C_x \\
&= \frac{C_x}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\left(\check{Z}_{X_\rho,t-1}+\frac{C_x}{T^{\kappa_x}}\Psi_{T,t-1}^{\rho}\right)X_{\rho,t-1}'\Phi_T^{-T}C_x \\
&= \frac{C_x}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\check{Z}_{X_\rho,t-1}X_{\rho,t-1}'\Phi_T^{-T}C_x
\end{aligned}$$

$$
\begin{aligned}
& +\frac{C_x}{T^{2\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}C_x\Psi_{T,t-1}^{\rho}X_{\rho,t-1}'\Phi_T^{-T}C_x \\
= & \frac{C_xC_x}{T^{2\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\Psi_{T,t-1}^{\rho}X_{\rho,t-1}'\Phi_T^{-T}C_x+o_p(1) \\
= & \frac{C_xC_x}{T^{2\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=k_T+k_T'+2}^{T}\Phi_T^{-T}\Psi_{T,t-1}^{\rho}X_{\rho,t-1}'\Phi_T^{-T}C_x+o_p(1) \\
= & \frac{C_xC_x}{T^{\kappa_x}}\sum_{t=k_T+k_T'+2}^{T}\Phi_T^{-(T-t)}\left(\frac{\Phi_T^{-t}}{T^{\kappa_x/2+\min\{\eta,\kappa_x\}}}\Psi_{T,t-1}^{\rho}\right)\left(\frac{\Phi_T^{-t}}{T^{\kappa_x/2}}X_{\rho,t-1}\right)'\Phi_T^{-(T-t)}C_x+o_p(1) \\
= & \frac{C_xC_xC_{z,\kappa_x,\eta}}{T^{\kappa_x}}\sum_{t=1}^{T}\Phi_T^{-(T-t)}Y_{C_x}^{\rho}Y_{C_x}^{\rho\prime}\Phi_T^{-(T-t)}C_x+o_p(1) \\
\Rightarrow & C_xC_{z,\kappa_x,\eta}\left(W_{C_x}^{(1)}\right)\equiv C_xC_{z,\kappa_x,\eta}\left(C_x\int_0^{\infty}e^{-pC_x}Y_{C_x}^{\rho}Y_{C_x}^{\rho\prime}e^{-pC_x}dpC_x\right),
\end{aligned}
$$

where $Y_{C_x}^{\rho}=N\left(0,\int_0^{\infty}e^{-pC_x}\Omega_{w,\rho}e^{-pC_x}dp\right)$, $\Omega_{w,\rho}=\sum_{j=-\infty}^{\infty}E[w_{\rho,t}w_{\rho,t-j}]$ and

$$
C_{z,\kappa_x,\eta}=\begin{cases}-C_z^{-1}, \text{ if } \kappa_x>\eta,\\ C_x^{-1}, \text{ if } \kappa_x<\eta,\\ (C_x-C_z)^{-1}, \text{ if } \kappa_x=\eta\end{cases}.
$$

Likewise, we can also obtain the following limits

$$
\begin{aligned}
\frac{c_uC_x}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{t-1}X_{\rho,t-1}'\Phi_T^{-T}C_x & \Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(2)} \\
& \equiv C_xC_{z,\kappa_x,\eta}\left(c_u\int_0^{\infty}e^{-pC_x}Y_{C_x}Y_{C_x}^{\rho\prime}e^{-pC_x}dpC_x\right), \\
\frac{c_uC_x}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{X_\rho,t-1}X_{t-1}'\Phi_T^{-T} & \Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(3)} \\
& \equiv C_xC_{z,\kappa_x,\eta}c_uC_x\int_0^{\infty}e^{-pC_x}Y_{C_x}^{\rho}Y_{C_x}'e^{-pC_x}dp, \\
\frac{c_u^2C_x}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{t-1}X_{t-1}'\Phi_T^{-T} & \Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(4)} \\
& \equiv C_xC_{z,\kappa_x,\eta}c_u^2\int_0^{\infty}e^{-pC_x}Y_{C_x}Y_{C_x}'e^{-pC_x}dp.
\end{aligned}
$$

And eventually, we have

$$\frac{T^{2\kappa_x}}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_t\tilde{X}_t'\Phi_T^{-T}\Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x},$$

where $W_{C_x}=W_{C_x}^{(1)}-W_{C_x}^{(2)}-W_{C_x}^{(3)}+W_{C_x}^{(4)}$. Similar to (20), we can write

$$\begin{aligned}
&\frac{T^{2\kappa_x}}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=P+1}^{T}\Phi_T^{-T}\tilde{Z}_t\tilde{Z}_t'\Phi_T^{-T}\\
=\ &\frac{T^{2\kappa_x}\Phi_T^{-T}}{T^{2\min\{\eta,\kappa_x\}}}\begin{pmatrix}\frac{C_x}{T^{2\kappa_x}}\sum_{t=P+1}^{T}Z_{X_\rho,t-1}X_{\rho,t-1}'C_x-\frac{c_uC_x}{T^{\kappa_x+\kappa_u}}\sum_{t=P+1}^{T}Z_{X_\rho,t-1}X_{t-1}'\\ -\frac{c_u}{T^{\kappa_x+\kappa_u}}\sum_{t=P+1}^{T}Z_{t-1}X_{\rho,t-1}'C_x+\frac{c_u^2}{T^{2\kappa_u}}\sum_{t=P+1}^{T}Z_{t-1}X_{t-1}'\end{pmatrix}\Phi_T^{-T}+o_p(1)\\
=\ &\frac{\Phi_T^{-T}}{T^{2\min\{\eta,\kappa_x\}}}\begin{bmatrix}C_x\sum_{t=P+1}^{T}Z_{X_\rho,t-1}Z_{X_\rho,t-1}'C_x-c_uC_x\sum_{t=P+1}^{T}Z_{X_\rho,t-1}Z_{t-1}'\\ -c_u\sum_{t=P+1}^{T}Z_{t-1}Z_{X_\rho,t-1}'C_x+c_u^2\sum_{t=P+1}^{T}Z_{t-1}Z_{t-1}'\end{bmatrix}\Phi_T^{-T}+o_p(1),
\end{aligned}$$

and

$$\begin{aligned}
&\frac{C_x}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=P+1}^{T}\Phi_T^{-T}Z_{X_\rho,t-1}Z_{X_\rho,t-1}'\Phi_T^{-T}C_x\\
=\ &\frac{C_x}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=P+1}^{T}\Phi_T^{-T}\left(\check{Z}_{X_\rho,t-1}+\frac{C_x}{T^{\kappa_x}}\Psi_{T,t-1}^{\rho}\right)\left(\check{Z}_{X_\rho,t-1}+\frac{C_x}{T^{\kappa_x}}\Psi_{T,t-1}^{\rho}\right)'\Phi_T^{-T}C_x\\
=\ &\frac{C_xC_x}{T^{\kappa_x}}\sum_{t=1}^{T}\Phi_T^{-(T-t)}\left(\frac{\Phi_T^{-t}}{T^{\kappa_x/2+\min\{\eta,\kappa_x\}}}\Psi_{T,t-1}^{\rho}\right)\left(\frac{\Phi_T^{-t}}{T^{\kappa_x/2+\min\{\eta,\kappa_x\}}}\Psi_{T,t-1}^{\rho}\right)'\Phi_T^{-T}C_xC_x+o_p(1)\\
\Rightarrow\ &C_xC_{z,\kappa_x,\eta}C_x\int_0^\infty e^{-pC_x}Y_{C_x}^{\rho}Y_{C_x}^{\rho\prime}e^{-pC_x}dpC_{z,\kappa_x,\eta}C_xC_x\\
=\ &C_xC_{z,\kappa_x,\eta}W_{C_x}^{(1)}C_{z,\kappa_x,\eta}C_x.
\end{aligned}$$

In a similar fashion, we can obtain that

$$\begin{aligned}
\frac{c_uC_x}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{X_\rho,t-1}Z_{t-1}'\Phi_T^{-T}\ &\Rightarrow\ C_{z,\kappa_x,\eta}c_uC_x\int_0^\infty e^{-pC_x}Y_{C_x}^{\rho}Y_{C_x}'e^{-pC_x}dpC_{z,\kappa_x,\eta},\\
&=\ C_xC_{z,\kappa_x,\eta}W_{C_x}^{(3)}C_{z,\kappa_x,\eta}C_x,
\end{aligned}$$

$$\frac{c_u}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{t-1}Z_{X_\rho,t-1}'\Phi_T^{-T}C_x\Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(2)}C_{z,\kappa_x,\eta}C_x,$$

$$\frac{c_u^2}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{t-1}Z_{t-1}'\Phi_T^{-T}\Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(4)}C_{z,\kappa_x,\eta}C_x.$$

The above joint convergence implies

$$\frac{T^{2\kappa_x}}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_t\tilde{Z}_t'\Phi_T^{-T}\Rightarrow C_x C_{z,\kappa_{x,\eta}} W_{C_x} C_{z,\kappa_{x,\eta}} C_x.$$

For $\sum_{t=2}^{T}\tilde{Z}_{t-1}z_{0,t}^{\mu}$, we have

$$\begin{aligned}
& \frac{1}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_{t-1}z_{0,t}^{\mu} \\
= & \frac{1}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\left(\frac{C_x}{T^{\kappa_x}}Z_{X_\rho,t-2}-\frac{c_u}{T^{\kappa_u}}Z_{t-2}+Z_{\rho,t-1}\right)z_{0,t}^{\mu} \\
= & \frac{C_x}{T^{\min\{\eta,\kappa_x\}+\kappa_x}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{X_\rho,t-2}z_{0,t}^{\mu}-\frac{c_u}{T^{\min\{\eta,\kappa_x\}+\kappa_u}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{t-2}z_{0,t}^{\mu}+ \\
& \frac{1}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{\rho,t-1}z_{0,t}^{\mu}. \qquad (21)
\end{aligned}$$

Since $\kappa_x=\kappa_u$, we can express

$$\begin{aligned}
\frac{T^{\kappa_x}}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_{t-1}z_{0,t}^{\mu} &= \frac{T^{\kappa_x}}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\left(\frac{C_x}{T^{\kappa_x}}Z_{X_\rho,t-2}-\frac{c_u}{T^{\kappa_u}}Z_{t-2}+Z_{\rho,t-1}\right)z_{0,t}^{\mu} \\
&= \frac{1}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\left(C_xZ_{X_\rho,t-2}-c_uZ_{t-2}+Z_{\rho,t-1}\right)z_{0,t}^{\mu} \\
&= \frac{1}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\left(C_xZ_{X_\rho,t-2}-c_uZ_{t-2}\right)z_{0,t}+o_p(1), \quad (22)
\end{aligned}$$

where the third equality is implied by Lemma 6 in Yang et al. (2020). For the first term in (22), we can show that

$$\begin{aligned}
& \frac{1}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\left(C_xZ_{X_\rho,t-2}-c_uZ_{t-2}\right)z_{0,t} \\
= & \frac{1}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\left(C_x\frac{C_x}{T^{\kappa_x}}\Psi_{T,t-2}^{\rho}-c_u\frac{C_x}{T^{\kappa_x}}\Psi_{T,t-2}\right)z_{0,t}+o_p(1) \\
= & \frac{1}{T^{\kappa_x/2}}\sum_{t=k_T+k_T'+2}^{T}\Phi_T^{t-T}\left[\frac{\Phi_T^{-t}}{T^{\kappa_x/2+\min\{\eta,\kappa_x\}}}\left(C_xC_x\Psi_{T,t-2}^{\rho}-c_uC_x\Psi_{T,t-2}\right)\right]z_{0,t}+o_p(1)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{T^{\kappa_x/2}} \sum_{t=k_T+k_T'+2}^{T} \Phi_T^{t-T} \left[\frac{\Phi_T^{-t}}{T^{\kappa_x/2+\min\{\eta,\kappa_x\}}} \left(C_x C_x \Psi_{T,t-2}^{\rho} - c_u C_x \Psi_{T,t-2} \right) \right] z_{0,t} + o_p(1) \\
&= \frac{1}{T^{\kappa_x/2}} \sum_{t=k_T+k_T'+2}^{T} \Phi_T^{t-T} z_{0,t} \left[C_x C_{z,\kappa_x,\eta} \left(C_x Y_{C_x}^{\rho} - c_u Y_{C_x} \right) \right] + o_p(1) \\
&= C_x C_{z,\kappa_x,\eta} \frac{1}{T^{\kappa_x/2}} \sum_{t=k_T+k_T'+2}^{T} \Phi_T^{t-T} z_{0,t} \left[\left(C_x Y_{C_x}^{\rho} - c_u Y_{C_x} \right) \right] + o_p(1) \\
&= C_x C_{z,\kappa_x,\eta} \frac{1}{T^{\kappa_x/2}} \sum_{t=1}^{T} \Phi_T^{t-T} z_{0,t} \left[\left(C_x Y_{C_x}^{\rho} - c_u Y_{C_x} \right) \right] + o_p(1) \\
&\Rightarrow C_x C_{z,\kappa_x,\eta} \times MN \left(0, \int_0^{\infty} e^{-pC_x} \left(C_x Y_{C_x}^{\rho} - c_u Y_{C_x} \right) \left(C_x Y_{C_x}^{\rho} - c_u Y_{C_x} \right)' e^{-pC_x} dp \times \sigma_z^2 \right),
\end{aligned}
$$

where the convergence in last line is obtained following the same steps as in Equations (22) to (26) of Magdalinos and Phillips (2009).

$$
W_{C_x} = \int_0^{\infty} e^{-pC_x} \left(C_x Y_{C_x}^{\rho} - c_u Y_{C_x} \right) \left(C_x Y_{C_x}^{\rho} - c_u Y_{C_x} \right)' e^{-pC_x} dp.
$$

Eventually, we have

$$
\frac{T^{\kappa_x}}{T^{\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} \tilde{Z}_{t-1} z_{0,t}^{\mu} \Rightarrow C_x C_{z,\kappa_x,\eta} \times MN \left(0, W_{C_x} \times \sigma_z^2 \right).
$$

Suppose that $\kappa_x > \kappa_u$. $\frac{c_u^2}{T^{2\kappa_u}} \sum_{t=p+1}^{T} Z_{t-1} X_{t-1}'$ and $\frac{c_u}{T^{\min\{\eta,\kappa_x\}+\kappa_u}} \sum_{t=p+1}^{T} \Phi_T^{-T} Z_{t-2} z_{0,t}^{\mu}$ asymptotically dominate the other terms in (19) and (21), respectively. Correspondingly, we have

$$
\begin{aligned}
\frac{T^{2\kappa_u}}{T^{\kappa_x+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} \tilde{Z}_t \tilde{X}_t' \Phi_T^{-T} &= \frac{c_u^2}{T^{\kappa_x+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} Z_{t-1} X_{t-1}' \Phi_T^{-T} + o_p(1) \\
&\Rightarrow C_x C_{z,\kappa_x,\eta} W_{C_x}^{(4)},
\end{aligned}
$$

$$
\begin{aligned}
\frac{T^{2\kappa_u}}{T^{2\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} \tilde{Z}_t \tilde{Z}_t' \Phi_T^{-T} &= \frac{c_u^2}{T^{2\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} Z_{t-1} Z_{t-1}' \Phi_T^{-T} + o_p(1) \\
&\Rightarrow C_x C_{z,\kappa_x,\eta} W_{C_x}^{(4)} C_{z,\kappa_x,\eta} C_x,
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{T^{\kappa_u}}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_{t-1}z_{0,t}^{\mu} &= \frac{c_u}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{t-2}z_{0,t}+o_p(1)\\
&\Rightarrow C_xC_{z,\kappa_x,\eta}\times MN\left(0,W_{C_x}^{(4)}\times\sigma_z^2\right).
\end{aligned}
$$

Similarly, if $\kappa_u > \kappa_x$, we can show

$$
\begin{aligned}
\frac{T^{2\kappa_x}}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_t\tilde{X}_t'\Phi_T^{-T} &= \frac{C_x}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}X_{\rho,t-1}Z_{X_\rho,t-1}'\Phi_T^{-T}C_x+o_p(1),\\
&\Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(1)},
\end{aligned}
$$

$$
\begin{aligned}
\frac{T^{\kappa_x}}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_t\tilde{Z}_t'\Phi_T^{-T} &= \frac{C_x}{T^{2\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{X_\rho,t-1}Z_{X_\rho,t-1}'\Phi_T^{-T}C_x+o_p(1),\\
&\Rightarrow C_xC_{z,\kappa_x,\eta}W_{C_x}^{(1)}C_{z,\kappa_x,\eta}C_x.
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{T^{\kappa x}}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_{t-1}z_{0,t}^{\mu} &= \frac{C_x}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}Z_{X_\rho,t-2}z_{0,t}+o_p(1)\\
&\Rightarrow C_xC_{z,\kappa_x,\eta}\times MN\left(0,W_{C_x}^{(1)}\times\sigma_z^2\right).
\end{aligned}
$$

This completes the proof of Lemma A.3. ■

B. Proof of main theorems

Before we prove Lemma 2.1, it is useful to present the following Lemma that is a more detailed version of Lemma 2.1.

Lemma A.4. *Under the same set of assumptions as in Lemma 2.1, as $T \to \infty$, we have following set of results:*
(i). if X_t is LUR and $u_{0,t}$ is stationary, then

$$
\begin{aligned}
T\left(\hat{\beta}-\beta\right) &\Rightarrow \Psi_x, &(23)\\
W_T &\Rightarrow \Psi_x'R'\left[R\left[\int_0^1 J_{C_x}^{\mu}(r)J_{C_x}^{\mu}(r)'dr\right]^{-1}\Omega_uR'\right]^{-1}R\Psi_x; &(24)
\end{aligned}
$$

(ii). if both X_t and $u_{0,t}$ are LUR, then

$$\begin{aligned} \hat{\beta} - \beta &\Rightarrow \hat{\beta}_\infty \equiv \left[\int_0^1 J^\mu_{C_x}(r) J^\mu_{C_x}(r)' dr\right]^{-1} \int_0^1 J^\mu_{C_x}(r) J^\mu_{c_u}(r) dr, \\ W_T &= O_p\left(\frac{T}{M_T}\right); \end{aligned} \tag{25}$$

(iii). if X_t is LUR and $u_{0,t}$ is MI, then

$$\begin{aligned} \frac{T^2}{T^{1+\kappa_u}}\left(\hat{\beta} - \beta\right) &\Rightarrow -\left[\int_0^1 J^\mu_{C_x}(r) J^\mu_{C_x}(r)' dr\right]^{-1} \times \\ &\quad \left(2\left(\Lambda'_{01} + \Sigma'_{01}\right) + \int_0^1 J^\mu_{C_x}(r) dB_0(r) - B_0(1) \int_0^1 J^\mu_{C_x}(r) dr\right), \\ W_T &= O_p\left(\frac{T^{\kappa_u}}{M_T}\right); \end{aligned} \tag{26}$$

(iv). if X_t is MI and $u_{0,t}$ is stationary, then

$$\begin{aligned} T^{\kappa_x}\left(\hat{\beta} - \beta\right) &\xrightarrow{p} V_{xx}^{-1} \Lambda_{\tilde{z}\varepsilon}, \\ T^{-1+\kappa_x} W_T &\xrightarrow{p} \left(R V_{xx}^{-1} \Lambda_{\tilde{z}\varepsilon}\right)' \left[V_{xx}^{-1} \Omega_u\right]^{-1} \left(R V_{xx}^{-1} \Lambda_{\tilde{z}\varepsilon}\right); \end{aligned} \tag{27}$$

furthermore, if $u_{0,t}$ is an iid random sequence with zero mean and finite variance, then

$$T^{(1+\kappa_x)/2}\left(\hat{\beta} - \beta\right) \Rightarrow N(0, V_{xx}^{-1} \Sigma_{00}), \tag{28}$$

$$W_T \Rightarrow \chi^2(q); \tag{29}$$

(v). if X_t is MI and $u_{0,t}$ is LUR, then

$$\begin{aligned} \hat{\beta} - \beta &\Rightarrow [C_x V_{xx}]^{-1} \left[\int J^\mu_{c_u}(r) dB_1(r) + 2\Lambda_{01} + \Sigma_{01}\right]', \\ W_T &= O_p\left(\frac{T^{\kappa_x}}{M_T}\right); \end{aligned}$$

(vi). Assume both X_t and $u_{0,t}$ are MI. If $\kappa_u < \kappa_x$, then

$$\begin{aligned} T^{\kappa_x - \kappa_u}\left(\hat{\beta} - \beta\right) &\xrightarrow{p} 2 V_{xx}^{-1} c_u^{-1} \left(\Lambda'_{01} + \Sigma'_{01}\right), \\ W_T &= O_p\left(\frac{T^{1-\kappa_x+\kappa_u}}{M_T}\right); \end{aligned} \tag{30}$$

if $\kappa_u = \kappa_x$, *then*

$$\begin{aligned} \hat{\beta} - \beta &\xrightarrow{p} 2V_{xx}^{-1} \left(C_x + c_u I_k\right)^{-1} \left(\Lambda_{01}' + \Sigma_{01}'\right), \\ W_T &= O_p\left(\frac{T}{M_T}\right) \xrightarrow{p} \infty; \end{aligned}$$

if $\kappa_x < \kappa_u$, *then*

$$\begin{aligned} \hat{\beta} - \beta &\xrightarrow{p} 2V_{xx}^{-1} C_x^{-1} \left(\Lambda_{01}' + \Sigma_{01}'\right), \\ W_T &= O_p\left(\frac{T^{1-\kappa_x+\kappa_u}}{M_T}\right); \end{aligned}$$

(vii). if X_t *is ME and* $u_{0,t}$ *is stationary, then*

$$\begin{aligned} T^{\kappa_x} \Phi_T^T \left(\hat{\beta} - \beta\right) &\Rightarrow MN\left(0, \Psi_{Y_{C_x}}^{-1} \Omega_u\right), \\ W_T &\Rightarrow \chi^2(q); \end{aligned} \tag{31}$$

(viii). if X_t *is ME and* $u_{0,t}$ *is LUR, then*

$$\begin{aligned} T^{(\kappa_x-1)/2} \Phi_T^T \left(\hat{\beta} - \beta\right) &\Rightarrow \Psi_{Y_{C_x}}^{-1} C_x^{-1} Y_{C_x} J_{c_u}^{\mu}(r), \\ W_T &= O_p\left(\frac{T^{\kappa_x}}{M_T}\right); \end{aligned}$$

(ix). Assume X_t *is ME and* $u_{0,t}$ *is MI. If* $\kappa_u = \kappa_x$, *then*

$$\begin{aligned} \Phi_T^T \left(\hat{\beta} - \beta\right) &\Rightarrow MN\left(0, \Xi\right), \\ W_T &= O_p\left(\frac{T^{\kappa_x}}{M_T}\right); \end{aligned}$$

if $\kappa_u > \kappa_x$, *then*

$$\begin{aligned} T^{(\kappa_x-\kappa_u)/2} \Phi_T^T \left(\hat{\beta} - \beta\right) &\Rightarrow MN\left(0, \frac{\left(C_x \Psi_{Y_{C_x}}\right)^{-1}}{-2c_u} Y_{C_x}^2 \left(\Psi_{Y_{C_x}} C_x\right)^{-1} \Omega_{00}\right), \\ W_T &= O_p\left(\frac{T^{\kappa_x}}{M_T}\right); \end{aligned}$$

if $\kappa_x > \kappa_u$, *then*

$$T^{\kappa_x-\kappa_u} \Phi_T^T \left(\hat{\beta} - \beta\right) \Rightarrow MN\left(0, \frac{\Psi_{Y_{C_x}}^{-1}}{2c_u^2} Y_{C_x}^2 C_x^2 \Psi_{Y_{C_x}}^{-1} \Omega_{00}\right),$$

$$W_T = O_p\left(\frac{T^{\kappa_u}}{M_T}\right).$$

In the above, we adopt following notations:

$$\begin{aligned}
J^{\mu}_{C_x}(r) &= J_{C_x}(r) - \int_0^1 J_{C_x}(r)dr, \\
J_{C_x}(r) &= \int_0^r e^{(r-s)C_x} dB_1(s), \\
\Psi_x &= \left[\int_0^1 J^{\mu}_{C_x}(r)J^{\mu}_{C_x}(r)'dr\right]^{-1}\left[\int_0^1 J_{C_x}(r)dB_0(r) + \Omega_{\varepsilon u} - \int_0^1 J_{C_x}(r)drB_0(1)\right], \\
\Psi_{Y_{C_x}} &= \int_0^{\infty} e^{-pC_x} Y_{C_x} Y'_{C_x} e^{-pC_x} dp, \\
Y_{C_x} &= N\left(0, \int_0^{\infty} e^{-pC_x}\Omega_{11}e^{-pC_x}dp\right) \\
\Xi &= \left[(c_u I_k + C_x)\Psi_{Y_{C_x}}\right]^{-1} Y^2_{C_x}\Omega_{00}\left[\frac{1}{-2c_u} + \frac{1}{2}C_x^2 + 2[c_u I_k + C_x]^{-1}\right]\left[(c_u I_k + C_x)\Psi_{Y_{C_x}}\right]^{-1},
\end{aligned}$$

where $[B_0(r), B_1(r)']'$ is a $(k+1)\times 1$ vector Brownian motion with variance matrix Ω.

Proof of Lemma 2.1 and A.4. Note that the results in Table 1 can be extracted from Lemma A.4, so it is sufficient to prove Lemma A.4. To estimate $\Omega_u = \Sigma_u + \Lambda_u + \Lambda'_u$ and $\Sigma_u = E\left[u_{0,t}u'_{0,t}\right]$, we use the following estimators

$$\hat{\Omega}_u = \frac{1}{T}\sum_{j=-M_T}^{M_T}\left(1 - \frac{j}{1+M_T}\right)\sum_{t=h+1}^{T} e_{0,t}e_{0,t-h}, \hat{\Sigma}_u = \frac{1}{T}\sum_{t=1}^{T} e^2_{0,t} \tag{32}$$

$$\hat{\Omega}_{01} = \frac{1}{T}\sum_{j=-M_T}^{M_T}\left(1 - \frac{j}{1+M_T}\right)\sum_{t=h+1}^{T} e_{0,t}e_{1,t-h}, \tag{33}$$

$$\hat{\Omega}_{11} = \frac{1}{T}\sum_{j=-M_T}^{M_T}\left(1 - \frac{j}{1+M_T}\right)\sum_{t=h+1}^{T} e_{1,t}e_{1,t-h}, \tag{34}$$

where M_T is the bandwidth with $M_T \to \infty$ and $M_T/T \to 0$, $e_{0,t}$ and $e_{1,t} = X^{\mu}_t - \hat{\Phi}_T X^{\mu}_{t-1}$ are the OLS residuals from the first and third equations in model (2), respectively. We can express the centered OLS estimator as

$$\hat{\beta} - \beta = \left[\sum_{t=1}^{T} X^{\mu}_{t-1}X^{\mu\prime}_{t-1}\right]^{-1}\sum_{t=1}^{T} X^{\mu}_{t-1}u^{\mu}_{0,t}. \tag{35}$$

(i). Under case 1, Lemma 3.1 in Phillips (1988) provides the following joint convergence,

$$\frac{1}{T^2}\sum_{t=1}^{T} X_{t-1}X'_{t-1} \;\Rightarrow\; \int_0^1 J_{C_x}(r)J_{C_x}(r)'dr, \tag{36}$$

$$\frac{1}{T}\sum_{t=1}^{T} X_{t-1}u_{0,t} \;\Rightarrow\; \int_0^1 J_{C_x}(r)dB_0(r)+\Lambda_{\varepsilon u}, \tag{37}$$

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T} X_{t-1} \;\Rightarrow\; \int_0^1 J_{C_x}(r)dr,$$

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_{0,t} \;\Rightarrow\; B_0(1), \tag{38}$$

The above results and continuous mapping theorem implies (23). To study the limit of Wald statistic W_T, let $Q_W \equiv \left[\sum_{t=1}^{T} X^{\mu}_{t-1}X^{\mu\prime}_{t-1}\right]^{-1}\hat{\Omega}_u$, where $\hat{\Omega}_u \xrightarrow{p} \Omega_u$. We have

$$T^2 Q_W \Rightarrow \left[\int_0^1 J^{\mu}_{C_x}(r)J^{\mu}_{C_x}(r)'dr\right]^{-1}\Omega_u. \tag{39}$$

Thus, $Q_W = O_p(T^{-2})$. (23) and (39) imply, under the null hypothesis $R\beta = r$, that

$$\begin{aligned} W_T &= \left(R\hat{\beta}-r\right)'\left[RQ_W R'\right]^{-1}\left(R\hat{\beta}-r\right) \\ &= \left(RT(\hat{\beta}-\beta)\right)'\left[RT^2Q_W R'\right]^{-1}RT(\hat{\beta}-\beta) \\ &\Rightarrow \Psi_x' R'\left[R\left[\int_0^1 J^{\mu}_{C_x}(r)J^{\mu}_{C_x}(r)'dr\right]^{-1}\Omega_u R'\right]^{-1}R\Psi_x. \end{aligned}$$

(ii) Under case 2, Lemma 3.1 in Phillips (1988) gives

$$\frac{1}{T^2}\sum_{t=1}^{T} X_{t-1}u_{0,t} \Rightarrow \int_0^1 J_{C_x}(r)J_{c_u}(r)dr. \tag{40}$$

With Lemma 3.1 in Phillips (1988) and the continuous mapping theorem, we can obtain

$$\hat{\beta}-\beta \Rightarrow \frac{\int_0^1 J^{\mu}_{C_x}(r)J^{\mu}_{c_u}(r)dr}{\int_0^1 J^{\mu}_{C_x}(r)J^{\mu}_{C_x}(r)'dr}.$$

For the Wald statistic W_T, let $\hat{\Omega}_u \equiv \sum_{h=-M+1}^{M}(1-\frac{h}{M+1})\hat{\gamma}_h$, $\hat{\gamma}_h \equiv \frac{1}{T}\sum_{t=h+1}^{T} e_{0,t}e_{0t-h}$, and $e_{0,t} \equiv y_t^\mu - \hat{\beta} X_{t-1}^\mu$ be the OLS residual. For $\hat{\gamma}_h$,

$$\begin{aligned}\hat{\gamma}_h &= \frac{1}{T}\sum_{t=1}^{T}\left(y_t^\mu - \hat{\beta}' X_{t-1}^\mu\right)\left(y_{t-h}^\mu - \hat{\beta}' X_{t-1-h}^\mu\right) \\ &= \frac{1}{T}\sum_{t=1}^{T}\left(u_{0,t}^\mu + \left(\beta - \hat{\beta}\right)' X_{t-1}^\mu\right)\left(u_{0,t-h}^\mu + \left(\beta - \hat{\beta}\right)' X_{t-1-h}^\mu\right) \qquad (41)\end{aligned}$$

Under the local-to-unit-root assumption, both $u_{0,t-h}^\mu$ and X_{t-1-h}^μ are $O_p(\sqrt{T})$. Thus,

$$\hat{\gamma}_h = O_p(T). \qquad (42)$$

Let $M = M_T$ where $M_T \to \infty$ as $T \to \infty$, and $M_T/T \to 0$. By construction, since the Bartlett kernel function includes $\hat{\gamma}_h$ up to $h = M_T - 1$, we have

$$\hat{\Omega}_u = O_p(M_T T). \qquad (43)$$

By (38), the continuous mapping theorem and (43), we have

$$Q_W = \left[\sum_{t=1}^{T} X_{t-1}^\mu X_{t-1}^{\mu\prime}\right]^{-1} \hat{\Omega}_u = O_p(T^{-2}) O_p(M_T T) = O_p\left(\frac{M_T}{T}\right). \qquad (44)$$

Eventually,

$$\begin{aligned}W_T &= \left(R\hat{\beta} - r\right)' \left[R Q_W R'\right]^{-1} \left(R\hat{\beta} - r\right) \\ &= O_p(1)\left[O_p\left(\frac{M_T}{T}\right)\right]^{-1} O_p(1) \\ &= O_p\left(\frac{T}{M_T}\right) \xrightarrow{p} \infty.\end{aligned}$$

(iii). For case 3, let us express $X_{t-1}^\mu u_{0,t}^\mu$ as

$$\begin{aligned}X_{t-1}^\mu u_{0,t}^\mu &= \left(\Phi_T X_{t-2}^\mu + \varepsilon_{1,t-1}^\mu\right)\left(\rho_T u_{0,t-1}^\mu + \varepsilon_{0,t-1}^\mu\right) \\ &= \Phi_T \rho_T X_{t-2}^\mu u_{0,t-1}^\mu + \Phi_T X_{t-2}^\mu \varepsilon_{0,t-1}^\mu \\ &\quad + \rho_T \varepsilon_{1,t-1}^\mu u_{0,t-1}^\mu + \varepsilon_{1,t-1}^\mu \varepsilon_{0,t-1}^\mu.\end{aligned}$$

Summing up the above expression over $t = 1, ..., T$ and subtracting both sides by

$\sum_{t=1}^{T} X_{t-2}^{\mu} u_{0,t-1}^{\mu}$ yield

$$
\begin{aligned}
X_{T-1}^{\mu} u_{0,T}^{\mu} &= (\Phi_T \rho_T - I_k) \sum_{t=1}^{T} X_{t-2}^{\mu} u_{0,t-1}^{\mu} + \Phi_T \sum_{t=1}^{T} X_{t-2}^{\mu} \varepsilon_{0,t-1}^{\mu} \\
&\quad + \rho_T \sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu} u_{0,t-1}^{\mu} + \sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu} \varepsilon_{0,t-1}^{\mu}. \qquad (45)
\end{aligned}
$$

Since $\Phi_T \rho_T - I_k = \left(\left(1 + \frac{C_x}{T}\right)\left(1 + \frac{c_u}{T^{\kappa_u}}\right) - I_k\right) = \frac{c_u}{T^{\kappa_u}} + O(T^{-1})$,

$$
\begin{aligned}
\left(\frac{c_u}{T^{\kappa_u}} + O(T^{-1})\right) \sum_{t=1}^{T} X_{t-2}^{\mu} u_{0,t-1}^{\mu} &= X_{T-1}^{\mu} u_{0,T}^{\mu} - \Phi_T \sum_{t=1}^{T} X_{t-2}^{\mu} \varepsilon_{0,t-1}^{\mu} \\
&\quad - \rho_T \sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu} u_{0,t-1}^{\mu} - \sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu} \varepsilon_{0,t-1}^{\mu}.
\end{aligned}
$$

By (38), Lemma 3.1 in Phillips (1988), Lemma A.1 in Lin and Tu (2020) and the ergodic theorem, we can obtain the following limits when $\Phi_T \to I_k$ and $\rho_T \to 1$

$$
\begin{aligned}
X_{T-1}^{\mu} u_{0,T}^{\mu} &= O_p(T^{1/2}) O_p(T^{\frac{\kappa_u}{2}}) = O_p(T^{\frac{1+\kappa_u}{2}}), \qquad (46) \\
\frac{1}{T} \sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu} \varepsilon_{0,t-1}^{\mu} &\xrightarrow{p} \Sigma_{01}', \qquad (47) \\
\frac{1}{T} \sum_{t=1}^{T} X_{t-2}^{\mu} \varepsilon_{0,t-1}^{\mu} &= \frac{1}{T} \sum_{t=1}^{T} X_{t-2}^{\mu} \left(\varepsilon_{0,t-1} - \frac{1}{T} \sum_{t=1}^{T} \varepsilon_{0,t} \right) \\
&= \frac{1}{T} \sum_{t=1}^{T} X_{t-2}^{\mu} \varepsilon_{0,t-1} - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_{0,t} \frac{1}{T^{3/2}} \sum_{t=1}^{T} X_{t-2}^{\mu} \\
&\Rightarrow \int_0^1 J_{C_x}^{\mu}(r) dB_0(r) + \Lambda_{01}' - B_0(1) \int_0^1 J_{C_x}^{\mu}(r) dr. \qquad (48)
\end{aligned}
$$

To study the limit of $\frac{1}{T} \sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu} u_{0,t-1}^{\mu}$, note that

$$
\begin{aligned}
\frac{1}{T} \sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu} u_{0,t-1}^{\mu} &= \frac{1}{T} \sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu} \left(\rho_T u_{0,t-2}^{\mu} + \varepsilon_{0,t-1}^{\mu} \right) \\
&= \rho_T \frac{1}{T} \sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu} u_{0,t-2}^{\mu} + \frac{1}{T} \sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu} \varepsilon_{0,t-1}^{\mu}, \qquad (49)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}u_{0,t-2}^{\mu} &= \frac{1}{T}\sum_{t=1}^{T}\left(\varepsilon_{1,t-1}-\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t}\right)\left(u_{0,t-2}-\frac{1}{T}\sum_{t=1}^{T}u_{0,t}\right) \\
&= \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}u_{0,t-2}-\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}\frac{1}{T}\sum_{t=1}^{T}u_{0,t}-\frac{1}{T}\sum_{t=1}^{T}u_{0,t-2}\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t} \\
&\quad +\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t}\frac{1}{T}\sum_{t=1}^{T}u_{0,t}. \qquad (50)
\end{aligned}
$$

From Lemma 3.1 and Lemma 3.3 in Magdalinos and Phillips (2009), we know that

$$
\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}u_{0,t-2}\xrightarrow{p}\Lambda_{01}'. \qquad (51)
$$

For the second term in (50), note that $\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}=O_p(T^{-1/2})$ from the functional central limit theorem. Also, writing

$$
\sum_{t=1}^{T}u_{0,t}=\rho_T\sum_{t=1}^{T}u_{0,t-1}+\sum_{t=1}^{T}\varepsilon_{0,t}
$$

leads to

$$
\begin{aligned}
u_{0,T}-u_{0,0} &= (\rho_T-1)\sum_{t=1}^{T}u_{0,t-1}+\sum_{t=1}^{T}\varepsilon_{0,t} \\
&= \left(\frac{c_u}{T^{\kappa_u}}\right)\sum_{t=1}^{T}u_{0,t-1}+\sum_{t=1}^{T}\varepsilon_{0,t}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
\frac{1}{T^{1/2+\kappa_u}}\sum_{t=1}^{T}u_{0,t-1} &= c^{-1}\frac{(u_{0,T}-u_{0,0})}{T^{1/2}}-c^{-1}\frac{1}{T^{1/2}}\sum_{t=1}^{T}\varepsilon_{0,t} \\
&= o_p(1)+O_p(1)=O_p(1). \qquad (52)
\end{aligned}
$$

Therefore,

$$
\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}\frac{1}{T}\sum_{t=1}^{T}u_{0,t}=O_p(T^{-1/2})T^{-1}O_p(T^{1/2+\kappa_u})=O_p(T^{\kappa_u-1}).
$$

Following similar steps, we can also show the third term $\frac{1}{T}\sum_{t=1}^{T}u_{0,t-2}\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t}=$

$O_p(T^{\kappa_u-1})$. Finally, for the last term, $\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t}\right)\left(\frac{1}{T}\sum_{t=1}^{T}u_{0,t}\right) = O_p(T^{\kappa_u-1})$. Therefore, we have

$$\frac{1}{T}\sum_{t=1}^{T}\varepsilon^{\mu}_{1,t-1}u^{\mu}_{0,t-2} = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}u_{0,t-2} + o_p(1) \xrightarrow{p} \Lambda'_{01}. \tag{53}$$

From (51), $\rho_T \to 1$, $\frac{1}{T}\sum_{t=1}^{T}\varepsilon^{\mu}_{1,t-1}\varepsilon^{\mu}_{0,t-1} \xrightarrow{p} \Sigma'_{01}$ because of the ergodic theorem. Therefore,

$$\frac{1}{T}\sum_{t=1}^{T}\varepsilon^{\mu}_{1,t-1}u^{\mu}_{0,t-1} \xrightarrow{p} \Lambda'_{01} + \Sigma'_{01}. \tag{54}$$

By (46), (47), (48) and (54), we have

$$\frac{c_u}{T^{\kappa_u}}\frac{1}{T}\sum_{t=1}^{T}X^{\mu}_{t-1}u^{\mu}_{0,t} \xrightarrow{p} -\left(2\Lambda'_{01} + 2\Sigma'_{01} + \int_0^1 J^{\mu}_{C_x}(r)dB_0(r) - B_0(1)\int_0^1 J^{\mu}_{C_x}(r)dr\right), \tag{55}$$

and with $\frac{1}{T^2}\sum_{t=1}^{T}X^{\mu}_{t-1}X^{\mu\prime}_{t-1} \Rightarrow \int_0^1 J^{\mu}_{C_x}(r)J^{\mu}_{C_x}(r)'dr$, we have

$$\begin{aligned}
&\frac{T^2}{T^{1+\kappa_u}}\left(\hat{\beta}-\beta\right) \\
= &\left[\frac{1}{T^2}\sum_{t=1}^{T}X^{\mu}_{t-1}X^{\mu\prime}_{t-1}\right]^{-1}\frac{c_u}{T^{\kappa_u}}\frac{1}{T}\sum_{t=1}^{T}X^{\mu}_{t-1}u^{\mu}_{0,t} \\
\Rightarrow &-\left[\int_0^1 J^{\mu}_{C_x}(r)J^{\mu}_{C_x}(r)'dr\right]^{-1}\left(2\Lambda'_{01} + 2\Sigma'_{01} + \int_0^1 J^{\mu}_{C_x}(r)dB_0(r) - B_0(1)\int_0^1 J^{\mu}_{C_x}(r)dr\right).
\end{aligned}$$

This establishes (26).

We now study the limit of the test statistic. For the scaled sample covariance $\frac{1}{T^{\kappa_u}}\hat{\gamma}_h = \frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}e_{0,t}e_{0t-h}$,

$$\begin{aligned}
&\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}e_{0,t}e_{0,t-h} \\
= &\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}\left(u^{\mu}_{0,t} + \left(\beta-\hat{\beta}\right)'X^{\mu}_{t-1}\right)\left(u^{\mu}_{0,t-h} + \left(\beta-\hat{\beta}\right)'X^{\mu}_{t-1-h}\right) \\
= &\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}u^{\mu}_{0,t}u^{\mu}_{0,t-h} + \left(\beta-\hat{\beta}\right)'\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}X^{\mu}_{t-1}X^{\mu\prime}_{t-1-h}\left(\beta-\hat{\beta}\right) \\
&+\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}u^{\mu}_{0,t}X^{\mu\prime}_{t-1-h}\left(\beta-\hat{\beta}\right) + \left(\beta-\hat{\beta}\right)'\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}X^{\mu}_{t-1}u^{\mu}_{0,t-h}.
\end{aligned} \tag{56}$$

We now study the orders of the terms appears in (56). For the first term,

$$\begin{aligned}
\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}u_{0,t}^{\mu}u_{0,t-h}^{\mu} &= \frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}\left(u_{0,t}-\frac{1}{T}\sum_{t=1}^{T}u_{0,t}\right)\left(u_{0,t-h}-\frac{1}{T}\sum_{t=1}^{T}u_{0,t}\right)\\
&= \frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}u_{0,t}u_{0,t-h}-\left(\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}u_{0,t}\right)\frac{1}{T}\sum_{t=1}^{T}u_{0,t}\\
&\quad -\left(\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}u_{0,t-h}\right)\frac{1}{T}\sum_{t=1}^{T}u_{0,t}+\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}\left(\frac{1}{T}\sum_{t=1}^{T}u_{0,t}\right)^2\\
&= \frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T}u_{0,t}u_{0,t-h}+O_p(T^{\kappa_u-1}),
\end{aligned}$$

where the third equality is obtained from applying (52). For the term $\sum_{t=h+1}^{T}u_{0,t}u_{0,t-h}$, note that we can write

$$\begin{aligned}
\sum_{t=h+1}^{T}u_{0,t}u_{0,t-h} &= \sum_{t=h+1}^{T}\left(\rho_T u_{0,t-1}+\varepsilon_{0,t}\right)\left(\rho_T u_{0,t-h-1}+\varepsilon_{0,t-h}\right)\\
&= \rho_T^2\sum_{t=h+1}^{T}u_{0,t-1}u_{0,t-h-1}+\rho_T\sum_{t=h+1}^{T}u_{0,t-1}\varepsilon_{0,t-h}\\
&\quad +\rho_T\sum_{t=h+1}^{T}\varepsilon_{0,t}u_{0,t-h-1}+\sum_{t=h+1}^{T}\varepsilon_{0,t}\varepsilon_{0,t-h}. \qquad (57)
\end{aligned}$$

Subtracting both sides of (57) by $\sum_{t=h+1}^{T}u_{0,t-1}u_{0,t-h-1}$ and multiplying them by T^{-1} give

$$\begin{aligned}
&\left(\rho_T^2-1\right)\frac{1}{T}\sum_{t=h+1}^{T}u_{0,t-1}u_{0,t-h-1} \qquad (58)\\
&= \frac{1}{T}\left(u_{0,T-1}u_{0,T-h-1}-u_{0,h}u_{0,0}\right)-\rho_T\frac{1}{T}\sum_{t=h+1}^{T}u_{0,t-1}\varepsilon_{0,t-h}\\
&\quad -\rho_T\frac{1}{T}\sum_{t=h+1}^{T}\varepsilon_{0,t}u_{0,t-h-1}-\frac{1}{T}\sum_{t=h+1}^{T}\varepsilon_{0,t}\varepsilon_{0,t-h}\\
&= -\rho_T\frac{1}{T}\sum_{t=h+1}^{T}u_{0,t-1}\varepsilon_{0,t-h}-\rho_T\frac{1}{T}\sum_{t=h+1}^{T}\varepsilon_{0,t}u_{0,t-h-1}
\end{aligned}$$

$$-\frac{1}{T}\sum_{t=h+1}^{T}\varepsilon_{0,t}\varepsilon_{0,t-h}+o_p(1), \tag{59}$$

where the second equality is obtained by $u_{0,t}=O_p(T^{\kappa_u/2})$. For the last term, it is clear that

$$\frac{1}{T}\sum_{t=h+1}^{T}\varepsilon_{0,t}\varepsilon_{0,t-h}\xrightarrow{p}E[\varepsilon_{0,t}\varepsilon_{0,t-h}]=\sigma_z^2\sum_{j=0}^{\infty}c_{0,j}c_{0,j-h}.$$

We now study the limit of $\frac{1}{T}\sum_{t=h+1}^{T}u_{0,t-1}\varepsilon_{0,t-h}$. Let $\tilde{z}_{0,t}=\sum_{j=0}^{\infty}\tilde{c}_{0,j}z_{0,t-j}$ and $\tilde{c}_{0,j}=\sum_{k=j+1}^{\infty}c_{0,j}$ by using the Beveridge-Nelson decomposition. We have

$$\begin{aligned}\frac{1}{T}\sum_{t=h+1}^{T}u_{0,t-1}\varepsilon_{0,t-h} &= \frac{C_0(1)}{T}\sum_{t=h+1}^{T}u_{0,t-1}z_{0,t-h}-\frac{1}{T}\sum_{t=h+1}^{T}u_{0,t-1}\Delta\tilde{z}_{0,t-h}\\ &= \frac{C_0(1)}{T}\sum_{t=h+1}^{T}u_{0,t-1}z_{0,t-h}-\left[\frac{1}{T}\left(u_{0,T-1}\tilde{z}_{0,T-h}-u_{0,h}\tilde{z}_{0,1}\right)\right]\\ &\quad+\frac{1}{T}\sum_{t=h+1}^{T}\tilde{z}_{0,t-h-1}\Delta u_{0,t-1} \qquad (60)\\ &= \frac{C_0(1)}{T}\sum_{t=h+1}^{T}u_{0,t-1}z_{0,t-h}+\frac{1}{T}\sum_{t=h+1}^{T}\tilde{z}_{0,t-h-1}\Delta u_{0,t-1}+o_p(1).\end{aligned}$$

For the first term in (61),

$$\begin{aligned}\frac{1}{T}\sum_{t=h+1}^{T}u_{0,t-1}z_{0,t-h} &= \frac{1}{T}\sum_{t=h+1}^{T}\left[\rho_T^h u_{0,t-h-1}+\sum_{j=t-h}^{t-1}\rho_T^{t-1-j}\varepsilon_{0,j}\right]z_{0,t-h} \qquad (61)\\ &= \rho_T^h\frac{1}{T}\sum_{t=h+1}^{T}u_{0,t-h-1}z_{0,t-h}+\frac{1}{T}\sum_{t=h+1}^{T}\left(\sum_{j=t-h}^{t-1}\rho_T^{t-1-j}\varepsilon_{0,j}\right)z_{0,t-h}.\end{aligned}$$

Since

$$\begin{aligned}&\frac{1}{T}\sum_{t=h+1}^{T}\left(\sum_{j=t-h}^{t-1}\rho_T^{t-1-j}\varepsilon_{0,j}\right)z_{0,t-h}\\ =\ &\frac{1}{T}\sum_{t=h+1}^{T}\varepsilon_{0,t-1}z_{0,t-h}+\rho_T\frac{1}{T}\sum_{t=h+1}^{T}\varepsilon_{0,t-2}z_{0,t-h}+...+\rho_T^{h-1}\frac{1}{T}\sum_{t=h+1}^{T}\varepsilon_{0,t-h}z_{0,t-h}\\ \xrightarrow{p}\ &\sum_{j=1}^{h}c_{0,h-j}E\left[z_{0,t-h}^2\right]=\sigma_z^2\sum_{j=1}^{h}c_{0,h-j}.\end{aligned}$$

For $\frac{1}{T}\sum_{t=h+1}^{T} u_{0,t-h-1} z_{0,t-h}$ in (61), applying Lemma 4.2 in Magdalinos and Phillips (2009) gives

$$\frac{1}{T^{\frac{1+\kappa_u}{2}}} \sum_{t=h+1}^{T} u_{0,t-h-1} z_{0,t-h} = O_p(1).$$

Thus, $\frac{1}{T}\sum_{t=h+1}^{T} u_{0,t-h-1} z_{0,t-h} = O_p(T^{(\kappa_u - 1)/2}) = o_p(1)$. For $\frac{1}{T}\sum_{t=h+1}^{T} \tilde{z}_{0,t-h-1} \Delta u_{0,t-1}$ in (61),

$$\begin{aligned} \frac{1}{T}\sum_{t=h+1}^{T} \tilde{z}_{0,t-h-1} \Delta u_{0,t-1} &= \frac{1}{T}\sum_{t=h+1}^{T} \tilde{z}_{0,t-h-1} \left(u_{0,t-1} + \varepsilon_{0,t-1} \right) \quad (62) \\ &= \frac{(\rho_T - 1)}{T}\sum_{t=h+1}^{T} \tilde{z}_{0,t-h-1} u_{0,t-1} + \frac{1}{T}\sum_{t=h+1}^{T} \tilde{z}_{0,t-h-1} \varepsilon_{0,t-1}. \end{aligned}$$

For the first term $\frac{1}{T}\sum_{t=h+1}^{T} \tilde{z}_{0,t-h-1} u_{0,t-1}$ in (63), note that the proofs of Lemma 4.2 in Phillips and Magdalinos (2009) and Theorem 3.2 in Phillips and Magdalinos (2007) are still applicable, we thus have

$$\frac{(\rho_T - 1)}{T}\sum_{t=h+1}^{T} \tilde{z}_{0,t-h-1} u_{0,t-1} = o_p(1).$$

For the second term, the ergodic theorem yields

$$\begin{aligned} \frac{1}{T}\sum_{t=h+1}^{T} \tilde{z}_{0,t-h-1} \varepsilon_{0,t-1} &= \frac{1}{T}\sum_{t=h+1}^{T} \left(\sum_{j=0}^{\infty} \tilde{c}_{0,j} z_{t-1-j-h} \right) \left(\sum_{j=0}^{\infty} c_{0,j} z_{t-1-j} \right) \\ &\xrightarrow{p} \sigma_z^2 \sum_{j=0}^{\infty} c_{0,j} c_{0,j-h} + \sigma_z^2 \sum_{j=0}^{\infty} \tilde{c}_{0,j} c_{0,j+h}. \end{aligned}$$

So, we can conclude that

$$\frac{1}{T}\sum_{t=h+1}^{T} u_{0,t-1} \varepsilon_{0,t-h} \xrightarrow{p} \sigma_z^2 \sum_{j=0}^{\infty} \left(c_{0,j} c_{0,j-h} + \tilde{c}_{0,j} c_{0,j+h} \right).$$

For $\frac{1}{T}\sum_{t=h+1}^{T} \varepsilon_{0,t} u_{0,t-h-1}$ in (59), by the Beveridge-Nelson decomposition and summation by parts, we have

$$\frac{1}{T}\sum_{t=h+1}^{T} \varepsilon_{0,t} u_{0,t-h-1} = \frac{C_0(1)}{T}\sum_{t=h+1}^{T} z_{0,t} u_{0,t-h-1} - \frac{1}{T}\sum_{t=h+1}^{T} u_{0,t-h-1} \Delta \tilde{z}_{0,t}$$

$$= \frac{C_0(1)}{T} \sum_{t=h+1}^{T} z_{0,t} u_{0,t-h-1} + \frac{1}{T} \sum_{t=h+1}^{T} \tilde{z}_{0,t} \Delta u_{0,t-h-1} + o_p(1).$$

For the first term $\frac{1}{T}\sum_{t=h+1}^{T} z_{0,t}u_{0,t-h-1}$, note that the proofs of Lemma 4.2 in Phillips and Magdalinos (2009) and Theorem 3.2 in Phillips and Magdalinos (2007) are applicable. We have

$$\frac{1}{T^{\frac{1+\kappa_u}{2}}} \sum_{t=h+1}^{T} z_{0,t} u_{0,t-h-1} = O_p(1).$$

For the second term $\frac{1}{T}\sum_{t=h+1}^{T} \tilde{z}_{0,t}\Delta u_{0,t-h-1}$, we have

$$\begin{aligned}
\frac{1}{T} \sum_{t=h+1}^{T} \tilde{z}_{0,t} \Delta u_{0,t-h-1} &= \frac{1}{T} \sum_{t=h+1}^{T} \tilde{z}_{0,t} \left(\left(\rho_T - 1 \right) u_{0,t-h-2} + \varepsilon_{0,t-h-1} \right) \\
&= \frac{(\rho_T - 1)}{T} \sum_{t=h+1}^{T} \tilde{z}_{0,t} u_{0,t-h-2} + \frac{1}{T} \sum_{t=h+1}^{T} \tilde{z}_{0,t} \varepsilon_{0,t-h-1} \\
&\xrightarrow{p} \sigma_z^2 \sum_{j=0}^{\infty} \tilde{c}_{0,h+1+j} c_{0,j}.
\end{aligned}$$

Eventually, since $\rho_T^2 - 1 = \frac{c_u^2}{T^{2\kappa_u}} + \frac{2c_u}{T^{\kappa_u}} = \frac{2c_u}{T^{\kappa_u}} + O\left(\frac{1}{T^{2\kappa_u}}\right)$ we can conclude that the first term in (56) has the following limit,

$$\frac{1}{T^{1+\kappa_u}} \sum_{t=h+1}^{T} u_{0,t} u_{0,t-h} \xrightarrow{p} -\frac{\sigma_z^2}{2c_u} \left(\sum_{j=0}^{\infty} \tilde{c}_{0,h+1+j} c_{0,j} + 2c_{0,j} c_{0,j-h} + \tilde{c}_{0,j} c_{0,j+h} \right).$$

For the second term in (56), note that from (26) and $X_t^{\mu} = O_p(\sqrt{T})$, we have

$$\begin{aligned}
& \left(\beta - \hat{\beta}\right)' \frac{1}{T^{1+\kappa_u}} \sum_{t=h+1}^{T} X_{t-1}^{\mu} X_{t-1-h}^{\mu\prime} \left(\beta - \hat{\beta}\right) \\
&= O_p\left(\frac{T^{1+\kappa_u}}{T^2}\right) \frac{1}{T^{1+\kappa_u}} O_p\left(T^2\right) O_p\left(\frac{T^{1+\kappa_u}}{T^2}\right) \\
&= O_p\left(\frac{T^{1+\kappa_u}}{T^2}\right) = o_p(1).
\end{aligned}$$

The same proving strategy to show (55) is applicable to the last two terms in (56). Thus, $\sum_{t=h+1}^{T} u_{0,t}^{\mu} X_{t-1-h}^{\mu\prime}$ and $\sum_{t=h+1}^{T} X_{t-1}^{\mu} u_{0,t-h}^{\mu}$ are both $O_p\left(T^{1+\kappa_u}\right)$. Therefore,

$$\frac{1}{T^{1+\kappa_u}} \sum_{t=h+1}^{T} u_{0,t}^{\mu} X_{t-1-h}^{\mu\prime} \left(\beta - \hat{\beta}\right) + \left(\beta - \hat{\beta}\right)' \frac{1}{T^{1+\kappa_u}} \sum_{t=h+1}^{T} X_{t-1}^{\mu} u_{0,t-h}^{\mu}$$

$$= \quad 2O_p(1)O_p\left(\frac{T^{1+\kappa_u}}{T^2}\right) = o_p(1).$$

We then have

$$\frac{1}{T^{1+\kappa_u}}\sum_{t=h+1}^{T} e_{0,t}e_{0,t-h} \xrightarrow{p} -\frac{\sigma_z^2}{2c_u}\left(\sum_{j=0}^{\infty}\tilde{c}_{0,h+1+j}c_{0,j} + 2c_{0,j}c_{0,j-h} + \tilde{c}_{0,j}c_{0,j+h}\right),$$

or equivalently,

$$\hat{\gamma}_h = O_p(T^{\kappa_u}). \tag{63}$$

It implies

$$\hat{\Omega}_u = O_p(M_T T^{\kappa_u}), \tag{64}$$

$$Q_W = \left[\sum_{t=1}^{T} X_{t-1}^{\mu} X_{t-1}^{\mu\prime}\right]^{-1}\hat{\Omega}_u = O_p(T^{-2})O_p(M_T T^{\kappa_u}) = O_p\left(\frac{M_T}{T^{2-\kappa_u}}\right), \tag{65}$$

and

$$\begin{aligned} W_T &= \left(R\hat{\beta} - r\right)' [RQ_W R']^{-1}\left(R\hat{\beta} - r\right) \\ &= O_p\left(\frac{T^{1+\kappa_u}}{T^2}\right) O_p\left(\frac{T^{2-\kappa_u}}{M_T}\right) O_p\left(\frac{T^{1+\kappa_u}}{T^2}\right) \\ &= O_p\left(\frac{T^{2+2\kappa_u}}{T^4}\right) O_p\left(\frac{T^{2-\kappa_u}}{M_T}\right) = O_p\left(\frac{T^{\kappa_u}}{M_T}\right). \end{aligned}$$

(iv). Note that we can express $\sum_{t=1}^{T} X_{t-1}^{\mu} u_{0,t}^{\mu}$ as

$$\begin{aligned} \sum_{t=1}^{T} X_{t-1}^{\mu} u_{0,t}^{\mu} &= \sum_{t=1}^{T}\left(X_{t-1} - \frac{1}{T}\sum_{t=1}^{T} X_t\right)\left(u_{0,t} - \frac{1}{T}\sum_{t=1}^{T} u_{0,t}\right) \\ &= \sum_{t=1}^{T} X_{t-1}u_{0,t} - \frac{1}{T}\sum_{t=1}^{T} X_{t-1}\sum_{t=1}^{T} u_{0,t} - \frac{1}{T}\sum_{t=1}^{T} X_t \sum_{t=1}^{T} u_{0,t} \\ &\quad + \frac{1}{T}\sum_{t=1}^{T} X_t \sum_{t=1}^{T} u_{0,t} \\ &= \sum_{t=1}^{T} X_{t-1}u_{0,t} + T^{-1}O_p(T^{1/2+\kappa_x})O_p(T^{1/2}) \\ &= \sum_{t=1}^{T} X_{t-1}u_{0,t} + O_p(T^{\kappa_x}). \end{aligned} \tag{66}$$

By the Beveridge-Nelson decomposition and summation by parts, we have

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} X_{t-1}^{\mu} u_{0,t}^{\mu} &= \frac{1}{T}\sum_{t=1}^{T} X_{t-1} u_{0,t} + O_p(T^{\kappa_x - 1}) \\
&= \frac{\Psi_0(1)}{T}\sum_{t=1}^{T} X_{t-1} z_{0,t} - \frac{1}{T}\sum_{t=1}^{T} X_{t-1}\Delta \tilde{z}_{0,t} + O_p(T^{\kappa_x - 1}) \\
&= \frac{\Psi_0(1)}{T}\sum_{t=1}^{T} X_{t-1} z_{0,t} + \frac{1}{T}\sum_{t=1}^{T} \Delta X_{t-1} \tilde{z}_{0,t} \\
&\quad - \left(\frac{1}{T}\left(X_{T-1}\tilde{z}_{0,T} - X_1 \tilde{z}_{0,0}\right)\right) + O_p(T^{\kappa_x - 1}) \\
&= \frac{\Psi_0(1)}{T}\sum_{t=1}^{T} X_{t-1} z_{0,t} + \frac{1}{T}\sum_{t=1}^{T} \Delta X_{t-1} \tilde{z}_{0,t} + o_p(1) \\
&= \frac{\Psi_0(1)}{T}\sum_{t=1}^{T} X_{t-1} z_{0,t} + \frac{1}{T}\sum_{t=1}^{T} \left(\left(\Phi_T - I_k\right) X_{t-2} + \varepsilon_{1,t}\right) \tilde{z}_{0,t} + o_p(1) \\
&= \frac{\Psi_0(1)}{T}\sum_{t=1}^{T} X_{t-1} z_{0,t} + \frac{C_x}{T^{\kappa_x}}\frac{1}{T}\sum_{t=1}^{T} X_{t-2}\tilde{z}_{0,t} + \frac{1}{T}\sum_{t=1}^{T} \tilde{z}_{0,t}\varepsilon_{1,t} + o_p(1).
\end{aligned}
$$

From the analogous arguments of Lemma 3.1 and Lemma 3.3 in Magdalinos and Phillips (2009), we have

$$
\begin{aligned}
\sum_{t=1}^{T} X_{t-1} z_{0,t} &= O_p(T^{(1+\kappa_u)/2}), \\
\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} X_{t-2}\tilde{z}_{0,t} &= O_p(T^{-\kappa_x/2}).
\end{aligned}
$$

Thus,

$$
\frac{1}{T}\sum_{t=1}^{T} X_{t-1}^{\mu} u_{0,t}^{\mu} = \frac{1}{T}\sum_{t=1}^{T} \tilde{z}_{0,t}\varepsilon_{1,t} + o_p(1) \xrightarrow{p} \Lambda_{\tilde{z}\varepsilon}, \tag{67}
$$

where $\Lambda_{\tilde{z}\varepsilon} = E[\tilde{z}_{0,t}\varepsilon_{1,t}]$. Applying Equation (7) to Magdalinos and Phillips (2009) gives

$$
\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} X_{t-1}^{\mu} X_{t-1}^{\mu\prime} = \frac{1}{T^{1+\kappa_u}}\sum_{t=1}^{T} X_{t-1} X_{t-1}^{\prime} + o_p(1) \xrightarrow{p} V_{xx}. \tag{68}
$$

We then have

$$T^{\kappa_x}\left(\hat{\beta}-\beta\right)=\left[\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{T}\sum_{t=1}^{T}X_{t-1}^{\mu}u_{0,t}^{\mu}\xrightarrow{p}V_{xx}^{-1}\Lambda_{\tilde{z}\varepsilon}.$$

Obtaining the limiting distribution, as in Magdalinos and Phillips (2009), requires re-centering $\sum_{t=1}^{T}X_{t-1}u_{0,t}$. Let

$$\psi=\frac{C_x}{T^{\kappa_x}}M_{0,T}+\Lambda_{\tilde{z}\varepsilon},\, M_{0,T}=\sum_{j=0}^{\infty}\tilde{c}_{0,j+1}\Sigma_{00}C_1(1)'\Phi_T^{j+1},$$

where $C_1(1)=\sum_{j=0}^{\infty}C_{1,j}$. Applying Lemma 3.4 in Magdalinos and Phillips (2009), we have

$$\begin{aligned}
&\frac{1}{T^{(1+\kappa_x)/2}}\sum_{t=1}^{T}X_{t-1}u_{0,t}-\frac{T}{T^{(1+\kappa_x)/2}}\psi\\
=\;&\frac{\Psi_0(1)}{T^{(1+\kappa_x)/2}}\sum_{t=1}^{T}X_{t-1}z_{0,t}+\frac{1}{T^{(1+\kappa_x)/2}}\sum_{t=1}^{T}\Delta X_{t-1}\tilde{z}_{0,t}-\frac{T}{T^{(1+\kappa_x)/2}}\psi+o_p(1)\\
=\;&\frac{\Psi_0(1)}{T^{(1+\kappa_x)/2}}\sum_{t=1}^{T}X_{t-1}z_{0,t}+\frac{C_x}{T^{\kappa_x}}\frac{1}{T^{(1+\kappa_x)/2}}\sum_{t=1}^{T}\left(X_{t-2}\tilde{z}_{0,t}-M_{0,T}\right)\\
&+\frac{1}{T^{(1+\kappa_x)/2}}\sum_{t=1}^{T}\left(\tilde{z}_{0,t}\varepsilon_{1,t}-\Lambda_{\tilde{z}\varepsilon}\right)+o_p(1)\\
\Rightarrow\;&\Psi_0(1)N(0,\Sigma_{00}V_{xx})=N(0,V_{xx}\Omega_{00}). \qquad (69)
\end{aligned}$$

Letting $\hat{\beta}_\psi=\hat{\beta}-T\left[\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}\psi$, we have

$$\begin{aligned}
T^{(1+\kappa_x)/2}\left(\hat{\beta}_\psi-\beta\right)\;&=\;\left[\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}\frac{\Psi_0(1)}{T^{(1+\kappa_x)/2}}\sum_{t=1}^{T}X_{t-1}z_{0,t}\\
&\Rightarrow\;V_{xx}^{-1}N(0,V_{xx}\Omega_{00})=N(0,V_{xx}^{-1}\Omega_{00}). \qquad (70)
\end{aligned}$$

We now proceed to study the sample covariance $\hat{\gamma}_h$. As will be shown, this helps us obtain the order of Wald statistic. From (41), we have

$$\hat{\gamma}_h=\frac{1}{T}\sum_{t=1}^{T}\left(u_{0,t}^{\mu}+\left(\beta-\hat{\beta}\right)'X_{t-1}^{\mu}\right)\left(u_{0,t-h}^{\mu}+\left(\beta-\hat{\beta}\right)'X_{t-1-h}^{\mu}\right).$$

Since $\beta-\hat{\beta}=O_p(T^{-\kappa_x})$, and $X_{t-1}^{\mu}=O_p(T^{\kappa_x/2})$, $\left(\beta-\hat{\beta}\right)'X_{t-1}^{\mu}=O_p(T^{-\kappa_x/2})=o_p(1)$, we have

$$\hat{\gamma}_h = \frac{1}{T}\sum_{t=1}^{T} u_{0,t}^{\mu} u_{0,t-h}^{\mu} + o_p(1) \xrightarrow{p} \gamma_h = E[u_{0,t}u_{0,t-h}].$$

By the standard asymptotics for the Newey-west estimator, we have $\hat{\Omega}_u \xrightarrow{p} \Omega_u$. Thus,

$$T^{1+\kappa_x} Q_W = \left[\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} X_{t-1}^{\mu} X_{t-1}^{\mu\prime}\right]^{-1} \hat{\Omega}_u \xrightarrow{p} V_{xx}^{-1}\Omega_u.$$

Since the Wald statistic uses the OLS estimator without re-centering, we can express

$$\begin{aligned} T^{-1+\kappa_x} W_T &= T^{\kappa_x}\left(R\hat{\beta} - r\right)' \left[RT^{1+\kappa_x} Q_W R'\right]^{-1} T^{\kappa_x}\left(R\hat{\beta} - r\right) \\ &\xrightarrow{p} \left(RV_{xx}^{-1}\Lambda_{\tilde{z}\varepsilon}\right)' \left[V_{xx}^{-1}\Omega_u\right]^{-1} \left(RV_{xx}^{-1}\Lambda_{\tilde{z}\varepsilon}\right). \end{aligned}$$

This implies that $W_T = O_p(T^{1-\kappa_x})$. For the second case, note that if $u_{0,t}$ is an mds, the OLS estimator needs not to be re-centered since $\frac{1}{T^{(1+\kappa_x)/2}}\sum_{t=1}^{T} X_{t-1}u_{0,t} = \frac{1}{T^{(1+\kappa_x)/2}}\sum_{t=1}^{T} X_{t-1}z_{0,t}$. (70) gives the limiting distribution for the OLS estimator and

$$\begin{aligned} W_T &= \left(R\hat{\beta}_\psi - r\right)' \left[RT^{1+\kappa_x} Q_W R'\right]^{-1} \left(R\hat{\beta}_\psi - r\right) \\ &= T^{(1+\kappa_x)/2} R\left(\hat{\beta}_\psi - \beta\right)' \left[RT^{1+\kappa_x} Q_W R'\right]^{-1} T^{(1+\kappa_x)/2} R\left(\hat{\beta}_\psi - \beta\right) \\ &\Rightarrow \chi^2(q). \end{aligned}$$

(v). In case 5, from (45), we have

$$\begin{aligned} \left(\Phi_T\rho_T - I_k\right)\frac{1}{T}\sum_{t=1}^{T} X_{t-2}^{\mu} u_{0,t-1}^{\mu} &= \frac{1}{T}X_{T-1}^{\mu}u_{0,T}^{\mu} - \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}\varepsilon_{0,t-1}^{\mu} \\ &\quad -\Phi_T\frac{1}{T}\sum_{t=1}^{T} X_{t-2}^{\mu}\varepsilon_{0,t-1}^{\mu} - \rho_T\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}u_{0,t-1}^{\mu}. \end{aligned}$$

Using the same arguments as (46) and (47), we have $\frac{1}{T}X_{T-1}^{\mu}u_{0,T}^{\mu} = o_p(1)$ and

$$\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}\varepsilon_{0,t-1}^{\mu} \xrightarrow{p} \Sigma_{01}'.$$

Moreover,

$$\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}u_{0,t-1}^{\mu} = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}\left(u_{0,t-2}^{\mu} + \varepsilon_{0,t-1}^{\mu}\right)$$

$$
\begin{aligned}
&= \frac{\rho_T}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}u_{0,t-2}^{\mu}+\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}\varepsilon_{0,t-1}^{\mu} \\
&\Rightarrow \left[\int J_{c_u}^{\mu}(r)dB_1(r)+\Lambda_{01}-B_1(1)\int J_{c_u}^{\mu}(r)dr\right]'+\Sigma_{01}',
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}X_{t-2}^{\mu}\varepsilon_{0,t-1}^{\mu} &= \frac{1}{T}\sum_{t=1}^{T}\left(X_{t-2}-\left(\frac{1}{T}\sum_{t=1}^{T}X_t\right)\right)\left(\varepsilon_{0,t-1}-\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{0,t}\right) \\
&= \frac{1}{T}\sum_{t=1}^{T}X_{t-2}\varepsilon_{0,t-1}+o_p(1) \\
&\xrightarrow{p} \Lambda_{01}',
\end{aligned}
$$

where the second equality is obtained by the fact that both $\frac{1}{T}\sum_{t=1}^{T}X_t$ and $\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{0,t}$ are $o_p(1)$ by (52) and the ergodic theorem. Since $\Phi_T\rho_T-I_k=\frac{C_x}{T^{\kappa_x}}+O\left(\frac{1}{T}\right)$, we can obtain

$$
\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-1}^{\mu}u_{0,t}^{\mu}\Rightarrow C_x^{-1}\left[\int J_{c_u}^{\mu}(r)dB_1(r)+2\Lambda_{01}+\Sigma_{01}-B_1(1)\int J_{c_u}^{\mu}(r)dr\right]'. \quad (71)
$$

Due to (68), we have

$$
\begin{aligned}
\hat{\beta}-\beta &= \left[\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-1}^{\mu}u_{0,t}^{\mu} \\
&\Rightarrow [C_xV_{xx}]^{-1}\left[\int J_{c_u}^{\mu}(r)dB_1(r)+2\Lambda_{01}+\Sigma_{01}-B_1(1)\int J_{c_u}^{\mu}(r)dr\right]'.
\end{aligned}
$$

As $u_{0,t}=O_p(\sqrt{T})$, we can show that $\hat{\Omega}_u=O_p(M_TT)$ using an argument similar to the derivation of (43). Thus,

$$
Q_W=\left[\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}\hat{\Omega}_u=O_p(T^{-(1+\kappa_x)})O_p(M_TT)=O_p\left(\frac{M_T}{T^{\kappa_x}}\right),
$$

$$
W_T=\left(R\hat{\beta}-r\right)'[RQ_WR']^{-1}\left(R\hat{\beta}-r\right)=O_p(1)\left[O_p\left(\frac{M_T}{T^{\kappa_x}}\right)\right]^{-1}O_p(1)=O_p\left(\frac{T^{\kappa_x}}{M_T}\right).
$$

(vi). In case 6, from (45), we have

$$
(\Phi_T\rho_T-I_k)\frac{1}{T}\sum_{t=1}^{T}X_{t-2}^{\mu}u_{0,t-1}^{\mu} = \frac{1}{T}X_{T-1}^{\mu}u_{0,T}^{\mu}-\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}\varepsilon_{0,t-1}^{\mu}
$$

$$-\Phi_T \frac{1}{T}\sum_{t=1}^{T} X_{t-2}^{\mu}\varepsilon_{0,t-1}^{\mu} - \rho_T \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}u_{0,t-1}^{\mu}.$$

Similar to the previous analysis, $\frac{1}{T}X_{T-1}^{\mu}u_{0,T}^{\mu} = o_p(1)$, $\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}\varepsilon_{0,t-1}^{\mu} \xrightarrow{p} \Sigma_{01}'$. As in (53) and (54), we have

$$\frac{1}{T}\sum_{t=1}^{T} X_{t-2}^{\mu}\varepsilon_{0,t-1}^{\mu} \xrightarrow{p} \Lambda_{01}'$$
$$\frac{1}{T}\sum_{t=1}^{T} \varepsilon_{1,t-1}^{\mu}u_{0,t-1}^{\mu} \xrightarrow{p} \Lambda_{01}' + \Sigma_{01}'.$$

Combining the above results and (68) we have

$$(\Phi_T\rho_T - I_k)\frac{1}{T}\sum_{t=1}^{T} X_{t-2}^{\mu}u_{0,t-1}^{\mu} \xrightarrow{p} -2\left(\Lambda_{01}' + \Sigma_{01}'\right), \tag{72}$$

and

$$\left[\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}(\Phi_T\rho_T - I_k)\frac{1}{T}\sum_{t=1}^{T} X_{t-2}^{\mu}u_{0,t-1}^{\mu} \xrightarrow{p} -2V_{xx}^{-1}\left(\Lambda_{01}' + \Sigma_{01}'\right)$$

Note that $\Phi_T\rho_T - I_k = \frac{C_x}{T^{\kappa_x}} + \frac{c_u}{T^{\kappa_u}} + O\left(\frac{1}{T^{\bar{\kappa}}}\right)$. Suppose $\kappa_x < \kappa_u$,

$$\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} X_{t-2}^{\mu}u_{0,t-1}^{\mu} \xrightarrow{p} -2C_x^{-1}\left(\Lambda_{01}' + \Sigma_{01}'\right), \tag{73}$$

we have

$$\begin{aligned}\hat{\beta} - \beta &= \left[\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} X_{t-2}^{\mu}u_{0,t-1}^{\mu} \\ &\xrightarrow{p} -2V_{xx}^{-1}C_x^{-1}\left(\Lambda_{01}' + \Sigma_{01}'\right).\end{aligned}$$

Suppose $\kappa_u < \kappa_x$,. Then we have

$$\frac{1}{T^{1+\kappa_u}}\sum_{t=1}^{T} X_{t-2}^{\mu}u_{0,t-1}^{\mu} \xrightarrow{p} -2c_u^{-1}\left(\Lambda_{01}' + \Sigma_{01}'\right) \tag{74}$$

and

$$
\begin{aligned}
T^{\kappa_x-\kappa_u}\left(\hat{\beta}-\beta\right) &= \left[\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{T^{1+\kappa_u}}\sum_{t=1}^{T}X_{t-2}^{\mu}u_{0,t-1}^{\mu} \\
&\xrightarrow{p} -2V_{xx}^{-1}c_u^{-1}\left(\Lambda_{01}'+\Sigma_{01}'\right).
\end{aligned}
$$

Finally, if $\kappa_u=\kappa_x$, we can deduce that

$$
\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-2}^{\mu}u_{0,t-1}^{\mu}\xrightarrow{p} -2\left(C_x+c_u\right)^{-1}\left(\Lambda_{01}'+\Sigma_{01}'\right), \tag{75}
$$

$$
\begin{aligned}
\hat{\beta}-\beta &= \left[\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-2}^{\mu}u_{0,t-1}^{\mu} \\
&\xrightarrow{p} -2V_{xx}^{-1}\left(C_x+c_u\right)^{-1}\left(\Lambda_{01}'+\Sigma_{01}'\right).
\end{aligned}
$$

From the above analysis, we can directly verify that the stochastic order of the residual $e_{0,t}$ is determined by $u_{0,t}$. We can show $\hat{\gamma}_h=O_p(T^{\kappa_u})$ and $\hat{\Omega}_u=O_p(M_TT^{\kappa_u})$. Eventually,

$$
Q_W=\left[\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}\hat{\Omega}_u=O_p(T^{-1-\kappa_x})O_p(M_TT^{\kappa_u})=O_p\left(\frac{M_T}{T^{1+\kappa_x-\kappa_u}}\right).
$$

We can therefore obtain the stochastic order of the Wald statistic,

$$
\begin{aligned}
W_T &= \left(R\hat{\beta}-r\right)'\left[RQ_WR'\right]^{-1}\left(R\hat{\beta}-r\right) \\
&= \begin{cases} O_p\left(\frac{T}{M_T}\right), & \text{if } \kappa_u=\kappa_x, \\ O_p\left(\frac{T^{1+\kappa_x-\kappa_u}}{M_T}\right), & \text{if } \kappa_u>\kappa_x, \\ O_p\left(\frac{T^{1-\kappa_x+\kappa_u}}{M_T}\right), & \text{otherwise.} \end{cases}
\end{aligned}
$$

(vii). (31) can be obtained by following the proof of Theorem 4.1 in Magdalinos and Phillips (2009), we therefore omit its proof. For the Wald statistic, note that demeaning X_{t-1} will not make a difference because

$$
\frac{1}{T^{2\kappa_x}}\sum_{t=1}^{T}\Phi_T^{-T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\Phi_T^{-T}
$$

$$
\begin{aligned}
&= \frac{1}{T^{2\kappa_x}}\sum_{t=1}^{T}\Phi_T^{-T}X_{t-1}X_{t-1}'\Phi_T^{-T} + \frac{1}{T^{2\kappa_x+1}}\sum_{t=1}^{T}\Phi_T^{-T}X_{t-1}\sum_{t=1}^{T}X_{t-1}'\Phi_T^{-T} \\
&= \frac{1}{T^{2\kappa_x}}\sum_{t=1}^{T}\Phi_T^{-T}X_{t-1}X_{t-1}'\Phi_T^{-T} \\
&\quad + \frac{T^{3\kappa_x}}{T^{2\kappa_x+1}}\left[T^{-\kappa_x}\sum_{t=1}^{T}\Phi_T^{-T}\left(\frac{X_{t-1}}{T^{\kappa_x/2}}\Phi_T^{-t}\right)\right]\left[T^{-\kappa_x}\sum_{t=1}^{T}\Phi_T^{-T}\left(\frac{X_{t-1}}{T^{\kappa_x/2}}\Phi_T^{-t}\right)\right]' \\
&= \frac{1}{T^{2\kappa_x}}\sum_{t=1}^{T}\Phi_T^{-T}X_{t-1}X_{t-1}'\Phi_T^{-T} + O_p(T^{\kappa_x-1}), \\
&\Rightarrow \int_0^{\infty} e^{-pC_x}Y_{C_x}Y_{C_x}'e^{-pC_x}dp = \Psi_{Y_{C_x}} \qquad (76)
\end{aligned}
$$

where we have applied Lemmas 2.2 in Phillips and Lee (2016) to obtain the second and third equality. The limiting distribution is achieved using equation (20) in Magdalinos and Phillips (2009).

Therefore, we can express

$$
\begin{aligned}
W_T &= \left(R\left(\hat{\beta}-\beta\right)\right)'\left[R\left[\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}R'\hat{\Omega}_u\right]^{-1}R\left(\left(\hat{\beta}-\beta\right)\right) \\
&= \left(RT^{\kappa_x}\Phi_T^{T}\left(\hat{\beta}-\beta\right)\right)'\left[R\left[\frac{1}{T^{2\kappa_x}}\sum_{t=1}^{T}\Phi_T^{-T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\Phi_T^{-T}\right]^{-1}R'\hat{\Omega}_u\right]^{-1} \\
&\quad \left(RT^{\kappa_x}\Phi_T^{T}\left(\hat{\beta}-\beta\right)\right) \\
&\Rightarrow \chi^2(q).
\end{aligned}
$$

(viii). Scaling (45) by $T^{-(\kappa_x+1)/2}\Phi_T^{-T}$, we have

$$
\begin{aligned}
&\left(\Phi_T\rho_T - I_k\right)T^{-(\kappa_x+1)/2}\Phi_T^{-T}\sum_{t=1}^{T}X_{t-2}^{\mu}u_{0,t-1}^{\mu} \\
&= \left(T^{-\kappa_x/2}\Phi_T^{-T}X_{T-1}^{\mu}\right)\frac{u_{0,T}^{\mu}}{T^{1/2}} - T^{-(\kappa_x+1)/2}\Phi_T^{-T+1}\sum_{t=1}^{T}X_{t-2}^{\mu}\varepsilon_{0,t-1}^{\mu} \\
&\quad -T^{-(\kappa_x+1)/2}\Phi_T^{-T}\rho_T\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}u_{0,t-1}^{\mu} - T^{-(\kappa_x+1)/2}\Phi_T^{-T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}\varepsilon_{0,t-1}^{\mu} \\
&= \left(T^{-\kappa_x/2}\Phi_T^{-T}X_{T-1}^{\mu}\right)\frac{u_{0,T}^{\mu}}{T^{1/2}} - T^{-(\kappa_x+1)/2}\Phi_T^{-T+1}O_p(T^{\kappa_x}\Phi_T^{T}) \\
&\quad -T^{-(\kappa_x+1)/2}\Phi_T^{-T}\rho_T O_p(T) - T^{-(\kappa_x+1)/2}\Phi_T^{-T}O_p(T)
\end{aligned}
$$

$$\begin{aligned}
&= \left(T^{-\kappa_x/2}\Phi_T^{-T}X^{\mu}_{T-1}\right)\frac{u^{\mu}_{0,T}}{T^{1/2}} + o_p(1) \\
&= \left(T^{-\kappa_x/2}\Phi_T^{-T}\left(X_{T-1} - \frac{1}{T}\sum_{t=1}^{T}X_{t-1}\right)\frac{u^{\mu}_{0,T}}{T^{1/2}}\right) + o_p(1) \\
&= T^{-\kappa_x/2}\Phi_T^{-T}X_{T-1}\frac{u^{\mu}_{0,T}}{T^{1/2}} + T^{-\kappa_x/2}\Phi_T^{-T}\frac{1}{T}\sum_{t=1}^{T}X_{t-1}\left(\frac{u^{\mu}_{0,T}}{T^{1/2}}\right)
\end{aligned}$$

Applying Lemma 4.1 in Phillips and Magdalinos (2009) gives

$$T^{-\kappa_x/2}\Phi_T^{-T}X_{T-1} = Y_{C_x,T} \Rightarrow Y_{C_x},$$

where $Y_{C_x,T} = T^{-\kappa_x/2}\sum_{j=1}^{T}\Phi_T^{-j}\varepsilon_{1,j}$. Since

$$\sum_{t=1}^{T}X_t = \sum_{t=1}^{T}X_{t-1} + \sum_{t=1}^{T}\varepsilon_{1,t},$$

$$\frac{C_x}{T^{\kappa_x}}\sum_{t=1}^{T}X_{t-1} = X_T - X_0 - \sum_{t=1}^{T}\varepsilon_{1,t},$$

we can write

$$\begin{aligned}
\frac{C_x}{T^{\kappa_x}}T^{-\kappa_x/2}\Phi_T^{-T}\sum_{t=1}^{T}X_{t-1} &= T^{-\kappa_x/2}\Phi_T^{-T}X_T - T^{-\kappa_x/2}\Phi_T^{-T}X_0 - T^{-\kappa_x/2}\Phi_T^{-T}\sum_{t=1}^{T}\varepsilon_{1,t} \\
&= Y_{C_x,T} + o_p(1) = O_p(1).
\end{aligned}$$

Therefore, $T^{-\kappa_x/2}\Phi_T^{-T}\frac{1}{T}\sum_{t=1}^{T}X_{t-1} = o_p(1)$, with $\frac{u^{\mu}_{0,T}}{T^{1/2}} \Rightarrow J^{\mu}_{c_u}(r)$ and $\Phi_T\rho_T - I_k = \frac{C_x}{T^{\kappa_x}} + o(1)$. We then have

$$\frac{\Phi_T^{-T}}{T^{(3\kappa_x+1)/2}}\sum_{t=1}^{T}X^{\mu}_{t-2}u^{\mu}_{0,t-1} \Rightarrow C_x^{-1}Y_{C_x}J^{\mu}_{c_u}(r). \tag{77}$$

(77) and (76) jointly give

$$\begin{aligned}
&T^{(\kappa_x-1)/2}\Phi_T^{T}\left(\hat{\beta}-\beta\right) \\
&= \left[\frac{1}{T^{2\kappa_x}}\Phi_T^{-T}\sum_{t=1}^{T}X^{\mu}_{t-1}X^{\mu\prime}_{t-1}\Phi_T^{-T}\right]^{-1}\frac{\Phi_T^{-T}}{T^{(3\kappa_x+1)/2}}\sum_{t=1}^{T}X^{\mu}_{t-2}u^{\mu}_{0,t-1} \\
&\Rightarrow \Psi^{-1}_{Y_{C_x}}C_x^{-1}Y_{C_x}J^{\mu}_{c_u}(r).
\end{aligned}$$

For the Wald statistic, note that since $\hat{\beta}-\beta$ converges in probability to zero at a rate faster than $\sqrt{T}$ and T, it can be easily shown that $\frac{1}{T}\sum_{t=1}^{T}\hat{u}_t\hat{u}_{t-j} = \frac{1}{T}\sum_{t=1}^{T}u_tu_{t-j}+o_p(1)$. Eventually, since $u_{0,t} = O_p(\sqrt{T})$, we can show that $\hat{\Omega}_u = O_p(M_TT)$. Thus,

$$\begin{aligned}
W_T &= \left(R\left(\hat{\beta}-\beta\right)\right)'\left[R\left[\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\right]^{-1}R'\hat{\Omega}_u\right]^{-1}\left(R\left(\hat{\beta}-\beta\right)\right) \\
&= O_p(\Phi_T^{-2T}T^{(1-\kappa_x)})O_p(T^{2\kappa_x}\Phi_T^{2T})O_p((M_TT)^{-1}) \\
&= O_p(T^{(1-\kappa_x)})O_p(T^{2\kappa_x})O_p((M_TT)^{-1}) \\
&= O_p\left(\frac{T^{\kappa_x}}{M_T}\right).
\end{aligned}$$

(ix). Pre-multiplying (45) by Φ_T^{-T}, we obtain

$$\begin{aligned}
&\left(\Phi_T\rho_T - I_k\right)\Phi_T^{-T}\sum_{t=1}^{T}X_{t-2}^{\mu}u_{0,t-1}^{\mu} \\
&= \Phi_T^{-T}X_{T-1}^{\mu}u_{0,T}^{\mu} - \Phi_T\Phi_T^{-T}\sum_{t=1}^{T}X_{t-2}^{\mu}\varepsilon_{0,t-1}^{\mu} - \rho_T\Phi_T^{-T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}u_{0,t-1}^{\mu} - \Phi_T^{-T}\sum_{t=1}^{T}\varepsilon_{1,t-1}^{\mu}\varepsilon_{0,t-1}^{\mu} \\
&= \Phi_T^{-T}X_{T-1}u_{0,T} - \Phi_T\Phi_T^{-T}\sum_{t=1}^{T}X_{t-2}\varepsilon_{0,t-1} + o_p(1) \\
&= \Phi_T^{-T}\left(\sum_{j=1}^{T-1}\Phi_T^{T-j}\varepsilon_{1,j}\right)\left(\sum_{t=1}^{T}\rho_T^{T-t}\varepsilon_{0,t}\right) - \Phi_T\Phi_T^{-T}\sum_{t=1}^{T}\left(\sum_{j=1}^{t-2}\Phi_T^{t-j}\varepsilon_{1,j}\right)\varepsilon_{0,t-1} + o_p(1) \quad (78)
\end{aligned}$$

Second equality above is due to the fact that demeaning the time series has no effect asymptotically; see Phillips and Lee (2016) for a discussion. Furthermore,

$$\begin{aligned}
\Phi_T^{-T}\sum_{t=1}^{T}\left(\sum_{j=1}^{t-2}\Phi_T^{t-j}\varepsilon_{1,j}\right)\varepsilon_{0,t-1} &= \Phi_T^{-T}\sum_{t=1}^{T}\left(\sum_{j=1}^{T-1}\Phi_T^{t-j}\varepsilon_{1,j}\right)\varepsilon_{0,t-1} \\
&\quad -\Phi_T^{-T}\sum_{t=1}^{T}\left(\sum_{j=t-1}^{T}\Phi_T^{t-j}\varepsilon_{1,j}\right)\varepsilon_{0,t-1},
\end{aligned}$$

since

$$\begin{aligned}
E\left\|\Phi_T^{-T}\sum_{t=1}^{T}\left(\sum_{j=t-1}^{T}\Phi_T^{t-j}\varepsilon_{1,j}\right)\varepsilon_{0,t-1}\right\| &= E\left\|\Phi_T^{-T}\sum_{t=1}^{T}\Phi_T^{t}\left(\sum_{j=t-1}^{T}\Phi_T^{-j}\varepsilon_{1,j}\right)\varepsilon_{0,t-1}\right\| \\
&\leq E\left\|\Phi_T^{-T}\right\|\left\|\sum_{t=1}^{T}\Phi_T^{t}\left(\sum_{j=t-1}^{T}\Phi_T^{-j}\varepsilon_{1,j}\right)\varepsilon_{0,t-1}\right\|
\end{aligned}$$

$$\leq \left\|\Phi_T^{-T}\right\| \left\|\sum_{t=1}^{T}\Phi_T^{t}\right\| \left\|\sum_{j=t-1}^{T}\Phi_T^{-j}\right\| E\left\|\varepsilon_{1,j}\varepsilon_{0,t-1}\right\|$$
$$= O(1)O(\Phi_T^{-T})O(T^{\kappa_x})O(T) = o(1).$$

We can express (78) as

$$\begin{aligned}
&\left(\Phi_T\rho_T - I_k\right)\Phi_T^{-T}\sum_{t=1}^{T}X_{t-2}^{\mu}u_{0,t-1}^{\mu}\\
=\;& \Phi_T^{-T}\left(\sum_{j=1}^{T-1}\Phi_T^{T-j}\varepsilon_{1,j}\right)\left(\sum_{t=1}^{T}\rho_T^{T-t}\varepsilon_{0,t}\right) - \Phi_T\Phi_T^{-T}\sum_{t=1}^{T}\left(\sum_{j=1}^{T-1}\Phi_T^{t-j}\varepsilon_{1,j}\right)\varepsilon_{0,t-1} + o_p(1) \qquad (79)\\
=\;& \left(\sum_{j=1}^{T-1}\Phi_T^{-j}\varepsilon_{1,j}\right)\left(\sum_{t=1}^{T}\rho_T^{T-t}\varepsilon_{0,t}\right) - \Phi_T\Phi_T^{-T}\sum_{t=1}^{T}\Phi_T^{t}\left(\sum_{j=1}^{T-1}\Phi_T^{-j}\varepsilon_{1,j}\right)\varepsilon_{0,t-1} + o_p(1)\\
=\;& \left[\sum_{t=1}^{T}\rho_T^{T-t}\varepsilon_{0,t} - \sum_{t=1}^{T}\Phi_T^{t+1-T}\varepsilon_{0,t-1}\right]\left(\sum_{j=1}^{T-1}\Phi_T^{-j}\varepsilon_{1,j}\right) + o_p(1)\\
=\;& \left[\sum_{t=1}^{T}\rho_T^{T-t}\varepsilon_{0,t} - \left(\sum_{t=1}^{T}\Phi_T^{t+2-T}\varepsilon_{0,t} + \Phi_T^{2-T}\varepsilon_{0,0} - \Phi_T^{2}\varepsilon_{0,T}\right)\right]\left(\sum_{j=1}^{T-1}\Phi_T^{-j}\varepsilon_{1,j}\right) + o_p(1) \qquad (80)\\
=\;& \left[\sum_{t=1}^{T}\left(\rho_T^{T-t} - \Phi_T^{t+2-T}\right)\varepsilon_{0,t}\right]\left(\sum_{j=1}^{T-1}\Phi_T^{-j}\varepsilon_{1,j}\right) - \left(\Phi_T^{2-T}\varepsilon_{0,0} - \Phi_T^{2}\varepsilon_{0,T}\right)\left(\sum_{j=1}^{T-1}\Phi_T^{-j}\varepsilon_{1,j}\right) + o_p(1).
\end{aligned}$$

Pre-multiplying (80) by $T^{-(\min\{\kappa_x,\kappa_u\}+\kappa_x)/2}$, we have

$$\begin{aligned}
\frac{\left(\Phi_T\rho_T - I_k\right)\Phi_T^{-T}}{T^{(\min\{\kappa_x,\kappa_u\}+\kappa_x)/2}}\sum_{t=1}^{T}X_{t-2}^{\mu}u_{0,t-1}^{\mu} &= \left[\sum_{t=1}^{T}\frac{\left(\rho_T^{T-t} - \Phi_T^{t+2-T}\right)\varepsilon_{0,t}}{T^{\max\{\kappa_x,\kappa_u\}/2}}\right]\left(T^{-\kappa_x/2}\sum_{j=1}^{T-1}\Phi_T^{-j}\varepsilon_{1,j}\right)\\
&\quad - \frac{\left(\Phi_T^{2-T}\varepsilon_{0,0} - \Phi_T^{2}\varepsilon_{0,T}\right)}{T^{\max\{\kappa_x,\kappa_u\}/2}}\left(T^{-\kappa_x/2}\sum_{j=1}^{T-1}\Phi_T^{-j}\varepsilon_{1,j}\right) + o_p(1)\\
&= \left[\sum_{t=1}^{T}\frac{\left(\rho_T^{T-t} - \Phi_T^{t+2-T}\right)\varepsilon_{0,t}}{T^{\max\{\kappa_x,\kappa_u\}/2}}\right]\left(T^{-\kappa_x/2}\sum_{j=1}^{T-1}\Phi_T^{-j}\varepsilon_{1,j}\right) + o_p(1)\\
&= \left[\sum_{t=1}^{T}\frac{\left(\rho_T^{T-t} - \Phi_T^{t+2-T}\right)\varepsilon_{0,t}}{T^{\max\{\kappa_x,\kappa_u\}/2}}\right]Y_{C_x,T} + o_p(1),
\end{aligned}$$

where we have applied Lemma 4.1 in Magdalinos and Phillips (2009) to obtain the last equality. We now focus on the term in the bracket. By the Beveridge-Nelson decomposition and summation by parts, we have

$$
\begin{aligned}
&\sum_{t=1}^{T} \frac{\left(\rho_T^{T-t}-\Phi_T^{t+2-T}\right)\varepsilon_{0,t}}{T^{\max\{\kappa_x,\kappa_u\}/2}} \\
=\;& \sum_{t=1}^{T} \frac{\left(\rho_T^{T-t}-\Phi_T^{t+2-T}\right)\left(C_0(1)z_{0,t}-\Delta\tilde{z}_{0,t}\right)}{T^{\max\{\kappa_x,\kappa_u\}/2}} \\
=\;& C_0(1)\sum_{t=1}^{T} \frac{\left(\rho_T^{T-t}-\Phi_T^{t+2-T}\right)z_{0,t}}{T^{\max\{\kappa_x,\kappa_u\}/2}} - \left[\frac{\left(I_k-\Phi_T^{2}\right)\tilde{z}_{0,T}}{T^{\max\{\kappa_x,\kappa_u\}/2}} - \frac{\left(\rho_T^{T-1}-\Phi_T^{3-T}\right)\tilde{z}_{0,1}}{T^{\max\{\kappa_x,\kappa_u\}/2}}\right] \\
&+\frac{1}{T^{\max\{\kappa_x,\kappa_u\}/2}}\sum_{t=1}^{T}\left(\rho_T^{T-t}\left(1-\rho_T^{-1}\right)-\Phi_T^{t+2-T}\left(I_k-\Phi_T^{-1}\right)\right)\tilde{z}_{0,t-1}.
\end{aligned}
$$

Since $\tilde{z}_{0,T}/T^{1/2}=o_p(1)$, $I_k-\Phi_T^2=\frac{2C_x}{T^{\kappa_x}}+O\left(\frac{1}{T^{2\kappa_x}}\right)$, $\frac{\left(I_k-\Phi_T^2\right)\tilde{z}_{0,T}}{T^{\max\{\kappa_x,\kappa_u\}/2}}=\tilde{z}_{0,T}O\left(\frac{1}{T^{\kappa_x+\max\{\kappa_x,\kappa_u\}/2}}\right)$ and $\kappa_x+\max\{\kappa_x,\kappa_u\}/2>1$, we have $\frac{\left(I_k-\Phi_T^2\right)\tilde{z}_{0,T}}{T^{\max\{\kappa_x,\kappa_u\}/2}}=o_p(1)$. Similarly, using the proof of Lemma 3.2 in Lui et al. (2021), we can show $\rho_T^T=\exp(\frac{c_uT}{T^{\kappa_u}})+o(1)$. Note that $\exp(c_uT^{1-\kappa_u})\to 0$ as $c_u<0$, and also that $\Phi_T^{-T}=o(1)$. Hence, we have $\frac{\left(\rho_T^{T-1}-\Phi_T^{3-T}\right)\tilde{z}_{0,1}}{T^{\max\{\kappa_x,\kappa_u\}/2}}=o_p(1)$. Moreover, as

$$
\frac{1}{T^{\max\{\kappa_x,\kappa_u\}/2}}\sum_{t=1}^{T}\left[\rho_T^{T-t}\tilde{z}_{0,t-1}\right]=O_p(1)
$$

by Lemma A.2.a in Lin and Tu (2020) and $\left(1-\rho_T^{-1}\right)=o_p(1)$, we have

$$
\frac{1}{T^{\max\{\kappa_x,\kappa_u\}/2}}\sum_{t=1}^{T}\left[\rho_T^{T-t}\tilde{z}_{0,t-1}\right]\left(1-\rho_T^{-1}\right)=o_p(1).
$$

Finally, note that $\frac{1}{T^{\kappa_x}}\sum_{t=1}^{T}\Phi_T^{-t}\tilde{z}_{0,t-1}=O_p(1)$ by Lemma 4.1 in Magdalinos and Phillips (2009). We then have

$$
\frac{\Phi_T^2}{T^{\max\{\kappa_x,\kappa_u\}/2}}\sum_{t=1}^{T}\Phi_T^{-(T+t)}\tilde{z}_{0,t-1}\left(I_k-\Phi_T^{-1}\right)=o_p(1).
$$

The above result implies that

$$
\sum_{t=1}^{T}\frac{\left(\rho_T^{T-t}-\Phi_T^{t+2-T}\right)\varepsilon_{0,t}}{T^{\max\{\kappa_x,\kappa_u\}/2}}=\frac{C_0(1)}{T^{\max\{\kappa_x,\kappa_u\}/2}}\sum_{t=1}^{T}\left(\rho_T^{T-t}-\Phi_T^{-(T-t)+2}\right)z_{0,t}+o_p(1).
$$

Following Magdalinos and Phillips (2009), we define a sequence $(K_T)_{T\in N}$ which increases to infinity such that

$$\|\Phi_T\|^{-K_T} \to 0, T^{\kappa_x} \|\Phi_T\|^{-(T-K_T)} \to 0.$$

We then have

$$\frac{C_0(1)}{T^{\max\{\kappa_x,\kappa_u\}/2}} \sum_{t=1}^{K_T} \left(\rho_T^{T-t} - \Phi_T^{-(T-t)+2}\right) z_{0,t} = o_p(1),$$

which leads to

$$\frac{(\Phi_T\rho_T - I_k)\,\Phi_T^{-T}}{T^{(\max\{\kappa_x,\kappa_u\}+\kappa_x)/2}} \sum_{t=1}^{T} X_{t-2}^{\mu} u_{0,t-1}^{\mu} = B_{K_T} Y_{C_x,T}(1+o_p(1)) + o_p(1), \tag{81}$$

where

$$\begin{aligned} B_{K_T} &= \frac{C_0(1)}{T^{\min\{\kappa_x,\kappa_u\}/2}} \sum_{t=K_T+1}^{T} \left(\rho_T^{T-t} - \Phi_T^{-(T-t)+2}\right) z_{0,t} \\ &= \frac{C_0(1)}{T^{\max\{\kappa_x,\kappa_u\}/2}} \sum_{t=1}^{T-K_T} \left(\rho_T^{T-K_T-t} - \Phi_T^{-(T-K_T-t)+2}\right) z_{0,t+K_T}. \end{aligned}$$

Let $\zeta_{T,t+K_T} = B_{K_T} Y_{C_x,T}$ and $\langle M_T \rangle_K \equiv \sum_{t=1}^{K} E_{F_T,t+K_T-1}\left(\zeta_{T,t+K_T}\zeta_{T,t+K_T}'\right)$. We have the following predictable quadratic variation of $\langle M_T \rangle_K$

$$\begin{aligned} \langle M_T \rangle_{T-K_T} &= Y_{C_x,T}^2 \Omega_{00} \frac{1}{T^{\max\{\kappa_x,\kappa_u\}}} \sum_{t=1}^{T-K_T} \left(\rho_T^{T-K_T-t} - \Phi_T^{-(T-K_T-t)+2}\right)^2 \\ &= \begin{cases} Y_{C_x,T}^2 \Omega_{00} \left[\frac{1}{-2c_u} + \frac{1}{2} C_x^2 + 2[c_u I_k + C_x]^{-1}\right], \text{ if } \kappa_u = \kappa_x, \\ Y_{C_x,T}^2 \Omega_{00} \frac{1}{T^{\kappa_u}} \left[\frac{T^{\kappa_u}}{-2c_u} + \frac{T^{\kappa_x}}{2} C_x^2 + T^{\kappa_x}[C_x]^{-1}\right], \text{ if } \kappa_u > \kappa_x, \\ Y_{C_x,T}^2 \Omega_{00} \frac{1}{T^{\kappa_x}} \left[\frac{T^{\kappa_u}}{-2c_u} + \frac{T^{\kappa_x}}{2} C_x^2 + \frac{T^{\kappa_u}}{c_u}\right], \text{ otherwise.} \end{cases} \\ &\Rightarrow \begin{cases} Y_{C_x}^2 \Omega_{00} \left[\frac{1}{-2c_u} + \frac{1}{2} C_x^2 + 2[c_u I_k + C_x]^{-1}\right], \text{ if } \kappa_u = \kappa_x, \\ \frac{1}{-2c_u} Y_{C_x}^2 \Omega_{00}, \text{ if } \kappa_u > \kappa_x, \\ \frac{1}{2} Y_{C_x}^2 C_x^2 \Omega_{00}, \text{ otherwise.} \end{cases} \end{aligned}$$

Note that $\Phi_T\rho_T - I_k = \frac{c_u I_k}{T^{\kappa_u}} + \frac{C_x}{T^{\kappa_x}} + o(1)$. Following the proof of part (iii) of Proposition A1 and Equation (22)-(26) of Magdalinos and Phillips (2009), we obtain that, as $T \to \infty$, if $\kappa_u = \kappa_x$.

$$\frac{(c_u + C_x)\,\Phi_T^{-T}}{T^{2\kappa_u}} \sum_{t=1}^{T} X_{t-2}^{\mu} u_{0,t-1}^{\mu} \Rightarrow MN\left(0, Y_{C_x}^2 \Omega_{00}\left[\frac{1}{-2c_u} + \frac{1}{2}C_x^2 + 2[c_u + C_x]^{-1}\right]\right). \tag{82}$$

If $\kappa_u > \kappa_x$,

$$\frac{C_x \Phi_T^{-T}}{T^{(\kappa_u + 3\kappa_x)/2}} \sum_{t=1}^{T} X_{t-2}^{\mu} u_{0,t-1}^{\mu} \Rightarrow MN\left(0, \frac{1}{-2c_u} Y_{C_x}^2 \Omega_{00}\right). \tag{83}$$

If $\kappa_x > \kappa_u$,

$$\frac{c_u \Phi_T^{-T}}{T^{\kappa_x + \kappa_u}} \sum_{t=1}^{T} X_{t-2}^{\mu} u_{0,t-1}^{\mu} \Rightarrow MN\left(0, \frac{1}{2} Y_{C_x}^2 C_x^2 \Omega_{00}\right). \tag{84}$$

By (82), (83), (84) and (76), if $\kappa_u = \kappa_x$, we have

$$\begin{aligned} & \Phi_T^T\left(\hat{\beta} - \beta\right) \\ \Rightarrow \quad & MN\left(\begin{array}{c} 0, \left[(c_u I_k + C_x)\,\Psi_{Y_{C_x}}\right]^{-1} Y_{C_x}^2 \\ \times\Omega_{00}\left[\frac{1}{-2c_u} + \frac{1}{2}C_x^2 + 2[c_u + C_x]^{-1}\right]\left[(c_u I_k + C_x)\,\Psi_{Y_{C_x}}\right]^{-1} \end{array}\right). \end{aligned}$$

If $\kappa_u > \kappa_x$,

$$T^{(\kappa_x - \kappa_u)/2}\Phi_T^T\left(\hat{\beta} - \beta\right) \Rightarrow MN\left(0, \frac{\left(C_x \Psi_{Y_{C_x}}\right)^{-1}}{-2c_u} Y_{C_x}^2 \left(\Psi_{Y_{C_x}} C_x\right)^{-1} \Omega_{00}\right).$$

If $\kappa_x > \kappa_u$,

$$T^{\kappa_x - \kappa_u}\Phi_T^T\left(\hat{\beta} - \beta\right) \Rightarrow MN\left(0, \frac{\Psi_{Y_{C_x}}^{-1}}{2c_u^2} Y_{C_x}^2 C_x^2 \Psi_{Y_{C_x}}^{-1} \Omega_{00}\right).$$

For the Wald statistic, following the proof of (64), we can show $\hat{\Omega} = O_p(T^{\kappa_u} M_T)$. If $\kappa_u = \kappa_x$, we have

$$\begin{aligned} \frac{M_T}{T^{\kappa_x}} W_T &= \frac{T^{\kappa_u} M_T}{T^{2\kappa_x}} W_T \\ &= \left[R\Phi_T^T\left(\hat{\beta} - \beta\right)\right]'\left[R\left(\frac{1}{T^{2\kappa_x}}\Phi_T^{-T}\sum_{t=1}^{T} X_{t-1}^{\mu} X_{t-1}^{\mu\prime}\Phi_T^{-T}\right)^{-1} R' \frac{\hat{\Omega}_u}{T^{\kappa_u} M_T}\right]^{-1} \\ &\quad \left[R\Phi_T^T\left(\hat{\beta} - \beta\right)\right] \\ &= O_p(1). \end{aligned}$$

Therefore, we have $W_T = O_p\left(\frac{T^{\kappa_x}}{M_T}\right)$. Similarly, we can also obtain $W_T = O_p\left(\frac{T^{\kappa_x}}{M_T}\right)$

when $\kappa_u > \kappa_x$ and $W_T = O_p\left(\frac{T^{\kappa_u}}{M_T}\right)$ when $\kappa_x > \kappa_u$. This completes the proof of Lemma 2.1. ■

Proof of Proposition 2.1. We first prove the result for the case where $u_{0,t}$ is an MI process with $\kappa_u > 1/3$. For the residual $e_{0,t}$, we can express

$$e_{0,t} = y_t^{\mu} - \hat{\beta}' X_{t-1}^{\mu} = u_{0,t}^{\mu} + \left(\beta - \hat{\beta}\right)' X_{t-1}^{\mu}.$$

Thus, $S_t = \sum_{i=1}^{t} u_{0,t}^{\mu} + (\beta - \hat{\beta})' \sum_{i=1}^{t} X_{t-1}^{\mu}$. For the first term in S_t, note that for $t \in \lfloor Tr \rfloor$

$$\begin{aligned} \frac{1-\rho_T}{T^{1/2}} \sum_{i=1}^{t} u_{0,t}^{\mu} &= \frac{1-\rho_T}{T^{1/2}} \sum_{i=1}^{t} \left(u_{0,t} - \frac{1}{T} \sum_{t=1}^{T} u_{0,t} \right) \\ &= \frac{1-\rho_T}{T^{1/2}} \sum_{i=1}^{t} u_{0,t} - \frac{t}{T} \frac{1-\rho_T}{T^{1/2}} \sum_{t=1}^{T} u_{0,t} \\ &\Rightarrow B_0(r) - rB_0(1). \end{aligned} \tag{85}$$

For the second term in S_t, note that

$$\begin{aligned} (\beta - \hat{\beta})' \frac{1-\rho_T}{T^{1/2}} \sum_{i=1}^{t} X_{t-1}^{\mu} &= (\beta - \hat{\beta})' (1-\rho_T) \sum_{i=1}^{t} \frac{X_{t-1}^{\mu}}{T^{1/2}} \\ &= c_u \frac{T^2}{T^{1+\kappa_u}} \left(\beta - \hat{\beta}\right)' \frac{1}{T^{3/2}} \sum_{i=1}^{t} X_{t-1}^{\mu} \\ &\Rightarrow c_u G_\beta \int_0^r J_{C_x}^{\mu}(s), \end{aligned} \tag{86}$$

where

$$G_\beta = -\left[\int_0^1 J_{C_x}^{\mu}(r) J_{C_x}^{\mu}(r)' dr\right]^{-1} \times \left(2\left(\Lambda_{01}' + \Sigma_{01}'\right) + \int_0^1 J_{C_x}^{\mu}(r) dB_0(r) - B_0(1) \int_0^1 J_{C_x}^{\mu}(r) dr\right).$$

Hence, we obtain the result from Lemma A.4(iii). The joint convergence of (85) and (86) gives

$$\frac{1-\rho_T}{T^{1/2}} S_t \Rightarrow B_0(r) - rB_0(1) + c_u G_\beta \int_0^r J_{C_x}^{\mu}(s) =: G(r).$$

From the continuous mapping theorem, we have

$$(1-\rho_T)^2 \frac{1}{T^2} \sum_{t=1}^{T} S_t^2 \Rightarrow \int_0^1 G(r) dr. \tag{87}$$

From (64), we have

$$\hat{\Omega}_u = O_p(M_T T^{\kappa_u}).$$

Eventually, we have

$$\begin{aligned} P\left(L_T > cv\right) &= P\left(\frac{1}{T^2}\sum_{t=1}^{T} S_t^2 > cv\hat{\Omega}_u\right) \\ &= P\left((1-\rho_T)^2 \frac{1}{T^2}\sum_{t=1}^{T} S_t^2 > c_u \frac{\hat{\Omega}_u}{T^{2\kappa_u}} cv\right). \end{aligned} \tag{88}$$

Since $\frac{\hat{\Omega}_u}{T^{2\kappa_u}} = O_p\left(\frac{M_T T^{\kappa_u}}{T^{2\kappa_u}}\right) = O_p\left(\frac{M_T}{T^{\kappa_u}}\right) \xrightarrow{p} 0$, from (87) and (88), we have $P\left(L_T > cv\right) \to 1$ as $T \to \infty$. If $u_{0,t}$ is a LUR process, we have

$$\begin{aligned} \frac{1}{T^{1/2}} e_{0,t} &= \frac{1}{T^{1/2}} u_{0,t}^{\mu} + \left(\beta - \hat{\beta}\right)' \frac{1}{T^{1/2}} X_{t-1}^{\mu} \\ &\Rightarrow J_{c_u}(r) + \hat{\beta}_{\infty} J_{C_x}(r), \end{aligned}$$

where $\hat{\beta}_{\infty}$ is defined in Lemma A.4(ii). In this case $e_{0,t}$ is $O_p(T^{1/2})$ and the remaining proof can be obtained following the proof of Proposition 1 in Müller (2005). This completes the proof of Proposition 2.1. ■

Before we prove Theorem 3.1, We first establish the following lemma, which provides a more refined and detailed version of Theorem 3.1.

Lemma A.5. *Under the same set of assumptions as in Theorem 3.1, as $T \to \infty$, we have the following results.*

(i). Suppose X_t is LUR or MI and $u_{0,t}$ is stationary.

If $\kappa_x > \eta$,

$$\begin{aligned} \hat{\beta}_{IVX} - \beta = O_p(T^{-\eta}) \quad &\textit{and} \quad W_{\hat{\beta}_{IVX}} = O_p(T^{1-\eta}), \\ \check{\beta}_{IVX} - \beta = O_p(T^{-(1+\min\{\kappa_x,\eta\})/2}) \quad &\textit{and} \quad W_{\check{\beta}_{IVX}} \Rightarrow \chi^2(q); \end{aligned}$$

if $\eta \geq \kappa_x$,

$$\begin{aligned} \hat{\beta}_{IVX} - \beta = O_p(T^{-\kappa_x}) \quad &\textit{and} \quad W_{\hat{\beta}_{IVX}} = O_p(T^{1-\kappa_x}), \\ \check{\beta}_{IVX} - \beta = O_p(T^{-(1+\min\{\kappa_x,\eta\})/2}) \quad &\textit{and} \quad W_{\check{\beta}_{IVX}} \Rightarrow \chi^2(q). \end{aligned}$$

(ii). If X_t is LUR or MI and $u_{0,t}$ is LUR, then

$$\hat{\beta}_{IVX} - \beta = O_p(T^{\eta-\min\{\kappa_x,\eta\}}) \textit{ and } W_{\hat{\beta}_{IVX}} = O_p(T^{2\eta-\min\{\kappa_x,\eta\}});$$

if $\kappa_x < \eta$,

$$\check{\beta}_{IVX} - \beta = O_p(1) \text{ and } W_{\check{\beta}_{IVX}} = O_p(T^{\eta - 1/3});$$

if $\kappa_x = \eta$,

$$\check{\beta}_{IVX} - \beta = O_p(1) \text{ and } W_{\check{\beta}_{IVX}} = O_p(T^{5/3 - \eta});$$

if $\kappa_x > \eta$,

$$\check{\beta}_{IVX} - \beta = O_p(T^{\eta - \kappa_x}) \text{ and } W_{\check{\beta}_{IVX}} = O_p(T^{\eta - 1/3 + 2(1-\kappa_x)}).$$

(iii). If X_t is LUR or MI and $u_{0,t}$ is MI, then

$$\begin{aligned} \hat{\beta}_{IVX} - \beta &= O_p(T^{\min\{\eta,\kappa_u\} - \min\{\kappa_x,\eta\}}), \\ \check{\beta}_{IVX} - \beta &= O_p(T^{\min\{\kappa_x,\kappa_u\} - \min\{\kappa_x,\eta\}}). \end{aligned}$$

If $\kappa_u \geq 2/3$, *then*

$$W_{\hat{\beta}_{IVX}} = \begin{cases} O_p(T^{[1+2\kappa_x+2\min\{\kappa_x,\kappa_u\}-3\eta-\kappa_u]}), & \text{if } \eta < \kappa_x \\ O_p(T^{[2\min\{\kappa_x,\kappa_u\}-\kappa_x+1-\kappa_u]}), & \text{if } \eta \geq \kappa_x \end{cases};$$

if $1/3 < \kappa_u \leq 2/3$, *then*

$$W_{\hat{\beta}_{IVX}} = \begin{cases} O_p\left(T^{1+2\min\{\eta,\kappa_u\}-\max\{\kappa_u,\eta-1/3\}-\eta}\right), & \text{if } \eta < \kappa_x \\ O_p\left(T^{2+2\min\{\eta,\kappa_u\}-\max\{\kappa_x-1/3+\eta,1+\kappa_u\}-\kappa_x}\right), & \text{if } \eta \geq \kappa_x \end{cases},$$

$$W_{\check{\beta}_{IVX}} = \begin{cases} O_p\left(T^{2\min\{\kappa_u,\eta\}-\eta+2/3-\kappa_u}\right), & \text{if } \eta < \kappa_x \\ O_p\left(T^{2\min\{\kappa_u,\eta\}-\kappa_x+2/3-\kappa_u}\right), & \text{if } \eta \geq \kappa_x \end{cases}.$$

(iv). Suppose X_t is ME with $\kappa_x > 1/2$ and $u_{0,t}$ is stationary, then

$$\begin{aligned} T^{\kappa_x}\Phi_T^T(\hat{\beta}_{IVX} - \beta) &\Rightarrow MN\left(0, \Psi_{Y_{C_x}}^{-1}\Omega_u\right), \\ T^{\kappa_x}\Phi_T^T(\check{\beta}_{IVX} - \beta) &\Rightarrow MN\left(0, \Psi_{Y_{C_x}}^{-1}\Omega_u\right), \end{aligned}$$

and

$$\begin{aligned} W_{\hat{\beta}_{IVX}} &\Rightarrow \frac{\Omega_u}{\Sigma_u}\chi^2(q), \\ W_{\check{\beta}_{IVX}} &\Rightarrow \chi^2(q). \end{aligned}$$

(v). Suppose X_t is ME with $\kappa_x > 1/2$ and $u_{0,t}$ is LUR, then

$$\begin{aligned}\hat{\beta}_{IVX} - \beta &= O_p(\Phi_T^{-T} T^{(\kappa_x+1)/2-\min\{\kappa_x,\eta\}}),\\ \check{\beta}_{IVX} - \beta &= O_p(\Phi_T^{-T} T^{(\kappa_x+1)/2-\min\{\kappa_x,\eta\}}),\end{aligned}$$

and

$$\begin{aligned}W_{\hat{\beta}_{IVX}} &= O_p(T^{3\kappa_x-2\min\{\kappa_x,\eta\}}),\\ W_{\check{\beta}_{IVX}} &= O_p(T^{3\kappa_x-2\min\{\kappa_x,\eta\}-1/3}).\end{aligned}$$

(vi). Suppose X_t is ME with $\kappa_x > 1/2$ and $u_{0,t}$ is MI, then

$$\hat{\beta}_{IVX} - \beta = \begin{cases} O_p\left(\Phi_T^{-T} T^{\kappa_x+\kappa_u-\eta}\right), & \textit{if } \kappa_x \geq \kappa_u \geq \eta,\\ O_p\left(\Phi_T^{-T} \frac{T^{(\kappa_u+3\kappa_x)/2+\min\{\kappa_u,\eta\}}}{T^{2\min\{\kappa_x,\eta\}}}\right), & \textit{otherwise.}\end{cases},$$

$$\check{\beta}_{IVX} - \beta = \begin{cases} O_p\left(\Phi_T^{-T} T^{\kappa_x+\kappa_u-\eta}\right), & \textit{if } \kappa_x \geq \kappa_u \geq \eta,\\ O_p\left(\Phi_T^{-T} \frac{T^{(\kappa_u+3\kappa_x)/2+\min\{\kappa_u,\eta\}}}{T^{2\min\{\kappa_x,\eta\}}}\right), & \textit{otherwise;}\end{cases} \tag{89}$$

if $\kappa_x \leq 1/3 + \kappa_u$,

$$W_{\hat{\beta}_{IVX}} = \begin{cases} O_p\left(T^{4\kappa_x-3\kappa_u}\right), & \textit{if } \kappa_x \geq \kappa_u \geq \eta,\\ O_p\left(T^{5\kappa_x+2\min\{\kappa_u,\eta\}-4\min\{\kappa_x,\eta\}}\right), & \textit{otherwise;}\end{cases}$$

if $\kappa_x > 1/3 + \kappa_u$, then

$$W_{\hat{\beta}_{IVX}} = \begin{cases} O_p\left(T^{3\kappa_x-1/3-2\kappa_u}\right), & \textit{if } \kappa_x > \kappa_u \geq \eta,\\ O_p\left(T^{3\kappa_x/2+\min\{\kappa_x,\eta\}-\min\{\kappa_u,\eta\}-\kappa_u/2-1/3}\right), & \textit{otherwise.}\end{cases}$$

Finally,

$$W_{\check{\beta}_{IVX}} = \begin{cases} O_p\left(T^{4\kappa_x+\kappa_u-2\eta-1/3}\right), & \textit{if } \kappa_x \geq \kappa_u \geq \eta,\\ O_p\left(T^{5\kappa_x+2\min\{\kappa_u,\eta\}-1/3-4\min\{\kappa_x,\eta\}}\right), & \textit{otherwise.}\end{cases}$$

Proof of Theorem 3.1 and Lemma A.5. Before proceeding to the proof, we first list several results provided in the Online Supplement of Kostakis et al. (2015). If $C_x \leq 0$,

$\rho_T = \rho$ and $|\rho| < 1$, we have

$$\frac{1}{T^{1+\eta}} \sum_{t=1}^{T} Z_{t-1} X^{\mu\prime}_{t-1} = \frac{1}{T^{1+\eta}} \sum_{t=1}^{T} Z_{t-1} X'_{t-1} + o_p(1), \text{ if } \eta < \kappa_x \in (0,1), \quad (90)$$

$$\frac{1}{T^{1+\kappa_x}} \sum_{t=1}^{T} Z_{t-1} X^{\mu\prime}_{t-1} = \frac{1}{T^{1+\kappa_x}} \sum_{t=1}^{T} Z_{t-1} X'_{t-1} + o_p(1), \text{if } \kappa_x \in (0,\eta], \quad (91)$$

$$\frac{1}{T^{\frac{1+\eta}{2}}} \sum_{t=1}^{T} Z_{t-1} u^{\mu}_{0,t} = \frac{1}{T^{\frac{1+\eta}{2}}} \sum_{t=1}^{T} Z_{t-1} u_{0,t} + o_p(1), \text{if } \eta < \kappa_x \in (0,1], \quad (92)$$

$$\frac{1}{T^{\frac{1+\kappa_x}{2}}} \sum_{t=1}^{T} Z_{t-1} u^{\mu}_{0,t} = \frac{1}{T^{\frac{1+\kappa_x}{2}}} \sum_{t=1}^{T} Z_{t-1} u_{0,t} + o_p(1), \text{if } \kappa_x \in (0,\eta]. \quad (93)$$

While in Kostakis et al. (2015), $u_{0,t}$ is an mds rather than a linear process, the above four results are still valid since $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_{0,t} = O_p(1)$ is sufficient to establish these results.

For the sake of the clarity of the presentation, we prove the claims for the IVX estimator and the re-centered IVX estimator separately in two parts. Now, We first prove the results for the IVX estimator.

(i). By construction,

$$\begin{aligned} Z_t &= \sum_{j=1}^{t} \Upsilon_T^{t-j} \Delta X_j \\ &= \sum_{j=1}^{t} \Upsilon_T^{t-j} \left((\Phi_T - 1) X_{j-1} + \varepsilon_{1,j} \right) \\ &= Z_t^* + (\Phi_T - 1) \Psi_{T,t}, \end{aligned} \quad (94)$$

where $Z_t^* = \sum_{j=1}^{t} \Upsilon_T^{t-j} \varepsilon_{1,j}$, $Z_0 = 0$, and $\Psi_{T,t} = \sum_{j=1}^{t} \Upsilon_T^{t-j} X_{j-1}$. Suppose that $2/3 < \eta < \min\{\kappa_x, 1\}$. Applying the analogous argument of Lemma 3.1(i) in Phillips and Magdalinos (2009) and (92), we have

$$\begin{aligned} \frac{1}{T^{\frac{1+\eta}{2}}} \sum_{t=2}^{T} Z_{t-1} u^{\mu}_{0,t} &= \frac{1}{T^{\frac{1+\eta}{2}}} \sum_{t=2}^{T} Z_{t-1} u_{0,t} + o_p(1) \\ &= \frac{1}{T^{(1+\eta)/2}} \sum_{t=2}^{T} Z^*_{t-1} u_{0,t} + o_p(1). \end{aligned}$$

Applying a similar argument to (69), and letting

$$\psi_{IVX}=\frac{C_z}{T^{\eta}}G_{0,T}+\Lambda_{01}, G_{0,T}=\sum_{j=0}^{\infty}\tilde{c}_{0,j+1}\Sigma_{00}C_1(1)'\Upsilon_T^{j+1},$$

we have

$$\begin{aligned}
&\frac{1}{T^{(1+\eta)/2}}\sum_{t=2}^{T}Z_{t-1}u_{0,t}^{\mu}-\frac{T}{T^{(1+\eta)/2}}\psi_{IVX}\\
=\ &\frac{1}{T^{(1+\eta)/2}}\sum_{t=2}^{T}Z_{t-1}u_{0,t}-\frac{T}{T^{(1+\eta)/2}}\psi_{IVX}+o_p(1)\\
=\ &\frac{\Psi_0(1)}{T^{(1+\eta)/2}}\sum_{t=2}^{T}Z_{t-1}^{*}z_{0,t}+\frac{1}{T^{(1+\eta)/2}}\sum_{t=2}^{T}\Delta Z_{t-1}^{*}\tilde{z}_{0,t}-\frac{T}{T^{(1+\eta)/2}}\psi_{IVX}+o_p(1)\\
=\ &\frac{\Psi_0(1)}{T^{(1+\eta)/2}}\sum_{t=2}^{T}Z_{t-1}^{*}z_{0,t}+\frac{C_x}{T^{\eta}}\frac{1}{T^{(1+\eta)/2}}\sum_{t=2}^{T}\left(Z_{t-2}^{*}\tilde{z}_{0,t}-G_{0,T}\right)\\
&+\frac{1}{T^{(1+\eta)/2}}\sum_{t=2}^{T}\left(\tilde{z}_{0,t}\varepsilon_{1,t}-\Lambda_{01}\right)+o_p(1)\\
\Rightarrow\ &N\left(0,V_{zz}\Omega_{00}\right),
\end{aligned}$$

where $V_{zz}=\int_0^{\infty}e^{rC_z}\Omega_{zz}e^{rC_z}dr$. Proposition A1 in Phillips and Magdalinos (2009) is used to obtain the last result.

Note that since $\psi_{IVX}=\Lambda_{01}+O\left(T^{-\eta}\right)$, we have

$$\frac{1}{T^{(1+\eta)/2}}\sum_{t=1}^{T}\left(Z_{t-1}u_{0,t}-\Lambda_{01}\right)\Rightarrow N(0,V_{zz}\Omega_{00}). \tag{95}$$

Thus, we can also obtain that

$$\frac{1}{T}\sum_{t=1}^{T}Z_{t-1}u_{0,t}=\Lambda_{01}+O_p(T^{(\eta-1)/2}).$$

For $\sum_{t=1}^{T}Z_{t-1}X_{t-1}^{\mu\prime}$, Equation (26) and (27) in the Online Supplement of Kostakis et al. (2015) give $\frac{1}{T^{1+\eta}}\sum_{t=1}^{T}Z_{t-1}X_{t-1}^{\mu\prime}=O_p(1)$. Therefore, we can obtain

$$\begin{aligned}
T^{\eta}\left(\hat{\beta}_{IVX}-\beta\right) &= \left[\frac{1}{T^{1+\eta}}\sum_{t=1}^{T}Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{T}\sum_{t=1}^{T}Z_{t-1}u_{0,t}^{\mu}\\
&= O_p(1).
\end{aligned}$$

And for the Wald statistic, since

$$\frac{1}{T^{1+\eta}}\sum_{t=1}^{T} Z_{t-1}Z_{t-1}' \xrightarrow{p} V_{zz}, \hat{\Sigma}_{00} \xrightarrow{p} \Sigma_{00}, \tag{96}$$

and noting that $TZ_{t-1}Z_{t-1}\hat{\Omega}_{FM}$ is asymptotic dominated by $\left[\sum_{t=1}^{T} Z_{t-1}Z_{t-1}'\right]\hat{\Sigma}_{00}$, we have $\frac{M}{T^{1+\eta}} \xrightarrow{p} V_{zz}\Sigma_{00}$. Therefore,

$$\begin{aligned} Q_{IVX} &= R\left[\sum_{t=1}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1} M \left[\sum_{t=1}^{T} X_{t-1}^{\mu}Z_{t-1}'\right]^{-1} R' \\ &= O_p\left(T^{-2-2\eta}\right) O_p(T^{1+\eta}) \\ &= O_p(T^{-1-\eta}) \end{aligned}$$

Consequently,

$$\begin{aligned} W_{\hat{\beta}_{IVX}} &= \left(R\hat{\beta}_{IVX} - r\right)' Q_{IVX}^{-1} \left(R\hat{\beta}_{IVX} - r\right) \\ &= \left(T^{-2\eta}\right) O_p(T^{1+\eta}) \\ &= O_p(T^{1-\eta}). \end{aligned}$$

If $\kappa_x \in (1/3, \eta)$, from (93) and (67) and Lemma 3.5 in Phillips and Magdalinos (2009), we have

$$\frac{1}{T}\sum_{t=1}^{T} Z_{t-1}u_{0,t}^{\mu} = \frac{1}{T}\sum_{t=1}^{T} X_{t-1}u_{0,t} + o_p(1) \xrightarrow{p} \Lambda_{\tilde{z}\varepsilon}. \tag{97}$$

(91), (68) and Lemma 3.5 in Phillips and Magdalinos (2009) gives

$$\begin{aligned} \frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} Z_{t-1}X_{t-1}^{\mu\prime} &= \frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} Z_{t-1}X_{t-1}' + o_p(1) \\ &= \frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} X_{t-1}X_{t-1}' + o_p(1) \\ &\xrightarrow{p} V_{xx}. \end{aligned} \tag{98}$$

Therefore, we have

$$T^{\kappa_x}\left(\hat{\beta}_{IVX} - \beta\right) = \left[\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1} \frac{1}{T}\sum_{t=1}^{T} Z_{t-1}u_{0,t}^{\mu}$$

$$= O_p(1).$$

For the Wald statistic, first note that

$$\frac{1}{T^{1+\kappa_x}} \sum_{t=1}^{T} Z_{t-1} Z'_{t-1} \xrightarrow{p} V_{xx}. \tag{99}$$

Also note that $\hat{\Sigma}_{00} \xrightarrow{p} \Sigma_{00}$, $\frac{M}{T^{1+\kappa_x}} \xrightarrow{p} V_{zz}\Sigma_{00}$. Therefore,

$$Q_{IVX} = O_p\left(T^{-2-2\kappa_x}\right) O_p(T^{1+\kappa_x}) = O_p(T^{-1-\kappa_x})$$

and as a result $W_{\hat{\beta}_{IVX}} = O_p(T^{1-\kappa_x})$.

For $\kappa_x = \eta$, (93) and Lemma 3.6 in Phillips and Magdalinos (2009) give

$$\frac{1}{T} \sum_{t=1}^{T} Z_{t-1} u^{\mu}_{0,t} \xrightarrow{p} \Lambda_{01}, \sum_{t=1}^{T} Z_{t-1} X^{\mu\prime}_{t-1} = O_p(T^{1+\kappa_x}) \tag{100}$$

and

$$\sum_{t=1}^{T} Z_{t-1} Z'_{t-1} = O_p(T^{1+\kappa_x}). \tag{101}$$

Following the previous analysis, we obtain $T^{\kappa_x}\left(\hat{\beta}_{IVX} - \beta\right) = O_p(1)$ and $W_{\hat{\beta}_{IVX}} = O_p(T^{1-\kappa_x})$.

(ii). We now consider the case that $\eta \leq \kappa_x$. Pre-multiplying Z_{t-1} by $u^{\mu}_{0,t}$ and summing over $t = 1, ..., T$, we have

$$\sum_{t=1}^{T} u^{\mu}_{0,t} Z_{t-1} = \sum_{t=1}^{T} u^{\mu}_{0,t} Z^*_{t-1} + (\Phi_T - I_k) \sum_{t=1}^{T} u^{\mu}_{0,t} \Psi_{T,t-1}. \tag{102}$$

Note that

$$\begin{aligned} u^{\mu}_{0,t} Z^*_{t-1} &= \left(\rho_T u^{\mu}_{0,t-1} + \varepsilon^{\mu}_{0,t}\right)\left(\Upsilon_T Z^*_{t-2} + \varepsilon_{1,t-1}\right) \\ &= \rho_T \Upsilon_T u^{\mu}_{0,t-1} Z^*_{t-2} + \rho_T u^{\mu}_{0,t-1} \varepsilon_{1,t-1} + \Upsilon_T Z^*_{t-2} \varepsilon^{\mu}_{0,t} + \varepsilon^{\mu}_{0,t} \varepsilon_{1,t-1}. \end{aligned}$$

Taking the average of the above equation over $t = 1, ..., T$, we have

$$\begin{aligned} &(\rho_T \Upsilon_T - I_k) \frac{1}{T} \sum_{t=1}^{T} u^{\mu}_{0,t-1} Z^*_{t-2} \qquad (103) \\ = \;& \frac{1}{T}\left(u^{\mu}_{0,T} Z^*_{T-1} - u^{\mu}_{0,1} Z^*_0\right) - \rho_T \frac{1}{T} \sum_{t=1}^{T} u^{\mu}_{0,t-1} \varepsilon_{1,t-1} - \Upsilon_T \frac{1}{T} \sum_{t=1}^{T} Z^*_{t-2} \varepsilon^{\mu}_{0,t} - \frac{1}{T} \sum_{t=1}^{T} \varepsilon^{\mu}_{0,t} \varepsilon_{1,t-1}. \end{aligned}$$

We rewrite (103) as

$$
\begin{aligned}
& (\rho_T \Upsilon_T - I_k) \frac{1}{T} \sum_{t=1}^{T} u_{0,t-1}^{\mu} Z_{t-2}^{*} \\
= \quad & -\rho_T \frac{1}{T} \sum_{t=1}^{T} u_{0,t-1}^{\mu} \varepsilon_{1,t-1} - \Upsilon_T \frac{1}{T} \sum_{t=1}^{T} Z_{t-2}^{*} \varepsilon_{0,t}^{\mu} - \frac{1}{T} \sum_{t=1}^{T} \varepsilon_{0,t}^{\mu} \varepsilon_{1,t-1} + o_p(1),
\end{aligned}
$$

where we the last equality is due to $Z_0^* = 0$, $Z_T^* = \sum_{j=1}^{T} \Phi_T^{T-j} u_{x,j} = O_p(T^{\eta/2})$, and $u_{0,T}^{\mu} = O_p(T^{1/2})$. Therefore, $\frac{1}{T}\left(u_{0,T}^{\mu} Z_{T-1}^{*} - u_{0,1}^{\mu} Z_0^{*}\right) \xrightarrow{p} 0$.

From (92) and using the identical argument that obtains (67), we have $\frac{1}{T}\sum_{t=1}^{T} Z_{t-2}^{*}\varepsilon_{0,t}^{\mu} \xrightarrow{p} \Lambda_{\tilde{z}\varepsilon} - \gamma_1$, where $\gamma_1 \equiv E[\varepsilon_{0,t}\varepsilon_{1,t-1}]$ and by ergodic theorem $\frac{1}{T}\sum_{t=1}^{T} \varepsilon_{0,t}^{\mu}\varepsilon_{1,t-1} \xrightarrow{as} \gamma_1$. Since $\rho_T \to 1$ and

$$
\begin{aligned}
\frac{1}{T} \sum_{t=1}^{T} u_{0,t-1}^{\mu} \varepsilon_{1,t-1} &= \frac{1}{T} \sum_{t=1}^{T} \left(\rho_T u_{0,t-2}^{\mu} + \varepsilon_{0,t-1}^{\mu}\right) \varepsilon_{1,t-1} \\
&= \rho_T \frac{1}{T} \sum_{t=1}^{T} u_{0,t-2}^{\mu} \varepsilon_{1,t-1} + \frac{1}{T} \sum_{t=1}^{T} \varepsilon_{0,t-1}^{\mu} \varepsilon_{1,t-1} \\
&\Rightarrow \int_0^1 J_{c_u}^{\mu}(r) dB_1(r) + \gamma_1.
\end{aligned}
$$

Eventually, we have

$$
(I_k - \rho_T \Upsilon_T) \frac{1}{T} \sum_{t=1}^{T} u_{0,t-1}^{\mu} Z_{t-2}^{*} \Rightarrow \int_0^1 J_{c_u}^{\mu}(r) dB_1(r) + \Lambda_{\tilde{z}\varepsilon} - \gamma_1.
$$

Since

$$
\begin{aligned}
1 - \rho_T \Upsilon_T &= 1 - \left(1 + \frac{C_z}{T^{\eta}} + \frac{c_u}{T} + \frac{C_z c_u}{T^{\eta+1}}\right) \\
&= -\frac{C_z}{T^{\eta}} + O(T^{-1}),
\end{aligned}
$$

we have

$$
\frac{1}{T^{1+\eta}} \sum_{t=1}^{T} u_{0,t}^{\mu} Z_{t-1}^{*} \Rightarrow -C_z^{-1} \left[\int_0^1 J_{c_u}^{\mu}(r) dB_1(r) + \Lambda_{\tilde{z}\varepsilon} - \gamma_1\right]. \qquad (104)
$$

For the second term in (102), note that

$$
\begin{aligned}
u_{0,t}^{\mu} \Psi_{T,t-1} &= \left(\rho_T u_{0,t-1}^{\mu} + \varepsilon_{0,t}^{\mu}\right) \left(\Upsilon_T \Psi_{T,t-2} + X_{t-1}\right) \\
&= \rho_T \Upsilon_T u_{0,t-1}^{\mu} \Psi_{T,t-2} + \rho_T u_{0,t-1}^{\mu} X_{t-1} + \Upsilon_T \Psi_{T,t-2} \varepsilon_{0,t}^{\mu} + X_{t-1} \varepsilon_{0,t}^{\mu}.
\end{aligned}
$$

The sum of the above expression over $t = 1, ..., T$ and dividing it by $T^{1+\kappa_x}$ lead to

$$\begin{aligned}
& (1-\rho_T\Upsilon_T)\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}u^{\mu}_{0,t-1}\Psi_{T,t-2} \\
= & \frac{1}{T^{1+\kappa_x}}\left(u^{\mu}_{0,1}\Psi_{T,0}-u^{\mu}_{0,T}\Psi_{T,T-1}\right)+\rho_T\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}u^{\mu}_{0,t-1}X_{t-1} \\
& +\Upsilon_T\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}\Psi_{T,t-2}\varepsilon^{\mu}_{0,t}+\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-1}\varepsilon^{\mu}_{0,t}. \qquad (105)
\end{aligned}$$

Applying Proposition A2 in Phillips and Magdalinos (2009) gives

$$\sup_{1\le t\le T} E||\Psi_{T,t}||^2 = O(T^{\max\{\kappa_x,\eta\}+2\min\{\kappa_x,\eta\}}).$$

This implies

$$\begin{aligned}
\sup_{1\le t\le T} E\left\|\frac{\Psi_{T,t}}{T^{1/2+\kappa_x}}\right\|^2 &= O(T^{\max\{\kappa_x,\eta\}+2\min\{\kappa_x,\eta\}-1-2\kappa_x}) \\
&= \begin{cases} O(T^{2\eta-1-\kappa_x}), \text{ if } \eta<\kappa_x, \\ O_p(T^{\eta-1}), \text{ if } \eta\ge\kappa_x. \end{cases} \\
&= o(1),
\end{aligned}$$

$$\frac{1}{T^{1+\kappa_x}}u^{\mu}_{0,t}\Psi_{T,t-1} = \frac{u^{\mu}_{0,t}}{T^{1/2}}\frac{\Psi_{T,t-1}}{T^{1/2+\kappa_x}} = O_p(1)o_p(1) = o_p(1),$$

$$\begin{aligned}
E\left\|\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}\Psi_{T,t-2}\varepsilon^{\mu}_{0,t}\right\|^2 &\le \frac{1}{T^{2+2\kappa_x}}\left(E\sum_{t=1}^{T}\|\Psi_{T,t-2}\|^2\right)^{1/2}\left(E\sum_{t=1}^{T}\varepsilon^{\mu 2}_{0,t}\right)^{1/2} \\
&= \frac{1}{T^{2+2\kappa_x}}O(T^{\frac{\max\{\kappa_x,\eta\}+2\min\{\kappa_x,\eta\}+1}{2}})O\left(T^{1/2}\right) \\
&= o(1), \qquad (106)
\end{aligned}$$

and following (37) and (67), one can deduce that $\sum_{t=1}^{T}X_{t-1}\varepsilon^{\mu}_{0,t} = O_p(T)$ when $\kappa_x \in (0,1]$, thus

$$\frac{1}{T^{1+\kappa_x}}\sum_{t=1}^{T}X_{t-1}\varepsilon^{\mu}_{0,t} = o_p(1) \text{ for } \kappa_x \in (0,1]. \qquad (107)$$

Hence, following (37) and (71), we can deduce that

$$
\begin{aligned}
\left(1-\rho_T \Upsilon_T\right) \frac{1}{T^{1+\kappa_x}} \sum_{t=1}^{T} u_{0,t-1}^{\mu} \Psi_{T,t-2} =& \rho_T \frac{1}{T^{1+\kappa_x}} \sum_{t=1}^{T} u_{0,t-1}^{\mu} X_{t-1} + o_p(1) \\
& \Rightarrow \begin{cases} \int_0^1 J_{c_u}^{\mu}(r) J_{C_x}(r) dr, \text{ if } \kappa_x = 1, \\ C_x^{-1}\left[\int_0^1 J_{c_u}^{\mu}(r) dB_1(r) + 2\Lambda_{01} + \Sigma_{01}\right]', \text{ if } \kappa_x \in (0,1) \end{cases}
\end{aligned}
\tag{108}
$$

Finally, since $1-\rho_T \Upsilon_T = -\frac{C_z}{T^{\eta}} + O(T^{-1})$, we have

$$
\frac{1}{T^{1+\kappa_x+\eta}} \sum_{t=1}^{T} u_{0,t-1}^{\mu} \Psi_{T,t-2} \Rightarrow \begin{cases} -C_z^{-1} \int_0^1 J_{c_u}^{\mu}(r) J_{C_x}(r) dr, \text{ if } \kappa_x = 1, \\ -C_z^{-1} C_x^{-1}\left[\int_0^1 J_{c_u}^{\mu}(r) dB_1(r) + 2\Lambda_{01} + \Sigma_{01}\right]', \text{ if } \kappa_x \in (0,1) \end{cases}
\tag{109}
$$

By (104) and (109) we have $\sum_{t=1}^{T} Z_{t-1} u_{0,t}^{\mu} = O_p(T^{1+\eta})$ because

$$
\begin{aligned}
& \frac{1}{T^{1+\eta}} \sum_{t=1}^{T} Z_{t-1} u_{0,t}^{\mu} \\
=& \frac{1}{T^{1+\eta}} \sum_{t=1}^{T} Z_{t-1}^{*} u_{0,t}^{\mu} + \left(\Phi_T - 1\right) \frac{1}{T^{1+\eta}} \sum_{t=1}^{T} \Psi_{T,t-1} u_{0,t}^{\mu} \\
=& \frac{1}{T^{1+\eta}} \sum_{t=1}^{T} Z_{t-1}^{*} u_{0,t}^{\mu} + \frac{C_x}{T^{1+\eta+\kappa_x}} \sum_{t=1}^{T} \Psi_{T,t-1} u_{0,t}^{\mu} \\
\Rightarrow & \begin{cases} -C_z^{-1}\left[\int_0^1 J_{c_u}^{\mu}(r) dJ_{c_u}(r) + \Lambda_{\tilde{z}\varepsilon} - \gamma_1\right], \text{ if } \kappa_x = 1, \\ -C_z^{-1}\left[\int_0^1 J_{c_u}^{\mu}(r) dB_1(r) + \Lambda_{\tilde{z}\varepsilon} - \gamma_1 + C_x^{-1}\left[\int_0^1 J_{c_u}^{\mu}(r) dB_1(r) + 2\Lambda_{01} + \Sigma_{01}\right]'\right], \text{ if } \kappa_x \in (0,1) \end{cases}
\end{aligned}
\tag{110}
$$

For the term $\sum_{t=1}^{T} X_{t-1}^{\mu} Z_{t-1}'$, by Equation (25) in the Online Supplement of Kostakis et al. (2015) and Lemma 3.1, 3.5 and 3.6 in Magdalinos and Phillips (2009), we have

$$
\frac{1}{T^{1+\min\{\kappa_x,\eta\}}} \sum_{t=1}^{T} X_{t-1}^{\mu} Z_{t-1}' = O_p(1),
\tag{111}
$$

Combining (110) and (111), we have

$$
\hat{\beta}_{IVX} - \beta = \left[\sum_{t=1}^{T} X_{t-1}^{\mu} Z_{t-1}'\right]^{-1} \sum_{t=1}^{T} Z_{t-1} u_{0,t}^{\mu} = O_p(T^{\eta-\min\{\kappa_x,\eta\}}).
$$

We proceed to analyze the Wald statistic and focus on the stochastic orders of $\hat{\Omega}_{FM}$, M and Q_{IVX}. For $\hat{\Omega}_{FM}$, note that from (42), we have $\hat{\Sigma}_{00} = O_p(T)$, $\hat{\Omega}_{11} \xrightarrow{p} \Omega_{11}$, for $\hat{\Omega}_{01}$, let $\hat{\gamma}_{01,h} = \frac{1}{T}\sum_{t=h}^{T} e_{0,t}e_{1,t-h}$, we have

$$\begin{aligned}
\hat{\gamma}_{01,h} &= \frac{1}{T}\sum_{t=h+1}^{T}\left(u_{0,t}^{\mu} + \left(\beta - \hat{\beta}\right)' X_{t-1}^{\mu}\right)\left(\varepsilon_{0,t-h}^{\mu} + \left(\Phi_T - \hat{R}_T\right)' X_{t-h-1}^{\mu}\right) \\
&= \frac{1}{T}\sum_{t=h+1}^{T} u_{0,t}^{\mu}\varepsilon_{0,t-h}^{\mu} + \frac{1}{T}\sum_{t=h+1}^{T} u_{0,t}^{\mu}X_{t-h-1}^{\mu\prime}\left(\Phi_T - \hat{R}_T\right) \qquad (112) \\
&\quad + \left(\beta - \hat{\beta}\right)'\frac{1}{T}\sum_{t=h+1}^{T} X_{t-1}^{\mu}\varepsilon_{0,t-h}^{\mu} + \left(\beta - \hat{\beta}\right)'\frac{1}{T}\sum_{t=h+1}^{T} X_{t-1}^{\mu}X_{t-h-1}^{\mu}\left(\Phi_T - \hat{R}_T\right) \\
&= O_p(1) + T^{-1}O_p(T^2)O_p(T^{-1}) + O_p(1)T^{-1}O_p(T) \\
&\quad + O_p(1)T^{-1}O_p(T^2)O_p(T^{-1}) \\
&= O_p(1).
\end{aligned}$$

Thus,

$$\hat{\Omega}_{01} = O_p(M_T) \qquad (113)$$

and

$$\hat{\Omega}_{FM} = O_p(T) + O_p(M_T^2) = O_p(T). \qquad (114)$$

For M, from (96), (100) and (42), we have

$$\sum_{t=1}^{T} Z_{t-1}Z_{t-1}' = O_p(T^{1+\min\{\kappa_x,\eta\}}), \hat{\Sigma}_{00} = O_p(T). \qquad (115)$$

Therefore,

$$\left[\sum_{t=1}^{T} Z_{t-1}Z_{t-1}'\right]\hat{\Sigma}_{00} = O_p(T^{2+\min\{\kappa_x,\eta\}}). \qquad (116)$$

For the orders of $T\bar{Z}_{t-1}\bar{Z}_{t-1}'$, the proof of Theorem 1 in the Online Appendix of Kostakis et al. (2015) shows that

$$T\bar{Z}_{t-1}\bar{Z}_{t-1}' = \begin{cases} O_p(T^{2\eta}), \text{ if } \eta < \kappa_x, \\ O_p(T^{2\kappa_x+\eta-1}), \text{ if } \eta \geq \kappa_x. \end{cases} \qquad (117)$$

Thus, (114) and (117) give

$$T\bar{Z}_{t-1}\bar{Z}_{t-1}'\hat{\Omega}_{FM} = \begin{cases} O_p(T^{2\eta+1}), \text{ if } \eta < \kappa_x, \\ O_p(T^{2\kappa_x+\eta}), \text{ if } \eta \geq \kappa_x. \end{cases}$$

and

$$
\begin{aligned}
M &= \begin{cases} O_p(T^{2+\eta}) + O_p(T^{2\eta+1}), & \text{if } \eta < \kappa_x, \\ O_p(T^{2+\kappa_x}) + O_p(T^{2\kappa_x+\eta}), & \text{if } \eta \geq \kappa_x. \end{cases} \\
&= O_p(T^{2+\min\{\kappa_x,\eta\}}).
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
Q_{IVX} &= R\left[\sum_{t=2}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1} M \left[\sum_{t=2}^{T} X_{t-1}^{\mu} Z_{t-1}'\right]^{-1} R' \\
&= \left[O_p(T^{1+\min\{\kappa_x,\eta\}})\right]^{-2} O_p(T^{2+\min\{\kappa_x,\eta\}}) \\
&= O_p(T^{-\min\{\kappa_x,\eta\}}). \qquad (118)
\end{aligned}
$$

and consequently

$$
\begin{aligned}
W_{\hat{\beta}_{IVX}} &= \left(R\hat{\beta}_{IVX} - r\right)' Q_{IVX}^{-1} \left(R\hat{\beta}_{IVX} - r\right) \\
&= O_p(T^{\eta-\min\{\kappa_x,\eta\}}) O_p(T^{\min\{\kappa_x,\eta\}}) O_p(T^{\eta-\min\{\kappa_x,\eta\}}) \\
&= O_p(T^{2\eta-\min\{\kappa_x,\eta\}}).
\end{aligned}
$$

(iii). Following (72), we have the following limit,

$$
(\rho_T \Upsilon_T - I_k) \frac{1}{T} \sum_{t=2}^{T} u_{0,t-1}^{\mu} Z_{t-2}^{*} \xrightarrow{p} -2(\Lambda_{01}' + \Sigma_{01}').
$$

Since $\rho_T \Upsilon_T - I_k = \left(1 + \frac{c_u}{T^{\kappa_u}}\right)\left(1 + \frac{C_z}{T^{\eta}}\right) = \frac{C_z}{T^{\eta}} + \frac{c_u}{T^{\kappa_u}} + O\left(\frac{1}{T^{\kappa_u+\eta}}\right)$, we have

$$
\left(\frac{C_z}{T^{\eta}} + \frac{c_u}{T^{\kappa_u}}\right) \frac{1}{T} \sum_{t=2}^{T} u_{0,t} Z_{t-1}^{*} \xrightarrow{p} -2(\Lambda_{01}' + \Sigma_{01}'). \qquad (119)
$$

Similar to (110), we can write

$$
\begin{aligned}
& \frac{1}{T^{1+\min\{\eta,\kappa_u\}}} \sum_{t=2}^{T} u_{0,t}^{\mu} Z_{t-1} \\
= & \frac{1}{T^{1+\min\{\eta,\kappa_u\}}} \sum_{t=2}^{T} u_{0,t}^{\mu} Z_{t-1}^{*} + \frac{C_x}{T^{1+\min\{\eta,\kappa_u\}+\kappa_x}} \sum_{t=2}^{T} u_{0,t}^{\mu} \Psi_{T,t-1} + o_p(1). \qquad (120)
\end{aligned}
$$

Similar to (108), we can show

$$\left(1-\rho_T\Upsilon_T\right)\frac{1}{T^{1+\kappa_x}}\sum_{t=2}^{T}u_{0,t-1}^{\mu}\Psi_{T,t-2}=\rho_T\frac{1}{T^{1+\kappa_x}}\sum_{t=2}^{T}u_{0,t-1}^{\mu}X_{t-1}+o_p(1).$$

Similar to (55), (73), (74) and (75), we can show

$$\frac{1}{T^{1+\kappa_x}}\sum_{t=2}^{T}u_{0,t-1}^{\mu}X_{t-1}=\begin{cases}o_p(1)\text{ if }\kappa_x=1,\\ O_p(1),\ \text{if }\kappa_x\in(1/3,1)\text{ and }\kappa_x\leq\kappa_u\\ o_p(1),\ \text{if }1>\kappa_x>\kappa_u.\end{cases}\tag{121}$$

Since κ_u and $\eta>0$, from (119), (120) and (121), we can obtain

$$\begin{aligned}\frac{1}{T^{1+\min\{\eta,\kappa_u\}}}\sum_{t=2}^{T}u_{0,t}^{\mu}Z_{t-1} &= \frac{1}{T^{1+\min\{\eta,\kappa_u\}}}\sum_{t=2}^{T}u_{0,t}Z_{t-1}^{*}+o_p(1)\\ &\xrightarrow{p}\begin{cases}-2C_z^{-1}\left(\Lambda_{01}'+\Sigma_{01}'\right)\text{ if }\kappa_x<\kappa_u,\\ -2c_u^{-1}\left(\Lambda_{01}'+\Sigma_{01}'\right),\ \text{if }\kappa_x>\kappa_u,\\ -2\left(C_z+c_uI_k\right)^{-1}\left(\Lambda_{01}'+\Sigma_{01}'\right),\ \text{if }\kappa_x=\kappa_u.\end{cases}\end{aligned}\tag{122}$$

From (111) and (122), we can conclude that

$$\begin{aligned}&T^{\min\{\kappa_x,\eta\}-\min\{\eta,\kappa_u\}}\left(\hat{\beta}_{IVX}-\beta\right)\\ =&\left[\frac{1}{T^{1+\min\{\kappa_x,\eta\}}}\sum_{t=2}^{T}Z_{t-1}X_{t-1}^{\mu}\right]^{-1}\left[\frac{1}{T^{1+\min\{\eta,\kappa_u\}}}\sum_{t=2}^{T}Z_{t-1}u_{0,t}^{\mu}\right]\\ =&\ O_p(1).\end{aligned}\tag{123}$$

We now proceed to analyze the Wald statistic via the stochastic orders of $\hat{\Omega}_{FM}$, M and Q_{IVX}. Note that

$$\begin{aligned}Q_{IVX} &= R\left[\sum_{t=2}^{T}Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1}M\left[\sum_{t=2}^{T}X_{t-1}^{\mu}Z_{t-1}'\right]^{-1}R',\\ M &= \left[\sum_{t=2}^{T}Z_{t-1}Z_{t-1}'\right]\hat{\Sigma}_{00}-T\bar{Z}_{t-1}\bar{Z}_{t-1}'\hat{\Omega}_{FM},\\ \hat{\Omega}_{FM} &= \hat{\Sigma}_{00}-\hat{\Omega}_{01}\hat{\Omega}_{11}^{-1}\hat{\Omega}_{01}',\ \bar{Z}_{t-1}=\frac{1}{T}\sum_{t=2}^{T}Z_{t-1}.\end{aligned}$$

From (63) and (115), $\hat{\Sigma}_{00} = O_p(T^{\kappa_u})$, $\sum_{t=1}^{T} Z_{t-1}Z_{t-1}' = O_p(T^{1+\min\{\kappa_x,\eta\}})$. Thus,

$$\left[\sum_{t=2}^{T} Z_{t-1}Z_{t-1}'\right]\hat{\Sigma}_{00} = O_p(T^{1+\min\{\kappa_x,\eta\}+\kappa_u}). \tag{124}$$

As in (112),

$$\begin{aligned}\hat{\gamma}_{01,h} &= \frac{1}{T}\sum_{t=h+1}^{T} u_{0,t}^{\mu}\varepsilon_{0,t-h}^{\mu} + \frac{1}{T}\sum_{t=h+1}^{T} u_{0,t}^{\mu}X_{t-h-1}^{\mu\prime}\left(\Phi_T - \hat{R}_T\right) \\ &\quad + \left(\beta-\hat{\beta}\right)'\frac{1}{T}\sum_{t=h+1}^{T} X_{t-1}^{\mu}\varepsilon_{0,t-h}^{\mu} + \left(\beta-\hat{\beta}\right)'\frac{1}{T}\sum_{t=h+1}^{T} X_{t-1}^{\mu}X_{t-h-1}^{\mu}\left(\Phi_T - \hat{R}_T\right).\end{aligned} \tag{125}$$

Note that

$$\frac{1}{T}\sum_{t=h+1}^{T} u_{0,t}^{\mu}X_{t-h-1}^{\mu\prime} = \rho_T^h\frac{1}{T}\sum_{t=h+1}^{T} u_{0,t-h}^{\mu}X_{t-h-1}^{\mu\prime} + \sum_{j=0}^{h-1}\rho_T^j\left(\frac{1}{T}\sum_{t=h+1}^{T}\varepsilon_{0,t-j}^{\mu}X_{t-h-1}^{\mu\prime}\right).$$

By Theorem 4.1 in Phillips (1988), Equation (17) in Phillips and Magdalinos (2007), Equation (10) in Magdalinos and Phillips (2009), Equation (71)-(75), we have

$$\begin{aligned}\frac{1}{T}\sum_{t=h+1}^{T} u_{0,t-h}^{\mu}X_{t-h-1}^{\mu\prime} &= O_p(T^{\min\{\kappa_u,\kappa_x\}}), \\ \frac{1}{T}\sum_{t=h+1}^{T}\varepsilon_{0,t-j}^{\mu}X_{t-h-1}^{\mu\prime} &= O_p(1), \\ \Phi_T - \hat{R}_T &= O_p(T^{-\kappa_x}),\end{aligned}$$

These results imply

$$\begin{aligned}\frac{1}{T}\sum_{t=h+1}^{T} u_{0,t}^{\mu}X_{t-h-1}^{\mu\prime}\left(\Phi_T - \hat{R}_T\right) &= \rho_T^h\frac{1}{T}\sum_{t=h+1}^{T} u_{0,t-h}^{\mu}X_{t-h-1}^{\mu\prime}\left(\Phi_T - \hat{R}_T\right) + o_p(1) \\ &= \begin{cases} o_p(1) \text{ if } \kappa_x > \kappa_u, \\ O_p(1), \text{ if } \kappa_x \leq \kappa_u, \end{cases}.\end{aligned}$$

It is straightforward to verify that $\frac{1}{T}\sum_{t=h+1}^{T} u_{0,t}^{\mu}\varepsilon_{0,t-h}^{\mu} = O_p(1)$ and the rest of the terms appear in (125) is either $O_p(1)$ or $o_p(1)$. Thus, we can obtain

$$\hat{\Omega}_{01} = O_p(M_T) = O_p(T^{1/3}) \tag{126}$$

and

$$\hat{\Omega}_{FM} = O_p(T^{\kappa_u}) + O_p(M_T^2) = O_p(T^{\kappa_u}) + O_p(T^{2/3}). \tag{127}$$

For the orders of $T\bar{Z}_{t-1}\bar{Z}'_{t-1}$, the proof of Theorem 1 in the Online Appendix of Kostakis et al. (2015) shows that

$$T\bar{Z}_{t-1}\bar{Z}'_{t-1} = \begin{cases} O_p(T^{2\eta}), \text{ if } \eta < \kappa_x, \\ O_p(T^{2\kappa_x+\eta-1}), \text{ if } \eta \geq \kappa_x. \end{cases} \tag{128}$$

If $\kappa_u \geq 2/3$, $\hat{\Omega}_{FM} = O_p(T^{\kappa_u})$, we have

$$T\bar{Z}_{t-1}\bar{Z}'_{t-1}\hat{\Omega}_{FM} = \begin{cases} O_p(T^{2\eta+\kappa_u}), \text{ if } \eta < \kappa_x, \\ O_p(T^{2\kappa_x+\eta-1+\kappa_u}), \text{ if } \eta \geq \kappa_x. \end{cases}$$

From (124), we can obtain

$$\begin{aligned} M &= \left[\sum_{t=2}^{T} Z_{t-1}Z'_{t-1}\right]\hat{\Sigma}_{00} - T\bar{Z}_{t-1}\bar{Z}'_{t-1}\hat{\Omega}_{FM} \\ &= \begin{cases} O_p(T^{1+\eta+\kappa_u}) + O_p(T^{2\eta+\kappa_u}), \text{ if } \eta < \kappa_x, \\ O_p(T^{1+\kappa_x+\kappa_u}) + O_p(T^{2\kappa_x+\eta-1+\kappa_u}), \text{ if } \eta \geq \kappa_x. \end{cases} \\ &= \begin{cases} O_p(T^{1+\eta+\kappa_u}), \text{ if } \eta < \kappa_x, \\ O_p(T^{1+\kappa_x+\kappa_u}), \text{ if } \eta \geq \kappa_x, \end{cases} \\ &= O_p(T^{1+\min\{\eta,\kappa_x\}+\kappa_u}). \end{aligned}$$

This implies that

$$\begin{aligned} Q_{IVX} &= R\left[\sum_{t=2}^{T} Z_{t-1}X^{\mu\prime}_{t-1}\right]^{-1} M \left[\sum_{t=2}^{T} X^{\mu}_{t-1}Z'_{t-1}\right]^{-1} R' \\ &= \left[O_p(T^{1+\min\{\kappa_x,\eta\}})\right]^{-2} O_p(T^{1+\min\{\eta,\kappa_x\}+\kappa_u}) \\ &= O_p(T^{\kappa_u-1-\min\{\eta,\kappa_x\}}). \end{aligned}$$

Using (123), we can show

$$\begin{aligned} W_{\hat{\beta}_{IVX}} &= \left(R\hat{\beta}_{IVX,T} - r\right)' Q_{IVX}^{-1}\left(R\hat{\beta}_{IVX,T} - r\right) \\ &= O_p(T^{2\min\{\eta,\kappa_u\}-2\min\{\kappa_x,\eta\}})O_p(T^{1+\min\{\eta,\kappa_x\}-\kappa_u}) \\ &= O_p(T^{1+2\min\{\eta,\kappa_u\}-\min\{\kappa_x,\eta\}-\kappa_u}). \end{aligned}$$

If $1/3 < \kappa_u < 2/3$, $\hat{\Omega}_{FM} = O_p(T^{2/3})$, we have

$$T\bar{Z}_{t-1}\bar{Z}'_{t-1}\hat{\Omega}_{FM} = \begin{cases} O_p(T^{2\eta+2/3}), & \text{if } \eta < \kappa_x, \\ O_p(T^{2\kappa_x+\eta-1/3}), & \text{if } \eta \geq \kappa_x. \end{cases} \tag{129}$$

From (124) and (129), straightforward calculation yields

$$\begin{aligned} M &= \left[\sum_{t=2}^{T} Z_{t-1}Z'_{t-1}\right]\hat{\Sigma}_{00} - T\bar{Z}_{t-1}\bar{Z}'_{t-1}\hat{\Omega}_{FM} \\ &= \begin{cases} O_p(T^{1+\eta+\kappa_u}) + O_p(T^{2\eta+2/3}), & \text{if } \eta < \kappa_x \\ O_p(T^{1+\kappa_x+\kappa_u}) + O_p(T^{2\kappa_x+\eta-1/3}), & \text{if } \eta \geq \kappa_x \end{cases} \\ &= \begin{cases} O_p(T^{1+\eta+\max\{\kappa_u,\eta-1/3\}}), & \text{if } \eta < \kappa_x \\ O_p(T^{\kappa_x+\max\{\kappa_x-1/3+\eta,1+\kappa_u\}}), & \text{if } \eta \geq \kappa_x \end{cases}. \end{aligned}$$

Eventually, we have

$$\begin{aligned} Q_{IVX} &= R\left[\sum_{t=2}^{T} Z_{t-1}X^{\mu\prime}_{t-1}\right]^{-1} M \left[\sum_{t=2}^{T} X^{\mu}_{t-1}Z'_{t-1}\right]^{-1} R' \\ &= \begin{cases} \left[O_p(T^{1+\min\{\kappa_x,\eta\}})\right]^{-2} O_p(T^{1+\eta+\max\{\kappa_u,\eta-1/3\}}), & \text{if } \eta < \kappa_x \\ \left[O_p(T^{1+\min\{\kappa_x,\eta\}})\right]^{-2} O_p(T^{\kappa_x+\max\{\kappa_x-1/3+\eta,1+\kappa_u\}}), & \text{if } \eta \geq \kappa_x \end{cases} \\ &= \begin{cases} O_p(T^{\max\{\kappa_u,\eta-1/3\}-1-\eta}), \text{if } \eta < \kappa_x \\ O_p(T^{\max\{\kappa_x-1/3+\eta,1+\kappa_u\}-2-\kappa_x}), & \text{if } \eta \geq \kappa_x \end{cases}. \end{aligned}$$

Therefore, we have the following stochastic order for the Wald statistic,

$$\begin{aligned} W_{\hat{\beta}_{IVX}} &= \left(R\hat{\beta}_{IVX,T} - r\right)' Q_{IVX}^{-1}\left(R\hat{\beta}_{IVX,T} - r\right) \\ &= \begin{cases} O_p\left(T^{2\min\{\eta,\kappa_u\}-2\min\{\kappa_x,\eta\}}\right) O_p(T^{1+\eta-\max\{\kappa_u,\eta-1/3\}}), & \text{if } \eta < \kappa_x \\ O_p\left(T^{2\min\{\eta,\kappa_u\}-2\min\{\kappa_x,\eta\}}\right) O_p(T^{2+\kappa_x-\max\{\kappa_x-1/3+\eta,1+\kappa_u\}}), & \text{if } \eta \geq \kappa_x \end{cases} \\ &= \begin{cases} O_p\left(T^{1+2\min\{\eta,\kappa_u\}-\max\{\kappa_u,\eta-1/3\}-\eta}\right), & \text{if } \eta < \kappa_x \\ O_p\left(T^{2+2\min\{\eta,\kappa_u\}-\max\{\kappa_x-1/3+\eta,1+\kappa_u\}-\kappa_x}\right), & \text{if } \eta \geq \kappa_x \end{cases}. \end{aligned}$$

It can be directly verify that $W_{\hat{\beta}_{IVX}} \xrightarrow{p} \infty$ given the above stochastic orders and our parameter settings.

(iv). The first result can be obtained directly from Theorem 2.2 in Phillips and Lee

(2016). For the second result, scaling the Wald statistic by $\frac{\hat{\Sigma}_{00}}{\hat{\Omega}_{00}}$ gives

$$\frac{\hat{\Sigma}_{00}}{\hat{\Omega}_{00}} W_{\hat{\beta}_{IVX}} \equiv \left(R\hat{\beta}_{IVX} - r\right)' \left[\frac{\hat{\Omega}_{00}}{\hat{\Sigma}_{00}} Q_{IVX}\right]^{-1} \left(R\hat{\beta}_{IVX} - r\right), \tag{130}$$

where

$$\begin{aligned}
\frac{\hat{\Omega}_{00}}{\hat{\Sigma}_{00}} Q_{IVX} &= R\left[\sum_{t=2}^{T} Z_{t-1} X_{t-1}^{\mu\prime}\right]^{-1} \left(\frac{\hat{\Omega}_{00}}{\hat{\Sigma}_{00}} M\right) \left[\sum_{t=2}^{T} X_{t-1}^{\mu} Z_{t-1}'\right]^{-1} R', \\
\frac{\hat{\Omega}_{00}}{\hat{\Sigma}_{00}} M &= \left[\sum_{t=2}^{T} Z_{t-1} Z_{t-1}'\right] \hat{\Omega}_{00} - \frac{\hat{\Omega}_{00}}{\hat{\Sigma}_{00}} T \bar{Z}_{t-1} \bar{Z}_{t-1}' \hat{\Omega}_{FM}.
\end{aligned}$$

Since $\hat{\Omega}_{00} \xrightarrow{p} \Omega_u$, $\hat{\Sigma}_{00} \xrightarrow{p} \Sigma_u$, and $T\bar{Z}_{t-1}\bar{Z}_{t-1}'\hat{\Omega}_{FM}$ is only used for finite sample correction, by applying Theorem 2.3 in Phillips and Lee (2016), we obtain

$$W_{\hat{\beta}_{IVX}} \Rightarrow \frac{\Omega_u}{\Sigma_u} \chi^2(q). \tag{131}$$

(v). We first focus on the second term in the right side of (102). Since $\Psi_{T,t} = \sum_{j=1}^{t} \Upsilon_T^{t-j} X_{j-1}$,

$$\Psi_{T,t} = \Upsilon_T \Psi_{T,t-1} + X_{t-1}, \Psi_{T,0} = X_0 = O_p(1).$$

We can express

$$\begin{aligned}
u_{0,t}^{\mu} \Psi_{T,t-1} &= (\rho_T u_{0,t-1}^{\mu} + \varepsilon_{0,t}^{\mu})(\Upsilon_T \Psi_{T,t-2} + X_{t-2}) \\
&= \rho_T \Upsilon_T u_{0,t-1}^{\mu} \Psi_{T,t-2} + \rho_T u_{0,t-1}^{\mu} X_{t-2} + \varepsilon_{0,t}^{\mu} \Upsilon_T \Psi_{T,t-2} + \varepsilon_{0,t}^{\mu} X_{t-2}.
\end{aligned}$$

Summing the above expression over $t = 1, ..., T$, we have

$$\begin{aligned}
(\rho_T \Upsilon_T - I_k) \sum_{t=1}^{T} u_{0,t-1}^{\mu} \Psi_{T,t-2} &= u_{0,T}^{\mu} \Psi_{T,T-1} - u_{0,0}^{\mu} \Psi_{T,1} - \rho_T \sum_{t=1}^{T} u_{0,t-1}^{\mu} X_{t-2} \\
&\quad - \Upsilon_T \sum_{t=1}^{T} \varepsilon_{0,t}^{\mu} \Psi_{T,t-2} - \sum_{t=1}^{T} \varepsilon_{0,t}^{\mu} X_{t-2}.
\end{aligned} \tag{132}$$

The proofs of Equation (26) in Magdalinos and Phillips (2009), Lemma 2.4 in Phillips and Lee (2016) and (77) show the following orders:

$$\sum_{t=1}^{T} u_{0,t-1}^{\mu} X_{t-2} = O_p\left(\Phi_T^T T^{(3\kappa_x+1)/2}\right),$$

$$\begin{aligned}\sum_{t=1}^{T} \varepsilon_{0,t}^{\mu} \Psi_{T,t-2} &= O_p\left(\Phi_T^T T^{\min\{\kappa_x,\eta\}}\right), \\ \sum_{t=1}^{T} \varepsilon_{0,t}^{\mu} X_{t-2} &= O_p\left(\Phi_T^T T^{\kappa_x}\right).\end{aligned}$$

Since $\kappa_x > 1/3$, $\sum_{t=1}^{T} u_{0,t-1}^{\mu} X_{t-2}$ asymptotically dominates the other two terms. From Lemma 2.3 in Phillips and Lee (2016) and $u_{0,t}^{\mu} = O_p(\sqrt{T})$, we can obtain

$$\frac{\Phi_T^{-T}}{T^{(3\kappa_x+1)/2}} \left(u_{0,T}^{\mu} \Psi_{T,T-1} - u_{0,0}^{\mu} \Psi_{T,1}\right) = o_p(1).$$

Therefore, applying (77) and noting that $\rho_T \Upsilon_T - I_k = \frac{C_x}{T^{\kappa_x}} + (1)$, we have

$$\begin{aligned}\frac{\Phi_T^{-T}}{T^{(5\kappa_x+1)/2}} \sum_{t=2}^{T} u_{0,t-1}^{\mu} \Psi_{T,t-2} &= \rho_T \frac{\Phi_T^{-T}}{T^{(3\kappa_x+1)/2}} \sum_{t=2}^{T} u_{0,t-1}^{\mu} X_{t-2} + o_p(1) \\ &\Rightarrow C_x^{-2} Y_{C_x} J_{c_u}^{\mu}(r).\end{aligned}$$

Thus,

$$(\Phi_T - I_k) \sum_{t=1}^{T} u_{0,t}^{\mu} \Psi_{T,t-1} = \frac{C_x}{T^{\kappa_x}} O_p(\Phi_T^T T^{(5\kappa_x+1)/2}) = O_p(\Phi_T^T T^{(3\kappa_x+1)/2}).$$

Since

$$\sum_{t=1}^{T} u_{0,t}^{\mu} \tilde{Z}_{t-1} = O_p(T^{1+\min\{\kappa_x,\eta\}}),$$

it is dominated by

$$(\Phi_T - I_k) \sum_{t=1}^{T} u_{0,t}^{\mu} \Psi_{T,t-1}$$

as $T \to \infty$. Eventually, we have

$$\begin{aligned}\frac{\Phi_T^{-T}}{T^{(3\kappa_x+1)/2}} \sum_{t=2}^{T} u_{0,t}^{\mu} Z_{t-1} &= \frac{\Phi_T^{-T}}{T^{(3\kappa_x+1)/2}} \sum_{t=2}^{T} u_{0,t}^{\mu} \tilde{Z}_{t-1} + (\Phi_T - I_k) \frac{\Phi_T^{-T}}{T^{(3\kappa_x+1)/2}} \sum_{t=2}^{T} u_{0,t}^{\mu} \Psi_{T,t-1} \\ &= (\Phi_T - I_k) \frac{\Phi_T^{-T}}{T^{(3\kappa_x+1)/2}} \sum_{t=2}^{T} u_{0,t}^{\mu} \Psi_{T,t-1} + o_p(1) \\ &= \frac{C_x \Phi_T^{-T}}{T^{(5\kappa_x+1)/2}} \sum_{t=2}^{T} u_{0,t}^{\mu} \Psi_{T,t-1} + o_p(1) \\ &\Rightarrow C_x^{-1} Y_{C_x} J_{c_u}^{\mu}(r). \qquad (133)\end{aligned}$$

By Lemma 2.4 of Phillips and Lee (2016), $\frac{1}{T^{\kappa_x+\min(\kappa_x,\eta)}}\sum_{t=2}^{T}\Phi_T^{-T}Z_{t-1}X_{t-1}^{\mu\prime}\Phi_T^{-T}=O_p(1)$. This result and (133) give

$$\begin{aligned}
&\Phi_T^{T}T^{\min(\kappa_x,\eta)-\frac{\kappa_x+1}{2}}\left(\hat{\beta}_{IVX}-\beta\right)\\
=\ &\Phi_T^{T}T^{\kappa_x+\min(\kappa_x,\eta)-(3\kappa_x+1)/2}\left[\sum_{t=2}^{T}Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1}\sum_{t=2}^{T}Z_{t-1}u_{0,t}^{\mu}\\
=\ &\left[\frac{1}{T^{\kappa_x+\min(\kappa_x,\eta)}}\sum_{t=2}^{T}\Phi_T^{-T}Z_{t-1}X_{t-1}^{\mu\prime}\Phi_T^{-T}\right]^{-1}\frac{\Phi_T^{-T}}{T^{(3\kappa_x+1)/2}}\sum_{t=2}^{T}Z_{t-1}u_{0,t}^{\mu}\\
=\ &O_p(1).
\end{aligned}$$

For the Wald statistic, Lemma A.2 in Phillips and Lee (2016) gives

$$\frac{1}{T^{2\min\{\kappa_x,\eta\}}}\sum_{t=2}^{T}\Phi_T^{-T}Z_{t-1}Z_{t-1}'\Phi_T^{-T}=O_p(1), \tag{134}$$

and

$$\begin{aligned}
\hat{\Sigma}_{00}=\frac{1}{T}\sum_{t=1}^{T}e_{0,t}^{2}\ =\ &\left(\beta-\hat{\beta}\right)'\frac{1}{T}\sum_{t=1}^{T}X_{t-1}^{\mu}X_{t-1}^{\mu\prime}\left(\beta-\hat{\beta}\right)'\\
&+2\left(\beta-\hat{\beta}\right)'\frac{1}{T}\sum_{t=1}^{T}X_{t-1}^{\mu}u_{0,t}^{\mu}+\frac{1}{T}\sum_{t=1}^{T}u_{0,t}^{\mu 2}.
\end{aligned}$$

Applying the results from (76), (77), Lemma A.4(viii) and Lemma 3.1 in Phillips (1988), we can show that the last term $\frac{1}{T}\sum_{t=1}^{T}u_{0,t}^{\mu 2}$ asymptotically dominates the other terms. Thus, we have

$$\hat{\Sigma}_{00}=\frac{1}{T}\sum_{t=1}^{T}u_{0,t}^{\mu 2}+o_p(1)=O_p(T).$$

For $\bar{Z}_t$, using Equation (13) in Phillips and Magdalinos (2009), we can express

$$\begin{aligned}
\bar{Z}_t\ &=\ \frac{1}{T}\sum_{t=1}^{T}Z_t=\frac{1}{T}\sum_{t=1}^{T}\left(Z_t^{*}+\frac{C_x}{T^{\kappa_x}}\Psi_{T,t}\right)\\
&=\ \frac{1}{T}\sum_{t=1}^{T}Z_t^{*}+\frac{C_x}{T^{1+\kappa_x}}\sum_{t=1}^{T}\Psi_{T,t}.
\end{aligned}$$

Applying Lemma 2.3 in Phillips and Lee (2016), we can express

$$\begin{aligned}\frac{\Phi_T^{-T}}{T^{\kappa_x+\kappa_x/2+\min\{\eta,\kappa_x\}}}\sum_{t=1}^{T}\Psi_{T,t} &= \frac{1}{T^{\kappa_x}}\sum_{t=1}^{T}\Phi_T^{-(T-t)}\left(\frac{1}{T^{\kappa_x/2+\min\{\eta,\kappa_x\}}}\Phi_T^{-t}\Psi_{T,t}\right)\\ &= \frac{1}{T^{\kappa_x}}\sum_{t=k_T+\kappa_T'}^{T}\Phi_T^{-(T-t)}\left(C_{z,\kappa_x,\eta}Y_{C_x}\right)+o_p(1)\\ &= \frac{1}{T^{\kappa_x}}C_{z,\kappa_x,\eta}\sum_{t=k_T+\kappa_T'}^{T}\Phi_T^{-(T-t)}Y_{C_x}+o_p(1)\\ &= C_{z,\kappa_x,\eta}\left(\sum_{t=1}^{T}\Phi_T^{-(T-t)}\right)Y_{C_x}+o_p(1)\\ &= \frac{1}{T^{\kappa_x}}O_p(T^{\kappa_x})=O_p(1).\end{aligned}$$

This implies $\frac{C_x}{T^{1+\kappa_x}}\sum_{t=1}^{T}\Psi_{T,t}=O_p(\Phi_T^{T}T^{\kappa_x/2+\min\{\eta,\kappa_x\}-1})$. Moreover, using the analogous steps in proving Equation (7) in Giraitis and Phillips (2006), one can easily show $\frac{1}{T}\sum_{t=1}^{T}Z_t^{*}=O_p(T^{\eta-1/2})$.Thus

$$\bar{Z}_{t-1}=O_p(\Phi_T^{T}T^{\kappa_x/2+\min\{\eta,\kappa_x\}-1}). \tag{135}$$

Finally, the order of $\hat{\Omega}_{FM}$ can be established using the same procedure as in (114) and we have

$$\hat{\Omega}_{FM}=O_p(T)+O_p(T^{2/3})=O_p(T). \tag{136}$$

Thus, we can establish the order of M as

$$\begin{aligned}&\frac{\Phi_T^{-T}M\Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+1}}\\ =\ &\left[\frac{1}{T^{2\min\{\kappa_x,\eta\}}}\sum_{t=2}^{T}\Phi_T^{-T}Z_{t-1}Z_{t-1}'\Phi_T^{-T}\right]\frac{\hat{\Sigma}_{00}}{T}-\frac{\Phi_T^{-T}T\bar{Z}_{t-1}\bar{Z}_{t-1}'\hat{\Omega}_{FM}\Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+1}}.\end{aligned}$$

From (135), (136) and since $\kappa_x<1$, we have

$$\frac{\Phi_T^{-T}\bar{Z}_{t-1}\bar{Z}_{t-1}'\hat{\Omega}_{FM}\Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+1}}=o_p(1).$$

This gives

$$\frac{\Phi_T^{-T}M\Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+1}}=\left[\frac{1}{T^{2\min\{\kappa_x,\eta\}}}\sum_{t=2}^{T}\Phi_T^{-T}Z_{t-1}Z_{t-1}'\Phi_T^{-T}\right]\frac{\hat{\Sigma}_{00}}{T}+o_p(1)=O_p(1),$$

Therefore, we have

$$\begin{aligned} T^{2\kappa_x-1}\Phi_T^T Q_{IVX}\Phi_T^T &= T^{2\kappa_x-1} R\left[\sum_{t=2}^T Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1} M\left[\sum_{t=2}^T X_{t-1}^{\mu} Z_{t-1}^{\prime}\right]^{-1} R^{\prime} \\ &= R\left[\frac{1}{T^{\kappa_x+\min\{\kappa_x,\eta\}}}\sum_{t=2}^T \Phi_T^{-T} Z_{t-1}X_{t-1}^{\mu\prime}\Phi_T^{-T}\right]^{-1}\frac{\Phi_T^{-T} M\Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+1}} \\ &\quad \times\left[\frac{1}{T^{\kappa_x+\min\{\kappa_x,\eta\}}}\sum_{t=2}^T \Phi_T^{-T} X_{t-1}^{\mu} Z_{t-1}^{\prime}\Phi_T^{-T}\right]^{-1} R^{\prime} \\ &= O_p(1). \end{aligned}$$

It leads to

$$\begin{aligned} T^{2\min(\kappa_x,\eta)-3\kappa_x} W_{\hat{\beta}_{IVX}} &= \frac{T^{2\min(\kappa_x,\eta)-\kappa_x-1}}{T^{2\kappa_x-1}}\left[R\left(\hat{\beta}_{IVX}-\beta\right)\right]^{\prime}\left[Q_{IVX}\right]^{-1} R\left(\hat{\beta}_{IVX}-\beta\right) \\ &= \left[R\Phi_T^T T^{\min(\kappa_x,\eta)-\frac{\kappa_x+1}{2}}\left(\hat{\beta}_{IVX}-\beta\right)\right]^{\prime}\left[T^{2\kappa_x-1}\Phi_T^T Q_{IVX}\Phi_T^T\right]^{-1} \\ &\quad \times R\Phi_T^T T^{\min(\kappa_x,\eta)-\frac{\kappa_x+1}{2}}\left(\hat{\beta}_{IVX}-\beta\right) \\ &= O_p(1). \end{aligned}$$

Eventually, $W_{\hat{\beta}_{IVX}} = O_p(T^{3\kappa_x-2\min(\kappa_x,\eta)})$.

(vi). As in the previous proof, we focus on the stochastic orders of the terms in (132). The proofs of Equation (26) in Magdalinos and Phillips (2009), Lemma 2.4 in Phillips and Lee (2016), (81), (82), (83) and (84) show the following orders:

$$\sum_{t=2}^T \varepsilon_{0,t}^{\mu}\Psi_{T,t-2} = O_p\left(\Phi_T^T T^{\min\{\kappa_x,\eta\}}\right), \sum_{t=2}^T \varepsilon_{0,t}^{\mu} X_{t-2} = O_p\left(\Phi_T^T T^{\kappa_x}\right),$$

$$\frac{1}{T^{\min\{\kappa_x,\kappa_u\}}}\frac{\Phi_T^{-T}}{T^{(\max\{\kappa_x,\kappa_u\}+\kappa_x)/2}}\sum_{t=2}^T X_{t-2}u_{0,t-1}^{\mu} = O_p(1). \tag{137}$$

Combining $u_{0,t}^{\mu} = O_p(T^{\kappa_u/2})$ from Lemma A.1 in Lin and Tu (2020) and Lemma 2.3 in Phillips and Lee (2016), we have

$$\frac{\Phi_T^{-T}}{T^{(\kappa_u+\kappa_x)/2+\min\{\kappa_x,\eta\}}}\left(u_{0,T}^{\mu}\Psi_{T,T-1} - u_{0,0}^{\mu}\Psi_{T,1}\right) = O_p(1). \tag{138}$$

Using the above stochastic orders, after some straightforward calculations, we can obtain

$$(\rho_T \Upsilon_T - I_k) \sum_{t=2}^{T} u_{0,t-1}^{\mu} \Psi_{T,t-2} = \begin{cases} O_p(\Phi_T^T T^{\kappa_x+\kappa_u}), \text{ if } \kappa_x \geq \kappa_u \geq \eta, \\ O_p(\Phi_T^T T^{(\kappa_u+3\kappa_x)/2}), \text{ otherwise.} \end{cases}$$

Or equivalently, given that

$$\rho_T \Upsilon_T - I_k = \frac{c_u I_k}{T^{\kappa_u}} + \frac{C_z}{T^{\eta}} + o(1) = O\left(\frac{1}{T^{\min\{\kappa_u,\eta\}}}\right),$$

we can write

$$\sum_{t=2}^{T} u_{0,t-1}^{\mu} \Psi_{T,t-2} = \begin{cases} O_p(\Phi_T^T T^{\kappa_x+\kappa_u+\eta}), \text{ if } \kappa_x \geq \kappa_u \geq \eta, \\ O_p(\Phi_T^T T^{(\kappa_u+3\kappa_x)/2+\min\{\kappa_u,\eta\}}), \text{ otherwise.} \end{cases}$$

Form (122), it is clear that $(\rho_T \Upsilon_T - I_k) \sum_{t=1}^{T} u_{0,t-1}^{\mu} \Psi_{T,t-2}$ dominates the first term on the right-hand side of (102). Therefore, when $\kappa_x \geq \kappa_u \geq \eta$, we have

$$\begin{aligned} & \frac{\Phi_T^{-T}}{T^{\kappa_x+\kappa_u+\eta}} \sum_{t=2}^{T} u_{0,t}^{\mu} Z_{t-1} \\ = \; & (\Phi_T - I_k) \frac{\Phi_T^{-T}}{T^{\kappa_x+\kappa_u+\eta}} \sum_{t=2}^{T} u_{0,t}^{\mu} \Psi_{T,t-1} + O_p\left(\frac{\Phi_T^{-T}}{T^{\kappa_x+\kappa_u+\eta-1}}\right). \end{aligned}$$

When $\kappa_x \geq \kappa_u \geq \eta$ does not hold, we have

$$\begin{aligned} & \frac{\Phi_T^{-T}}{T^{(\kappa_u+3\kappa_x)/2+\min\{\kappa_u,\eta\}}} \sum_{t=2}^{T} u_{0,t}^{\mu} Z_{t-1} \\ = \; & (\Phi_T - I_k) \frac{\Phi_T^{-T}}{T^{(\kappa_u+3\kappa_x)/2+\min\{\kappa_u,\eta\}}} \sum_{t=2}^{T} u_{0,t}^{\mu} \Psi_{T,t-1} + O_p\left(\frac{\Phi_T^{-T}}{T^{(\kappa_u+3\kappa_x)/2+\min\{\kappa_u,\eta\}-1}}\right). \end{aligned}$$

Since $\Phi_T - I_k = \frac{C_x}{T^{\kappa_x}}$, we can obtain

$$\sum_{t=2}^{T} u_{0,t}^{\mu} Z_{t-1} = \begin{cases} O_p(\Phi_T^T T^{\kappa_u+\eta}), \text{ if } \kappa_x \geq \kappa_u \geq \eta, \\ O_p(\Phi_T^T T^{(\kappa_u+\kappa_x)/2+\min\{\kappa_u,\eta\}}), \text{ otherwise.} \end{cases} \tag{139}$$

Lemma 2.4 of Phillips and Lee (2016) and (139) give

$$\hat{\beta}_{IVX} - \beta = \begin{cases} O_p(\Phi_T^{-T} T^{\kappa_u+\eta-\kappa_x-\min\{\kappa_x,\eta\}}), \text{ if } \kappa_x \geq \kappa_u \geq \eta, \\ O_p(\Phi_T^{-T} T^{(\kappa_u+\kappa_x)/2+\min\{\kappa_u,\eta\}-\kappa_x-\min\{\kappa_x,\eta\}}), \text{ otherwise.} \end{cases}$$

For the Wald statistic, following (63) and (127), we can obtain $\hat{\Sigma}_{00} = O_p(T^{\kappa_u})$ and $\hat{\Omega}_{FM} = O_p(T^{\kappa_u}) + O_p(T^{2/3})$. If $\kappa_x \leq 1/3 + \kappa_u$, by straightforward calculation, we can show

$$\frac{\Phi_T^{-T} T \bar{Z}_{t-1} \bar{Z}_{t-1}' \hat{\Omega}_{FM} \Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+\kappa_u}} = o_p(1).$$

We can establish the order of M as

$$\begin{aligned}
\frac{\Phi_T^{-T} M \Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+\kappa_u}} &= \left[\frac{1}{T^{2\min\{\kappa_x,\eta\}}} \sum_{t=2}^{T} \Phi_T^{-T} Z_{t-1} Z_{t-1}' \Phi_T^{-T}\right] \frac{\hat{\Sigma}_{00}}{T^{\kappa_u}} - \frac{\Phi_T^{-T} T \bar{Z}_{t-1} \bar{Z}_{t-1}' \hat{\Omega}_{FM} \Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+\kappa_u}} \\
&= \left[\frac{1}{T^{2\min\{\kappa_x,\eta\}}} \sum_{t=2}^{T} \Phi_T^{-T} Z_{t-1} Z_{t-1}' \Phi_T^{-T}\right] \frac{\hat{\Sigma}_{00}}{T^{\kappa_u}} + o_p(1) \\
&= O_p(1).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
T^{2\kappa_x-\kappa_u} \Phi_T^T Q_{IVX} \Phi_T^T &= T^{2\kappa_x-\kappa_u} \Phi_T^T R \left[\sum_{t=2}^{T} Z_{t-1} X_{t-1}^{\mu\prime}\right]^{-1} M \left[\sum_{t=2}^{T} X_{t-1}^{\mu} Z_{t-1}'\right]^{-1} R' \Phi_T^T \\
&= R \left[\frac{1}{T^{\kappa_x+\min\{\kappa_x,\eta\}}} \sum_{t=2}^{T} \Phi_T^{-T} Z_{t-1} X_{t-1}^{\mu\prime} \Phi_T^{-T}\right]^{-1} \frac{\Phi_T^{-T} M \Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+\kappa_u}} \\
&\quad \times \left[\frac{1}{T^{\kappa_x+\min\{\kappa_x,\eta\}}} \sum_{t=2}^{T} \Phi_T^{-T} X_{t-1}^{\mu} Z_{t-1}' \Phi_T^{-T}\right]^{-1} R' \\
&= O_p(1).
\end{aligned}$$

If that $\kappa_x \geq \kappa_u \geq \eta$, we can obtain

$$\begin{aligned}
& T^{3\kappa_u-4\kappa_x+2\eta-2\min\{\kappa_x,\eta\}} W_{\hat{\beta}_{IVX}} \\
=\ & \frac{T^{2(\kappa_u+\eta-\kappa_x-\min\{\kappa_x,\eta\})}}{T^{2\kappa_x-\kappa_u}} \left[R\left(\hat{\beta}_{IVX} - \beta\right)\right]' \left[Q_{IVX}\right]^{-1} R\left(\hat{\beta}_{IVX} - \beta\right) \\
=\ & \left[R \Phi_T^{-T} T^{\kappa_u+\eta-\kappa_x-\min\{\kappa_x,\eta\}} \left(\hat{\beta}_{IVX} - \beta\right)\right]' \left[T^{2\kappa_x-\kappa_u} \Phi_T^T Q_{IVX} \Phi_T^T\right]^{-1} \\
& \times R \Phi_T^{-T} T^{\kappa_u+\eta-\kappa_x-\min\{\kappa_x,\eta\}} \left(\hat{\beta}_{IVX} - \beta\right) \\
=\ & O_p(1).
\end{aligned}$$

If $\kappa_x \geq \kappa_u \geq \eta$ does not hold, we have

$$\frac{1}{T^{5\kappa_x+2\min\{\kappa_u,\eta\}-4\min\{\kappa_x,\eta\}}} W_{\hat{\beta}_{IVX}}$$

$$
\begin{aligned}
&= \frac{1}{T^{5\kappa_x+2\min\{\kappa_u,\eta\}-4\min\{\kappa_x,\eta\}}}\left(R\left(\hat{\beta}_{IVX}-\beta\right)\right)' Q_{IVX}^{-1}\left(R\left(\hat{\beta}_{IVX}-\beta\right)\right) \\
&= \left(R\Phi_T^T\frac{T^{2\min\{\kappa_x,\eta\}}}{T^{(\kappa_u+3\kappa_x)/2+\min\{\kappa_u,\eta\}}}\left(\hat{\beta}_{IVX}-\beta\right)\right)'\left[T^{2\kappa_x-\kappa_u}\Phi_T^T Q_{IVX}\Phi_T^T\right]^{-1} \\
&\quad \left(R\Phi_T^T\frac{T^{2\min\{\kappa_x,\eta\}}}{T^{(\kappa_u+3\kappa_x)/2+\min\{\kappa_u,\eta\}}}\left(\hat{\beta}_{IVX}-\beta\right)\right) \\
&= O_p(1).
\end{aligned}
$$

Therefore, we can obtain

$$
W_{\hat{\beta}_{IVX}} = \begin{cases} O_p\left(T^{4\kappa_x-3\kappa_u+2\min\{\kappa_x,\eta\}-2\eta}\right), & \text{if } \kappa_x \geq \kappa_u \geq \eta, \\ O_p\left(T^{5\kappa_x+2\min\{\kappa_u,\eta\}-4\min\{\kappa_x,\eta\}}\right), & \text{otherwise.} \end{cases}
$$

If $\kappa_x > 1/3+\kappa_u$, $T\bar{Z}_{t-1}\bar{Z}_{t-1}'\hat{\Omega}_{FM}$ dominates $\left[\sum_{t=2}^T Z_{t-1}Z_{t-1}'\right]\hat{\Sigma}_{00}$ in M. Hence, we have

$$
\begin{aligned}
\frac{\Phi_T^{-T}M\Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+\kappa_x-1/3}} &= \frac{1}{T^{\kappa_x-1/3-\kappa_u}}\left[\frac{1}{T^{2\min\{\kappa_x,\eta\}}}\sum_{t=2}^T \Phi_T^{-T}Z_{t-1}Z_{t-1}'\Phi_T^{-T}\right]\frac{\hat{\Sigma}_{00}}{T^{\kappa_u}} \\
&\quad -\frac{\Phi_T^{-T}T\bar{Z}_{t-1}\bar{Z}_{t-1}'\hat{\Omega}_{FM}\Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+\kappa_x-1/3}} \\
&= -\frac{\Phi_T^{-T}T\bar{Z}_{t-1}\bar{Z}_{t-1}'\hat{\Omega}_{FM}\Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+\kappa_x-1/3}} + o_p(1) \\
&= O_p(1).
\end{aligned}
$$

We now follow a similar procedure to show the limit of $W_{\hat{\beta}_{IVX}}$ under $\kappa_x > 1/3+\kappa_u$. We have

$$
\begin{aligned}
T^{\kappa_x-1/3}\Phi_T^T Q_{IVX}\Phi_T^T &= T^{\kappa_x-1/3}\Phi_T^T R\left[\sum_{t=2}^T Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1} M\left[\sum_{t=2}^T X_{t-1}^{\mu}Z_{t-1}'\right]^{-1} R'\Phi_T^T \\
&= R\left[\frac{1}{T^{\kappa_x+\min\{\kappa_x,\eta\}}}\sum_{t=2}^T \Phi_T^{-T}Z_{t-1}X_{t-1}^{\mu\prime}\Phi_T^{-T}\right]^{-1}\frac{\Phi_T^{-T}M\Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+\kappa_x-1/3}} \\
&\quad \times\left[\frac{1}{T^{\kappa_x+\min\{\kappa_x,\eta\}}}\sum_{t=2}^T \Phi_T^{-T}X_{t-1}^{\mu}Z_{t-1}'\Phi_T^{-T}\right]^{-1} R' \\
&= O_p(1).
\end{aligned}
$$

Eventually, if $\kappa_x > \kappa_u \geq \eta$, we have

$$
T^{2\kappa_u-3\kappa_x+1/3}W_{\hat{\beta}_{IVX}} = T^{2\kappa_u-3\kappa_x+1/3}\left[R\left(\hat{\beta}_{IVX}-\beta\right)\right]' Q_{IVX}^{-1}R\left(\hat{\beta}_{IVX}-\beta\right)
$$

$$
\begin{aligned}
&= \left[R\Phi_T^{-T} T^{\kappa_u-\kappa_x}\left(\hat{\beta}_{IVX}-\beta\right)\right]'\left[T^{\kappa_x-1/3}\Phi_T^T Q_{IVX}\Phi_T^T\right]^{-1} \\
&\quad \times R\Phi_T^{-T} T^{\kappa_u-\kappa_x}\left(\hat{\beta}_{IVX}-\beta\right) \\
&= O_p(1).
\end{aligned}
$$

If $\kappa_x > \kappa_u \geq \eta$ does not hold, we have

$$
\begin{aligned}
& T^{\kappa_u/2+\min\{\kappa_u,\eta\}-\frac{3}{2}\kappa_x-\min\{\kappa_x,\eta\}+1/3} W_{\hat{\beta}_{IVX}} \\
=\; & T^{\kappa_u/2+\min\{\kappa_u,\eta\}-\frac{3}{2}\kappa_x-\min\{\kappa_x,\eta\}+1/3}\left[R\left(\hat{\beta}_{IVX}-\beta\right)\right]'\left[Q_{IVX}\right]^{-1} R\left(\hat{\beta}_{IVX}-\beta\right) \\
=\; & \left[R\Phi_T^T T^{(\kappa_u+\kappa_x)/2+\min\{\kappa_u,\eta\}-\kappa_x-\min\{\kappa_x,\eta\}}\left(\hat{\beta}_{IVX}-\beta\right)\right]'\left[T^{\kappa_x-1/3}\Phi_T^T Q_{IVX}\Phi_T^T\right]^{-1} \\
& \times R\Phi_T^T T^{(\kappa_u+\kappa_x)/2+\min\{\kappa_u,\eta\}-\kappa_x-\min\{\kappa_x,\eta\}}\left(\hat{\beta}_{IVX}-\beta\right) \\
=\; & O_p(1).
\end{aligned}
$$

We now proceed to prove the claims for the re-centered estimators.

(i). Since $2/3 < \eta < \min\{\kappa_x, 1\}$, from (95) and Lemma A0 in Phillips and Magdalinos (2009), let

$$
\hat{\Lambda}_{01} = \frac{1}{T}\sum_{h=0}^{M_T}\left(1-\frac{h}{M_T+1}\right)\sum_{t=h+1}^{T} e_{0,t}e_{1,t-h}, \tag{140}
$$

we have

$$
\hat{\Lambda}_{01} - \Lambda_{01} = o_p(T^{-\frac{1-\eta}{2}}),
$$

and

$$
\begin{aligned}
\frac{1}{T^{(1+\eta)/2}}\sum_{t=2}^{T}\left(Z_{t-1}u_{0,t}^{\mu}-\hat{\Lambda}_{01}\right) &= \frac{1}{T^{(1+\eta)/2}}\sum_{t=2}^{T}\left(Z_{t-1}u_{0,t}-\hat{\Lambda}_{01}\right)+o_p(1) \\
&= \frac{1}{T^{(1+\eta)/2}}\sum_{t=2}^{T}\left(Z_{t-1}u_{0,t}-\Lambda_{01}\right)+o_p(1) \\
&\Rightarrow N(0, V_{zz}\Omega_{00}).
\end{aligned} \tag{141}
$$

If $\kappa_x = 1$, Theorem A in the Online Appendix of Kostakis et al. (2015) gives

$$
\frac{1}{T^{1+\eta}}\sum_{t=2}^{T} Z_{t-1}X_{t-1}^{\mu'} \Rightarrow \Psi_{zz} \equiv -C_z^{-1}\left(\Omega_{00}+\int_0^1 J_{C_x}^{\mu'}(r)dJ_{C_x}^{\mu}(r)\right). \tag{142}
$$

where the last convergence in distribution can be obtained using the analogous argument in (69). The joint convergence of (142) and (141)[2] implies

$$T^{(1+\eta)/2}\left(\check{\beta}_{IVX}-\beta\right) = \left[\frac{1}{T^{1+\eta}}\sum_{t=2}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{T^{(1+\eta)/2}}\sum_{t=2}^{T}\left(Z_{t-1}u_{0,t}^{\mu}-\hat{\Lambda}_{01}\right)$$
$$\Rightarrow MN(0,\Psi_{zz}^{-1}V_{zz}\Omega_u\Psi_{zz}^{-1\prime}).$$

If $\eta < \kappa_x < 1$, from (90) and Equation (20) in Phillips and Magdalinos (2009), we can obtain

$$\frac{1}{T^{1+\eta}}\sum_{t=2}^{T} Z_{t-1}X_{t-1}^{\mu\prime} \xrightarrow{p} -C_z^{-1}\Omega_{11}.$$

Thus, we have

$$T^{(1+\eta)/2}\left(\check{\beta}_{IVX}-\beta\right) = \left[\frac{1}{T^{1+\eta}}\sum_{t=2}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{T^{(1+\eta)/2}}\sum_{t=2}^{T}\left(Z_{t-1}u_{0,t}^{\mu}-\hat{\Lambda}_{01}\right)$$
$$\Rightarrow \Omega_{11}^{-1}C_zN(0,V_{zz}\Omega_u) = N(0,\Omega_{11}^{-1}C_zV_{zz}\Omega_{00}C_z\Omega_{11}^{-1})$$

If $\eta = \kappa_x < 1$, (91), (93) and Lemma 3.6 in Phillips and Magdalinos (2009) give

$$T^{(1+\kappa_x)/2}\left(\check{\beta}_{IVX}-\beta\right) = \left[\frac{1}{T^{1+\kappa_x}}\sum_{t=2}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{T^{(1+\kappa_x)/2}}\sum_{t=2}^{T}\left(Z_{t-1}u_{0,t}^{\mu}-\hat{\Lambda}_{01}\right)$$
$$\Rightarrow V_{xz}^{\prime -1}C_x^{-1}N\left(0,\int_0^{\infty} e^{sC_z}(C_xV_{xz}C_z+C_zV_{xz}^{\prime}C_x)e^{sC_z}ds\Omega_u\right).$$

Finally, if $\eta > \kappa_x$, by (93), (69), (95) and Lemma 3.5 in Phillips and Magdalinos (2009), we have

$$T^{(1+\kappa_x)/2}\left(\check{\beta}_{IVX}-\beta\right) = \left[\frac{1}{T^{1+\kappa_x}}\sum_{t=2}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{T^{(1+\kappa_x)/2}}\sum_{t=2}^{T}\left(Z_{t-1}u_{0,t}^{\mu}-\hat{\Lambda}_{01}\right)$$
$$\Rightarrow V_{xx}^{-1}N\left(0,V_{xx}\Omega_u\right) = N\left(0,\Omega_uV_{xx}^{-1}\right).$$

For the Wald statistic, since $T\bar{Z}_{t-1}\bar{Z}_{t-1}^{\prime}\check{\Omega}_{FM}$ in $\check{M}$ is used for finite sample correction that is asymptotically negligible, we treat $\check{M} = \left[\sum_{t=2}^{T} Z_{t-1}Z_{t-1}^{\prime}\right]\hat{\Omega}_{00}$ to shorten the proof. From Equation (14), Lemma 3.5, 3.6 in Phillips and Magdalinos (2009) an by

[2]The joint convergence is proven by Proposition A1. in Phillips and Magdalinos (2009).

$\hat{\Omega}_{00} \xrightarrow{p} \Omega_u$, we have

$$\frac{\check{M}}{T^{1+\min\{\eta,\kappa_x\}}} \xrightarrow{p} \begin{cases} V_{zz}\Omega_u, \text{ if } \eta < \kappa_x, \\ \int_0^\infty e^{sC_z}(C_x V_{xz} C_z + C_z V_{xz}' C_x) e^{sC_z} ds \Omega_u, \text{ if } \eta = \kappa_x, \\ V_{xx}\Omega_u, \text{ if } \eta > \kappa_x. \end{cases}$$

Therefore,

$$\begin{aligned} & \frac{1}{T^{1+\min\{\eta,\kappa_x\}}} \check{Q}_{IVX} \\ = & \left[\frac{1}{T^{1+\min\{\eta,\kappa_x\}}} \sum_{t=2}^T Z_{t-1} X_{t-1}^{\mu\prime}\right]^{-1} \frac{M}{T^{1+\min\{\eta,\kappa_x\}}} \left[\frac{1}{T^{1+\min\{\eta,\kappa_x\}}} \sum_{t=2}^T X_{t-1}^{\mu} Z_{t-1}'\right]^{-1} \\ \Rightarrow & \begin{cases} \Omega_{11} C_z^{-1} V_{zz} \Omega_u C_z \Omega_{11}^{-1}, \text{ if } \eta < \kappa_x, \\ V_{xz}^{\prime -1} C_x^{-1} \int_0^\infty e^{sC_z}(C_x V_{xz} C_z + C_z V_{xz}' C_x) e^{sC_z} ds \Omega_u C_x^{-1} V_{xz}^{-1}, \text{ if } \eta = \kappa_x, \\ \Omega_u V_{xx}^{-1}, \text{ if } \eta > \kappa_x. \end{cases} \end{aligned}$$

We have

$$\begin{aligned} & W_{\check{\beta}_{IVX}} \\ = & \left(R\check{\beta}_{IVX} - r\right)' \check{Q}_{IVX}^{-1} \left(R\check{\beta}_{IVX} - r\right) \\ = & \left(T^{(1+\min\{\eta,\kappa_x\})/2} R\left(\check{\beta}_{IVX} - \beta\right)\right)' \left(R \frac{\check{Q}_{IVX}}{T^{1+\min\{\eta,\kappa_x\}}} R'\right)^{-1} \left(T^{(1+\min\{\eta,\kappa_x\})/2} \left(\check{\beta}_{IVX} - \beta\right)\right) \\ \Rightarrow & \chi^2(q). \end{aligned}$$

(ii). Note that

$$\check{\beta}_{IVX} - \beta = \hat{\beta}_{IVX} - \beta - \left(\sum_{t=2}^T Z_{t-1} X_{t-1}^{\mu\prime}\right)^{-1} T\hat{\Lambda}_{01}. \tag{143}$$

We consider the stochastic order of $\left(\sum_{t=2}^T Z_{t-1} X_{t-1}^{\mu\prime}\right)^{-1} T\hat{\Lambda}_{01}$.For $\hat{\Lambda}_{01}$, from (112), we can deduce that $\hat{\Lambda}_{01} = O_p(M_T)$. From (111), we have $\sum_{t=1}^T Z_{t-1} X_{t-1}^{\mu\prime} = O_p(T^{1+\min\{\eta,\kappa_x\}})$. Using the results in Theorem 3.1(ii), we have

$$\begin{aligned} & T^{\min\{\eta,\kappa_x\}-\eta} \left(\check{\beta}_{IVX} - \beta\right) \\ = & \left(\hat{\beta}_{IVX} - \beta\right) - T^{\min\{\eta,\kappa_x\}-\eta} \left(\sum_{t=2}^T Z_{t-1} X_{t-1}^{\mu\prime}\right)^{-1} T\hat{\Lambda}_{01} \end{aligned}$$

$$\begin{aligned}
&= T^{\min\{\eta,\kappa_x\}-\eta}\left(\hat{\beta}_{IVX}-\beta\right) - T^{\min\{\eta,\kappa_x\}-\eta} O_p\left(T^{-1-\min\{\eta,\kappa_x\}}\right) T O_p(M_T) \\
&= T^{\min\{\eta,\kappa_x\}-\eta}\left(\hat{\beta}_{IVX}-\beta\right) - O_p\left(T^{1/3-\eta}\right) \\
&= T^{\min\{\eta,\kappa_x\}-\eta}\left(\hat{\beta}_{IVX}-\beta\right) + o_p(1),
\end{aligned}$$

where we have the last equality because $M_T = T^{1/3}$ and $\eta, \kappa_x > 1/3$. This implies that the stochastic order of $\check{\beta}_{IVX} - \beta$ is that of $\hat{\beta}_{IVX} - \beta$ in Theorem 3.1.2. For the Wald statistic, from (42), we have $\hat{\gamma}_h = O_p(T)$. Thus, $\hat{\Omega}_{00} = O_p(TM_T)$. And

$$\begin{aligned}
\check{\Omega}_{FM} &= O_p(TM_T) + O_p(M_T)O_p(1)O_p(M_T) \\
&= O_p(TM_T). \qquad (144)
\end{aligned}$$

If $\eta < \kappa_x$, (128) and $\check{\Omega}_{FM} = O_p(TM_T)$ imply that $T\bar{Z}_{t-1}\bar{Z}'_{t-1}\check{\Omega}_{FM} = O_p(T^{2\eta})O_p(TM_T) = O_p(T^{2\eta+4/3})$, and

$$\begin{aligned}
\check{M} &= \left[\sum_{t=2}^{T} Z_{t-1}Z'_{t-1}\right]\hat{\Omega}_{00} - T\bar{Z}_{t-1}\bar{Z}'_{t-1}\check{\Omega}_{FM} \\
&= O_p(T^{1+\eta})O_p(T^{4/3}) + O_p(T^{2\eta+4/3}) \\
&= O_p(T^{1+\eta+4/3}).
\end{aligned}$$

Eventually,

$$\begin{aligned}
\check{Q}_{IVX} &= R\left[\sum_{t=2}^{T} Z_{t-1}X^{\mu\prime}_{t-1}\right]^{-1}\check{M}\left[\sum_{t=2}^{T} X^{\mu}_{t-1}Z'_{t-1}\right]^{-1}R' \\
&= O_p(T^{-2-2\eta})O_p(T^{1+\eta+4/3}) \\
&= O_p\left(T^{1/3-\eta}\right),
\end{aligned}$$

and

$$\begin{aligned}
W_{\check{\beta}_{IVX}} &\equiv \left(R\check{\beta}_{IVX} - r\right)'\check{Q}^{-1}_{IVX}\left(R\check{\beta}_{IVX} - r\right) \\
&= O_p(1)O_p\left(\frac{T^{\eta}}{M_T}\right) = O_p\left(T^{\eta-1/3}\right) \xrightarrow{p} \infty.
\end{aligned}$$

If $\eta \geq \kappa_x$, (128) and $\check{\Omega}_{FM} = O_p(TM_T)$ imply that

$$T\bar{Z}_{t-1}\bar{Z}'_{t-1}\check{\Omega}_{FM} = O_p(T^{2\kappa_x+\eta-1})O_p(TM_T) = O_p(T^{2\kappa_x+\eta-1/3}).$$

Thus

$$\begin{aligned}\check{Q}_{IVX} &= O_p(T^{-2(1+\min\{\kappa_x,\eta\})})O_p(T^{2\kappa_x+\eta-1/3}) \\ &= O_p(T^{-2-2\kappa_x})O_p(T^{2\kappa_x+\eta-1/3}) = O_p(T^{\eta-5/3})\end{aligned}$$

and

$$\begin{aligned}W_{\check{\beta}_{IVX}} &= O_p(T^{2\eta-2\min\{\eta,\kappa_x\}})O_p(T^{5/3-\eta}) \\ &= O_p(T^{2\eta-2\kappa_x})O_p(T^{5/3-\eta}) \\ &= O_p(T^{\eta+5/3-2\kappa_x}).\end{aligned}$$

(iii). Note that $\check{\beta}_{IVX} - \beta = \hat{\beta}_{IVX} - \beta - \left(\sum_{t=1}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right)^{-1} T\hat{\Lambda}_{01}$. Similar to (126), we can deduce that $\hat{\Lambda}_{01} = O_p(M_T)$. Thus, $T\hat{\Lambda}_{01} = O_p(TM_T)$ and

$$\left(\sum_{t=1}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right)^{-1} T\hat{\Lambda}_{01} = O_p(T^{-\min\{\kappa_x,\eta\}-1})O_p(TM_T) =$$

$O_p\left(\frac{M_T}{T^{\min\{\kappa_x,\eta\}}}\right)$. From Theorem 3.1(iii), $\hat{\beta}_{IVX} - \beta = O_p\left(T^{\min\{\kappa_u,\eta\}-\min\{\kappa_x,\eta\}}\right)$, we can write

$$\begin{aligned}&T^{\min\{\kappa_x,\eta\}-\min\{\eta,\kappa_u\}}\left(\check{\beta}_{IVX}-\beta\right) \\ =\ & T^{\min\{\kappa_x,\eta\}-\min\{\eta,\kappa_u\}}\left(\hat{\beta}_{IVX}-\beta\right) - T^{\min\{\kappa_x,\eta\}-\min\{\eta,\kappa_u\}}\left(\sum_{t=1}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right)^{-1} T\hat{\Lambda}_{01} \\ =\ & T^{\min\{\kappa_x,\eta\}-\min\{\eta,\kappa_u\}}\left(\hat{\beta}_{IVX}-\beta\right) - T^{\min\{\kappa_x,\eta\}-\min\{\eta,\kappa_u\}}O_p\left(\frac{M_T}{T^{\min\{\kappa_x,\eta\}}}\right) \\ =\ & T^{\min\{\kappa_x,\eta\}-\min\{\eta,\kappa_u\}}\left(\hat{\beta}_{IVX}-\beta\right) - O_p\left(\frac{M_T}{T^{\min\{\kappa_u,\eta\}}}\right) \\ =\ & T^{\min\{\kappa_x,\eta\}-\min\{\eta,\kappa_u\}}\left(\hat{\beta}_{IVX}-\beta\right) + o_p(1). \qquad (145)\end{aligned}$$

For the Wald statistic, from (63), we can deduce that $\hat{\Omega}_{00} = O_p(M_T T^{\kappa_u})$, and thus

$$\left[\sum_{t=2}^{T} Z_{t-1}Z_{t-1}'\right]\hat{\Omega}_{00} = O_p(T^{1+\min\{\kappa_x,\eta\}})O_p(M_T T^{\kappa_u}) = O_p(T^{1+\kappa_u+\min\{\kappa_x,\eta\}+1/3}).$$

For $\check{\Omega}_{FM}$, similar to $\hat{\Omega}_{FM}$ in (127), we can deduce

$$\check{\Omega}_{FM} = O_p(M_T T^{\kappa_u}) + O_p(M_T^2) = O_p(M_T T^{\kappa_u}) = O_p(T^{\kappa_u+1/3}). \qquad (146)$$

(146) and (128) jointly imply

$$T\bar{Z}_{t-1}\bar{Z}'_{t-1}\check{\Omega}_{FM} = \begin{cases} O_p(T^{2\eta+\kappa_u+1/3}), \text{ if } \eta < \kappa_x, \\ O_p(T^{2\kappa_x+\eta+\kappa_u-2/3}), \text{ if } \eta \geq \kappa_x. \end{cases} \tag{147}$$

Therefore, if $\eta < \kappa_x$, we have

$$\begin{aligned} \check{M} &= \left[\sum_{t=2}^{T} Z_{t-1}Z'_{t-1}\right]\hat{\Omega}_{00} - T\bar{Z}_{t-1}\bar{Z}'_{t-1}\check{\Omega}_{FM} \\ &= O_p(T^{1+\kappa_u+\eta+1/3}) + O_p(T^{2\eta+\kappa_u+1/3}) \\ &= O_p(T^{1+\kappa_u+\eta+1/3}), \end{aligned}$$

and

$$\begin{aligned} \check{Q}_{IVX} &= R\left[\sum_{t=2}^{T} Z_{t-1}X^{\mu\prime}_{t-1}\right]^{-1}\check{M}\left[\sum_{t=2}^{T} X^{\mu}_{t-1}Z'_{t-1}\right]^{-1}R' \\ &= O_p(T^{-2-2\eta})O_p(T^{1+\kappa_u+\eta+1/3}) \\ &= O_p(T^{\kappa_u-\eta-2/3}). \end{aligned}$$

From Theorem 3.1.(iii), we have

$$\begin{aligned} W_{\check{\beta}_{IVX}} &= \left(R\check{\beta}_{IVX} - r\right)'\check{Q}^{-1}_{IVX}\left(R\check{\beta}_{IVX} - r\right) \\ &= O_p\left(T^{2\min\{\kappa_u,\eta\}-2\eta}\right)O_p(T^{2/3-\kappa_u+\eta}) \\ &= O_p\left(T^{2\min\{\kappa_u,\eta\}-\eta+2/3-\kappa_u}\right). \end{aligned}$$

If $\eta \geq \kappa_x$, (115), (146) and (147) give

$$\begin{aligned} \check{M} &= \left[\sum_{t=2}^{T} Z_{t-1}Z'_{t-1}\right]\hat{\Omega}_{00} - T\bar{Z}_{t-1}\bar{Z}'_{t-1}\check{\Omega}_{FM} \\ &= O_p(T^{1+\kappa_x})O_p(T^{\kappa_u+1/3}) + O_p(T^{2\kappa_x+\eta+\kappa_u-2/3}) \\ &= O_p(T^{1+\kappa_x})O_p(T^{\kappa_u+1/3}) \\ &= O_p(T^{4/3+\kappa_x+\kappa_u}), \end{aligned}$$

and

$$\check{Q}_{IVX} = R\left[\sum_{t=2}^{T} Z_{t-1}X^{\mu\prime}_{t-1}\right]^{-1}\check{M}\left[\sum_{t=2}^{T} X^{\mu}_{t-1}Z'_{t-1}\right]^{-1}R'$$

$$
\begin{aligned}
&= O_p(T^{-2-2\kappa_x})O_p(T^{4/3+\kappa_x+\kappa_u}) \\
&= O_p(T^{\kappa_u-2/3-\kappa_x}).
\end{aligned}
$$

Eventually,

$$
\begin{aligned}
W_{\check{\beta}_{IVX}} &= O_p\left(T^{2\min\{\kappa_u,\eta\}-2\min\{\eta,\kappa_x\}}\right) O_p(T^{2/3+\kappa_x-\kappa_u}) \\
&= O_p\left(T^{2\min\{\kappa_u,\eta\}-2\kappa_x}\right) O_p(T^{2/3+\kappa_x-\kappa_u}) \\
&= O_p\left(T^{2\min\{\kappa_u,\eta\}-\kappa_x+2/3-\kappa_u}\right).
\end{aligned}
$$

It can be directly verified that $W_{\check{\beta}_{IVX}} \xrightarrow{p} \infty$.

(iv). Given the consistency of $\hat{\beta}$ as shown in (31), we can show $\hat{\Lambda}_{01} \xrightarrow{p} \Lambda_{01}$. Applying Lemma 2.4 of Phillips and Lee (2016), we can show

$$
\left[\left(\sum_{t=2}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right)^{-1} T\hat{\Lambda}_{01}\right] = O_p\left(\frac{\Phi_T^{-2T}}{T^{\kappa_x+\min\{\kappa_x,\eta\}-1}}\right) = o_p(1).
$$

Therefore, $\check{\beta}_{IVX} - \beta = \hat{\beta}_{IVX} - \beta + o_p(1)$. The result in Theorem 3.1.4 is applicable. For the Wald statistic, since we use $\hat{\Omega}_{00}$ to estimate Ω_u and $\hat{\Omega}_{00} \xrightarrow{p} \Omega_u$, following the proof of (131), we can easily show $W_{\check{\beta}_{IVX}} \Rightarrow \chi^2(q)$.

(v). As in (43) and (113), we can show $\hat{\Omega}_{00} = O_p(M_T T) = O_p(T^{4/3})$, $\hat{\Omega}_{01} = O_p(M_T) = O_p(T^{1/3})$, and $\hat{\Lambda}_{01} = O_p(T^{1/3})$ given $\hat{\Omega}_{01} = O_p(T^{1/3})$. Applying Lemma 2.4 in Phillips and Lee (2016), we can show the bias correction term has the following stochastic order

$$
\left(\sum_{t=2}^{T} Z_{t-1}X_{t-1}^{\mu\prime}\right)^{-1} T\hat{\Lambda}_{01} = O_p(\Phi_T^{-2T}T^{1+1/3-\kappa_x-\min\{\kappa_x,\eta\}}) = o_p(1).
$$

Again, $\check{\beta}_{IVX} - \beta = \hat{\beta}_{IVX} - \beta + o_p(1)$. Thus, the order of $\check{\beta}_{IVX} - \beta$ can be established using Theorem 3.1.4. For the Wald statistic, given $\hat{\Omega}_{00} = O_p(T^{4/3})$ and $\frac{1}{T^{2\min\{\kappa_x,\eta\}}}\sum_{t=2}^{T}\Phi_T^{-T}Z_{t-1}Z_{t-1}'\Phi_T^{-T} = O_p(1)$, by Lemma A2 of Phillips and Lee (2016), we have

$$
\left[\sum_{t=1}^{T} Z_{t-1}Z_{t-1}'\right]\hat{\Omega}_{00} = O_p(\Phi_T^{2T}T^{4/3+2\min\{\kappa_x,\eta\}}). \tag{148}
$$

For $T\bar{Z}_{t-1}\bar{Z}_{t-1}'\check{\Omega}_{FM}$, note that $\check{\Omega}_{FM} = O_p(TM_T)$ can be shown using the steps that prove (144). From (135), $\bar{Z}_{t-1} = O_p(\Phi_T^{T}T^{\kappa_x/2+\min\{\eta,\kappa_x\}-1})$. Thus,

$$
T\bar{Z}_{t-1}\bar{Z}_{t-1}'\check{\Omega}_{FM} = TO_p(\Phi_T^{2T}T^{\kappa_x+2\min\{\eta,\kappa_x\}-2})O_p(TM_T)
$$

$$= O_p(\Phi_T^{2T} T^{\kappa_x + 2\min\{\eta,\kappa_x\}+1/3}),$$

since $\kappa_x < 1$, $T\bar{Z}_{t-1}\bar{Z}'_{t-1}\check{\Omega}_{FM}$ is asymptotically dominated by $\left[\sum_{t=1}^{T} Z_{t-1}Z'_{t-1}\right]\hat{\Omega}_{00}$. Using Lemma 2.4 of Phillips and Lee (2016) and (148), we can obtain

$$\begin{aligned}
& T^{2\kappa_x - 4/3}\Phi_T^T \check{Q}_{IVX}\Phi_T^T \\
= \; & T^{2\kappa_x-4/3}\Phi_T^T R\left[\sum_{t=2}^{T} Z_{t-1}X^{\mu\prime}_{t-1}\right]^{-1}\frac{\check{M}}{T^{4/3}}\left[\sum_{t=2}^{T} X^{\mu}_{t-1}Z'_{t-1}\right]^{-1} R'\Phi_T^T \\
= \; & R\left[\frac{1}{T^{\kappa_x+\min\{\kappa_x,\eta\}}}\sum_{t=2}^{T}\Phi_T^{-T}Z_{t-1}X^{\mu\prime}_{t-1}\Phi_T^{-T}\right]^{-1}\frac{\Phi_T^{-T}\check{M}\Phi_T^{-T}}{T^{2\min\{\kappa_x,\eta\}+4/3}} \\
& \left[\frac{1}{T^{\kappa_x+\min\{\kappa_x,\eta\}}}\sum_{t=2}^{T}\Phi_T^{-T}X^{\mu}_{t-1}Z'_{t-1}\Phi_T^{-T}\right]^{-1} R' \\
= \; & O_p(1).
\end{aligned}$$

Thus,

$$\begin{aligned}
& \frac{1}{T^{3\kappa_x-2\min\{\kappa_x,\eta\}-1/3}} W_{\check{\beta}_{IVX}} \\
= \; & \frac{1}{T^{3\kappa_x-2\min\{\kappa_x,\eta\}-1/3}}\left(R\left(\hat{\beta}_{IVX}-\beta\right)\right)'\left[\check{Q}_{IVX}\right]^{-1}\left(R\left(\hat{\beta}_{IVX}-\beta\right)\right) \\
= \; & \left(R\frac{\Phi_T^T}{T^{(\kappa_x+1)/2-\min\{\kappa_x,\eta\}}}\left(\hat{\beta}_{IVX}-\beta\right)\right)'\left[T^{2\kappa_x-4/3}\Phi_T^T\check{Q}_{IVX}\Phi_T^T\right]^{-1} \\
& \times\left(R\frac{\Phi_T^T}{T^{(\kappa_x+1)/2-\min\{\kappa_x,\eta\}}}\left(\hat{\beta}_{IVX}-\beta\right)\right) \\
= \; & O_p(1).
\end{aligned}$$

Eventually, $W_{\check{\beta}_{IVX}} = O_p(T^{3\kappa_x-2\min\{\kappa_x,\eta\}-1/3})$.

(vi). As in (64) and (126), we can show $\hat{\Omega}_{00} = O_p(M_T T^{\kappa_u}) = O_p(T^{\kappa_u+1/3})$ and $\hat{\Omega}_{01} = O_p(M_T) = O_p(T^{1/3})$, thus $\hat{\Lambda}_{01} = O_p(T^{1/3})$. The bias correction term has the stochastic order

$$\left(\sum_{t=2}^{T} Z_{t-1}X^{\mu\prime}_{t-1}\right)^{-1} T\hat{\Lambda}_{01}O_p(\Phi_T^{-2T}T^{1+1/3-\kappa_x-\min\{\kappa_x,\eta\}}) = o_p(1).$$

Thus, the bias correction term vanishes as $T\to\infty$. As $\check{\beta}_{IVX}-\beta = \hat{\beta}_{IVX}-\beta+o_p(1)$, we obtain the order of $\check{\beta}_{IVX} - \beta$. For the Wald statistic, note that

$$\left[\sum_{t=2}^{T} Z_{t-1} Z_{t-1}'\right] \hat{\Omega}_{00} = O_p(\Phi_T^{2T} T^{\kappa_u + 1/3 + 2\min\{\kappa_x, \eta\}}).$$

We can deduce $\check{\Omega}_{FM} = O_p(T^{\kappa_u + 1/3})$ as in (146). Thus,

$$\begin{aligned} T\bar{Z}_{t-1}\bar{Z}_{t-1}'\check{\Omega}_{FM} &= TO_p(\Phi_T^{2T} T^{\kappa_x + 2\min\{\eta,\kappa_x\} - 2}) O_p(T^{\kappa_u + 1/3}) \\ &= O_p(\Phi_T^{2T} T^{\kappa_x + \kappa_u + 2\min\{\eta,\kappa_x\} - 2/3}). \end{aligned}$$

Since $\kappa_x < 1$, $T\bar{Z}_{t-1}\bar{Z}_{t-1}'\check{\Omega}_{FM}$ is asymptotically dominated by $\left[\sum_{t=2}^{T} Z_{t-1} Z_{t-1}'\right] \hat{\Omega}_{00}$ and thus $\check{M} = O_p(\Phi_T^{2T} T^{\kappa_u + 1/3 + 2\min\{\kappa_x, \eta\}})$. Hence, we can express

$$\begin{aligned} & T^{2\kappa_x - \kappa_u - 1/3} \Phi_T^T \check{Q}_{IVX} \Phi_T^T \\ = \ & T^{2\kappa_x - \kappa_u - 1/3} \Phi_T^T R \left[\sum_{t=2}^{T} Z_{t-1} X_{t-1}^{\mu\prime}\right]^{-1} \check{M} \left[\sum_{t=2}^{T} X_{t-1}^{\mu} Z_{t-1}'\right]^{-1} R' \Phi_T^T \\ = \ & R \left[\frac{1}{T^{\kappa_x + \min\{\kappa_x,\eta\}}} \sum_{t=2}^{T} \Phi_T^{-T} Z_{t-1} X_{t-1}^{\mu\prime} \Phi_T^{-T}\right]^{-1} \frac{\Phi_T^{-T} \check{M} \Phi_T^{-T}}{T^{\kappa_u + 1/3 + 2\min\{\kappa_x,\eta\}}} \\ & \times \left[\frac{1}{T^{\kappa_x + \min\{\kappa_x,\eta\}}} \sum_{t=2}^{T} \Phi_T^{-T} X_{t-1}^{\mu} Z_{t-1}' \Phi_T^{-T}\right]^{-1} R' \\ = \ & O_p(1). \end{aligned}$$

If $\kappa_x \geq \kappa_u \geq \eta$, we have

$$\begin{aligned} \frac{1}{T^{4\kappa_x + \kappa_u - 2\eta - 1/3}} W_{\check{\beta}_{IVX}} &= \frac{1}{T^{4\kappa_x + \kappa_u - 2\eta - 1/3}} \left(R\left(\hat{\beta}_{IVX} - \beta\right)\right)' \check{Q}_{IVX}^{-1} \left(R\left(\hat{\beta}_{IVX} - \beta\right)\right) \\ &= \left(R \frac{\Phi_T^T}{T^{\kappa_x + \kappa_u - \eta}} \left(\hat{\beta}_{IVX} - \beta\right)\right)' \left[T^{2\kappa_x - \kappa_u - 1/3} \Phi_T^T \check{Q}_{IVX} \Phi_T^T\right]^{-1} \\ &\quad \left(R \frac{\Phi_T^T}{T^{\kappa_x + \kappa_u - \eta}} \left(\hat{\beta}_{IVX} - \beta\right)\right) \\ &= O_p(1). \end{aligned}$$

If $\kappa_x \geq \kappa_u \geq \eta$ does not hold, we have

$$\begin{aligned} & \frac{1}{T^{5\kappa_x + 2\min\{\kappa_u,\eta\} - 1/3 - 4\min\{\kappa_x,\eta\}}} W_{\check{\beta}_{IVX}} \\ = \ & \frac{1}{T^{5\kappa_x + 2\min\{\kappa_u,\eta\} - 1/3 - 4\min\{\kappa_x,\eta\}}} \left(R\left(\hat{\beta}_{IVX} - \beta\right)\right)' \check{Q}_{IVX}^{-1} \left(R\left(\hat{\beta}_{IVX} - \beta\right)\right) \end{aligned}$$

$$
\begin{aligned}
&= \left(R\Phi_T^T \frac{T^{2\min\{\kappa_x,\eta\}}}{T^{(\kappa_u+3\kappa_x)/2+\min\{\kappa_u,\eta\}}}\left(\hat{\beta}_{IVX}-\beta\right)\right)' \left[T^{2\kappa_x-\kappa_u-1/3}\Phi_T^T\check{Q}_{IVX}\Phi_T^T\right]^{-1} \\
&\quad \left(R\Phi_T^T \frac{T^{2\min\{\kappa_x,\eta\}}}{T^{(\kappa_u+3\kappa_x)/2+\min\{\kappa_u,\eta\}}}\left(\hat{\beta}_{IVX}-\beta\right)\right) \\
&= O_p(1).
\end{aligned}
$$

Therefore, we have

$$
W_{\breve{\beta}_{IVX}} = \begin{cases} O_p\left(T^{4\kappa_x+\kappa_u-2\eta-1/3}\right), & \text{if } \kappa_x \geq \kappa_u \geq \eta, \\ O_p\left(T^{5\kappa_x+2\min\{\kappa_u,\eta\}-1/3-4\min\{\kappa_x,\eta\}}\right), & \text{otherwise} \end{cases} .
$$

This completes the proof of Theorem 3.1 and Lemma A.5. ■

We are now ready to show the limiting distribution of $W_{\tilde{\beta}_{IVX}}$ under the null hypothesis that $\beta = 0$.

Proof of Theorem 4.1. We first prove this theorem under the case that $\Phi_T = I_k + \frac{C_x}{T^{\kappa_x}}, C_x \leq 0, \kappa_x \in (0,1]$ and $\rho_{1,T} = 1 + \frac{c_u}{T^{\kappa_u}}$, $c_u < 0, \kappa_u \in (0,1]$. Applying the results from Lemma A.2 and A.3(i), suppose that $\min\{\eta,\kappa_x\} < 2\kappa_u$,

$$
\begin{aligned}
\sqrt{T}\tilde{\beta}_{IVX} &= \left[\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}\tilde{X}_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}z_{0,t}^{\mu} + o_p(1) \\
&\Rightarrow E[w_{\rho,t}w_{\rho,t}']^{-1}N(0,\sigma_z^2E[w_{\rho,t}w_{\rho,t}']) \\
&= N(0,\sigma_z^2E[w_{\rho,t}w_{\rho,t}']^{-1}).
\end{aligned}
$$

For the Wald statistic $W_{\tilde{\beta}_{IVX}}$,

$$
W_{\tilde{\beta}_{IVX}} = \left(\sqrt{T}\tilde{\beta}_{IVX}\right)'[TQ_{IVX}]^{-1}\left(\sqrt{T}\tilde{\beta}_{IVX}\right).
$$

Note that

$$
\tilde{M} = \left[\sum_{t=p+1}^{T}\tilde{Z}_{t-1}\tilde{Z}_{t-1}'\right]\tilde{\Omega}_{00} - T\breve{Z}_{t-1}\breve{Z}_{t-1}'\tilde{\Omega}_{FM},
$$

since $T\breve{Z}_{t-1}\breve{Z}_{t-1}'\tilde{\Omega}_{FM}$ is asymptotically dominated by $\left[\sum_{t=p+1}^{T}\tilde{Z}_{t-1}\tilde{Z}_{t-1}'\right]\tilde{\Omega}_{00}$ as shown in Equation (33) in the online appendix of Kostakis et al. (2015), to simplify the proof, we assume

$$
\tilde{M} = \left[\sum_{t=p+1}^{T}\tilde{Z}_{t-1}\tilde{Z}_{t-1}'\right]\tilde{\Omega}_{00.}
$$

Then, we have

$$TQ_{IVX} = \left[\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}\tilde{X}_{t-1}^{\mu\prime}\right]^{-1}\left[\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}\tilde{Z}_{t-1}'\right]\tilde{\Omega}_{00}\left[\frac{1}{T}\sum_{t=p+1}^{T}\tilde{X}_{t-1}^{\mu}\tilde{Z}_{t-1}'\right]^{-1}$$
$$\xrightarrow{p} E[w_{\rho,t}w_{\rho,t}']^{-1}E[w_{\rho,t}w_{\rho,t}']\sigma_z^2 E[w_{\rho,t}w_{\rho,t}']^{-1}$$
$$= \sigma_z^2 E[w_{\rho,t}w_{\rho,t}']^{-1}.$$

From Slutsky's theorem, we have $W_{\tilde{\beta}_{IVX}} = \left(\sqrt{T}\tilde{\beta}_{IVX}\right)'[TQ_{IVX}]^{-1}\left(\sqrt{T}\tilde{\beta}_{IVX}\right) \Rightarrow \chi^2(k)$.

If $\min\{\eta,\kappa_x\} = 2\kappa_u$ and $\kappa_x > \eta$, we have

$$\sqrt{T}\tilde{\beta}_{IVX} = \left[\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}\tilde{X}_{t-1}^{\mu\prime}\right]^{-1}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}z_{0,t}^{\mu} \tag{149}$$
$$\Rightarrow \left[E[w_{\rho,t}w_{\rho,t}'] - c_u^2 C_z^{-1}\left(\int_0^1 J_{C_x}dB_x' + \Omega_{xx}\right)'\right]^{-1}\left(B_{w_\rho}(1) - c_u U_z(1)\right),$$

and

$$TQ_{IVX} = \left[\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}\tilde{X}_{t-1}^{\mu\prime}\right]^{-1}\left[\frac{1}{T}\sum_{t=p+1}^{T}\tilde{Z}_{t-1}\tilde{Z}_{t-1}'\right]\tilde{\Omega}_{00}\left[\frac{1}{T}\sum_{t=p+1}^{T}\tilde{X}_{t-1}^{\mu}\tilde{Z}_{t-1}'\right]^{-1}$$
$$\Rightarrow \left[E[w_{\rho,t}w_{\rho,t}'] - c_u^2 C_z^{-1}\left(\int_0^1 J_{C_x}dB_x' + \Omega_{xx}\right)'\right]^{-1} \tag{150}$$
$$\times\left[\sigma_z^2 E[w_{\rho,t}w_{\rho,t}'] + c_u^2\sigma_z^2 V_{zz}^x\right]\left[E[w_{\rho,t}w_{\rho,t}'] - \left(\int_0^1 J_{C_x}dB_x' + \Omega_{xx}\right)c_u^2 C_z^{-1}\right]^{-1}.$$

Note that the variance of $B_{w_\rho}(1) - c_u U_z(1)$ is $\sigma_z^2 E[w_{\rho,t}w_{\rho,t}'] + c_u^2\sigma_z^2 V_{zz}^x$. The joint convergence of (150) and (151) gives $W_{\tilde{\beta}_{IVX}} \Rightarrow \chi^2(k)$. The claims for $\kappa_x < \eta$ and $\kappa_x = \eta$ and the case under $\min\{\eta,\kappa_x\} > 2\kappa_u$ can be established using the similar arguments and the results from Lemma A.3(i).

We now consider the case when $\Phi_T = I_k + \frac{C_x}{T^{\kappa_x}}, C_x > 0, \kappa_x \in (0.5, 1)$ and $\rho_{1,T} = 1 + \frac{c_u}{T^{\kappa_u}}$, $c_u < 0, \kappa_u \in (0,1)$. If $\kappa_x = \kappa_u$, applying the results from A.3(ii) and proposition 2.4 in Phillips and Lee (2016), we have

$$\Phi_T^T\tilde{\beta}_{IVX} = \left[\frac{T^{2\kappa_x}}{T^{\kappa_x+\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_{t-1}\tilde{X}_{t-1}^{\mu\prime}\Phi_T^{-T}\right]^{-1}\frac{T^{\kappa_x}}{T^{\min\{\eta,\kappa_x\}}}\sum_{t=p+1}^{T}\Phi_T^{-T}\tilde{Z}_{t-1}z_{0,t}^{\mu}$$
$$\Rightarrow [C_x C_{z,\kappa_x,\eta}W_{C_x}]^{-1}C_x C_{z,\kappa_x,\eta}\times MN(0,\sigma_z^2 W_{C_x}). \tag{151}$$

Moreover, we have

$$
\begin{aligned}
&\Phi_T^T Q_{IVX} \Phi_T^T \\
= &\left[\frac{T^{2\kappa_x}}{T^{\kappa_x+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} \tilde{Z}_{t-1} \tilde{X}_{t-1}^{\mu\prime} \Phi_T^{-T}\right]^{-1} \left[\frac{T^{2\kappa_x}}{T^{2\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} \tilde{Z}_{t-1} \tilde{Z}_{t-1}^{\prime} \Phi_T^{-T}\right] \tilde{\Omega}_{00} \\
&\times \left[\frac{T^{2\kappa_x}}{T^{\kappa_x+\min\{\eta,\kappa_x\}}} \sum_{t=p+1}^{T} \Phi_T^{-T} \tilde{X}_{t-1}^{\mu} \tilde{Z}_{t-1}^{\prime} \Phi_T^{-T}\right]^{-1} \\
\Rightarrow &\ [C_x C_{z,\kappa_x,\eta} W_{C_x}]^{-1} [C_x C_{z,\kappa_x,\eta} W_{C_x} C_{z,\kappa_x,\eta} C_x] \sigma_z^2 [W_{C_x} C_x C_{z,\kappa_x,\eta}]^{-1}.
\end{aligned} \tag{152}
$$

The joint convergence of (151) and (152) implies $W_{\tilde{\beta}_{IVX}} \Rightarrow \chi^2(k)$. For the cases where $\kappa_x > \kappa_u$ and $\kappa_x < \kappa_u$, applying the results from A.3.(ii), we can replace W_{C_x} in (151) and (152) by $W_{C_x}^{(4)}$ and $W_{C_x}^{(1)}$, respectively. By a similar argument, we have $W_{\tilde{\beta}_{IVX}} \Rightarrow \chi^2(k)$ under these two cases. ■

C. Limit Theory for Bonferroni Confidence Interval and Cauchy Estimator

C.1 Bonferroni confidence interval

Consider the case when X_t is a scalar time series. Cavanagh et al. (1995) proposed to use the Bonferroni confidence interval (CI) for β under the assumption that X_t is a local-to-unit root (LUR) time series. As the localizing parameter in the AR root can not be consistently estimated, the Bonferroni method, which has an extra layer of confidence interval associated with the localizing parameter of the regressor (C_x), is often used. This approach is well discussed in the context of predictive regression in Cavanagh et al. (1995), Campbell and Yogo (2006) and Phillips (2014). Campbell and Yogo (2006) and Phillips (2014) show that the Bonferroni corrected confidence interval can have a well-controlled size for the slope parameter in the single-variate predictive regression model (1) when the regressors has an LUR AR root.

To fix idea, we consider the following Bonferroni CI with a nominal coverage of at least $100(1-a)\%$ considered in Cavanagh et al. (1995):

$$
\begin{aligned}
C_\beta(\alpha) &= \bigcup_{c_x \in C_{C_x}(a_1)} C_{\beta|C_x}(a_2), \\
C_{\beta|C_x}(a_2) &= \left\{\rho : c_{1,C_x} \leq t_{\hat{\beta}_{T,\beta}} \leq c_{2,C_x}\right\}
\end{aligned} \tag{153}
$$

where $t_{\hat{\beta}_T,\beta} \equiv \frac{\hat{\beta}_T - \beta}{se(\hat{\beta}_T)}$, $C_{C_x}(a_1)$ is the $100(1-a_1)\%$ confidence interval for C_x, $C_{\rho|C_x}(a_2)$ is the $100(1-a_2)\%$ confidence interval for ρ given C_x, $\alpha = a_1 + a_2$, and c_{1,C_x} and c_{2,C_x} are the $a_2/2$ and $1-a_2/2$ quantiles for the random variable

$$\tau_{C_x} \equiv \delta \tau_{1C_x} + (1-\delta^2)^{1/2} z,$$

where $\delta \equiv \frac{\Omega_{01}}{(\Omega_{11}\Omega_{00})^{1/2}}$, $\tau_{1C_x} \equiv \left(\int_0^1 J_{C_x}^{\mu 2}(r)dr\right)^{-1/2} \int_0^1 J_{C_x}^{\mu}(r)dB_1(r)$, z is a standard normal random variable that is independent of τ_{1C_x}.

In practice, the Bonferonni CI confirms the predictiveness of a regressor when $C_\beta(\alpha)$ excludes zero. The following theorem shows the asymptotic behavior of $C_\beta(\alpha)$ under model (2).

Theorem C.1. *Under the same set of assumptions as in Lemma 2.1. As $T \to \infty$, if $|t_{\hat{\beta}_T,\beta}| \xrightarrow{p} \infty$, we have*

$$\Pr(\beta \in C_\beta(\alpha)) \to 0.$$

Proof of Theorem C.1. Letting $G_{T,c_x}(y) = \Pr(\tau_c \le y)$, we can express the confidence interval $C_{\beta|C_x}(a_2)$ as

$$\begin{aligned} C_{\beta|c_x}(a_2) &= \left\{\beta : G_{T,c_x}(c_{1,c_x}) \le G_{T,c_x}(t_{\hat{\beta}_T,\beta}) \le G_{T,c_x}(c_{2,c_x})\right\} \\ &= \left\{\beta : a_2/2 \le G_{T,c_x}(t_{\hat{\beta}_T,\beta}) \le 1 - a_2/2\right\}. \end{aligned}$$

We have

$$\begin{aligned} G_{T,c_x}(t_{\hat{\beta}_T,\beta}) &= \Pr(\tau_c \le t_{\hat{\beta}_T,\beta}) \\ &= \Pr(0 \le t_{\hat{\beta}_T,\beta} - \tau_c). \end{aligned}$$

Note that $|t_{\hat{\beta}_T,\beta} - \tau_c| \xrightarrow{p} \infty$, as $\tau_c = O(1)$ and $t_{\hat{\beta}_T,\beta}$ diverges. We have either $G_{T,c_x}(t_{\hat{\beta}_T,\beta}) \xrightarrow{p} 1$ or $G_{T,c_x}(t_{\hat{\beta}_T,\beta}) \xrightarrow{p} 0$. Consequently, $C_{\beta|C_x}(a_2)$ asymptotically becomes an empty set as $1 > 1 - a_2/2$ and $0 < a_2/2$, so does the union set $C_\beta(\alpha) = \bigcup_{c_x \in C_{c_x}(a_1)} C_{\beta|c_x}(a_2)$. ■

Theorem C.1 suggests that the probability for the CI to contain the true parameter β shrinks to zero whenever the test statistic diverges. Lemma 2.1 provides the scenarios when W_T (thus $t_{\hat{\beta}_T,\beta}$) diverges. Theorem C.1 shows that the limiting coverage probability of β is not at least $100(1-a)\%$ asymptotically. Moreover, when $\beta = 0$, as $C_\beta(\alpha)$ excludes zero when T is large. Consequently, using the Bonferroni CI constructed as (153) may lead to the spurious conclusion of predictiveness of X_{t-1} under the cases where the test statistic diverges.

C.2 Cauchy Estimator

In addition to the Bonferroni CI, Cauchy estimator has also been proposed as an alternative for the inference in predictive regression with univariate LUR predictor. Choi et al. (2016) show that the test statistic built upon Cauchy estimator asymptotically distributed as a standard normal random variable. As a consequence, a straightforward test for the predictability can be obtained. To be precise, we define the Cauchy estimator of β as

$$\hat{\beta}_{Cau} = \frac{\sum_{t=1}^{T} sgn\left(X_{t-1}\right) y_t}{\sum_{t=1}^{T} \left|X_{t-1}\right|}, \tag{154}$$

where $sgn(.)$ denotes sign function such that $sgn(x) = 1$ if $x \geq 0$, and $sgn(x) = -1$ if $x < 0$.

The t statistic for $\hat{\beta}_{Cau}$ is given as

$$t_{\hat{\beta}_{Cau}} = \frac{\hat{\beta}_{Cau}}{se(\hat{\beta}_{Cau})}, \text{ with } se(\hat{\beta}_{Cau}) = \hat{\Omega}_u^{1/2}\sqrt{T}\left(\sum_{t=1}^{T}\left|X_{t-1}\right|\right)^{-1},$$

where $\hat{\Omega}_u$ is a Newey-West-type estimator with the bandwidth M_T. For this statistic, we have following asymptotic results.

Theorem C.2. *Under model (2) with $\alpha = 0$, we have the following asymptotic distributions for $\hat{\beta}_{Cau}$ and $t_{\hat{\beta}_{Cau}}$ under various assumptions.*

(i). If $\Phi_T = 1 + \frac{C_x}{T}, \rho_T = \rho$, and $|\rho| < 1$,

$$\begin{aligned} \hat{\beta}_{cau} - \beta &= O_p(T^{-1/2}), \\ t_{\hat{\beta}_{Cau}} &\Rightarrow N(0,1). \end{aligned}$$

(ii). If $\Phi_T = 1 + \frac{C_x}{T}, \rho_T = 1 + \frac{c_u}{T}$,

$$\begin{aligned} \hat{\beta}_{cau} - \beta &\Rightarrow \frac{\int_0^1 sgn\left(J_{C_x}(r)\right) J_{c_u}(r)dr}{\int_0^1 \left|J_{C_x}(r)\right| dr}, \\ t_{\hat{\beta}_{Cau}} &= O_p\left(\frac{T}{M_T}\right)^{1/2}. \end{aligned}$$

(iii). If $\Phi_T = 1 + \frac{C_x}{T}, \rho_T = 1 + \frac{c_u}{T^{\kappa_u}}$,

$$T^{1-\kappa_u}\left(\hat{\beta}_{cau,T} - \beta\right) \Rightarrow \frac{-c_u^{-1}\int_0^1 sgn\left(J_{C_x}(r)\right) dB_0(r)}{\int_0^1 \left|J_{C_x}(r)\right| dr},$$

$$t_{\hat{\beta}_{Cau}} = O_p\left(\frac{T^{\kappa_u}}{M_T}\right)^{1/2}.$$

Proof of Theorem C.2. (i). For the numerator in (154), note that $sgn\left(X_{t-1}\right) = sgn\left(X_{t-1}/T\right)$. Using the Beveridge-Nelson decomposition and summation by part, we have

$$\begin{aligned}
& \frac{1}{\sqrt{T}}\sum_{t=1}^{T} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)u_{0,t} \\
= & \frac{\Psi_0(1)}{\sqrt{T}}\sum_{t=1}^{T} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)z_{0,t} - \frac{1}{\sqrt{T}}\sum_{t=1}^{T} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)\Delta\tilde{z}_{0,t} \\
= & \frac{\Psi_0(1)}{\sqrt{T}}\sum_{t=1}^{T} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)z_{0,t} + o_p(1) \\
\Rightarrow & \Psi_0(1)N(0,\sigma^2) = N(0,\Omega_u). \qquad (155)
\end{aligned}$$

For the second equality above, note that $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)\Delta\tilde{z}_{0,t}$. Taking summation by part yields

$$\begin{aligned}
\frac{1}{\sqrt{T}}\sum_{t=1}^{T} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)\Delta\tilde{z}_{0,t} &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) - sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right)\right)\tilde{z}_{0,t} + o_p(1) \\
&\leq \sup_t |\tilde{z}_{0,t}| \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left|sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) - sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right)\right|.
\end{aligned}$$

Note that $\left|sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) - sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right)\right| = 1$ if $sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) \neq sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right)$. A necessary condition is $\left|\frac{X_{t-1}-X_{t-2}}{\sqrt{T}}\right| > 0$ or asymptotically equivalently $|u_{1,t}/\sqrt{T}| + o_p(1) > 0$. Thus, applying the Beveridge-Nelson decomposition to $u_{1,t}$, we have

$$\begin{aligned}
& \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left|sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) - sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right)\right| \\
\leq & \frac{1}{\sqrt{T}}\sum_{t=1}^{T} I\left(\left|\frac{X_{t-1}-X_{t-2}}{\sqrt{T}}\right| > 0\right) \\
= & \frac{1}{\sqrt{T}}\sum_{t=1}^{T} I\left(|C_1(1)|\left|\frac{z_{1,t}}{\sqrt{T}}\right| + o_p(1) > 0\right) \xrightarrow{p} 0,
\end{aligned}$$

where $I(\cdot)$ denotes indicator function and the last convergence in probability is ensured by the mean square convergence. Since $z_{1,t}$ is i.i.d., we have

$$\begin{aligned}
& E\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} I\left(|C_1(1)|\left|\frac{z_{1,t}}{\sqrt{T}}\right| + o_p(1) > 0\right)\right]^2 \\
= & \frac{1}{T} T E\left[I\left(|C_1(1)|\left|\frac{z_{1,t}}{\sqrt{T}}\right| + o_p(1) > 0\right)\right] \\
= & P\left(|C_1(1)|\left|\frac{z_{1,t}}{\sqrt{T}}\right| + o_p(1) > 0\right) \to 0,
\end{aligned}$$

where the second equality is due to the fact that $\sup_t |\tilde{z}_{0,t}| = O_p(1)$. For the denominator in (154), we have

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T} |X_{t-1}| = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{X_{t-1}}{\sqrt{T}}\right| \Rightarrow \int_0^1 |J_{c_x}(r)|\, dr. \tag{156}$$

Combining (155) and (156) we have

$$T(\hat{\beta}_{cau,T} - \beta) = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) u_{0,t}}{\frac{1}{T^{3/2}}\sum_{t=1}^{T} |X_{t-1}|} \Rightarrow \frac{N(0, \Omega_u)}{\int_0^1 |J_{c_x}(r)|\, dr}.$$

To obtain the limiting distribution of the test statistic $t_{\hat{\beta}_{cau,T}}$, note that $\hat{\Omega}_u \xrightarrow{p} \Omega_u$ and that

$$t_{\hat{\beta}_{cau,T}} = \frac{\sum_{t=1}^{T} sgn\,(X_{t-1})\, u_{0,t}}{\hat{\Omega}_u^{1/2}\sqrt{T}} \Rightarrow N(0,1).$$

(ii). For the numerator in (154), noting that $sgn\,(X_{t-1}) = sgn\left(X_{t-1}/\sqrt{T}\right)$, we have

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T} sgn\,(X_{t-1})\, u_{0,t} = \frac{1}{T}\sum_{t=1}^{T} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)\frac{u_{0,t}}{\sqrt{T}}.$$

By the continuous mapping theorem, the joint convergence of $sgn\left(\frac{x_t}{\sqrt{T}}\right)$, $\frac{u_{0,t}}{\sqrt{T}}$, (36) and (38), we have

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T} sgn\,(X_{t-1})\, u_{0,t} \Rightarrow \int_0^1 sgn\,(J_{c_x}(r))\, J_{c_u}(r) dr. \tag{157}$$

Combining (157) and (156), we obtain

$$\hat{\beta}_{cau,T} - \beta \Rightarrow \frac{\int_0^1 sgn\,(J_{c_x}(r))\, J_{c_u}(r) dr}{\int_0^1 |J_{c_x}(r)|\, dr}.$$

This proves the first claim of Theorem C.2. For the test statistic $t_{\hat{\beta}_{cau,T}}$, note that we can express $t_{\hat{\beta}_{cau,T}}$ as $\frac{\sum_{t=1}^{T} sgn(X_{t-1})u_{0,t}}{\hat{\Omega}_u^{1/2}\sqrt{T}}$. Also from (43), we have $\hat{\Omega}_u = O_p(M_T T)$. Thus, we can deduce that

$$t_{\hat{\beta}_{cau,T}} = \frac{\sum_{t=1}^{T} sgn\left(X_{t-1}\right) u_{0,t}}{\hat{\Omega}_u^{1/2}\sqrt{T}} = \frac{O_p\left(T^{3/2}\right)}{O_p(M_T^{1/2}T^{1/2})T^{1/2}} = O_p\left(\frac{T}{M_T}\right)^{1/2}.$$

(iii). Note that

$$\begin{aligned} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) u_{0,t} &= sgn\left(\Phi_T \frac{X_{t-2}}{\sqrt{T}} + \varepsilon_{1,t-1}\right)\left(\rho_T u_{0,t-1} + \varepsilon_{0,t}\right) \\ &= \rho_T sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) u_{0,t-1} + sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)\varepsilon_{0,t} \qquad (158) \\ &= \rho_T sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right) u_{0,t-1} + sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)\varepsilon_{0,t} \qquad (159) \\ &\quad + \rho_T\left(sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) - sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right)\right) u_{0,t-1}, \end{aligned}$$

Subtracting both sides of (158) by $sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right) u_{0,t-1}$ and summing over $t = 2, ..., T$, we have

$$\begin{aligned} & sgn\left(\frac{X_{T-1}}{\sqrt{T}}\right) u_{0,T} - sgn\left(\frac{X_0}{\sqrt{T}}\right) u_{0,1} \\ = \; & (\rho_T - 1)\sum_{t=2}^{T} sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right) u_{0,t-1} + \sum_{t=2}^{T} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)\varepsilon_{0,t} + \\ & \rho_T \sum_{t=2}^{T}\left(sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) - sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right)\right) u_{0,t-1}. \end{aligned}$$

Scaling the above expression by $T^{-1/2}$, and using $\rho_T - 1 = c_u/T^{\kappa_u}$, we can obtain

$$\begin{aligned} & \frac{c_u}{T^{1/2+\kappa_u}}\sum_{t=2}^{T} sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right) u_{0,t-1} \\ = \; & -\sum_{t=1}^{T} sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right)\frac{\varepsilon_{0,t}}{\sqrt{T}} + sgn\left(\frac{X_{T-1}}{\sqrt{T}}\right)\frac{u_{0,T}}{\sqrt{T}} - sgn\left(\frac{X_0}{\sqrt{T}}\right)\frac{u_{0,1}}{\sqrt{T}} \\ & -\rho_T\sum_{t=2}^{T}\left(sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) - sgn\left(\frac{X_{t-2}}{\sqrt{T}}\right)\right)\frac{u_{0,t-1}}{\sqrt{T}} \\ \Rightarrow \; & -\int_0^1 sgn\left(J_{c_x}(r)\right) dB_0(r) - \int_0^1 \left[sgn\left(J_{c_x}(r)\right) - sgn\left(J_{c_x}(r)\right)\right] dB_0(r) \end{aligned}$$

$$= -\int_0^1 sgn\left(J_{c_x}(r)\right) dB_0(r).$$

Combining this result with (156), we have

$$\begin{aligned} T^{1-\kappa_u}\left(\hat{\beta}_{cau,T} - \beta\right) &= \frac{\frac{1}{T^{1/2+\kappa_u}} \sum_{t=2}^T sgn\left(\frac{X_{t-1}}{\sqrt{T}}\right) u_{0,t}}{\frac{1}{T^{3/2}} \sum_{t=1}^T |X_{t-1}|} \\ &\Rightarrow \frac{-c_u^{-1} \int_0^1 sgn\left(J_{c_x}(r)\right) dB_0(r)}{\int_0^1 |J_{c_x}(r)|\, dr} \end{aligned}$$

For the t statistic, we have

$$\frac{\sum_{t=1}^T sgn\left(X_{t-1}\right) u_{0,t}}{\hat{\Omega}_u^{1/2} \sqrt{T}} = \frac{O_p(T^{1/2+\kappa_u})}{O_p(M_T^{1/2} T^{\kappa_\mu/2}) T^{1/2}} = \frac{O_p(T^{\kappa_u/2})}{O_p(M_T^{1/2})} = O_p\left(\frac{T^{\kappa_u}}{M_T}\right)^{1/2},$$

where $\hat{\Omega}_u = O_p(M_T T^{\kappa_\mu})$ from (64) is used. This completes the proof of Theorem C.2. ■

Theorem C.2 shows that the Cauchy estimator and its corresponding test statistic also lead to spurious inference as OLS, although the asymptotic distributions are quite different.

D. Additional simulation results

Table 9: Empirical sizes under DGP1 with $T = 100$

		c_x						
		-20	-10	-5	-1	0	1	3
$\rho_T = 1 - 30/T$	IVX	0.666	0.602	0.540	0.420	0.384	0.351	0.362
	IVX-AR	0.073	0.074	0.076	0.074	0.076	0.078	0.089
	Modified IVX-AR	0.036	0.032	0.031	0.034	0.040	0.044	0.046
	Modified IVX-AR (BIC)	0.033	0.031	0.030	0.035	0.040	0.044	0.047
$\rho_T = 1 - 10/T$	IVX	0.732	0.756	0.749	0.698	0.664	0.626	0.621
	IVX-AR	0.056	0.066	0.073	0.076	0.084	0.093	0.144
	Modified IVX-AR	0.043	0.042	0.041	0.039	0.039	0.045	0.045
	Modified IVX-AR (BIC)	0.039	0.039	0.038	0.037	0.037	0.042	0.044
$\rho_T = 1 - 5/T$	IVX	0.660	0.734	0.764	0.760	0.743	0.710	0.713
	IVX-AR	0.053	0.059	0.064	0.075	0.077	0.094	0.188
	Modified IVX-AR	0.045	0.044	0.045	0.045	0.047	0.045	0.049
	Modified IVX-AR (BIC)	0.040	0.039	0.040	0.041	0.044	0.042	0.048
$\rho_T = 1$	IVX	0.508	0.625	0.687	0.776	0.799	0.825	0.859
	IVX-AR	0.054	0.056	0.059	0.066	0.074	0.096	0.256
	Modified IVX-AR	0.048	0.049	0.046	0.050	0.052	0.050	0.054
	Modified IVX-AR (BIC)	0.043	0.043	0.041	0.044	0.045	0.044	0.050

Notes: This table reports the empirical rejection rates of original IVX test proposed in Kostakis et al. (2015), IVX-AR test proposed in Yang et al. (2020) as well as the modified IVX-AR we propose based on BIC. The data is generated from DGP 1 and the sample size is 100.

Table 10: Empirical sizes under DGP2 with $T = 100$

		c_x						
		-20	-10	-5	-1	0	1	3
$\rho_T = 1 - 30/T$	IVX	0.290	0.301	0.300	0.275	0.305	0.344	0.424
	IVX-AR	0.077	0.078	0.087	0.089	0.086	0.081	0.081
	Modified IVX-AR	0.048	0.045	0.050	0.058	0.056	0.053	0.054
	Modified IVX-AR (BIC)	0.047	0.046	0.048	0.056	0.056	0.052	0.053
$\rho_T = 1 - 10/T$	IVX	0.492	0.532	0.538	0.544	0.605	0.645	0.696
	IVX-AR	0.102	0.128	0.148	0.088	0.092	0.098	0.117
	Modified IVX-AR	0.050	0.050	0.051	0.049	0.049	0.050	0.050
	Modified IVX-AR (BIC)	0.050	0.050	0.050	0.046	0.047	0.048	0.049
$\rho_T = 1 - 5/T$	IVX	0.566	0.626	0.671	0.666	0.699	0.738	0.760
	IVX-AR	0.123	0.172	0.202	0.108	0.118	0.139	0.173
	Modified IVX-AR	0.054	0.058	0.058	0.048	0.050	0.050	0.048
	Modified IVX-AR (BIC)	0.054	0.059	0.058	0.046	0.048	0.049	0.048
$\rho_T = 1$	IVX	0.742	0.815	0.842	0.795	0.773	0.762	0.724
	IVX-AR	0.174	0.270	0.342	0.216	0.246	0.253	0.251
	Modified IVX-AR	0.060	0.070	0.065	0.052	0.054	0.054	0.055
	Modified IVX-AR (BIC)	0.060	0.072	0.064	0.050	0.052	0.052	0.052

Notes: This table reports the empirical rejection rates of original IVX test proposed in Kostakis et al. (2015), IVX-AR test proposed in Yang et al. (2020) as well as the modified IVX-AR we propose based on BIC. The data is generated from DGP 2 and the sample size is 100.