Estimating and Inferring the Nonlinear Autoregressive Distributed Lag Model by Ordinary Least Squares

JIN SEO CHO

School of Economics, Yonsei University, Seodaemun-gu, Seoul 03722, Korea

jinseocho@yonsei.ac.kr

June 2024

Abstract

This study examines the large sample behavior of an ordinary least squares (OLS) estimator when a nonlinear autoregressive distributed lag (NARDL) model is correctly specified for nonstationary data. Although the OLS estimator suffers from an asymptotically singular matrix problem, it is consistent for unknown model parameters, and follows a mixed normal distribution asymptotically. We also examine the large sample behavior of the standard Wald test defined by the OLS estimator for asymmetries in long- and short-run NARDL parameters, and further supplement it by noting that the long-run parameter estimator is not super-consistent. Using Monte Carlo simulations, we then affirm the theory on the Wald test. Finally, using the U.S. GDP and exogenous fiscal shock data provided by Romer and Romer (2010, *American Economic Review*), we find statistical evidence for long-and short-run symmetries between tax increase and decrease in relation to the U.S. GDP.

Key Words: Nonlinear autoregressive distributed lag model; OLS estimation; Singular matrix; Limit distribution; Wald test; Exogenous fiscal shocks; GDP.

JEL Classifications: C12, C13, C22, E62.

Acknowledgements: The author has benefited from discussions with Matthew Grrnwood-Nimmo and Yoncheol Shin. Cho acknowledges the research grant provided by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF2020S1A5A2A01040235).

1 Introduction

The nonlinear autoregressive distributed lag (NARDL) model is popularly applied to estimate the asymmetric cointegrating relationship between nonstationary variables. After the NARDL model was introduced by Shin, Yu, and Greenwood-Nimmo (2014), many long-run relationships have been revisited, and linear ones have been modified using different slope coefficients depending on the variables' signs. For example, Borenstein, Cameron, and Gilbert (1997) identified the so-called rockets and feathers in gasoline prices, showing that upward cost shocks in crude oil prices pass through faster than downward shocks, affecting the other economic variables asymmetrically. Chesnes (2016) empirically affirms this feature using the NARDL model.

Despite its popularity, estimating the NARDL model by ordinary least squares (OLS) has not yet been theoretically established. Cho, Greenwood-Nimmo, and Shin (2023c) point out that the OLS method suffers from an asymptotically singular matrix problem, although many empirical studies estimate unknown parameters by OLS and compare the standard Wald test using critical values obtained from mixed chi-squared distribution, ignoring the asymptotically singular matrix problem.

This study mainly proposes to revisit the OLS estimator and provide its large sample theory. Although the estimator suffers from the asymptotically singular matrix problem, we find that it provides consistent estimation results for unknown parameters, thus enabling us to derive its asymptotic distribution and present a theoretical ground to apply the Wald test principle to OLS for the NARDL hypothesis. Indeed, the OLS estimator follows a mixed normal distribution under some mild regularity conditions. Although the long-run parameter estimator is not super-consistent, this feature asymptotically validates the popular use of the standard Wald test for empirical data.

Furthermore, we demonstrate the proper use of OLS estimation for NARDL by using empirical data provided by Romer and Romer (2010) and examining the long- and short-run relationships between the U.S. GDP and fiscal exogenous shocks. Romer and Romer (2010) use narrative records such as presidential speeches and *Congressional reports* to measure legislated exogenous tax changes pertaining to the U.S. GDP. Using the NARDL model, we infer the long- and short-run relationships between tax increase and decrease in relation to the U.S. GDP. This serves the additional purpose of illustrating our method in a standard setting.

This study bridges the gap in the NARDL model estimation literature. Indeed, this is not the first work to overcome the asymptotically singular matrix problem associated with NARDL. Cho et al. (2023c) explain the asymptotically singular matrix problem and overcome it by estimating the NARDL parameters using a two-step procedure. They estimate the long- and short-run parameters separately as in Engle and Granger (1987), avoiding the asymptotically singular matrix problem. Referring to this procedure as two-step NARDL (2SNARDL) estimation, they demonstrate that the long-run parameter estimator is super-consistent. Thus, the current study provides a platform to compare the OLS method with 2SNARDL estimation, while resolving the methodological issues related to OLS.

The problem of asymptotically singular matrix may be difficult to resolve through higher-order expansion. Note that the problem may be related to an unidentified model problem. As for the maximum-likelihood (ML) and nonlinear least squares (NLS) estimations, the lack of model identification can be tackled through higher-order model expansion, as in the literature (e.g., Teräsvirta, 1994; Cho and White, 2007, 2010; Cho, Ishida, and White, 2011; Cho and Ishida, 2012; Baek, Cho, and Phillips, 2015; Cho and Phillips, 2018; White and Cho, 2012; Seong, Cho, and Teräsvirta, 2022, among others). However, applying higher-order expansion directly to NARDL for OLS estimation can be challenging. The model already assumes linearity, unlike ML or NLS estimators, which means that higher-order expansion has to be applied to the estimator itself, and not to the model. This is particularly challenging because higher-order expansion has to be applied to the inverse matrix determinant associated with OLS, given that the dimension of the inverse matrix can be arbitrary.

We address this by obtaining the asymptotic distribution of the OLS estimator indirectly as follows. We first represent the OLS estimator as a transform of other primitive estimators that do not suffer from the asymptotically singular matrix problem, and then derive their weak limits, to obtain the desired limit distribution. Through this limit distribution, the OLS convergence rate becomes lower than that without the singular matrix problem, not letting the long-run parameter estimator become super-consistent. This lower convergence rate typically occurs when higher-order expansion is used to derive the asymptotic distribution of an estimator.

The limit distribution thus obtained indirectly can be used to infer the unknown parameter. We can also use it to examine the large sample behavior of the Wald test as defined by OLS, and further supplement this by noting that the OLS estimator has a lower convergence rate. As discussed below, the primitive estimator used for long-run parameter estimation is super-consistent, implying that we can define another Wald statistic to test a long-run parameter.

The empirical illustration of this study presents a suitable environment to apply NARDL. We are primarily interested in identifying how the long- and short-run relationships between the U.S. GDP and tax decreases differ from those between the U.S. GDP and tax increases. Romer and Romer (2010) classify the legislated tax changes into exogenous and endogenous ones in terms of GDP based on narrative data. The literature presents two types of exogenous tax changes: those for deficit reduction, and those for long-run growth. All tax changes for deficit reduction are related to tax increases, while most tax changes for long-run growth are related to tax decreases. Thus, by applying the NARDL model, we can infer the long- and short-run relationships between the tax changes for deficit reduction and those for long-run growth in terms of U.S. GDP. This investigation provides statistical evidence that long- and short-run parameters are symmetrical. That is, the tax changes for deficit reduction and those for long-run growth affect the U.S. GDP similarly. Moreover, our findings indicate that a 1% exogenous GDP tax decrease increases the log real GDP by about 3% in the long run. This is close to the estimation results in Romer and Romer (2010). While drawing this evidence, we illustrate our methodology using OLS along with 2SNARDL for comparison.

This study is structured as follows. Section 2 overviews the NARDL model and discusses the asymptotically singular matrix problem associated with OLS. Section 3 defines primitive estimators and presents the OLS estimator as a bilinear transform of other estimators. Section 4 discusses the limit distribution of an OLS estimator with different distributions depending on parameter values and the specific conditions for limit distributions. Section 5 examines the large sample properties of the standard Wald test for the NARDL hypothesis. This section also discusses another Wald test for supplementary purposes, examining its large sample behavior. Section 6 conducts Monte Carlo simulations for the Wald tests, while Section 7 presents the empirical illustration. Finally, the concluding remarks are presented in Section 8. All mathematical proofs are provided in the Online Supplement.

Before concluding this section, we present the notation used throughout the study. We provide the weak limit of an estimator by a stochastic integral. Denoting the weak limit by $\int \mathscr{B}$ or $\int d\mathscr{B}$ means $\int_0^1 \mathscr{B}(u)du$ or $\int_0^1 d\mathscr{B}(u)$, respectively, where $\mathscr{B}(\cdot)$ is a Brownian motion.

2 Motivation and the NARDL Model in the Literature

This section briefly summarizes NARDL and motivates the current study by relating OLS to the asymptotically singular matrix problem. In addition, the study relates the problem to the literature by associating the singular matrix problem with an unidentified model.

Consider a NARDL(p, q) process augmented by a time drift:

$$y_t = \alpha_* + \xi_* t + \sum_{j=1}^{p} \phi_{j*} y_{t-j} + \sum_{j=0}^{q} (\boldsymbol{\theta}_{j*}^{+\prime} \boldsymbol{x}_{t-j}^{+} + \boldsymbol{\theta}_{j*}^{-\prime} \boldsymbol{x}_{t-j}^{-}) + e_t, \tag{1}$$

where $\boldsymbol{x}_t \in \mathbb{R}^k$, $\boldsymbol{x}_t^+ := \sum_{j=1}^t \Delta\boldsymbol{x}_j^+$, $\boldsymbol{x}_t^- := \sum_{j=1}^t \Delta\boldsymbol{x}_j^-$, $\Delta\boldsymbol{x}_t^+ := \max[\mathbf{0}, \Delta\boldsymbol{x}_t]$, $\Delta\boldsymbol{x}_t^- := \min[\mathbf{0}, \Delta\boldsymbol{x}_t]$, $\{e_t, \mathscr{F}_t\}$ is a martingale difference array (MDA), and $\mathscr{F}_t$ is the smallest σ-algebra driven by $\{y_{t-1}, \boldsymbol{x}_t^+, \boldsymbol{x}_t^-, y_{t-2}, \boldsymbol{x}_{t-1}^+, \boldsymbol{x}_{t-1}^-, \ldots\}$ such that $\Delta\boldsymbol{x}_t$ is a stationary process. The NARDL process in (1) is more general than thpse defined by Shin et al. (2014) and Cho, Greenwood-Nimmo, and Shin (2023b), because the latter ones do not allow for a time trend on the right-hand side.

Note that (1) can be rewritten in the error-correction form as follows:

$$\Delta y_t = \rho_* y_{t-1} + \boldsymbol{\theta}_*^{+\prime}\boldsymbol{x}_{t-1}^{+} + \boldsymbol{\theta}_*^{-\prime}\boldsymbol{x}_{t-1}^{-} + \xi_*(t-1) + \alpha_* + \sum_{j=1}^{p-1}\varphi_{j*}\Delta y_{t-j} + \sum_{j=0}^{q-1}\left(\boldsymbol{\pi}_{j*}^{+\prime}\Delta\boldsymbol{x}_{t-j}^{+} + \boldsymbol{\pi}_{j*}^{-\prime}\Delta\boldsymbol{x}_{t-j}^{-}\right) + e_t, \tag{2}$$

by letting ρ_*, $\boldsymbol{\theta}_*^{+}$, $\boldsymbol{\theta}_*^{-}$, φ_{j*} ($j = 1, 2, \ldots, p-1$), $\boldsymbol{\pi}_{j*}^{+}$, and $\boldsymbol{\pi}_{j*}^{-}$ ($j = 0, 1, \ldots, q-1$) be such that $\rho_* := \sum_{k=1}^{p}\phi_{j*} - 1$, $\boldsymbol{\theta}_*^{+} := \sum_{j=0}^{q}\boldsymbol{\theta}_{j*}^{+}$, $\boldsymbol{\theta}_*^{-} := \sum_{j=0}^{q}\boldsymbol{\theta}_{j*}^{-}$, $\boldsymbol{\pi}_{0*}^{+} := \boldsymbol{\theta}_{0*}^{+}$, $\boldsymbol{\pi}_{0*}^{-} := \boldsymbol{\theta}_{0*}^{-}$, and for $\ell = 1, 2, \ldots, p-1$ and $j = 1, 2, \ldots, q-1$, $\varphi_{\ell*} := -\sum_{i=\ell+1}^{p}\phi_{i*}$, $\boldsymbol{\pi}_{j*}^{+} := -\sum_{i=j+1}^{q}\boldsymbol{\theta}_{i*}^{+}$, and $\boldsymbol{\pi}_{j*}^{-} := -\sum_{i=j+1}^{q}\boldsymbol{\theta}_{i*}^{-}$. If y_t is cointegrated with $(\boldsymbol{x}_t^{+\prime}, \boldsymbol{x}_t^{-\prime})'$, we may rewrite (2) as

$$\Delta y_t = \rho_* u_{t-1} + \gamma_* + \sum_{j=1}^{p-1}\varphi_{j*}\Delta y_{t-j} + \sum_{j=0}^{q-1}\left(\boldsymbol{\pi}_{j*}^{+\prime}\Delta\boldsymbol{x}_{t-j}^{+} + \boldsymbol{\pi}_{j*}^{-\prime}\Delta\boldsymbol{x}_{t-j}^{-}\right) + e_t \tag{3}$$

such that the cointegration error u_{t-1} can be defined as $u_{t-1} := y_{t-1} - \boldsymbol{\beta}_*^{+\prime}\boldsymbol{x}_{t-1}^{+} - \boldsymbol{\beta}_*^{-\prime}\boldsymbol{x}_{t-1}^{-} - \zeta_*(t-1) - \nu_*$ with $\boldsymbol{\beta}_*^{+} := -(\boldsymbol{\theta}_*^{+}/\rho_*)$, $\boldsymbol{\beta}_*^{-} := -(\boldsymbol{\theta}_*^{-}/\rho_*)$, and $\zeta_* := -(\xi_*/\rho_*)$, where $\gamma_* := \alpha_* + \xi_* + \rho_*\nu_*$. Here, we introduce ν_* such that $\mathbb{E}[u_t] = 0$, and let u_t be a stationary process that can be correlated with $\Delta\boldsymbol{x}_t$.

The NARDL process captures an asymmetric cointegrating relationship between nonstationary processes. If $\boldsymbol{\mu}_*^{+} := \mathbb{E}[\Delta\boldsymbol{x}_t^{+}]$ and $\boldsymbol{\mu}_*^{-} := \mathbb{E}[\Delta\boldsymbol{x}_t^{-}]$, then $\boldsymbol{\mu}_*^{+} + \boldsymbol{\mu}_*^{-} \equiv \mathbb{E}[\Delta\boldsymbol{x}_t]$ by construction, as $\Delta\boldsymbol{x}_t \equiv \Delta\boldsymbol{x}_t^{+} + \Delta\boldsymbol{x}_t^{-}$. Therefore, if further $\boldsymbol{s}_t^{+} := \Delta\boldsymbol{x}_t^{+} - \boldsymbol{\mu}_*^{+}$ and $\boldsymbol{s}_t^{-} := \Delta\boldsymbol{x}_t^{-} - \boldsymbol{\mu}_*^{-}$, then it follows that

$$\boldsymbol{x}_t^{+} = \boldsymbol{\mu}_*^{+}t + \mathbf{m}_t^{+} \quad \text{and} \quad \boldsymbol{x}_t^{-} = \boldsymbol{\mu}_*^{-}t + \mathbf{m}_t^{-} \quad \text{by letting} \quad \mathbf{m}_t^{+} := \sum_{j=1}^{t}\boldsymbol{s}_j^{+} \quad \text{and} \quad \mathbf{m}_t^{-} := \sum_{j=1}^{t}\boldsymbol{s}_j^{-}. \tag{4}$$

From (4), it is clear that $\boldsymbol{x}_t^{+}$ and $\boldsymbol{x}_t^{-}$ are unit-root processes with non-zero time drifts. Moreover, Δy_t is not necessarily distributed around zero even when $\boldsymbol{x}_t$ is a unit-root process. From (3), we obtain

$$\delta_* := \mathbb{E}[\Delta y_t] = \frac{1}{\varrho_*}\left[\gamma_* + \sum_{j=0}^{q-1}\boldsymbol{\pi}_{j*}^{+\prime}\boldsymbol{\mu}_*^{+} + \sum_{j=0}^{q-1}\boldsymbol{\pi}_{j*}^{-\prime}\boldsymbol{\mu}_*^{-}\right], \quad \text{where} \quad \varrho_* := 1 - \sum_{j=1}^{p-1}\varphi_{j*},$$

such that if $d_t := \Delta y_t - \delta_*$, then

$$y_t = \delta_* t + \sum_{j=1}^{t} d_j. \tag{5}$$

This implies that y_t is a unit-root process with deterministic time drift. These findings indicate that (3) can capture a cointegrating relationship between y_t and $(\boldsymbol{x}_t^{+\prime}, \boldsymbol{x}_t^{-\prime})'$.

OLS estimation of the unknown parameters in (2) is not straightforward, although this approach is popular in the empirical literature. This is mainly because it suffers from an asymptotically singular matrix problem. For the examination of this issue and notational simplicity, we first assume that

$$
\begin{aligned}
\boldsymbol{z}_t &:= [\ \boldsymbol{z}_{1t}' \ \vdots\ \boldsymbol{z}_{2t}'\ ]' \\
&:= [\ y_{t-1} \quad \boldsymbol{x}_{t-1}^{+\prime} \quad \boldsymbol{x}_{t-1}^{-\prime} \quad (t-1) \quad 1 \ \vdots\ \Delta \boldsymbol{y}_{t-1}' \quad \Delta \boldsymbol{x}_t^{+\prime} \quad \ldots \quad \Delta \boldsymbol{x}_{t-q+1}^{+\prime} \quad \Delta \boldsymbol{x}_t^{-\prime} \quad \ldots \quad \Delta \boldsymbol{x}_{t-q+1}^{-\prime}\ ]',
\end{aligned}
$$

where $\Delta \boldsymbol{y}_{t-1} := [\Delta y_{t-1}, \Delta y_{t-2}, \ldots, \Delta y_{t-p+1}]'$. Note that $\boldsymbol{z}_t \in \mathbb{R}^{2+p+2k(1+q)}$ is partitioned into two variables, $\boldsymbol{z}_{1t}$ and $\boldsymbol{z}_{2t}$, such that $\boldsymbol{z}_{1t} \in \mathbb{R}^{3+2k}$ and $\boldsymbol{z}_{2t} \in \mathbb{R}^{p+2kq-1}$. Here, $\boldsymbol{z}_{1t}$ and $\boldsymbol{z}_{2t}$ collect the variables in the long- and short-run equations, respectively. Furthermore, we also assume that

$$
\boldsymbol{\alpha}_* := [\ \boldsymbol{\alpha}_{1*}' \ \vdots\ \boldsymbol{\alpha}_{2*}'\ ]' := [\ \rho_* \quad \boldsymbol{\theta}_*^{+\prime} \quad \boldsymbol{\theta}_*^{-\prime} \quad \xi_* \quad \alpha_* \ \vdots\ \boldsymbol{\varphi}_*' \quad \boldsymbol{\pi}_*^{+\prime} \quad \boldsymbol{\pi}_*^{-\prime}\ ]',
$$

where $\boldsymbol{\varphi}_* := [\varphi_{1*}, \varphi_{2*}, \ldots, \varphi_{p-1*}]'$, $\boldsymbol{\pi}_*^{+} := [\boldsymbol{\pi}_{0*}^{+\prime}, \boldsymbol{\pi}_{1*}^{+\prime}, \ldots, \boldsymbol{\pi}_{q-1*}^{+\prime}]'$, and $\boldsymbol{\pi}_*^{-} := [\boldsymbol{\pi}_{0*}^{-\prime}, \boldsymbol{\pi}_{1*}^{-\prime}, \ldots, \boldsymbol{\pi}_{q-1*}^{-\prime}]'$. From this, we can rewrite the OLS estimator as

$$
\widehat{\boldsymbol{\alpha}}_T := [\ \widehat{\rho}_T \quad \widehat{\boldsymbol{\theta}}_T^{+\prime} \quad \widehat{\boldsymbol{\theta}}_T^{-\prime} \quad \widehat{\xi}_T \quad \widehat{\alpha}_T \ \vdots\ \widehat{\boldsymbol{\varphi}}_T' \quad \widehat{\boldsymbol{\pi}}_T^{+\prime} \quad \widehat{\boldsymbol{\pi}}_T^{-\prime}\ ]' := \left(\sum_{t=1}^{T} \boldsymbol{z}_t \boldsymbol{z}_t'\right)^{-1} \left(\sum_{t=1}^{T} \boldsymbol{z}_t \Delta y_t\right).
$$

To discuss the asymptotically singular matrix problem related to OLS estimation, we first introduce some mild regularity conditions for the data:

Assumption 1. *(i)* $\{(\Delta \boldsymbol{x}_t', u_t, e_t)' \in \mathbb{R}^{k+2} : t = \ldots, -1, 0, 1, \ldots\}$ *is a strictly stationary mixing process of size* $-r/(2(r-1))$ *or* α *of size* $-r/(r-2)$ *and* $r > 2$*;*

(ii) $\mathbb{E}[|\Delta x_{ti}|^r] < \infty$ $(i = 1, 2, \ldots, k)$, $\mathbb{E}[|u_t|^r] < \infty$, $\mathbb{E}[|e_t|^2] < \infty$, *and* $\delta_* \neq 0$, *where* x_{ti} *is the* i*-th row element of* $\boldsymbol{x}_t$*;*

(iii) $\boldsymbol{\Sigma}_* := \lim_{T\to\infty} var[T^{-1/2} \sum_{t=1}^{T} \boldsymbol{w}_t]$ *is positive definite, where*

$$
\boldsymbol{w}_t := [\ \boldsymbol{w}_{1t}' \ \vdots\ \boldsymbol{w}_{2t}'\ ] := [\ \boldsymbol{s}_{t-1}^{+\prime} \quad \boldsymbol{s}_{t-1}^{-\prime} \quad u_{t-1} \ \vdots\ e_t \quad e_t u_{t-1} \quad e_t \boldsymbol{z}_{2t}'\ ]';
$$

(iv) for some $\boldsymbol{\alpha}_*$ *with* $\rho_* < 0$, Δy_t *is generated by* (2) *such that* $|L_*| > 1$, *where* $1 - \sum_{j=1}^{p} \phi_{j*} L_*^j \equiv 0$*;*

(v) $\{e_t, \mathscr{F}_t\}$ *is an MDA.* □

Remarks. **(a)** Assumptions 1 (*i* and *ii*) assume mixing and moment conditions to apply the functional central limit theorem (FCLT) to partial-sum data process. FCLT is popularly applied to derive the limit distribution of the OLS estimator applied to estimate a cointegrating relationship (e.g., Phillips and Hansen, 1990; Phillips, 1991; White, 2001, chapter 7).

(b) Assumption 1 (*iii*) assumes a positive-definite matrix condition for the OLS limit distribution.

(c) Assumption 1 (*iv*) assumes $\rho_* < 0$ for a cointegrating relationship between y_t and $(\boldsymbol{x}_t^+, \boldsymbol{x}_t^-)$. If we let $\boldsymbol{\beta}_*^+ = \boldsymbol{\beta}_*^- = \mathbf{0}$ and $\zeta_* = 0$, then from (3) it follows that

$$y_t = \alpha_* + \sum_{j=1}^{p} \phi_{j*} y_{t-j} + \sum_{j=0}^{q-1} \left(\boldsymbol{\pi}_{j*}^{+\prime} \Delta \boldsymbol{x}_{t-j}^{+} + \boldsymbol{\pi}_{j*}^{-\prime} \Delta \boldsymbol{x}_{t-j}^{-} \right) + e_t,$$

while Assumption 1 (*iv*) implies that y_t is a stationary process, contradicting the assumption that y_t is a unit-root process, indicating that we need to assume that $\rho_* = 0$ for $\boldsymbol{\beta}_*^+ = \boldsymbol{\beta}_*^- = \mathbf{0}$ and $\zeta_* = 0$. In this study, we do not consider this, but focus on the limit behavior of the OLS estimator to estimate the cointegration system. As pointed out by Pesaran, Shin, and Smith (2001) and Banerjee, Dolado, and Mestre (1998) who examine the F-test and t-test for $\rho_* = 0$, the limit behavior of the estimate under cointegration is different from that without cointegration. Our primary interest is in examining the estimate under cointegration.

(d) Assumption 1 (*v*) is a standard condition for the error term in the autoregressive distributed lag (ARDL) and NARDL processes (e.g., Pesaran and Shin, 1998; Pesaran et al., 2001; Shin et al., 2014; Cho et al., 2023c). □

Next, we examine the asymptotically singular matrix problem. For this, we need the following lemma.

Lemma 1. *Under Assumption 1,*

(i) if $\mathbf{D}_1 := \mathit{diag}[T^{3/2}\mathbf{I}_{2+2k}, T^{1/2}]$, $\mathbf{D}_1^{-1}\left(\sum_{t=1}^{T} \boldsymbol{z}_{1t}\boldsymbol{z}_{1t}'\right)\mathbf{D}_1^{-1} \xrightarrow{\mathbb{P}} \mathbf{M}_{11}$, *where*

$$\mathbf{M}_{11} := \begin{bmatrix} \frac{1}{3}\delta_*^2 & \frac{1}{3}\delta_*\boldsymbol{\mu}_*^{+\prime} & \frac{1}{3}\delta_*\boldsymbol{\mu}_*^{-\prime} & \frac{1}{3}\delta_* & \frac{1}{2}\delta_* \\ \frac{1}{3}\delta_*\boldsymbol{\mu}_*^{+} & \frac{1}{3}\boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{+\prime} & \frac{1}{3}\boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{-\prime} & \frac{1}{3}\boldsymbol{\mu}_*^{+} & \frac{1}{2}\boldsymbol{\mu}_*^{+} \\ \frac{1}{3}\delta_*\boldsymbol{\mu}_*^{-} & \frac{1}{3}\boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{+\prime} & \frac{1}{3}\boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{-\prime} & \frac{1}{3}\boldsymbol{\mu}_*^{-} & \frac{1}{2}\boldsymbol{\mu}_*^{-} \\ \frac{1}{3}\delta_* & \frac{1}{3}\boldsymbol{\mu}_*^{+\prime} & \frac{1}{3}\delta_*\boldsymbol{\mu}_*^{-\prime} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2}\delta_* & \frac{1}{2}\boldsymbol{\mu}_*^{+\prime} & \frac{1}{2}\boldsymbol{\mu}_*^{-\prime} & \frac{1}{2} & 1 \end{bmatrix};$$

(ii) if $\mathbf{D}_2 := \mathit{diag}[T^{1/2}\mathbf{I}_{p+2kq-1}]$, $\mathbf{D}_2^{-1}\left(\sum_{t=1}^{T} \boldsymbol{z}_{2t}\boldsymbol{z}_{2t}'\right)\mathbf{D}_2^{-1} \xrightarrow{\mathbb{P}} \mathbf{M}_{22} := \mathbb{E}[\boldsymbol{z}_{2t}\boldsymbol{z}_{2t}']$*;*

(iii) $\mathbf{D}_2^{-1}\left(\sum_{t=1}^{T} \boldsymbol{z}_{2t}\boldsymbol{z}_{1t}'\right)\mathbf{D}_1^{-1} \xrightarrow{\mathbb{P}} \mathbf{M}_{21}$, *where*

$$\mathbf{M}_{21} := \begin{bmatrix} \frac{1}{2}\delta_*^2\boldsymbol{\iota}_{p-1} & \frac{1}{2}\delta_*\boldsymbol{\iota}_{p-1}\boldsymbol{\mu}_*^{+\prime} & \frac{1}{2}\delta_*\boldsymbol{\iota}_{p-1}\boldsymbol{\mu}_*^{-\prime} & \frac{1}{2}\delta_*\boldsymbol{\iota}_{p-1} & \delta_*\boldsymbol{\iota}_{p-1} \\ \frac{1}{2}\delta_*\boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*^{+} & \frac{1}{2}\boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{+\prime} & \frac{1}{2}\boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{-\prime} & \frac{1}{2}\boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*^{+} & \boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*^{+} \\ \frac{1}{2}\delta_*\boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*^{-} & \frac{1}{2}\boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*^{+}\frac{1}{2}\boldsymbol{\mu}_*^{-\prime} & \frac{1}{2}\boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{-\prime} & \frac{1}{2}\boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*^{-} & \boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*^{-} \end{bmatrix};$$

(iv) if $\mathbf{D} := diag[\mathbf{D}_1, \mathbf{D}_2]$ *and* $\mathbf{M}_{12} := \mathbf{M}_{21}'$, $\mathbf{D}^{-1}\left(\sum_{t=1}^{T} \boldsymbol{z}_t \boldsymbol{z}_t'\right)\mathbf{D}^{-1} \xrightarrow{\mathbb{P}} \mathbf{M}$, *where*

$$\mathbf{M} := \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{bmatrix},$$

which is singular. □

Note that $\mathbf{M}$ is singular because its first $(1 + 2k)$ columns are proportional to the $(2 + 2k)$-th column. The proof of Lemma 1 is provided in the Online Supplement by extending lemma 1 of Cho et al. (2023c), demonstrating another singular matrix problem when there is no time trend in (1). Augmenting the time trend on the right-hand side does not eliminate the asymptotically singular matrix problem.

The asymptotically singular matrix problem implies not the absence of OLS estimator limit distribution, but the fact that the limit distribution has to be obtained differently from the standard case. For ML and NLS estimations, the limit distribution is typically obtained by applying higher-order approximation to the nonlinear model, bringing the convergence rate of the estimator lower than the standard case (e.g., Teräsvirta, 1994; Cho and White, 2007, 2010; Cho et al., 2011; Baek et al., 2015; Cho and Phillips, 2018; White and Cho, 2012; Seong et al., 2022).

Applying higher-order expansion theory to the OLS estimator is not straightforward here. This is mainly because the expansion involves higher-order approximation of the determinant of $(\sum_{t=1}^{T} \boldsymbol{z}_t \boldsymbol{z}_t')$, which is challenging because its dimension is determined by k, p, and q, and these can be arbitrarily given.

3 An Alternative Representation of the OLS Estimator

Owing to the complexity of higher-order expansion, we obtain the desired limit distribution using a different approach compared to the standard case. Representing the OLS estimator as a bilinear transformation of other primitive estimators that do not suffer from an asymptotic singularity problem, we obtain the OLS limit distribution from their limit distributions. This brings the OLS convergence rate lower than that of $\mathbf{D}$.

We proceed with our following discussions in three steps. First, we estimate the long-run parameters using OLS. That is,

$$\widehat{\boldsymbol{\upsilon}}_T := (\widehat{\boldsymbol{\beta}}_T^{+\prime}, \widehat{\boldsymbol{\beta}}_T^{-\prime}, \widehat{\zeta}_T, \widehat{\nu}_T)' := \underset{\boldsymbol{\beta}^+, \boldsymbol{\beta}^-, \zeta, \nu}{\arg\min} \sum_{t=1}^{T} (y_{t-1} - \boldsymbol{\beta}^{+\prime}\boldsymbol{x}_{t-1}^{+} - \boldsymbol{\beta}^{-\prime}\boldsymbol{x}_{t-1}^{-} - \zeta(t-1) - \nu)^2 \quad \text{and}$$

$$\widehat{\boldsymbol{\upsilon}}_T = \left(\sum_{t=1}^{T} \widetilde{\boldsymbol{r}}_{t-1}\widetilde{\boldsymbol{r}}_{t-1}'\right)^{-1} \left(\sum_{t=1}^{T} \widetilde{\boldsymbol{r}}_{t-1} y_{t-1}\right),$$

where $\widetilde{\boldsymbol{r}}_t := [\boldsymbol{x}_t^{+\prime}, \boldsymbol{x}_t^{-\prime}, t, 1]'$. We then assume that $\widehat{u}_t := y_t - \widetilde{\boldsymbol{r}}_t' \widehat{\boldsymbol{v}}_T$. Here, we obtain $\widehat{\boldsymbol{v}}_T$ from the following equation,

$$y_t = \boldsymbol{\beta}_*^{+\prime} \boldsymbol{x}_t^+ + \boldsymbol{\beta}_*^{-\prime} \boldsymbol{x}_t^- + \zeta_* t + \nu_* + u_t. \tag{6}$$

When $\boldsymbol{v}_* := [\boldsymbol{\beta}_*^{+\prime}, \boldsymbol{\beta}_*^{-\prime}, \zeta_*, \nu_*]'$, $\widehat{\boldsymbol{v}}_T$ estimates $\boldsymbol{v}_*$.

In the second step, we specify model

$$\begin{aligned} \Delta y_t = \rho_* u_{t-1} + \boldsymbol{\eta}_*^{+\prime} \boldsymbol{x}_{t-1}^+ + \boldsymbol{\eta}_*^{-\prime} \boldsymbol{x}_{t-1}^- + \psi_*(t-1) + \gamma_* \\ + \sum_{j=1}^{p-1} \varphi_{j*} \Delta y_{t-j} + \sum_{j=0}^{q-1} \left(\boldsymbol{\pi}_{j*}^{+\prime} \Delta \boldsymbol{x}_{t-j}^+ + \boldsymbol{\pi}_{j*}^{-\prime} \Delta \boldsymbol{x}_{t-j}^- \right) + e_t \end{aligned} \tag{7}$$

by combining (2) and (6), where $\boldsymbol{\eta}_*^+ := \boldsymbol{\theta}_*^+ + \rho_* \boldsymbol{\beta}_*$, $\boldsymbol{\eta}_*^- := \boldsymbol{\theta}_*^- + \rho_* \boldsymbol{\beta}_*^-$, and $\psi_* := \xi_* + \rho_* \zeta_*$ for OLS estimation of the parameters in (7):

$$\widetilde{\boldsymbol{\omega}}_T := \left(\sum_{t=1}^{T} \widetilde{\boldsymbol{z}}_t \widetilde{\boldsymbol{z}}_t' \right)^{-1} \left(\sum_{t=1}^{T} \widetilde{\boldsymbol{z}}_t \Delta y_t \right).$$

Here, we assume that $\widetilde{\boldsymbol{z}}_t := [\ \widehat{u}_{t-1} \ \vdots \ \widetilde{\boldsymbol{z}}_{1t}' \ \vdots \ \widetilde{\boldsymbol{z}}_{2t}' \ ]' := [\ \widehat{u}_{t-1} \ \vdots \ \widetilde{\boldsymbol{r}}_{t-1}' \ \vdots \ \boldsymbol{z}_{2t}' \ ]'$. Note that we replace u_{t-1} by $\widehat{u}_{t-1}$, and from the definitions of $\boldsymbol{\beta}_*^+$, $\boldsymbol{\beta}_*^-$, and ζ_*, we have $\boldsymbol{\eta}_* = \mathbf{0}$, $\boldsymbol{\eta}_*^- = \mathbf{0}$, and $\psi_* = 0$, but their corresponding variables included as auxiliary regressors, namely, $\boldsymbol{x}_{t-1}^+$, $\boldsymbol{x}_{t-1}$, and $t-1$. By this inclusion, we relate $\widetilde{\boldsymbol{\omega}}_T$ to $\widehat{\boldsymbol{\alpha}}_T$, that is,

$$\widehat{\boldsymbol{\alpha}}_T = \mathbf{R}_T \widetilde{\boldsymbol{\omega}}_T, \quad \text{where} \quad \mathbf{R}_T := \begin{bmatrix} \mathbf{R}_T^{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{p+2kq-1} \end{bmatrix}, \quad \text{and} \quad \mathbf{R}_T^{11} := \begin{bmatrix} 1 & \mathbf{0}_{1\times(2+2k)} \\ -\widehat{\boldsymbol{v}}_T & \mathbf{I}_{2+2k} \end{bmatrix}. \tag{8}$$

Now, $\widehat{\boldsymbol{\alpha}}_T$ is a bilinear transformation of $\widehat{\boldsymbol{v}}_T$ and $\widetilde{\boldsymbol{\omega}}_T$. We obtain this by applying Supplement Lemma A.1 to (7) by letting y_t, $\boldsymbol{x}_t$, and $\boldsymbol{z}_t$ of Lemma A.1 be Δy_t, y_{t-1}, and $[\widetilde{\boldsymbol{r}}_{t-1}', \boldsymbol{z}_{2t}']'$, respectively. Here, $\widehat{\boldsymbol{v}}_t$ in Lemma A.1 can be identified as $\widehat{u}_{t-1}$.

As the final step, we represent $\widehat{\boldsymbol{v}}_T$ and $\widetilde{\boldsymbol{\omega}}_T$ using other primitive estimators having no asymptotically singular matrix problem. This step is essential because both $\widehat{\boldsymbol{v}}_T$ and $\widetilde{\boldsymbol{\omega}}_T$ also suffer from asymptotically singular matrix problems as Supplement Lemma A.3 demonstrates. We therefore represent them as estimators defined by other estimators not suffering from a singularity problem. Note that (4) and (6) imply that

$$y_t = \boldsymbol{\beta}_*^{+\prime} \mathbf{m}_t^+ + \boldsymbol{\beta}_*^{-\prime} \mathbf{m}_t^- + \vartheta_* t + \nu_* + u_t, \tag{9}$$

where $\vartheta_* := \boldsymbol{\beta}_*^{+\prime} \boldsymbol{\mu}_*^+ + \boldsymbol{\beta}_*^{-\prime} \boldsymbol{\mu}_*^- + \zeta_*$. Therefore, we can estimate the coefficients in (9), when y_t is regressed against $\mathbf{m}_t^+$, $\mathbf{m}_t^-$, t, and 1. However, as $\mathbf{m}_t^+$ and $\mathbf{m}_t^-$ are not observable, we estimate them by regressing $\boldsymbol{x}_t^+$ and $\boldsymbol{x}_t^-$ against t according to (4): $\widehat{\boldsymbol{\mu}}_T^+ := \left(\sum_{t=1}^{T-1} t^2\right)^{-1} \left(\sum_{t=1}^{T-1} t\boldsymbol{x}_t^+\right)$ and $\widehat{\boldsymbol{\mu}}_T^- := \left(\sum_{t=1}^{T-1} t^2\right)^{-1} \left(\sum_{t=1}^{T-1} t\boldsymbol{x}_t^-\right)$, to obtain the regression residual by $\widehat{\mathbf{m}}_t^+ := \boldsymbol{x}_t^+ - t\widehat{\boldsymbol{\mu}}_T^+$ and $\widehat{\mathbf{m}}_t^- := \boldsymbol{x}_t^- - t\widehat{\boldsymbol{\mu}}_T^-$. We then assume that $\dot{\boldsymbol{r}}_t := [\widehat{\mathbf{m}}_t^{+\prime}, \widehat{\mathbf{m}}_t^{-\prime}, t, 1]'$, and regress y_t against $\dot{\boldsymbol{r}}_t$, to obtain

$$\widetilde{\boldsymbol{\upsilon}}_T := [\widetilde{\boldsymbol{\beta}}_T^{+\prime}, \widetilde{\boldsymbol{\beta}}_T^{-\prime}, \widetilde{\vartheta}_T, \widetilde{\nu}_T]' := \left(\sum_{t=1}^{T} \dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}'\right)^{-1} \left(\sum_{t=1}^{T} \dot{\boldsymbol{r}}_{t-1}y_{t-1}\right)$$

by estimating the coefficients in (9), namely, $\bar{\boldsymbol{\upsilon}}_* := [\boldsymbol{\beta}_*^{+\prime}, \boldsymbol{\beta}_*^{-\prime}, \vartheta_*, \nu_*]'$. From now, for notational simplicity, we let $\mathbf{m}_t := [\mathbf{m}_t^{+\prime}, \mathbf{m}_t^{-\prime}]'$, $\boldsymbol{x}_t := [\boldsymbol{x}_t^{+\prime}, \boldsymbol{x}_t^{-\prime}]'$, $\boldsymbol{\beta} := [\boldsymbol{\beta}^{+\prime}, \boldsymbol{\beta}^{-\prime}]'$, $\boldsymbol{\theta} := [\boldsymbol{\theta}^{+\prime}, \boldsymbol{\theta}^{-\prime}]'$, $\boldsymbol{\eta} := [\boldsymbol{\eta}^{+\prime}, \boldsymbol{\eta}^{-\prime}]'$, and $\boldsymbol{\mu} := [\boldsymbol{\mu}^{+\prime}, \boldsymbol{\mu}^{-\prime}]'$.

This estimator $\widetilde{\boldsymbol{\upsilon}}_T$ is specifically related to $\widehat{\boldsymbol{\upsilon}}_T$ as follows:

$$\widehat{\boldsymbol{\upsilon}}_T := [\widehat{\boldsymbol{\beta}}_T^{+\prime}, \widehat{\boldsymbol{\beta}}_T^{-\prime}, \widehat{\zeta}_T, \widehat{\nu}_T]' = \mathbf{P}_T\widetilde{\boldsymbol{\upsilon}}_T = [\widetilde{\boldsymbol{\beta}}_T^{+\prime}, \widetilde{\boldsymbol{\beta}}_T^{-\prime}, \widetilde{\vartheta}_T - \widehat{\boldsymbol{\mu}}_T'\widetilde{\boldsymbol{\beta}}_T, \widetilde{\nu}_T]', \tag{10}$$

where

$$\mathbf{P}_T := \begin{bmatrix} \mathbf{I}_{2k} & \mathbf{0} \\ \mathbf{P}_T^{21} & \mathbf{I}_2 \end{bmatrix} \quad \text{and} \quad \mathbf{P}_T^{21} := \begin{bmatrix} -\widehat{\boldsymbol{\mu}}_T' \\ \mathbf{0}_{1\times 2k} \end{bmatrix}.$$

We establish this by once again applying Supplement Lemma A.2 . If we let $(\boldsymbol{x}_{t-1}^{+\prime}, \boldsymbol{x}_{t-1}^{-\prime})'$, $(t-1)$, and 1 be $\boldsymbol{x}_t$, $\boldsymbol{z}_t$, and $\boldsymbol{w}_t$ of Lemma A.2, respectively, then (10) follows. We employ both $\widehat{\boldsymbol{\mu}}_T^+$ and $\widetilde{\boldsymbol{\upsilon}}_T$ to represent $\widehat{\boldsymbol{\upsilon}}_T$ in the form given in (10). The primary purpose of this is to exploit the fact that the limit behaviors of primitive estimators on the right-hand side of (10) do not suffer from an asymptotically singular matrix problem, which we verify in Lemma 2 below.

Furthermore, we represent $\widetilde{\boldsymbol{\omega}}_T$ using other primitive estimators not suffering from asymptotically singular matrix problem. We first combine (4) and (7) to obtain

$$\Delta y_t = \rho_* u_{t-1} + \boldsymbol{\eta}_*'\mathbf{m}_{t-1} + \varsigma_*(t-1) + \gamma_* + \sum_{j=1}^{p-1}\varphi_{j*}\Delta y_{t-j} + \sum_{j=0}^{q-1}\left(\boldsymbol{\pi}_{j*}^{+\prime}\Delta\boldsymbol{x}_{t-j}^+ + \boldsymbol{\pi}_{j*}^{-\prime}\Delta\boldsymbol{x}_{t-j}^-\right) + e_t, \tag{11}$$

where $\varsigma_* := \boldsymbol{\mu}_*'\boldsymbol{\eta}_* + \psi_*$, and then we regress Δy_t against $\dot{\boldsymbol{z}}_t := [\ \widehat{u}_{t-1} \mid \dot{\boldsymbol{z}}_{1t}' \mid \dot{\boldsymbol{z}}_{2t}'\ ]' := [\ \widehat{u}_{t-1} \mid \dot{\boldsymbol{r}}_{t-1}' \mid \boldsymbol{z}_{2t}'\ ]'$ to obtain $\dot{\boldsymbol{\tau}}_T := \left(\sum_{t=1}^{T} \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\right)^{-1} \left(\sum_{t=1}^{T} \dot{\boldsymbol{z}}_t\Delta y_t\right)$, estimating the coefficients in (11). Note that we replaced $\mathbf{m}_{t-1}$ with $\widehat{\mathbf{m}}_{t-1}$ as earlier. For notational simplicity, we partition $\dot{\boldsymbol{\tau}}_T$ as follows:

$$[\ \dot{\rho}_T \mid \dot{\boldsymbol{\eta}}_T' \quad \dot{\varsigma}_T \quad \dot{\gamma}_T \mid \dot{\boldsymbol{\tau}}_{2T}'\ ]' := [\ \dot{\rho}_T \mid \dot{\boldsymbol{\tau}}_{1T}' \mid \dot{\boldsymbol{\tau}}_{2T}'\ ]' := \dot{\boldsymbol{\tau}}_T,$$

estimating $\boldsymbol{\tau}_* := [\ \rho_* \ \vdots\ \boldsymbol{\tau}_{1*}' \ \vdots\ \boldsymbol{\tau}_{2*}'\ ]' := [\ \rho_* \ \vdots\ \boldsymbol{\eta}_*' \quad \varsigma_* \quad \gamma_* \ \vdots\ \boldsymbol{\varphi}_*' \quad \boldsymbol{\pi}_*^{+\prime} \quad \boldsymbol{\pi}_*^{-\prime}\ ]'$. Here, we separately estimate ψ_* by $\dot{\psi}_T := -\widehat{\boldsymbol{\mu}}_T'\dot{\boldsymbol{\eta}}_T + \dot{\varsigma}_T$, from the definition of ς_*. Using these definitions, we relate $\dot{\boldsymbol{\tau}}_T$ to $\widetilde{\boldsymbol{\omega}}_T$ as follows:

$$\widetilde{\boldsymbol{\omega}}_T = \mathbf{Q}_T\dot{\boldsymbol{\tau}}_T, \tag{12}$$

where

$$\mathbf{Q}_T := \begin{bmatrix} \mathbf{I}_{1+2k} & \mathbf{0} \\ \mathbf{Q}_T^{21} & \mathbf{I}_{2kq+1} \end{bmatrix} \quad \text{and} \quad \mathbf{Q}_T^{21} := \begin{bmatrix} 0 & -\widehat{\boldsymbol{\mu}}_T' \\ \mathbf{0}_{(p+2kq)\times 1} & \mathbf{0}_{(p+2kq)\times 2k} \end{bmatrix}.$$

This representation is based on Supplement Lemma A.2. By letting $(\boldsymbol{x}_{t-1}^{+\prime}, \boldsymbol{x}_{t-1}^{-\prime})'$, $(t-1)$, and $[\widehat{u}_{t-1}, 1, \boldsymbol{z}_{2t}']'$ be $\boldsymbol{x}_t$, $\boldsymbol{z}_t$, and $\boldsymbol{w}_t$ of Lemma A.2, respectively, we obtain the desired representation. From Lemma 2 below, $\dot{\boldsymbol{\tau}}_T$ does not suffer from an asymptotically singular matrix problem. Here, $\mathbf{Q}_T\dot{\boldsymbol{\tau}}_T$ is almost identical to $\dot{\boldsymbol{\tau}}_T$, with the only difference in the $(2+2k)$-th rows between $\mathbf{Q}_T\dot{\boldsymbol{\tau}}_T$ and $\dot{\boldsymbol{\tau}}_T$. The $(2+2k)$-th row element of $\mathbf{Q}_T\dot{\boldsymbol{\tau}}_T$ is equal to $\dot{\psi}_T := -\widehat{\boldsymbol{\mu}}_T'\dot{\boldsymbol{\eta}}_T + \dot{\varsigma}_T$, whereas the corresponding element of $\dot{\boldsymbol{\tau}}_T$ is $\dot{\varsigma}_T$. In addition, (12) implies that $\widetilde{\psi}_T = \dot{\psi}_T$, where $\widetilde{\psi}_T$ denotes the $(2+2k)$-th row of $\widetilde{\boldsymbol{\omega}}_T$.

Using these alternative forms for $\widehat{\boldsymbol{v}}_T$ and $\widetilde{\boldsymbol{\omega}}_T$, we represent $\widehat{\boldsymbol{\alpha}}_T$ as a bilinear transform of the other primitive OLS estimators. That is, when we combine (8) and (12), we obtain

$$\widehat{\boldsymbol{\alpha}}_T = \mathbf{T}_T\dot{\boldsymbol{\tau}}_T, \quad \text{where} \quad \mathbf{T}_T := \mathbf{R}_T\mathbf{Q}_T. \tag{13}$$

Here, $\mathbf{T}_T$ is specifically structured as

$$\mathbf{R}_T\mathbf{Q}_T = \mathbf{T}_T := \begin{bmatrix} \mathbf{T}_T^{11} & \mathbf{0} \\ \mathbf{T}_T^{21} & \mathbf{I}_{p+2kq+1} \end{bmatrix},$$

where

$$\mathbf{T}_T^{11} := \begin{bmatrix} 1 & \mathbf{0}_{1\times 2k} \\ -\widehat{\boldsymbol{\beta}}_T & \mathbf{I}_{2k} \end{bmatrix} \quad \text{and} \quad \mathbf{T}_T^{21} := \begin{bmatrix} -\widehat{\zeta}_T & -\widehat{\boldsymbol{\mu}}_T' \\ -\widehat{\nu}_T & \mathbf{0}_{1\times 2k} \\ \mathbf{0}_{(p+2kq-1)\times 1} & \mathbf{0}_{(p+2kq-1)\times 2k} \end{bmatrix}.$$

Thus, we obtain the OLS estimator as follows:

$$\widehat{\boldsymbol{\alpha}}_T = \left[\dot{\rho}_T, -\dot{\rho}_T\widehat{\boldsymbol{\beta}}_T' + \dot{\boldsymbol{\eta}}_T', -\dot{\rho}_T\widehat{\zeta}_T - \widehat{\boldsymbol{\mu}}_T'\dot{\boldsymbol{\eta}}_T + \dot{\varsigma}_T, -\dot{\rho}_T\widehat{\nu}_T + \dot{\gamma}_T, \dot{\boldsymbol{\tau}}_{2T}'\right]'. \tag{14}$$

Remarks. **(a)** The OLS estimator is consistent owing to the consistency of the primitive estimators. If

$\widehat{\boldsymbol{\upsilon}}_T$ and $\dot{\boldsymbol{\tau}}_T$ are consistent for $\boldsymbol{\upsilon}_*$ and $\boldsymbol{\tau}_*$, respectively, then

$$\begin{aligned}\widehat{\boldsymbol{\alpha}}_T \xrightarrow{\mathbb{P}} & \left[\rho_*, -\rho_*\boldsymbol{\beta}_*' + \boldsymbol{\eta}_*', -\rho_*\zeta_* - \boldsymbol{\mu}_*'\boldsymbol{\eta}_* + \varsigma_*, -\rho_*\nu_* + \gamma_*, \boldsymbol{\tau}_{2*}'\right]' \\ &= \left[\rho_*, -\rho_*\boldsymbol{\beta}_*', -\rho_*\zeta_* + \varsigma_*, -\rho_*\nu_* + \gamma_*, \boldsymbol{\tau}_{2*}'\right]' \end{aligned} \tag{15}$$

because $\boldsymbol{\eta}_* = \mathbf{0}$. Also note that the final limit is identical to $[\rho_*, \boldsymbol{\theta}_*', \xi_*, \alpha_*, \boldsymbol{\varphi}_*', \boldsymbol{\pi}_*^{+\prime}, \boldsymbol{\pi}_*^{-\prime}]'$ by the definitions of $\boldsymbol{\theta}_*$, ξ_*, α_*, and $\boldsymbol{\tau}_{2*}$. That is, $\widehat{\boldsymbol{\alpha}}_T$ is consistent owing to the consistency of $\widehat{\boldsymbol{\upsilon}}_T$ and $\dot{\boldsymbol{\tau}}_T$.

(b) We can represent the limit distribution of $\widehat{\boldsymbol{\alpha}}_T$ as the weak limits of the primitive estimators. Theorem 1 demonstrates this below.

(c) As the convergence rates of the primitive estimators are not identical in terms of the parameters, we can demonstrate that the convergence rate of $\widehat{\boldsymbol{\alpha}}_T$ is lower than $\mathbf{D}$. This is the same effect anticipated when applying higher-order expansion to the estimator. For example, from (14), we have $\widehat{\boldsymbol{\theta}}_T = -\dot{\rho}_T\widehat{\boldsymbol{\beta}}_T + \dot{\boldsymbol{\eta}}_T$, while Lemmas 3 and 4 given below detail the conditions for the convergence rates of $\dot{\rho}_T$ and $(\widehat{\boldsymbol{\beta}}_T, \dot{\boldsymbol{\eta}}_T)$ to be $T^{1/2}$ and T, respectively. From this, the convergence rate of $\widehat{\boldsymbol{\theta}}_T$ is determined as $T^{1/2}$, because the limit distribution of $\widehat{\boldsymbol{\theta}}_T$ is determined by an estimator with lower convergence rate. This argument also applies to the other estimators in $\widehat{\boldsymbol{\alpha}}_T$, making the convergence rate lower than $\mathbf{D}$. Thus, we can obtain the limit distribution of $\widehat{\boldsymbol{\alpha}}_T$ without employing higher-order expansion for the OLS estimator. $\square$

Next, we discuss why $\widehat{\boldsymbol{\upsilon}}_T$ and $\dot{\boldsymbol{\tau}}_T$ do not suffer from an asymptotically singular matrix problem. We first define the notation needed for an efficient provision for them. We assume

$$\boldsymbol{\mathscr{B}}(\cdot) := [\boldsymbol{\mathscr{B}}_m(\cdot)', \mathscr{B}_u(\cdot), \mathscr{B}_e(\cdot), \mathscr{B}_{ue}(\cdot), \boldsymbol{\mathscr{B}}_{\boldsymbol{z}e}(\cdot)']' := \boldsymbol{\Sigma}_*^{1/2}\boldsymbol{\mathscr{W}}(\cdot),$$

where $\boldsymbol{\mathscr{W}}(\cdot)$ is a vector of $(2+p+2k(1+q))$ independent standard Wiener processes, and $\boldsymbol{\Sigma}_*$ is the global covariance matrix provided in Assumption 1. Here, $\boldsymbol{\mathscr{B}}(\cdot)$ is the Brownian motion obtained by applying FCLT to $\mathbf{B}_T(\cdot) := T^{-1/2}\sum_{t=1}^{\lfloor(\cdot)T\rfloor}\boldsymbol{w}_t$; Lemma B.1 establishes this in the Online Supplement. For later purpose, we also partition $\mathbf{B}_T(\cdot)$ similarly to $\boldsymbol{\mathscr{B}}(\cdot)$:

$$\mathbf{B}_T(\cdot) := [\mathbf{B}_{mT}(\cdot)', B_{uT}(\cdot), B_{eT}(\cdot), B_{ueT}(\cdot), \mathbf{B}_{\boldsymbol{z}eT}(\cdot)']' := \frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor(\cdot)T\rfloor}[\boldsymbol{s}_{t-1}', u_{t-1}, e_t, u_{t-1}e_t, \boldsymbol{z}_{2t}'e_t]'$$

by noting that $\boldsymbol{w}_t := [\begin{array}{cc:ccc} \boldsymbol{s}_{t-1}' & u_{t-1} & e_t & u_{t-1}e_t & \boldsymbol{z}_{2t}'e_t \end{array}]'$, where $\boldsymbol{s}_t := [\boldsymbol{s}_t^{+\prime}, \boldsymbol{s}_t^{-\prime}]'$.

The following lemma examines the asymptotic behaviors of statistics constituting $\dot{\boldsymbol{\tau}}_T$ and $\widetilde{\boldsymbol{\upsilon}}_T$.

Lemma 2. *Given Assumption 1,*

(i) if $\dot{\mathbf{D}}_1 := diag[T\mathbf{I}_{2k}, T^{3/2}, T^{1/2}]$, $T^{-1/2}\left(\sum_{t=1}^T \widehat{u}_{t-1}\dot{\boldsymbol{z}}_{1t}\right)\dot{\mathbf{D}}_1^{-1} \Rightarrow \dot{\mathbf{M}}_{1u} := \mathbf{0}_{(2k+2)\times 1}$;

(ii) $\dot{\mathbf{D}}_1^{-1}\left(\sum_{t=1}^T \dot{\boldsymbol{z}}_{1t}\dot{\boldsymbol{z}}_{1t}'\right)\dot{\mathbf{D}}_1^{-1} \Rightarrow \dot{\boldsymbol{\mathscr{M}}}_{11}$, *where*

$$\dot{\boldsymbol{\mathscr{M}}}_{11} := \begin{bmatrix} \int \bar{\boldsymbol{\mathscr{B}}}_m\bar{\boldsymbol{\mathscr{B}}}_m' & \mathbf{0}_{k\times 1} & \int \bar{\boldsymbol{\mathscr{B}}}_m \\ \mathbf{0}_{1\times k} & \frac{1}{3} & \frac{1}{2} \\ \int \bar{\boldsymbol{\mathscr{B}}}_m' & \frac{1}{2} & 1 \end{bmatrix}$$

and $\bar{\boldsymbol{\mathscr{B}}}_m(\cdot) := \boldsymbol{\mathscr{B}}_m(\cdot) - 3(\cdot)\int r\boldsymbol{\mathscr{B}}_m$;

(iii) if $\dot{\mathbf{D}}_2 := T^{1/2}\mathbf{I}_{p+2kq-1}$, $T^{-1/2}\left(\sum_{t=1}^T \widehat{u}_{t-1}\dot{\boldsymbol{z}}_{2t}\right)\dot{\mathbf{D}}_2^{-1} \xrightarrow{\mathbb{P}} \dot{\mathbf{M}}_{2u} := \mathbb{E}[u_{t-1}\boldsymbol{z}_{2t}]$;

(iv) $\dot{\mathbf{D}}_2^{-1}\left(\sum_{t=1}^T \dot{\boldsymbol{z}}_{2t}\dot{\boldsymbol{z}}_{1t}'\right)\dot{\mathbf{D}}_1^{-1} \Rightarrow \dot{\boldsymbol{\mathscr{M}}}_{21}$, *where*

$$\dot{\boldsymbol{\mathscr{M}}}_{21} := \begin{bmatrix} \delta_*\boldsymbol{\iota}_{p-1}\int \bar{\boldsymbol{\mathscr{B}}}_m' & \frac{1}{2}\delta_*\boldsymbol{\iota}_{p-1} & \delta_*\boldsymbol{\iota}_{p-1} \\ \boldsymbol{\iota}_q\otimes\boldsymbol{\mu}_*\int \bar{\boldsymbol{\mathscr{B}}}_m' & \frac{1}{2}\boldsymbol{\iota}_q\otimes\boldsymbol{\mu}_* & \boldsymbol{\iota}_q\otimes\boldsymbol{\mu}_* \end{bmatrix};$$

(v) $\dot{\mathbf{D}}_2^{-1}\left(\sum_{t=1}^T \dot{\boldsymbol{z}}_{2t}\dot{\boldsymbol{z}}_{2t}'\right)\dot{\mathbf{D}}_2^{-1} \xrightarrow{\mathbb{P}} \dot{\mathbf{M}}_{22} := \mathbf{M}_{22}$;

(vi) if $\dot{\mathbf{D}} := diag[T^{1/2}, \dot{\mathbf{D}}_1, \dot{\mathbf{D}}_2]$, $\dot{\mathbf{D}}^{-1}\left(\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\right)\dot{\mathbf{D}}^{-1} \Rightarrow \dot{\boldsymbol{\mathscr{M}}}$, *where*

$$\dot{\boldsymbol{\mathscr{M}}} := \begin{bmatrix} \sigma_u^2 & \dot{\mathbf{M}}_{u1} & \dot{\mathbf{M}}_{u2} \\ \dot{\mathbf{M}}_{1u} & \dot{\boldsymbol{\mathscr{M}}}_{11} & \dot{\boldsymbol{\mathscr{M}}}_{12} \\ \dot{\mathbf{M}}_{2u} & \dot{\boldsymbol{\mathscr{M}}}_{21} & \dot{\mathbf{M}}_{22} \end{bmatrix},$$

$\dot{\boldsymbol{\mathscr{M}}}_{12} := \dot{\boldsymbol{\mathscr{M}}}_{21}'$, $\dot{\mathbf{M}}_{u1} := \dot{\mathbf{M}}_{1u}'$, $\dot{\mathbf{M}}_{u2} := \dot{\mathbf{M}}_{2u}'$, *and* $\sigma_u^2 := \mathbb{E}[u_t^2]$; *and*

(vii) $\dot{\mathbf{D}}_1^{-1}\left(\sum_{t=1}^T \dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}'\right)\dot{\mathbf{D}}_1^{-1} \Rightarrow \dot{\boldsymbol{\mathscr{M}}}_{11}$. □

Note that $\dot{\boldsymbol{\tau}}_T$ and $\widetilde{\boldsymbol{\upsilon}}_T$ can be defined by $(\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t')^{-1}$ and $(\sum_{t=1}^T \dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}')^{-1}$, respectively, while Lemmas 2 (*vi* and *vii*) confirm that their respective weak limits are obtained without involving the asymptotically singular matrix problem. This implies that the asymptotic singular matrix problem is not due to $\dot{\boldsymbol{\tau}}_T$ and $\widetilde{\boldsymbol{\upsilon}}_T$.

4 Limit Distribution of OLS

In this section, we derive the limit distribution of $\widehat{\boldsymbol{\alpha}}_T$ using the weak limits of primitive estimators. We further show the consistency of $(\widehat{\boldsymbol{\upsilon}}_T, \dot{\boldsymbol{\tau}}_T)$ for $(\boldsymbol{\upsilon}_*, \boldsymbol{\tau}_*)$, from which the consistency of $\widehat{\boldsymbol{\alpha}}_T$ follows.

We first examine the limit distribution of $\widehat{\boldsymbol{\upsilon}}_T$. For this, we note that $\widehat{\boldsymbol{\upsilon}}_T = \mathbf{P}_T\widetilde{\boldsymbol{\upsilon}}_T$ from (10). The limit distribution of $\widehat{\boldsymbol{\upsilon}}_T$ is then determined by each element on the right-hand side. If we let $\mathbf{P}$ be the limit of $\mathbf{P}_T$, that is,

$$\mathbf{P} := \begin{bmatrix} \mathbf{I}_{2k} & \mathbf{0} \\ \mathbf{P}^{21} & \mathbf{I}_2 \end{bmatrix} \quad \text{and} \quad \mathbf{P}^{21} := \begin{bmatrix} -\boldsymbol{\mu}_*' \\ \mathbf{0}_{1\times 2k} \end{bmatrix},$$

its consistency follows from the consistency of $\widehat{\boldsymbol{\mu}}_T$ for $\boldsymbol{\mu}_*$. Next, to obtain the limit distribution of $\widetilde{\boldsymbol{\upsilon}}_T$, we first note that $\mathbf{m}_t = \widehat{\mathbf{m}}_t + (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)t$ from the definitions of $\widehat{\mathbf{m}}_t$ and $\widehat{\boldsymbol{\mu}}_T$. If we rewrite (9) as

$$y_t = \boldsymbol{\beta}_*'\widehat{\mathbf{m}}_t + \vartheta_{T*}t + \nu_* + u_t, \tag{16}$$

where $\vartheta_{T*} := \boldsymbol{\beta}_*'(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) + \vartheta_*$, it follows that

$$\widetilde{\boldsymbol{\upsilon}}_T := \left(\sum_{t=1}^{T} \dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}'\right)^{-1}\left(\sum_{t=1}^{T} \dot{\boldsymbol{r}}_{t-1}y_{t-1}\right) = \bar{\boldsymbol{\upsilon}}_{T*} + \left(\sum_{t=1}^{T} \dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}'\right)^{-1}\left(\sum_{t=1}^{T} \dot{\boldsymbol{r}}_{t-1}u_{t-1}\right), \tag{17}$$

where $\bar{\boldsymbol{\upsilon}}_{T*} := [\boldsymbol{\beta}_*', \vartheta_{T*}, \nu_*]'$. By exploiting this arrangement, we can obtain the limit distributions of $\widetilde{\boldsymbol{\upsilon}}_T$ and $\widehat{\boldsymbol{\upsilon}}_T$, which are contained in the following lemma:

Lemma 3. *Let* $\boldsymbol{\varrho}_{m*} := \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T}\sum_{\tau=1}^{t-1}\mathbb{E}[\boldsymbol{s}_\tau u_t]$. *Given Assumption 1,*

(i) $\dot{\mathbf{D}}_1(\widetilde{\boldsymbol{\upsilon}}_T - \bar{\boldsymbol{\upsilon}}_{T*}) \Rightarrow [\dot{\boldsymbol{\mathscr{L}}}_{11}', \dot{\boldsymbol{\mathscr{L}}}_{12}', \dot{\mathscr{L}}_{13}, \dot{\mathscr{L}}_{14}]' := \dot{\boldsymbol{\mathscr{L}}}_1 := \dot{\boldsymbol{\mathscr{M}}}_{11}^{-1}\dot{\boldsymbol{\mathscr{S}}}_1$, *where*

$$\dot{\boldsymbol{\mathscr{S}}}_1 := \begin{bmatrix} \dot{\boldsymbol{\mathscr{S}}}_{11} \\ \dot{\boldsymbol{\mathscr{S}}}_{12} \\ \hdashline \dot{\mathscr{S}}_{13} \\ \dot{\mathscr{S}}_{14} \end{bmatrix} := \begin{bmatrix} \int \boldsymbol{\mathscr{B}}_m d\mathscr{B}_u - 3\int r\boldsymbol{\mathscr{B}}_m \int r d\mathscr{B}_u + \boldsymbol{\varrho}_{m*} \\ \hdashline \int r d\mathscr{B}_u \\ \int d\mathscr{B}_u \end{bmatrix}$$

such that $\dot{\boldsymbol{\mathscr{L}}}_{11}$, $\dot{\boldsymbol{\mathscr{L}}}_{12}$, $\dot{\boldsymbol{\mathscr{S}}}_{11}$, *and* $\dot{\boldsymbol{\mathscr{S}}}_{12} \in \mathbb{R}^k$; *and*

(ii) $\dot{\mathbf{D}}_\dagger(\widehat{\boldsymbol{\upsilon}}_T - \boldsymbol{\upsilon}_*) \Rightarrow [\dot{\boldsymbol{\mathscr{L}}}_{11}', \dot{\boldsymbol{\mathscr{L}}}_{12}', -\boldsymbol{\mu}_*^{+\prime}\dot{\boldsymbol{\mathscr{L}}}_{11} - \boldsymbol{\mu}_*^{-\prime}\dot{\boldsymbol{\mathscr{L}}}_{12}, \dot{\mathscr{L}}_{14}]'$, *where* $\dot{\mathbf{D}}_\dagger :=$ *diag*$[T\mathbf{I}_{2k}, T, T^{1/2}]$. □

Remarks. **(a)** We prove Lemma 3 (*i*) as follows. From (17), it follows that $\dot{\mathbf{D}}_1(\widetilde{\boldsymbol{\upsilon}}_T - \bar{\boldsymbol{\upsilon}}_{T*}) = (\dot{\mathbf{D}}_1^{-1}\sum_{t=1}^{T} \dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}'\dot{\mathbf{D}}_1^{-1})^{-1}\dot{\mathbf{D}}_1^{-1}\sum_{t=1}^{T}\dot{\boldsymbol{r}}_{t-1}u_{t-1}$. We focus on deriving the weak limit of $\dot{\mathbf{D}}_1^{-1}\sum_{t=1}^{T}\dot{\boldsymbol{r}}_{t-1}u_{t-1}$, because Lemma 2 (*vii*) already provides the weak limit of the inverse matrix. To prove Lemma 3 (*ii*), we note that $\widehat{\boldsymbol{\upsilon}}_T - \boldsymbol{\upsilon}_* = (\mathbf{P}_T - \mathbf{P})(\widetilde{\boldsymbol{\upsilon}}_T - \bar{\boldsymbol{\upsilon}}_{T*}) + \mathbf{P}(\widetilde{\boldsymbol{\upsilon}}_T - \bar{\boldsymbol{\upsilon}}_{T*}) + (\mathbf{P}_T - \mathbf{P})(\bar{\boldsymbol{\upsilon}}_{T*} - \bar{\boldsymbol{\upsilon}}_*) + \mathbf{P}(\bar{\boldsymbol{\upsilon}}_{T*} - \bar{\boldsymbol{\upsilon}}_*) + (\mathbf{P}_T - \mathbf{P})\bar{\boldsymbol{\upsilon}}_*$ and derive the weak limit of each component on the right-hand side.

(b) From Lemma 3, the consistency of parameter estimators follows. Lemmas 3 (*i* and *ii*) imply that $\widetilde{\boldsymbol{\upsilon}}_T - \bar{\boldsymbol{\upsilon}}_{T*} \xrightarrow{\mathbb{P}} \mathbf{0}$ and $\widehat{\boldsymbol{\upsilon}}_T \xrightarrow{\mathbb{P}} \boldsymbol{\upsilon}_*$, respectively. We also have $\bar{\boldsymbol{\upsilon}}_{T*} \xrightarrow{\mathbb{P}} \bar{\boldsymbol{\upsilon}}_*$, as $\vartheta_{T*} := \boldsymbol{\beta}_*'(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) + \vartheta_* \xrightarrow{\mathbb{P}} \vartheta_*$, because $\widehat{\boldsymbol{\mu}}_T = \boldsymbol{\mu}_* + O_{\mathbb{P}}(T^{-1/2})$, as established by Supplement Lemma B.7. This also implies that $\mathbf{P}_T \xrightarrow{\mathbb{P}} \mathbf{P}$, so that $\widehat{\boldsymbol{\upsilon}}_T = \mathbf{P}_T\widetilde{\boldsymbol{\upsilon}}_T \to \mathbf{P}\bar{\boldsymbol{\upsilon}}_*$, which is identical to $\boldsymbol{\upsilon}_*$. From this, it follows that $\widehat{\boldsymbol{\upsilon}}_T \xrightarrow{\mathbb{P}} \boldsymbol{\upsilon}_*$.

(c) The weak limit of $\widehat{\boldsymbol{\upsilon}}_T$ in Lemma 3 (*ii*) implies that $\widetilde{\boldsymbol{\beta}}_T$ and $\widetilde{\vartheta}_T$ are linearly correlated asymptotically. Specifically, their weak limits are obtained as $(\dot{\boldsymbol{\mathscr{L}}}_{11}, \dot{\boldsymbol{\mathscr{L}}}_{12})$ and $-\boldsymbol{\mu}_*^{+\prime}\dot{\boldsymbol{\mathscr{L}}}_{11} - \boldsymbol{\mu}_*^{-\prime}\dot{\boldsymbol{\mathscr{L}}}_{12}$, respectively. By this, $\widehat{\boldsymbol{\upsilon}}_T$ suffers from the asymptotically singular matrix problem, although $\widetilde{\boldsymbol{\upsilon}}_T$ does not. Supplement Lemma A.3 also examines the asymptotic behavior of $\sum_{t=1}^T \widetilde{\boldsymbol{z}}_t\widetilde{\boldsymbol{z}}_t'$ and confirm the asymptotic singular matrix problem. $\square$

We next examine the limit behavior of $\dot{\boldsymbol{\tau}}_T$. We first note that

$$u_t = \widehat{u}_t + (\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)'\widehat{\mathbf{m}}_t + (\widetilde{\vartheta}_T - (\boldsymbol{\beta}_*'\widehat{\boldsymbol{\mu}}_T + \zeta_*))t + (\widetilde{\nu}_T - \nu_*)$$

from (16) because $\widehat{u}_t := y_t - \widetilde{\boldsymbol{r}}_t'\widehat{\boldsymbol{\upsilon}}_T = y_t - \dot{\boldsymbol{r}}_t'\widetilde{\boldsymbol{\upsilon}}_T$. Therefore, from (11), we have $\Delta y_t = \boldsymbol{\tau}_{T*}'\dot{\boldsymbol{z}}_t + e_t$, assuming that $\boldsymbol{\tau}_{T*} := [\rho_*, \boldsymbol{\tau}_{1T}', \boldsymbol{\tau}_{2*}']'$ and

$$\boldsymbol{\tau}_{1T} := [(\boldsymbol{\eta}_* + \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*))', \psi_* + \boldsymbol{\eta}_*'\widehat{\boldsymbol{\mu}}_T + \rho_*(\widetilde{\vartheta}_T - \vartheta_{T*}), \gamma_* + \rho_*(\widetilde{\nu}_T - \nu_*)]'.$$

From this, we have

$$\dot{\boldsymbol{\tau}}_T := \left(\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\right)^{-1}\left(\sum_{t=1}^T \dot{\boldsymbol{z}}_t\Delta y_t\right) = \boldsymbol{\tau}_{T*} + \left(\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\right)^{-1}\left(\sum_{t=1}^T \dot{\boldsymbol{z}}_t e_t\right). \tag{18}$$

We present the limit distribution of $\dot{\boldsymbol{\tau}}_T$ in the following lemma:

Lemma 4. *Given Assumption 1,*

(i) $\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) \Rightarrow [\dot{\mathscr{L}}_1, \dot{\boldsymbol{\mathscr{L}}}_2', \dot{\boldsymbol{\mathscr{L}}}_3', \dot{\mathscr{L}}_4, \dot{\mathscr{L}}_5, \dot{\boldsymbol{\mathscr{L}}}_6']' := \dot{\boldsymbol{\mathscr{L}}} := \dot{\boldsymbol{\mathscr{M}}}^{-1}\dot{\boldsymbol{\mathscr{S}}}$, *where*

$$\dot{\boldsymbol{\mathscr{S}}} := \left[\begin{array}{c:cc:ccc} \dot{\mathscr{S}}_1 & \dot{\boldsymbol{\mathscr{S}}}_2' & \dot{\boldsymbol{\mathscr{S}}}_3' & \dot{\mathscr{S}}_4 & \dot{\mathscr{S}}_5 & \dot{\boldsymbol{\mathscr{S}}}_6' \end{array}\right]' := \left[\begin{array}{c:c:ccc} \int d\mathscr{B}_{ue} & \int \bar{\boldsymbol{\mathscr{B}}}_m' d\mathscr{B}_e & \int r d\mathscr{B}_e & \int d\mathscr{B}_e & \int d\boldsymbol{\mathscr{B}}_{ze}' \end{array}\right]'$$

such that $\dot{\boldsymbol{\mathscr{L}}}_2$, $\dot{\boldsymbol{\mathscr{L}}}_3$, $\dot{\boldsymbol{\mathscr{S}}}_2$, *and* $\dot{\boldsymbol{\mathscr{S}}}_3 \in \mathbb{R}^k$*; and*

(ii) $\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*) \Rightarrow \dot{\boldsymbol{\mathscr{L}}} + \rho_*[0, \dot{\boldsymbol{\mathscr{L}}}_{11}', \dot{\boldsymbol{\mathscr{L}}}_{12}', \dot{\mathscr{L}}_{13}, \dot{\mathscr{L}}_{14}, \mathbf{0}']'$. $\square$

Remarks. **(a)** From the definition of $\dot{\boldsymbol{\mathscr{L}}}$, $\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*})$ asymptotically follows a mixed normal distribution.

(b) To prove Lemma 4, we examine the limit behavior of each component on the right-hand side of equation (18). First, we note that $\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) = (\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\dot{\mathbf{D}}^{-1})^{-1}\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t e_t$. As Lemma 2 (*vi*) already shows that $\dot{\mathbf{D}}^{-1}(\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t')\dot{\mathbf{D}}^{-1} \Rightarrow \dot{\boldsymbol{\mathscr{M}}}$, we focus on deriving the weak limit of $\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t e_t$. Noting that $\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) = \dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*) + [0, -\rho_* T(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)', -\rho_* T^{3/2}(\widetilde{\vartheta}_T - \vartheta_{T*}), -\rho_* T^{1/2}(\widetilde{\nu}_T - \nu_*), \mathbf{0}']'$, we exploit Lemmas 3 (*ii*) and 4 (*i*), to obtain the weak limit of the right-hand side.

(c) We derive Lemma 4 assuming that $\rho_* < 0$. Therefore, we cannot use Lemma 4 (*ii*) to obtain the null limit distribution of the t-statistic testing $\rho_* = 0$. $\square$

We finally provide the limit distribution of $\widehat{\boldsymbol{\alpha}}_T$. We note that (14) and (15) imply that

$$(\widehat{\boldsymbol{\alpha}}_T - \boldsymbol{\alpha}_*) = (\mathbf{T}_T - \mathbf{T})\boldsymbol{\tau}_* + \mathbf{T}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*) + (\mathbf{T}_T - \mathbf{T})(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*), \tag{19}$$

where $\mathbf{T}$ is the probability limit of $\mathbf{T}_T$, namely,

$$\mathbf{T} := \begin{bmatrix} \mathbf{T}^{11} & \mathbf{0} \\ \mathbf{T}^{21} & \mathbf{I}_{p+2kq+1} \end{bmatrix}, \quad \mathbf{T}^{11} := \begin{bmatrix} 1 & \mathbf{0}_{1\times 2k} \\ -\boldsymbol{\beta}_* & \mathbf{I}_{2k} \end{bmatrix},$$

and

$$\mathbf{T}^{21} := \begin{bmatrix} -\zeta_* & -\boldsymbol{\mu}_*' \\ -\nu_* & \mathbf{0}_{1\times 2k} \\ \mathbf{0}_{(p+2kq-1)\times 1} & \mathbf{0}_{(p+2kq-1)\times 2k} \end{bmatrix},$$

from which we obtain that $\sqrt{T}(\widehat{\boldsymbol{\alpha}}_T - \boldsymbol{\alpha}_*) = \boldsymbol{c}_*\sqrt{T}(\dot{\rho}_T - \rho_*) + \sqrt{T}\boldsymbol{d}_T + o_{\mathbb{P}}(1)$, where $\boldsymbol{c}_* := [1, -\boldsymbol{\beta}_*', -\zeta_*, -\nu_*, \mathbf{0}']'$ and $\boldsymbol{d}_T := [0, \mathbf{0}_{2k\times 1}', 0, \{(\dot{\gamma}_T - \gamma_*) - \rho_*(\widetilde{\nu}_T - \nu_*)\}, (\dot{\boldsymbol{\tau}}_{2T} - \boldsymbol{\tau}_{2*})']'$. We next provide the weak limit of $\widehat{\boldsymbol{\alpha}}_T$ using the weak limits of $\sqrt{T}(\dot{\rho}_T - \rho_*)$ and $\sqrt{T}\boldsymbol{d}_T$ in the following theorem.

Theorem 1. *Given Assumption 1,*

(i) if for each $j = 1, 2, \ldots, k$, $\beta_{j*}^+ \neq 0$, $\beta_{j*}^- \neq 0$, and $\zeta_* \neq 0$, then $\sqrt{T}(\widehat{\boldsymbol{\alpha}}_T - \boldsymbol{\alpha}_*) \Rightarrow \boldsymbol{c}_*\dot{\mathscr{L}}_1 + [0, \mathbf{0}_{k\times 1}', \mathbf{0}_{k\times 1}', 0, \dot{\mathscr{L}}_5, \dot{\boldsymbol{\mathscr{L}}}_6']'$, where β_{j*}^+ and β_{j*}^- are the j-th row element of $\boldsymbol{\beta}_*^+$ and $\boldsymbol{\beta}_*^-$, respectively;

(ii) if $\boldsymbol{\beta}_*^+ = \mathbf{0}$, but for each $j = 1, 2, \ldots, k$, $\beta_{j*}^- \neq 0$, and $\zeta_* \neq 0$, then $\dot{\mathbf{D}}_+(\widehat{\boldsymbol{\alpha}}_T - \boldsymbol{\alpha}_*) \Rightarrow [\dot{\mathscr{L}}_1, \dot{\boldsymbol{\mathscr{L}}}_2', -\boldsymbol{\beta}_*^{-\prime}\dot{\mathscr{L}}_1, -\zeta_*\dot{\mathscr{L}}_1, \dot{\mathscr{L}}_5 - \nu_*\dot{\mathscr{L}}_1, \dot{\boldsymbol{\mathscr{L}}}_6']'$, where $\dot{\mathbf{D}}_+ := diag[T^{1/2}, T\mathbf{I}_k, T^{1/2}\mathbf{I}_{k+2}, \dot{\mathbf{D}}_2]$;

(iii) if $\boldsymbol{\beta}_*^- = \mathbf{0}$, but for each $j = 1, 2, \ldots, k$, $\beta_{j*}^+ \neq 0$, and $\zeta_* \neq 0$, then $\dot{\mathbf{D}}_-(\widehat{\boldsymbol{\alpha}}_T - \boldsymbol{\alpha}_*) \Rightarrow [\dot{\mathscr{L}}_1, -\boldsymbol{\beta}_*^{+\prime}\dot{\mathscr{L}}_1, \dot{\boldsymbol{\mathscr{L}}}_3', -\zeta_*\dot{\mathscr{L}}_1, \dot{\mathscr{L}}_5 - \nu_*\dot{\mathscr{L}}_1, \dot{\boldsymbol{\mathscr{L}}}_6']'$, where $\dot{\mathbf{D}}_- := diag[T^{1/2}\mathbf{I}_{k+1}, T\mathbf{I}_k, T^{1/2}\mathbf{I}_2, \dot{\mathbf{D}}_2]$; and

(iv) if for each $j = 1, 2, \ldots, k$, $\beta_{j*}^+ \neq 0$, $\beta_{j*}^- \neq 0$, but $\zeta_* = 0$, $\dot{\mathbf{D}}_\odot(\widehat{\boldsymbol{\alpha}}_T - \boldsymbol{\alpha}_*) \Rightarrow [\dot{\mathscr{L}}_1, -\boldsymbol{\beta}_*^{+\prime}\dot{\mathscr{L}}_1, -\boldsymbol{\beta}_*^{-\prime}\dot{\mathscr{L}}_1, -\boldsymbol{\mu}_*^{+\prime}\dot{\boldsymbol{\mathscr{L}}}_2 - \boldsymbol{\mu}_*^{-\prime}\dot{\boldsymbol{\mathscr{L}}}_3, \dot{\mathscr{L}}_5 - \nu_*\dot{\mathscr{L}}_1, \dot{\boldsymbol{\mathscr{L}}}_6']'$, where $\dot{\mathbf{D}}_\odot := diag[T^{1/2}\mathbf{I}_{2k+1}, T, T^{1/2}, \dot{\mathbf{D}}_2]$. $\square$

Remarks. **(a)** Although the time trend and nonstationary regressors are included as regressors on the right-hand side, the convergence rate of $\widehat{\boldsymbol{\alpha}}_T$ is slower than $\mathbf{D}$ by Theorem 1 (*i*).

(b) In the Online Supplement, we prove Theorem 1 (*i*) by deriving the weak limit of each component on the right-hand side of (19). Note that the limit distribution of $\widehat{\boldsymbol{\alpha}}_T$ differs from the parameter estimation for ARDL models without an asymptotically singular matrix problem. For example, if we focus on

$\widehat{\boldsymbol{\theta}}_T$, as shown in the Online Supplement,

$$(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_*) = -\boldsymbol{\beta}_*(\dot{\rho}_T - \rho_*) + (\dot{\boldsymbol{\eta}}_T - \boldsymbol{\eta}_*) - \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*), \tag{20}$$

where $(\dot{\rho}_T - \rho_*) = O_{\mathbb{P}}(T^{-1/2})$, and $\dot{\boldsymbol{\eta}}_T - \boldsymbol{\eta}_* - \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*) = O_{\mathbb{P}}(T^{-1})$, and so Lemmas 3 and 4, it follows that $\sqrt{T}(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_*) \Rightarrow -\boldsymbol{\beta}_*\dot{\mathscr{L}}_1$, leading to Theorem 1 (*i*).

(c) By Theorem 1 (*i*), the weak limit of $[\widehat{\rho}_T, \widehat{\boldsymbol{\theta}}'_T, \widehat{\xi}_T]'$ is $[1, \boldsymbol{\beta}'_*, -\zeta_*]'\dot{\mathscr{L}}_1$, indicating that the estimates are linearly correlated at the limit. This result highlights how the asymptotically singular matrix problem affects the convergence rate of $\widehat{\boldsymbol{\alpha}}_T$. Although $\dot{\mathscr{L}}_1$ is not associated with an asymptotically singular matrix problem, its bilinear transform affects $\widehat{\boldsymbol{\alpha}}_T$ owing to the problem.

(d) Despite the asymptotic singularity problem associated with $\widehat{\boldsymbol{\alpha}}_T$, its weak limit given in Theorem 1 (*i*) is determined by $\dot{\mathscr{L}}_1$, $\dot{\mathscr{L}}_5$, and $\dot{\boldsymbol{\mathscr{L}}}_6$. This implies the following. First, $\widehat{\boldsymbol{\alpha}}_T$ follows a mixed normal distribution by Lemma 4. Therefore, if the standard t-test applies to $\widehat{\boldsymbol{\alpha}}_T$, it follows a mixed normal distribution under the null hypothesis and the condition in Theorem 1 (*i*). Second, both $\dot{\mathscr{L}}_1$ and $\dot{\boldsymbol{\mathscr{L}}}_6$ are weak limits of the last two OLS estimators obtained by regressing Δy_t against $(1, u_{t-1}, \boldsymbol{z}'_{2t})'$. Furthermore, they are also weak limits of the corresponding OLS estimators obtained by replacing the regressor with $(1, \widehat{u}_{t-1}, \boldsymbol{z}'_{2t})'$. Therefore, the null weak limit of the t-statistic testing $\rho_* < 0$ is equivalent to the weak limits of t-statistics testing long-run parameters $\boldsymbol{\beta}^+_*$, $\boldsymbol{\beta}^-_*$, and ξ_* under their respective null hypotheses.

(e) If the zero coefficient condition in Theorem 1 (*ii*) holds, the limit distribution of $\widehat{\boldsymbol{\alpha}}_T$ is determined by the next-order term in (19). For instance, if $\boldsymbol{\beta}^+_* = \mathbf{0}$, then (20) implies that

$$T(\widehat{\boldsymbol{\theta}}^+_T - \boldsymbol{\theta}^+_*) = T\{(\dot{\boldsymbol{\eta}}^+_T - \boldsymbol{\eta}^+_*) - \rho_*(\widetilde{\boldsymbol{\beta}}^+_T - \boldsymbol{\beta}^+_*)\} \Rightarrow \dot{\boldsymbol{\mathscr{L}}}_2$$

by Lemma 4 (*i*). In parallel, if $\boldsymbol{\beta}^-_* = \mathbf{0}$, then $T(\widehat{\boldsymbol{\theta}}^-_T - \boldsymbol{\theta}^-_*) = T\{(\dot{\boldsymbol{\eta}}^-_T - \boldsymbol{\eta}^-_*) - \rho_*(\widetilde{\boldsymbol{\beta}}^-_T - \boldsymbol{\beta}^-_*)\} \Rightarrow \dot{\boldsymbol{\mathscr{L}}}_3$; and if $\zeta_* = 0$, then $T(\widehat{\xi}_T - \xi_*) = -\boldsymbol{\mu}'_* T\{(\dot{\boldsymbol{\eta}}_T - \boldsymbol{\eta}_*) - \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)\} + O_{\mathbb{P}}(T^{-1/2}) \Rightarrow -\boldsymbol{\mu}^{+\prime}_*\dot{\boldsymbol{\mathscr{L}}}_2 - \boldsymbol{\mu}^{-\prime}_*\dot{\boldsymbol{\mathscr{L}}}_3$. Note that the weak limits of $\widehat{\boldsymbol{\theta}}^+_T$, $\widehat{\boldsymbol{\theta}}^-_T$, and $\widehat{\xi}_T$ are determined by $\dot{\boldsymbol{\mathscr{L}}}_2$ and $\dot{\boldsymbol{\mathscr{L}}}_3$ that follow mixed normal distributions centered at zero. This implies that if the standard t-test applies to $\widehat{\boldsymbol{\theta}}^-_T$ or $\widehat{\xi}_T$, its null weak limit is a mixed normal distribution centered at zero.

(f) Caution is required for Theorems 1 (*ii*, *iii*, and *iv*). If $\boldsymbol{\beta}^+_* = \boldsymbol{\beta}^-_* = \mathbf{0}$ and $\zeta_* = 0$, then $\rho_* = 0$ by the remark given below Assumption 1, contradicting the assumption that $\rho_* < 0$. Theorems 1 (*ii*, *iii*, and *iv*) assume an environment where all of $\boldsymbol{\beta}^+_*$, $\boldsymbol{\beta}^-_*$ and ζ_* are not zero. As mentioned in **(e)**, Theorems 1 (*iii* and *iv*) also determine the OLS estimator limit distribution by next-order approximation. This implies that the null limit distribution of t-test testing the zero coefficient of $\boldsymbol{\beta}^+_*$, $\boldsymbol{\beta}^-_*$, or ζ_* has to be determined by the limit distributions in Theorems 1 (*ii*, *iii*, and *iv*).

(g) Pesaran et al. (2001) and Banerjee et al. (1998) provide the asymptotic critical values of the F- and t-statistics testing $\rho_* = 0$, showing that their null limit distributions cannot be approximated by a mixed normal distribution. □

5 Hypotheses Testing

In this section, we examine testing hypotheses associated with NARDL by applying the Wald test principle. In particular, we suppose that a cointegrating relationship holds between y_t and $(\boldsymbol{x}_t^+, \boldsymbol{x}_t^-)$ under Assumption 1 because $\boldsymbol{\theta}_*^+ \neq \mathbf{0}$ and/or $\boldsymbol{\theta}_*^- \neq \mathbf{0}$.

The NARDL process reduces to the ARDL process when $\boldsymbol{\theta}_*^+ = \boldsymbol{\theta}_*^-$ and $\boldsymbol{\pi}_*^+ = \boldsymbol{\pi}_*^-$. Therefore, under these conditions, it would be inefficient to estimate the parameters by NARDL, making it necessary to test the symmetry conditions. For this, we specify the following three hypothesis systems:

$$\mathscr{H}_0' : \boldsymbol{\theta}_*^+ = \boldsymbol{\theta}_*^- \quad \text{vs.} \quad \mathscr{H}_1' : \boldsymbol{\theta}_*^+ \neq \boldsymbol{\theta}_*^-;$$

$$\mathscr{H}_0'' : \boldsymbol{\pi}_*^+ = \boldsymbol{\pi}_*^- \quad \text{vs.} \quad \mathscr{H}_1'' : \boldsymbol{\pi}_*^+ \neq \boldsymbol{\pi}_*^-;$$

$$\mathscr{H}_0''' : \boldsymbol{\theta}_*^+ = \boldsymbol{\theta}_*^- \ \text{ and } \ \boldsymbol{\pi}_*^+ = \boldsymbol{\pi}_*^- \quad \text{vs.} \quad \mathscr{H}_1''' : \boldsymbol{\theta}_*^+ \neq \boldsymbol{\theta}_*^- \ \text{ or } \ \boldsymbol{\pi}_*^+ \neq \boldsymbol{\pi}_*^-.$$

We examine testing of each hypothesis. We first apply the Wald test principle to testing $\mathscr{H}_0'$. The standard Wald test applied to OLS is defined as follows:

$$W_T^{(1)} := \widehat{\boldsymbol{\alpha}}_T' \dot{\mathbf{R}}_1' \left(\dot{\mathbf{W}}_T^{(1)}\right)^{-1} \dot{\mathbf{R}}_1 \widehat{\boldsymbol{\alpha}}_T, \quad \text{where} \quad \dot{\mathbf{W}}_T^{(1)} := \dot{\sigma}_{e,T}^2 \dot{\mathbf{R}}_1 \left(\sum_{t=1}^T \boldsymbol{z}_t \boldsymbol{z}_t'\right)^{-1} \dot{\mathbf{R}}_1'$$

and $\dot{\mathbf{R}}_1 := [\mathbf{0}_{k\times 1}, \mathbf{I}_k, -\mathbf{I}_k, \mathbf{0}_{k\times(1+p+2kq)}]$. Note that $\dot{\mathbf{R}}_1 \widehat{\boldsymbol{\alpha}}_T = (\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-)$, so that $\sqrt{T}(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) \Rightarrow (\boldsymbol{\beta}_*^- - \boldsymbol{\beta}_*^+)\dot{\mathscr{L}}_1$, which is equal to $\mathbf{0}$ under $\mathscr{H}_0'$ by noting that $\mathscr{H}_0'$ implies that $\boldsymbol{\beta}_*^- = \boldsymbol{\beta}_*^+$. Thus, the limit distribution in Theorem 1 (*i*) is not useful to obtain the null limit distribution of the Wald test. The null limit behavior has to be obtained by the next order term of $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-)$. Specifically, (14) implies that $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) = (\dot{\boldsymbol{\eta}}_T^+ - \dot{\boldsymbol{\eta}}_T^-) - \dot{\rho}_T(\widehat{\boldsymbol{\beta}}_T^+ - \widehat{\boldsymbol{\beta}}_T^-) = (\dot{\boldsymbol{\eta}}_T^+ - \dot{\boldsymbol{\eta}}_T^-) - (\boldsymbol{\eta}_*^+ - \boldsymbol{\eta}_*^-) - \rho_*(\widehat{\boldsymbol{\beta}}_T^+ - \widehat{\boldsymbol{\beta}}_T^-) - (\dot{\rho}_T - \rho_*)(\widehat{\boldsymbol{\beta}}_T^+ - \widehat{\boldsymbol{\beta}}_T^-) = (\dot{\boldsymbol{\eta}}_T^+ - \dot{\boldsymbol{\eta}}_T^-) - (\boldsymbol{\eta}_{T*}^+ - \boldsymbol{\eta}_{T*}^-) + o_{\mathbb{P}}(T^{-1})$ under $\mathscr{H}_0'$, where $(\boldsymbol{\eta}_{T*}^{+\prime}, \boldsymbol{\eta}_{T*}^{-\prime})'$ denotes the vector formed by the first $2k$-row elements of $\boldsymbol{\tau}_{1T}$. That is, the null limit distribution of $W_T^{(1)}$ has to be determined not by Theorem 1, but by Lemma 4. Therefore, $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) = O_{\mathbb{P}}(T^{-1})$. However, the alternative behavior of $\dot{\mathbf{R}}_1 \widehat{\boldsymbol{\alpha}}_T = (\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-)$ is determined by $O_{\mathbb{P}}(T^{-1/2})$ terms.

We supplement the unbalanced convergence rates under the two hypotheses by further applying the Wald test principle to the primitive estimator $\widetilde{\boldsymbol{\upsilon}}_T$. This is mainly because $\widetilde{\boldsymbol{\upsilon}}_T$ enables us to define the Wald test using the next-order terms. Note that because testing $\mathscr{H}_0'$ is equivalent to testing $\mathbb{H}_0' : \boldsymbol{\beta}_*^+ - \boldsymbol{\beta}_*^- = \mathbf{0}$, we can test $\mathbb{H}_0'$ by further applying the Wald test principle to the estimator for $\boldsymbol{\beta}_*$. If $\ddot{\mathbf{R}}_1 := [\mathbf{I}_k, -\mathbf{I}_k, \mathbf{0}_{k\times 2}]$, $\ddot{\mathbf{R}}_1\bar{\boldsymbol{\upsilon}}_{T*} = \boldsymbol{\beta}_*^+ - \boldsymbol{\beta}_*^-$, we can estimate it $\bar{\boldsymbol{\upsilon}}_{T*}$ by $\widehat{\boldsymbol{\beta}}_T^+ - \widehat{\boldsymbol{\beta}}_T^-$, which is $\ddot{\mathbf{R}}_1\widetilde{\boldsymbol{\upsilon}}_T$. Next, we assume that

$$\mathscr{W}_T^{(1)} := \widetilde{\boldsymbol{\upsilon}}_T'\dot{\mathbf{D}}_1\ddot{\mathbf{R}}_1'(\ddot{\mathbf{W}}_T^{(1)})^{-1}\ddot{\mathbf{R}}_1\dot{\mathbf{D}}_1\widetilde{\boldsymbol{\upsilon}}_T, \quad \text{where} \quad \ddot{\mathbf{W}}_T^{(1)} := \widehat{\sigma}_{u,T}^2\ddot{\mathbf{R}}_1\dot{\mathbf{D}}_1\left(\sum_{t=1}^T \dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}'\right)^{-1}\dot{\mathbf{D}}_1\ddot{\mathbf{R}}_1'.$$

Note that $\mathscr{W}_T^{(1)}$ is the standard Wald test obtained from (9) because $\ddot{\mathbf{R}}_1\dot{\mathbf{D}}_1\widetilde{\boldsymbol{\upsilon}}_T = T(\widetilde{\boldsymbol{\beta}}_T^+ - \widetilde{\boldsymbol{\beta}}_T^-)$. Here, $(\widetilde{\boldsymbol{\beta}}_T^+ - \widetilde{\boldsymbol{\beta}}_T^-)$ is not exactly the same as $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-)$. The main difference between the two test bases is captured by $(\dot{\boldsymbol{\eta}}_T - \dot{\boldsymbol{\eta}}_T)$ and $\boldsymbol{\beta}_*(\dot{\rho}_T - \rho_*)$, and they do not contribute to power. The Wald test $\mathscr{W}_T^{(1)}$ is defined only by the test base with power, while the limit behaviors of the test are determined by the $O_{\mathbb{P}}(T^{-1})$ term under both hypotheses.

We next apply the Wald test principle to test $\mathscr{H}_0''$. The standard Wald test applied to OLS is as follows:

$$W_T^{(2)} := \widehat{\boldsymbol{\alpha}}_T'\dot{\mathbf{R}}_2'(\dot{\mathbf{W}}_T^{(2)})^{-1}\dot{\mathbf{R}}_2\widehat{\boldsymbol{\alpha}}_T, \quad \text{where} \quad \dot{\mathbf{W}}_T^{(2)} := \dot{\sigma}_{e,T}^2\dot{\mathbf{R}}_2\left(\sum_{t=1}^T \boldsymbol{z}_t\boldsymbol{z}_t'\right)^{-1}\dot{\mathbf{R}}_2'$$

and $\dot{\mathbf{R}}_2 := [\mathbf{0}_{kq\times(2+p+2k)}, \mathbf{I}_{kq}, -\mathbf{I}_{kq}]$. We also define another supplementary Wald test using the primitive estimator. From (13), it follows that $(\widehat{\boldsymbol{\varphi}}_T', \widehat{\boldsymbol{\pi}}_T^{+\prime}, \widehat{\boldsymbol{\pi}}_T^{-\prime})' = (\dot{\boldsymbol{\varphi}}_T', \dot{\boldsymbol{\pi}}_T^{+\prime}, \dot{\boldsymbol{\pi}}_T^{-\prime})' := \dot{\boldsymbol{\tau}}_{2T}$, implying that we can test $\mathscr{H}_0''$ by using $\dot{\boldsymbol{\tau}}_T$. Note that $\dot{\boldsymbol{\tau}}_T$ is not associated with an asymptotically singular matrix problem as Lemma 2 (*vi*) affirms. From this, we define the following Wald test:

$$\mathscr{W}_T^{(2)} := \dot{\boldsymbol{\tau}}_T'\dot{\mathbf{D}}\ddot{\mathbf{R}}_2'(\ddot{\mathbf{W}}_T^{(2)})^{-1}\ddot{\mathbf{R}}_2\dot{\mathbf{D}}\dot{\boldsymbol{\tau}}_T, \quad \text{where} \quad \ddot{\mathbf{W}}_T^{(2)} := \dot{\sigma}_{e,T}^2\ddot{\mathbf{R}}_2\dot{\mathbf{D}}\left(\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\right)^{-1}\dot{\mathbf{D}}\ddot{\mathbf{R}}_2'$$

and $\ddot{\mathbf{R}}_2 := \dot{\mathbf{R}}_2$. Note that $\mathscr{W}_T^{(2)}$ is defined by applying the Wald test principle to (11) because $\ddot{\mathbf{R}}_2\dot{\mathbf{D}}\dot{\boldsymbol{\tau}}_T = \sqrt{T}(\dot{\boldsymbol{\pi}}_T^+ - \dot{\boldsymbol{\pi}}_T^-)$. Although the test base $\sqrt{T}(\dot{\boldsymbol{\pi}}_T^+ - \dot{\boldsymbol{\pi}}_T^-)$ is identical to $\sqrt{T}(\widehat{\boldsymbol{\pi}}_T^+ - \widehat{\boldsymbol{\pi}}_T^-)$, the weighting matrix $\ddot{\mathbf{W}}_T^{(2)}$ is constructed from $\dot{\boldsymbol{\tau}}_T$, defining $\mathscr{W}_T^{(2)}$ directly. The supplementary Wald test also has a balanced convergence rate under both hypotheses.

Finally, we apply the Wald test principle to test $\mathscr{H}_0'''$. For this, we define the following test:

$$W_T^{(3)} := \widehat{\boldsymbol{\alpha}}_T'\dot{\mathbf{R}}_3'(\dot{\mathbf{W}}_T^{(3)})^{-1}\dot{\mathbf{R}}_3\widehat{\boldsymbol{\alpha}}_T, \quad \text{where} \quad \dot{\mathbf{W}}_T^{(3)} := \dot{\sigma}_{e,T}^2\dot{\mathbf{R}}_3\left(\sum_{t=1}^T \boldsymbol{z}_t\boldsymbol{z}_t'\right)^{-1}\dot{\mathbf{R}}_3'$$

and $\dot{\mathbf{R}}_3 := \dot{\mathbf{R}}_3 := \text{diag}[\dot{\mathbf{R}}_1, \dot{\mathbf{R}}_2]$. We also reformulate $\mathscr{H}_0'''$ against $\mathscr{H}_1'''$ into $\mathbb{H}_0''' : \boldsymbol{\beta}_*^+ = \boldsymbol{\beta}_*^-$ and $\boldsymbol{\pi}_*^+ = \boldsymbol{\pi}_*^-$ against $\mathbb{H}_1''' : \boldsymbol{\beta}_*^+ \neq \boldsymbol{\beta}_*^{-1}$ or $\boldsymbol{\pi}_*^+ \neq \boldsymbol{\pi}_*^-$, and define the following supplementary Wald test:

$$\mathscr{W}_T^{(3)} := \ddot{\boldsymbol{\delta}}_T' \bar{\mathbf{D}} \ddot{\mathbf{R}}_3' (\ddot{\mathbf{W}}_T^{(3)})^{-1} \ddot{\mathbf{R}}_3 \bar{\mathbf{D}} \ddot{\boldsymbol{\delta}}_T,$$

where $\ddot{\boldsymbol{\delta}}_T := (\widetilde{\boldsymbol{\upsilon}}_T', \dot{\boldsymbol{\tau}}_T')'$, $\bar{\mathbf{D}} := \text{diag}[\dot{\mathbf{D}}_1, \dot{\mathbf{D}}]$, $\ddot{\mathbf{R}}_3 := \text{diag}[\ddot{\mathbf{R}}_1, \ddot{\mathbf{R}}_2]$, and $\ddot{\mathbf{W}}_T^{(3)} := \text{diag}[\ddot{\mathbf{W}}_T^{(1)}, \ddot{\mathbf{W}}_T^{(2)}]$. Note that $\mathscr{W}_T^{(3)}$ is defined by applying the Wald test principle to both $\widetilde{\boldsymbol{\upsilon}}_T$ and $\dot{\boldsymbol{\tau}}_T$, and by noting that $\ddot{\mathbf{R}}_3 \bar{\mathbf{D}} \ddot{\boldsymbol{\delta}}_T = (T(\widetilde{\boldsymbol{\beta}}_T^+ - \widetilde{\boldsymbol{\beta}}_T^{-1})', \sqrt{T}(\dot{\boldsymbol{\pi}}_T^+ - \dot{\boldsymbol{\pi}}_T^-)')'$. We can easily confirm that $\mathscr{W}_T^{(3)} = \mathscr{W}_T^{(1)} + \mathscr{W}_T^{(2)}$ because $\ddot{\mathbf{W}}_T^{(3)}$ is block-diagonal. For notational simplicity, we further partition $\mathbb{H}_0'''$ into $\mathbb{H}_{01}''' : \boldsymbol{\beta}_*^+ = \boldsymbol{\beta}_*^-$ and $\mathbb{H}_{02}''' : \boldsymbol{\pi}_*^+ = \boldsymbol{\pi}_*^-$. In parallel, we also partition $\mathbb{H}_1'''$ into $\mathbb{H}_{11}''' : \boldsymbol{\beta}_*^+ = \boldsymbol{\beta}_*^-$ and $\mathbb{H}_{12}''' : \boldsymbol{\pi}_*^+ \neq \boldsymbol{\pi}_*^-$. These sub-hypotheses are separately defined because the test power depends on them.

We obtain the limit behaviors of the Wald tests from the limit behaviors of $\widetilde{\boldsymbol{\upsilon}}_T$ and $\dot{\boldsymbol{\tau}}_T$, which are already given in Lemmas 3 and 4. We summarize their limit behaviors in the following theorem, assuming that $\sigma_e^2 := \mathbb{E}[e_t^2]$:

Theorem 2. *Given Assumption 1, if $\boldsymbol{\theta}_*^+ \neq \mathbf{0}$ and/or $\boldsymbol{\theta}_*^- \neq \mathbf{0}$,*

(i) *(a)* $W_T^{(1)} \Rightarrow \dot{\mathscr{L}}' \mathbf{R}_1' (\sigma_e^2 \mathbf{R}_1 \dot{\mathscr{M}}^{-1} \mathbf{R}_1')^{-1} \mathbf{R}_1 \dot{\mathscr{L}}$ *under $\mathscr{H}_0'$, where* $\mathbf{R}_1 := [\mathbf{0}_{k \times 1}, \mathbf{I}_k, -\mathbf{I}_k, \mathbf{0}_{k \times (1+p+2kq)}]$;
(b) $W_T^{(2)} \Rightarrow \dot{\mathscr{L}}' \dot{\mathbf{R}}_2' (\sigma_e^2 \dot{\mathbf{R}}_2 \dot{\mathscr{M}}^{-1} \dot{\mathbf{R}}_2')^{-1} \dot{\mathbf{R}}_2 \dot{\mathscr{L}}$ *under $\mathscr{H}_0''$;*
(c) $W_T^{(3)} \Rightarrow \dot{\mathscr{L}}' \mathbf{R}_3' (\sigma_e^2 \mathbf{R}_3 \dot{\mathscr{M}}^{-1} \mathbf{R}_3')^{-1} \mathbf{R}_3 \dot{\mathscr{L}}$ *under $\mathscr{H}_0'''$, where* $\mathbf{R}_3 :=$ *diag*$[\mathbf{R}_1, \dot{\mathbf{R}}_2]$;

(ii) *(a) for any $c_T' = o(T)$,* $\lim_{T\to\infty} \mathbb{P}(W_T^{(1)} > c_T) = 0$ *under $\mathscr{H}_1'$;*
(b) for any $c_T' = o(T)$, $\lim_{T\to\infty} \mathbb{P}(W_T^{(2)} > c_T') = 0$ *under $\mathscr{H}_1''$;*
(c) for any $c_T' = o(T)$, $\lim_{T\to\infty} \mathbb{P}(W_T^{(3)} > c_T') = 0$ *under $\mathscr{H}_1'''$;*

(iii) *(a)* $\mathscr{W}_T^{(1)} \Rightarrow \dot{\mathscr{L}}_1' \ddot{\mathbf{R}}_1' (\sigma_u^2 \ddot{\mathbf{R}}_1 \dot{\mathscr{M}}_{11}^{-1} \ddot{\mathbf{R}}_1')^{-1} \ddot{\mathbf{R}}_1 \dot{\mathscr{L}}_1$ *under $\mathbb{H}_0'$;*
(b) $\mathscr{W}_T^{(2)} \Rightarrow \dot{\mathscr{L}}' \ddot{\mathbf{R}}_2' (\sigma_e^2 \ddot{\mathbf{R}}_2 \dot{\mathscr{M}}^{-1} \ddot{\mathbf{R}}_2')^{-1} \ddot{\mathbf{R}}_2 \dot{\mathscr{L}}$ *under $\mathscr{H}_0''$;*
(c) $\mathscr{W}_T^{(3)} \Rightarrow \dot{\mathscr{L}}_1' \ddot{\mathbf{R}}_1' (\sigma_u^2 \ddot{\mathbf{R}}_1 \dot{\mathscr{M}}_{11}^{-1} \ddot{\mathbf{R}}_1')^{-1} \ddot{\mathbf{R}}_1 \dot{\mathscr{L}}_1 + \dot{\mathscr{L}}' \ddot{\mathbf{R}}_2' (\sigma_e^2 \ddot{\mathbf{R}}_2 \dot{\mathscr{M}}^{-1} \ddot{\mathbf{R}}_2')^{-1} \ddot{\mathbf{R}}_2 \dot{\mathscr{L}}$ *under $\mathbb{H}_0'''$; and*

(iv) *(a) for any $c_T = o(T^2)$,* $\lim_{T\to\infty} \mathbb{P}(\mathscr{W}_T^{(1)} > c_T) = 0$ *under $\mathbb{H}_1'$;*
(b) for any $c_T' = o(T)$, $\lim_{T\to\infty} \mathbb{P}(\mathscr{W}_T^{(2)} > c_T') = 0$ *under $\mathscr{H}_1''$;*
(c) for any $c_T = o(T^2)$ and $c_T' = o(T)$, $\lim_{T\to\infty} \mathbb{P}(\mathscr{W}_T^{(3)} > c_T) = 0$ *under* $\mathbb{H}_{11}''' \bigcap \mathbb{H}_{02}'''$*;* $\lim_{T\to\infty} \mathbb{P}(\mathscr{W}_T^{(3)} > c_T') = 0$ *under* $\mathbb{H}_{01}''' \bigcap \mathbb{H}_{12}'''$*; and* $\lim_{T\to\infty} \mathbb{P}(\mathscr{W}_T^{(3)} > c_T) = 0$ *under* $\mathbb{H}_{11}''' \bigcap \mathbb{H}_{12}'''$. □

Remarks. (**a**) The asymptotic behaviors of $W_T^{(1)}$, $W_T^{(2)}$, and $W_T^{(3)}$ are determined by $(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*})$. For example, note that $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) = (\boldsymbol{\theta}_*^+ - \boldsymbol{\theta}_*^-) - (\dot{\rho}_T - \rho_*)(\boldsymbol{\beta}_*^+ - \boldsymbol{\beta}_*^-) + (\dot{\boldsymbol{\eta}}_T^+ - \boldsymbol{\eta}_{T*}^+) - (\dot{\boldsymbol{\eta}}_T^- - \boldsymbol{\eta}_{T*}^-) + o_{\mathbb{P}}(T^{-1})$. Thus, it follows that $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) = (\dot{\boldsymbol{\eta}}_T^+ - \boldsymbol{\eta}_{T*}^+) - (\dot{\boldsymbol{\eta}}_T^- - \boldsymbol{\eta}_{T*}^-) + o_{\mathbb{P}}(T^{-1})$ under $\mathscr{H}_0'$, namely, $\dot{\mathbf{R}}_1 \widehat{\boldsymbol{\alpha}}_T = \mathbf{R}_1(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) + o_{\mathbb{P}}(T^{-1})$. We can determine the null limit distribution of $W_T^{(1)}$ from that of $(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*})$. Theorem 2 (*i.a*) reports the null limit distribution obtained from this. However,

$(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) = (\boldsymbol{\theta}_*^+ - \boldsymbol{\theta}_*^-) - (\dot{\rho}_T - \rho_*)(\boldsymbol{\beta}_*^+ - \boldsymbol{\beta}_*^-) + o_{\mathbb{P}}(T^{-1/2})$ under $\mathscr{H}_1'$, where $(\dot{\rho}_T - \rho_*) = O_{\mathbb{P}}(T^{-1/2})$. Theorem 2 (*ii.a*) reports the consequence of this. Similar arguments apply to $W_T^{(2)}$ and $W_T^{(3)}$.

(b) From $\ddot{\mathbf{R}}_1 \dot{\mathscr{L}}_1 = \dot{\mathscr{L}}_{11} - \dot{\mathscr{L}}_{12}$, the null weak limit in Theorem 2 (*iii.a*) is given as $(\dot{\mathscr{L}}_{11} - \dot{\mathscr{L}}_{12})'(\sigma_u^2 \ddot{\mathbf{R}}_1 \dot{\mathscr{M}}_{11}^{-1} \ddot{\mathbf{R}}_1')^{-1}(\dot{\mathscr{L}}_{11} - \dot{\mathscr{L}}_{12})$. In parallel, if we assume that $[\dot{\mathscr{L}}_{61}', \dot{\mathscr{L}}_{62}', \dot{\mathscr{L}}_{63}']' := \dot{\mathscr{L}}_6$ such that $\dot{\mathscr{L}}_{61} \in \mathbb{R}^{p-1}$, $\dot{\mathscr{L}}_{62} \in \mathbb{R}^{kq}$, and $\dot{\mathscr{L}}_{63} \in \mathbb{R}^{kq}$, then $\ddot{\mathbf{R}}_2 \dot{\mathscr{L}} = \dot{\mathscr{L}}_{62} - \dot{\mathscr{L}}_{63}$, and so the weak the limit in Theorem 2 (*iii.b*) can be rewritten as $(\dot{\mathscr{L}}_{62} - \dot{\mathscr{L}}_{63})'(\sigma_e^2 \ddot{\mathbf{R}}_2 \dot{\mathscr{M}}^{-1} \ddot{\mathbf{R}}_2')^{-1}(\dot{\mathscr{L}}_{62} - \dot{\mathscr{L}}_{63})$.

(c) The null limit distributions of $\mathscr{W}_T^{(1)}$, $\mathscr{W}_T^{(2)}$, and $\mathscr{W}_T^{(3)}$ critically depend on the null hypotheses. We can rewrite the null weak of $\mathscr{W}_T^{(1)}$ as $\dot{\mathscr{S}}_1' \dot{\mathscr{M}}_{11}^{-1} \ddot{\mathbf{R}}_1' (\sigma_u^2 \ddot{\mathbf{R}}_1 \dot{\mathscr{M}}_{11}^{-1} \ddot{\mathbf{R}}_1')^{-1} \ddot{\mathbf{R}}_1 \dot{\mathscr{M}}_{11}^{-1} \dot{\mathscr{S}}_1$ using the definition of $\dot{\mathscr{L}}_1$. Similarly, the null weak limit of $\mathscr{W}_T^{(2)}$ becomes $\dot{\mathscr{S}}' \dot{\mathscr{M}}^{-1} \ddot{\mathbf{R}}_1' (\sigma_e^2 \ddot{\mathbf{R}}_1 \dot{\mathscr{M}}^{-1} \ddot{\mathbf{R}}_1')^{-1} \ddot{\mathbf{R}}_1 \dot{\mathscr{M}}^{-1} \dot{\mathscr{S}}$. Note that the null weak limit of $\mathscr{W}_T^{(2)}$ is a mixed chi-squared distribution from the definition of $\dot{\mathscr{S}}$. However, $\mathscr{W}_T^{(1)}$ and $\mathscr{W}_T^{(3)}$ do not follow a mixed chi-squared distribution under the null because $\dot{\mathscr{S}}_1$ does not follow a mixed normal distribution. Instead, we can use the bootstrap method to obtain their critical values. We demonstrate this in Section 6. □

6 Monte Carlo Simulations

In this section, we conduct simulations to examine the finite sample properties of the Wald tests.

For our simulation, we assume the following data-generating process (DGP) condition:

$$y_{t-1} = \nu_* + \beta_*^+ x_{t-1}^+ + \beta_*^- x_{t-1}^- + \zeta_*(t-1) + u_{t-1} \quad \text{and}$$

$$\Delta y_t = \alpha_* + \rho_* u_{t-1} + \varphi_* \Delta y_{t-1} + \pi_*^+ \Delta x_t^+ + \pi_*^- \Delta x_t^- + e_t,$$

where $\Delta x_t = v_t$, and $(e_t, v_t)' \sim \text{IID}N(\mathbf{0}_2, \mathbf{I}_2)$. We also set $(\nu_*, \zeta_*, \alpha_*, \rho_*, \varphi_*) = (0, 0, 0, -1/2, 0)$ throughout the simulation, but adjust the value of $(\beta_*^+, \beta_*^-, \pi_*^+, \pi_*^-)$, depending on the hypotheses of interest. According to the NARDL condition, it must hold that $\theta_*^+ = -\rho_* \beta_*^+$ and $\theta_*^- = -\rho_* \beta_*^-$.

We follow the next procedure to define the Wald tests. First, we estimate the unknown parameters using primitive parameter estimators. For this, we estimate $\bar{\boldsymbol{\upsilon}}_{T*}$ and $\boldsymbol{\tau}_{T*}$ separately by specifying the following models:

$$y_t = \beta^+ \widehat{m}_t^+ + \beta^- \widehat{m}_t^- + \vartheta t + \nu + u_t \quad \text{and}$$

$$\Delta y_t = \rho \widehat{u}_{t-1} + \eta^+ \widehat{m}_{t-1}^+ + \eta^- \widehat{m}_{t-1}^- + \varsigma(t-1) + \gamma + \varphi \Delta y_{t-1} + \pi^+ \Delta x_t^+ + \pi^- \Delta x_t^- + e_t,$$

where we set $\widehat{m}_t^+ := x_t^+ - t\widehat{\mu}_T^+$, $\widehat{m}_t^- := x_t^- - t\widehat{\mu}_T^-$, and $\widehat{u}_t := y_t - \widetilde{r}_t' \widehat{\boldsymbol{\upsilon}}_T$ with $\widehat{\mu}_T^+ := (\sum_{t=1}^{T-1} t^2)^{-1} \sum_{t=1}^{T-1} t x_t^+$, $\widehat{\mu}_T^- := (\sum_{t=1}^{T-1} t^2)^{-1} \sum_{t=1}^{T-1} t x_t^-$, $\widehat{\boldsymbol{\upsilon}}_T := (\sum_{t=1}^{T} \widetilde{\boldsymbol{r}}_{t-1} \widetilde{\boldsymbol{r}}_{t-1}')^{-1} \sum_{t=1}^{T} \widetilde{\boldsymbol{r}}_{t-1} y_{t-1}$, and $\widetilde{\boldsymbol{r}}_t := [x_t^+, x_t^-, t, 1]'$.

Second, we specify the following hypothesis systems:

$$\mathscr{H}_0' : \theta_*^+ = \theta_*^- \text{ vs. } \mathscr{H}_1' : \theta_*^+ \neq \theta_*^-; \quad \mathscr{H}_0'' : \pi_*^+ = \pi_*^- \text{ vs. } \mathscr{H}_1'' : \pi_*^+ \neq \pi_*^-;$$

$$\mathscr{H}_0''' : \theta_*^+ = \theta_*^- \text{ and } \pi_*^+ = \pi_*^- \text{ vs. } \mathscr{H}_1''' : \theta_*^+ \neq \theta_*^- \text{ or } \pi_*^+ \neq \pi_*^-;$$

$$\mathbb{H}_0' : \beta_* = 0 \text{ vs. } \mathbb{H}_1' : \beta_* \neq 0; \quad \text{and} \quad \mathbb{H}_0''' : \beta_* = 0 \text{ and } \pi_*^+ = \pi_*^- \text{ vs. } \mathbb{H}_1''' : \beta_* \neq 0 \text{ or } \pi_*^+ \neq \pi_*^-.$$

Finally, we compute the Wald tests $\mathscr{W}_T^{(1)}$, $\mathscr{W}_T^{(2)}$, $\mathscr{W}_T^{(3)}$, $W_T^{(1)}$, $W_T^{(2)}$, and $W_T^{(3)}$ as stated in Section 5.

We conduct simulations under the following two DGP conditions. First, we set $\beta_*^+ = \beta_*^- = 1$ and $\pi_*^+ = \pi_*^- = 1/2$ to generate data. Note that this parameter condition satisfies the ARDL condition. From this, we examine the finite sample properties of Wald tests under the null. Second, we set $\beta_*^+ = 1/4$, $\beta_*^- = -1/4$, $\pi_*^+ = 1/8$, and $\pi_*^- = -1/8$, and use this to examine the power.

A bootstrap method can be used for the testing. Note that the null limit distributions of $\mathscr{W}_T^{(1)}$ and $\mathscr{W}_T^{(3)}$ are not standard, as Theorem 2 reveals. We apply the following residual bootstrap method.

S1: After computing the Wald tests, we estimate the ARDL model by regressing Δy_t against y_{t-1}, x_{t-1}, $(t-1)$, 1, Δy_{t-1} and Δx_t. We let the estimated linear coefficient be $(\widehat{\rho}_T, \widehat{\theta}_T, \widehat{\xi}_T, \widehat{\alpha}_T, \widehat{\varphi}_T, \widehat{\pi}_T)$. We also let the residual be $\widehat{e}_t := \Delta y_t - \widehat{\rho}_T y_{t-1} - \widehat{\theta}_T x_{t-1} - \widehat{\xi}_T(t-1) - \widehat{\alpha}_T - \widehat{\varphi}_T \Delta y_{t-1} - \widehat{\pi}_T \Delta x_t$.

S2: We construct resampled series as follows. First, we resample $\widehat{e}_t$ with replacement and denote it as e_t^b. Next, we let

$$\Delta y_t^b := \widehat{\rho}_T y_{t-1}^b + \widehat{\theta}_T x_{t-1} + \widehat{\xi}_T(t-1) + \widehat{\alpha}_T + \widehat{\varphi}_T \Delta y_{t-1}^b + \widehat{\pi}_T \Delta x_t + e_t^b,$$

where y_t^b is the cumulative sum of Δy_t^b. Note that we do not resample Δx_t. Using the resampled series, we compute the Wald tests. For example, we let $\mathscr{W}_T^{(1),b}$ denote the bootstrapped $\mathscr{W}_T^{(1)}$.

S3: We iterate the second step B times in total and compute the empirical p-value of the test. For example, we let the empirical p-value be $p_T^{(1)} := B^{-1} \sum_{b=1}^{B} \mathbb{1}\{\mathscr{W}_T^{(1),b} > \mathscr{W}_T^{(1)}\}$ for $\mathscr{W}_T^{(1)}$. If $p_T^{(1)}$ is less than the significance level, we reject the null hypothesis. □

The bootstrap procedure is applicable even when the Wald test null limit distribution is mixed chi-squared.

Our simulation follows the above procedure. Assuming that $B = 500$, we independently iterate the above experiment 5,000 times for $T = 100, 200, 300, 400$, and 500. The simulation results are reported in Tables 1 and 2. We also evaluate the Wald tests by mixed chi-squared distributions and report the empirical rejection rates. The simulation results are summarized as follows:

(a) Table 1 shows the empirical Wald test rejection rates using the null DGP condition allowing $\beta_*^+ = \beta_*^- = 1$ and $\pi_*^+ = \pi_*^- = 1/2$. The left-hand side panel shows the empirical rejection rates obtained using the residual bootstrap method. The rejection rates on the right-hand side are based on the mixed chi-squared distribution. In sum, first, when using the bootstrap method, for each T, the empirical rejection rates are very close to the nominal significance levels. This implies that this method effectively controls the test levels. Second, when considering the asymptotic critical values from mixed chi-squared distribution, we find no significant level distortion for $\mathscr{W}_T^{(2)}$, although $\mathscr{W}_T^{(1)}$ and $\mathscr{W}_T^{(3)}$ experience level distortions. For the current DGP, $\mathscr{W}_T^{(1)}$ and $\mathscr{W}_T^{(3)}$ reject $\mathbb{H}_0'$ and $\mathbb{H}_0'''$ more often than the significance levels. This confirms that the null behaviors of $\mathscr{W}_T^{(1)}$ and $\mathscr{W}_T^{(3)}$ differ from mixed chi-squared distribution. However, $\mathscr{W}_T^{(2)}$ does not suffer from level distortions, indicating that $\dot{\pi}_T^+$ and $\dot{\pi}_T^-$ follow mixed normal distributions, as Lemma 4 (*i*) establishes. Finally, $W_T^{(1)}$, $W_T^{(2)}$, and $W_T^{(3)}$ suffer no level distortions when using the mixed chi-squared distribution method.

(b) Table 2 shows the empirical Wald test rejection rates under the alternative DGP. The summary results are as follows. First, with the bootstrap method, the empirical rejection rates of $\mathscr{W}_T^{(1)}$, $\mathscr{W}_T^{(2)}$, and $\mathscr{W}_T^{(3)}$ tend toward 100% as T increases, implying their consistency against $\mathbb{H}_1'$, $\mathscr{H}_1''$, and $\mathbb{H}_1'''$, respectively. Second, when applying the critical values obtained from the mixed chi-squared distribution, the empirical rejection rates of $\mathscr{W}_T^{(1)}$, $\mathscr{W}_T^{(2)}$, and $\mathscr{W}_T^{(3)}$ also converge toward 100% as T increases. However, it is difficult to control their sizes, as Table 1 describes. Third, the standard Wald tests are also consistently powerful. Fourth, as for the bootstrap method, $\mathscr{W}_T^{(1)}$ and $\mathscr{W}_T^{(3)}$ are more powerful than $W_T^{(1)}$ and $W_T^{(3)}$, respectively, for small T, but this relationship is reversed as T increases. In contrast, both $\mathscr{W}_T^{(2)}$ and $W_T^{(2)}$ show similar power patterns. □

We conduct another simulation assuming that Δx_t is serially correlated. Instead of $\Delta x_t = v_t$, we set $\Delta x_t = \kappa_* \Delta x_{t-1} + v_t$, and apply the residual bootstrap method as earlier. Then, assuming $\kappa_* = 1/2$, we conduct simulation by setting $B = 500$ and independently iterating the experiment 5,000 times for $T = 100$, 200, 300, 400, and 500. The simulation results are presented in Tables 3 and 4 in the same format used in Tables 1 and 2, respectively. The simulation results are summarized as follows:

(a) Table 3 presents the empirical Wald test rejection rates obtained using the null DGP condition. The left-hand side panel gives the empirical rejection rates obtained using the residual bootstrap method, while the right-hand side panel shows the values obtained using the mixed chi-squared distribution method. For each value of T, the empirical rejection rates are very close to the nominal significance levels obtained with the residual bootstrap method. In contrast, the asymptotic critical values obtained using mixed chi-squared distribution introduce level distortions for $\mathscr{W}_T^{(1)}$ and $\mathscr{W}_T^{(3)}$, as observed earlier. These simulation results are overall similar to those in Table 1.

(b) Table 4 presents the empirical Wald test rejection rates obtained using the alternative DGP condition. The results are as follows. First, when the residual bootstrap method is used, the empirical rejection rates of $\mathscr{W}_T^{(1)}$, $\mathscr{W}_T^{(2)}$, and $\mathscr{W}_T^{(3)}$ tend toward 100% as T increases, implying their consistency against $\mathbb{H}_1'$, $\mathscr{H}_1''$, and $\mathbb{H}_1'''$, respectively. Second, when the asymptotic critical values obtained from the mixed chi-squared distribution are used, the empirical rejection rates also converge toward 100% with the increase in T, although the level distortions are difficult to control for $\mathscr{W}_T^{(1)}$ and $\mathscr{W}_T^{(3)}$ under the null. Third, standard Wald tests also exhibit consistent power. As T increases, $W_T^{(1)}$, $W_T^{(2)}$, and $W_T^{(3)}$ reject the null with rejection rates tending toward 100% with the increase in T. Finally, $\mathscr{W}_T^{(1)}$ and $\mathscr{W}_T^{(3)}$ are more powerful than $W_T^{(1)}$ and $W_T^{(3)}$, respectively, for small T. Meanwhile, both $\mathscr{W}_T^{(2)}$ and $W_T^{(2)}$ exhibit similar power patterns. $\square$

7 Empirical Application

This section examines the empirical data provided by Romer and Romer (2010) when measuring exogenous fiscal shocks with respect to the U.S. GDP. We review the literature and apply the NARDL model to examine the long- and short-run symmetries in data.

7.1 Literature Review and Empirical Motivation

Estimating the fiscal policy impact on output is challenging because many fiscal factors leading to tax changes are correlated with the output, causing the OLS estimator to suffer from the endogenous bias problem. Although not all fiscal factors are endogenous in terms of output growth, using all tax changes to regress GDP growth can lead to biased OLS estimates. Blanchard and Perotti (2002) address this problem by using structural vector autoregression (SVAR). They assume that policymakers do not respond to shocks contemporaneously, but use information on the elasticity of revenue to create cyclically adjusted revenues. This implies that the effect of a tax cut on GDP in the U.S. is around 1%. However, Romer and Romer (2010) and Cloyne (2013) argue that the structural assumptions used in the SVAR model may be unrealistic in estimating the impact of fiscal policy on output.

Romer and Romer (2010) try to disentangle the problem of effects of tax changes correlated with GDP differently. Performing a narrative analysis, they identify the motivations behind each tax change from 1945 to 2007. Using sources such as *the Economic Report of the President* and *the Congressional Record*, they classify legislated tax changes, which changed tax liabilities from one quarter to the next, into four categories: (i) tax changes to counteract changes in government spending, (ii) tax changes to offset other factors affecting near-term output, (iii) tax changes to address inherited budget deficits, and (iv) tax changes to promote long-term growth. This classification is based on the motivations behind the tax changes, with the first two categories considered countercyclical and motivated by restoring the output growth reduced by other factors. This feature makes it difficult to categorize the first two tax changes as purely exogenous. In contrast, the last two categories are based on policymakers' perceptions of prudent fiscal policies or focus on increasing long-term growth, and hence may be classified as exogenous. In consequence, they identified 54 exogenous tax changes through narrative analysis during the same period. Using exogenous fiscal shocks, they developed a time-series model and estimated that the GDP will increase by approximately 3% over three years following a tax cut of 1% of GDP. This estimate differs significantly from the estimate of Blanchard and Perotti (2002).

Narrative analysis is a popular methodology for uncovering the impact of fiscal shocks on GDP. For example, Cloyne (2013) applies narrative analysis to U.K. legislation and estimates the impact of exogenous fiscal shocks on GDP. The study finds that a 1% tax cut, as a percentage of GDP, raises output by nearly 2.5% over the next three years, as in the U.S. Mertens and Olea (2018) use narrative analysis to estimate the short-run tax elasticity of income as about 1.2% in the U.S. by measuring the exogenous variations in marginal tax rate. Gunter, Riera-Crichton, Vegh, and Vuletin (2021) extend narrative analysis to estimate the value-added-tax multipliers for 51 countries, to find that the effect of tax changes on output is highly nonlinear.

Narrative analysis is also used to specialize the time-series model in Romer and Romer (2010) for specific models under different economic environments. For instance, Mertens and Ravn (2012) distinguish between surprise and anticipated tax changes to examine the dynamic effect of tax change on GDP, and report that anticipated tax cuts lead to contraction in GDP. Demirel (2021) and Ghassibe and Zanetti (2021) examine the state-dependent impact of exogenous tax changes on GDP, allowing for different tax multiplier estimations in recessions and expansions. Narrative analysis on the effect of tax change on GDP is widely applied to other fields, allowing for comparison with the outcomes obtained by conventional analyses.

The standpoint of the current studies motivates the extension of the estimation beyond short-run relationships to focus on estimating the long-run relationship between GDP and fiscal shocks. Studies have primarily focused on the short-term effects of fiscal policy. For example, one of the models specified by Romer and Romer (2010) is given as

$$\Delta y_t = \gamma_* + \sum_{j=0}^{q-1} \pi_{j*} \Delta\tau_{t-i} + \sum_{j=1}^{p-1} \varphi_{j*} \Delta y_{t-j} + e_t \tag{21}$$

to examine how GDP responds to exogenous tax changes, where y_t is the logarithm of real GDP and $\Delta\tau_t$ is the exogenous log tax change.

As both y_t and τ_t are observable, we can estimate their long-run relationship by applying cointegration analysis, but to our knowledge, no prior work has used this approach for narrative data. We, therefore, augment the cointegration error on the right-hand side of (21) as follows:

$$\Delta y_t = \gamma_* + \rho_* u_{t-1} + \sum_{j=0}^{q-1} \pi_{j*}\Delta\tau_{t-i} + \sum_{j=1}^{p-1} \varphi_{j*}\Delta y_{t-j} + e_t, \tag{22}$$

where $u_t := y_t - \beta_*\tau_t - \zeta_* t - \nu_*$. Note that the long-run relationship between y_t and τ_t can be found by estimating β_*, and the short-run relationship can be further revealed by estimating the coefficients of π_{j*} and φ_{j*}. We can also convert (22) into

$$\Delta y_t = \alpha_* + \rho_* y_{t-1} + \theta_*\tau_{t-1} + \xi_* t + \sum_{j=0}^{q-1} \pi_{j*}\Delta\tau_{t-i} + \sum_{j=1}^{p-1} \varphi_{j*}\Delta y_{t-j} + e_t \tag{23}$$

and estimate the unknown parameters in (23) by OLS.

For this estimation, we first examine the partial sum processes of exogenous tax changes used in our empirical analysis. Figure 1 illustrates the partial sum processes due to exogenous tax changes.[1] The solid and dashed lines represent the partial sum processes of tax changes for deficit reduction (τ_{1t}) and long-run growth (τ_{2t}), respectively, while the dotted line represents the partial sum process of their sum (τ_t).[2]

Exogenous tax changes exhibit characteristics suitable for NARDL analysis. We explain these characteristics one by one. First, as tax changes for budget deficit always lead to tax increases, $\Delta\tau_{1t}$ is always positive. Second, most tax changes for long-run economic growth are related to tax decreases. Out of 31 legislated tax changes for long-run growth, only 6 result in tax increases. This means that the partial sum processes for deficit reduction and long-run growth remain in positive and negative regions, respectively. Although the NARDL model assumptions do not perfectly align with the characteristics of exogenous fiscal shocks, we use the approximation of $\Delta\tau_t^+ := \max[0, \Delta\tau_t]$ for tax changes due to budget deficit and $\Delta\tau_t^- := \min[0, \Delta\tau_t]$ for tax changes due to long-run growth.

For this purpose, we specify the following NARDL model and estimate the long- and short-run parameters using the methodology of the current study:

$$\Delta y_t = \gamma_* + \rho_* u_{t-1} + \sum_{j=0}^{q-1} \left(\pi_{j*}^+\Delta\tau_{t-i}^+ + \pi_{j*}^-\Delta\tau_{t-i}^-\right) + \sum_{j=1}^{p-1} \varphi_{j*}\Delta y_{t-j} + e_t, \tag{24}$$

[1] Data can be obtained from the following URL: https://eml.berkeley.edu/~cromer/ (Accessed: Feb. 10, 2023).

[2] We obtain the partial sum processes by first converting the nominal tax changes into consistent values over the period 1947Q1 to 2007Q4. For this, we first discount the nominal values with price index implied by the nominal GDP and the quantity index for GDP in the data set, and then apply a log transformation. We find that $\Delta\tau_t = \Delta\tau_{1t} + \Delta\tau_{2t}$, $\Delta\tau_{1t} := \mathrm{sgn}(\Delta T_{1t})\log(|\Delta T_{1t}|/p_t)$, $\Delta\tau_{2t} := \mathrm{sgn}(\Delta T_{2t})\log(|\Delta T_{2t}|/p_t)$, and $p_t := NY_t/Y_t$, where ΔT_{1t} and ΔT_{2t} represent the nominal tax changes for budget deficit and long-run growth, respectively, and NY_t and Y_t represent the nominal GDP and quantity GDP index, respectively. In case $\Delta T_{1t} = 0$ or $\Delta T_{2t} = 0$, we let $\Delta\tau_{1t} = 0$ or $\Delta\tau_{2t} = 0$, respectively. The partial sum processes in Figure 1 represent τ_t, τ_{1t}, and τ_{2t}.

where $u_t := y_t - \beta_*^+ \tau_t^+ - \beta_*^- \tau_t^- - \zeta_* t - \nu_*$. This can be rewritten as

$$\Delta y_t = \alpha_* + \rho_* y_{t-1} + \theta_*^+ \tau_{t-1}^+ + \theta_*^- \tau_{t-1}^- + \xi_* t + \sum_{j=0}^{q-1} \left(\pi_{j*}^+ \Delta \tau_{t-i}^+ + \pi_{j*}^- \Delta \tau_{t-i}^- \right) + \sum_{j=1}^{p-1} \varphi_{j*} \Delta y_{t-j} + e_t, \quad (25)$$

which we estimate by OLS. If $\beta_*^+ = \beta_*^-$ (or $\theta_*^+ = \theta_*^-$) and $\pi_{j*}^+ = \pi_{j*}^-$, (24) reduces to (22), which economically implies that the relationship between tax changes aimed at deficit reduction and long-run growth in real GDP is roughly symmetrical in both the long and short run. We can use the Wald tests defined in Section 5 for this inference.

7.2 Empirical Results

This section presents the estimation and inference results. Our discussion is divided into two parts. The first part presents the estimation results obtained using the tax change data outlined in the previous section. The second part uses the tax change data utilized by Romer and Romer (2010). We have reduced the sample period to 1947Q1–2007Q4 by excluding missing observations.

7.2.1 Tax Changes Measured by Log Transformation of Tax

Before presenting the estimation and inference results, we provide the basic statistical characteristics of the data. The logarithm of GDP quantity index multiplied by 100 is represented by y_t, while τ_t, $\tau_{1,t}$, and $\tau_{2,t}$ are defined as in footnote 2. The descriptive statistics of Δy_t, $\Delta \tau_{1t}$, $\Delta \tau_{2t}$, and $\Delta \tau_t$ can be found in Supplement Table A.1. Furthermore, our unit-root test on y_t, τ_{1t}, τ_{2t}, and τ_t follows the method of Phillips and Perron (1988), including or excluding the time trend. The test results indicate that we cannot reject the unit-root hypothesis for the series.

We report the estimation results in Table 5. Columns marked “Exo” give the parameter estimates obtained by OLS for NARDL and ARDL models. In other words, the unknown parameters in equations (25) and (23) are estimated by OLS. Orders for the NARDL model are based on the Akaike information criterion (AIC), with $p = 3$ and $q = 1$ for both the NARDL and ARDL models. Standard errors are listed in parentheses below the parameter estimates. Except for the coefficient of y_{t-1}, we use the asymptotic critical values from mixed normal distribution for t-tests. For t-test on the coefficient of y_{t-1}, we use the asymptotic critical values provided by Banerjee et al. (1998). Furthermore, we test whether all coefficients of y_{t-1}, τ_{t-1}^+, and τ_{t-1}^- are equal to zero by applying the F-test from Pesaran et al. (2001). Finally, we test the hypotheses of symmetry between long-run parameters, short-run parameters, or both using the Wald tests defined in Section 5. The results are presented in the two bottom panels. We summarize the results in

Table 5 as follows:

(a) The coefficient of y_{t-1} in the t-test is significant at the 10% and 5% levels for the NARDL and ARDL models, respectively. Although the hypothesis of no cointegration cannot be rejected by the F-test for the NARDL model, it is significant at the 10% level for the ARDL model. Our analysis using the 2SNARDL model detailed below suggests that we can reasonably conclude that a cointegrating relationship exists between the log real GDP and exogenous log tax shock.

(b) The NARDL model estimation results indicate that an increase in exogenous tax shock measured by τ^{+}_{t-1} reduces the log real GDP. In contrast, a decrease in exogenous tax shock measured by τ^{-}_{t-1} increases the log of real GDP. The ARDL model also reveals the same relationship between exogenous log tax shock and log real GDP.

(c) The estimated coefficients of τ^{+}_{t-1} and τ^{-}_{t-1} are almost equal in magnitude, suggesting no long-run asymmetry between log real GDP and exogenous log tax shock. This conclusion is further confirmed by the Wald test. Both $\mathscr{W}^{(1)}_{T}$ and $W^{(1)}_{T}$ provide p-values that make it difficult to reject the symmetry hypothesis.

(d) Short-run symmetry is confirmed by $\mathscr{W}^{(2)}_{T}$ and $W^{(2)}_{T}$ as they do not reject the symmetry hypothesis. Moreover, both the long-run and short-run symmetry hypotheses are not challenged by either $\mathscr{W}^{(3)}_{T}$ or $W^{(3)}_{T}$. From this, we can conclude that the ARDL model is appropriate for studying the relationship between log real GDP and exogenous log tax shock.

(e) We present the estimation results for endogenous log tax shock, calculated in the same way as for exogenous log tax shock. Columns labeled "Endo" give the estimation and inference results obtained using the endogenous log tax shock data in Romer and Romer (2010). Similarly, the columns labeled "Sum" display the estimation and inference results obtained using both exogenous and endogenous log tax shocks. The estimated signs of τ^{+}_{t-1}, τ^{-}_{t-1}, and τ_{t-1} are inconsistent with economic theory for endogenous log tax shock. As for using both exogenous and endogenous log tax shocks jointly, there is little evidence of cointegration. These estimation results indicate that only exogenous log tax shocks reveal the impact on fiscal shock GDP, with the coefficients having signs consistent with economic theory. □

Next, we estimate the NARDL and ARDL models using the 2SNARDL method proposed by Cho, Greenwood-Nimmo, and Shin (2023a). For the ARDL model, we apply 2SNARDL estimation by imposing the short- and long-run parameter symmetry conditions. The results are presented in Table 6, structured in parallel to Table 5. This is to validate the inference results in Table 5. We summarize the results in Table 6.

(a) The NARDL and ARDL models estimated using 2SNARDL show that the coefficient of u_{t-1} for exogenous tax shock is significant. The significance levels are 10% and 5% for the NARDL and

ARDL models, respectively. Moreover, we use a unit root test as in Phillips and Perron (1988) to test the cointegration residuals obtained from both models. The results reject the unit root hypothesis. The p-values are 10.19% and 2.58% for the NARDL and ARDL models, respectively. This confirms the cointegrating relationship between log real GDP and exogenous log tax shock.

(b) As regards exogenous log tax shock, the NARDL model indicates that the long-run log real GDP coefficient for log tax increase is -0.4329%, while that for log tax decrease is -0.3202%. These signs are consistent with economic theory, and both are significant at the 1% and 10% levels, respectively. For the ARDL model, the long-run coefficient for log tax increase is -0.2328%. This sign is also consistent with economic theory, and the estimated coefficient is significant at the 1% level. Given that 2SNARDL is super-consistent, these estimates are likely to be more precise than those OLS.

(c) Our findings for endogenous and aggregate log tax shocks align with the previous results in Table 5. Specifically, the coefficients of τ_{t-1}^{+}, τ_{t-1}^{-}, and τ_{t-1} for endogenous log tax shock are statistically significant, but their signs are inconsistent with economic theory. However, for aggregate log tax shock, none of these coefficients are significant. Moreover, the coefficient of u_{t-1} is insignificant owing to the asymptotic critical values provided by Banerjee et al. (1998). These inconsistent results are believed due to the correlation between the endogenous shock and structural error as Romer and Romer (2010) point out. □

The results in Tables 5 and 6 suggest that by using exogenous log tax shocks for model estimation, we can properly identify the relationship between log real GDP and fiscal shock. The findings indicate limited statistical support for asymmetry between tax shock for deficit reduction and long-run growth. Moreover, OLS and 2SNARDL estimations produce qualitatively similar results.

7.2.2 Tax Changes Measured by Tax to GDP Ratio

This section extends Romer and Romer (2010) to investigate the long- and short-run relationships between fiscal shock and real GDP. Rather than the tax change logarithm (τ_t), we employ $\Delta r_t := (\Delta T_{1t} + \Delta T_{2t})/NY_t$, which represents the tax change to nominal GDP ratio, to specify the models corresponding to (22), (23), (24), and (25). As with Tables 5 and 6, we estimate the models using OLS and 2SNARDL. The estimation and inference results are presented in Tables 7 and 8, respectively. We summarize them as follows:

(a) For exogenous tax change, we find the coefficient of y_{t-1} in the t-test of Table 7 significant at the 25% level for the NARDL and ARDL models. The F-test does not reject the hypothesis of no cointegration. However, the coefficient of u_{t-1} in the t-test of Table 8 is significant at the 25% and 10% levels for the NARDL and ARDL models, respectively. The inference results in Table 8 have more weight since the

2SNARDL estimation has a faster convergence rate than OLS estimation. Neither of the Wald tests $\mathscr{W}_T^{(1)}$ and $W_T^{(1)}$ rejects the symmetry hypothesis in long-run parameters, indicating that the ARDL model is more relevant to the data, and that a cointegrating relationship exists between y_t and r_t.

(b) The ARDL model results indicate that an increase in exogenous tax shock measured by r_{t-1} reduces the long-run log real GDP by about 3%. This is remarkably close to the overall estimation results in Romer and Romer (2010), estimating that the GDP would increase by approximately 3% over three years following a tax cut of 1% of GDP.

(c) The Wald tests $\mathscr{W}_T^{(2)}$ and $W_T^{(2)}$ do not reject the hypothesis of symmetric short-run parameters. Furthermore, neither of the Wald tests $\mathscr{W}_T^{(3)}$ and $W_T^{(3)}$ rejects the hypothesis of long- and short-run symmetry. This confirms that the ARDL model is appropriate for the long- and short-run relationships between y_t and r_t.

(d) As for endogenous and aggregate tax changes, there is negligible evidence of cointegration between the real GDP and tax changes. Neither of the t- and F-tests in Table 7 rejects the hypothesis of no cointegration. Furthermore, none of the coefficient of u_{t-1} in Table 8 is statistically significant. Specifically for aggregate tax change, we cannot confirm that r_t is nonstationary from the unit root test in the Online Supplement. From this, we can estimate the long- and short-run relationships properly only by using exogenous tax change. □

In summary, the empirical results obtained with the specification in Romer and Romer (2010) provide qualitatively the same results as in Section 7.2.1. In particular, the long-run relationship between y_t and r_t captured by the cointegration coefficient is close to their estimate.

8 Conclusion

OLS estimation suffers from an asymptotically singular matrix problem when performed with the NARDL model. Although the limit behavior of OLS is unknown, it is popularly used in the empirical literature.

This study investigates the large sample behavior of the OLS estimator by addressing the asymptotically singular matrix problem. Specifically, we find the OLS estimator consistent for unknown NARDL parameters, following a mixed normal distribution asymptotically under some mild regularity conditions. This implies that the standard t- and Wald test principles can be applied to the OLS estimator as if no asymptotically singular matrix problem existed. For this, we first derive the large sample distribution of the OLS estimator by representing it as a bilinear transform of other primitive estimators that do not suffer from the asymptotically singular matrix problem. This representation would make the application of higher-order expansion to OLS unnecessary for limit distribution.

Furthermore, we examine the large sample behavior of the Wald tests for the NARDL hypothesis. Besides the standard Wald tests defined by OLS, we develop other supplementary Wald tests using primitive estimators to test for asymmetric long- and/or short-run parameters. The null limit distributions of the standard Wald tests are mixed chi-squared, but the supplementary Wald tests have different ones when testing for asymmetry in long-run parameters. Applying the residual bootstrap method here, we find through Monte Carlo simulation that the supplementary Wald tests perform better than the standard Wald tests overall.

Finally, we illustrate the proper use of the NARDL model by estimating the long- and short-run relationships between GDP and exogenous fiscal shocks due to deficit reduction and long-run growth. For this, we use the empirical data provided by Romer and Romer (2010). As all tax changes for deficit reduction represent tax increases and most tax changes for long-run growth are tax decreases, we estimate the NARDL model and examine whether the long- and short-run relationships between tax increase and decrease are symmetric. Consequently, we find no evidence of asymmetric long- and short-run relationships between tax increase and decrease. We also find that a 1% exogenous GDP tax increase reduces the log real GDP by about 3% in the long run. This is consistent overall with the estimation result in Romer and Romer (2010).

References

BAEK, Y. I., J. S. CHO, AND P. C. B. PHILLIPS (2015): "Testing Linearity Using Power Transforms of Regressors," *Journal of Econometrics*, 187, 376–384.

BANERJEE, A., J. DOLADO, AND R. MESTRE (1998): "Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework," *Journal of Time Series Analysis*, 19, 267–283.

BLANCHARD, O. AND R. PEROTTI (2002): "An Empirical Characterization of the Dynamic of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics*, 117, 1329–1368.

BORENSTEIN, S., C. A. CAMERON, AND R. GILBERT (1997): "Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes?" *Quarterly Journal of Economics*, 112, 305–339.

CHESNES, M. (2016): "Asymmetric Pass-Through in U.S. Gasoline Prices," *Energy Journal*, 37, 153–180.

CHO, J. S., M. GREENWOOD-NIMMO, AND Y. SHIN (2023a): "Nonlinear Autoregressive Distributed Lag Model Estimation Using Data with Drifts Applied to the Relationship between Income Inequality and Economic Openness," Discussion paper, School of Economics, Yonsei Univresity.

——— (2023b): "Recent Developments of the Autoregressive Distributed Lag Modelling Framework," *Journal of Economic Surveys*, 37, 7–32.

——— (2023c): “Two-Step NARDL Estimation of the Relationship between R&D Intensity and Investment,” Discussion paper, School of Economics, Yonsei Univresity.

CHO, J. S. AND I. ISHIDA (2012): “Testing for the Effects of Omitted Power Transformations,” *Economics Letters*, 117, 287–290.

CHO, J. S., I. ISHIDA, AND H. WHITE (2011): “Revisiting Tests for Neglected Nonlinearity Using Artificial Neural Networks,” *Neural Computation*, 23, 1133–1186.

CHO, J. S. AND P. C. B. PHILLIPS (2018): “Sequentially Testing Polynomial Model Hypotheses Using Power Transforms of Regressors,” *Journal of Applied Econometrics*, 33, 141–159.

CHO, J. S. AND H. WHITE (2007): “Testing for Regime Switching,” *Econometrica*, 75, 1671–1720.

——— (2010): “Testing for Unobserved Heterogeneity in Exponential and Weibull Duration Models,” *Journal of Econometrics*, 157, 458–480.

CLOYNE, J. (2013): “Discretionary Tax Changes and the Macroeconomy: New Narrative Evidence from the United Kingdom,” *American Economic Review*, 103, 1507–1528.

DE JONG, R. M. AND J. DAVIDSON (2000): “The Functional Central Limit Theorem and Weak Convergence to Stochastic Integrals I: Weakly Dependent Processes,” *Econometric Theory*, 16, 621–642.

DEMIREL, U. D. (2021): “The Short-Term Effects of Tax Changes: The Role of State Dependence,” *Journal of Monetary Economics*, 117, 918–934.

ENGLE, R. F. AND C. W. GRANGER (1987): “Co-integration and Error Correction: Representation, Estimation and Testing,” *Econometrica*, 55, 251–276.

GHASSIBE, M. AND F. ZANETTI (2021): “State Dependence of Fiscal Multipliers: The Source of Fluctuations Natters,” *Journal of Monetary Economics*, 132, 1–23.

GUNTER, S., D. RIERA-CRICHTON, C. A. VEGH, AND G. VULETIN (2021): “Non-Linear Effects of Tax Changes on Output: The Role of the Initial Level of Taxation,” *Journal of International Economics*, 131, 103450.

MERTENS, K. AND J. L. M. OLEA (2018): “Marginal Tax Rates and Income: New Time Series Evidence,” *Quarterly Journal of Economics*, 133, 1803–1884.

MERTENS, K. AND M. O. RAVN (2012): “Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated US Tax Policy Shocks,” *American Economic Journal: Economic Policy*, 4, 145–181.

PESARAN, M. H. AND Y. SHIN (1998): “An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis,” in *Econometrics and Economic Theory: The Ragnar Frisch Centennial Symposium*, ed. by S. Strom, Cambridge: Cambridge University Press, Econometric Society Monographs, 371–413.

PESARAN, M. H., Y. SHIN, AND R. J. SMITH (2001): “Bounds Testing Approaches to the Analysis of Level Relationships,” *Journal of Applied Econometrics*, 16, 289–326.

PHILLIPS, P. (1991): “Optimal Inference in Cointegrated Systems,” *Econometrica*, 59, 283–306.

PHILLIPS, P. AND P. PERRON (1988): “Testing for a Unit Root in Time Series Regression,” *Biometrika*, 75, 335–346.

PHILLIPS, P. C. AND B. E. HANSEN (1990): “Statistical Inference in Instrumental Variable Regression with I(1) Processes,” *Review of Economic Studies*, 57, 99–125.

ROMER, C. D. AND D. H. ROMER (2010): “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks,” *American Economic Review*, 100, 763–801.

SEONG, D., J. S. CHO, AND T. TERÄSVIRTA (2022): “Comprehensively Testing Linearity Hypothesis Using the Smooth Transition Autoregressive Model,” *Econometric Reviews*, 41, 966–984.

SHIN, Y., B. YU, AND M. J. GREENWOOD-NIMMO (2014): “Modelling Asymmetric Cointegration and Dynamic Multipliers in a Nonlinear ARDL Framework,” in *Festschrift in Honor of Peter Schmidt: Econometric Methods and Applications*, ed. by W. Horrace and R. Sickles, New York (NY): Springer Science & Business Media, 281–314.

TERÄSVIRTA, T. (1994): “Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models,” *Journal of the American Statistical Association*, 89, 208–218.

WHITE, H. (2001): *Asymptotic Theory for Econometricians*, San Diego (CA): Academic Press.

WHITE, H. AND J. S. CHO (2012): “Higher-Order Approximations for Testing Neglected Nonlinearity,” *Neural Computation*, 24, 273–287.

Wald test	Method	Bootstrap Method					Mixed Chi-squared Distribution				
	$\alpha \setminus T$	100	200	300	400	500	100	200	300	400	500
$\mathscr{W}_T^{(1)}$	10%	10.70	10.78	10.80	9.58	10.54	38.16	40.26	40.48	40.24	41.02
	5%	5.52	5.48	5.38	4.68	5.12	30.38	31.30	32.50	31.52	32.46
	1%	1.28	1.00	0.94	1.12	0.98	17.46	18.38	19.48	18.34	19.24
$\mathscr{W}_T^{(2)}$	10%	9.46	10.26	9.94	9.46	10.26	11.12	11.24	10.58	9.96	10.52
	5%	4.74	5.10	5.00	4.90	4.94	5.80	5.68	5.38	5.20	5.36
	1%	1.00	1.02	1.06	0.98	1.04	1.44	1.28	1.26	0.94	1.02
$\mathscr{W}_T^{(3)}$	10%	10.26	11.02	10.54	9.84	10.80	33.96	35.14	36.18	35.62	36.64
	5%	5.40	5.62	4.98	4.62	5.44	26.32	27.08	27.90	26.74	27.94
	1%	1.24	1.04	1.00	1.00	1.04	14.28	14.92	15.84	14.28	15.70
$W_T^{(1)}$	10%	10.94	9.90	10.18	9.74	10.44	14.98	12.14	11.48	10.68	11.36
	5%	5.18	4.84	4.98	4.86	4.96	8.70	6.24	5.92	5.46	5.68
	1%	0.88	0.82	1.14	0.84	0.86	2.18	1.42	1.44	1.16	1.04
$W_T^{(2)}$	10%	10.74	9.16	9.58	9.22	10.02	12.68	10.04	10.10	9.60	10.62
	5%	5.06	4.62	5.02	4.58	5.14	6.86	5.20	5.28	4.96	5.28
	1%	1.06	0.96	1.20	1.08	1.14	1.80	1.24	1.30	1.12	1.22
$W_T^{(3)}$	10%	10.78	8.92	9.80	9.42	10.06	14.92	11.46	11.04	10.52	10.90
	5%	4.98	4.46	4.82	4.92	4.94	8.36	5.82	5.64	5.48	5.44
	1%	0.98	0.90	1.16	0.90	1.20	2.42	1.52	1.42	1.24	1.46

Table 1: EMPIRICAL REJECTION RATES OF THE WALD TESTS (IN PERCENT). This table shows the empirical rejection rates of the Wald statistics testing $\mathbb{H}_0' : \beta_* = 0$, $\mathscr{H}_0'' : \pi_*^+ = \pi_*^-$, and $\mathbb{H}_0''' : \beta_* = 0$ and $\pi_*^+ = \pi_*^-$. The total number of repetitions is 5,000, and the bootstrap iteration is 500. DGP: $\Delta y_t = \rho_* u_{t-1} + \pi_*^+ \Delta x_t^+ + \pi_*^- \Delta x_t^- + e_t$, $u_t = y_t - \beta_*^+ x_t^+ - \beta_*^- x_t^-$, $\Delta x_t = v_t$, and $(e_t, v_t)' \sim$ IID $N(\mathbf{0}_2, \mathbf{I}_2)$ with $(\rho_*, \pi_*^+, \pi_*^-, \beta_*^+, \beta_*^-) = (-1/2, 1/2, 1/2, 1, 1)$. Here, $\mathscr{W}_T^{(1)}$, $\mathscr{W}_T^{(2)}$, and $\mathscr{W}_T^{(3)}$ denote the Wald tests in Section 5, and $W_T^{(1)}$, $W_T^{(2)}$, and $W_T^{(3)}$ are the standard Wald tests.

Wald test	Method	Bootstrap Method					Mixed Chi-squared Distribution				
	$\alpha \setminus T$	100	200	300	400	500	100	200	300	400	500
$\mathscr{W}_T^{(1)}$	10%	49.52	88.04	97.98	99.88	99.96	70.78	96.20	99.48	99.96	100.0
	5%	37.34	81.12	96.62	99.60	99.92	65.14	94.50	99.30	99.96	99.98
	1%	15.54	61.22	90.30	98.22	99.64	52.50	90.24	98.64	99.88	99.98
$\mathscr{W}_T^{(2)}$	10%	17.62	25.96	33.14	40.40	48.24	20.18	27.28	34.70	41.24	48.94
	5%	10.58	17.32	21.98	28.90	36.24	12.70	18.60	23.42	30.02	37.34
	1%	2.96	6.00	8.44	11.88	16.30	4.60	6.98	9.44	12.96	17.54
$\mathscr{W}_T^{(3)}$	10%	47.82	86.70	97.84	99.84	99.96	67.46	94.86	99.22	99.96	99.98
	5%	35.50	80.26	96.38	99.54	99.92	60.74	92.78	99.00	99.96	99.98
	1%	15.22	59.92	89.70	97.98	99.62	48.80	87.72	97.88	99.86	99.96
$W_T^{(1)}$	10%	44.88	86.62	98.28	99.62	99.98	53.84	89.18	98.54	99.68	99.98
	5%	33.16	79.72	96.54	99.42	99.94	43.76	83.78	97.32	99.56	99.98
	1%	14.12	59.88	90.36	98.16	99.76	25.30	68.94	92.64	98.56	99.80
$W_T^{(2)}$	10%	17.06	24.66	33.86	40.02	48.20	19.80	26.00	34.62	41.44	48.92
	5%	9.66	15.64	22.94	28.38	35.82	11.72	17.24	24.26	29.34	37.16
	1%	2.90	4.74	8.90	11.78	16.50	4.18	6.04	10.06	13.12	17.64
$W_T^{(3)}$	10%	39.80	82.48	97.08	99.42	99.98	48.72	85.30	97.70	99.50	99.98
	5%	28.60	74.36	94.76	99.02	99.94	38.42	78.44	95.74	99.08	99.94
	1%	11.76	53.86	86.66	97.42	99.52	20.60	61.92	89.80	97.94	99.66

Table 2: EMPIRICAL REJECTION RATES OF THE WALD TESTS (IN PERCENT). This table shows the empirical rejection rates of the Wald statistics testing $\mathbb{H}_0' : \beta_* = 0$, $\mathscr{H}_0'' : \pi_*^+ = \pi_*^-$, and $\mathbb{H}_0''' : \beta_* = 0$ and $\pi_*^+ = \pi_*^-$. The total number of repetitions is 5,000, and the bootstrap iteration is 500. DGP: $\Delta y_t = \rho_* u_{t-1} + \pi_*^+ \Delta x_t^+ + \pi_*^- \Delta x_t^- + e_t$, $u_t = y_t - \beta_*^+ x_t^+ - \beta_*^- x_t^-$, $\Delta x_t = v_t$, and $(e_t, v_t)' \sim$ IID $N(\mathbf{0}_2, \mathbf{I}_2)$ with $(\rho_*, \pi_*^+, \pi_*^-, \beta_*^+, \beta_*^-) = (-1/2, 1/8, -1/8, 1/4, -1/4)$. Here, $\mathscr{W}_T^{(1)}$, $\mathscr{W}_T^{(2)}$, and $\mathscr{W}_T^{(3)}$ denote the Wald tests in Section 5, and $W_T^{(1)}$, $W_T^{(2)}$, and $W_T^{(3)}$ are the standard Wald tests.

Wald test	Method	Bootstrap Method					Mixed Chi-squared Distribution				
	$\alpha \setminus T$	100	200	300	400	500	100	200	300	400	500
$\mathscr{W}_T^{(1)}$	10%	10.72	10.98	10.44	10.24	9.58	50.64	53.24	52.44	52.16	51.76
	5%	5.72	6.04	5.20	5.16	4.80	42.80	45.64	44.82	45.54	44.54
	1%	1.00	1.14	0.90	0.86	1.16	29.24	32.36	32.20	32.72	31.04
$\mathscr{W}_T^{(2)}$	10%	10.18	10.30	10.04	9.90	10.38	12.70	11.16	10.86	10.62	10.90
	5%	4.84	5.36	4.88	4.76	5.12	6.78	6.22	5.42	5.26	5.48
	1%	0.84	1.08	0.92	0.84	1.04	1.70	1.44	1.12	0.98	1.10
$\mathscr{W}_T^{(3)}$	10%	10.26	11.12	9.98	10.12	9.20	47.14	49.32	48.42	48.74	47.70
	5%	5.40	5.78	5.28	5.28	4.62	38.56	41.22	40.00	41.20	39.10
	1%	0.94	1.00	0.86	0.80	1.26	25.36	28.38	27.32	27.82	26.70
$W_T^{(1)}$	10%	9.76	9.70	9.48	9.62	9.78	14.04	11.92	10.68	10.82	10.46
	5%	4.82	5.02	4.78	4.60	5.14	8.08	6.66	5.46	5.50	5.90
	1%	1.08	0.98	0.74	1.08	0.98	2.36	1.54	1.20	1.32	1.20
$W_T^{(2)}$	10%	10.14	10.04	9.96	10.26	9.80	12.80	11.56	10.76	10.82	10.12
	5%	5.36	5.08	4.88	5.12	5.14	7.14	6.04	5.78	5.54	5.34
	1%	1.20	1.00	0.92	0.88	0.86	1.96	1.30	1.18	1.16	1.12
$W_T^{(3)}$	10%	10.48	10.14	9.92	9.64	9.96	15.02	12.58	11.18	10.72	11.12
	5%	5.10	4.72	4.90	5.04	4.82	9.14	6.48	6.00	5.98	5.58
	1%	0.90	0.96	0.84	0.92	0.84	2.44	1.46	1.34	1.24	1.00

Table 3: EMPIRICAL REJECTION RATES OF THE WALD TESTS (IN PERCENT). This table shows the empirical rejection rates of the Wald statistics testing $\mathbb{H}_0' : \beta_* = 0$, $\mathscr{H}_0'' : \pi_*^+ = \pi_*^-$, and $\mathbb{H}_0''' : \beta_* = 0$ and $\pi_*^+ = \pi_*^-$. The total number of repetitions is 5,000, and the bootstrap iteration is 500. DGP: $\Delta y_t = \rho_* u_{t-1} + \pi_*^+ \Delta x_t^+ + \pi_*^- \Delta x_t^- + e_t$, $u_t = y_t - \beta_*^+ x_t^+ - \beta_*^- x_t^-$, $\Delta x_t = \kappa_* \Delta x_{t-1} + v_t$, and $(e_t, v_t)' \sim$ IID $N(\mathbf{0}_2, \mathbf{I}_2)$ with $(\kappa_*, \rho_*, \pi_*^+, \pi_*^-, \beta_*^+, \beta_*^-) = (1/2, -1/2, 1/2, 1/2, 1, 1)$. Here, $\mathscr{W}_T^{(1)}$, $\mathscr{W}_T^{(2)}$, and $\mathscr{W}_T^{(3)}$ denote the Wald tests in Section 5, and $W_T^{(1)}$, $W_T^{(2)}$, and $W_T^{(3)}$ are the standard Wald tests.

Wald test	Method	Bootstrap Method					Mixed Chi-squared Distribution				
	$\alpha \setminus T$	100	200	300	400	500	100	200	300	400	500
$\mathscr{W}_T^{(1)}$	10%	70.42	98.08	99.96	100.0	100.0	86.72	99.54	100.0	100.0	100.0
	5%	58.52	96.24	99.94	99.96	100.0	83.04	99.36	100.0	100.0	100.0
	1%	31.88	88.28	99.36	99.92	99.98	74.30	98.56	99.98	100.0	100.0
$\mathscr{W}_T^{(2)}$	10%	19.34	31.56	42.16	51.48	60.40	23.44	34.30	43.90	53.00	62.24
	5%	11.94	21.60	30.82	38.88	48.20	15.34	24.10	33.22	41.02	49.90
	1%	4.06	7.56	12.20	18.60	25.84	6.20	9.86	14.74	21.34	28.06
$\mathscr{W}_T^{(3)}$	10%	68.38	97.80	99.96	99.98	100.0	83.84	99.40	100.0	100.0	100.0
	5%	56.20	95.96	99.94	99.96	100.0	79.80	99.00	100.0	100.0	100.0
	1%	30.42	87.80	99.34	99.90	99.98	70.72	97.96	99.98	99.96	100.0
$W_T^{(1)}$	10%	68.52	97.96	99.92	100.0	100.0	76.54	98.62	99.94	100.0	100.0
	5%	56.32	95.98	99.80	100.0	100.0	68.12	97.34	99.90	100.0	100.0
	1%	32.50	89.18	99.38	100.0	99.98	48.98	92.80	99.62	100.0	100.0
$W_T^{(2)}$	10%	19.96	31.52	41.64	51.20	60.28	23.56	33.88	43.52	52.62	61.46
	5%	12.58	21.60	30.30	39.14	48.26	15.88	24.22	32.42	41.20	49.70
	1%	3.98	8.30	13.18	18.26	25.12	6.28	10.00	15.90	21.02	27.66
$W_T^{(3)}$	10%	61.92	96.50	99.86	100.0	100.0	70.90	97.36	99.90	100.0	100.0
	5%	49.76	94.32	99.74	100.0	100.0	62.10	95.48	99.82	100.0	100.0
	1%	27.10	85.10	99.14	100.0	99.96	42.90	89.64	99.44	100.0	99.98

Table 4: EMPIRICAL REJECTION RATES OF THE WALD TESTS (IN PERCENT). This table shows the empirical rejection rates of the Wald statistics testing $\mathbb{H}_0' : \beta_* = 0$, $\mathscr{H}_0'' : \pi_*^+ = \pi_*^-$, and $\mathbb{H}_0''' : \beta_* = 0$ and $\pi_*^+ = \pi_*^-$. The total number of repetitions is 5,000, and the bootstrap iteration is 500. DGP: $\Delta y_t = \rho_* u_{t-1} + \pi_*^+ \Delta x_t^+ + \pi_*^- \Delta x_t^- + e_t$, $u_t = y_t - \beta_*^+ x_t^+ - \beta_*^- x_t^-$, $\Delta x_t = \kappa_* \Delta x_{t-1} + v_t$, and $(e_t, v_t)' \sim$ IID $N(\mathbf{0}_2, \mathbf{I}_2)$ with $(\kappa_*, \rho_*, \pi_*^+, \pi_*^-, \beta_*^+, \beta_*^-) = (1/2, -1/2, 1/8, -1/8, 1/4, -1/4)$. Here, $\mathscr{W}_T^{(1)}$, $\mathscr{W}_T^{(2)}$, and $\mathscr{W}_T^{(3)}$ denote the Wald tests in Section 5, and $W_T^{(1)}$, $W_T^{(2)}$, and $W_T^{(3)}$ are the standard Wald tests.

NARDL Model				ARDL Model			
Variables \ Tax	Exo.	Endo.	Sum.	Variables \ Tax	Exo.	Endo.	Sum.
y_{t-1}	-0.0683	-0.0817	-0.0497	y_{t-1}	-0.070	-0.0755	-0.0404
	(0.0191)	(0.0223)	(0.0158)		(0.0185)	(0.0220)	(0.0148)
τ_{t-1}^{+}	-0.0123	0.0158^{**}	0.0016	τ_{t-1}	-0.0139	0.0142	-0.0053
	(0.0078)	(0.0078)	(0.0093)		(0.0060)	(0.0077)	(0.0074)
τ_{t-1}^{-}	-0.0123	0.0149	-0.0093				
	(0.0083)	(0.0176)	(0.0090)				
Trend	0.0549^{***}	0.0631^{***}	0.0324^{**}	Trend	0.0570^{***}	0.0584^{***}	0.0333^{***}
	(0.0177)	(0.0176)	(0.0127)		(0.0151)	(0.0169)	(0.0120)
Constant	0.7588^{***}	0.6974^{***}	0.8357^{***}	Constant	0.7826^{***}	0.7542^{***}	0.8543^{***}
	(0.1874)	(0.1644)	(0.1744)		(0.1539)	(0.1604)	(0.1561)
Δy_{t-1}	0.3129^{***}	0.3118^{***}	0.3135^{***}	Δy_{t-1}	0.3091^{***}	0.3035^{***}	0.3031^{***}
	(0.0633)	(0.0658)	(0.0646)		(0.0630)	(0.0657)	(0.0648)
Δy_{t-2}	0.1265^{*}	0.1446^{**}	0.1190^{*}	Δy_{t-2}	0.1304^{**}	0.1331^{**}	0.1091^{*}
	(0.0646)	(0.0658)	(0.0645)		(0.0643)	(0.0653)	(0.0647)
$\Delta\tau_t^{+}$	-0.0029	0.0890^{*}	0.03984	$\Delta\tau_t$	-0.0444	0.0434	-0.0122
	(0.0435)	(0.0465)	(0.0341)		(0.0272)	(0.0362)	(0.0226)
$\Delta\tau_t^{-}$	-0.0780^{*}	-0.0302	-0.0683^{*}				
	(0.0391)	(0.0598)	(0.0347)				
AIC	-6.5459	-6.5372	-6.5332	AIC	-6.5561	-6.5434	-6.5284
BIC	-6.4225	-6.4099	-6.5332	BIC	-6.4602	-6.4475	-6.4378
t-test	-3.5680^{*}	-3.6519^{*}	$-3.1308^{\dagger}$	t-test	-3.8220^{**}	-3.4319^{*}	-2.7229
F-test	4.5387	4.9164	3.8509	F-test	7.3222^{*}	6.5583^{*}	4.7046
$\mathscr{W}_T^{(1)}$					9.8862	0.0048	9.8910
					(0.4344)	(0.9696)	(0.5000)
$\mathscr{W}_T^{(2)}$					1.4822	3.8085	5.2908
					(0.7472)	(0.1902)	(0.6567)
$\mathscr{W}_T^{(3)}$					30.9528	1.4114	32.3642
					(0.2737)	(0.2659)	(0.2718)
$W_T^{(1)}$					0.0000	0.0039	1.0285
					(0.9982)	(0.9604)	(0.4485)
$W_T^{(2)}$					1.5426	2.5180	4.4950
					(0.2265)	(0.1259)	(0.0398)
$W_T^{(3)}$					1.5527	2.5326	5.2070
					(0.5474)	(0.3706)	(0.1525)

Table 5: OLS ESTIMATION OF THE NARDL AND ARDL MODELS. This table presents the OLS estimation using quarterly data from Romer and Romer (2010). The left and right panels display estimated parameters for (25) and (23), respectively. Figures in parentheses indicate standard errors of the OLS estimates. At the bottom of the top panels, AIC, BIC, t-test, and Pesaran et al.'s (2001) F-test are reported. †, *, **, and *** indicate significance at 25%, 10%, 5%, and 1% levels, respectively. Wald tests in the last two bottom panels show the Wald tests in Section 5 and the standard Wald tests. Figures in parentheses below the Wald tests show p-values. They are obtained from 100,000 bootstrap iterations.

	NARDL Model				ARDL Model			
	Variables \ Tax	Exo.	Endo.	Sum.	Variables \ Tax	Exo.	Endo.	Sum.
Long-Run	Constant	4.2360	3.7413***	4.9730**	Constant	6.0785***	4.8528***	5.8407**
		(2.5919)	(1.2787)	(2.3861)		(0.0709)	(1.3852)	(2.8526)
	τ^+_{t-1}	-0.4329***	0.2715***	0.2759	τ_{t-1}	-0.2328***	0.2523***	0.1564
		(0.13189)	(0.0546)	(0.1723)		(0.0709)	(0.0591)	(0.1716)
	τ^-_{t-1}	-0.3202*	0.3276*	0.2598				
		(0.1722)	(0.1757)	(0.1826)				
	Trend	0.8340***	0.8461***	0.8436***	Trend	0.8261***	0.8367***	0.8287***
		(0.0184)	(0.0090)	(0.0169)		(0.0127)	(0.0098)	(0.0202)
Short-Run	u_{t-1}	-0.0683*	-0.0817*	-0.0497†	u_{t-1}	-0.0708**	-0.0755*	-0.0404
		(0.0191)	(0.0223)	(0.0158)		(0.0185)	(0.0220)	(0.0148)
	Constant	0.6734***	0.6918***	0.5683***	Constant	0.5752***	0.6590***	0.6295***
		(0.1424)	(0.1458)	(0.1472)		(0.5752)	(0.1392)	(0.1404)
	Δy_{t-1}	0.3129***	0.3118***	0.3135***	Δy_{t-1}	0.3091***	0.3035***	0.3031***
		(0.0633)	(0.0658)	(0.0646)		(0.0630)	(0.0657)	(0.0648)
	Δy_{t-2}	0.1265*	0.1446**	0.1190*	Δy_{t-2}	0.1304**	0.1331**	0.1091*
		(0.0646)	(0.0658)	(0.0645)		(0.0643)	(0.0653)	(0.0647)
	$\Delta\tau^+_t$	-0.0029	0.0890*	0.0398	$\Delta\tau_t$	-0.0444	0.0434	-0.0122
		(0.0435)	(0.0465)	(0.0341)		(0.0272)	(0.0362)	(0.0226)
	$\Delta\tau^-_t$	-0.0780**	-0.0302	-0.0683*				
		(0.0391)	(0.0598)	(0.0347)				

Table 6: 2SNARDL ESTIMATION OF THE NARDL AND ARDL MODELS. This table presents the 2SNARDL estimation using the quarterly data from Romer and Romer (2010). The left and right panels display estimated parameters for (24) and (22), respectively. †, *, **, and *** imply that the tests are significant at 25%, 10%, 5%, and 1%, respectively.

NARDL Model				ARDL Model			
Variables \ Tax	Exo. ratio	Endo. ratio	Sum. ratio	Variables \ Tax	Exo. ratio	Endo. ratio	Sum. ratio
y_{t-1}	-0.0609	-0.0633	-0.0460	y_{t-1}	-0.0552	-0.0658	-0.0400
	(0.0184)	(0.0251)	(0.0177)		(0.0163)	(0.0247)	(0.0160)
r^+_{t-1}	-0.1992	0.0850	0.0037	r_{t-1}	-0.1125	0.1003	-0.0351
	(0.1349)	(0.0954)	(0.0789)		(0.0743)	(0.0925)	(0.0618)
r^-_{t-1}	-0.0970	0.0401	-0.0566				
	(0.0801)	(0.1345)	(0.0645)				
Trend	0.0499***	0.0488**	0.0324**	Trend	0.0419***	0.0520***	0.0317***
	(0.0173)	(0.0200)	(0.0136)		(0.0127)	(0.0194)	(0.0135)
Constant	0.2683	0.3673	0.3685*	Constant	0.4054**	0.4700***	0.5261***
	(0.2363)	(0.2380)	(0.2047)		(0.1787)	(0.1663)	(0.1543)
Δy_{t-1}	0.3095***	0.3144***	0.3049***	Δy_{t-1}	0.3042***	0.3115***	0.2921***
	(0.0639)	(0.0672)	(0.0657)		(0.0636)	(0.0660)	(0.0650)
Δy_{t-2}	0.1223*	0.1265**	0.1139*	Δy_{t-2}	0.1168*	0.1242*	0.1073*
	(0.0652)	(0.0660)	(0.0651)		(0.0644)	(0.0657)	(0.0648)
Δr^+_t	0.1638	0.2465	0.1778	Δr_t	0.1725	0.2553	0.1467
	(0.6215)	(0.2788)	(0.2621)		(0.618)0	(0.2776)	(0.2609)
Δr^-_t	-0.2706	-0.1477	-0.2350				
	(0.2734)	(0.3051)	(0.2087)				
AIC	2.6831	2.6925	2.6895	AIC	2.6739	2.6780	2.6814
BIC	2.8133	2.8226	2.8196	BIC	2.7751	2.7792	2.7826
t-test	-3.3087†	-2.5198	-2.5917	t-test	-3.3867†	-2.6602	-2.4960
F-test	3.9276	3.5933	3.4873	F-test	5.7417	5.2684	4.8210
$\mathscr{W}_T^{(1)}$					52.7574	2.1358	48.9740
					(0.1294)	(0.7290)	(0.2581)
$\mathscr{W}_T^{(2)}$					0.0722	0.8121	0.4782
					(0.8015)	(0.4021)	(0.4981)
$\mathscr{W}_T^{(3)}$					52.8296	2.9480	49.4522
					(0.1340)	(0.7595)	(0.2610)
$W_T^{(1)}$					0.5384	0.3344	0.8601
					(0.5697)	(0.6463)	(0.5035)
$W_T^{(2)}$					0.4079	0.9926	1.5118
					(0.5278)	(0.3177)	(0.2279)
$W_T^{(3)}$					1.0869	1.2200	2.1652
					(0.6618)	(0.6150)	(0.4681)

Table 7: OLS ESTIMATION OF THE NARDL AND ARDL MODELS. This table presents the OLS estimation using quarterly data from Romer and Romer (2010). The left and right panels display estimated parameters for (25) and (23) using r_t instead of τ_t, respectively. Figures in parentheses indicate standard errors of the OLS estimates. At the bottom of the top panels, AIC, BIC, t-test, and Pesaran et al.'s (2001) F-test are reported. †, *, **, and *** indicate significance at 25%, 10%, 5%, and 1% levels, respectively. Wald tests in the last two bottom panels show the Wald tests in Section 5 and the standard Wald tests. Figures in parentheses below the Wald tests show p-values. They are obtained from 100,000 bootstrap iterations.

	NARDL Model				ARDL Model			
	Variables \ Tax	Exo. ratio	Endo. ratio	Sum. ratio	Variables \ Tax	Exo. ratio	Endo. ratio	Sum. ratio
Long-Run	Constant	-2.5602	-2.5888**	-2.9531	Constant	-0.3468	-2.2311*	-1.0492
		(2.4603)	(1.1990)	(2.5202)		(2.3897)	(1.2927)	(2.8199)
	r^+_{t-1}	-6.7407***	4.2616***	4.5455***	r_{t-1}	-3.0752**	3.3960***	2.2608*
		(2.0794)	(0.5100)	(1.2686)		(1.2428)	(0.5463)	(1.2356)
	r^-_{t-1}	-2.8360*	5.1176***	2.762**3				
		(1.5211)	(0.9583)	(1.2373)				
	Trend	0.8364***	0.8391***	0.8435***	Trend	0.8242***	0.8370***	0.8295***
		(0.0166)	(0.0081)	(0.0170)		(0.0161)	(0.0087)	(0.0190)
Short-Run	u_{t-1}	-0.0609†	-0.0633	-0.0460	u_{t-1}	-0.0558*	-0.0627	-0.0378
		(0.0184)	(0.0251)	(0.0177)		(0.0162)	(0.0248)	(0.0160)
	Constant	0.6763***	0.6693***	0.6719***	Constant	0.5730***	0.6828***	0.6517***
		(0.1484)	(0.1510)	(0.1534)		(0.1470)	(0.1468)	(0.1464)
	Δy_{t-1}	0.3095***	0.3144***	0.3049***	Δy_{t-1}	0.3038***	0.3127***	0.3007***
		(0.0639)	(0.0672)	(0.0657)		(0.0635)	(0.0669)	(0.0656)
	Δy_{t-2}	0.1223*	0.1265*	0.1139*	Δy_{t-2}	0.1149*	0.1220*	0.1070
		(0.0652)	(0.0660)	(0.0651)		(0.0644)	(0.0658)	(0.0648)
	Δr^+_t	0.1638	0.2465	0.1778	Δr_t	-0.2094	0.0801	-0.0737
		(0.6215)	(0.2788)	(0.2621)		(0.2435)	(0.2087)	(0.1596)
	Δr^-_t	-0.2706	-0.1477	-0.2350				
		(0.2734)	(0.3051)	(0.2087)				

Table 8: 2SNARDL estimation of the NARDL and ARDL Models. This table presents the 2SNARDL estimation using the quarterly data from Romer and Romer (2010). The left and right panels display estimated parameters for (24) and (22), respectively. †, *, **, and *** imply that the tests are significant at 25%, 10%, 5%, and 1%, respectively.

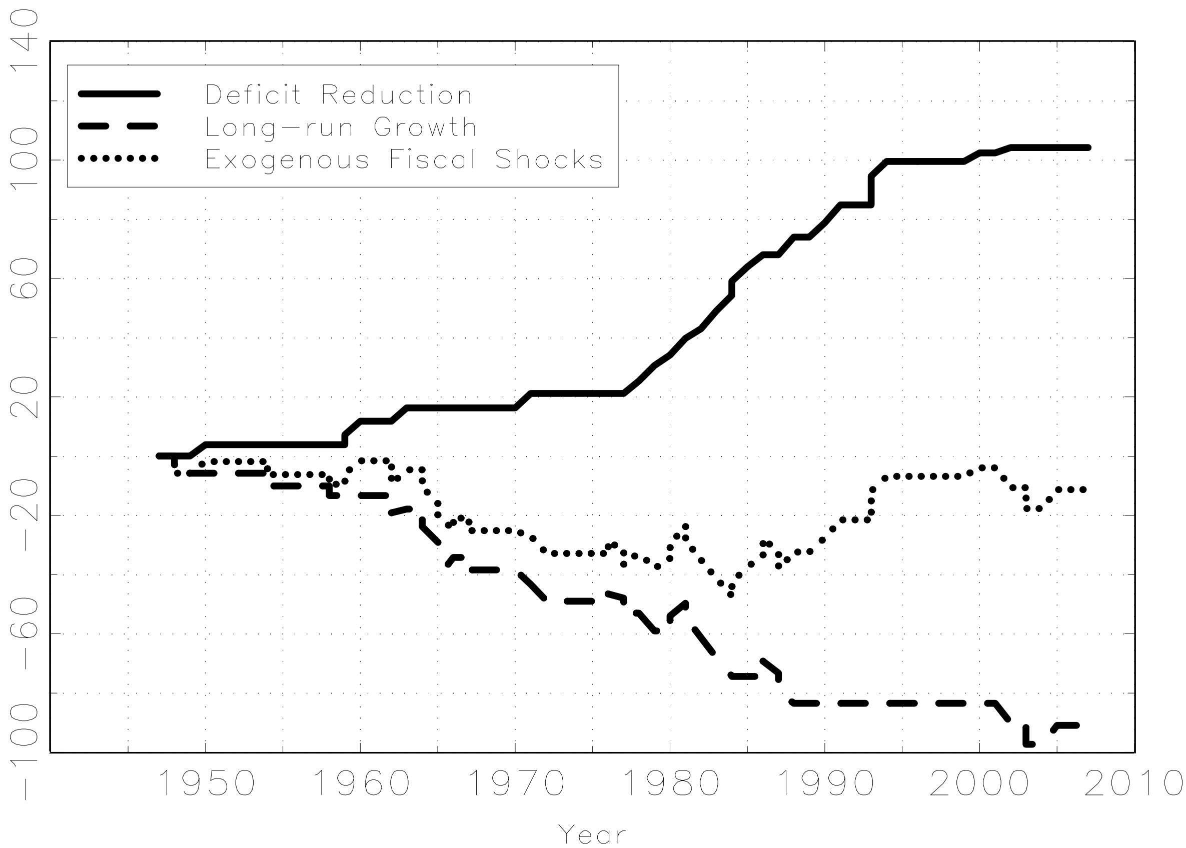

Figure 1: PARTIAL SUM PROCESS FORMED BY EXOGENOUS FISCAL SHOCKS. The solid, dashed, and dotted lines represent deficit reduction, long-run growth, and exogenous fiscal shocks, respectively.

Online Supplement for

'Estimating and Inferring the Nonlinear Autoregressive Distributed Lag Model by Ordinary Least Squares'[1]

by

Jin Seo Cho

Yonsei University

This Online Supplement is an Appendix that provides proofs of all the results in the paper, including the lemmas.

A Supplements

A.1 Supplementary Lemmas

Before proving the main claims in the text, we first provide some preliminary lemmas to prove the main claims efficiently.

The following two lemmas provide alternative representations of the OLS estimators, which we define under different environments. We first suppose that $y_t \in \mathbb{R}$ is a dependent variable and $(\boldsymbol{x}_t', \boldsymbol{z}_t')' \in \mathbb{R}^{(s+k)}$ is an explanatory variable, and the OLS estimator is obtained by regressing y_t against $(\boldsymbol{x}_t', \boldsymbol{z}_t')'$. Given this, we provide alternative forms of the OLS estimator in the following lemmas:

Lemma A.1. *Suppose that $\{(y_t, \boldsymbol{x}_t', \boldsymbol{z}_t')' \in \mathbb{R}^{1+s+k} : t = 1, 2, \ldots, T\}$. If the OLS estimators are obtained as follows: for $j = 1, 2, \ldots, s$,*

$$(\widehat{\boldsymbol{\beta}}_T, \widehat{\boldsymbol{\gamma}}_T) := \arg\min_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \sum_{t=1}^{T} \left(y_t - \boldsymbol{x}_t \boldsymbol{\beta} - \boldsymbol{z}_t' \boldsymbol{\gamma} \right)^2, \quad \widehat{\boldsymbol{\phi}}_{jT} := \arg\min_{\boldsymbol{\phi}_j} \sum_{t=1}^{T} \left(x_{jt} - \boldsymbol{z}_t' \boldsymbol{\phi}_j \right)^2, \quad \textit{and}$$

$$(\widehat{\boldsymbol{\xi}}_T, \widehat{\boldsymbol{\delta}}_T) := \arg\min_{\boldsymbol{\xi}, \boldsymbol{\delta}} \sum_{t=1}^{T} \left(y_t - \widehat{\boldsymbol{v}}_t' \boldsymbol{\xi} - \boldsymbol{z}_t' \boldsymbol{\delta} \right)^2,$$

where for each t, $\widehat{\boldsymbol{v}}_t := \boldsymbol{x}_t - \widehat{\boldsymbol{\phi}}_T' \boldsymbol{z}_t$ and $\widehat{\boldsymbol{\phi}}_T := (\widehat{\boldsymbol{\phi}}_{1T}, \ldots, \widehat{\boldsymbol{\phi}}_{sT})$, then $\widehat{\boldsymbol{\beta}}_T = \widehat{\boldsymbol{\xi}}_T$ and $\widehat{\boldsymbol{\gamma}}_T = \widehat{\boldsymbol{\delta}}_T - \widehat{\boldsymbol{\phi}}_T \widehat{\boldsymbol{\xi}}_T$. $\square$

[1]The author has benefited from discussions with Matthew Grrnwood-Nimmo and Yoncheol Shin. Cho acknowledges the research grant provided by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF2020S1A5A2A01040235).

Therefore, we can obtain the OLS estimator $(\widehat{\beta}_T, \widehat{\gamma}_T)$ by combining the two OLS estimators $\widehat{\boldsymbol{\phi}}_T$ and $(\widehat{\boldsymbol{\xi}}_T, \widehat{\boldsymbol{\delta}}_T)$ obtained from the first- and second-step OLS estimators, respectively.

The following lemma considers a different environment. We let the OLS estimator be obtained by regressing y_t against $(\boldsymbol{x}_t, \boldsymbol{z}_t', \boldsymbol{w}_t')' \in \mathbb{R}^{s+k+d}$ and provide an alternative form of the OLS estimator:

Lemma A.2. *Suppose that* $\{(y_t, \boldsymbol{x}_t', \boldsymbol{z}_t', \boldsymbol{w}_t')' \in \mathbb{R}^{1+s+k+d} : t = 1, 2, \ldots, T\}$. *If the OLS estimators are obtained as follows: for* $j = 1, 2, \ldots, s$*,*

$$(\widehat{\boldsymbol{\beta}}_T, \widehat{\boldsymbol{\gamma}}_T, \widehat{\boldsymbol{\alpha}}_T) := \arg\min_{\beta, \boldsymbol{\gamma}, \boldsymbol{\alpha}} \sum_{t=1}^{T} \left(y_t - \boldsymbol{x}_t'\boldsymbol{\beta} - \boldsymbol{z}_t'\boldsymbol{\gamma} - \boldsymbol{w}_t'\boldsymbol{\alpha}\right)^2, \quad \widehat{\boldsymbol{\phi}}_{jT} := \arg\min_{\boldsymbol{\phi}} \sum_{t=1}^{T} \left(x_{jt} - \boldsymbol{z}_t'\boldsymbol{\phi}_j\right)^2, \quad \textit{and}$$

$$(\widehat{\boldsymbol{\xi}}_T, \widehat{\boldsymbol{\delta}}_T, \widehat{\boldsymbol{\theta}}_T) := \arg\min_{\boldsymbol{\xi}, \boldsymbol{\delta}, \boldsymbol{\theta}} \sum_{t=1}^{T} \left(y_t - \widehat{\boldsymbol{v}}_t'\boldsymbol{\xi} - \boldsymbol{z}_t'\boldsymbol{\delta} - \boldsymbol{w}_t'\boldsymbol{\theta}\right)^2,$$

where for each t, $\widehat{\boldsymbol{v}}_t := \boldsymbol{x}_t - \widehat{\boldsymbol{\phi}}_T'\boldsymbol{z}_t$ *and* $\widehat{\boldsymbol{\phi}}_T := (\widehat{\boldsymbol{\phi}}_{1T}, \ldots, \widehat{\boldsymbol{\phi}}_{sT})$*, then* $\widehat{\boldsymbol{\beta}}_T = \widehat{\boldsymbol{\xi}}_T$*,* $\widehat{\boldsymbol{\gamma}}_T = \widehat{\boldsymbol{\delta}}_T - \widehat{\boldsymbol{\phi}}_T\widehat{\boldsymbol{\xi}}_T$*, and* $\widehat{\boldsymbol{\alpha}}_T = \widehat{\boldsymbol{\theta}}_T$. □

Note that $\boldsymbol{w}_t$ is added as an additional regressor to the regressors given in Lemma A.1 and that the nuisance parameter estimator $\widehat{\boldsymbol{\alpha}}_T$ is the same as the nuisance parameter estimator $\widehat{\boldsymbol{\theta}}_T$ obtained in the second step.

We now prove Lemmas A.1 and A.2. For notational simplicity, we let

$$\mathbf{Y} := [y_1, y_2, \ldots, y_T]', \quad \mathbf{X} := [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T]', \quad \mathbf{Z} := [\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_T]', \quad \widehat{\mathbf{V}} := [\widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_T]',$$

and $\mathbf{W} := [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_T]'$.

Proof of Lemma A.1. From the definition of $(\widehat{\beta}_T, \widehat{\gamma}_T)$, we first note that

$$\begin{bmatrix} \widehat{\beta}_T \\ \widehat{\gamma}_T \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}'\mathbf{Y} \\ \mathbf{Z}'\mathbf{Y} \end{bmatrix} = \begin{bmatrix} (\mathbf{X}'\mathbf{Q}\mathbf{X})^{-1}\mathbf{X}'\mathbf{Q}\mathbf{Y} \\ (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{Q}\mathbf{X})^{-1}\mathbf{X}'\mathbf{Q}]\mathbf{Y} \end{bmatrix}, \tag{A.1}$$

where $\mathbf{Q} := \mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$. Next, we note that

$$\widehat{\mathbf{V}} = \mathbf{X} - \mathbf{Z}\widehat{\phi}_T = \mathbf{X} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X} = \mathbf{Q}\mathbf{X}. \tag{A.2}$$

Therefore, $\mathbf{Z}'\widehat{\mathbf{V}} = \mathbf{0}$ by noting that $\mathbf{Z}'\mathbf{Q} = \mathbf{0}$. Third, we note that

$$\begin{bmatrix} \widehat{\xi}_T \\ \widehat{\delta}_T \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{V}}'\widehat{\mathbf{V}} & \widehat{\mathbf{V}}'\mathbf{Z} \\ \mathbf{Z}'\widehat{\mathbf{V}} & \mathbf{Z}'\mathbf{Z} \end{bmatrix}^{-1} \begin{bmatrix} \widehat{\mathbf{V}}'\mathbf{Y} \\ \mathbf{Z}'\mathbf{Y} \end{bmatrix} = \begin{bmatrix} (\widehat{\mathbf{V}}'\widehat{\mathbf{V}})^{-1}\widehat{\mathbf{V}}'\mathbf{Y} \\ (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y} \end{bmatrix}$$

using the fact that $\mathbf{Z}'\widehat{\mathbf{V}} = \mathbf{0}$. Therefore,

$$\widehat{\boldsymbol{\xi}}_T = (\widehat{\mathbf{V}}'\widehat{\mathbf{V}})^{-1}\widehat{\mathbf{V}}'\mathbf{Y} = (\mathbf{X}'\mathbf{Q}\mathbf{X})^{-1}\mathbf{X}'\mathbf{Q}\mathbf{Y} \tag{A.3}$$

using (A.2). This shows that $\widehat{\boldsymbol{\beta}}_T = \widehat{\boldsymbol{\xi}}_T$. Finally, we note that

$$\widehat{\boldsymbol{\delta}}_T - \widehat{\boldsymbol{\phi}}_T\widehat{\boldsymbol{\xi}}_T = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y} - (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}(\widehat{\mathbf{V}}'\widehat{\mathbf{V}})^{-1}\widehat{\mathbf{V}}'\mathbf{Y} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{Q}\mathbf{X})^{-1}\mathbf{X}'\mathbf{Q}]\mathbf{Y} = \widehat{\boldsymbol{\gamma}}_T,$$

which follows from (A.1), where the second equality follows from (A.3). Thus, $\widehat{\boldsymbol{\delta}}_T - \widehat{\boldsymbol{\phi}}_T\widehat{\boldsymbol{\xi}}_T = \widehat{\boldsymbol{\gamma}}_T$. This completes the proof. ■

Proof of Lemma A.2. To prove the claim, we represent the OLS estimators in different forms. If we let

$$(\widehat{\boldsymbol{\beta}}_T(\boldsymbol{\alpha}), \widehat{\boldsymbol{\gamma}}_T(\boldsymbol{\alpha})) := \arg\min_{\boldsymbol{\beta},\boldsymbol{\gamma}} \sum_{t=1}^{T} \left(y_t - \boldsymbol{x}_t'\boldsymbol{\beta} - \boldsymbol{z}_t'\boldsymbol{\gamma} - \boldsymbol{w}_t'\boldsymbol{\alpha}\right)^2,$$

then

$$\widehat{\boldsymbol{\alpha}}_T = \arg\min_{\boldsymbol{\alpha}} \sum_{t=1}^{T} \left(y_t - \boldsymbol{x}_t\widehat{\boldsymbol{\beta}}_T(\boldsymbol{\alpha}) - \boldsymbol{z}_t'\widehat{\boldsymbol{\gamma}}_T(\boldsymbol{\alpha}) - \boldsymbol{w}_t'\boldsymbol{\alpha}\right)^2,$$

and $(\widehat{\boldsymbol{\beta}}_T(\widehat{\boldsymbol{\alpha}}_T), \widehat{\boldsymbol{\gamma}}_T(\widehat{\boldsymbol{\alpha}}_T)) = (\widehat{\boldsymbol{\beta}}_T, \widehat{\boldsymbol{\gamma}}_T)$. Likewise, if we let

$$(\widehat{\boldsymbol{\xi}}_T(\boldsymbol{\theta}), \widehat{\boldsymbol{\delta}}_T(\boldsymbol{\theta})) := \arg\min_{\xi,\boldsymbol{\delta}} \sum_{t=1}^{T} \left(y_t - \widehat{\boldsymbol{v}}_t'\boldsymbol{\xi} - \boldsymbol{z}_t'\boldsymbol{\delta} - \boldsymbol{w}_t'\boldsymbol{\theta}\right)^2,$$

then

$$\widehat{\boldsymbol{\theta}}_T = \arg\min_{\boldsymbol{\theta}} \sum_{t=1}^{T} \left(y_t - \widehat{\boldsymbol{v}}_t'\widehat{\boldsymbol{\xi}}_T(\boldsymbol{\theta}) - \boldsymbol{z}_t'\widehat{\boldsymbol{\delta}}_T(\boldsymbol{\theta}) - \boldsymbol{w}_t'\boldsymbol{\theta}\right)^2,$$

and $(\widehat{\boldsymbol{\xi}}_T(\widehat{\boldsymbol{\theta}}_T), \widehat{\boldsymbol{\delta}}_T(\widehat{\boldsymbol{\theta}}_T)) = (\widehat{\boldsymbol{\xi}}_T, \widehat{\boldsymbol{\delta}}_T)$.

Here, for each $\boldsymbol{\alpha}$, if we let $y_t(\boldsymbol{\alpha}) := y_t - \boldsymbol{w}_t'\boldsymbol{\alpha}$,

$$(\widehat{\boldsymbol{\beta}}_T(\cdot), \widehat{\boldsymbol{\gamma}}_T(\cdot)) := \arg\min_{\boldsymbol{\beta},\boldsymbol{\gamma}} \sum_{t=1}^{T} \left(y_t(\cdot) - \boldsymbol{x}_t'\boldsymbol{\beta} - \boldsymbol{z}_t'\boldsymbol{\gamma}\right)^2, \quad \text{and}$$

$$(\widehat{\boldsymbol{\xi}}_T(\cdot), \widehat{\boldsymbol{\delta}}_T(\cdot)) := \arg\min_{\boldsymbol{\xi},\boldsymbol{\delta}} \sum_{t=1}^{T} \left(y_t(\cdot) - \widehat{\boldsymbol{v}}_t'\boldsymbol{\xi} - \boldsymbol{z}_t'\boldsymbol{\delta}\right)^2,$$

so that Lemma A.1 implies that $\widehat{\boldsymbol{\beta}}_T(\cdot) = \widehat{\boldsymbol{\xi}}_T(\cdot)$ and $\widehat{\boldsymbol{\gamma}}_T(\cdot) = \widehat{\boldsymbol{\delta}}_T(\cdot) - \widehat{\boldsymbol{\phi}}_T \widehat{\boldsymbol{\xi}}_T(\cdot)$. Therefore,

$$\sum_{t=1}^{T} \left(y_t(\cdot) - \widehat{\boldsymbol{v}}_t' \widehat{\boldsymbol{\xi}}_T(\cdot) - \boldsymbol{z}_t' \widehat{\boldsymbol{\delta}}_T(\cdot) \right)^2 = \sum_{t=1}^{T} \left(y_t(\cdot) - (\boldsymbol{x}_t - \widehat{\boldsymbol{\phi}}_T' \boldsymbol{z}_t)' \widehat{\boldsymbol{\xi}}_T(\cdot) - \boldsymbol{z}_t' \widehat{\boldsymbol{\delta}}_T(\cdot) \right)^2$$
$$= \sum_{t=1}^{T} \left(y_t(\cdot) - \boldsymbol{x}_t' \widehat{\boldsymbol{\xi}}_T(\cdot) - \boldsymbol{z}_t' (\widehat{\boldsymbol{\delta}}_T(\cdot) - \widehat{\boldsymbol{\phi}}_T \widehat{\boldsymbol{\xi}}_T(\cdot)) \right)^2 = \sum_{t=1}^{T} \left(y_t(\cdot) - \boldsymbol{x}_t' \widehat{\boldsymbol{\beta}}_T(\cdot) - \boldsymbol{z}_t' \widehat{\boldsymbol{\gamma}}_T(\cdot) \right)^2,$$

implying that

$$\arg\min_{\boldsymbol{\alpha}} \sum_{t=1}^{T} \left(y_t(\boldsymbol{\alpha}) - \widehat{\boldsymbol{v}}_t' \widehat{\boldsymbol{\xi}}_T(\boldsymbol{\alpha}) - \boldsymbol{z}_t' \widehat{\boldsymbol{\delta}}_T(\boldsymbol{\alpha}) \right)^2 = \arg\min_{\boldsymbol{\theta}} \sum_{t=1}^{T} \left(y_t(\boldsymbol{\theta}) - \boldsymbol{x}_t' \widehat{\boldsymbol{\beta}}_T(\boldsymbol{\theta}) - \boldsymbol{z}_t' \widehat{\boldsymbol{\gamma}}_T(\boldsymbol{\theta}) \right)^2,$$

viz., $\widehat{\boldsymbol{\alpha}}_T = \widehat{\boldsymbol{\theta}}_T$. Thus, it follows that $\widehat{\boldsymbol{\beta}}_T(\widehat{\boldsymbol{\alpha}}_T) = \widehat{\boldsymbol{\xi}}_T(\widehat{\boldsymbol{\theta}}_T)$ and $\widehat{\boldsymbol{\gamma}}_T(\widehat{\boldsymbol{\alpha}}_T) = \widehat{\boldsymbol{\delta}}_T(\widehat{\boldsymbol{\theta}}_T) - \widehat{\boldsymbol{\phi}}_T \widehat{\boldsymbol{\xi}}_T(\widehat{\boldsymbol{\theta}}_T)$. That is, $\widehat{\boldsymbol{\beta}}_T = \widehat{\boldsymbol{\xi}}_T$ and $\widehat{\boldsymbol{\gamma}}_T = \widehat{\boldsymbol{\delta}}_T - \widehat{\boldsymbol{\phi}}_T \widehat{\boldsymbol{\xi}}_T$. This completes the proof. ■

The following lemma shows that $\widehat{\boldsymbol{v}}_T$ and $\widetilde{\boldsymbol{\omega}}_T$ defined in Section 3 suffer from asymptotically singular matrix problems.

Lemma A.3. *Given Assumption 1,*

(i) $T^{-1} \sum_{t=1}^{T} \widehat{u}_{t-1}^2 \xrightarrow{\mathbb{P}} \sigma_u^2 := \mathbb{E}[u_t^2]$;

(ii) if we let $\widetilde{\mathbf{D}}_1 := diag[T^{3/2}\mathbf{I}_{2k}, T^{3/2}, T^{1/2}]$, $T^{-1/2} \left(\sum_{t=1}^{T} \widehat{u}_{t-1} \widetilde{\boldsymbol{z}}_{1t} \right) \widetilde{\mathbf{D}}_1^{-1} \Rightarrow \widetilde{\mathbf{M}}_{1u} := \mathbf{0}_{1\times(2k+2)}$;

(iii) $\widetilde{\mathbf{D}}_1^{-1} \left(\sum_{t=1}^{T} \widetilde{\boldsymbol{z}}_{1t} \widetilde{\boldsymbol{z}}_{1t}' \right) \widetilde{\mathbf{D}}_1^{-1} \Rightarrow \widetilde{\mathcal{M}}_{11}$, *where*

$$\widetilde{\mathcal{M}}_{11} := \begin{bmatrix} \frac{1}{3}\boldsymbol{\mu}_*\boldsymbol{\mu}_*' & \frac{1}{3}\boldsymbol{\mu}_* & \frac{1}{2}\boldsymbol{\mu}_* \\ \frac{1}{3}\boldsymbol{\mu}_*' & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2}\boldsymbol{\mu}_*' & \frac{1}{2} & 1 \end{bmatrix};$$

(iv) if we let $\widetilde{\mathbf{D}}_2 := diag[T^{1/2}\mathbf{I}_{p+2kq-1}]$, $T^{-1/2} \left(\sum_{t=1}^{T} \widehat{u}_{t-1} \widetilde{\boldsymbol{z}}_{2t} \right) \widetilde{\mathbf{D}}_2^{-1} \xrightarrow{\mathbb{P}} \widetilde{\mathbf{M}}_{2u} := \mathbb{E}[u_{t-1}\boldsymbol{z}_{2t}]$;

(v) $\widetilde{\mathbf{D}}_2^{-1} \left(\sum_{t=1}^{T} \widetilde{\boldsymbol{z}}_{2t} \widetilde{\boldsymbol{z}}_{1t}' \right) \widetilde{\mathbf{D}}_1^{-1} \Rightarrow \widetilde{\mathcal{M}}_{21}$, *where*

$$\widetilde{\mathcal{M}}_{21} := \begin{bmatrix} \frac{1}{2}\delta_*\boldsymbol{\iota}_{p-1}\boldsymbol{\mu}_*' & \frac{1}{2}\delta_*\boldsymbol{\iota}_{p-1} & \delta_*\boldsymbol{\iota}_{p-1} \\ \frac{1}{2}\boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_*\boldsymbol{\mu}_*' & \frac{1}{2}\boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_* & \boldsymbol{\iota}_q \otimes \boldsymbol{\mu}_* \end{bmatrix};$$

(vi) $\widetilde{\mathbf{D}}_2^{-1} \left(\sum_{t=1}^{T} \widetilde{\boldsymbol{z}}_{2t} \widetilde{\boldsymbol{z}}_{2t}' \right) \widetilde{\mathbf{D}}_2^{-1} \xrightarrow{\mathbb{P}} \widetilde{\mathbf{M}}_{22} := \mathbf{M}_{22}$;

(vii) *if we let* $\widetilde{\mathbf{D}} := diag[T^{1/2}, \widetilde{\mathbf{D}}_1, \widetilde{\mathbf{D}}_2]$, $\widetilde{\mathbf{D}}^{-1}\left(\sum_{t=1}^{T} \widetilde{\boldsymbol{z}}_t \widetilde{\boldsymbol{z}}_t'\right)\widetilde{\mathbf{D}}^{-1} \Rightarrow \widetilde{\mathscr{M}}$, *where*

$$\widetilde{\mathscr{M}} := \begin{bmatrix} \sigma_u^2 & \widetilde{\mathbf{M}}_{u1} & \widetilde{\mathbf{M}}_{u2} \\ \widetilde{\mathbf{M}}_{1u} & \widetilde{\mathscr{M}}_{11} & \widetilde{\mathscr{M}}_{12} \\ \widetilde{\mathbf{M}}_{2u} & \widetilde{\mathscr{M}}_{21} & \widetilde{\mathbf{M}}_{22} \end{bmatrix},$$

which is singular, $\widetilde{\mathscr{M}}_{12} := \widetilde{\mathscr{M}}_{21}'$, $\widetilde{\mathbf{M}}_{u1} := \widetilde{\mathbf{M}}_{1u}'$, *and* $\widetilde{\mathbf{M}}_{u2} := \widetilde{\mathbf{M}}_{2u}'$; *and*

(viii) $\widetilde{\mathbf{D}}_1^{-1}\left(\sum_{t=1}^{T} \widetilde{\boldsymbol{r}}_{t-1}\widetilde{\boldsymbol{r}}_{t-1}'\right)\widetilde{\mathbf{D}}_1^{-1} \xrightarrow{\mathbb{P}} \widetilde{\mathscr{M}}_{11}$, *which is singular.* □

Both $\widehat{\boldsymbol{v}}_T$ and $\widetilde{\boldsymbol{\omega}}_T$ suffer from the asymptotically singular matrix problem by Lemma A.3 (*vii* and *viii*). Specifically, every column from the second to $(1+k)$-th columns of $\widetilde{\mathscr{M}}$ is proportional to the $(2+2k)$-th column. Likewise, every column from the first to k-th columns in $\widetilde{\mathscr{M}}_{11}$ is proportional to the $(1+2k)$-th column of $\widetilde{\mathscr{M}}_{11}$.

We now prove Lemma A.3.

Proof of Lemma A.3. (*i*) This follows from Lemma B.6 (*iv*) given in Section A.2.

(*ii*) This follows from Lemmas B.3 (*v*), B.4 (*v*), and B.5 (*iv*) given in Section A.2.

(*iii*) This follows from Lemmas B.3 (*i*, *ii*), B.4 (*i*, *ii*), and B.5 (*i*) given in Section A.2.

(*iv*) This follows from Lemma B.2 (*vii*) given in Section A.2.

(*v*) This follows from Lemma B.2 (*ii*, *iii*, and *iv*) given in Section A.2.

(*vi*) This follows from Lemma B.2 (*i*) given in Section A.2.

(*vii*) The given weak convergence follows from Lemma A.3 (*i*, *ii*, *iii*, *iv*, *v*, and *vi*) given in Section A.2, and the singularity follows from the structure of $\widetilde{\mathscr{M}}$.

(*viii*) This follows from the definition of $\widetilde{\boldsymbol{r}}_{t-1} := \widetilde{\boldsymbol{z}}_{1t}$ and Lemma A.3 (*iii*) given in Section A.2. ■

A.2 Preliminary Lemmas

We next provide preliminary lemmas to prove the main claims efficiently.

Lemma B.1. *Given Assumption 1,* $\mathbf{B}_T(\cdot) := T^{-1/2}\sum_{t=1}^{\lfloor(\cdot)T\rfloor} \boldsymbol{w}_t \Rightarrow \mathscr{B}(\cdot)$. □

Lemma B.2. *Given Assumption 1,*

(i) $T^{-1}\sum_{t=1}^{T} \boldsymbol{z}_{2t}\boldsymbol{z}_{2t}' \xrightarrow{\mathbb{P}} \mathbb{E}[\boldsymbol{z}_{2t}\boldsymbol{z}_{2t}']$;

(ii) $T^{-1}\sum_{t=1}^{T} \boldsymbol{z}_{2t} \xrightarrow{\mathbb{P}} [\delta_* \boldsymbol{\iota}_{p-1}', \boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*']'$;

(iii) $T^{-2}\sum_{t=1}^{T} (t-1)\boldsymbol{z}_{2t} \xrightarrow{\mathbb{P}} [\frac{1}{2}\delta_* \boldsymbol{\iota}_{p-1}', \frac{1}{2}\boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*']'$;

(iv) $T^{-2}\sum_{t=1}^{T} \boldsymbol{x}_{t-1}\boldsymbol{z}_{2t}' \xrightarrow{\mathbb{P}} [\frac{1}{2}\delta_* \boldsymbol{\mu}_* \boldsymbol{\iota}_{p-1}', \frac{1}{2}\boldsymbol{\mu}_* \boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*']'$;

(v) $T^{-3/2}\sum_{t=1}^{T}\widehat{\mathbf{m}}_{t-1}\boldsymbol{z}_{2t}' \Rightarrow [\delta_* \int \bar{\boldsymbol{\mathscr{B}}}_m \boldsymbol{\iota}_{p-1}', \int \bar{\boldsymbol{\mathscr{B}}}_m \boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*']'$;

(vi) $T^{-2}\sum_{t=1}^{T} y_{t-1}\boldsymbol{z}_{2t}' \xrightarrow{\mathbb{P}} [\frac{1}{2}\delta_*^2\boldsymbol{\iota}_{p-1}', \frac{1}{2}\delta_*\boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*']'$;

(vii) $T^{-1}\sum_{t=1}^{T}\widehat{u}_{t-1}\boldsymbol{z}_{2t} \xrightarrow{\mathbb{P}} \mathbb{E}[u_{t-1}\boldsymbol{z}_{2t}]$. □

Lemma B.3. *Given Assumption 1,*

(i) $T^{-2}\sum_{t=1}^{T} t \to \frac{1}{2}$;

(ii) $T^{-2}\sum_{t=1}^{T}\boldsymbol{x}_t \xrightarrow{\mathbb{P}} \frac{1}{2}\boldsymbol{\mu}_*$;

(iii) $T^{-3/2}\sum_{t=1}^{T}\widehat{\mathbf{m}}_t \Rightarrow \int \bar{\boldsymbol{\mathscr{B}}}_m$;

(iv) $T^{-2}\sum_{t=1}^{T} y_t \xrightarrow{\mathbb{P}} \frac{1}{2}\delta_*$;

(v) $\sum_{t=1}^{T}\widehat{u}_{t-1} \equiv 0$. □

Lemma B.4. *Given Assumption 1,*

(i) $T^{-3}\sum_{t=1}^{T} t^2 \to \frac{1}{3}$;

(ii) $T^{-3}\sum_{t=1}^{T} t\boldsymbol{x}_t \xrightarrow{\mathbb{P}} \frac{1}{3}\boldsymbol{\mu}_*$;

(iii) $\sum_{t=1}^{T}(t-1)\widehat{\mathbf{m}}_{t-1} \equiv \mathbf{0}$;

(iv) $T^{-3}\sum_{t=1}^{T} t y_t \xrightarrow{\mathbb{P}} \frac{1}{3}\delta_*$;

(v) $\sum_{t=1}^{T}(t-1)\widehat{u}_{t-1} \equiv 0$. □

Lemma B.5. *Given Assumption 1,*

(i) $T^{-3}\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t' \xrightarrow{\mathbb{P}} \frac{1}{3}\boldsymbol{\mu}_*\boldsymbol{\mu}_*'$;

(ii) $T^{-2}\sum_{t=1}^{T}\boldsymbol{x}_{t-1}\widehat{\mathbf{m}}_{t-1}' \Rightarrow \int \boldsymbol{\mathscr{B}}_m\bar{\boldsymbol{\mathscr{B}}}_m'$;

(iii) $T^{-3}\sum_{t=1}^{T}\boldsymbol{x}_t y_t \xrightarrow{\mathbb{P}} \frac{1}{3}\delta_*\boldsymbol{\mu}_*$;

(iv) $\sum_{t=1}^{T}\boldsymbol{x}_{t-1}\widehat{u}_{t-1} \equiv \mathbf{0}$. □

Remark. $\int \boldsymbol{\mathscr{B}}_m\bar{\boldsymbol{\mathscr{B}}}_m' = \int \bar{\boldsymbol{\mathscr{B}}}_m\bar{\boldsymbol{\mathscr{B}}}_m' = \int \boldsymbol{\mathscr{B}}_m\boldsymbol{\mathscr{B}}_m' - 3\int r\boldsymbol{\mathscr{B}}_m \int r\boldsymbol{\mathscr{B}}_m'$ from the definition of $\bar{\boldsymbol{\mathscr{B}}}_m(s) = \boldsymbol{\mathscr{B}}_m(s) - 3s\int r\boldsymbol{\mathscr{B}}_m$. □

Lemma B.6. *Given Assumption 1,*

(i) $T^{-2}\sum_{t=1}^{T}\widehat{\mathbf{m}}_{t-1}\widehat{\mathbf{m}}_{t-1}' \Rightarrow \int \boldsymbol{\mathscr{B}}_m\bar{\boldsymbol{\mathscr{B}}}_m'$;

(ii) $\sum_{t=1}^{T}\widehat{\mathbf{m}}_{t-1}\widehat{u}_{t-1} \equiv \mathbf{0}$;

(iii) $T^{-3}\sum_{t=1}^{T} y_t^2 \xrightarrow{\mathbb{P}} \frac{1}{3}\delta_*^2$;

(iv) $\widehat{\sigma}_{u,T}^2 := T^{-1}\sum_{t=1}^{T}\widehat{u}_t^2 \xrightarrow{\mathbb{P}} \sigma_u^2 := \mathbb{E}[u_t^2]$. □

Lemma B.7. *Let* $\boldsymbol{\varrho}_{m*} := \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{t-1}\mathbb{E}[\boldsymbol{s}_\tau^+ u_t]$. *Given Assumption 1,*

(i) $\sqrt{T}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) \Rightarrow 3\int r\boldsymbol{\mathscr{B}}_m$;

(ii) *for every t,* $\widetilde{\boldsymbol{r}}_t'\widehat{\boldsymbol{\upsilon}}_T = \dot{\boldsymbol{r}}_t'\widetilde{\boldsymbol{\upsilon}}_T$;

(iii) $T^{-1}\sum_{t=1}^{T}\mathbf{m}_{t-1}u_t \Rightarrow \int \boldsymbol{\mathscr{B}}_m d\mathscr{B}_u + \boldsymbol{\varrho}_{m*}$;

(iv) $T^{-1}\sum_{t=1}^{T}\mathbf{m}_{t-1}e_t \Rightarrow \int \boldsymbol{\mathscr{B}}_m d\mathscr{B}_e$;

(v) $\dot{\sigma}_{e,T}^2 := T^{-1}\sum_{t=1}^{T}(\Delta y_t - \dot{\boldsymbol{z}}_t'\dot{\boldsymbol{\tau}}_T)^2 \xrightarrow{\mathbb{P}} \sigma_e^2 := \mathbb{E}[e_t^2]$. □

We now prove the preliminary Lemmas B.1 to B.7.

Proof of Lemma B.1. This trivially follows theorem 7.30 of White (2001). ■

Proof of Lemma B.2. (*i*) It holds by the ergodic theorem.

(*ii*) It holds by the ergodic theorem and the fact that $\mathbb{E}[\boldsymbol{z}_{2t}] = [\delta_* \boldsymbol{\iota}'_{p-1}, \boldsymbol{\iota}'_q \otimes \boldsymbol{\mu}'_*]'$ because $\delta_* \boldsymbol{\iota}_{p-1} = \mathbb{E}[\Delta \boldsymbol{y}_{t-}]$, $\boldsymbol{\iota}'_q \otimes \boldsymbol{\mu}^{+\prime}_* = \mathbb{E}[(\Delta \boldsymbol{x}^{+\prime}_t, \ldots, \Delta \boldsymbol{x}^{+\prime}_{t-q+1})]$, and $\boldsymbol{\iota}'_q \otimes \boldsymbol{\mu}^{-\prime}_* = \mathbb{E}[(\Delta \boldsymbol{x}^{-\prime}_t, \ldots, \Delta \boldsymbol{x}^{-\prime}_{t-q+1})]$.

(*iii*) We note that $T^{-2} \sum_{t=1}^T (t-1)\boldsymbol{z}_{2t} = T^{-1} \sum_{t=1}^T ((t-1)/T)\boldsymbol{z}_{2t}$, and we let $\boldsymbol{s}_t := ((t-1)/T)\boldsymbol{z}_{2t}$ for notational simplicity, which is a heterogeneous process. Therefore, $T^{-1}\sum_{t=1}^T (\boldsymbol{s}_t - \mathbb{E}[\boldsymbol{s}_t]) \xrightarrow{\mathbb{P}} \mathbf{0}$. In addition, $T^{-1} \sum_{t=1}^T \mathbb{E}[\boldsymbol{s}_t] = T^{-2} \sum_{t=1}^T (t-1)\mathbb{E}[\boldsymbol{z}_{2t}] \to \frac{1}{2}\mathbb{E}[\boldsymbol{z}_{2t}]$, implying that $T^{-1}\sum_{t=1}^T \boldsymbol{s}_t \xrightarrow{\mathbb{P}} \frac{1}{2}\mathbb{E}[\boldsymbol{z}_{2t}]$ by White (2001, theorem 3.47), given the DGP condition in Assumption 1. We further note that $\mathbb{E}[\boldsymbol{z}_{2t}] = [\delta_* \boldsymbol{\iota}'_{p-1}, \boldsymbol{\iota}'_q \otimes \boldsymbol{\mu}'_*]'$, leading to the desired result.

(*iv*) We first note that $\boldsymbol{x}_{t-1} = \boldsymbol{\mu}_*(t-1) + \mathbf{m}_{t-1}$. Therefore, $\sum_{t=1}^T \boldsymbol{x}_{t-1}\boldsymbol{z}'_{2t} = \sum_{t=1}^T \boldsymbol{\mu}_*(t-1)\boldsymbol{z}'_{2t} + \sum_{t=1}^T \mathbf{m}_{t-1}\boldsymbol{z}'_{2t}$. The proof of Lemma B.2 (*iii*) already shows that $T^{-2}\sum_{t=1}^T (t-1)\boldsymbol{z}_{2t} \xrightarrow{\mathbb{P}} \frac{1}{2}\mathbb{E}[\boldsymbol{z}_{2t}]$. Furthermore, $\sum_{t=1}^T \mathbf{m}_{t-1}\boldsymbol{z}'_{2t} = O_{\mathbb{P}}(T^{3/2})$ as shown in the proof of Lemma B.2 (*v*). Therefore, it follows that $T^{-2}\sum_{t=1}^T \boldsymbol{x}_{t-1}\boldsymbol{z}'_{2t} \xrightarrow{\mathbb{P}} \frac{1}{2}\boldsymbol{\mu}_*\mathbb{E}[\boldsymbol{z}_{2t}]$, as desired.

(*v*) Note that $\widehat{\mathbf{m}}_{t-1} = \mathbf{m}_{t-1} - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)(t-1)$. Therefore, $\sum_{t=1}^T \widehat{\mathbf{m}}_{t-1}\boldsymbol{z}'_{2t} = \sum_{t=1}^T \mathbf{m}_{t-1}\boldsymbol{z}'_{2t} - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)\sum_{t=1}^T (t-1)\boldsymbol{z}'_{2t}$. We here note that $\sum_{t=1}^T \mathbf{m}_{t-1}\boldsymbol{z}'_{2t} = \sum_{t=1}^T \mathbf{m}_{t-1}\mathbb{E}[\boldsymbol{z}'_{2t}] + \sum_{t=1}^T \mathbf{m}_{t-1}(\boldsymbol{z}'_{2t} - \mathbb{E}[\boldsymbol{z}'_{2t}])$. The proof of Lemma A.3 (*iv*) implies that $T^{-3/2}\sum_{t=1}^T \mathbf{m}_{t-1} \Rightarrow \int \boldsymbol{\mathscr{B}}_m$, and $\sum_{t=1}^T \mathbf{m}_{t-1}(\boldsymbol{z}'_{2t} - \mathbb{E}[\boldsymbol{z}'_{2t}]) = o_{\mathbb{P}}(T^{3/2})$ by noting that $\mathbf{m}_{t-1} = O_{\mathbb{P}}(T^{1/2})$ and $\sum_{t=1}^T (\boldsymbol{z}'_{2t} - \mathbb{E}[\boldsymbol{z}'_{2t}]) = O_{\mathbb{P}}(T^{1/2})$. Therefore, $T^{-3/2}\sum_{t=1}^T \mathbf{m}_{t-1}\boldsymbol{z}'_{2t} \Rightarrow \int \boldsymbol{\mathscr{B}}_m \mathbb{E}[\boldsymbol{z}'_{2t}]$. Next, $(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)\sum_{t=1}^T (t-1)\boldsymbol{z}'_{2t} = \sqrt{T}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)T^{-2}\sum_{t=1}^T (t-1)\boldsymbol{z}'_{2t}$, and Lemma B.7 (*i*) implies that $\sqrt{T}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) \Rightarrow 3\int r\boldsymbol{\mathscr{B}}_m$. In addition to this, Lemma B.2 (*iii*) shows that $T^{-2}\sum_{t=1}^T (t-1)\boldsymbol{z}'_{2t} \xrightarrow{\mathbb{P}} \frac{1}{2}\mathbb{E}[\boldsymbol{z}'_{2t}]$. Therefore, $T^{-3/2}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)\sum_{t=1}^T (t-1)\boldsymbol{z}'_{2t} \Rightarrow \frac{3}{2}\int r\boldsymbol{\mathscr{B}}_m$. Hence, if we combine all these, it follows that $\sum_{t=1}^T \widehat{\mathbf{m}}_{t-1}\boldsymbol{z}'_{2t} \Rightarrow (\int \boldsymbol{\mathscr{B}}_m - \frac{3}{2}\int r\boldsymbol{\mathscr{B}}_m)\mathbb{E}[\boldsymbol{z}'_{2t}] = \int \bar{\boldsymbol{\mathscr{B}}}_m \mathbb{E}[\boldsymbol{z}'_{2t}]$.

(*vi*) Note that $y_{t-1} = \delta_*(t-1) + \sum_{j=1}^{t-1} d_j$. Therefore, $\sum_{t=1}^T y_{t-1}\boldsymbol{z}'_{2t} = \sum_{t=1}^T \delta_*(t-1)\boldsymbol{z}'_{2t} + \sum_{t=1}^T (\sum_{j=1}^{t-1} d_j)\boldsymbol{z}'_{2t}$. The proof of Lemma B.2 (*iii*) already showed that $T^{-2}\sum_{t=1}^T (t-1)\boldsymbol{z}_{2t} \xrightarrow{\mathbb{P}} \frac{1}{2}\mathbb{E}[\boldsymbol{z}_{2t}]$. Furthermore, we note that $\sum_{t=1}^T (\sum_{j=1}^{t-1} d_j)\boldsymbol{z}'_{2t} = O_{\mathbb{P}}(T^{3/2})$. Therefore, $T^{-2}\sum_{t=1}^T y_{t-1}\boldsymbol{z}'_{2t} \xrightarrow{\mathbb{P}} \frac{1}{2}\delta_*\mathbb{E}[\boldsymbol{z}_{2t}]$, as desired.

(*vii*) In the proof of Lemma B.6 (*iv*), we show that $\widehat{u}_t = u_t + O_{\mathbb{P}}(T^{-1/2})$. Therefore, it follows that $T^{-1}\sum_{t=1}^T \widehat{u}_{t-1}\boldsymbol{z}_{2t} = T^{-1}\sum_{t=1}^T u_{t-1}\boldsymbol{z}_{2t} + o_{\mathbb{P}}(1) \xrightarrow{\mathbb{P}} \mathbb{E}[u_{t-1}\boldsymbol{z}_{2t}]$ by the ergodic theorem. This completes the proof. ■

Proof of Lemma B.3. (*i*) $\sum_{t=1}^T t = T(T+1)/2$, leading to the desired result.

(*ii*) Note that $\boldsymbol{x}_t = \boldsymbol{\mu}_* t + \mathbf{m}_t$, so that $\sum_{t=1}^T \boldsymbol{x}_t = \boldsymbol{\mu}_*\sum_{t=1}^T t + \sum_{t=1}^T \mathbf{m}_t$. Here, $T^{-2}\sum_{t=1}^T t \to \frac{1}{2}$

by Lemma B.3 (*i*), and Lemma B.1 implies that $T^{-3/2}\sum_{t=1}^{T}\mathbf{m}_t = T^{-1}\sum_{t=1}^{T}\frac{1}{\sqrt{T}}\mathbf{m}_t = \int \mathbf{B}_m(r)dr \Rightarrow \int \boldsymbol{\mathscr{B}}_m$. Therefore, $T^{-3/2}\sum_{t=1}^{T}\boldsymbol{x}_t = \boldsymbol{\mu}_* T^{-3/2}\sum_{t=1}^{T} t + o_{\mathbb{P}}(1) \xrightarrow{\mathbb{P}} \frac{1}{2}\boldsymbol{\mu}_*^{+}$.

(*iii*) Note that $\widehat{\mathbf{m}}_t = \mathbf{m}_t + (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)t$. Thus, $T^{-3/2}\sum_{t=1}^{T}\widehat{\mathbf{m}}_t = T^{-3/2}\sum_{t=1}^{T}\mathbf{m}_t + \sqrt{T}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)T^{-2}\sum_{t=1}^{T} t$. Lemma B.7 (*i*) implies that $\sqrt{T}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) \Rightarrow 3\int r\boldsymbol{\mathscr{B}}_m$. Lemma B.3 (*i*) implies that $T^{-2}\sum_{t=1}^{T} t \to \frac{1}{2}$. In addition to these, $T^{-3/2}\sum_{t=1}^{T}\mathbf{m}_t = T^{-1}\sum_{t=1}^{T}\frac{1}{\sqrt{T}}\mathbf{m}_t = \int \mathbf{B}_m(r)dr \Rightarrow \int \boldsymbol{\mathscr{B}}_m$ by Lemma B.1. Thus, $T^{-3/2}\sum_{t=1}^{T}\widehat{\mathbf{m}}_t \Rightarrow \int \boldsymbol{\mathscr{B}}_m(r) - 3\int rdr\int s\boldsymbol{\mathscr{B}}_m(s)ds = \int \bar{\boldsymbol{\mathscr{B}}}_m$ by the definition of $\bar{\boldsymbol{\mathscr{B}}}_m(\cdot)$, viz., $\bar{\boldsymbol{\mathscr{B}}}_m(\cdot) = \boldsymbol{\mathscr{B}}_m(\cdot) - 3(\cdot)\int s\boldsymbol{\mathscr{B}}_m(s)ds$.

(*iv*) From (5), $T^{-2}\sum_{t=1}^{T} y_t = \delta_* T^{-2}\sum_{t=1}^{T} t + T^{-2}\sum_{t=1}^{T}\sum_{j=1}^{t} d_j = \frac{1}{2}\delta_* + o_{\mathbb{P}}(1)$.

(*v*) We note that $\widehat{u}_{t-1} := y_{t-1} - \widetilde{\boldsymbol{r}}_{t-1}'\widehat{\boldsymbol{v}}_T$, so that $\sum_{t=1}^{T}\widehat{u}_{t-1}\widetilde{\boldsymbol{r}}_{t-1} \equiv \mathbf{0}$. We here note that $\widetilde{\boldsymbol{r}}_{t-1} := [\boldsymbol{x}_{t-1}', (t-1), 1]'$. This completes the proof. ■

Proof of Lemma B.4. (*i*) $\sum_{t=1}^{T} t^2 = T(T+1)(2T+1)/6$, leading to the desired result.

(*ii*) $T^{-3}\sum_{t=1}^{T} t\boldsymbol{x}_t = \boldsymbol{\mu}_* T^{-3}\sum_{t=1}^{T} t^2 + T^{-3}\sum_{t=1}^{T} t\mathbf{m}_t$. Here, it follows that $T^{-3}\sum_{t=1}^{T} t^2 \to \frac{1}{3}$ and $T^{-5/2}\sum_{t=1}^{T} t\mathbf{m}_t = T^{-1}\sum_{t=1}^{T}(\frac{t}{T})\frac{1}{\sqrt{T}}\mathbf{m}_t = \int r\mathbf{B}_{mT}(r)dr \Rightarrow \int r\boldsymbol{\mathscr{B}}_m$. Therefore, $T^{-3}\sum_{t=1}^{T} t\boldsymbol{x}_t = \frac{1}{3}\boldsymbol{\mu}_* + o_{\mathbb{P}}(1)$.

(*iii*) Note that $\widehat{\mathbf{m}}_t = \boldsymbol{x}_t - t\widehat{\boldsymbol{\mu}}_T$ and $\widehat{\boldsymbol{\mu}}_T = (\sum_{t=1}^{T-1} t^2)^{-1}\sum_{t=1}^{T-1} t\boldsymbol{x}_t$. Therefore, $\sum_{t=1}^{T}(t-1)\widehat{\mathbf{m}}_{t-1} \equiv \mathbf{0}$.

(*iv*) From (5), $T^{-3}\sum_{t=1}^{T} ty_t = \delta_* T^{-3}\sum_{t=1}^{T} t^2 + T^{-3}\sum_{t=1}^{T} t\sum_{j=1}^{t} d_j = \frac{1}{3}\delta_* + o_{\mathbb{P}}(1)$.

(*v*) The proof of Lemma B.3 (*v*) already shows the given claim. ■

Proof of Lemma B.5. (*i*) Using the fact that $\boldsymbol{x}_t = \boldsymbol{\mu}_* t + \mathbf{m}_t$, $\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t' = \sum_{t=1}^{T}(\boldsymbol{\mu}_* t + \mathbf{m}_t)(\boldsymbol{\mu}_* t + \mathbf{m}_t)' = \sum_{t=1}^{T}\boldsymbol{\mu}_*\boldsymbol{\mu}_*' t^2 + \sum_{t=1}^{T}\boldsymbol{\mu}_*\mathbf{m}_t' t + \sum_{t=1}^{T}\mathbf{m}_t\boldsymbol{\mu}_*' t + \sum_{t=1}^{T}\mathbf{m}_t\mathbf{m}_t'$. Here, $T^{-3}\sum_{t=1}^{T}\boldsymbol{\mu}_*\boldsymbol{\mu}_*' t^2 \to \frac{1}{3}\boldsymbol{\mu}_*\boldsymbol{\mu}_*'$ by Lemma B.4 (*i*), and $T^{-5/2}\sum_{t=1}^{T}\boldsymbol{\mu}_*\mathbf{m}_t' t \Rightarrow \boldsymbol{\mu}_*\int r\boldsymbol{\mathscr{B}}_m$ as shown in the proof of Lemma B.4 (*ii*). Furthermore, $T^{-2}\sum_{t=1}^{T}\mathbf{m}_t\mathbf{m}_t' = T^{-1}\sum_{t=1}^{T}\frac{1}{\sqrt{T}}\mathbf{m}_t\frac{1}{\sqrt{T}}\mathbf{m}_t' = \int \mathbf{B}_{mT}(r)\mathbf{B}_{mT}(r)'dr \Rightarrow \int \boldsymbol{\mathscr{B}}_n\boldsymbol{\mathscr{B}}_m'$. Thus, $T^{-3}\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t' \xrightarrow{\mathbb{P}} \frac{1}{3}\boldsymbol{\mu}_*\boldsymbol{\mu}_*'$.

(*ii*) We first note that $\sum_{t=1}^{T}\boldsymbol{x}_{t-1}\widehat{\mathbf{m}}_{t-1}' = \sum_{t=1}^{T}(\boldsymbol{\mu}_*(t-1) + \mathbf{m}_{t-1})\widehat{\mathbf{m}}_{t-1}' = \boldsymbol{\mu}_*\sum_{t=1}^{T}(t-1)\widehat{\mathbf{m}}_{t-1}' + \sum_{t=1}^{T}\mathbf{m}_{t-1}\widehat{\mathbf{m}}_{t-1}' = \sum_{t=1}^{T}\mathbf{m}_{t-1}\widehat{\mathbf{m}}_{t-1}'$ by Lemma B.4 (*iii*). We further note that $\sum_{t=1}^{T}\mathbf{m}_{t-1}\widehat{\mathbf{m}}_{t-1}' = \sum_{t=1}^{T}\mathbf{m}_{t-1}(\mathbf{m}_{t-1} - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)(t-1))'$ using the fact that $\widehat{\mathbf{m}}_t = \mathbf{m}_t - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)t$. Here, we note that $T^{-2}\sum_{t=1}^{T}\mathbf{m}_t\mathbf{m}_t' \Rightarrow \int \boldsymbol{\mathscr{B}}_m\boldsymbol{\mathscr{B}}_m'$ as shown in the proof of Lemma B.5 (*i*), and $T^{-2}\sum_{t=1}^{T} t\mathbf{m}_t(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)' = T^{-5/2}\sum_{t=1}^{T} t\mathbf{m}_t\sqrt{T}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)' \Rightarrow 3\int r\boldsymbol{\mathscr{B}}_m\int r\boldsymbol{\mathscr{B}}_m$ by Lemma B.7 (*ii*) and the fact that $T^{-5/2}\sum_{t=1}^{T} t\mathbf{m}_t = T^{-1}\sum_{t=1}^{T}\frac{t}{T}\frac{1}{\sqrt{T}}\mathbf{m}_t = \int r\mathbf{B}_m(r)dr \Rightarrow \int r\boldsymbol{\mathscr{B}}_m$ as shown in the proof of Lemma B.4 (*ii*). Therefore, $T^{-2}\sum_{t=1}^{T}\boldsymbol{x}_{t-1}\widehat{\mathbf{m}}_{t-1}' \Rightarrow \int \boldsymbol{\mathscr{B}}_m(\boldsymbol{\mathscr{B}}_m - 3r\int s\boldsymbol{\mathscr{B}}_m)' = \int \boldsymbol{\mathscr{B}}_m\bar{\boldsymbol{\mathscr{B}}}_m'$. We further note that $\int \boldsymbol{\mathscr{B}}_m\bar{\boldsymbol{\mathscr{B}}}_m' = \int \bar{\boldsymbol{\mathscr{B}}}_m\bar{\boldsymbol{\mathscr{B}}}_m'$ by the definition of $\boldsymbol{\mathscr{B}}_m\bar{\boldsymbol{\mathscr{B}}}_m$ and the fact that $\int r^2 = \frac{1}{3}$.

(*iii*) Using the fact that $\boldsymbol{x}_t = \boldsymbol{\mu}_* t + \mathbf{m}_t$ and $y_t = \delta_* t + \sum_{j=1}^{t} d_j$, $\sum_{t=1}^{T}\boldsymbol{x}_t y_t = \boldsymbol{\mu}_*\delta_*\sum_{t=1}^{T} t^2 + \delta_*\sum_{t=1}^{T} t\mathbf{m}_t + \boldsymbol{\mu}_*\sum_{t=1}^{T} t\sum_{j=1}^{t} d_j + \sum_{t=1}^{T}\mathbf{m}_t\sum_{j=1}^{T} d_j$. Here, $T^{-3}\sum_{t=1}^{T} t^2 \to \frac{1}{3}$ by Lemma B.4 (*i*), and $T^{-5/2}\sum_{t=1}^{T} t\mathbf{m}_t \Rightarrow \int r\boldsymbol{\mathscr{B}}_m$ as shown in the proof of Lemma B.4 (*ii*). Furthermore, we note that $\sum_{t=1}^{T} t\sum_{j=1}^{t} d_j = O_{\mathbb{P}}(T^{5/2})$ and $\sum_{t=1}^{T}\mathbf{m}_t\sum_{j=1}^{T} d_j = O_{\mathbb{P}}(T^2)$. Therefore, $T^{-3}\sum_{t=1}^{T}\boldsymbol{x}_t y_t \xrightarrow{\mathbb{P}} \frac{1}{3}\delta_*\boldsymbol{\mu}_*$.

(*iv*) The proof of Lemma B.3 (*v*) already shows the given claim. ■

Proof of Lemma B.6. (*i*) We first note that $\sum_{t=1}^{T}\widehat{\mathbf{m}}_{t-1}\widehat{\mathbf{m}}_{t-1}' = \sum_{t=1}^{T}(\mathbf{m}_{t-1}-(\widehat{\boldsymbol{\mu}}_T-\boldsymbol{\mu}_*)(t-1))\widehat{\mathbf{m}}_{t-1}' = \sum_{t=1}^{T}\mathbf{m}_{t-1}\widehat{\mathbf{m}}_{t-1}' - (\widehat{\boldsymbol{\mu}}_T-\boldsymbol{\mu}_*)\sum_{t=1}^{T}(t-1)\widehat{\mathbf{m}}_{t-1}' = \sum_{t=1}^{T}\mathbf{m}_{t-1}\widehat{\mathbf{m}}_{t-1}'$ by Lemma B.4 (*iii*). We next note that $\sum_{t=1}^{T}\mathbf{m}_{t-1}\widehat{\mathbf{m}}_{t-1}' = \sum_{t=1}^{T}\mathbf{m}_{t-1}(\mathbf{m}_{t-1}-(\widehat{\boldsymbol{\mu}}_T-\boldsymbol{\mu}_*)(t-1))'$ using the fact that $\widehat{\mathbf{m}}_t = \mathbf{m}_t-(\widehat{\boldsymbol{\mu}}_T-\boldsymbol{\mu}_*)t$. Here, $T^{-2}\sum_{t=1}^{T}\mathbf{m}_t\mathbf{m}_t' \Rightarrow \int\boldsymbol{\mathscr{B}}_m\boldsymbol{\mathscr{B}}_m'$ as shown in the proof of Lemma B.5 (*i*), and $T^{-2}\sum_{t=1}^{T} t\mathbf{m}_t(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)' \Rightarrow 3\int r\boldsymbol{\mathscr{B}}_m\int r\boldsymbol{\mathscr{B}}_m$ as shown in the proof of Lemma B.5 (*ii*). Therefore, $T^{-2}\sum_{t=1}^{T}\widehat{\mathbf{m}}_{t-1}\widehat{\mathbf{m}}_{t-1}' \Rightarrow \int\boldsymbol{\mathscr{B}}_m(\boldsymbol{\mathscr{B}}_m-3r\int s\boldsymbol{\mathscr{B}}_m)' = \int\boldsymbol{\mathscr{B}}_m\bar{\boldsymbol{\mathscr{B}}}_m'$.

(*ii*) As Lemma B.7 (*iii*) shows, for each t, $\widetilde{\boldsymbol{r}}_t'\widehat{\boldsymbol{\upsilon}}_T = \dot{\boldsymbol{r}}_t'\widetilde{\boldsymbol{\upsilon}}_T$, so that $\widehat{u}_t := y_t - \widetilde{\boldsymbol{r}}_t'\widehat{\boldsymbol{\upsilon}}_T = y_t - \dot{\boldsymbol{r}}_t'\widetilde{\boldsymbol{\upsilon}}_T$. We further note that $\widetilde{\boldsymbol{\upsilon}}_T := (\sum_{t=1}^{T}\dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}')^{-1}\sum_{t=1}^{T}\dot{\boldsymbol{r}}_{t-1}y_{t-1}$, so that $\sum_{t=1}^{T}\dot{\boldsymbol{r}}_{t-1}\widehat{u}_{t-1} = \mathbf{0}$. We now note that $\dot{\boldsymbol{r}}_{t-1} := [\widehat{\mathbf{m}}_{t-1}', (t-1), 1]'$, leading to that $\sum_{t=1}^{T}\widehat{\mathbf{m}}_{t-1}\widehat{u}_{t-1} = \mathbf{0}$.

(*iii*) From (5), $y_t = \delta_* t + \sum_{j=1}^{t} d_j$. We here note that $\sum_{t=1}^{T} y_t^2 = \sum_{t=1}^{T}(\delta_* t + \sum_{j=1}^{t} d_j)^2 = \sum_{t=1}^{T}\delta_*^2 t^2 + 2\sum_{t=1}^{T}\delta_* t\sum_{j=1}^{t} d_j + \sum_{t=1}^{T}(\sum_{j=1}^{t} d_j)^2$. Furthermore, $T^{-3}\sum_{t=1}^{T} t^2 \to \frac{1}{3}$ by Lemma B.4 (*i*), $\sum_{t=1}^{T}\delta_* t\sum_{j=1}^{t} d_j = O_{\mathbb{P}}(T^{5/2})$, and $\sum_{t=1}^{T}(\sum_{j=1}^{t} d_j)^2 = O_{\mathbb{P}}(T^2)$, implying that $T^{-3}\sum_{t=1}^{T} y_t^2 \xrightarrow{\mathbb{P}} \frac{1}{3}\delta_*^2$.

(*iv*) Note that $\widehat{u}_t := y_t - \widetilde{\boldsymbol{r}}_t'\widehat{\boldsymbol{\upsilon}}_T$, and $\widetilde{\boldsymbol{r}}_t'\widehat{\boldsymbol{\upsilon}}_T = \boldsymbol{x}_t'\widehat{\boldsymbol{\beta}}_T + t\widehat{\zeta}_T + \widehat{\nu}_T$ with $\boldsymbol{x}_t = \widehat{\mathbf{m}}_t + \widehat{\boldsymbol{\mu}}_T t$. Furthermore, $y_t = \boldsymbol{\beta}_*'(\widehat{\mathbf{m}}_t + \widehat{\boldsymbol{\mu}}_T t) + \zeta_* t + \nu_* + u_t$ using (6). Hence, $\widehat{u}_t = u_t - (\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)'\widehat{\mathbf{m}}_t - (\widetilde{\vartheta}_T - \vartheta_{T*})t - (\widetilde{\nu}_T - \nu_*)$. We now note that Lemma 3, $\widehat{\mathbf{m}}_t = O_{\mathbb{P}}(T^{1/2})$, and $t = O(T)$ imply that $\widehat{u}_t = u_t + O_{\mathbb{P}}(T^{-1/2})$. Therefore, $T^{-1}\sum_{t=1}^{T}\widehat{u}_t^2 = T^{-1}\sum_{t=1}^{T} u_t^2 + o_{\mathbb{P}}(1)$, and $T^{-1}\sum_{t=1}^{T} u_t^2 \xrightarrow{\mathbb{P}} \mathbb{E}[u_t^2]$ by the ergodic theorem, implying that $T^{-1}\sum_{t=1}^{T}\widehat{u}_t^2 \xrightarrow{\mathbb{P}} \sigma_u^2 := \mathbb{E}[u_t^2]$. ■

Proof of Lemma B.7. (*i*) Note that $\widehat{\boldsymbol{\mu}}_T = (\sum_{t=1}^{T-1} t^2)^{-1}\sum_{t=1}^{T-1} t\boldsymbol{x}_t$ and $\boldsymbol{x}_t = \boldsymbol{\mu}_* t + \mathbf{m}_t$. Therefore, $\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_* = (\sum_{t=1}^{T-1} t^2)^{-1}\sum_{t=1}^{T-1} t\mathbf{m}_t$, so that $\sqrt{T}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) = (T^{-3}\sum_{t=1}^{T-1} t^2)^{-1} T^{-5/2}\sum_{t=1}^{T-1} t\mathbf{m}_t$. Lemma B.4 (*i*) implies that $T^{-3}\sum_{t=1}^{T} t^2 \to \frac{1}{3}$. Furthermore, $T^{-5/2}\sum_{t=1}^{T-1} t\mathbf{m}_t = T^{-1}\sum_{t=1}^{T-1}\frac{t}{T}\frac{1}{\sqrt{T}}\mathbf{m}_t = \int r\mathbf{B}_{mT}(r)dr \Rightarrow \int r\boldsymbol{\mathscr{B}}_m$. Hence, $\sqrt{T}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) \Rightarrow 3\int r\boldsymbol{\mathscr{B}}_m$.

(*ii*) Note that $\widetilde{\boldsymbol{r}}_t'\widehat{\boldsymbol{\upsilon}}_T = \boldsymbol{x}_t'\widehat{\boldsymbol{\beta}}_T + t\widehat{\zeta}_T + \widehat{\nu}_T$ with $\boldsymbol{x}_t = \widehat{\mathbf{m}}_t + \widehat{\boldsymbol{\mu}}_T t$. Therefore, $\widetilde{\boldsymbol{r}}_t'\widehat{\boldsymbol{\upsilon}}_T = \widehat{\mathbf{m}}_t'\widehat{\boldsymbol{\beta}}_T + t(\widehat{\boldsymbol{\mu}}_T'\widehat{\boldsymbol{\beta}}_T + \widehat{\zeta}_T) + \widehat{\nu}_T = \widehat{\mathbf{m}}_t'\widetilde{\boldsymbol{\beta}}_T + t\widetilde{\vartheta}_T + \widetilde{\nu}_T = \dot{\boldsymbol{r}}_t'\widetilde{\boldsymbol{\upsilon}}_T$, where the second last equality holds by (10), and the last equality follows from the definition of $\dot{\boldsymbol{r}}_t$.

(*iii*) Note that $T^{-1}\sum_{t=1}^{T}\mathbf{m}_{t-1}u_t = T^{-1}\sum_{t=1}^{T} T^{-1/2}\mathbf{m}_{t-1}\sqrt{T}(B_{uT}(t/T) - B_{uT}((t-1)/T)) = \int\mathbf{B}_{mT}(r)dB_{uT}(r)$. We here note that $\int\mathbf{B}_{mT}(r)dB_{uT}(r) \Rightarrow \int\boldsymbol{\mathscr{B}}_m d\mathscr{B}_u + \boldsymbol{\varrho}_{m*}$ by applying theorem 4 of de Jong and Davidson (2000).

(*iv*) We first note that $T^{-1}\sum_{t=1}^{T}\mathbf{m}_{t-1}e_t = T^{-1}\sum_{t=1}^{T} T^{-1/2}\mathbf{m}_{t-1}\sqrt{T}(B_{eT}(t/T) - B_{eT}((t-1)/T)) = \int\mathbf{B}_{mT}(r)dB_{eT}(r)$. Note that $\int\mathbf{B}_{mT}(r)dB_{eT}(r) \Rightarrow \int\boldsymbol{\mathscr{B}}_m d\mathscr{B}_e$ by applying theorem 4 of de Jong and

Davidson (2000) and noting that $\mathbb{E}[\boldsymbol{s}_\tau e_t] = \mathbf{0}$ for each $\tau < t$.

(v) Note that if we let $\dot{e}_t := \Delta y_t - \dot{\boldsymbol{z}}_t \dot{\boldsymbol{\tau}}_T$, it follows that $\dot{e}_t = -\dot{\boldsymbol{z}}_t'(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) + e_t$ from the fact that $\Delta y_t = \dot{\boldsymbol{z}}_t' \boldsymbol{\tau}_{T*} + e_t$, implying that

$$\sum_{t=1}^T \dot{e}_t^2 = (\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*})'\dot{\mathbf{D}} \left(\dot{\mathbf{D}}^{-1} \sum_{t=1}^T \dot{\boldsymbol{z}}_t \dot{\boldsymbol{z}}_t' \dot{\mathbf{D}}^{-1} \right) \dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) - 2 \left(\sum_{t=1}^T e_t \dot{\boldsymbol{z}}_t \right) \dot{\mathbf{D}}^{-1}\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) + \sum_{t=1}^T e_t^2.$$

We examine the asymptotic behavior of each element on the right-hand side. First, Lemmas 3 (vi) and 4 (i) imply that $\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) = O_{\mathbb{P}}(1)$ and $\dot{\mathbf{D}}^{-1} \sum_{t=1}^T \dot{\boldsymbol{z}}_t \dot{\boldsymbol{z}}_t' \dot{\mathbf{D}}^{-1} = O_{\mathbb{P}}(1)$. Second, from the definitions of $\dot{\boldsymbol{z}}_t$ and $\dot{\mathbf{D}}$, $\sum_{t=1}^T e_t \dot{\boldsymbol{z}}_t \dot{\mathbf{D}}^{-1} = [T^{-1/2} \sum_{t=1}^T e_t \widehat{u}_{t-1}, T^{-1} \sum_{t=1}^T e_t \widehat{\mathbf{m}}_{t-1}', T^{-3/2} \sum_{t=1}^T e_t (t-1), T^{-1/2} \sum_{t=1}^T e_t,$ $T^{-1/2} \sum_{t=1}^T e_t \boldsymbol{z}_{2t}']'$. We verify that each element on the right-hand side is $O_{\mathbb{P}}(1)$. We first note that $T^{-3/2} \sum_{t=1}^T e_t(t-1) = O_{\mathbb{P}}(1)$, $T^{-1/2} \sum_{t=1}^T e_t = O_{\mathbb{P}}(1)$, and $T^{-1/2} \sum_{t=1}^T e_t \boldsymbol{z}_{2t} = O_{\mathbb{P}}(1)$ by the martingale difference CLT based upon the fact that $\{e_t, \mathscr{F}_t\}$ is an MDA. In addition, $T^{-1} \sum_{t=1}^T e_t \widehat{\mathbf{m}}_{t-1} = T^{-1} \sum_{t=1}^T e_t \mathbf{m}_{t-1} - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) T^{-1} \sum_{t=1}^T e_t(t-1)$ by noting that $\widehat{\mathbf{m}}_{t-1} = \mathbf{m}_{t-1} - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)(t-1)$ as given in the proof of Lemma B.6(i). Here, Lemmas B.7 (i and iv) imply that $(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) = O_{\mathbb{P}}(T^{-1/2})$ and $T^{-1} \sum_{t=1}^T e_t \mathbf{m}_{t-1} = O_{\mathbb{P}}(1)$, respectively, so that $T^{-1} \sum_{t=1}^T e_t \widehat{\mathbf{m}}_{t-1} = O_{\mathbb{P}}(1)$. Finally, $T^{-1/2} \sum_{t=1}^T e_t \widehat{u}_{t-1}$ $= T^{-1/2} \sum_{t=1}^T e_t u_{t-1} + o_{\mathbb{P}}(1)$ using the fact that $\widehat{u}_{t-1} = u_{t-1} + O_{\mathbb{P}}(T^{-1/2})$ as given in the proof of Lemma B.6 (iv). All these facts imply that $\sum_{t=1}^T e_t \dot{\boldsymbol{z}}_t \dot{\mathbf{D}}^{-1} = O_{\mathbb{P}}(1)$. By these two facts, it follows that $\sum_{t=1}^T \dot{e}_t^2 = \sum_{t=1}^T e_t^2 + O_{\mathbb{P}}(1)$, implying that $\dot{\sigma}_{e,T}^2 := T^{-1} \sum_{t=1}^T \dot{e}_t^2 = T^{-1} \sum_{t=1}^T e_t^2 + O_{\mathbb{P}}(T^{-1})$. The desired result follows from the ergodic theorem, and this completes the proof. ∎

A.3 Proofs

Proof of Lemma 1. (i) This follows from Lemmas B.3 (i, ii, iv), B.4 (i, ii, iv), B.5 (i, iii), B.6 (iii), and the remarks below Lemmas B.3, B.4, and B.5.

(ii) This follows from Lemma B.2 (i).

(iii) This follows from Lemmas B.2 (ii, iii, iv, vi) and the remark below Lemma B.2.

(iv) This follows from Lemmas 1 (i, ii, iii) and the structure of $\mathbf{M}$. ∎

Proof of Lemma 2. (i) This follows from Lemmas B.3 (v), B.4 (v), and B.6 (ii).

(ii) This follows from Lemmas B.3 (i, ii, iii), B.4 (i, ii, iii), and B.6 (i).

(iii) This follows from Lemma B.2 (vii).

(iv) This follows from Lemmas B.2 (ii, iii, iv, and vi).

(v) This follows from Lemma B.2 (i).

(vi) This follows from Lemma A.3 (i) and Lemmas 2 (i, ii, iii, iv, v).

(*vii*) This follows from the definition of $\dot{\boldsymbol{z}}_{1t} := \dot{\boldsymbol{r}}_{t-1}$ and Lemma 2 (*ii*). ■

Proof of Lemma 3. (*i*) We note that $\dot{\mathbf{D}}_1(\widetilde{\boldsymbol{v}}_T - \bar{\boldsymbol{v}}_{T*}) = (\dot{\mathbf{D}}_1^{-1}\sum_{t=1}^T \dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}'\dot{\mathbf{D}}_1^{-1})^{-1}\dot{\mathbf{D}}_1^{-1}\sum_{t=1}^T \dot{\boldsymbol{r}}_{t-1}u_{t-1}$ from (17), and Lemma 2 (*vii*) implies that $\dot{\mathbf{D}}_1^{-1}(\sum_{t=1}^T \dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}')\dot{\mathbf{D}}_1^{-1} \Rightarrow \dot{\boldsymbol{\mathscr{M}}}_{11}$. We therefore focus on the limit distribution of $\dot{\mathbf{D}}_1^{-1}\sum_{t=1}^T \dot{\boldsymbol{r}}_{t-1}u_{t-1}$. Note that

$$\dot{\mathbf{D}}_1^{-1}\sum_{t=1}^T \dot{\boldsymbol{r}}_{t-1}u_{t-1} = \left[T^{-1}\sum_{t=1}^T \widehat{\mathbf{m}}_{t-1}'u_{t-1}, T^{-3/2}\sum_{t=1}^T (t-1)u_{t-1}, T^{-1/2}\sum_{t=1}^T u_{t-1}\right]'.$$

We now examine the asymptotic behavior of each element on the right-hand side. First, we note that $\widehat{\mathbf{m}}_{t-1} = \mathbf{m}_{t-1} - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)(t-1)$. Therefore, $T^{-1}\sum_{t=1}^T \widehat{\mathbf{m}}_{t-1}u_{t-1} = T^{-1}\sum_{t=1}^T \mathbf{m}_{t-2}u_{t-1} + T^{-1}\sum_{t=1}^T \boldsymbol{s}_{t-1}u_{t-1} - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)T^{-1}\sum_{t=1}^T (t-1)u_{t-1}$. We here note that $T^{-1}\sum_{t=1}^T \mathbf{m}_{t-2}u_{t-1} \Rightarrow \int \boldsymbol{\mathscr{B}}_m d\mathscr{B}_u + +\boldsymbol{\varrho}_{m*}$ by Lemma B.7 (*iii*), and $\sqrt{T}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) \Rightarrow 3\int r\boldsymbol{\mathscr{B}}_m$ by Lemma B.7 (*i*). In addition to this, $T^{-3/2}\sum_{t=1}^T (t-1)u_{t-1} = T^{-1}\sum_{t=1}^T \frac{(t-1)}{T}\sqrt{T}(B_{uT}((t-1)/T) - B_{uT}((t-2)/T)) = \int r dB_{uT}(r) \Rightarrow \int r d\mathscr{B}_u$. Hence, it follows that $T^{-1}\sum_{t=1}^T \widehat{\mathbf{m}}_{t-1}u_{t-1} \Rightarrow (\dot{\boldsymbol{\mathscr{S}}}_{11}', \dot{\boldsymbol{\mathscr{S}}}_{12}')' := \int \boldsymbol{\mathscr{B}}_m d\mathscr{B}_u + \boldsymbol{\varrho}_{m*} - 3\int r\boldsymbol{\mathscr{B}}_m \int r d\mathscr{B}_u$. Next, it is already showed that $T^{-3/2}\sum_{t=1}^T (t-1)u_{t-1} \Rightarrow \dot{\mathscr{S}}_{13} := \int r d\mathscr{B}_u$. Third, note that $T^{-1/2}\sum_{t=1}^T u_{t-1} = T^{-1}\sum_{t=1}^T \sqrt{T}(B_{uT}((t-1)/T) - B_{uT}((t-2)/T)) = \int r dB_{uT}(r) \Rightarrow \dot{\mathscr{S}}_{14} := \int d\mathscr{B}_u$. We now combine the first to third facts to obtain that $\dot{\mathbf{D}}_1^{-1}\sum_{t=1}^T \dot{\boldsymbol{r}}_{t-1}u_{t-1} \Rightarrow \dot{\boldsymbol{\mathscr{S}}}_1$, leading to that $\dot{\mathbf{D}}_1(\widetilde{\boldsymbol{v}}_T - \bar{\boldsymbol{v}}_{T*}) \Rightarrow \dot{\boldsymbol{\mathscr{M}}}_{11}^{-1}\dot{\boldsymbol{\mathscr{S}}}_1$, as desired.

(*ii*) We first note that $\widehat{\boldsymbol{v}}_T - \boldsymbol{v}_* = \mathbf{P}_T\widetilde{\boldsymbol{v}}_T - \mathbf{P}\bar{\boldsymbol{v}}_*$. Therefore, $\widehat{\boldsymbol{v}}_T - \boldsymbol{v}_* = (\mathbf{P}_T - \mathbf{P})(\widetilde{\boldsymbol{v}}_T - \bar{\boldsymbol{v}}_{T*}) + \mathbf{P}(\widetilde{\boldsymbol{v}}_T - \bar{\boldsymbol{v}}_{T*}) + (\mathbf{P}_T - \mathbf{P})(\bar{\boldsymbol{v}}_{T*} - \bar{\boldsymbol{v}}_*) + \mathbf{P}(\bar{\boldsymbol{v}}_{T*} - \bar{\boldsymbol{v}}_*) + (\mathbf{P}_T - \mathbf{P})\bar{\boldsymbol{v}}_*$. From the definition of $\mathbf{P}_T$ and Lemma 3 (*i*), we note that $(\mathbf{P}_T - \mathbf{P}) = O_{\mathbb{P}}(T^{-1/2})$, $(\widetilde{\boldsymbol{v}}_T - \bar{\boldsymbol{v}}_{T*}) = O_{\mathbb{P}}(\dot{\mathbf{D}}^{-1})$, and $(\mathbf{P}_T - \mathbf{P})(\bar{\boldsymbol{v}}_{T*} - \bar{\boldsymbol{v}}_*) = \mathbf{0}$. Furthermore, $\mathbf{P}(\bar{\boldsymbol{v}}_{T*} - \bar{\boldsymbol{v}}_*) = [\mathbf{0}', \mathbf{0}', (\vartheta_{T*} - \vartheta_*), 0]'$ such that $(\vartheta_{T*} - \vartheta_*) = \boldsymbol{\beta}_*'(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)$, and $(\mathbf{P}_T - \mathbf{P})\bar{\boldsymbol{v}}_* = [\mathbf{0}, \mathbf{0}, -\boldsymbol{\beta}_*'(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*), 0]'$, so that $\mathbf{P}(\bar{\boldsymbol{v}}_{T*} - \bar{\boldsymbol{v}}_*) + (\mathbf{P}_T - \mathbf{P})\bar{\boldsymbol{v}}_* = \mathbf{0}$. Hence, it now follows that $\widehat{\boldsymbol{v}}_T - \boldsymbol{v}_* = (\mathbf{P}_T - \mathbf{P})(\widetilde{\boldsymbol{v}}_T - \bar{\boldsymbol{v}}_{T*}) + \mathbf{P}(\widetilde{\boldsymbol{v}}_T - \bar{\boldsymbol{v}}_{T*}) = \mathbf{P}(\widetilde{\boldsymbol{v}}_T - \bar{\boldsymbol{v}}_{T*}) + O_{\mathbb{P}}(T^{-3/2})$, so that

$$\begin{aligned}\dot{\mathbf{D}}_\dagger(\widehat{\boldsymbol{v}}_T - \boldsymbol{v}_*) &= [\ T(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)' \ \vdots\ -\boldsymbol{\mu}_*'T(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*) \ \vdots\ \sqrt{T}(\widetilde{\nu}_T - \nu_*)\ ]' + o_{\mathbb{P}}(1)\\ &\Rightarrow [\ \dot{\mathscr{L}}_{11}' \quad \dot{\mathscr{L}}_{12}' \ \vdots\ -\boldsymbol{\mu}_*^{+\prime}\dot{\mathscr{L}}_{11}' - \boldsymbol{\mu}_*^{-\prime}\dot{\mathscr{L}}_{12}' \ \vdots\ \dot{\mathscr{L}}_{14}\ ]'\end{aligned}$$

by Lemma 3 (*i*). ■

Proof of Lemma 4. (*i*) We first note that $\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) = (\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\dot{\mathbf{D}}^{-1})^{-1}\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t e_t$ from (18), and Lemma 2 (*vi*) implies that $\dot{\mathbf{D}}^{-1}(\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t')\dot{\mathbf{D}}^{-1} \Rightarrow \dot{\boldsymbol{\mathscr{M}}}$. We therefore focus on the limit distribution of $\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t e_t$. We note that $\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t e_t = \dot{\mathbf{D}}^{-1}\sum_{t=1}^T [\widehat{u}_{t-1}e_r, \widehat{\mathbf{m}}_{t-1}'e_t, (t-1)e_t, e_t, \boldsymbol{z}_{2t}'e_t]'$.

We now investigate the asymptotic behavior of each element on the right-hand side. First, we already showed in the proof of Lemma B.6 (*iv*) that $\widehat{u}_t = u_t + O_{\mathbb{P}}(T^{-1/2})$. Therefore, $T^{-1/2}\sum_{t=1}^{T}\widehat{u}_{t-1}e_t = T^{-1/2}\sum_{t=1}^{T}u_{t-1}e_t + o_{\mathbb{P}}(1)$. In addition, $T^{-1/2}\sum_{t=1}^{T}u_{t-1}e_t = T^{-1}\sum_{t=1}^{T}\sqrt{T}(B_{ueT}(t/T) - B_{ueT}((t-1)/T)) = \int r dB_{ueT}(r) \Rightarrow \dot{\mathscr{S}}_1 := \int d\mathscr{B}_{ue}$. Second, we note that $\widehat{\mathbf{m}}_{t-1} = \mathbf{m}_{t-1} - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)(t-1)$. Therefore, $T^{-1}\sum_{t=1}^{T}\widehat{\mathbf{m}}_{t-1}e_t = T^{-1}\sum_{t=1}^{T}\mathbf{m}_{t-1}e_t - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)T^{-1}\sum_{t=1}^{T}(t-1)e_t$. We here note that $T^{-1}\sum_{t=1}^{T}\mathbf{m}_{t-1}e_t \Rightarrow \int \boldsymbol{\mathscr{B}}_m d\mathscr{B}_e$ by Lemma B.7 (*iv*) and $\sqrt{T}(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) \Rightarrow 3\int r\boldsymbol{\mathscr{B}}_m$ by Lemma B.7 (*i*). In addition to this, $T^{-3/2}\sum_{t=1}^{T}(t-1)e_t = T^{-1}\sum_{t=1}^{T}\frac{(t-1)}{T}\sqrt{T}(B_{eT}((t-1)/T) - B_{eT}((t-2)/T)) = \int r dB_{eT}(r) \Rightarrow \int r d\mathscr{B}_e$. Hence, it follows that $T^{-1}\sum_{t=1}^{T}\widehat{\mathbf{m}}_{t-1}e_t \Rightarrow (\dot{\boldsymbol{\mathscr{S}}}_2', \dot{\boldsymbol{\mathscr{S}}}_3')' := \int \boldsymbol{\mathscr{B}}_m d\mathscr{B}_e - 3\int r\boldsymbol{\mathscr{B}}_m \int r d\mathscr{B}_u = \int \bar{\boldsymbol{\mathscr{B}}}_m d\mathscr{B}_e$. Third, it is already showed that $T^{-3/2}\sum_{t=1}^{T}(t-1)e_{t-1} \Rightarrow \dot{\mathscr{S}}_4 := \int r d\mathscr{B}_e$. Fourth, we note that $T^{-1/2}\sum_{t=1}^{T}e_t = T^{-1}\sum_{t=1}^{T}\sqrt{T}(B_{eT}(t/T) - B_{eT}((t-1)/T)) = \int dB_{eT}(r) \Rightarrow \dot{\mathscr{S}}_5 := \int d\mathscr{B}_e$. Fifth, note that $T^{-1/2}\sum_{t=1}^{T}\boldsymbol{z}_{2t}e_t = T^{-1}\sum_{t=1}^{T}\sqrt{T}(\mathbf{B}_{zeT}(t/T) - \mathbf{B}_{zeT}((t-1)/T)) = \int d\mathbf{B}_{zeT}(r) \Rightarrow \dot{\boldsymbol{\mathscr{S}}}_6 := \int d\boldsymbol{\mathscr{B}}_{ze}$. We next combine the first to fifth facts to obtain that $\dot{\mathbf{D}}^{-1}\sum_{t=1}^{T}\dot{\boldsymbol{z}}_t e_t \Rightarrow \dot{\boldsymbol{\mathscr{S}}}$, leading to that $\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) \Rightarrow \dot{\boldsymbol{\mathscr{M}}}^{-1}\dot{\boldsymbol{\mathscr{S}}}$, as desired.

(*ii*) From the definitions of $\boldsymbol{\tau}_{T*} := [\rho_*, \boldsymbol{\tau}_{1T}', \boldsymbol{\tau}_{2*}']'$ and $\boldsymbol{\tau}_{1T} := [(\boldsymbol{\eta}_* + \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*))', \varsigma_* + \boldsymbol{\eta}_*'(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*) + \rho_*(\widetilde{\vartheta}_T - \vartheta_{T*}), \gamma_* + \rho_*(\widetilde{\nu}_T - \nu_*)]'$, we obtain that $\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) = \dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*) + [0, -\rho_* T(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)', -\rho_* T^{3/2}(\widetilde{\vartheta}_T - \vartheta_{T*}), -\rho_* T^{1/2}(\widetilde{\nu}_T - \nu_*), \mathbf{0}']' \Rightarrow \dot{\boldsymbol{\mathscr{L}}}$ by Lemma 4 (*i*) and noting that $\boldsymbol{\eta}_* = \mathbf{0}$. In addition, Lemma 3 implies that $T(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*) \Rightarrow (\dot{\boldsymbol{\mathscr{L}}}_{11}', \dot{\boldsymbol{\mathscr{L}}}_{12}')'$, and $T^{1/2}(\widetilde{\nu}_T - \nu_*) \Rightarrow \dot{\mathscr{L}}_{14}$. Furthermore, the proof of Lemma 3 (*i*) shows that $T^{3/2}(\widetilde{\vartheta}_T - \vartheta_{T*}) \Rightarrow \dot{\mathscr{L}}_{13}$. Therefore, $\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*) \Rightarrow \dot{\boldsymbol{\mathscr{L}}} + \rho_*[0, \dot{\boldsymbol{\mathscr{L}}}_{11}', \dot{\boldsymbol{\mathscr{L}}}_{12}', \dot{\mathscr{L}}_{13}, \dot{\mathscr{L}}_{14}, \mathbf{0}']'$. This completes the proof. ■

Proof of Theorem 1. (*i*) We note that $(\widehat{\boldsymbol{\alpha}}_T - \boldsymbol{\alpha}_*) = (\mathbf{T}_T - \mathbf{T})\boldsymbol{\tau}_* + \mathbf{T}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*) + (\mathbf{T}_T - \mathbf{T})(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*) = (\mathbf{T}_T - \mathbf{T})\boldsymbol{\tau}_* + \mathbf{T}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*) + o_{\mathbb{P}}(\mathbf{T}_T - \mathbf{T})$ using (14) and (15). We further note that

$$
\begin{aligned}
&(\mathbf{T}_T - \mathbf{T})\boldsymbol{\tau}_* + \mathbf{T}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*)\\
&= \begin{bmatrix}
(\dot{\rho}_T - \rho_*)\\
-\boldsymbol{\beta}_*(\dot{\rho}_T - \rho_*) + (\dot{\boldsymbol{\eta}}_T - \boldsymbol{\eta}_*) - \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)\\
(\dot{\varsigma}_T - \varsigma_*) - \boldsymbol{\mu}_*'(\dot{\boldsymbol{\eta}}_T - \boldsymbol{\eta}_*) - \zeta_*(\dot{\rho}_T - \rho_*) - \rho_*(\widehat{\zeta}_T - \zeta_*) - \boldsymbol{\eta}_*'(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)\\
(\dot{\gamma}_T - \gamma_*) - \rho_*(\widehat{\nu}_T - \nu_*) - \nu_*(\dot{\rho}_T - \rho_*)\\
(\dot{\boldsymbol{\tau}}_{2T} - \boldsymbol{\tau}_{*2})
\end{bmatrix}
\end{aligned}
$$

and the fact that $\widehat{\boldsymbol{\beta}}_T = \widetilde{\boldsymbol{\beta}}_T$, $\widehat{\zeta}_T = -\widehat{\boldsymbol{\mu}}_T'\widetilde{\boldsymbol{\beta}}_T + \widetilde{\vartheta}_T$ and $\zeta_* = -\boldsymbol{\mu}_*'\boldsymbol{\beta}_* + \vartheta_*$, so that $(\widehat{\zeta}_T - \zeta_*) = (\widetilde{\vartheta}_T - \vartheta_{T*}) - \boldsymbol{\mu}_*'(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*) - (\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)'(\widehat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)$, where $(\widetilde{\vartheta}_T - \vartheta_{T*}) = O_{\mathbb{P}}(T^{-3/2})$ and $(\widehat{\boldsymbol{\mu}}_T - \boldsymbol{\mu}_*)'(\widehat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*) =$

$O_{\mathbb{P}}(T^{-3/2})$ by Lemmas 4 (*i*), B.7 (*i*), and 3 (*ii*). Therefore, it now follows that

$$
\begin{aligned}
(\mathbf{T}_T - \mathbf{T})\boldsymbol{\tau}_* + &\mathbf{T}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*) \\
&= (\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) - (\dot{\rho}_T - \rho_*) \begin{bmatrix} 0 \\ \boldsymbol{\beta}_* \\ \zeta_* \\ \nu_* \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{0} \\ -\boldsymbol{\mu}_*'\{(\dot{\boldsymbol{\eta}}_T - \boldsymbol{\eta}_*) - \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)\} + O_{\mathbb{P}}(T^{-3/2}) \\ 0 \\ \mathbf{0} \end{bmatrix}. \quad \text{(A.4)}
\end{aligned}
$$

We here use the fact that $\boldsymbol{\eta}_* = \mathbf{0}$. We further note that $\sqrt{T}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) \Rightarrow [\dot{\mathscr{L}}_1, \mathbf{0}', \mathbf{0}', 0, \dot{\mathscr{L}}_5, \dot{\boldsymbol{\mathscr{L}}}_6]'$ by Lemma 4 (*i*) and $-\sqrt{T}(\dot{\rho}_T - \rho_*)\left[0, \boldsymbol{\beta}_*', \zeta_*, \nu_*, \mathbf{0}'\right] \Rightarrow -\dot{\mathscr{L}}_1\left[0, \boldsymbol{\beta}_*', \zeta_*, \nu_*, \mathbf{0}'\right]$. In addition, we note that $-\boldsymbol{\mu}_*'\sqrt{T}\{(\dot{\boldsymbol{\eta}}_T - \boldsymbol{\eta}_*) - \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)\} + O_{\mathbb{P}}(T^{-1}) = o_{\mathbb{P}}(1)$ because $\sqrt{T}\{(\dot{\boldsymbol{\eta}}_T - \boldsymbol{\eta}_*) - \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)\} = O_{\mathbb{P}}(T^{-1/2})$ by Lemma 4 (*i*). Therefore, it follows that $\sqrt{T}(\widehat{\boldsymbol{\alpha}}_T - \boldsymbol{\alpha}_*) \Rightarrow [\dot{\mathscr{L}}_1, -\boldsymbol{\beta}_*'\dot{\mathscr{L}}_1, -\zeta_*\dot{\mathscr{L}}_1, \dot{\mathscr{L}}_5, -\nu_*\dot{\mathscr{L}}_1, \dot{\boldsymbol{\mathscr{L}}}_6']'$. We finally note that the derived weak limit is identical to $\boldsymbol{c}_*\dot{\mathscr{L}}_1 + [0, \mathbf{0}', \mathbf{0}', 0, \dot{\mathscr{L}}_5, \dot{\boldsymbol{\mathscr{L}}}_6']'$.

(*ii*, *iii*, and *iv*) We prove (*ii*, *iii*, and *iv*) together by supposing that $\boldsymbol{\beta}_*^+ = \boldsymbol{\beta}_*^- = \mathbf{0}$ and $\zeta_* = 0$. If so, it follows from (A.4) that

$$
\begin{aligned}
(\widehat{\boldsymbol{\alpha}}_T - \boldsymbol{\alpha}_*) &= (\mathbf{T}_T - \mathbf{T})\boldsymbol{\tau}_* + \mathbf{T}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_*) + o_{\mathbb{P}}(\mathbf{T}_T - \mathbf{T}) \\
&= (\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) + \begin{bmatrix} 0 \\ \mathbf{0} \\ -\boldsymbol{\mu}_*'\{(\dot{\boldsymbol{\eta}}_T - \boldsymbol{\eta}_*) - \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)\} + O_{\mathbb{P}}(T^{-3/2}) \\ -\nu_*(\dot{\rho}_T - \rho_*) \\ \mathbf{0} \end{bmatrix} + o_{\mathbb{P}}(\mathbf{T}_T - \mathbf{T}).
\end{aligned}
$$

Here, Lemma 4 (*i*) implies that $\ddot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) \Rightarrow [\dot{\mathscr{L}}_1, \dot{\boldsymbol{\mathscr{L}}}_2', \dot{\boldsymbol{\mathscr{L}}}_3', 0, \dot{\mathscr{L}}_5, \dot{\boldsymbol{\mathscr{L}}}_6]'$ and $T\{(\dot{\boldsymbol{\eta}}_T - \boldsymbol{\eta}_*) - \rho_*(\widetilde{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_*)\} \Rightarrow (\dot{\boldsymbol{\mathscr{L}}}_2', \dot{\boldsymbol{\mathscr{L}}}_3')'$, so that $\ddot{\mathbf{D}}(\widehat{\boldsymbol{\alpha}}_T - \boldsymbol{\alpha}_*) \Rightarrow [\dot{\mathscr{L}}_1, \dot{\boldsymbol{\mathscr{L}}}_2', \dot{\boldsymbol{\mathscr{L}}}_3', -\boldsymbol{\mu}_*^{+\prime}\dot{\boldsymbol{\mathscr{L}}}_2 - \boldsymbol{\mu}_*^{-\prime}\dot{\boldsymbol{\mathscr{L}}}_3, \dot{\mathscr{L}}_5 - \nu_*\dot{\mathscr{L}}_1, \dot{\boldsymbol{\mathscr{L}}}_6']'$, where $\ddot{\mathbf{D}} := \mathrm{diag}[\sqrt{T}, T\mathbf{I}_{2k+1}, \sqrt{T}, \dot{\mathbf{D}}_2]$. From this result, (*ii*, *iii*, and *iv*) follow. ■

Proof of Theorem 2. (*i*) We show the null limit distribution of each test using Lemma 4.

(*i.a*) Note that $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) = (\boldsymbol{\theta}_*^+ - \boldsymbol{\theta}_*^-) - (\dot{\rho}_T - \rho_*)(\boldsymbol{\beta}_*^+ - \boldsymbol{\beta}_*^-) + (\dot{\boldsymbol{\eta}}_T^+ - \boldsymbol{\eta}_{T*}^+) - (\dot{\boldsymbol{\eta}}_T^- - \boldsymbol{\eta}_{T*}^-) + o_{\mathbb{P}}(T^{-1})$, so that it follows that $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) = (\dot{\boldsymbol{\eta}}_T^+ - \boldsymbol{\eta}_{T*}^+) - (\dot{\boldsymbol{\eta}}_T^- - \boldsymbol{\eta}_{T*}^-) + o_{\mathbb{P}}(T^{-1})$ under $\mathscr{H}_0'$. Therefore, $\dot{\mathbf{R}}_1\widehat{\boldsymbol{\alpha}}_T =$

$\mathbf{R}_1(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) + o_{\mathbb{P}}(T^{-1})$, and it follows that

$$W_T^{(1)} = (\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*})'\dot{\mathbf{D}}\mathbf{R}_1' \left(\widehat{\sigma}_{e,T}^2 \mathbf{R}_1 \left(\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\dot{\mathbf{D}}^{-1}\right)^{-1}\mathbf{R}_1'\right)^{-1}\mathbf{R}_1\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) + o_{\mathbb{P}}(1)$$
$$\Rightarrow \dot{\mathscr{L}}'\mathbf{R}_1'\left(\sigma_e^2\mathbf{R}_1\dot{\mathscr{M}}^{-1}\mathbf{R}_1'\right)^{-1}\mathbf{R}_1\dot{\mathscr{L}}$$

by Lemmas 4 (i) and B.7 (v).

($i.b$) We note that $(\widehat{\boldsymbol{\pi}}_T^+, \widehat{\boldsymbol{\pi}}_T^-) = (\dot{\boldsymbol{\pi}}_T^+, \dot{\boldsymbol{\pi}}_T^-)$. Therefore, it follows that

$$W_T^{(2)} = (\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*})'\dot{\mathbf{D}}\dot{\mathbf{R}}_2' \left(\widehat{\sigma}_{e,T}^2 \dot{\mathbf{R}}_2 \left(\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\dot{\mathbf{D}}^{-1}\right)^{-1}\dot{\mathbf{R}}_2'\right)^{-1}\dot{\mathbf{R}}_2\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) + o_{\mathbb{P}}(1)$$
$$\Rightarrow \dot{\mathscr{L}}'\dot{\mathbf{R}}_2'\left(\sigma_e^2\dot{\mathbf{R}}_2\dot{\mathscr{M}}^{-1}\dot{\mathbf{R}}_2'\right)^{-1}\dot{\mathbf{R}}_2\dot{\mathscr{L}}$$

by Lemmas 4 (i) and B.7 (v).

($i.c$) From the notice given in the proofs of ($i.a$ and $i.b$), it follows that

$$W_T^{(3)} = (\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*})'\dot{\mathbf{D}}\mathbf{R}_3' \left(\widehat{\sigma}_{e,T}^2 \mathbf{R}_3 \left(\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\dot{\mathbf{D}}^{-1}\right)^{-1}\mathbf{R}_3'\right)^{-1}\mathbf{R}_3\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) + o_{\mathbb{P}}(1)$$
$$\Rightarrow \dot{\mathscr{L}}'\mathbf{R}_3'\left(\sigma_e^2\mathbf{R}_3\dot{\mathscr{M}}^{-1}\mathbf{R}_3'\right)^{-1}\mathbf{R}_3\dot{\mathscr{L}}$$

by Lemmas 4 (i) and B.7 (v).

(ii) We show the power behavior of each test using Lemma 4.

($ii.a$) We show that $W_T^{(1)} = O_{\mathbb{P}}(T)$. Note that $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) = (\boldsymbol{\theta}_*^+ - \boldsymbol{\theta}_*^-) - (\dot{\rho}_T - \rho_*)(\boldsymbol{\beta}_*^+ - \boldsymbol{\beta}_*^-) + o_{\mathbb{P}}(T^{-1/2})$ under $\mathscr{H}_1'$ and $(\dot{\rho}_T - \rho_*) = O_{\mathbb{P}}(T^{-1/2})$, so that $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) - (\boldsymbol{\theta}_*^+ - \boldsymbol{\theta}_*^-) = O_{\mathbb{P}}(T^{-1/2})$. Therefore, $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) = O_{\mathbb{P}}(T^{1/2})$, and this implies that $W_T^{(1)} = O_{\mathbb{P}}(T)$.

($ii.b$) We show that $W_T^{(2)} = O_{\mathbb{P}}(T)$. Note that Lemma 4 ($i$) implies that $\dot{\boldsymbol{\pi}}_T^+ - \dot{\boldsymbol{\pi}}_T^- = \boldsymbol{\pi}_*^+ - \boldsymbol{\pi}_*^- + O_{\mathbb{P}}(T^{-1/2})$. Therefore, $\dot{\boldsymbol{\pi}}_T^+ - \dot{\boldsymbol{\pi}}_T^- = O_{\mathbb{P}}(T^{1/2})$, implying that $W_T^2 = O_{\mathbb{P}}(T)$.

($ii.c$) From the proofs of ($ii.a$) and ($ii.b$), $(\widehat{\boldsymbol{\theta}}_T^+ - \widehat{\boldsymbol{\theta}}_T^-) = O_{\mathbb{P}}(T^{1/2})$ and $(\dot{\boldsymbol{\pi}}_T^+ - \dot{\boldsymbol{\pi}}_T^-) = O_{\mathbb{P}}(T^{1/2})$. Therefore, it trivially follows that $W_T^{(3)} = O_{\mathbb{P}}(T)$.

(iii) We show the null limit distribution of each test using Lemmas 3 and 4.

($iii.a$) First, Lemma B.6 (iv) implies that $\widehat{\sigma}_{u,T}^2 \xrightarrow{\mathbb{P}} \sigma_u^2$, and $\dot{\mathbf{D}}_1^{-1}\sum_{t=1}^T \dot{\boldsymbol{r}}_{t-1}\dot{\boldsymbol{r}}_{t-1}'\dot{\mathbf{D}}_1^{-1} \Rightarrow \dot{\mathscr{M}}_{11}$ from Lemma 2 (vii). Therefore, $\ddot{\mathbf{W}}_T^{(1)} \Rightarrow \sigma_u^2\ddot{\mathbf{R}}_1\dot{\mathscr{M}}_{11}^{-1}\ddot{\mathbf{R}}_1'$. Second, $\mathbb{H}_0'$ implies that $\ddot{\mathbf{R}}_1\dot{\mathbf{D}}_1\widetilde{\boldsymbol{v}}_T = T(\widetilde{\boldsymbol{\beta}}_T^+ - \boldsymbol{\beta}_*^+) \Rightarrow (\dot{\mathscr{L}}_{11} - \dot{\mathscr{L}}_{12})$. If we combine these two facts, it follows that $\mathscr{W}_T^{(1)} \Rightarrow (\dot{\mathscr{L}}_{11} - \dot{\mathscr{L}}_{12})'(\sigma_u^2\ddot{\mathbf{R}}_1\dot{\mathscr{M}}_{11}^{-1}\ddot{\mathbf{R}}_1')^{-1}(\dot{\mathscr{L}}_{11}$

$-\dot{\mathscr{L}}_{12})$ that is identical to the given weak limit by the definitions of $\ddot{\mathbf{R}}_1$ and $\dot{\mathscr{L}}_1$.

(*iii.b*) First, Lemma B.7 (*v*) implies that $\widehat{\sigma}^2_{e,T} \overset{\mathbb{P}}{\to} \sigma^2_e$, and $\dot{\mathbf{D}}^{-1}\sum_{t=1}^T \dot{\boldsymbol{z}}_t\dot{\boldsymbol{z}}_t'\dot{\mathbf{D}}^{-1} \Rightarrow \dot{\mathscr{M}}$ from Lemma 2 (*vi*). Therefore, $\ddot{\mathbf{W}}^{(2)}_T \Rightarrow \sigma^2_e\ddot{\mathbf{R}}_2\dot{\mathscr{M}}^{-1}\ddot{\mathbf{R}}_2'$. Second, $\ddot{\mathbf{R}}_2\dot{\mathbf{D}}\dot{\boldsymbol{\tau}}_T = \sqrt{T}(\dot{\boldsymbol{\pi}}^+_T - \dot{\boldsymbol{\pi}}^-_T) \Rightarrow \dot{\mathscr{L}}_{62} - \dot{\mathscr{L}}_{63} = \ddot{\mathbf{R}}_2\dot{\mathscr{L}}$. If we combine these two facts, it follows that $\mathscr{W}^{(2)}_T \Rightarrow \dot{\mathscr{L}}'\ddot{\mathbf{R}}_2'(\sigma^2_2\ddot{\mathbf{R}}_2\dot{\mathscr{M}}^{-1}\ddot{\mathbf{R}}_2')^{-1}\ddot{\mathbf{R}}_2\dot{\mathscr{L}}$.

(*iii.c*) We note that from the definition of $\mathscr{W}^{(3)}_T$, it follows that $\mathscr{W}^{(3)}_T = \mathscr{W}^{(1)}_T + \mathscr{W}^{(2)}_T$, and it follows from (*i.a* and *i.b*) that $\mathscr{W}^{(3)}_T \Rightarrow \dot{\mathscr{L}}_1'\ddot{\mathbf{R}}_1'(\sigma^2_u\ddot{\mathbf{R}}_1\dot{\mathscr{M}}^{-1}_{11}\ddot{\mathbf{R}}_1')^{-1}\ddot{\mathbf{R}}_1\dot{\mathscr{L}}_1 + \dot{\mathscr{L}}'\ddot{\mathbf{R}}_2'(\sigma^2_2\ddot{\mathbf{R}}_2\dot{\mathscr{M}}^{-1}\ddot{\mathbf{R}}_2')^{-1}\ddot{\mathbf{R}}_2\dot{\mathscr{L}}$ under $\mathbb{H}_0'''$.

(*iv*) We next show the power behavior of each test.

(*iv.a*) To show the claim, we show that $\mathscr{W}^{(1)}_T = O_{\mathbb{P}}(T^2)$ under $\mathbb{H}_1'$. Given that $\ddot{\mathbf{W}}^{(1)}_T \Rightarrow \sigma^2_u\ddot{\mathbf{R}}_1\dot{\mathscr{M}}^{-1}_{11}\ddot{\mathbf{R}}_1'$, we focus on the limit behavior of $\ddot{\mathbf{R}}_1\dot{\mathbf{D}}_1\widetilde{\boldsymbol{v}}_T$. Note that $\ddot{\mathbf{R}}_1\dot{\mathbf{D}}_1(\widetilde{\boldsymbol{v}}_T - \bar{\boldsymbol{v}}_{T*}) \Rightarrow (\dot{\mathscr{L}}_{11} - \dot{\mathscr{L}}_{12})$ by Lemma 3 (*i*) and $\ddot{\mathbf{R}}_1\dot{\mathbf{D}}_1\bar{\boldsymbol{v}}_{T*} = T\boldsymbol{\beta}_* = O(T)$. Therefore, $\ddot{\mathbf{R}}_1\dot{\mathbf{D}}_1\widetilde{\boldsymbol{v}}_T = O_{\mathbb{P}}(T)$, implying that $\mathscr{W}^{(1)}_T = O_{\mathbb{P}}(T^2)$, leading to the desired result.

(*iv.b*) We show that $\mathscr{W}^{(2)}_T = O_{\mathbb{P}}(T)$ under $\mathscr{H}_1''$. Given that $\ddot{\mathbf{W}}^{(2)}_T \Rightarrow \sigma^2_e\ddot{\mathbf{R}}_2\dot{\mathscr{M}}^{-1}\ddot{\mathbf{R}}_2'$, we focus on the limit behavior of $\ddot{\mathbf{R}}_2\dot{\mathbf{D}}\dot{\boldsymbol{\tau}}_T$. We note that $\ddot{\mathbf{R}}_2\dot{\mathbf{D}}(\dot{\boldsymbol{\tau}}_T - \boldsymbol{\tau}_{T*}) = \sqrt{T}[(\dot{\boldsymbol{\pi}}^+_T - \dot{\boldsymbol{\pi}}^-_T) - (\boldsymbol{\pi}^+_* - \boldsymbol{\pi}^-_*)] \Rightarrow \ddot{\mathbf{R}}_2\dot{\mathscr{L}}$ by Lemma 4 (*i*) and $\ddot{\mathbf{R}}_2\dot{\mathbf{D}}\boldsymbol{\tau}_{T*} = \sqrt{T}(\boldsymbol{\pi}^+_* - \boldsymbol{\pi}^-_*) = O(T^{1/2})$. Therefore, $\ddot{\mathbf{R}}_2\dot{\mathbf{D}}\dot{\boldsymbol{\tau}}_T = O_{\mathbb{P}}(T^{1/2})$, implying that $\mathscr{W}^{(2)}_T = O_{\mathbb{P}}(T)$.

(*iv.c*) We show that $\mathscr{W}^{(3)}_T = O_{\mathbb{P}}(T^2)$ under $\mathbb{H}_1'''$. Note that $\mathscr{W}^{(3)}_T = \mathscr{W}^{(1)}_T + \mathscr{W}^{(2)}_T$, and $\mathscr{W}^{(1)}_T$ and $\mathscr{W}^{(2)}_T$ are $O_{\mathbb{P}}(T^2)$ and $O_{\mathbb{P}}(T)$ by (*ii.a*) and (*ii.b*), respectively. Therefore, $\mathscr{W}^{(3)}_T = O_{\mathbb{P}}(T^2)$ under $\mathbb{H}_{11}'''$ and $\mathscr{W}^{(3)}_T = O_{\mathbb{P}}(T)$ under $\mathbb{H}_{01}''' \bigcap \mathbb{H}_{12}'''$. This completes the proof. ■

A.4 Additional Empirical Supplements

In this section, we provide additional empirical supplements.

Two tables are provided. First, Table A.1 provides the descriptive statistics of the variables examined in Sections 7.2.1 and 7.2.2. The sample period is from 1947Q1 to 2007:Q4.

Second, Table A.2 provides the testing results using Phillips and Perron's (1988) unit root test applied to the partial sum processes for Tables 5 and 7. As we apply the unit-root testing by including both constant and trend or including only constant, two testing results are provided for each variable. Except r_t, the test results show that nonstationary data analysis has to be conducted for the other variables.

	Δy_t	Exo.			Endo.			Sum.		
		$\Delta\tau_{1t}$	$\Delta\tau_{2t}$	$\Delta\tau_t$	$\Delta\tau_{1t}$	$\Delta\tau_{2t}$	$\Delta\tau_t$	$\Delta\tau_{1t}$	$\Delta\tau_{2t}$	$\Delta\tau_t$
Mean	0.8256	0.4240	-0.4704	-0.0464	0.3773	-0.1841	0.1932	0.7564	-0.6228	0.1336
Median	0.7876	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Maximum	4.0198	6.4312	0.0000	6.4312	6.7617	0.0000	6.7617	6.7617	0.0000	6.7617
Minimum	-2.7525	0.0000	-7.0965	-7.0965	0.0000	-7.1122	-7.1122	0.0000	-7.1122	-7.1122
Std. Dev.	0.9780	1.3755	1.5073	2.1364	1.3317	1.0117	1.7136	1.7775	1.7316	2.6653
Skewness	-0.0501	3.0654	-3.0598	-0.2762	3.4094	-5.4276	0.4394	2.0432	-2.5281	-0.1450
Kurtosis	4.3614	10.8514	10.8759	6.3787	13.2115	31.1514	10.4245	5.4941	7.7131	4.1776
Obs.	243	243	243	243	243	243	243	243	243	243
	Δy_t	Exo. ratio			Endo. ratio			Sum. ratio		
		Δr_{1t}	Δr_{2t}	Δr_t	Δr_{1t}	Δr_{2t}	Δr_t	Δr_{1t}	Δr_{2t}	Δr_t
Mean	0.8256	0.0246	-0.0517	-0.0271	0.0471	-0.0260	0.0211	0.0679	-0.0738	-0.0059
Median	0.7876	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Maximum	4.0198	0.6977	0.0000	0.6977	1.9542	0.0000	1.9542	1.9542	0.0000	1.9542
Minimum	-2.7525	0.0000	-1.8706	-1.8706	0.0000	-2.8214	-2.8214	0.0000	-2.8214	-2.8214
Std. Dev.	0.9780	0.0963	0.2171	0.2428	0.2245	0.2028	0.3066	0.2351	0.2876	0.3847
Skewness	-0.0501	4.6231	-5.3926	-3.7013	6.0062	-11.5395	-1.0526	5.1164	-5.9045	-1.3783
Kurtosis	4.3614	25.4621	35.2975	25.0905	41.7414	152.4675	44.1557	33.0743	45.6366	21.6860
Obs.	243	243	243	243	243	243	243	243	243	243

Table A.1: DESCRIPTIVE STATISTICS. This table shows the descriptive statistics used in Sections 7.2.1 and 7.2.2.

PP test	y_t	Exo.			Endo.			Sum.		
		τ_{1t}	τ_{2t}	τ_t	τ_{1t}	τ_{2t}	τ_t	τ_{1t}	τ_{2t}	τ_t
PP test w/o trend	-1.4767	0.7551	-0.7188	-1.6243	-2.4701	-0.3499	-2.2743	-1.3064	-0.4173	-2.0303
p-value	0.5439	0.9931	0.8387	0.4686	0.1241	0.9140	0.1812	0.6270	0.9028	0.2738
PP test w/ trend	-2.3488	-1.8278	-1.3653	-1.3483	-0.7094	-2.2995	-0.8156	-0.9285	-2.0326	-2.7414
p-value	0.4056	0.6883	0.8686	0.8732	0.9707	0.4322	0.9618	0.9500	0.5801	0.2210
PP test	y_t	Exo. ratio			Endo. ratio			Sum. ratio		
		r_{1t}	r_{2t}	r_t	r_{1t}	r_{2t}	r_t	r_{1t}	r_{2t}	r_t
PP test w/o trend	-1.4767	0.7288	-0.7859	-1.7004	-2.5409	-0.9379	-3.0710	-1.7735	-1.8303	-2.9448
p-value	0.5439	0.9926	0.8210	0.4299	0.1071	0.7749	0.0301	0.3932	0.3652	0.0418
PP test w/ trend	-2.3488	-1.8466	-2.3558	-2.2056	-1.5667	-2.0336	-2.6499	-1.3375	-2.3690	-3.4114
p-value	0.4056	0.6790	0.4019	0.4840	0.8033	0.5796	0.2587	0.8761	0.3949	0.0521

Table A.2: PHILLIPS AND PERRON'S (1988) UNIT-ROOT TESTS APPLIED TO THE QUARTERLY DATA IN ROMER AND ROMER (2010). Two Phillips and Perron's test statistics are computed using the data in Table A.1 by including the time trend and/or the constant. The lag lengths are selected by BIC.