9-1-2013

Testing Linearity Using Power Transforms of Regressors

Yae In Baek

Jin Seo Cho

Peter C.B. Phillips

TESTING LINEARITY USING POWER TRANSFORMS OF REGRESSORS

By

Yae In Baek, Jin Seo Cho and Peter C.B. Phillips

September 2013

COWLES FOUNDATION DISCUSSION PAPER NO. 1917

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Testing Linearity Using Power Transforms of Regressors*

YAE IN BAEK
School of Economics
Yonsei University
50 Yonsei-ro, Seodaemun-gu, Seoul, 120-749, Korea

JIN SEO CHO
School of Economics
Yonsei University
50 Yonsei-ro, Seodaemun-gu, Seoul, 120-749, Korea

PETER C.B. PHILLIPS
Yale University, University of Auckland
Singapore Management University & University of Southampton

First version: December, 2012. This version: August, 2013

Abstract

We develop a method of testing linearity using power transforms of regressors, allowing for stationary processes and time trends. The linear model is a simplifying hypothesis that derives from the power transform model in three different ways, each producing its own identification problem. We call this modeling difficulty the *trifold identification problem* and show that it may be overcome using a test based on the quasi-likelihood ratio (QLR) statistic. More specifically, the QLR statistic may be approximated under each identification problem and the separate null approximations may be combined to produce a composite approximation that embodies the linear model hypothesis. The limit theory for the QLR test statistic depends on a Gaussian stochastic process. In the important special case of a linear time trend regressor and martingale difference errors asymptotic critical values of the test are provided. The paper also considers generalizations of the Box-Cox transformation, which are associated with the QLR test statistic.

Key Words: Box Cox transform; Gaussian stochastic process; Neglected nonlinearity; Power transformation; Quasi-likelihood ratio test; Trend exponent; Trifold identification problem.

JEL Classification: C12, C18, C46, C52.

*The authors have benefited from discussions with Stan Hurn, Isao Ishida, Valentin Zelenyuk, and other participants at the NZESG meeting in Auckland, 2013. Baek and Cho acknowledge research support from the second stage BK21 project of the Korean Research Foundation, and Phillips acknowledges support from a Kelly Fellowship and the NSF under Grant No. SES 12-58258.

1 Introduction

Linear models are a natural starting point in empirical work, offering advantages of straightforward computation and convenient interpretation, particularly with respect to the estimation of marginal effects. These models also relate in a fundamental way to underlying Gaussian assumptions and the use of wide sense conditional expectations. Testing linearity is therefore a familiar practice in applications, with major potential modeling implications in the case of rejection. In consequence, linearity is a central concern in theoretical work on model specification tests and much specification testing commences from simple linear model assumptions.

Power transformations are popularly used as alternatives to linearity. Tukey (1957, 1977) provides several rationales for the use of power transformations as alternatives to linear models and Box and Cox (1964) further developed their use in nonlinear modeling. The Box-Cox transformation, in particular, successfully implements the so-called Tukey ladder of power option in modeling by means of its flexible parametric form. In time series applications, some studies (notably, Wu (1981) and Phillips (2007)) have considered power transforms of time-trend regressors, providing limit theories that are relevant in estimating the parameters of these nonlinear systems. Such models are useful alternatives to the linear trend stationary specifications that are common in applied work.

Power transformations can be used to form test statistics that deliver consistent power against arbitrary alternatives to the linear model assumption. As Stinchcombe and White (1998) showed, any non-polynomial analytic function can be used to construct what are known as generically comprehensively revealing (GCR) test statistics, in the sense that they can reveal misspecification in regression relationships. This property, as an activation function for GCR tests, motivates use of power transforms for constructing test statistics with omnibus power.

In spite of this apparently useful property, testing linearity using power transforms is largely undeveloped in the literature, mainly because of the identification problem that arises under the null of the linear model assumption. As detailed below, the linear model hypothesis can be deduced from a power transformation in three different ways, each of which involves its own identification problem, a feature that we call the *trifold identification problem*. To our knowledge, this problem has never before been addressed in the literature.

The primary goal of the present work is therefore to resolve this complex threefold problem. Our focus is pragmatic and involves constructing mechanisms for testing linearity using power transformations. We concentrate on models involving power transforms of a strictly stationary variable or a time trend. While this focus excludes some possibilities, such as potential nonlinear transforms of nonstationary variates like random walks (e.g. Park and Phillips, 1999, and Shi and Phillips, 2012), the range of potential applications is large and includes both microeconometric and time series data.

This paper restricts attention to a particular statistic, the quasi-likelihood ratio (QLR) statistic. In part, this is because the QLR statistic assists in overcoming the difficulties of identification. As we demonstrate in the paper, the QLR statistic may be approximated in different ways under the linearity null and these null approximations may be usefully combined to produce a composite form that embodies the linear model hypothesis. An additional benefit from focusing on the QLR test is its relationship to the Box-Cox transformation. When testing linearity using the power transformation, the score of the test turns out to be related to an augmented form of Box-Cox transform. This structure provides some additional implications of the Box-Cox transformation that are explored in the paper.

Our approach to developing a null approximation of the QLR test extends the methodology of Cho and Ishida (2012), who studied how to test the effects of omitted power transformations. We advance that work and compare our null approximation with QLR test statistics that are popular in the artificial neural network (ANN) literature. Our approach also exploits the properties of time-trend power transforms and regressions studied recently in Phillips (2007). Time trend regressors and their power transforms have very different properties from those of stationary regressors in view of the asymptotic degeneracy of the signal matrix. This asymptotic multicollinearity of the induced regressors gives convergence rates that are case-dependent and involve slowly varying factors. These considerations violate the regularity conditions that typically operate for stationary variables where there is no degeneracy and common rates of convergence apply.

The remainder of the paper is organized as follows. Section 2 examines power transformations of a stationary process and tests for neglected nonlinearity. The relevant previous literature is overviewed, specific motivations for the present work are provided, and null approximations of the QLR test under each identification problem are developed, leading to the composite trifold identification problem. This Section also provides asymptotic theory for the tests in the stationary case. Section 3 extends the discussion and asymptotic results to power transforms of a time-trend regressor. Concluding remarks are given in Section 4. All proofs are in the Appendix.

For an arbitrary function f and $j = 1, 2, \ldots$, we let $(d^j/d^jx)f(0)$ denote $(d^j/dx^j)f(x)|_{x=0}$ for notational simplicity. Other notation is standard.

2 Testing for Neglected Power Transforms of a Stationary Regressor

2.1 Transform Models and Trifold Identification

We consider the following (maintained) model for $\mathbb{E}[Y_t|\mathbf{Z}_t]$:

$$\mathcal{M} := \{m_t(\cdot) : \mathbf{\Omega} \mapsto \mathbb{R} : m_t(\alpha, \boldsymbol{\delta}, \beta, \gamma) := \alpha + \mathbf{W}_t'\boldsymbol{\delta} + \beta X_t^{\gamma}\},$$

where $(Y_t, \mathbf{W}_t')' := (Y_t, X_t, \mathbf{D}_t')' \in \mathbb{R}^{2+k}$ ($k \in \mathbb{N}$) is strictly stationary and an absolutely regular mixing process with mixing coefficients β_ℓ that satisfy

$$\sum_{\ell=1}^{\infty} \ell^{1/(r-1)}\beta_\ell < \infty \text{ for some } r > 1; \tag{1}$$

X_t is positively valued; $\mathbf{Z}_t := (1, \mathbf{W}_t')'$; n is the sample size; $\mathbf{Z}'\mathbf{Z} = \sum_{t=1}^{n} \mathbf{Z}_t\mathbf{Z}_t'$ is nonsingular; and $\mathbf{\Omega}$ denotes the parameter space of $\boldsymbol{\omega} := (\alpha, \boldsymbol{\delta}', \beta, \gamma)'$. We also let $\boldsymbol{\delta} := (\xi, \boldsymbol{\eta}')'$ so that $\mathbf{W}_t'\boldsymbol{\delta} = \xi X_t + \mathbf{D}_t'\boldsymbol{\eta}$. In Section 3, X_t is the time trend t. In the stationary case considered here, the mixing condition (1) is imposed to satisfy the functional central limit theorem (FCLT) of Doukhan, Masart, and Rio (1995).

Our interest is in testing the effective form of X_t to $\mathbb{E}[Y_t|\mathbf{Z}_t]$, and we consider the following explicit hypotheses for this purpose:

$$\mathcal{H}_0 : \exists(\alpha_*, \boldsymbol{\delta}_*), \mathbb{E}[Y_t|\mathbf{Z}_t] = \alpha_* + \mathbf{W}_t'\boldsymbol{\delta}_* \text{ w.p. } 1 \text{ ; vs. } \mathcal{H}_1 : \forall(\alpha, \boldsymbol{\delta}), \mathbb{E}[Y_t|\mathbf{Z}_t] = \alpha + \mathbf{W}_t'\boldsymbol{\delta} \text{ w.p. } < 1. \tag{2}$$

The affix '$*$' is used to parameterize $\mathbb{E}[Y_t|\mathbf{Z}_t]$, so that for some α_o and β_o, $(\alpha_*, \beta_*, \gamma_*) \in \{(\alpha,\beta,\gamma) : \alpha + \beta X_t^{\gamma} = \alpha_o \text{ or } \alpha + \beta X_t^{\gamma} = \beta_o X_t\}$ under $\mathcal{H}_0$.

Testing the linear model hypothesis using a maintained model with a nonlinear component is common practice in the literature. Such tests may be regarded as a variant of Bierens' (1990) specification test. Similarly, Stinchcombe and White's (1998) GCR tests are constructed to test for the effect of a nonlinear component. A power transform is particularly popular for the nonlinear component. For example, Tukey (1957, 1977) introduced power transform flexible nonlinear models, which motivated the Box-Cox transformation. More specifically, Box and Cox (1964) found that this transformation accords with Tukey's (1957) 'ladder of power' and it has been widely applied in empirical work because of its convenient flexibility. The popularity of this methodology is well documented by Sakia (1992).

Power transforms are nonpolynomial and analytic, so that a statistic used to test the effect of the power component generally satisfies the criteria to be GCR, as pointed out by Stinchcombe and White (1998). Such tests can therefore consistently detect arbitrary alternatives to the linear model hypothesis. The literature contains many other variations of power transforms such as those used in Ramsey's (1969) specification test, where model specification is tested by using transforms in which the power coefficients are fixed to some known numbers.

Nothwithstanding this considerable interest in power transforms and specification tests, the hypotheses given in (2) have not been formally examined in the literature mainly because testing the hypotheses in (2) cannot be conducted in a standard way. More precisely, there are three different identification problems that arise under $\mathcal{H}_0$. If $\beta_* = 0$, γ_* is not identified and Davies' (1977, 1987) identification problem arises. On the other hand, if $\gamma_* = 0$, $\alpha_* + \beta_*$ is identified, but neither α_* nor β_* is separately identified. Furthermore, if $\gamma_* = 1$ and $\boldsymbol{\delta}_*$ is conformably partitioned as $(\xi_*, \eta_*)'$, $\xi_* + \beta_*$ is identified although neither ξ_* nor β_* is identified. Thus, *three* different identification problems arise under the linear model hypothesis. We denote these three hypotheses as

$$\mathcal{H}_0' : \beta_* = 0; \quad \mathcal{H}_0'' : \gamma_* = 0; \quad \text{and} \quad \mathcal{H}_0''' : \gamma_* = 1.$$

We call this construct the *trifold identification problem*.

To the best of our knowledge, the literature presently approaches the trifold identification problem only in a limited way. Hansen (1996), for instance, provided a testing methodology that employs the weighted bootstrap to treat the first null hypothesis $\mathcal{H}_0'$. Alternatively, the power coefficient might be fixed at some known number as in Ramsey's (1969) specification test, so that the identification problems under $\mathcal{H}_0''$ and $\mathcal{H}_0'''$ are avoided. Accordingly, the main goal of the current study is to provide a tractable test statistic that is able to handle the trifold identification problem within a unified framework and obtain a feasible asymptotic null distribution that can be used for inference about power transforms.

There are related identification problems that have appeared in the literature although there are important differences in the details. Cho, Ishida, and White (2011, 2013) test for neglected nonlinearity using ANN models and find that two different identification problems arise under the null of linearity. They show how this twofold identification problem may be addressed using the QLR test. Cho and Ishida (2012) similarly test for effects of power transforms using the same QLR statistic but their focus of interest differs from ours and their model has only a twofold identification problem. None of this work considers nonlinear trend effects.

The approach taken in the current work is to extend the analysis of Cho, Ishida, and White (2011, 2013) and Cho and Ishida (2012). The maximum order involved in the null approximation used in Cho, Ishida, and White (2011) is the fourth order, whereas that used in Cho, Ishida, and White (2013) is the sixth order. These authors observe that the maximum order is dependent on the activation function used in constructing the test. On the other hand, Cho and Ishida (2012) use a second-order approximation as is common in econometric practice. The present paper examines how these asymptotic approximations are modified by the trifold identification problem that appears here.

As in this earlier research, we therefore follow ongoing practice in the literature and examine the QLR test defined as

$$QLR_n := n(1 - \hat{\sigma}^2_{n,A}/\hat{\sigma}^2_{n,0}),$$

where $\hat{\sigma}^2_{n,A} := \inf_{\alpha,\beta,\gamma,\boldsymbol{\delta}} \frac{1}{n}\sum_{t=1}^n (Y_t - \alpha - \mathbf{W}_t'\boldsymbol{\delta} - \beta X_t^{\gamma})^2$ and $\hat{\sigma}^2_{n,0} := \inf_{\alpha,\boldsymbol{\delta}} \frac{1}{n}\sum_{t=1}^n (Y_t - \alpha - \mathbf{W}_t'\boldsymbol{\delta})^2$. The following subsections separately examine the asymptotic approximations of the QLR statistic that apply under $\mathcal{H}_0'$, $\mathcal{H}_0''$, and $\mathcal{H}_0'''$.

Before proceeding it is convenient to collect the assumptions above into the following formal statement of conditions.

Assumption 1. *(i)* $(Y_t, \mathbf{W}_t')' := (Y_t, X_t, \mathbf{D}_t')' \in \mathbb{R}^{2+k}$ *(*$k \in \mathbb{N}$*) is a strictly stationary and absolutely regular process with mixing coefficients* β_ℓ *such that for some* $r > 1$, $\sum_{j=1}^{\infty} \ell^{1/(r-1)}\beta_\ell < \infty$, $\mathbb{E}[|Y_t|] < \infty$, *and* X_t *is positively valued with probability 1;*

(ii) $\mathbb{E}[Y_t|\mathbf{Z}_t]$ *is specified as* $\mathcal{M} := \{m_t(\cdot) : \boldsymbol{\Omega} \mapsto \mathbb{R} : m_t(\alpha, \boldsymbol{\delta}, \beta, \gamma) := \alpha + \mathbf{W}_t'\boldsymbol{\delta} + \beta X_t^{\gamma}\}$, *where* $\boldsymbol{\Omega}$ *is the parameter space of* $\boldsymbol{\omega} := (\alpha, \boldsymbol{\delta}', \beta, \gamma)'$*;* $\mathbf{Z}_t := (1, \mathbf{W}_t')'$*; and* n *is the sample size;*

(iii) $\boldsymbol{\Omega} = \boldsymbol{A} \times \boldsymbol{\Delta} \times \boldsymbol{B} \times \boldsymbol{\Gamma}$ *such that* $\boldsymbol{A}$, $\boldsymbol{\Delta}$, $\boldsymbol{B}$, *and* $\boldsymbol{\Gamma}$ *are convex and compact parameter spaces in* $\mathbb{R}$, $\mathbb{R}^{k+1}$, $\mathbb{R}$, *and* $\mathbb{R}$, *respectively, such that 0 and 1 are interior elements of* $\boldsymbol{\Gamma}$*; and*

(iv) $\mathbf{Z}'\mathbf{Z} = \sum_{t=1}^n \mathbf{Z}_t\mathbf{Z}_t'$ *is nonsingular with probability 1.* □

2.2 QLR Statistic under $\mathcal{H}_0' : \beta_* = 0$

We first examine the asymptotic null approximation of the QLR test under $\mathcal{H}_0' : \beta_* = 0$. As γ_* is not identified under $\mathcal{H}_0'$, we approximate the model with respect to the other parameters and treat γ as an unidentified parameter as in Davies (1977, 1987). For notational simplicity, let the quasi-likelihood (QL) and concentrated QL (CQL) be denoted as

$$L_n(\alpha, \beta, \gamma, \boldsymbol{\delta}) := -\sum_{t=1}^n (Y_t - \alpha - \beta X_t^{\gamma} - \mathbf{W}_t'\boldsymbol{\delta})^2 \quad \text{and} \quad L_n(\beta; \gamma) := L_n(\hat{\alpha}_n(\beta;\gamma), \beta, \gamma, \hat{\boldsymbol{\delta}}_n(\beta;\gamma)),$$

respectively, where $(\hat{\alpha}_n(\beta;\gamma), \hat{\boldsymbol{\delta}}_n(\beta;\gamma)')' := \arg\max_{\alpha,\boldsymbol{\delta}} L_n(\alpha,\beta,\gamma,\boldsymbol{\delta})$. Since γ_* is unidentified under $\mathcal{H}_0'$, we fix γ and maximize the QL function with respect to $(\alpha, \boldsymbol{\delta})$. The resulting CQL has the specific form

$$L_n(\beta;\gamma) = -\{\mathbf{Y} - \beta\mathbf{X}(\gamma)\}'\mathbf{M}\{\mathbf{Y} - \beta\mathbf{X}(\gamma)\}, \tag{3}$$

where $\mathbf{Y} := (Y_1, Y_2, \ldots, Y_n)'$, $\mathbf{M} := \mathbf{I}_n - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$, $\mathbf{X}(\gamma) := (X_1^\gamma \ldots X_n^\gamma)'$, $\mathbf{Z} := [\mathbf{Z}_1', \ldots, \mathbf{Z}_n']'$ with $\mathbf{Z}_t := [1, \mathbf{W}_t']'$. Under $\mathcal{H}_0$, $\mathbf{MY} = \mathbf{MU}$ and $\mathbf{U} := (U_1, U_2, \ldots, U_n)'$ so that, for each t, $U_t := Y_t - \mathbb{E}[Y_t|\mathbf{Z}_t]$, and we suppose that $\{U_t, \mathcal{F}_t\}$ is a martingale difference sequence (MDS), where $\mathcal{F}_t$ is the adapted smallest σ-field generated by $\{\mathbf{Z}_{t+1}, U_t, \mathbf{Z}_t, U_{t-1}, \cdots\}$. We can maximize the CQL with respect to β, giving

$$\sup_\beta\{L_n(\beta;\gamma) - L_n(0;\gamma)\} = \frac{\{\mathbf{X}(\gamma)'\mathbf{MU}\}^2}{\mathbf{X}(\gamma)'\mathbf{MX}(\gamma)},$$

which leads to the following QLR statistic

$$QLR_n^{(\beta=0)} := \sup_\gamma \sup_\beta n\left\{1 - \frac{L_n(\beta;\gamma)}{L_n(0;\gamma)}\right\} = \sup_\gamma \frac{\{\mathbf{X}(\gamma)'\mathbf{MU}\}^2}{\hat{\sigma}_{n,0}^2\mathbf{X}(\gamma)'\mathbf{MX}(\gamma)}. \tag{4}$$

The given approximation is asymptotically bounded in probability under some mild regularity conditions. Specifically, under Assumption 2 below, the functional central limit theorem (FCLT) and uniform law of large numbers (ULLN) can be applied to $n^{-1/2}\mathbf{X}(\cdot)'\mathbf{MU}$ and $n^{-1}\mathbf{X}(\cdot)'\mathbf{MX}(\cdot)$, respectively, so both are bounded in probability. That is, $n^{-1/2}\mathbf{X}(\cdot)'\mathbf{MU} \Rightarrow \mathcal{G}(\cdot)$ and

$$\sup_{\gamma\in\boldsymbol{\Gamma}} |n^{-1}\mathbf{X}(\gamma)'\mathbf{MX}(\gamma) - (\mathbb{E}[X_t^{2\gamma}] - \mathbb{E}[X_t^\gamma\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t']^{-1}\mathbb{E}[\mathbf{Z}_tX_t^\gamma])| = o_p(1),$$

where $\mathcal{G}(\cdot)$ is a Gaussian process such that for each $\gamma, \gamma' \in \boldsymbol{\Gamma}$, $\mathbb{E}[\mathcal{G}(\gamma)] = 0$ and

$$\begin{aligned}\mathbb{E}[\mathcal{G}(\gamma)\mathcal{G}(\gamma')] =& \mathbb{E}[U_t^2X_t^{\gamma+\gamma'}] - \mathbb{E}[U_t^2X_t^\gamma\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t']^{-1}\mathbb{E}[\mathbf{Z}_tX_t^{\gamma'}] - \mathbb{E}[U_t^2X_t^{\gamma'}\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t']^{-1}\mathbb{E}[\mathbf{Z}_tX_t^\gamma] \\ &+ \mathbb{E}[X_t^\gamma\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t']^{-1}\mathbb{E}[U_t^2\mathbf{Z}_t\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t']^{-1}\mathbb{E}[\mathbf{Z}_tX_t^{\gamma'}] =: \kappa\left(\gamma, \gamma'\right).\end{aligned}$$

This Gaussian process $\mathcal{G}(\cdot)$ is particularly interesting as its sample path is smooth almost surely, a property that affects later results and inference. The covariance kernel $\kappa(\gamma, \gamma')$ is composed of analytic functions under mild moment conditions that assure use of dominated convergence, as given below, so it is smoothly second-order differentiable. This feature is important when obtaining the asymptotic null distribution of the QLR test.

The relatively simple covariance kernel $\kappa(\gamma, \gamma')$ is obtained because U_t is an MDS. If U_t exhibits conditional homoskedasticity with $\mathbb{E}\left(U_t^2|\mathcal{F}_{t-1}\right) = \sigma_*^2$, the covariance structure further simplifies to

$$\mathbb{E}[\mathcal{G}(\gamma)\mathcal{G}(\gamma')] = \sigma_*^2\{\mathbb{E}[X_t^{\gamma+\gamma'}] - \mathbb{E}[X_t^{\gamma}\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t']^{-1}\mathbb{E}[\mathbf{Z}_t X_t^{\gamma'}]\}.$$

Nonetheless, if $\gamma = 0$ or 1, $\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U} \equiv 0$ and $\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma) \equiv 0$ by the definition of $\mathbf{M}$, and so $QLR_n^{(\beta=0)}$ may not be bounded in probability under $\mathcal{H}_{01}$. For the moment, therefore, it is convenient to bound the parameter space of γ away from zero and unity by setting $\mathbf{\Gamma}(\epsilon) := \mathbf{\Gamma} \setminus ((-\epsilon, \epsilon) \cup (1-\epsilon, 1+\epsilon))$ and to redefine the QLR test as

$$QLR_n^{(\beta=0)}(\epsilon) := \sup_{\gamma \in \mathbf{\Gamma}(\epsilon)} \sup_{\beta} n\left\{1 - \frac{L_n(\gamma;\beta)}{L_n(0;\beta)}\right\} = \sup_{\gamma \in \mathbf{\Gamma}(\epsilon)} \frac{\{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}^2}{\hat{\sigma}_{n,0}^2 \mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma)}. \tag{5}$$

The statistic $QLR_n^{(\beta=0)}(\epsilon)$ is now bounded in probability under $\mathcal{H}_{01}$. Later in this Section we will consider behavior at the limits of the domain of definition as $\epsilon \to 0$. The following assumption provides sufficient conditions for the application of the FCLT and ULLN.

Assumption 2. *(i) For each $\epsilon > 0$, the following square matrices are positive definite uniformly on $\mathbf{\Gamma}(\epsilon)$:*

$$\begin{bmatrix} \mathbb{E}[X_t^{2\gamma}] & \mathbb{E}[X_t^{\gamma}\mathbf{Z}_t'] \\ \mathbb{E}[X_t^{\gamma}\mathbf{Z}_t] & \mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'] \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} \mathbb{E}[U_t^2 X_t^{2\gamma}] & \mathbb{E}[U_t^2 X_t^{\gamma}\mathbf{Z}_t'] \\ \mathbb{E}[U_t^2 X_t^{\gamma}\mathbf{Z}_t] & \mathbb{E}[U_t^2\mathbf{Z}_t\mathbf{Z}_t'] \end{bmatrix};$$

(ii) $\{U_t, \mathcal{F}_t\}$ is an MDS, where $\mathcal{F}_t$ is the adapted smallest σ-field generated by $\{\mathbf{Z}_{t+1}, U_t, \mathbf{Z}_t, U_{t-1}, \cdots\}$;

(iii) There is a strictly stationary and ergodic sequence $\{M_t\}$ such that $|U_t| \le M_t$, and for $j = 1, 2, \cdots, k+1$, $|D_{t,j}| \le M_t$ such that $\mathbb{E}[M_t^{4r}] < \infty$, where r is the same r as that defined in Assumption 1 and $D_{t,j}$ is the j-th row element of $\mathbf{D}_t$; and

(iv) $\sup_{\gamma \in \mathbf{\Gamma}} |X_t^{\gamma}| \le M_t$ and $\sup_{\gamma \in \mathbf{\Gamma}} |X_t^{\gamma}\log(X_t)| \le M_t$. □

Assumption 2 is imposed mainly to apply the FCLT and ULLN. Although Assumption 2(*i, ii*) is standard, Assumption 2(*iii, iv*) is different from standard conditions. This condition is imposed to show that $n^{-1/2}\mathbf{X}(\cdot)'\mathbf{M}\mathbf{U}$ is tight under the mixing condition given in Assumption 1 mainly by using the arguments of Doukhan, Massart, and Rio (1995). Furthermore, $n^{-1}\mathbf{X}(\cdot)'\mathbf{M}\mathbf{X}(\cdot)$ obeys the ULLN by the moment condition in Assumption 2(*iv*).

The main result of this subsection now follows.

Theorem 1. *Given Assumptions 1, 2, and $\mathcal{H}_0'$, for each $\epsilon > 0$,*

(i) $QLR_n^{(\beta=0)}(\epsilon) = \sup_{\gamma \in \mathbf{\Gamma}(\epsilon)} \{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}^2 / \{\hat{\sigma}_{n,0}^2 \mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma)\}$; and

(ii) $QLR_n^{(\beta=0)}(\epsilon) \Rightarrow \sup_{\gamma\in\mathbf{\Gamma}(\epsilon)} \mathcal{Z}(\gamma)^2$*, where for each* $\gamma \in \mathbf{\Gamma}(\epsilon)$*,* $\mathcal{Z}(\gamma) \sim N(0,1)$*, and for each pair* (γ, γ')

$$\mathbb{E}[\mathcal{Z}(\gamma)\mathcal{Z}(\gamma')] = \frac{\mathbb{E}[\mathcal{G}(\gamma)\mathcal{G}(\gamma')]}{\sqrt{\sigma^2(\gamma,\gamma)}\sqrt{\sigma^2(\gamma',\gamma')}}$$

with $\sigma^2(\gamma,\gamma) := \sigma_*^2(\mathbb{E}[X_t^{2\gamma}] - \mathbb{E}[X_t^{\gamma}\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t']^{-1}\mathbb{E}[\mathbf{Z}_t X_t^{\gamma}])$. □

2.3 QLR Statistic under $\mathcal{H}_0'' : \gamma_* = 0$

We next develop the asymptotic null approximation under $\mathcal{H}_0'' : \gamma_* = 0$. As mentioned earlier, if $\gamma_* = 0$, α_* and β_* are not separately identified although the combined parameter $\alpha_* + \beta_*$ is identified. To resolve this difficulty, our discussion proceeds in two ways. First, we may fix β, identify α_*, and obtain the asymptotic null approximation. Alternatively, we may fix α and identify β_*. We examine each case separately in what follows.

2.3.1 When β_* Is Not Identified

First fix β, approximate the CQL with respect to $(\alpha, \boldsymbol{\delta})$ as before, and then optimize the CQL with respect to β in the final step. For this purpose, define the CQL

$$L_n(\gamma;\beta) := L_n(\hat{\alpha}_n(\gamma;\beta), \beta, \gamma, \hat{\boldsymbol{\delta}}_n(\gamma;\beta)),$$

where $(\hat{\alpha}_n(\gamma;\beta), \hat{\boldsymbol{\delta}}_n(\gamma;\beta)')' := \arg\max_{\alpha,\boldsymbol{\delta}} L_n(\alpha,\beta,\gamma,\boldsymbol{\delta})$. This CQL differs slightly from the CQL given as $L_n(\beta;\gamma)$. Here we view $L_n(\cdot;\beta)$ as a function defined on $\mathbf{\Gamma}$, whereas $L_n(\cdot;\gamma)$ is defined on $\boldsymbol{B}$. We can write $L_n(\gamma;\beta)$ explicitly as

$$L_n(\gamma;\beta) = -\{\mathbf{Y} - \beta\mathbf{X}(\gamma)\}'\mathbf{M}\{\mathbf{Y} - \beta\mathbf{X}(\gamma)\}. \tag{6}$$

The right side (6) is identical to the right side of (3) although the treatment of the two arguments is different. Specifically, the nuisance parameter of (3) is γ, while the nuisance parameter of (6) is β.

Applying a second-order Taylor expansion to this function and optimizing with respect to γ, we have

$$\sup_{\gamma}\{L_n(\gamma;\beta) - L_n(0;\beta)\} = -\frac{\{L_n^{(1)}(0;\beta)\}^2}{2L_n^{(2)}(0;\beta)} + o_p(1) = \frac{\{\beta\mathbf{L}_1'\mathbf{M}\mathbf{U}\}^2}{\beta^2\mathbf{L}_1'\mathbf{M}\mathbf{L}_1 - \beta\mathbf{L}_2'\mathbf{M}\mathbf{U}} + o_p(1), \tag{7}$$

where $L_n^{(1)}(0;\beta) := (d/d\gamma)L_n^{(1)}(0;\beta) = 2\beta\mathbf{L}_1'\mathbf{M}\mathbf{U}$, $L_n^{(2)}(0;\beta) := (d^2/d\gamma^2)L_n^{(1)}(0;\beta) = 2\beta\mathbf{L}_2'\mathbf{M}\mathbf{U} - 2\beta^2\mathbf{L}_1'\mathbf{M}\mathbf{L}_1$, $\mathbf{L}_1 := [\log X_1, \ldots, \log X_n]'$, and $\mathbf{L}_2 := [\log^2(X_1), \ldots, \log^2(X_n)]'$.

Although the approximation (7) is a consequence of a conventional second-order approximation, it differs from those in the ANN literature. Importantly, $L_n^{(1)}(0;\beta)$ is not necessarily equal to zero and we can apply a central limit theorem (CLT) to this derivative. In the ANN literature, it is common to have zero first-order derivatives, so that higher-order approximations are needed for model approximations (e.g., Cho, Ishida, and White, 2011, 2013; and White and Cho, 2012). This difference mainly arises because the nonlinear functions in Cho, Ishida, and White (2011, 2013) and White and Cho (2012) have nuisance parameters that are multiplicative to X_t, whereas in the present case the nuisance parameter enters through the power coefficient.

We approximate the QLR test using a second-order Taylor expansion. Specifically, we note that the left side of (7) is free of β by scaling up to this order of approximation, provided $\mathbf{L}_2'\mathbf{MU} = o_p(n)$, which readily holds under mild regularity conditions such as $\mathbb{E}[\log^2(X_t)U_t] = 0$ and $\mathbb{E}[\mathbf{Z}_t U_t] = \mathbf{0}$ by virtue of the MDS property of $\{U_t, \mathcal{F}_t\}$. Under this condition, the right side of (7) simplifies, giving the following asymptotic approximation of the QLR statistic

$$QLR_n^{(\gamma=0;\beta)} := \sup_{\beta}\sup_{\gamma} n\left\{1 - \frac{L_n(\gamma;\beta)}{L_n(0;\beta)}\right\} = \frac{\{n^{-1/2}\mathbf{L}_1'\mathbf{MU}\}^2}{\hat{\sigma}_{n,0}^2\{n^{-1}\mathbf{L}_1'\mathbf{ML}_1\}} + o_p(1). \tag{8}$$

Here, we used the fact that $L_n(0;\beta) = -n\hat{\sigma}_{n,0}^2$. Applying a CLT and the ergodic theorem to the numerator and denominator, we find that $QLR_n^{(\gamma=0)}$ weakly converges to a noncentral chi-squared variate. For this purpose, we impose the following conditions.

Assumption 3. *(i) The following square matrices are positive definite:*

$$\begin{bmatrix} \mathbb{E}[\log^2(X_t)] & \mathbb{E}[\log(X_t)\mathbf{Z}_t'] \\ \mathbb{E}[\log(X_t)\mathbf{Z}_t] & \mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'] \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} \mathbb{E}[U_t^2\log^2(X_t)] & \mathbb{E}[U_t^2\log(X_t)\mathbf{Z}_t'] \\ \mathbb{E}[U_t^2\log(X_t)\mathbf{Z}_t] & \mathbb{E}[U_t^2\mathbf{Z}_t\mathbf{Z}_t'] \end{bmatrix};$$

(ii) $\{U_t, \mathcal{F}_t\}$ is an MDS, where $\mathcal{F}_t$ is the adapted smallest σ-field generated by $\{\mathbf{Z}_{t+1}, U_t, \mathbf{Z}_t, U_{t-1}, \cdots\}$;

(iii) There is a strictly stationary and ergodic sequence $\{M_t\}$ such that, for $j = 1, 2, \cdots, k+1$, $|W_{t,j}| \leq M_t$, $|U_t| \leq M_t$, and $|\log(X_t)| \leq M_t$ with $\mathbb{E}[M_t^4] < \infty$, where $W_{t,j}$ is the j-th row element of $\mathbf{W}_t$. □

The following lemma formalizes the result.

Lemma 1. *Given Assumptions 1, 3, and $\mathcal{H}_0''$,*

(i) $QLR_n^{(\gamma=0;\beta)} = \{\mathbf{L}_1'\mathbf{MU}\}^2/\{\hat{\sigma}_{n,0}^2(\mathbf{L}_1'\mathbf{ML}_1)\} + o_p(1)$; and

(ii) $QLR_n^{(\gamma=0;\beta)} = O_p(1)$. □

2.3.2 When α_* Is Not Identified

The model can be identified in another way when $\gamma_* = 0$. In this case, we can fix α and identify $(\beta_*, \boldsymbol{\delta}_*)$. For this purpose, let $(\hat{\beta}_n(\gamma;\alpha), \hat{\boldsymbol{\delta}}_n(\gamma;\alpha)')' := \arg\max_{\beta,\boldsymbol{\delta}} L_n(\alpha,\beta,\gamma,\boldsymbol{\delta})$, whose specific form is

$$[\hat{\beta}_n(\gamma;\alpha), \hat{\boldsymbol{\delta}}_n(\gamma;\alpha)']' = [\mathbf{Q}(\gamma)'\mathbf{Q}(\gamma)]^{-1}\mathbf{Q}(\gamma)'\mathbf{P}(\alpha),$$

and obtain the CQL as

$$L_n(\gamma;\alpha) := L_n(\alpha, \hat{\beta}_n(\gamma;\alpha), \gamma, \hat{\boldsymbol{\delta}}_n(\gamma;\alpha)) = -\mathbf{P}(\alpha)'[\mathbf{I} - \mathbf{Q}(\gamma)[\mathbf{Q}(\gamma)'\mathbf{Q}(\gamma)]^{-1}\mathbf{Q}(\gamma)']\mathbf{P}(\alpha),$$

where $\mathbf{P}(\alpha) := \mathbf{Y} - \alpha\boldsymbol{\iota}$, $\mathbf{Q}(\gamma) := [\mathbf{X}(\gamma) \vdots \mathbf{W}]$, and $\boldsymbol{\iota}$ is the $n \times 1$ vector of ones.

We approximate this CQL function with respect to γ at $\gamma_* = 0$. Define $\mathbf{K}_j := [\mathbf{L}_j \vdots \mathbf{0}_{n\times k}]$ for each $j = 1, 2, \ldots, k+1$. The first two derivatives are given in the following lemma, whose proof is in the Appendix.

Lemma 2. *Given Assumptions 1, 3, and $\mathcal{H}_0''$,*

(i) $L_n^{(1)}(0;\alpha) = 2(\alpha_* - \alpha)\mathbf{L}_1'\mathbf{M}\mathbf{U} + 2\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - \mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$*;*

(ii) $L_n^{(1)}(0;\alpha) = 2(\alpha_* - \alpha)\mathbf{L}_1'\mathbf{M}\mathbf{U} + o_p(\sqrt{n})$*; and*

(iii) $L_n^{(2)}(0;\alpha) = -2(\alpha_* - \alpha)^2\mathbf{L}_1'\mathbf{M}\mathbf{L}_1 + o_p(n)$. □

The derivation of Lemma 2(i) involves some algebra but the result simplifies because the last two terms on the right side of (*i*) are $O_p(1)$ by virtue of the regularity conditions in Assumption 3. Then Lemma 2(ii) follows directly.

Lemma 2 provides the components needed for the second-order expansion. Combining these components delivers the QLR test as follows:

$$\sup_{\gamma}\{L_n(\gamma;\alpha) - L_n(0;\alpha)\} = -\frac{\{L_n^{(1)}(0;\alpha)\}^2}{2L_n^{(2)}(0;\alpha)} + o_p(1) = \frac{\{2(\alpha_* - \alpha)n^{-1/2}\mathbf{L}_1'\mathbf{M}\mathbf{U}\}^2}{4(\alpha_* - \alpha)^2 n^{-1}\mathbf{L}_1'\mathbf{M}\mathbf{L}_1} + o_p(1).$$

The unidentified parameter α is cancelled just as in the previous subsection, so that the QLR test is approximated as

$$QLR_n^{(\gamma=0;\alpha)} = \sup_{\alpha}\sup_{\gamma} n\left\{1 - \frac{L_n(\gamma;\alpha)}{L_n(0;\alpha)}\right\} = \frac{\left\{n^{-1/2}\mathbf{L}_1'\mathbf{M}\mathbf{U}\right\}^2}{\hat{\sigma}_{n,0}^2\{n^{-1}\mathbf{L}_1'\mathbf{M}\mathbf{L}_1\}} + o_p(1). \tag{9}$$

The result is formalized in the following lemma.

Lemma 3. *Given Assumptions 1, 3, and $\mathcal{H}_0''$,* *(i)* $QLR_n^{(\gamma=0;\alpha)} = \{\mathbf{L}_1'\mathbf{MU}\}^2/\{\hat{\sigma}_{n,0}^2(\mathbf{L}_1'\mathbf{ML}_1)\} + o_p(1)$*; and*

(ii) $QLR_n^{(\gamma=0;\alpha)} = O_p(1)$. □

Combining Lemmas 2 and 3 has implications for the QLR test under $\mathcal{H}_0'' : \gamma_* = 0$. In particular, the asymptotics obtained by first fixing β are identical to those obtained by first fixing α. From this equivalence we conclude that these different identification problems yield the same asymptotic approximation. The result is stated formally in the following theorem.

Theorem 2. *Given Assumptions 1, 3, and* $\mathcal{H}_0''$*,*

(i) $QLR_n^{(\gamma=0)} = \{\mathbf{L}_1'\mathbf{MU}\}^2/\{\hat{\sigma}_{n,0}^2(\mathbf{L}_1'\mathbf{ML}_1)\} + o_p(1)$ *under* $\mathcal{H}_0'' : \gamma_* = 0$*, where* $QLR_n^{(\gamma=0)}$ *denotes the QLR statistic testing* $\mathcal{H}_0'' : \gamma_* = 0$*; and*

(ii) $QLR_n^{(\gamma=0)} = O_p(1)$. □

The asymptotic null approximation of the QLR test is driven by $\mathbf{L}_1$, a feature that is intuitively associated with the Box-Cox transformation. Passing the parameter of the Box-Cox transform to zero gives

$$\frac{d}{d\gamma}X_t^\gamma\bigg|_{\gamma=0} = \lim_{\gamma\to 0}\frac{X_t^\gamma - X_t^0}{\gamma - 0} = \lim_{\gamma\to 0}\frac{X_t^\gamma - 1}{\gamma} = \log X_t.$$

Thus, the Box-Cox transform with $\gamma = 0$ is associated with the first-order derivative which forms the primary component constituting the score of the QLR test. Additionally, the Box-Cox transform approximates

$$\mathbb{E}[Y_t|\mathbf{Z}_t] = (\alpha_* + \beta_*) + \xi_* X_t + \mathbf{D}_t'\boldsymbol{\eta}_* + \beta_*\gamma_*\frac{(X_t^{\gamma_*} - 1)}{\gamma_*}$$

by $\alpha_* + \xi_* X_t + \mathbf{D}_t'\boldsymbol{\eta}_* + \beta_*\gamma_*\log(X_t)$ when γ_* is sufficiently close to zero. For such a case, $\mathbf{L}_1'\mathbf{MU}$ is the primary score of standard statistics obtained under the null that $\beta_*\gamma_*$ is zero. This also implies that the Box-Cox transformation can be understood as an alternative to the constant function hypothesis in the context of the QLR test.

2.4 QLR Statistic under $\mathcal{H}_0''' : \gamma_* = 1$

We repeat the procedure to obtain the asymptotic null approximation under $\gamma_* = 1$. If $\gamma_* = 1$, ξ_* and β_* are not separately identified although the sum $\xi_* + \beta_*$ is identified. The reason is that under $\mathcal{H}_0''' : \gamma_* = 1$, $Y_t = \alpha_* + \mathbf{D}_t'\boldsymbol{\eta}_* + (\xi_* + \beta_*)X_t + U_t$, so that $\xi_* + \beta_*$ exists as the identifiable coefficient of X_t.

We first fix β to identify ξ_* and obtain the null approximation. Alternatively, we can fix ξ and identify β_* to obtain the null approximation. These two different approximations are separately examined in the following subsections.

2.4.1 When β_* Is Not Identified

The procedure to obtain the asymptotic approximation is similar to that shown in Section 2.3. As β_* is not identified, we first fix β at some particular value and concentrate the QL with respect to $(\alpha, \boldsymbol{\delta}')'$. The CQL obtained in this way is already given in (6). We now expand (6) with respect to γ around $\gamma_* = 1$ using a second-order expansion and optimize with respect to γ, leading to

$$\sup_{\gamma}\{L_n(\gamma;\beta) - L_n(1;\beta)\} = -\frac{\{L_n^{(1)}(1;\beta)\}^2}{2L_n^{(2)}(1;\beta)} + o_p(1) = \frac{\beta^2\left\{\mathbf{C}_1'\mathbf{M}\mathbf{U}\right\}^2}{\beta^2\mathbf{C}_1'\mathbf{M}\mathbf{C}_1 - \beta\mathbf{C}_2'\mathbf{M}\mathbf{U}} + o_p(1),$$

where $\mathbf{C}_1 := [X_1\log(X_1), \ldots, X_n\log(X_n)]'$, and $\mathbf{C}_2 := [X_1\log^2(X_1), \ldots X_n\log^2(X_n)]'$. Under conditions for which $\mathbf{C}_2'\mathbf{M}\mathbf{U} = o_p(n)$, it follows trivially that

$$\sup_{\gamma}\{L_n(\gamma;\beta) - L_n(1;\beta)\} = \frac{\{n^{-1/2}\mathbf{C}_1'\mathbf{M}\mathbf{U}\}^2}{n^{-1}\mathbf{C}_1'\mathbf{M}\mathbf{C}_1} + o_p(1).$$

The next assumption provides regularity assumptions for this result to hold.

Assumption 4. *(i) The following square matrices are positive definite:*

$$\begin{bmatrix} \mathbb{E}[X_t^2\log^2(X_t)] & \mathbb{E}[X_t\log(X_t)\mathbf{Z}_t'] \\ \mathbb{E}[X_t\log(X_t)\mathbf{Z}_t] & \mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'] \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} \mathbb{E}[U_t^2X_t^2\log^2(X_t)] & \mathbb{E}[U_t^2X_t\log(X_t)\mathbf{Z}_t'] \\ \mathbb{E}[U_t^2X_t\log(X_t)\mathbf{Z}_t] & \mathbb{E}[U_t^2\mathbf{Z}_t\mathbf{Z}_t'] \end{bmatrix};$$

(ii) $\{U_t, \mathcal{F}_t\}$ *is an MDS, where* $\mathcal{F}_t$ *is the adapted smallest* σ*-field generated by* $\{\mathbf{Z}_{t+1}, U_t, \mathbf{Z}_t, U_{t-1}, \cdots\}$;

(iii) There is a strictly stationary and ergodic sequence $\{M_t, S_t\}$ *such that for* $j = 1, 2, \cdots, k+1$, $|D_{t,j}| \leq M_t$, $\mathbb{E}[M_t^4] < \infty$, $\mathbb{E}[S_t^8] < \infty$, *and*

(iii.a) $|U_t| \leq M_t$, $|X_t| \leq S_t$, *and* $|\log[X_t]| \leq S_t$;

(iii.b) $|X_t| \leq M_t$, $|U_t| \leq S_t$, *and* $|\log[X_t]| \leq S_t$; *or*

(iii.c) $|\log[X_t]| \leq M_t$, $|X_t| \leq S_t$, *and* $|U_t| \leq S_t$. □

The moment condition in Assumption 4(*iii*) is stronger than that of Assumption 3(*iii*). Hence, properties implied by Assumption 3(*iii*) continue to apply under Assumption 4(*iii*). Note that the moment condition in Assumption 4(*iii.a*) does not imply Assumption 4(*iii.b*) or vice versa. If at least one of these separate conditions holds, however, we can obtain the desired results given in Lemma 4 below.

Using these regularity condition, we obtain the following approximation of the QLR statistic:

$$QLR_n^{(\gamma=1;\beta)} := \sup_{\beta}\sup_{\gamma}\sup_{\alpha,\boldsymbol{\delta}} n\left\{1 - \frac{L_n(\alpha,\beta,\gamma,\boldsymbol{\delta})}{L_n(1;\beta)}\right\} = \frac{\left\{n^{-1/2}\mathbf{C}_1'\mathbf{M}\mathbf{U}\right\}^2}{\hat{\sigma}_{n,0}^2\{n^{-1}\mathbf{C}_1'\mathbf{M}\mathbf{C}_1\}} + o_p(1).$$

Asymptotic results are summarized as follows.

Lemma 4. *Given Assumptions 1, 4, and $\mathcal{H}_0'''$,*

(i) $QLR_n^{(\gamma=1;\beta)} = \{\mathbf{C}_1'\mathbf{M}\mathbf{U}\}^2/\{\hat{\sigma}_{n,0}^2(\mathbf{C}_1'\mathbf{M}\mathbf{C}_1)\} + o_p(1)$*; and*

(ii) $QLR_n^{(\gamma=1;\beta)} = O_p(1)$. □

2.4.2 When ξ_* Is Not Identified

We now reverse the plan of identification. We first fix ξ and identify the other parameters $(\alpha_*, \beta_*, \boldsymbol{\eta}_*')'$. For notational simplicity, let $\boldsymbol{\theta} := (\beta, \boldsymbol{\eta}')'$ and $\mathbf{S}_t(\gamma) := (X_t^\gamma, \mathbf{D}_t')'$, so that $\boldsymbol{\theta}_* := (\beta_*, \boldsymbol{\eta}_*')'$. We regress $Y_t - \xi X_t$ on $(1, \mathbf{S}_t(\gamma)')$, using the following model

$$Y_t - \xi X_t = \alpha + \mathbf{S}_t(\gamma)'\boldsymbol{\theta} + U_t,$$

which yields the estimater $[\widehat{\alpha}_n(\gamma;\xi), \widehat{\boldsymbol{\theta}}_n(\gamma;\xi)']' := [\widetilde{\mathbf{Q}}(\gamma)'\widetilde{\mathbf{Q}}(\gamma)]^{-1}\widetilde{\mathbf{Q}}(\gamma)'\widetilde{\mathbf{P}}(\xi)$, where $\widetilde{\mathbf{P}}(\xi) := \mathbf{Y} - \xi\mathbf{X}$, $\widetilde{\mathbf{Q}}(\gamma) := [\boldsymbol{\iota} \vdots \mathbf{S}(\gamma)]$, $\mathbf{X} := (X_1, \ldots, X_n)'$, and $\mathbf{S}(\gamma) := [\mathbf{S}_1(\gamma), \ldots, \mathbf{S}_n(\gamma)]'$. The CQL then follows as

$$L_n(\gamma;\xi) := L_n(\widehat{\alpha}_n(\gamma;\xi), \widehat{\boldsymbol{\theta}}_n(\gamma;\xi), \gamma, \xi) = -\widetilde{\mathbf{P}}(\xi)'[\mathbf{I} - \widetilde{\mathbf{Q}}(\gamma)[\widetilde{\mathbf{Q}}(\gamma)'\widetilde{\mathbf{Q}}(\gamma)]^{-1}\widetilde{\mathbf{Q}}(\gamma)']\widetilde{\mathbf{P}}(\xi).$$

We again approximate the CQL with respect to γ at $\gamma_* = 1$ by means of a second-order approximation. Define the $n \times (k+2)$ matrix $\mathbf{J}_j := [\mathbf{0}_{n\times 1} \vdots \mathbf{C}_j \vdots \mathbf{0}_{n\times k}]$ for $j = 1, 2, \ldots k+1$. The following lemma, which is proved in the Appendix, provides the first two derivatives.

Lemma 5. *Given Assumptions 1, 4, and $\mathcal{H}_0'''$,*

(i) $L_n^{(1)}(1;\xi) = 2(\xi_* - \xi)\mathbf{C}_1'\mathbf{M}\mathbf{U} + 2\mathbf{U}'\mathbf{J}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - \mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{J}_1 + \mathbf{J}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$*;*

(ii) $L_n^{(1)}(1;\xi) = 2(\xi_* - \xi)\mathbf{C}_1'\mathbf{M}\mathbf{U} + o_p(\sqrt{n})$*; and*

(iii) $L_n^{(2)}(1;\xi) = -2(\xi_* - \xi)^2\mathbf{C}_1'\mathbf{M}\mathbf{C}_1 + o_p(n)$. □

As before, the first-order derivative is not necessarily equal to zero, which implies that we can approximate the CQL by a second-order Taylor expansion. Using the components provided in Lemma 5, we obtain the following expansion:

$$\sup_{\gamma}\{L_n(\gamma;\xi) - L_n(1;\xi)\} = -\frac{\{L_n^{(1)}(1;\xi)\}^2}{2L_n^{(2)}(1;\xi)} + o_p(1) = \frac{\{2(\xi_* - \xi)n^{-1/2}\mathbf{C}_1'\mathbf{M}\mathbf{U}\}^2}{4(\xi_* - \xi)^2 n^{-1}\mathbf{C}_1'\mathbf{M}\mathbf{C}_1} + o_p(1).$$

The unidentified parameter ξ again asymptotically cancels and the asymptotic approximation of the QLR test is simply

$$QLR_n^{(\gamma=1,\xi)} = \sup_{\xi}\sup_{\gamma} n\left\{1-\frac{L_n(\gamma;\xi)}{L_n(1;\xi)}\right\} = \frac{\{n^{-1/2}\mathbf{C}_1'\mathbf{MU}\}^2}{\widehat{\sigma}_{n,0}^2\{n^{-1}\mathbf{C}_1'\mathbf{MC}_1\}} + o_p(1).$$

We formalize this result in the next lemma.

Lemma 6. *Given Assumptions 1, 4, and* $\mathcal{H}_0'''$*,*

(i) $QLR_n^{(\gamma=1;\xi)} = \{\mathbf{C}_1'\mathbf{MU}\}^2/\{\hat{\sigma}_{n,0}^2(\mathbf{C}_1'\mathbf{MC}_1)\} + o_p(1)$*; and*

(ii) $QLR_n^{(\gamma=1;\xi)} = O_p(1)$. □

Combining Lemmas 5 and 6 has useful implications for the QLR test under $\mathcal{H}_0''' : \gamma_* = 1$. As in the earlier case, the asymptotics obtained by first fixing β are identical to the asymptotics obtained by first fixing ξ. So, these different identification problems yield the same asymptotic approximation. The result is stated formally in the following theorem.

Theorem 3. *Given Assumptions 1, 4, and* $\mathcal{H}_0'''$*,*

(i) $QLR_n^{(\gamma=1)} = \{\mathbf{C}_1'\mathbf{MU}\}^2/\{\hat{\sigma}_{n,0}^2(\mathbf{C}_1'\mathbf{MC}_1)\} + o_p(1)$*, where* $QLR_n^{(\gamma=1)}$ *denotes the QLR statistic testing* $\mathcal{H}_0''' : \gamma_* = 1$*; and*

(ii) $QLR_n^{(\gamma=1)} = O_p(1)$. □

The asymptotic null distribution is driven by $\mathbf{C}_1$ and, as before, this link can be associated with the Box-Cox transformation. In particular

$$\frac{d}{d\gamma}X_t^{\gamma}\bigg|_{\gamma=1} = \lim_{\gamma\to 1}\frac{X_t^{\gamma}-X_t}{\gamma-1}.$$

So modifying the Box-Cox transform as

$$ABC_t(\gamma) := \begin{cases} (X_t^{\gamma}-X_t)/(\gamma-1), & \text{if } \gamma \neq 1; \\ X_t\log[X_t], & \text{if } \gamma = 1, \end{cases}$$

we see that $X_t\log(X_t)$ is the typical element of $\mathbf{C}_1$, implying that $QLR_n^{(\gamma=1)}$ effectively tests the above transformation, giving an interpretation of the test in terms of the Box-Cox transformation. That is, when γ_* is believed to be sufficiently close to one in the data generating process with conditional mean function

$$\mathbb{E}[Y_t|\mathbf{Z}_t] = \alpha_* + (\xi_* + \beta_*)X_t + \mathbf{D}_t'\boldsymbol{\eta}_* + \beta_*(\gamma_*-1)\frac{(X_t^{\gamma_*}-X_t)}{\gamma_*-1},$$

the augmented Box-Cox transformation approximates the mean function by $\alpha_* + (\xi_* + \beta_*)X_t + \mathbf{D}_t'\boldsymbol{\eta}_* + \beta_*(\gamma_* - 1)X_t \log(X_t)$. For such a case, the primary score of standard statistics is constructed using $\mathbf{C}_1'\mathbf{MU}$ under the null hypothesis that $\beta_*(\gamma_* - 1)$ is zero. This approximation also implies that the augmented Box-Cox transformation can be understood as an alternative to the linear function hypothesis in the context of the QLR test.

2.5 Interrelationships of the QLR Statistics under $\mathcal{H}_0$

The separate weak limits obtained in the previous subsections are not independent. The stochastic relationships can be studied by letting γ converge to zero and unity in the test components studied in subsections 2.2, 2.3, and 2.4. To wit, define $N_n(\gamma)$ and $D_n(\gamma)$ as

$$N_n(\gamma) := \{\mathbf{X}(\gamma)'\mathbf{MU}\}^2 \quad \text{and} \quad D_n(\gamma) := \widehat{\sigma}_{n,0}^2 \mathbf{X}(\gamma)'\mathbf{MX}(\gamma),$$

representing the numerator and denominator of the right side of (4), respectively. First, consider the case where $\gamma \to 0$, which shows that

$$\operatorname*{plim}_{\gamma\to 0} N_n(\gamma) = 0 \quad \text{and} \quad \operatorname*{plim}_{\gamma\to 0} D_n(\gamma) = 0$$

because $\text{plim}_{\gamma\to 0}\mathbf{X}(\gamma) = \boldsymbol{\iota}$ and $\mathbf{M}$ is the idempotent projector constructed from the observations $\mathbf{Z}_t := [1, \mathbf{W}_t']' = [1, X_t, \mathbf{D}_t']'$. First order use of L'Hôpital's rule fails due to the further degeneracy

$$\operatorname*{plim}_{\gamma\to 0}(d/d\gamma)N_n(\gamma) = \operatorname*{plim}_{\gamma\to 0} 2\{\mathbf{X}(\gamma)'\mathbf{MU}\}\left\{(d/d\gamma)\mathbf{X}(\gamma)'\mathbf{MU}\right\} = 0,$$

$$\operatorname*{plim}_{\gamma\to 0}(d/d\gamma)D_n(\gamma) = \operatorname*{plim}_{\gamma\to 0} 2\widehat{\sigma}_{n,0}^2\left\{(d/d\gamma)\mathbf{X}(\gamma)'\mathbf{MX}(\gamma)\right\} = 0.$$

Hence, it is necessary to apply l'Hôpital's rule a further time to remove the degeneracy. Second, consider the case in which γ converges to one, which shows that

$$\operatorname*{plim}_{\gamma\to 1} N_n(\gamma) = 0 \quad \text{and} \quad \operatorname*{plim}_{\gamma\to 1} D_n(\gamma) = 0$$

because $\text{plim}_{\gamma\to 1}\mathbf{X}(\gamma) = \mathbf{X}$ and $\mathbf{X}'\mathbf{M} = \mathbf{0}$. Again, first order use of l'Hôpital fails because

$$\operatorname*{plim}_{\gamma\to 1}(d/d\gamma)N_n(\gamma) = \operatorname*{plim}_{\gamma\to 1} 2\{\mathbf{X}(\gamma)'\mathbf{MU}\}\left\{(d/d\gamma)\mathbf{X}(\gamma)'\mathbf{MU}\right\} = 0,$$

$$\operatorname*{plim}_{\gamma\to 1}(d/d\gamma)D_n(\gamma) = \operatorname*{plim}_{\gamma\to 1} 2\widehat{\sigma}^2_{n,0}\left\{(d/d\gamma)\mathbf{X}(\gamma)\right\}'\mathbf{M}\mathbf{X}(\gamma) = 0,$$

and a further application is needed to remove the degeneracy. The required further derivatives are provided in the following lemma.

Lemma 7. *Given Assumption 1,*

(i) plim$_{\gamma\to 0}$ $N_n^{(2)}(\gamma) = 2\{\mathbf{L}_1\mathbf{M}\mathbf{U}\}^2$ *and plim*$_{\gamma\to 0}$ $D_n^{(2)}(\gamma) = 2\widehat{\sigma}^2_{n,0}\mathbf{L}_1\mathbf{M}\mathbf{L}_1$*; and*

(ii) plim$_{\gamma\to 1}$ $N_n^{(2)}(\gamma) = 2\{\mathbf{C}_1\mathbf{M}\mathbf{U}\}^2$ *and plim*$_{\gamma\to 1}$ $D_n^{(2)}(\gamma) = 2\widehat{\sigma}^2_{n,0}\mathbf{C}_1\mathbf{M}\mathbf{C}_1$*, where for* $j = 1$ *and* 2*,* $N_n^{(j)}(\gamma) := (\partial^j/\partial\gamma^j)N_n(\gamma)$ *and* $D_n^{(j)}(\gamma) := (\partial^j/\partial\gamma^j)D_n(\gamma)$*.*

Lemma 7 implies that

$$\operatorname*{plim}_{\gamma\to 0}\frac{N_n(\gamma)}{D_n(\gamma)} = \frac{\{\mathbf{L}_1\mathbf{M}\mathbf{U}\}^2}{\widehat{\sigma}^2_{n,0}\mathbf{L}_1\mathbf{M}\mathbf{L}_1} \quad \text{and} \quad \operatorname*{plim}_{\gamma\to 1}\frac{N_n(\gamma)}{D_n(\gamma)} = \frac{\{\mathbf{C}_1\mathbf{M}\mathbf{U}\}^2}{\widehat{\sigma}^2_{n,0}\mathbf{C}_1\mathbf{M}\mathbf{C}_1}. \tag{10}$$

The limits in (10) are the same null approximation limits as those obtained in Theorems 2 and 3, respectively. This equivalence implies that we can obtain the null approximations directly by letting the parameter γ in $\{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}^2/\{\hat{\sigma}^2_{n,0}\mathbf{X}(\gamma)'\,\mathbf{M}\mathbf{X}(\gamma)\}$ pass to zero or unity and this is so even though Theorem 1 explicitly removes zero and unity from $\mathbf{\Gamma}(\epsilon)$. In addition, this result also implies that

$$\sup_{\gamma\in\mathbf{\Gamma}}\frac{\{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}^2}{\hat{\sigma}^2_{n,0}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma)} \geq \max\left[\frac{\{\mathbf{L}_1\mathbf{M}\mathbf{U}\}^2}{\hat{\sigma}^2_{n,0}\mathbf{L}_1\mathbf{M}\mathbf{L}_1}, \frac{\{\mathbf{C}_1\mathbf{M}\mathbf{U}\}^2}{\hat{\sigma}^2_{n,0}\mathbf{C}_1\mathbf{M}\mathbf{C}_1}\right].$$

Therefore, the asymptotic null approximations provided in Theorems 2 and 3 can be combined with the null approximation in Theorem 1 so that the null approximation of the QLR test can be delivered under $\mathcal{H}_0$. For this purpose it is necessary to combine the regularity conditions of Theorems 2 and 3, as in the following assumption.

Assumption 5. *(i) For each* $\epsilon > 0$*,* $\mathbf{A}(\gamma)$ *and* $\mathbf{B}(\gamma)$ *are positive definite uniformly on* $\mathbf{\Gamma}(\epsilon)$*, where*

$$\mathbf{A}(\gamma) := \begin{bmatrix} \mathbb{E}[X_t^{2\gamma}] & \mathbb{E}[X_t^{1+\gamma}\log(X_t)] & \mathbb{E}[X_t^{\gamma}\log(X_t)] & \mathbb{E}[X_t^{\gamma}\mathbf{Z}_t'] \\ \mathbb{E}[X_t^{1+\gamma}\log(X_t)] & \mathbb{E}[X_t^2\log^2(X_t)] & \mathbb{E}[X_t\log^2(X_t)] & \mathbb{E}[X_t\log(X_t)\mathbf{Z}_t'] \\ \mathbb{E}[X_t^{\gamma}\log(X_t)] & \mathbb{E}[X_t\log^2(X_t)] & \mathbb{E}[\log^2(X_t)] & \mathbb{E}[\log(X_t)\mathbf{Z}_t'] \\ \mathbb{E}[X_t^{\gamma}\mathbf{Z}_t] & \mathbb{E}[X_t\log(X_t)\mathbf{Z}_t] & \mathbb{E}[\log(X_t)\mathbf{Z}_t] & \mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'] \end{bmatrix} \quad \textit{and}$$

$$\mathbf{B}(\gamma) := \begin{bmatrix} \mathbb{E}[U_t^2 X_t^{2\gamma}] & \mathbb{E}[U_t^2 X_t^{1+\gamma}\log(X_t)] & \mathbb{E}[U_t^2 X_t^{\gamma}\log(X_t)] & \mathbb{E}[U_t^2 X_t^{\gamma}\mathbf{Z}_t'] \\ \mathbb{E}[U_t^2 X_t^{1+\gamma}\log(X_t)] & \mathbb{E}[U_t^2 X_t^{2}\log^2(X_t)] & \mathbb{E}[U_t^2 X_t\log^2(X_t)] & \mathbb{E}[U_t^2 X_t\log(X_t)\mathbf{Z}_t'] \\ \mathbb{E}[U_t^2 X_t^{\gamma}\log(X_t)] & \mathbb{E}[U_t^2 X_t\log^2(X_t)] & \mathbb{E}[U_t^2\log^2(X_t)] & \mathbb{E}[U_t^2\log(X_t)\mathbf{Z}_t'] \\ \mathbb{E}[U_t^2 X_t^{\gamma}\mathbf{Z}_t] & \mathbb{E}[U_t^2 X_t\log(X_t)\mathbf{Z}_t] & \mathbb{E}[U_t^2\log(X_t)\mathbf{Z}_t] & \mathbb{E}[U_t^2\mathbf{Z}_t\mathbf{Z}_t'] \end{bmatrix};$$

(ii) $\{U_t, \mathcal{F}_t\}$ *is an MDS, where* $\mathcal{F}_t$ *is the adapted smallest* σ*-field generated by* $\{\mathbf{Z}_{t+1}, U_t, \mathbf{Z}_t, U_{t-1}, \cdots\}$*;*

(iii) There is a strictly stationary and ergodic sequence $\{M_t, S_t\}$ *such that for* $j = 1, 2, \cdots, k+1$ *and for some* $r > 1$*,* $|D_{t,j}| \leq M_t$*,* $\mathbb{E}[M_t^{4r}] < \infty$*,* $\mathbb{E}[S_t^8] < \infty$*, and*

(iii.a) $|U_t| \leq M_t$*,* $|X_t| \leq S_t$*, and* $|\log[X_t]| \leq S_t$*;*

(iii.b) $|X_t| \leq M_t$*,* $|U_t| \leq S_t$*, and* $|\log[X_t]| \leq S_t$*; or*

(iii.c) $|\log[X_t]| \leq M_t$*,* $|X_t| \leq S_t$*, and* $|U_t| \leq S_t$*.*

(iv) $\sup_{\gamma\in\mathbf{\Gamma}} |X_t^{\gamma}| \leq M_t$ *and* $\sup_{\gamma\in\mathbf{\Gamma}} |X_t^{\gamma}\log(X_t)| \leq M_t$*.* □

Assumption 5 is stronger than Assumptions 2, 3, and 4, each of which separately holds under Assumption 5. Specifically, the square matrices in Assumption 5(*i*) are obtained by combining the square matrices in Assumptions 2(*i*), 3(*i*), and 4(*i*). Furthermore, Assumption 5(*iii*) is exactly the same as Assumption 4(*iii*). This is because Assumption 4(*iii*) is stronger than Assumption 3(*iii*), so that if Assumption 4(*iii*) only is assumed, this is already sufficient for obtaining the results implied by Assumption 3(*iii*) and Assumption 4(*iii*). Finally, Assumption 5(*iv*) is the same condition as Assumption 2(*iv*). This condition is imposed again to ensure the tightness property of the given process.

Using these conditions we have the following result.

Theorem 4. *Given Assumptions 1, 5, and* $\mathcal{H}_0$*,*

(i) $QLR_n = \sup_{\gamma\in\mathbf{\Gamma}} \{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}^2 / \{\hat{\sigma}_{n,0}^2 \mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma)\}$*; and*

(ii) $QLR_n \Rightarrow \sup_{\gamma\in\mathbf{\Gamma}} \mathcal{Z}(\gamma)^2$*.* □

This result gives the asymptotic null approximation of the QLR test under $\mathcal{H}_0$ and its limiting form as a functional of a Gaussian process $\mathcal{Z}(\cdot)$. Importantly, the process $\mathcal{Z}(\cdot)$ is not continuous on $\mathbf{\Gamma}$. In particular, the process is discontinuous at $\gamma = 0$ and 1 with probability 1. For each γ, defining

$$Z_n(\gamma) := \frac{\widetilde{N}_n(\gamma)}{\widetilde{D}_n(\gamma)},$$

where $\widetilde{N}_n(\gamma) := \mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}$ and $\widetilde{D}_n(\gamma) := \{\hat{\sigma}_{n,0}^2 \mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma)\}^{1/2}$, we can regard $\mathcal{Z}(\cdot)$ as the weak limit of $Z_n(\cdot)$. Observe that

$$\lim_{\gamma\downarrow 0} Z_n(\gamma) = -\lim_{\gamma\uparrow 0} Z_n(\gamma) \quad \text{and} \quad \lim_{\gamma\downarrow 1} Z_n(\gamma) = -\lim_{\gamma\uparrow 1} Z_n(\gamma)$$

with probability 1 by virtue of l'Hôpital's rule, so that $\mathcal{Z}(\cdot)$ is discontinuous at 0 and 1. Furthermore, neither $Z_n(0)$ nor $Z_n(1)$ are defined, a property that continues to hold when n tends to infinity. Thus, it follows that

$$\lim_{\gamma\downarrow 0} \mathcal{Z}(\gamma) = -\lim_{\gamma\uparrow 0} \mathcal{Z}(\gamma) \quad \text{and} \quad \lim_{\gamma\downarrow 1} \mathcal{Z}(\gamma) = -\lim_{\gamma\uparrow 1} \mathcal{Z}(\gamma) \tag{11}$$

with probability 1. A typical sample path is illustrated in Figure 1[1], wherein it is clear that $\mathcal{Z}(0)$ and $\mathcal{Z}(1)$ are undefined under $\mathcal{H}_0$. In view of (11), the squared process $\mathcal{Z}(\gamma)^2$ has equal left-hand and right-hand side limits as γ tends to 0 and 1, i.e.,

$$\lim_{\gamma\downarrow 0} \mathcal{Z}(\gamma)^2 = \lim_{\gamma\uparrow 0} \mathcal{Z}(\gamma)^2 \quad \text{and} \quad \lim_{\gamma\downarrow 1} \mathcal{Z}(\gamma)^2 = \lim_{\gamma\uparrow 1} \mathcal{Z}(\gamma)^2$$

with probability 1. If $\mathcal{Z}(0)$ and $\mathcal{Z}(1)$ are defined by these limits, it follows that $\mathcal{Z}(\cdot)^2$ is continuous on $\boldsymbol{\Gamma}$ with probability 1. Likewise, if $Z_n(0)^2$ and $Z_n(1)^2$ are defined by $\lim_{\gamma\to 0} Z_n(\gamma)^2$ and $\lim_{\gamma\to 1} Z_n(\gamma)^2$, we can define $Z_n(\cdot)$ to be continuous on $\boldsymbol{\Gamma}$ with probability 1. This feature of the process is exploited when computing the QLR test statistic under $\mathcal{H}_0$. More specifically, Theorem 4(*i*) implies that QLR_n is obtained as $\sup_{\gamma\in\boldsymbol{\Gamma}} Z_n(\gamma)^2$, and unless $Z_n(\cdot)^2$ is continuous on $\boldsymbol{\Gamma}$, the QLR test statistic may not be well defined.

Theorem 4 has the following main implications. First, the asymptotic null approximation of the QLR test addresses the trifold identification problem and, under the regularity conditions for each case, ensures that the limiting null distribution exists for each form of the null hypothesis. Second, the QLR test simultaneously satisfies these separate conditions, thereby accommodating trifold identification issues. With this property, the QLR test has the capacity to test linearity within a unified framework that accommodates all possibilities. Finally, the null approximation is obtained by using only second-order approximations. This aspect of the test differs from the ANN literature, in which higher-order approximations are usually required for testing linearity.

[1]Due to the smooth sample path property of $\mathcal{G}(\cdot)$, $\mathcal{Z}(\cdot)$ is also smooth except at $\gamma = 0$ and 1. Furthermore, we can apply l'Hôpital's rule even to $\mathcal{Z}(\cdot)$ by the second-order differentiability condition of $\kappa(\cdot,\cdot)$.

3 Testing for Power Transforms of a Trend Regressor

3.1 Asymptotically Collinear Trends

We now extend our discussion to the case where the dependent variable $\{Y_t\}$ is a trend stationary process with a deterministic time trend. This type of model is particularly important in analyzing nonstationary time series and trend removal procedures. More specifically, we suppose that Y_t is nonstationary and $\mathbb{E}[Y_t|\mathbf{D}_t]$ is a function of both t and $\mathbf{D}_t$, where $\{\mathbf{D}_t\}$ is, as before, a strictly stationary sequence satisfying certain mixing condition. Primary attention now focuses on testing whether $\mathbb{E}[Y_t|\mathbf{D}_t]$ is a linear function of $(1, \mathbf{D}_t', t)'$. For such a test, we consider the formulation

$$\mathcal{M}' := \{m_t(\cdot) : \mathbf{\Omega} \mapsto \mathbb{R} : m_t(\alpha, \boldsymbol{\delta}, \beta, \gamma) := \alpha + \mathbf{D}_t'\boldsymbol{\eta} + \xi t + \beta t^{\gamma}\}$$

as an application of the previous model. We note that the only difference between $\mathcal{M}$ and $\mathcal{M}'$ arises from the replacement of X_t with t. The explanatory regressor $\mathbf{D}_t$ may be present in the conditional mean and may be used to capture temporal dependence in the data that is not embodied in the nonlinear time trend t^{γ}.

In spite of this correspondence with the earlier model, the QLR test cannot be straightforwardly analyzed as in the previous section. The main reason is that the regularity condition in Assumption 5 no longer holds. More specifically, the positive definite matrix condition in Assumptions 5(*i*) fails and the (implied) regressors are asymptotically collinear. The following lemma states the property in a precise way.

Lemma 8. *If $\{\mathbf{D}_t\}$ is strictly stationary and ergodic such that for each $j = 1, 2, \dots k$, $\mathbb{E}[D_{t,j}^2] < \infty$, then for each $\gamma \in \mathbf{\Gamma}(\epsilon)$, $\mathbf{F}_n^{-1}\sum_{t=1}^n \mathbf{H}_t(\gamma)\mathbf{H}_t(\gamma)'\mathbf{F}_n^{-1} \stackrel{a.s.}{\to} \mathbf{\Xi}(\gamma)$, where $\mathbf{H}_t(\gamma) := [t^{\gamma}, t\log(t), \log(t), 1, t, \mathbf{D}_t']'$,*

$$\mathbf{\Xi}(\gamma) := \begin{bmatrix} \frac{1}{2\gamma+1} & \frac{1}{\gamma+2} & \frac{1}{\gamma+1} & \frac{1}{\gamma+1} & \frac{1}{\gamma+2} & \frac{1}{\gamma+1}\mathbb{E}[\mathbf{D}_t'] \\ \frac{1}{\gamma+2} & \frac{1}{3} & \frac{1}{2} & \frac{1}{2} & \frac{1}{3} & \frac{1}{2}\mathbb{E}[\mathbf{D}_t'] \\ \frac{1}{\gamma+1} & \frac{1}{2} & 1 & 1 & \frac{1}{2} & \mathbb{E}[\mathbf{D}_t'] \\ \frac{1}{\gamma+1} & \frac{1}{2} & 1 & 1 & \frac{1}{2} & \mathbb{E}[\mathbf{D}_t'] \\ \frac{1}{\gamma+2} & \frac{1}{3} & \frac{1}{2} & \frac{1}{2} & \frac{1}{3} & \frac{1}{2}\mathbb{E}[\mathbf{D}_t'] \\ \frac{1}{\gamma+1}\mathbb{E}[\mathbf{D}_t] & \frac{1}{2}\mathbb{E}[\mathbf{D}_t] & \mathbb{E}[\mathbf{D}_t] & \mathbb{E}[\mathbf{D}_t] & \frac{1}{2}\mathbb{E}[\mathbf{D}_t] & \mathbb{E}[\mathbf{D}_t\mathbf{D}_t'] \end{bmatrix},$$

$\mathbf{F}_n :=$ diag$[n^{\frac{1}{2}+\gamma}, n^{\frac{3}{2}}\log(n), n^{\frac{1}{2}}\log(n), n^{\frac{1}{2}}, n^{\frac{3}{2}}, n^{\frac{1}{2}}\boldsymbol{\iota}_k]$, and $\boldsymbol{\iota}_k$ is a $k \times 1$ vector of ones. □

The proof of Lemma 8 follows straightforwardly using equation (26) of Phillips (2007) and the monotone and dominated convergence theorems. Note that $\mathbf{F}_n^{-1}\sum_{t=1}^n \mathbf{H}_t(\gamma)\mathbf{H}_t(\gamma)'\mathbf{F}_n^{-1}$ is a (matrix normalized) sample analog of the square matrix considered in Assumption 5(*i*). Here X_t is replaced by time trend functions and each element of the square matrix is rescaled appropriately so that it is not negligible in probability and also bounded in probability. Since time trends are involved, the scaling rates of the components are different from the standard stationary variable case and are parameter dependent on γ.

The limit of the square signal matrix in Lemma 8 is a singular matrix. Specifically, the second column of the limit matrix is identical to the fifth column, and its third column is identical to the fourth column. Importantly, this singularity does not necessarily imply that the asymptotic null distribution of the QLR test does not exist. However, the limit theory for the QLR test cannot easily be revealed from the framework of Section 2. Instead, it is convenient to use a different approach based on Phillips (2007) in what follows.

3.2 QLR Statistic under the Null Hypothesis

The asymptotic null distribution of the QLT test can be found by reformulating the model. Instead of $\mathcal{M}'$, we use the following 'weak trend' formulation involving the trend fraction $\frac{t}{n}$ and power functions of $\frac{t}{n}$

$$\mathcal{M}'' := \left\{ m_t(\cdot) : \mathbf{\Omega} \mapsto \mathbb{R} : m_t(\alpha, \boldsymbol{\delta}, \beta, \gamma) := \alpha + \mathbf{D}_t'\boldsymbol{\eta} + \xi_n \left(\frac{t}{n}\right) + \lambda_n(\beta, \gamma) \left(\frac{t}{n}\right)^{\gamma} \right\},$$

where $\xi_n := \xi n$ and $\lambda_n(\beta, \gamma) := \beta n^{\gamma}$. For notational simplicity, define $s_{n,t} := t/n$, so that $s_{n,t} \in (0, 1]$. This model is an equivalent specification that captures the nonstationary aspect of Y_t by permitting the parameters to be functions of the sample size and converting the unbounded time trend t into the uniformly bounded regressor $\frac{t}{n}$. This weak trend has asymptotics closely related to those of a stationary regressor. Linearity is obtained from $\mathcal{M}'$ by setting $\beta = 0$, $\gamma = 0$, or $\gamma = 1$. On the other hand, linearity is obtained from $\mathcal{M}''$ by setting $\lambda_n(\cdot) = 0$ for any n, $\gamma = 0$, or $\gamma = 1$. Furthermore, $\beta = 0$ if and only if $\lambda_n(\cdot) = 0$. Thus, if $(\alpha_*, \boldsymbol{\eta}_*, \xi_*, \beta_*, \gamma_*)$ is such that $\mathbb{E}[Y_t|\mathbf{D}_t] = \alpha_* + \mathbf{D}_t'\boldsymbol{\eta}_* + \xi_* t + \beta_* t^{\gamma_*}$, the null hypothesis is given as

$$\widetilde{\mathcal{H}}_0 : \ \exists(\alpha_*, \boldsymbol{\eta}_*, \xi_*), \mathbb{E}[Y_t|\mathbf{D}_t] = \alpha_* + \mathbf{D}_t'\boldsymbol{\eta}_* + \xi_* t \ \text{ w.p. } 1,$$

which can be formulated in terms of the following specific hypotheses

$$\widetilde{\mathcal{H}}_0' : \lambda_n(\beta_*, \gamma_*) = 0; \quad \widetilde{\mathcal{H}}_0'' : \gamma_* = 0; \quad \text{and} \quad \widetilde{\mathcal{H}}_0''' : \gamma_* = 1.$$

Using this modification of the model, the asymptotic null behavior of the QLR test can be obtained under appropriate conditions using methods similar to those in the last section. We start with the following assumptions:

Assumption 6. *(i)* $(Y_t, \mathbf{D}_t')' \in \mathbb{R}^{1+k}$ *($k \in \mathbb{N}$) is given, and* $\{\mathbf{D}_t\}$ *is a* ϕ*-mixing process with mixing decay rate* $-m/2(m-1)$ *with* $m \geq 2$ *or an* α*-mixing process with mixing decay rate* $-m/(m-2)$ *with* $m > 2$*, and* Y_t *is a time-trend stationary process;*

(ii) $\mathbb{E}[Y_t|\mathbf{D}_t]$ *is specified as* $\mathcal{M}'' := \{m_t(\cdot) : \mathbf{\Omega}_n \mapsto \mathbb{R} : m_t(\alpha, \boldsymbol{\eta}, \xi_n, \lambda_n, \gamma) := \alpha + \mathbf{D}_t'\boldsymbol{\eta} + \xi_n s_{n,t} + \lambda_n s_{n,t}^{\gamma}\}$*, where* $\mathbf{\Omega}_n$ *is the parameter space of* $\boldsymbol{\omega}_n := (\alpha, \boldsymbol{\eta}', \xi_n, \lambda_n, \gamma)'$*;* $\mathbf{Z}_{n,t} := (1, s_{n,t}, \mathbf{D}_t')'$*; and* n *is the sample size;*

(iii) $\mathbf{\Omega} = \boldsymbol{A} \times \mathbf{H} \times \mathbf{\Xi}_n \times \mathbf{\Lambda}_n \times \mathbf{\Gamma}$ *such that* $\boldsymbol{A}$*,* $\mathbf{\Delta}$*, and* $\mathbf{\Gamma}$ *are convex and compact parameter spaces in* $\mathbb{R}$*,* $\mathbb{R}^k$*, and* $\mathbb{R}$*, respectively, such that 0 and 1 are interior elements of* $\mathbf{\Gamma}$ *with* $\gamma_o := \inf \mathbf{\Gamma} > -1/2$*, and for each* n*,* $\mathbf{\Xi}_n$ *and* $\mathbf{\Lambda}_n$ *are convex and compact parameter spaces in* $\mathbb{R}$*; and*

(iv) $\mathbf{Z}'\mathbf{Z} = \sum_{t=1}^{n} \mathbf{Z}_{n,t}\mathbf{Z}_{n,t}'$ *is nonsingular with probability 1.* □

Assumption 6 corresponds to Assumption 1 with some differences in the mixing condition. Although tightness is needed to obtain the asymptotic null distribution of the QLR test, the mixing condition in Assumption 1 is not necessary, as discussed below. We next modify Assumption 5 to fit the structure of the time-trend model.

Assumption 7. *(i) For each* $\epsilon > 0$*, the following square matrices are positive definite uniformly on* $\mathbf{\Gamma}(\epsilon)$*:*

$$\widetilde{\mathbf{A}}(\gamma) := \begin{bmatrix} \frac{1}{2\gamma+1} & -\frac{1}{(\gamma+2)^2} & -\frac{1}{(\gamma+1)^2} & \frac{1}{\gamma+1} & \frac{1}{\gamma+2} & \frac{1}{1+\gamma}\mathbb{E}[\mathbf{D}_t'] \\ -\frac{1}{(\gamma+2)^2} & \frac{2}{27} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{9} & -\frac{1}{4}\mathbb{E}[\mathbf{D}_t'] \\ -\frac{1}{(\gamma+1)^2} & \frac{1}{4} & 2 & -1 & -\frac{1}{4} & -\mathbb{E}[\mathbf{D}_t'] \\ \frac{1}{\gamma+1} & -\frac{1}{4} & -1 & 1 & \frac{1}{2} & \mathbb{E}[\mathbf{D}_t'] \\ \frac{1}{\gamma+2} & -\frac{1}{9} & -\frac{1}{4} & \frac{1}{2} & \frac{1}{3} & \frac{1}{2}\mathbb{E}[\mathbf{D}_t'] \\ \frac{1}{\gamma+1}\mathbb{E}[\mathbf{D}_t] & -\frac{1}{4}\mathbb{E}[\mathbf{D}_t] & -\mathbb{E}[\mathbf{D}_t] & \mathbb{E}[\mathbf{D}_t] & \frac{1}{2}\mathbb{E}[\mathbf{D}_t] & \mathbb{E}[\mathbf{D}_t\mathbf{D}_t'] \end{bmatrix} \quad \textit{and}$$

$$\widetilde{\mathbf{B}}(\gamma) := \begin{bmatrix} \frac{1}{2\gamma+1}\sigma_*^2 & -\frac{1}{(\gamma+2)^2}\sigma_*^2 & -\frac{1}{(\gamma+1)^2}\sigma_*^2 & \frac{1}{\gamma+1}\sigma_*^2 & \frac{1}{\gamma+2}\sigma_*^2 & \frac{1}{1+\gamma}\mathbb{E}[U_t^2\mathbf{D}_t'] \\ -\frac{1}{(\gamma+2)^2}\sigma_*^2 & \frac{2}{27}\sigma_*^2 & \frac{1}{4}\sigma_*^2 & -\frac{1}{4}\sigma_*^2 & -\frac{1}{9}\sigma_*^2 & -\frac{1}{4}\mathbb{E}[U_t^2\mathbf{D}_t'] \\ -\frac{1}{(\gamma+1)^2}\sigma_*^2 & \frac{1}{4}\sigma_*^2 & 2\sigma_*^2 & -\sigma_*^2 & -\frac{1}{4}\sigma_*^2 & -\mathbb{E}[U_t^2\mathbf{D}_t'] \\ \frac{1}{\gamma+1}\sigma_*^2 & -\frac{1}{4}\sigma_*^2 & -\sigma_*^2 & \sigma_*^2 & \frac{1}{2}\sigma_*^2 & \mathbb{E}[U_t^2\mathbf{D}_t'] \\ \frac{1}{\gamma+2}\sigma_*^2 & -\frac{1}{9}\sigma_*^2 & -\frac{1}{4}\sigma_*^2 & \frac{1}{2}\sigma_*^2 & \frac{1}{3}\sigma_*^2 & \frac{1}{2}\mathbb{E}[U_t^2\mathbf{D}_t'] \\ \frac{1}{\gamma+1}\mathbb{E}[U_t^2\mathbf{D}_t] & -\frac{1}{4}\mathbb{E}[U_t^2\mathbf{D}_t] & -\mathbb{E}[U_t^2\mathbf{D}_t] & \mathbb{E}[U_t^2\mathbf{D}_t] & \frac{1}{2}\mathbb{E}[U_t^2\mathbf{D}_t] & \mathbb{E}[U_t^2\mathbf{D}_t\mathbf{D}_t'] \end{bmatrix},$$

where $U_t := Y_t - \mathbb{E}[Y_t|\mathbf{D}_t]$ *and* $\sigma_*^2 := \mathbb{E}[U_t^2]$*;*

(ii) $\{U_t, \mathcal{F}_t\}$ *is an MDS, where* $\mathcal{F}_t$ *is the adapted smallest* σ*-field generated by* $\{\mathbf{D}_{t+1}, U_t, \mathbf{D}_t, U_{t-1}, \cdots\}$*; and*

(iii) There is a strictly stationary and ergodic sequence $\{M_t\}$ *such that for* $j = 1, 2, \cdots, k$, $|D_{t,j}| \leq M_t$, $|U_t| \leq M_t$, *and for some* $r > 1$, $\mathbb{E}[M_t^{4r}] < \infty$. □

Some discussion of Assumptions 6 and 7 is warranted. First, the mixing condition in Assumption 1 is relaxed to Assumption 6(*i*). We show in the proof of Theorem 5 given below that tightness of $\{n^{-1/2} \sum_{t=1}^{n} s_{n,t}^{(\cdot)} U_t\}$ follows without invoking the arguments of Doukhan, Massart, and Rio (1995). For this demonstration, we use the fact that $\{s_{n,t}\}$ is a sequence of non-random positive numbers uniformly bounded by unity. Second, the square matrices given in Assumption 7(*i*) are the probability limits of the following square matrices: $n^{-1} \sum \mathbf{G}_{n,t}(\gamma)\mathbf{G}_{n,t}(\gamma)'$ and $n^{-1} \sum U_t^2 \mathbf{G}_{n,t}(\gamma)\mathbf{G}_{n,t}(\gamma)'$, where for each γ, $\mathbf{G}_{n,t}(\gamma) := [s_{n,t}^{\gamma}, s_{n,t} \log(s_{n,t}), \log(s_{n,t}), 1, s_{n,t}$ which corresponds to $\mathbf{H}_t(\gamma)$. The limit matrices are obtained by replacing $\mathbb{E}[\cdot]$ and X_t in Assumption 5(*i*) with limits of corresponding averages of components involving the weak trend function $s_{n,t}$. The limit matrices are assumed to be nonsingular in Assumption 7(*i*). Third, the nonsingular matrix condition of $\widetilde{\mathbf{A}}(\gamma)$ is identical to the condition that $\mathbf{D}_t$ has a nonsingular covariance matrix. More specifically, note that for each $\gamma \in \mathbf{\Gamma}(\epsilon)$, the first five principal minors of $\widetilde{\mathbf{A}}(\gamma)$ have strictly positive determinants. Therefore, $\widetilde{\mathbf{A}}(\gamma)$ is positive definite if and only if $\mathbb{E}[\mathbf{D}_t\mathbf{D}_t'] - \widetilde{\mathbf{A}}^{(2,1)}(\gamma)\{\widetilde{\mathbf{A}}^{(1,1)}(\gamma)\}^{-1}\widetilde{\mathbf{A}}^{(1,2)}(\gamma)$ is positive definite, where we partition $\widetilde{\mathbf{A}}(\gamma)$ as

$$\widetilde{\mathbf{A}}(\gamma) \equiv \begin{bmatrix} \widetilde{\mathbf{A}}^{(1,1)}(\gamma) & \widetilde{\mathbf{A}}^{(1,2)}(\gamma) \\ \widetilde{\mathbf{A}}^{(2,1)}(\gamma) & \mathbb{E}[\mathbf{D}_t\mathbf{D}_t'] \end{bmatrix}.$$

The final entry is identical to $\mathbb{E}[\mathbf{D}_t\mathbf{D}_t'] - \mathbb{E}[\mathbf{D}_t]\mathbb{E}[\mathbf{D}_t']$ by the definition of $\widetilde{\mathbf{A}}(\gamma)$, which is the covariance matrix of $\mathbf{D}_t$. Fourth, although the model is modified to $\mathcal{M}''$ from $\mathcal{M}'$, the QLR test obtained by using $\mathcal{M}'$ is identical to that of $\mathcal{M}''$. This follows directly from the invariance principle of maximum likelihood: reparameterization does not modify the level of the maximized quasi-likelihood. Finally, Assumption 7(*ii*) does not impose conditional homoskedasticity. The asymptotic null distribution of the QLR test continues to hold under conditional heteroskedasticity.

Our main result now follows. We define the relevant matrices in the same way as in Section 2. That is, for each $\gamma \in \mathbf{\Gamma}$, $\mathbf{T}(\gamma) := [s_{n,1}^{\gamma}, \ldots, s_{n,n}^{\gamma}]'$ and $\mathbf{M} := \mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$ with $\mathbf{Z}_{n,t}'$ as the t-th row vector of $\mathbf{Z}$.

Theorem 5. *Given Assumptions 6, 7, and* $\widetilde{\mathcal{H}}_0$,

(i) $QLR_n = \sup_{\gamma \in \mathbf{\Gamma}} \{\mathbf{T}(\gamma)'\mathbf{M}\mathbf{U}\}^2 / \{\hat{\sigma}_{n,0}^2 \mathbf{T}(\gamma)'\mathbf{M}\mathbf{T}(\gamma)\}$;

(ii) $QLR_n \Rightarrow \sup_{\gamma \in \mathbf{\Gamma}} \widetilde{\mathcal{Z}}(\gamma)^2$, *where* $\widetilde{\mathcal{Z}}(\cdot)$ *is a Gaussian process with zero mean and covariance kernel* $\tilde{\kappa}(\gamma, \gamma')$ *such that for each* γ, $\gamma' \in \mathbf{\Gamma} \setminus \{0, 1\}$,

$$\tilde{\kappa}\left(\gamma, \gamma'\right) = \mathbb{E}[\widetilde{\mathcal{Z}}(\gamma)\widetilde{\mathcal{Z}}(\gamma')] = c(\gamma, \gamma')\frac{(1 + 2\gamma)^{1/2}(1 + 2\gamma'^{1/2}}{(1 + \gamma + \gamma')}, \tag{12}$$

where for each $\gamma \in \mathbf{\Gamma}$*,* $c(\gamma, \gamma') := \gamma\gamma'(\gamma - 1)(\gamma' - 1)/|\gamma\gamma'(\gamma - 1)(\gamma' - 1)|$. □

The proof of Theorem 5(*i*) follows that of Theorem 4(*i*). In particular, if $\mathbf{X}(\gamma)$ in Theorem 4(*i*) is replaced by $\mathbf{T}(\gamma)$, the arguments in the proof of Theorem 4(*i*) can be used in the proof of Theorem 5(*i*), and are not repeated. However, the Appendix does prove the weak convergence of the QLR test given in Theorem 5(*ii*) and derives the covariance kernel (12). Weak convergence follows because the sequence $\{n^{-1/2} \sum s_{n,t}^{(\cdot)} U_t\}$ is tight, which is straightforward because $s_{n,t}^{(\cdot)}$ is deterministic and the MDS U_t satisfies the mixing condition of Assumption 6(*i*). Theorem 5(*ii*) shows that reparameterizing the model $\mathcal{M}'$ as $\mathcal{M}''$ gives the asymptotic null behavior of the QLR test and reveals that it may again be represented as a functional of a Gaussian process.

The covariance structure of the associated Gaussian process is independent of the joint distribution of $(U_t, \mathbf{D}_t')$. Further, the same covariance structure applies irrespective of whether there is conditional heteroskedasticity in the residuals. We call the Gaussian process $\widetilde{\mathcal{Z}}$ with covariance kernel (12) the *power Gaussian process*, noting that $\widetilde{\mathcal{Z}}$ is obtained while testing for neglected nonlinearity using the power transform of a trend, and therefore differs from the process $\mathcal{Z}$ in Section 2. In particular, if the residual is serially correlated, the covariance structure of the associated Gaussian process will generally differ from that of the power Gaussian process and its form will depend on the serial correlation. We note further that the power Gaussian process is not continuous at $\gamma = 0$ and 1 as is evident from the functional form of $c(\cdot, \cdot)$. Thus, if $(\gamma, \gamma') \in \{(\gamma, \gamma') : \gamma, \gamma' \in (-0.5, 0.0), (0.0, 1.0), (1.0, \infty)\} \cup \{(\gamma, \gamma') : \gamma \in (1, \infty), \gamma' \in (-0.5, 1.0)\} \cup \{(\gamma, \gamma') : \gamma \in (-0.5, 1), \gamma' \in (1.0, \infty)\}$, $c(\gamma, \gamma') = 1$. On the other hand, if $(\gamma, \gamma') \in \{(\gamma, \gamma') : \gamma, \in (0.0, 1), \gamma' \in (-0.5, 0.0)\} \cup \{(\gamma, \gamma') : \gamma \in (1, \infty), \gamma' \in (0.0, 1.0)\} \cup \{(\gamma, \gamma') : \gamma \in (-0.5, 0.0), \gamma' \in (0.0, 1.0)\} \cup \{(\gamma, \gamma') : \gamma \in (-0.5, 0.0), \gamma' \in (0.0, 1.0)\}$, $c(\gamma, \gamma') = -1$.

The null distribution of the QLR test can be represented in terms of another Gaussian process. For this purpose, let

$$\bar{\mathcal{Z}}(\gamma) := \sum_{j=2}^{\infty} \left[\frac{\gamma^4}{(\gamma + 1)^2 (2\gamma + 1)} \right]^{-1/2} \left(\frac{\gamma}{\gamma + 1} \right)^j G_j, \tag{13}$$

where $G_j \sim$ IID $N(0, 1)$. When $\gamma > -0.5$, $[\gamma/(1 + \gamma)]^j \to 0$ geometrically as $j \to \infty$, so that the covariance structure of this Gaussian process is well defined on the given parameter space. This process coincides with the Gaussian process that appeared in Cho and White (2010) for testing unobserved heterogeneity in duration data. Their work tested unobserved heterogeneity by considering the mixture hypothesis of two exponential conditional distributions and the likelihood ratio (LR) test. The asymptotic null distribution of their LR test was obtained by exploiting the features of $\bar{\mathcal{Z}}(\cdot)$. Notice that

$$\mathbb{E}[\bar{\mathcal{Z}}(\gamma)\bar{\mathcal{Z}}(\gamma')] = \frac{(1+2\gamma)^{1/2}(1+2\gamma')^{1/2}}{(1+\gamma+\gamma')}, \tag{14}$$

and this covariance structure differs from that of $\widetilde{\mathcal{Z}}(\cdot)$. We call the Gaussian process with the covariance structure (14) the *exponential Gaussian process*. As mentioned earlier, the power Gaussian process is discontinuous at 0 and 1 with probability 1. On the other hand, the exponential Gaussian process is continuous on $\mathbf{\Gamma}$ with probability 1. Notwithstanding this difference, the squared process $\bar{\mathcal{Z}}(\cdot)^2$ is distributionally equivalent to $\widetilde{\mathcal{Z}}(\cdot)^2$ because for any (γ, γ'), $c(\gamma, \gamma')^2 \equiv 1$, so that the covariance kernels of the two processes are the same. The next result follows immediately by this equivalence and continuous mapping.

Theorem 6. *Given Assumptions 6, 7, and* $\widetilde{\mathcal{H}}_0$, $QLR_n \Rightarrow \sup_{\gamma\in\mathbf{\Gamma}} \bar{\mathcal{Z}}(\gamma)^2$.

The exponential Gaussian process $\bar{\mathcal{Z}}$ can be easily simulated from (13) using a sequence of IID standard normal random variables $\{G_j\}$ and truncating the summation as in

$$\bar{\mathcal{Z}}_m(\gamma) := \sum_{j=2}^{m} \left[\frac{\gamma^4}{(\gamma+1)^2(2\gamma+1)}\right]^{-1/2} \left(\frac{\gamma}{\gamma+1}\right)^j G_j. \tag{15}$$

for some large m. Table 1 reports the asymptotic critical values of the QLR test obtained by implementing this simulation. We consider three different levels of significance: 1%, 5%, and 10% and let the parameter spaces be $[-0.20, 1.50]$, $[-0.10, 1.50]$, $[0.00, 1.50]$, $[0.10, 1.50]$, $[-0.20, 2.50]$, $[-0.10, 2.50]$, $[0.00, 2.50]$, and $[0.10,\ 2.50]$. Specifically, we let m be 500 and simulate $\sup_{\gamma\in\mathbf{\Gamma}} \bar{\mathcal{Z}}_m(\gamma)^2$ 100,000 times to obtain the asymptotic critical values. We can see that the asymptotic critical values uniformly increase as the size of parameter space gets bigger. For example, the critical values of $\mathbf{\Gamma} = [-0.20, 1.50]$ are uniformly less than those of $\mathbf{\Gamma} = [-0.20, 2.50]$.

4 Simulations

This section reports Monte Carlo experiments conducted to explore the finite sample properties of the QLR test and assess the adequacy of the asymptotic theory. The asymptotic null distributions are studied under two different modeling environments, which are examined separately in the following subsections.

4.1 Testing for Power Transforms of a Stationary Regressor

4.1.1 When the Asymptotic Null Distribution is Used

First, let the data $\{(Y_t, X_t) : t = 1, 2, \ldots, n\}$ be generated by

$$Y_t = \alpha_* + \xi_* X_t + U_t,$$

where $X_t := \exp(-\lambda_* H_t)$, $U_t \sim$ IID $N(0, \sigma_*^2)$, and $H_t \sim$ IID Exp(λ_*) such that U_t is independent of H_t and $(\alpha_*, \xi_*, \sigma_*^2, \lambda_*) = (1, 1, 1, 1)$. Second, given this DGP, we specify the following model for $\mathbb{E}[Y_t|X_t]$

$$\mathcal{M} = \{m_t(\cdot) : m_t(\alpha, \xi, \beta, \gamma) = \alpha + \xi X_t + \beta X_t^{\gamma}, \gamma \in \mathbf{\Gamma}\}.$$

The other parameters besides γ are not constrained to a convex and compact parameter space as the asymptotic null distribution of the QLR test is not dependent upon these other parameters. We consider the eight different parameter spaces used for obtaining the critical values in Table 1. The parameter spaces can be classified into two different groups: we let $[-0.20, 1.50]$, $[-0.10, 1.50]$, $[0.00, 1.50]$, and $[0.10, 1.50]$ belong to the first group, and $[-0.20, 2.50]$, $[-0.10, 2.50]$, $[0.00, 2.50]$, and $[0.10, 2.50]$ belong to the second group. These two groups are classified by their upper bounds. These different parameter spaces are considered to examine how they influence the null distributions of the QLR test. In particular, the two parameter spaces $[0.10, 1.50]$ and $[0.10, 2.50]$ do not contain zero. These two parameter spaces therefore reduce the scope of the trifold identification problem, because eliminating $\gamma_* = 0.0$ implies that the number of unidentified model cases is reduced.

Theorem 4 provides the asymptotic null distribution of the QLR test statistic, and the associated Gaussian process obtained under our DGP condition has the same covariance structure as that of the power Gaussian process. This simple covariance structure is straightforwardly obtained as U_t is IID with conditionally homoskedastic variance and H_t follows an exponential distribution.

The results are given in Table 2 which contains the empirical rejection rates of the null hypothesis obtained from 10,000 replications. We consider sample sizes $n = 50, 100, 200, 300, 400, 500$. The findings are as follows. First, for each parameter constellation, the empirical rejection rates approach the nominal levels as n increases, a result that is corroborated in Fig. 2 which graphs the empirical and asymptotic null distributions of the QLR test in selected cases showing close conformity of the distributions. Second, convergence to the nominal levels tends to be slower when the lower bound of $\mathbf{\Gamma}$ is closer to -0.50 and the upper bound is the same. For example, when the upper bound is 1.50 and $n = 500$, the empirical rejection rates of the QLR test with $\mathbf{\Gamma} = [-0.20, 1.50]$ are worse than those with $\mathbf{\Gamma} = [0.10, 1.50]$. Level distortion in the test can therefore be reduced by raising the lower bound of $\mathbf{\Gamma}$ from the minimum. Third, convergence to the nominal level improves as the upper bound of $\mathbf{\Gamma}$ increases, with the same lower bound. For example, when the lower bound is 0.10 and $n = 500$, the empirical rejection rates of the QLR test with $\mathbf{\Gamma} = [0.10, 2.50]$ are closer to the nominal than those with $\mathbf{\Gamma} = [0.10, 1.50]$. Thus level distortion may be attenuated by using a higher upper bound of $\mathbf{\Gamma}$.

4.1.2 When the Asymptotic Null Distribution cannot be Used

Let the DGP for $\{(Y_t, X_t, Z_t) : t = 1, 2, \ldots, n\}$ be

$$Y_t = \alpha_* + \xi_* X_t + \pi_* Z_t + \cos(Z_t) U_t, \tag{16}$$

where X_t is generated in to the same way as in Section 4.1.1, and $(U_t, Z_t)' \sim$ IID $N(\mathbf{0}, \sigma_*^2 \mathbf{I}_2)$ such that $(\alpha_*, \xi_*, \pi_*, \sigma_*^2, \lambda_*) = (1, 1, 1, 1, 1)$. This DGP is different from that in Section 4.1.1 as Z_t is included as a regressor and appears in the conditional variance of the disturbance. Given this DGP, the following model for $\mathbb{E}[Y_t|X_t]$ is specified

$$\mathcal{M} = \{m_t(\cdot) : m_t(\alpha, \xi, \beta, \gamma) = \alpha + \xi X_t + \pi Z_t + \beta X_t^{\gamma}, \gamma \in \mathbf{\Gamma}\}.$$

As the disturbance exhibits conditional heteroskedasticity, we do not use the critical values in Table 1 for the QLR test. Instead, we apply the weighted bootstrap in Hansen (1996) to obtain critical values. The previous literature (e.g., Hansen, 1996; and Cho and White, 2010) studies applications of the weighted bootstrap to tests with unidentified parameters, similar to our model here, and readers are referred to Cho, Ishida, and White (2011, 2013) for discussion of the procedure.

We conduct experiments using the QLR test under the DGP and using model $\mathcal{M}$. As before, we let the sample size be 50, 100, 200, 300, 400, and 500. The number of iterations is also 10,000. We use test levels of 1%, 5%, and 10%. Results are reported in Table 3. The main findings are as follows. First, when the sample size is greater than or equal to 100, the empirical rejection rates are very close to the desired nominal levels, giving a better outcome than the asymptotic critical values. In particular, this feature applies irrespective of the size of the parameter space, implying that the weighted bootstrap gives reliable results for test size. Although the results are not reported here, similar findings were obtained when the disturbance is conditionally homoskedasticity. Second, P-P plots of the QLR tests are shown in Fig. 3 and seen to be close to the 45$^{\text{o}}$ line for every parameter space considered, implying that the asymptotic null behavior of the QLR test is well delivered by the weighted bootstrap.

4.2 Testing for Power Transforms of a Trend Regressor

4.2.1 When Asymptotic Null Distribution is Used

Let the DGP of $\{(Y_t, D_t) : t = 1, 2, \ldots, n\}$ be

$$Y_t = \alpha_* + \eta_* D_t + \xi_* t + U_t,$$

where $D_t = \rho_* D_{t-1} + V_t$ such that $(U_t, V_t)' \sim$ IID $N(\mathbf{0}, \sigma_*^2 \mathbf{I}_2)$ and $(\alpha_*, \eta_*, \xi_*, \rho_* \sigma_*^2) = (1, 1, 0, 0.5, 1)$. To generate stationary D_t, we generate data from $t = -100$ with $D_{-100} = 0$ and discard observations prior to $t = 1$. The model is therefore a serially correlated time series with no trend. Given this DGP, we specify the following model with both linear and nonlinear time trends, given in weak trend form as

$$\mathcal{M}'' = \left\{ m_{n,t}(\cdot) : m_t(\alpha, \xi_n, \eta, \beta, \gamma) = \alpha + \eta D_t + \xi_n \left(\frac{t}{n}\right) + \lambda_n(\beta, \gamma) \left(\frac{t}{n}\right)^{\gamma}, \gamma \in \mathbf{\Gamma} \right\},$$

where $\xi_n := \xi n$ and $\lambda_n(\beta, \gamma) := \beta n^{\gamma}$. As before, we do not constrain the parameter space of α, ξ, η, and β. The only parameter space influencing the asymptotic null distribution of the QLR test is $\mathbf{\Gamma}$. As for the earlier experiment and for similar reasons, we consider the same eight parameter spaces: $[-0.20, 1.50]$, $[-0.10, 1.50]$, $[0.00, 1.50]$, $[0.10, 1.50]$, $[-0.20, 2.50]$, $[-0.10, 2.50]$, $[0.00, 2.50]$, and $[0.10, 2.50]$. The explanatory variable D_t is included in the regressors to take account of the induced serial correlation in Y_t.

Theorem 5(*ii*) shows that the asymptotic null distribution of the QLR test is a functional of the power Gaussian process. We therefore apply the same critical values of Table 1 to the QLR statistic for testing neglected nonlinearity.

The simulation results are shown in Table 4, which follows the same format as Table 2 for the case of a stationary regressor. The findings are as follows. First, the empirical rejection rates approach nominal levels in each case as the sample size increases, corroborating Theorem 5(*ii*). Fig. 4 provides further confirmation, showing that the empirical and asymptotic null distributions of the QLR test are close, as in Figure 2. Second, convergence to nominal levels is again slower when the lower bound of $\mathbf{\Gamma}$ is close to -0.50 for the same upper bound, but improves as the upper bound of $\mathbf{\Gamma}$ is larger for a given lower bound. These results imply that level distortion can be reduced in practical work, as before, by using larger upper bounds and lower bounds distant from -0.5.

4.2.2 When Asymptotic Null Distribution cannot be Applied

We now suppose that U_t does not obey the conditional homoskedasticity condition $\mathbb{E}[U_t^2\mathbf{Z}_t] \neq \sigma_*^2\mathbb{E}[\mathbf{Z}_t]$ and modify the DGP to the following

$$Y_t = \alpha_* + \eta_* D_t + \xi_* t + \cos(D_t)U_t, \qquad (17)$$

where the conditional error variance is $\cos(D_t)^2$, and the other elements of the model, viz., $D_t = \rho_* D_{t-1} + V_t$, $(U_t, V_t) \sim$ IID $N(\mathbf{0}, \sigma_*^2\mathbf{I}_2)$ and $(\alpha_*, \eta_*, \sigma_*^2, \rho_*) = (1, 1, 0, 1, 0.5)$, are as before. Stationarity of D_t is assured in the same way as before.

Although the residual in (17) is conditionally heteroskedastic, we can apply the critical values in Table 1 to the QLR test. Theorem 5 shows that the asymptotic null distribution of the QLR test is determined by the distribution of the power Gaussian process. Simulation results are reported in Table 5 and Fig. 5 and are very close to those in Table 4 and Fig. 4, corroborating Theorem 5.

We may also use the weighted bootstrap when the time trend is included in the regression, following the same procedure as earlier and simply replacing X_t with the sample fraction t/n. The results are shown in Table 6 and Fig. 6. The estimated p-values of the QLR test are close to nominal levels even for moderate values of n, such as $n = 100$, which improves on the use of asymptotic critical values. From Fig. 6, the P-P plots of the QLR test statistics are close to the 45$^{\text{o}}$ line, again implying that the asymptotic null behavior of the QLR test is well delivered by the weighted bootstrap.

5 Conclusion

Linear models continue to be the mainstay of much empirical research, making specification tests of linearity an important feature of model robustness checks. Power transforms offer a natural alternative to linearity and provide a more general framework than simple polynomial specifications. However, as this paper demonstrates, tests of linearity in models that allow for power transforms of regressors raise critical issues of identification, producing what we have called a trifold identification problem that affects hypothesis testing. The approach adopted here resolves these issues by using a quasi-likelihood ratio statistic to provide a unified mechanism for capturing the trifold forms of the null hypothesis. The QLR statistic deals with the identification issues and delivers a convenient test for use in practical work with both microeconometric and time series data.

Under some weak regularity conditions, the asymptotic null distribution of the QLR test statistic is shown to be a simple functional of a Gaussian stochastic process. The methodology and limit theory for the stationary regressor case is extended to a model with a time trend and stationary regressors, facilitating tests for neglected nonlinearity with respect to trend. For such cases, the QLR test has an asymptotic null distribution that takes the form of a functional of a power Gaussian process when the disturbances form a martingale difference sequence. Asymptotic critical values of the QLR test are obtained and an alternative weighted bootstrap approach is explored to improve size control in testing. Simulations confirm the asymptotic theory and strongly affirm the use of the weighted bootstrap in reducing level distortion in tests of linearity both with stationary regressors and time trends.

6 Appendix: Proofs

Proof of Theorem 1: (*i*) This part is proved in the text.

(*ii*) We consider the numerator and denominator separately. The scaled numerator is $n^{-1/2}\mathbf{X}(\cdot)'\mathbf{M}\mathbf{U}$ and the uniform law of large numbers (ULLN) can be applied to $\{n^{-1}\sum_{t=1}^{n} X_t^{\gamma}\mathbf{Z}_t\}$, so that for each $j = 1, 2, \ldots, 2+k$,

$$\sup_{\gamma\in\mathbf{\Gamma}}\left|n^{-1}\sum_{t=1}^{n} X_t^{\gamma} Z_{t,j} - \mathbb{E}[X_t^{\gamma} Z_{t,j}]\right| \overset{\mathbb{P}}{\to} 0, \tag{18}$$

where $Z_{t,j}$ is the j-th row element of $\mathbf{Z}_t$. This result mainly follows from theorem 3(a) of Andrews (1992). In particular, Assumption 1(*iii*) implies that $\mathbf{\Gamma}$ is totally bounded; for $j = 1, 2, \cdots, k+2$, $\mathbb{E}[|X_t^{\gamma} Z_{t,j}|] \leq \mathbb{E}[M_t^2] < \infty$ by Assumption 2(*iii* and *iv*), so that for each $\gamma \in \mathbf{\Gamma}$, the ergodic theorem holds for $n^{-1}\sum_{t=1}^{n} X_t^{\gamma}$ $Z_{t,j}$; and finally $X_t^{(\cdot)} Z_{t,j}$ is Lipschitz continuous because for each j,

$$|X_t^{\gamma} Z_{t,j} - X_t^{\gamma'} Z_{t,j}| \leq \sup_{\gamma\in\mathbf{\Gamma}} |X_t^{\gamma}\log(X_t)| \cdot |Z_{t,j}| \cdot |\gamma - \gamma'| \leq M_t^2 |\gamma - \gamma'|, \tag{19}$$

where $M_t^2 = O_p(1)$. These three conditions are the assumptions required in theorem 3(a) of Andrews (1992) to prove the ULLN. This also implies that $\mathbb{E}[X_t^{(\cdot)}\mathbf{Z}_t]$ is continuous on $\mathbf{\Gamma}$. Since $n^{-1}\sum_{t=1}^{n}\mathbf{Z}_t\mathbf{Z}_t' \overset{\mathbb{P}}{\to} \mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t']$ by ergodicity we obtain

$$\sup_{\gamma\in\mathbf{\Gamma}}\left|n^{-1/2}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U} - n^{-1/2}\{\mathbf{X}(\gamma)'\mathbf{U} - \mathbb{E}[X_t^{\gamma}\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t']^{-1}\mathbf{Z}'\mathbf{U}\}\right| = o_p(1).$$

We can therefore show that $n^{-1/2}\{\mathbf{X}(\cdot)'\mathbf{U} - \mathbb{E}[X_t^{(\cdot)}\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t']^{-1}\mathbf{Z}'\mathbf{U}\} \Rightarrow \mathcal{G}(\cdot)$. For this, we apply the CLT to $n^{-1/2}\mathbf{Z}'\mathbf{U}$, so that $n^{-1/2}\mathbf{Z}'\mathbf{U} \overset{\mathrm{A}}{\sim} N(\mathbf{0}, \mathbb{E}[U_t^2\mathbf{Z}_t\mathbf{Z}_t'])$. Next, $X_t^{(\cdot)}U_t$ is Lipschitz continuous, so that

$|X_t^\gamma U_t - X_t^{\gamma'} U_t| \leq \sup_{\gamma\in\mathbf{\Gamma}} |X_t^\gamma \log(X_t)| \cdot |U_t| \cdot |\gamma - \gamma'| \leq M_t^2 |\gamma-\gamma'|$ by Assumption 2(*iii* and *iv*), so that

$$\mathbb{E}\left[\sup_{|\gamma-\gamma'|\leq\eta} |X_t^\gamma U_t - X_t^{\gamma'} U_t|^{2r}\right]^{\frac{1}{2r}} \leq \mathbb{E}[M_t^{4r}]^{\frac{1}{2r}}\eta. \tag{20}$$

This implies that $\{n^{-1/2}\mathbf{X}(\cdot)'\mathbf{U}\}$ is tight because Ossiander's L^{2r} entropy is finite by theorem 1 of Doukhan, Massart, and Rio (1995). We further note that (19) implies that for some $c > 0$,

$$\mathbb{E}\left[\sup_{|\gamma-\gamma'|<\eta} |\mathbb{E}[(X_t^\gamma - X_t^{\gamma'})\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}\mathbf{Z}_t U_t|^{2r}\right]^{\frac{1}{2r}} \leq c\mathbb{E}[M_t^{4r}]^{\frac{1}{2r}}\mathbb{E}[M_t^2]\eta, \tag{21}$$

implying that $\{n^{-1/2}\mathbb{E}[X_t^{(\cdot)}\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}\mathbf{Z}'\mathbf{U}\}$ is tight. Hence $\{n^{-1/2}(\mathbf{X}(\cdot)'\mathbf{U} - \mathbb{E}[X_t^{(\cdot)}\mathbf{Z}_t']\,\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}\mathbf{Z}'\mathbf{U})\}$ is also tight. Furthermore, the finite-dimensional multivariate CLT holds by the martingale CLT. It follows that $n^{-1/2}\{\mathbf{X}(\cdot)'\mathbf{U} - \mathbb{E}[X_t^{(\cdot)}\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}\mathbf{Z}'\mathbf{U}\} \Rightarrow \mathcal{G}(\cdot)$, implying that $n^{-1/2}\mathbf{X}(\cdot)'\mathbf{M}\mathbf{U} \Rightarrow \mathcal{G}(\cdot)$.

Second, we apply the ULLN to $n^{-1}\mathbf{X}(\cdot)'\mathbf{M}\mathbf{X}(\cdot)$. We separate our proof into two parts: we first show that $\sup_{\gamma\in\mathbf{\Gamma}} |n^{-1}\mathbf{X}(\gamma)'\mathbf{X}(\gamma) - \mathbb{E}[X_t^{2\gamma}]| = o_p(1)$ and next show that $\sup_{\gamma\in\mathbf{\Gamma}} |n^{-1}\mathbf{X}(\gamma)'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}(\gamma) - \mathbb{E}[X_t^\gamma\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}\mathbb{E}[\mathbf{Z}_t\mathbf{X}_t^\gamma]| = o_p(1)$. It then follows that

$$\sup_{\gamma\in\mathbf{\Gamma}}\left|n^{-1}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma) - \mathbb{E}[X_t^{2\gamma}] + (\mathbb{E}[X_t^\gamma\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}\mathbb{E}[\mathbf{Z}_t X_t^\gamma])\right| = o_p(1).$$

For this goal, we first note that $X_t^{2(\cdot)}$ is Lipschitz continuous, so that

$$|X_t^{2\gamma} - X_t^{2\gamma'}| \leq 2\sup_{\gamma\in\mathbf{\Gamma}}|X_t^{2\gamma}\log(X_t)|\cdot|\gamma-\gamma'| \leq 2\sup_{\gamma\in\mathbf{\Gamma}}|X_t^\gamma\log(X_t)|\cdot\sup_{\gamma\in\mathbf{\Gamma}}|X_t^\gamma|\cdot|\gamma-\gamma'| \leq 2M_t^2|\gamma-\gamma'|,$$

and $2M_t^2 = O_p(1)$ by Assumption 2(*iii* and *iv*). Theorem 3 of Andrews (1992) now shows that the ULLN holds for $\{n^{-1}\sum_{t=1}^n X_t^{2(\cdot)} - \mathbb{E}[X_t^{2(\cdot)}]\}$. We next note that

$$\begin{aligned}
&\sup_{\gamma\in\mathbf{\Gamma}}\left|n^{-1}\mathbf{X}(\gamma)'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}(\gamma) - \mathbb{E}[X_t^\gamma\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}\mathbb{E}[\mathbf{Z}_t\mathbf{X}_t^\gamma]\right| \\
&\quad\leq \sup_{\gamma\in\mathbf{\Gamma}}\left|(n^{-1}\mathbf{X}(\gamma)'\mathbf{Z} - \mathbb{E}[X_t^\gamma\mathbf{Z}_t'^{-1}\mathbf{Z}'\mathbf{Z})^{-1}n^{-1}\mathbf{Z}'\mathbf{X}(\gamma)\right| \\
&\qquad+ \sup_{\gamma\in\mathbf{\Gamma}}\left|\mathbb{E}[X_t^\gamma\mathbf{Z}_t'^{-1}\mathbf{Z}'\mathbf{Z})^{-1} - \mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1})n^{-1}\mathbf{Z}'\mathbf{X}(\gamma)\right| \\
&\qquad\quad+ \sup_{\gamma\in\mathbf{\Gamma}}\left|\mathbb{E}[X_t^\gamma\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}(n^{-1}\mathbf{Z}'\mathbf{X}(\gamma) - \mathbb{E}[\mathbf{Z}_t X_t^\gamma])\right|.
\end{aligned}$$

Hence, $\sup_{\gamma\in\mathbf{\Gamma}}\left|(n^{-1}\mathbf{X}(\gamma)'\mathbf{Z} - \mathbb{E}[X_t^\gamma\mathbf{Z}_t'])\right| = o_p(1)$ by (18), and $(n^{-1}\mathbf{Z}'\mathbf{Z})^{-1} - \mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1} = o_p(1)$ by

Assumption 2(*i* and *iii*) and ergodicity. Furthermore, $\sup_{\gamma\in\mathbf{\Gamma}} \left|n^{-1}\mathbf{X}(\gamma)'\mathbf{Z}\right| = O_p(1)$ by Assumption 2(*iii* and *iv*), so that $\sup_{\gamma\in\mathbf{\Gamma}} |\mathbb{E}[X_t^{\gamma}\mathbf{Z}_t]| = O(1)$. Therefore,

$$\sup_{\gamma\in\mathbf{\Gamma}} \left|(n^{-1}\mathbf{X}(\gamma)'\mathbf{Z} - \mathbb{E}[X_t^{\gamma}\mathbf{Z}_t'^{-1}\mathbf{Z}'\mathbf{Z})^{-1}n^{-1}\mathbf{Z}'\mathbf{X}(\gamma)\right|$$
$$\leq \sup_{\gamma\in\mathbf{\Gamma}} \left|(n^{-1}\mathbf{X}(\gamma)'\mathbf{Z} - \mathbb{E}[X_t^{\gamma}\mathbf{Z}_t'])\right| \cdot \left|(n^{-1}\mathbf{Z}'\mathbf{Z})^{-1}\right| \cdot \sup_{\gamma\in\mathbf{\Gamma}} \left|n^{-1}\mathbf{Z}'\mathbf{X}(\gamma)\right| = o_p(1),$$

where for an arbitrary function $\mathbf{f}(x) := [f_{i,j}(x)]$, we let $\sup_x |\mathbf{f}(x)| := [\sup_x |f_{i,j}(x)|]$. In a similar manner, it follows that

$$\sup_{\gamma\in\mathbf{\Gamma}} \left|\mathbb{E}[X_t^{\gamma}\mathbf{Z}_t'^{-1}\mathbf{Z}'\mathbf{Z})^{-1} - \mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1})n^{-1}\mathbf{Z}'\mathbf{X}(\gamma)\right|$$
$$\leq \sup_{\gamma\in\mathbf{\Gamma}} \left|\mathbb{E}[X_t^{\gamma}\mathbf{Z}_t']\right| \cdot \left|((n^{-1}\mathbf{Z}'\mathbf{Z})^{-1} - \mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1})\right| \cdot \sup_{\gamma\in\mathbf{\Gamma}} \left|n^{-1}\mathbf{Z}'\mathbf{X}(\gamma)\right| = o_p(1),$$

and

$$\sup_{\gamma\in\mathbf{\Gamma}} \left|\mathbb{E}[X_t^{\gamma}\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}(n^{-1}\mathbf{Z}'\mathbf{X}(\gamma) - \mathbb{E}[\mathbf{Z}_t X_t^{\gamma}])\right|$$
$$\leq \sup_{\gamma\in\mathbf{\Gamma}} \left|\mathbb{E}[X_t^{\gamma}\mathbf{Z}_t']\right| \cdot \left|\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}\right| \sup_{\gamma\in\mathbf{\Gamma}} \left|(n^{-1}\mathbf{Z}'\mathbf{X}(\gamma) - \mathbb{E}[\mathbf{Z}_t X_t^{\gamma}])\right| = o_p(1).$$

These two facts imply that $\sup_{\gamma\in\mathbf{\Gamma}} \left|n^{-1}\mathbf{X}(\gamma)'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}(\gamma) - \mathbb{E}[X_t^{\gamma}\mathbf{Z}_t']\mathbb{E}[\mathbf{Z}_t\mathbf{Z}_t'^{-1}\mathbb{E}[\mathbf{Z}_t\mathbf{X}_t^{\gamma}]\right| = o_p(1)$. Use of the continuous mapping theorem completes the proof. ■

Before proving Lemmas 1 and 2, we provide a supplementary lemma to assist in proving the main claims more efficiently.

Lemma A1. *Given Assumptions 1 and 3,*

(i) $\mathbf{L}_1'\mathbf{U} = O_p(\sqrt{n})$, $\mathbf{Z}'\mathbf{U} = O_p(\sqrt{n})$, $\mathbf{K}_1'\mathbf{U} = O_p(\sqrt{n})$;

(ii) $\mathbf{L}_1'\mathbf{Z} = O_p(n)$, $\mathbf{Z}'\mathbf{Z} = O_p(n)$, $\mathbf{K}_1'\mathbf{Z} = O_p(n)$;

(iii) $\mathbf{L}_1'\mathbf{L}_1 = O_p(n)$, $\mathbf{L}_1'\mathbf{K}_1 = O_p(n)$, $\mathbf{L}_2'\mathbf{U} = O_p(n)$, $\mathbf{L}_2'\mathbf{Z} = O_p(n)$, $\mathbf{K}_1'\mathbf{Z} = O_p(n)$, $\mathbf{K}_1'\mathbf{K}_1 = O_p(n)$, $\mathbf{K}_2'\mathbf{U} = O_p(n)$, *and* $\mathbf{K}_2'\mathbf{Z} = O_p(n)$*; and*

(iv) $\mathbf{L}_2'\mathbf{U} = o_p(n)$ *and* $\mathbf{K}_2'\mathbf{U} = o_p(n)$. □

Proof of Lemma A1: (*i*) By the definition of $\mathbf{K}_1 := [\mathbf{L}_1 \vdots \mathbf{0}]$, we note that if $\mathbf{L}_1'\mathbf{U} = O_p(\sqrt{n})$, $\mathbf{K}_1'\mathbf{U} = O_p(\sqrt{n})$. We, therefore, focus on proving that $\mathbf{L}_1'\mathbf{U} = O_p(\sqrt{n})$ and $\mathbf{Z}'\mathbf{U} = O_p(\sqrt{n})$. We also note that the structures of $\mathbf{L}_1'\mathbf{U}$ and $\mathbf{Z}'\mathbf{U}$ are identical. Accordingly, we let $\mathbf{R}$ be generic notation for $\mathbf{L}_1$ and $\mathbf{Z}$ and prove the given claims using $\mathbf{R}'\mathbf{U}$.

If we let $\mathbf{R} = [\mathbf{R}_{tj}]$, $\mathbf{R}'\mathbf{U} = \sum R_{tj}U_t$, which obeys the CLT if $\mathbb{E}[R_{tj}^2 U_t^2] < \infty$. We note that $\mathbb{E}[R_{tj}^2 U_t^2] \leq \mathbb{E}[R_{tj}^4]^{1/2}\mathbb{E}[U_t^4]^{1/2}$ by Cauchy-Schwarz, so the desired result follows since $\mathbb{E}[Z_{tj}^4] < \infty$, $\mathbb{E}[\log^4(X_t)] < \infty$, and $\mathbb{E}[U_t^4] < \infty$ by Assumption 3.

(*ii*) As in (*i*), if $\mathbf{L}_1'\mathbf{Z} = O_p(n)$, $\mathbf{K}_1'\mathbf{Z} = O_p(n)$ by the definition of $\mathbf{K}_1$. As before, we let $\mathbf{R}$ be generic notation for $\mathbf{L}_1$ and $\mathbf{Z}$ and prove the given claims using $\mathbf{R}'\mathbf{Z}$. As $\mathbf{R}'\mathbf{Z} = [\sum R_{tj}Z_{ti}]$, the result follows by ergodicity if $\mathbb{E}[|R_{tj}Z_{ti}|] < \infty$, which holds by virtue of Cauchy-Schwarz and the fact that $\mathbb{E}[\log^2(X_t)] < \infty$ and $\mathbb{E}[Z_{ti}^2] < \infty$ by Assumption 3.

(*iii*) By the definitions of $\mathbf{K}_1$ and $\mathbf{K}_2 := [\mathbf{L}_2 \vdots \mathbf{0}]$, if $\mathbf{L}_1'\mathbf{L}_1 = O_p(n)$, $\mathbf{L}_2'\mathbf{U} = O_p(n)$, $\mathbf{L}_2'\mathbf{Z} = O_p(n)$, and $\mathbf{L}_1'\mathbf{Z} = O_p(n)$ then $\mathbf{L}_1'\mathbf{K}_1 = O_p(n)$, $\mathbf{K}_1'\mathbf{Z} = O_p(n)$, $\mathbf{K}_2'\mathbf{U} = O_p(n)$, $\mathbf{K}_1'\mathbf{K}_1 = O_p(n)$, and $\mathbf{K}_2'\mathbf{Z} = O_p(n)$. We have already shown that $\mathbf{L}_1'\mathbf{Z} = O_p(n)$ in (*ii*). We, therefore, focus on proving $\mathbf{L}_1'\mathbf{L}_1 = O_p(n)$, $\mathbf{L}_2'\mathbf{U} = O_p(n)$, and $\mathbf{L}_2'\mathbf{Z} = O_p(n)$. Let $\mathbf{R}$ and $\mathbf{F}$ be generic notations for $\mathbf{L}_1$ or $\mathbf{L}_2$; and $\mathbf{L}_1$, $\mathbf{U}$, or $\mathbf{Z}$, respectively. For brevity, only $\mathbf{R}'\mathbf{F} = O_p(n)$ is proved and this follows in the same way by ergodicity, Cauchy-Schwarz and the moment conditions in Assumption 3 which ensure that $\mathbb{E}[\log^2(X_t)] < \infty$, $\mathbb{E}[\log^4(X_t)] < \infty$, $\mathbb{E}[U_t^2] < \infty$, and $\mathbb{E}[Z_{ti}^2] < \infty$.

(*iv*) From (*iii*), we note that the ergodic theorem applies to $n^{-1}\mathbf{L}_2'\mathbf{U}$ and $n^{-1}\mathbf{K}_2'\mathbf{U}$ and $\mathbb{E}[\log^2(X_t)U_t]$ $= 0$, so that $n^{-1}\mathbf{L}_2'\mathbf{U} = o_p(1)$ and $n^{-1}\mathbf{K}_2'\mathbf{U} = o_p(1)$, completing the proof. ■

Proof of Lemma 1: (*i*) This part is already proved in the text.

(*ii*) We partition the proof into three components. First, from the fact that $\mathbf{L}_1'\mathbf{M}\mathbf{U} = \mathbf{L}_1'\mathbf{U} - \mathbf{L}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}$ $\mathbf{Z}'\mathbf{U}$, Lemma A1(*i* and *ii*) and Assumption 3(*i*) imply that $\mathbf{L}_1'\mathbf{M}\mathbf{U} = O_p(\sqrt{n})$. Second, we note that $\mathbf{L}_1'\mathbf{M}\mathbf{L}_1 = \mathbf{L}_1'\mathbf{L}_1 - \mathbf{L}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{L}_1$, so that Lemma A1(*ii* and *iii*) and Assumption 3(*i*) imply that $\mathbf{L}_1'\mathbf{M}\mathbf{L}_1 = O_p(n)$. Third, $\mathbf{L}_2'\mathbf{M}\mathbf{U} = \mathbf{L}_2'\mathbf{U} - \mathbf{L}_2'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$. Lemma A1(*ii* and *iii*) and Assumption 3(*i*) imply that $\mathbf{L}_2'\mathbf{M}\mathbf{U} = O_p(n)$. Further, $\mathbf{L}_2'\mathbf{M}\mathbf{U} = \mathbf{L}_2'\mathbf{U} - \mathbf{L}_2'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$. Thus, $\mathbf{L}_2'\mathbf{M}\mathbf{U} = o_p(n)$ by Lemma A1(*iv*). Given these results, it now follows that the RHS of (8) is $O_p(1)$ as desired. ■

Proof of Lemma 2: (*i*) We can obtain the first-order derivative with respect to γ as follows:

$$L_n^{(1)}(0;\alpha) = 2\mathbf{P}(\alpha)'\mathbf{Q}(0)[\mathbf{Q}(0)'\mathbf{Q}(0)]^{-1}\mathbf{K}_1\mathbf{P}(\alpha) + \mathbf{P}(\alpha)'\mathbf{Q}(0)(d/d\gamma)[\mathbf{Q}(0)'\mathbf{Q}(0)]^{-1}\mathbf{Q}(0)'\mathbf{P}(\alpha).$$

We also note that

$$(d/d\gamma)[\mathbf{Q}(0)'\mathbf{Q}(0)]^{-1} = -(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}, \tag{22}$$

and that $\mathbf{P}(\alpha) = \mathbf{Y} - \alpha\boldsymbol{\iota} = \mathbf{Z}[\alpha_* - \alpha, \xi_*]' + \mathbf{U} = \mathbf{Z}\boldsymbol{\kappa}(\alpha) + \mathbf{U}$ by letting that $\boldsymbol{\kappa}(\alpha) := [\alpha_* - \alpha, \xi_*]'$. Going forward we suppress α of $\boldsymbol{\kappa}(\alpha)$ for notational simplicity. It follows that

$$\begin{aligned} L_n^{(1)}(0;\alpha) =& \underbrace{2(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})}_{(*)} \\ & \underbrace{-(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})}_{(**)}. \end{aligned}$$

We now examine each component on the right side. The first component (*) can be expressed as a sum of four other components: (a) $2\boldsymbol{\kappa}'\mathbf{Z}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{Z}\boldsymbol{\kappa} = 2\boldsymbol{\kappa}'\mathbf{K}_1'\mathbf{Z}\boldsymbol{\kappa}$; (b) $2\boldsymbol{\kappa}'\mathbf{K}_1'\mathbf{U}$; (c) $2\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{Z}\boldsymbol{\kappa} = 2\boldsymbol{\kappa}'\mathbf{Z}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$; and (d) $2\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{U}$. Next, the second component (**) can also be expressed as a sum of four components: (a) $-\boldsymbol{\kappa}'\mathbf{Z}'\mathbf{K}_1\boldsymbol{\kappa} - \boldsymbol{\kappa}'\mathbf{K}_1'\mathbf{Z}\boldsymbol{\kappa} = -2\boldsymbol{\kappa}'\mathbf{K}_1'\mathbf{Z}\boldsymbol{\kappa}$; (b) $-\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\,\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{K}_1\boldsymbol{\kappa} - \boldsymbol{\kappa}'\mathbf{K}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} = -2\boldsymbol{\kappa}'\mathbf{K}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$; (c) $-\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{Z}\boldsymbol{\kappa} - \boldsymbol{\kappa}'\mathbf{Z}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} = -2\,\boldsymbol{\kappa}'\mathbf{Z}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$; and (d) $-\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\,\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$. Adding and organizing all of these according to their orders of convergence yields the following

- (a) $2\boldsymbol{\kappa}'\mathbf{K}_1'\mathbf{Z}\boldsymbol{\kappa} - 2\boldsymbol{\kappa}'\mathbf{K}_1'\mathbf{Z}\boldsymbol{\kappa} = 0$;
- (b, c) $2\boldsymbol{\kappa}'\{\mathbf{K}_1' + \mathbf{Z}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}' - \mathbf{K}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}' - \mathbf{Z}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\}\mathbf{U} = 2(\alpha_* - \alpha)\mathbf{L}_1'\mathbf{M}\mathbf{U}$;
- (d) $2\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{U} - \mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$,

so that the first-order derivative is now obtained as

$$L_n^{(1)}(0;\alpha) = 2(\alpha_* - \alpha)\mathbf{L}_1'\mathbf{M}\mathbf{U} + 2\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - \mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}.$$

(*ii*) Given the result in (*i*), we note that $\mathbf{L}_1'\mathbf{M}\mathbf{U} = \mathbf{L}_1'\mathbf{U} - \mathbf{L}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$, and Lemma A1(*i* and *ii*) implies that $\mathbf{L}_1'\mathbf{M}\mathbf{U} = O_p(\sqrt{n})$. We also note that $\mathbf{K}_1'\mathbf{U} = [\mathbf{L}_1'\mathbf{U} \vdots \mathbf{0}] = O_p(\sqrt{n})$, so that Lemma A1(*i* and *ii*) implies that $\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} = O_p(1)$. Furthermore, Lemma A1(*i* and *ii*) implies that $\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} = O_p(1)$. Therefore,

$$\begin{aligned} L_n^{(1)}(0;\alpha) &= 2(\alpha_* - \alpha)\underbrace{\mathbf{L}_1'\mathbf{M}\mathbf{U}}_{O_p(\sqrt{n})} + 2\underbrace{\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}}_{O_p(1)} - \underbrace{\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}}_{O_p(1)} \\ &= 2(\alpha_* - \alpha)\mathbf{L}_1'\mathbf{M}\mathbf{U} + o_p(\sqrt{n}). \end{aligned}$$

(*iii*) The second-order derivative is

$$L_n^{(2)}(0;\alpha) = 2\mathbf{P}(\alpha)'\mathbf{K}_1[\mathbf{Q}(0)'\mathbf{Q}(0)]^{-1}\mathbf{K}_1'\mathbf{P}(\alpha) + 4\mathbf{P}(\alpha)'\mathbf{Q}(0)(d/d\gamma)[\mathbf{Q}(0)'\mathbf{Q}(0)]^{-1}\mathbf{K}_1'\mathbf{P}(\alpha)$$
$$+ 2\mathbf{P}(\alpha)'\mathbf{Q}(0)[\mathbf{Q}(0)'\mathbf{Q}(0)]^{-1}\mathbf{K}_2'\mathbf{P}(\alpha) + \mathbf{P}(\alpha)'\mathbf{Q}(0)(d^2/d\gamma^2)[\mathbf{Q}(0)'\mathbf{Q}(0)]^{-1}\mathbf{Q}(0)'\mathbf{P}(\alpha),$$

where

$$(d^2/d\gamma^2)[\mathbf{Q}(0)'\mathbf{Q}(0)]^{-1} = 2\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$$
$$- (\mathbf{Z}'\mathbf{Z})^{-1}(2\mathbf{K}_1'\mathbf{K}_1 + \mathbf{Z}'\mathbf{K}_2 + \mathbf{K}_2'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}, \tag{23}$$

and (22) already provides the specific form of $(d/d\gamma)[\mathbf{Q}(0)'\mathbf{Q}(0)]^{-1}$. Using these results and arranging them, we obtain the following second-order derivative:

$$L_n^{(2)}(0;\alpha) = 2(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})'\{\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1' + \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_2'\}(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})$$
$$- 4(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})$$
$$+ 2(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})$$
$$- (\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U})'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(2\mathbf{K}_1'\mathbf{K}_1 + \mathbf{Z}'\mathbf{K}_2 + \mathbf{K}_2'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{Z}\boldsymbol{\kappa} + \mathbf{U}). \tag{24}$$

We again organize this expression into three terms according to their orders:

- $2\boldsymbol{\kappa}'\{\mathbf{Z}'\mathbf{K}_1'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1' + \mathbf{K}_2'\}\mathbf{Z}\boldsymbol{\kappa} - 4\boldsymbol{\kappa}'(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{Z}\boldsymbol{\kappa} + 2\boldsymbol{\kappa}'(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})\boldsymbol{\kappa} - \boldsymbol{\kappa}'(2\mathbf{K}_1'\mathbf{K}_1 + \mathbf{Z}'\mathbf{K}_2 + \mathbf{K}_2'\mathbf{Z})\boldsymbol{\kappa} = 2\boldsymbol{\kappa}'\mathbf{K}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{K}_1\boldsymbol{\kappa} - 2\boldsymbol{\kappa}'\mathbf{K}_1'\mathbf{K}_1\boldsymbol{\kappa} = -2(\alpha_* - \alpha)^2\mathbf{L}_1'\mathbf{M}\mathbf{L}_1$;

- $4\boldsymbol{\kappa}'\mathbf{Z}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{U} - 4\boldsymbol{\kappa}'(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{U} - 4\boldsymbol{\kappa}'\mathbf{Z}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} + 2\boldsymbol{\kappa}'\mathbf{K}_2'\mathbf{U} + 2\boldsymbol{\kappa}'\mathbf{Z}'\mathbf{K}_2(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} + 4\boldsymbol{\kappa}'(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - 2\boldsymbol{\kappa}'(2\mathbf{K}_1'\mathbf{K}_1 + \mathbf{Z}'\mathbf{K}_2 + \mathbf{K}_2'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} = 2(\alpha_* - \alpha)[\mathbf{L}_2'\mathbf{M}\mathbf{U} - 2\mathbf{L}_1'\mathbf{M}\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - 2\mathbf{L}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{M}\mathbf{U}]$; and

- $2[\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{U} + \mathbf{U}'\mathbf{K}_2(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - 2\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}] + 2\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}[(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z}) - \mathbf{K}_1'\mathbf{K}_1 - \mathbf{Z}'\mathbf{K}_2](\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$.

Next apply Lemma A1 to each term. First, the proof of Lemma 1 has already shown that $\mathbf{L}_1'\mathbf{M}\mathbf{L}_1 = O_p(n)$ and $\mathbf{L}_2'\mathbf{M}\mathbf{U} = o_p(n)$. Second, $\mathbf{L}_1'\mathbf{M}\mathbf{K}_1 = \mathbf{L}_1'\mathbf{K}_1 - \mathbf{L}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{K}_1$. Assumption 3 and Lemma A1(*ii*, *iii*, and *iv*) now imply that $\mathbf{L}_1'\mathbf{M}\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} = o_p(n)$. Third, $\mathbf{K}_1'\mathbf{M}\mathbf{U} = \mathbf{K}_1'\mathbf{U} -$

$\mathbf{K}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} = o_p(n)$ by Lemma A1(*i* and *iv*), so that $\mathbf{L}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{M}\mathbf{U} = o_p(n)$ by Lemma A1(*ii* and *iii*). Therefore, $\mathbf{L}_2'\mathbf{M}\mathbf{U} - 2\mathbf{L}_1'\mathbf{M}\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - 2\mathbf{L}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{M}\mathbf{U} = o_p(n)$. Finally, we combine all components in Lemma A1 and obtain that $\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{U} + \mathbf{U}'\mathbf{K}_2(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - 2\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} + \mathbf{U}'\mathbf{Z}\ (\mathbf{Z}'\mathbf{Z})^{-1}[(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\ (\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z}) - \mathbf{K}_1'\mathbf{K}_1 - \mathbf{Z}'\mathbf{K}_2](\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} = o_p(n)$. Thus, the first, third, and final facts now imply that $L_n^{(2)}(0;\alpha) = -2(\alpha_* - \alpha)^2\mathbf{L}_1'\mathbf{M}\mathbf{L}_1 + o_p(n)$. This completes the proof. ■

Before proving Lemmas 4 and 5, we provide a supplementary lemma to assist in an efficient proof.

Lemma A2. *Given Assumptions 1 and 4,*

(i) $\mathbf{C}_1'\mathbf{U} = O_p(\sqrt{n})$, $\mathbf{Z}'\mathbf{U} = O_p(\sqrt{n})$, $\mathbf{J}_1'\mathbf{U} = O_p(\sqrt{n})$*;*

(ii) $\mathbf{C}_1'\mathbf{Z} = O_p(n)$, $\mathbf{Z}'\mathbf{Z} = O_p(n)$, $\mathbf{J}_1'\mathbf{Z} = O_p(n)$*;*

(iii) $\mathbf{C}_1'\mathbf{C}_1 = O_p(n)$, $\mathbf{C}_1'\mathbf{J}_1 = O_p(n)$, $\mathbf{C}_2'\mathbf{U} = O_p(n)$, $\mathbf{C}_2'\mathbf{Z} = O_p(n)$, $\mathbf{J}_1'\mathbf{Z} = O_p(n)$, $\mathbf{J}_1'\mathbf{J}_1 = O_p(n)$, $\mathbf{J}_2'\mathbf{U} = O_p(n)$, *and* $\mathbf{J}_2'\mathbf{Z} = O_p(n)$*; and*

(iv) $\mathbf{C}_2'\mathbf{U} = o_p(n)$ *and* $\mathbf{J}_2'\mathbf{U} = o_p(n)$. □

Proof of Lemma A2: (*i*) The plan of this proof is similar to that of Lemma A1. By the definition of $\mathbf{J}_1 := [\mathbf{0} \vdots \mathbf{C}_1 \vdots \mathbf{0}]$, we note that if $\mathbf{C}_1'\mathbf{U} = O_p(\sqrt{n})$, $\mathbf{J}_1'\mathbf{U} = O_p(\sqrt{n})$. We also note that the moment condition in Assumption 4(*iii*) is stronger than that of Assumption 3(*iii*). This implies that $\mathbf{Z}'\mathbf{U} = O_p(\sqrt{n})$ follows from Lemma A1(*i*). We therefore focus on proving $\mathbf{C}_1'\mathbf{U} = O_p(\sqrt{n})$.

From the definition of $\mathbf{C}_1'\mathbf{U}$, we note that $n^{-1/2}\mathbf{C}_1'\mathbf{U} = n^{-1/2}\sum_{t=1}^n X_t\log(X_t)U_t$, and we can apply the CLT if $\mathbb{E}[X_t^2\log^2(X_t)U_t^2] < \infty$. Note that $\mathbb{E}[X_t^2\log^2(X_t)U_t^2] \leq \mathbb{E}[X_t^4\log^4(X_t)]^{1/2}\ \mathbb{E}[U_t^4]^{1/2} \leq \mathbb{E}[X_t^8]^{1/4}\mathbb{E}[\log^8(X_t)]^{1/4}\mathbb{E}[U_t^4]^{1/2}$ by applying Cauchy-Schwarz. Each element in the right side is finite by Assumption 4(*iii.a*), so that $\mathbb{E}[X_t^2\log^2(X_t)U_t^2] < \infty$. Alternatively, $\mathbb{E}[X_t^2\log^2(X_t)U_t^2] \leq \mathbb{E}[X_t^4]^{1/2}\mathbb{E}[\log^4(X_t)U_t^4]^{1/2} \leq \mathbb{E}[X_t^4]^{1/2}\mathbb{E}[\log^8(X_t)]^{1/4}\mathbb{E}[U_t^8]^{1/4}$, and Assumption 4(*iii.b*) implies that the right side is finite. Finally, we note that $\mathbb{E}[X_t^2\log^2(X_t)U_t^2] \leq \mathbb{E}[\log^4(X_t)]^{1/2}\mathbb{E}[X_t^4U_t^4]^{1/2} \leq \mathbb{E}[\log^4(X_t)]^{1/2}\ \mathbb{E}[X_t^8]^{1/4}\mathbb{E}[U_t^8]^{1/4}$, and Assumption 4(*iii.c*) implies that the right side is finite. Thus, $\mathbf{C}_1'\mathbf{U} = O_p(\sqrt{n})$.

(*ii*) As in (*i*), if $\mathbf{C}_1'\mathbf{Z} = O_p(n)$, $\mathbf{J}_1'\mathbf{Z} = O_p(n)$ by the definition of $\mathbf{J}_1$. Furthermore, Lemma A1(*ii*) already shows that $\mathbf{Z}'\mathbf{Z} = O_p(n)$, and the current moment condition is stronger than Assumption 3(*iii*), so that $\mathbf{Z}'\mathbf{Z} = O_p(n)$. We therefore focus on proving $\mathbf{C}_1'\mathbf{Z} = O_p(n)$. By definition $n^{-1}\mathbf{C}_1'\mathbf{Z} = [n^{-1}\sum X_t\log(X_t)\ W_{t,j}]$, so that if $\mathbb{E}[|X_t\log(X_t)W_{t,j}|] < \infty$, the egodict theorem holds, giving the desired result. We first consider the case where $X_t = W_{t,j}$. If so, $\mathbb{E}[|X_t\log(X_t)W_{t,j}|] = \mathbb{E}[|X_t^2\log(X_t)|] \leq \mathbb{E}[X_t^4]^{1/2}\mathbb{E}[\log^2(X_t)]^{1/2} < \infty$ by Cauchy-Schwarz and Assumption 4(*iii*). Next consider the case where $X_t \neq W_{t,j}$: (*a*) $\mathbb{E}[|X_t\log$

$(X_t)W_{t,j}|] \le \mathbb{E}[|X_t \log(X_t)|^2]^{1/2}\mathbb{E}[W_{t,j}^2]^{1/2} \le \mathbb{E}[X_t^4]^{1/4}\mathbb{E}[|\log^4(X_t)]^{1/4}\mathbb{E}[W_{t,j}^2]^{1/2}$; (*b*) $\mathbb{E}[|X_t \log(X_t) W_{t,j}|] \le \mathbb{E}[|X_t W_{t,j}|^2]^{1/2}\mathbb{E}[\log^2(X_t)]^{1/2} \le \mathbb{E}[X_t^4]^{1/4}\mathbb{E}[W_{t,j}^4]^{1/4}\mathbb{E}[\log^2(X_t)]^{1/2}$; and finally (*c*) $\mathbb{E}[|X_t \log (X_t)W_{t,j}|] \le \mathbb{E}[|\log(X_t)W_{t,j}|^2]^{1/2}\mathbb{E}[X_t^2]^{1/2} \le \mathbb{E}[\log^4(X_t)]^{1/4}\mathbb{E}[W_{t,j}^4]^{1/4}\mathbb{E}[X_t^2]^{1/2}$ by Cauchy-Schwarz. Note that the elements on the right side of (*a*), (*b*), and (*c*) are finite by Assumption 4(*iii*).

(*iii*) By the definition of $\mathbf{J}_1$ and $\mathbf{J}_2 := [\mathbf{0} \vdots \mathbf{C}_2 \vdots \mathbf{0}]$, if $\mathbf{C}_1'\mathbf{C}_1 = O_p(n)$, $\mathbf{C}_2'\mathbf{U} = O_p(n)$, $\mathbf{C}_2'\mathbf{Z} = O_p(n)$, and $\mathbf{C}_1'\mathbf{Z} = O_p(n)$, then $\mathbf{C}_1'\mathbf{J}_1 = O_p(n)$, $\mathbf{J}_1'\mathbf{Z} = O_p(n)$, $\mathbf{J}_2'\mathbf{U} = O_p(n)$, $\mathbf{J}_1'\mathbf{J}_1 = O_p(n)$, and $\mathbf{J}_2'\mathbf{Z} = O_p(n)$. We have already shown that $\mathbf{C}_1'\mathbf{Z} = O_p(n)$ in (*ii*). We therefore focus on proving $\mathbf{C}_1'\mathbf{C}_1 = O_p(n)$, $\mathbf{C}_2'\mathbf{U} = O_p(n)$, and $\mathbf{C}_2'\mathbf{Z} = O_p(n)$.

We examine each case in turn. (*a*) Note that $n^{-1}\mathbf{C}_1'\mathbf{C}_1 = n^{-1}\sum X_t^2 \log^2(X_t)$, so that if $\mathbb{E}[X_t^2 \log^2(X_t)] < \infty$, the ergodic theorem holds. We also note that $\mathbb{E}[X_t^2 \log^2(X_t)] \le \mathbb{E}[X_t^4]^{1/2}\mathbb{E}[\log^4(X_t)]^{1/2}$, and the right side is finite by Assumption 4(*iii*). (*b*) Note that $n^{-1}\mathbf{C}_2'\mathbf{U} = n^{-1}\sum X_t \log^2(X_t) U_t$ and the ergodic theorem holds if $\mathbb{E}[|X_t \log^2(X_t)U_t|] < \infty$. Furthermore, we note that (*b.i*) $\mathbb{E}[|X_t \log^2(X_t)U_t|] \le \mathbb{E}[|X_t \log^2(X_t)|^2]^{1/2}\mathbb{E}[U_t^2]^{1/2} \le \mathbb{E}[X_t^4]^{1/4}\mathbb{E}[\log^8(X_t)]^{1/4}\mathbb{E}[U_t^2]^{1/2}$; (*b.ii*) $\mathbb{E}[|X_t \log^2(X_t)U_t|] \le \mathbb{E}[|U_t \log^2 (X_t)|^2]^{1/2}\mathbb{E}[X_t^2]^{1/2} \le \mathbb{E}[U_t^4]^{1/4}\mathbb{E}[\log^8(X_t)]^{1/4}\mathbb{E}[X_t^2]^{1/2}$; and finally (*b.iii*) $\mathbb{E}[|X_t \log^2(X_t)U_t|] \le \mathbb{E}[|U_t X_t|^2]^{1/2}\mathbb{E}[\log^4(X_t)]^{1/2} \le \mathbb{E}[|U_t|^4]^{1/4}\mathbb{E}[X_t^4]^{1/4}\mathbb{E}[\log^4(X_t)]^{1/2}$. We further note that each element forming the right sides of these upper bounds is finite by Assumption 4(*iii.a*), 4(*iii.b*), and 4(*iii.c*), respectively. Thus, $\mathbb{E}[|X_t \log^2(X_t)U_t|] < \infty$. (*c*) Finally, we note that $n^{-1}\mathbf{C}_2'\mathbf{Z} = [n^{-1}\sum X_t \log^2 (X_t)W_{t,j}]$, so that if $\mathbb{E}[|X_t \log^2(X_t)W_{t,j}|] < \infty$, the ergodic theorem applies. First, if $W_{t,j} = X_t$, the proof is the same as that for $\mathbb{E}[X_t^2 \log^2(X_t)] < \infty$, which we have just proved. Second, if $W_{t,j} \neq X_t$, by the same argument as in (*b*), (*c.i*) $\mathbb{E}[|X_t \log^2(X_t)W_{t,j}|] \le \mathbb{E}[X_t^4]^{1/4}\mathbb{E}[\log^8(X_t)]^{1/4}\ \mathbb{E}[W_{t,j}^2]^{1/2}$; (*c.ii*) $\mathbb{E}[|X_t \log^2(X_t)W_{t,j}|] \le \mathbb{E}[W_{t,j}^4]^{1/4}\mathbb{E}[\log^8(X_t)]^{1/4}\mathbb{E}[X_t^2]^{1/2}$; and (*c.iii*) $\mathbb{E}[|X_t \log^2(X_t)W_{t,j}|] \le \mathbb{E}[W_{t,j}^4]^{1/4}\mathbb{E}[X_t^4]^{1/4}\mathbb{E}[\log^4(X_t)\ ]^{1/2}$. Given these, the right sides in (*c.i*), (*c.ii*), and (*c.iii*) are finite if Assumption 4(*iii.a*) or 4(*iii.b*) holds; and furthermore, the right side in 4(*iii.c*) is finite if Assumption 4(*iii.c*) holds. Thus, $\mathbb{E}[|X_t \log^2(X_t)W_{t,j}|] < \infty$.

(*iv*) From the proof of (*iii*), the ergodic theorem applies to $n^{-1}\mathbf{C}_2'\mathbf{U}$ and $n^{-1}\mathbf{J}_2'\mathbf{U}$. Furthermore, $\mathbb{E}[X_t \log^2(X_t)U_t] = 0$, so that $n^{-1}\mathbf{C}_2'\mathbf{U} = o_p(1)$ and $n^{-1}\mathbf{J}_2'\mathbf{U} = o_p(1)$. This completes the proof. ■

Proof of Lemma 4: (*i*) This part is already proved in the text.

(*ii*) We partition the proof into three components. First, from the fact that $\mathbf{C}_1'\mathbf{M}\mathbf{U} = \mathbf{C}_1'\mathbf{U} - \mathbf{C}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$, Lemma A2(*i* and *ii*) and Assumption 4(*i*) imply that $\mathbf{C}_1'\mathbf{M}\mathbf{U} = O_p(\sqrt{n})$. Second, we note that $\mathbf{C}_1'\mathbf{M}\mathbf{C}_1 = \mathbf{C}_1'\mathbf{C}_1 - \mathbf{C}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{C}_1$, so that Lemma A2(*ii* and *iii*) and Assumption 4(*i*) imply that $\mathbf{C}_1'\mathbf{M}\mathbf{C}_1 = O_p(n)$. Third, $\mathbf{C}_2'\mathbf{M}\mathbf{U} = \mathbf{C}_2'\mathbf{U} - \mathbf{C}_2'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$. Lemma A2(*ii* and *iii*) and Assumption 4(*i*) imply that $\mathbf{C}_2'\mathbf{M}\mathbf{U} = O_p(n)$. Further, $\mathbf{C}_2'\mathbf{M}\mathbf{U} = \mathbf{C}_2'\mathbf{U} - \mathbf{C}_2'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}$. Thus, $\mathbf{C}_2'\mathbf{M}\mathbf{U} = o_p(n)$

by Lemma A2(*iv*). Given these findings, the desired result now follows. ■

Proof of Lemma 5: (*i*) The first-order derivative with respect to γ is

$$\frac{\partial}{\partial\gamma}L_n(\gamma;\xi) = 2\widetilde{\mathbf{P}}(\xi)'\widetilde{\mathbf{Q}}(\gamma)[\widetilde{\mathbf{Q}}(\gamma)'\widetilde{\mathbf{Q}}(\gamma)]^{-1}\frac{\partial}{\partial\gamma}\widetilde{\mathbf{Q}}(\gamma)'\widetilde{\mathbf{P}}(\xi) + \widetilde{\mathbf{P}}(\xi)'\widetilde{\mathbf{Q}}(\gamma)\frac{\partial}{\partial\gamma}[\widetilde{\mathbf{Q}}(\gamma)'\widetilde{\mathbf{Q}}(\gamma)]^{-1}\widetilde{\mathbf{Q}}(\gamma)'\widetilde{\mathbf{P}}(\xi).$$

When $\gamma = 1$, we can write the derivative as follows:

$$L_n^{(1)}(1;\xi) = 2\widetilde{\mathbf{P}}(\xi)'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{J}_1'\widetilde{\mathbf{P}}(\xi) + \widetilde{\mathbf{P}}(\xi)'\mathbf{Z}(d/d\gamma)[\widetilde{\mathbf{Q}}(1)'\widetilde{\mathbf{Q}}(1)]^{-1}\mathbf{Z}'\widetilde{\mathbf{P}}(\xi).$$

We also note that

$$(d/d\gamma)[\widetilde{\mathbf{Q}}(1)'\widetilde{\mathbf{Q}}(1)]^{-1} = -(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{J}_1 + \mathbf{J}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1} \tag{25}$$

and that $\widetilde{\mathbf{P}}(\xi) = (\mathbf{Y} - \xi\mathbf{X}) = \mathbf{Z}[\alpha_*, \xi_* - \xi, \boldsymbol{\eta}_*']' + \mathbf{Z}\mathbf{U} = \mathbf{Z}\boldsymbol{\zeta}(\xi) + \mathbf{U}$ by letting $\boldsymbol{\zeta}(\xi) := [\alpha_*, \xi_* - \xi, \boldsymbol{\eta}_*']'$. Going forward, we suppress ξ in $\boldsymbol{\zeta}(\xi)$ for notational simplicity. Then, it follows that

$$L_n^{(1)}(1;\xi) = 2(\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U})'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{J}_1'(\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U}) - (\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U})'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{J}_1+\mathbf{J}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U}).$$

(*ii*) We note that the form of $L_n^{(1)}(1;\xi)$ is identical to the form of $L_n^{(1)}(0;\alpha)$ in Lemma 2(*i*), provided that $(\xi_* - \xi)$, $\mathbf{C}_1$, and $\mathbf{J}_1$ are replaced by $(\alpha_* - \alpha)$, $\mathbf{L}_1$, and $\mathbf{K}_1$, respectively. Furthermore, the contents of Lemma A2 are also identical to those of Lemma A1, provided that $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{J}_1$, and $\mathbf{J}_2$ are replaced by $\mathbf{L}_1$, $\mathbf{L}_2$, $\mathbf{K}_1$, and $\mathbf{K}_2$, respectively. Thus, we can repeat the proof of Lemma 2(*ii*) for the proof here because Lemma 2(*ii*) holds as a corollary of Lemma A1.

(*iii*) We now examine the second-order derivative. We obtain

$$\begin{aligned} L_n^{(2)}(1;\xi) =& 2\widetilde{\mathbf{P}}(\xi)'\mathbf{J}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{J}_1'\widetilde{\mathbf{P}}(\xi) + 4\widetilde{\mathbf{P}}(\xi)'\mathbf{Z}(d/d\gamma)[\widetilde{\mathbf{Q}}(1)'\widetilde{\mathbf{Q}}(1)]^{-1}\mathbf{J}_1'\widetilde{\mathbf{P}}(\xi) \\ &+ 2\widetilde{\mathbf{P}}(\xi)'\mathbf{Z}(d/d\gamma)[\widetilde{\mathbf{Q}}(1)'\widetilde{\mathbf{Q}}(1)]^{-1}\mathbf{J}_1'\widetilde{\mathbf{P}}(\xi) + \widetilde{\mathbf{P}}(\xi)'\mathbf{Z}(d^2/d\gamma^2)[\widetilde{\mathbf{Q}}(1)'\widetilde{\mathbf{Q}}(1)]^{-1}\mathbf{Z}'\widetilde{\mathbf{P}}(\xi), \end{aligned}$$

where

$$\begin{aligned} (d^2/d\gamma^2)[\widetilde{\mathbf{Q}}(1)'\widetilde{\mathbf{Q}}(1)]^{-1} =& 2(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{J}_1 + \mathbf{J}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{J}_1 + \mathbf{J}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1} \\ &- (\mathbf{Z}'\mathbf{Z})^{-1}(2\mathbf{J}_1'\mathbf{J}_1 + \mathbf{Z}'\mathbf{J}_2 + \mathbf{J}_2'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}, \end{aligned} \tag{26}$$

and (25) already provides the form of $(d/d\gamma)[\widetilde{\mathbf{Q}}(1)'\widetilde{\mathbf{Q}}(1)]^{-1}$. Using these expressions and rearranging, we obtain the following second-order derivative:

$$
\begin{aligned}
L_n^{(2)}(1;\xi) = {} & 2(\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U})'\{\mathbf{J}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{J}_1' + \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{J}_2'\}(\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U}) \\
& - 4(\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U})'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{J}_1+\mathbf{J}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{J}_1'(\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U}) \\
& + 2(\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U})'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{J}_1+\mathbf{J}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{J}_1+\mathbf{J}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U}) \\
& - (\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U})'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(2\mathbf{J}_1'\mathbf{J}_1+\mathbf{Z}'\mathbf{J}_2+\mathbf{J}_2'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{Z}\boldsymbol{\zeta}+\mathbf{U}).
\end{aligned}
$$

We again note that the form of $L_n^{(2)}(1;\xi)$ is identical to that of $L_n^{(2)}(0;\alpha)$ in (24), provided that $\mathbf{J}_1$, $\mathbf{J}_2$, and $\boldsymbol{\zeta}$ are replaced by $\mathbf{K}_1$, $\mathbf{K}_2$, and $\boldsymbol{\kappa}$, respectively. Given Lemma A2, we may again repeat the proof of Lemma 2(*iii*) for the proof here as in the proof of (*ii*). ■

Proof of Lemma 7: (*i*) We examine the second-order derivative of $N_n(\gamma)$ and $D_n(\gamma)$ and let γ converge to zero. That is,

$$\operatorname*{plim}_{\gamma\to 0} N_n^{(2)}(\gamma) = \operatorname*{plim}_{\gamma\to 0} 2\left\{(d/d\gamma)\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\right\}^2 + 2\{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}\left\{(d^2/d\gamma^2)\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\right\} = 2\{\mathbf{L}_1\mathbf{M}\mathbf{U}\}^2$$

because $\operatorname{plim}_{\gamma\to 0}(d/d\gamma)\mathbf{X}(\gamma) = \mathbf{L}_1$ and $\operatorname{plim}_{\gamma\to 0}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U} = \boldsymbol{\iota}'\mathbf{M}\mathbf{U} = \mathbf{0}$. We further note that

$$\operatorname*{plim}_{\gamma\to 0}(d^2/d\gamma^2)D_n(\gamma) = \operatorname*{plim}_{\gamma\to 0} 2\left\{(d^2/d\gamma^2)\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma) + (d/d\gamma)\mathbf{X}(\gamma)'\mathbf{M}(d/d\gamma)\mathbf{X}(\gamma)\right\} = 2\mathbf{L}_1\mathbf{M}\mathbf{L}_1$$

because $\operatorname{plim}_{\gamma\to 0}(d^2/d\gamma^2)\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma) = \mathbf{L}_2'\mathbf{M}\boldsymbol{\iota} = \mathbf{0}$ and $\operatorname{plim}_{\gamma\to 0}(d/d\gamma)\mathbf{X}(\gamma) = \mathbf{L}_1$.

(*ii*) We now examine the second-order derivative of $N_n(\gamma)$ and $D_n(\gamma)$ and let γ converge to one. That is,

$$\operatorname*{plim}_{\gamma\to 1} N_n^{(2)}(\gamma) = \operatorname*{plim}_{\gamma\to 1} 2\left\{(d/d\gamma)\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\right\}^2 + 2\{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}\left\{(d^2/d\gamma^2)\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\right\} = 2\{\mathbf{C}_1\mathbf{M}\mathbf{U}\}^2$$

because $\operatorname{plim}_{\gamma\to 1}(d/d\gamma)\mathbf{X}(\gamma) = \mathbf{C}_1$ and $\operatorname{plim}_{\gamma\to 1}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U} = \mathbf{X}'\mathbf{M}\mathbf{U} = \mathbf{0}$. We also note that

$$\operatorname*{plim}_{\gamma\to 1} D_n^{(2)}(\gamma) = \operatorname*{plim}_{\gamma\to 1} 2\left\{(d^2/d\gamma^2)\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma) + (d/d\gamma)\mathbf{X}(\gamma)'\mathbf{M}(d/d\gamma)\mathbf{X}(\gamma)\right\} = 2\mathbf{C}_1\mathbf{M}\mathbf{C}_1,$$

from the fact that $\operatorname{plim}_{\gamma\to 1}(d^2/d\gamma^2)\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma) = \mathbf{C}_2'\mathbf{M}\mathbf{X} = \mathbf{0}$ and $\operatorname{plim}_{\gamma\to 1}(d/d\gamma)\mathbf{X}(\gamma) = \mathbf{C}_1$. ■

Proof of Theorem 4: (*i*, *ii*) These results hold as a corollary of Theorems 1, 2, and 3. ■

Before proving the main claims in Section 3, we provide the following supplementary lemmas to assist in delivering an efficient proof.

Lemma A3. *(i)* $(n\log(n))^{-1}\sum_{t=1}^{n}\log(t)\to 1$*;*

(ii) $(n\log^2(n))^{-1}\sum_{t=1}^{n}\log^2(t)\to 1$*;*

(iii) for each $\gamma\in(-1/2,\infty)$, $(n^{1+2\gamma}\log(n))^{-1}\sum_{t=1}^{n}t^{2\gamma}\log(t)\to 1/(2\gamma+1)$*; and*

(iv) for each $\gamma\in(-1/2,\infty)$, $(n^{1+2\gamma}\log^2(n))^{-1}\sum_{t=1}^{n}t^{2\gamma}\log^2(t)\to 1/(2\gamma+1)$. □

Proof of Lemma A3: (*i* and *ii*) This immediately follows from equation (26) of Phillips (2007) by letting his $L(\cdot)$ be $\log(\cdot)$.

(*iii* and *iv*) This also immediately follows from equation (55) of Phillips (2007). ■

Lemma A4. *Given the definition of* $s_{n,t}:=(t/n)$,

(i) for each $\gamma>-1$, $\frac{1}{n}\sum s_{n,t}^{\gamma}\to\int_0^1 s^{\gamma}ds=\frac{1}{1+\gamma}$*;*

(ii) for each $\gamma>-1$, $\frac{1}{n}\sum s_{n,t}^{\gamma}\log(s_{n,t})\to\int_0^1 s^{\gamma}\log(s)ds=-\frac{1}{(1+\gamma)^2}$*;*

(iii) for each $\gamma>-1$, $\frac{1}{n}\sum s_{n,t}^{\gamma}\log^2(s_{n,t})\to\int_0^1 s^{\gamma}\log^2(s)ds=-\frac{2}{(1+\gamma)^3}$*; and*

(iv) $\{n^{-1}\sum s_{n,t}^{(\cdot)}:\mathbf{\Gamma}\mapsto\mathbb{R}\}$ *is equicontinuous, where* $\mathbf{\Gamma}$ *is a convex and compact set in* $\mathbb{R}$. □

Proof of Lemma A4: (*i*, *ii*, and *iii*) These results are elementary.

(*iv*) We note that for some $\bar{\gamma}$ between γ and γ',

$$\left|\frac{1}{n}\sum s_{n,t}^{\gamma}-\frac{1}{n}\sum s_{n,t}^{\gamma'}\right|\le\frac{1}{n}\sum|s_{n,t}^{\bar{\gamma}}|\cdot|\log(s_{n,t})|\cdot|\gamma-\gamma'|\le\frac{1}{n}\sum|s_{n,t}|^{\gamma_o}\cdot|\log(s_{n,t})|\cdot|\gamma-\gamma'|,$$

where $\gamma_o:=\inf\mathbf{\Gamma}$. Also, $\frac{1}{n}\sum|s_{n,t}|^{\gamma_o}\cdot|\log(s_{n,t})|\to\frac{1}{\gamma_o+2}$. Therefore, for any $\epsilon>0$, if we let δ be $\epsilon(\gamma_o+2)$ and $|\gamma-\gamma'|<\delta$, $\limsup_{n\to\infty}|n^{-1}\sum s_{n,t}^{\gamma}-n^{-1}\sum s_{n,t}^{\gamma'}|\le\epsilon$. This completes the proof. ■

Lemma A5. *For a strictly stationary process* $\{\mathcal{Z}_t\}$ *and a deterministic sequence* $\{\xi_{n,t}\}$*, if we suppose that* $\mathbb{E}[|\mathcal{Z}_t|]<\infty$ *and* $\lim_{n\to\infty}\sum_{t=1}^{n}\xi_{n,t}=\xi_o\in(-\infty,\infty)$, $\sum_{t=1}^{n}\mathcal{X}_{n,t}\overset{a.s.}{\to}\xi_o\mathbb{E}[\mathcal{Z}_t]$*, where* $\mathcal{X}_{n,t}:=\xi_{n,t}\mathcal{Z}_t$. □

Proof of Lemma A5: We can apply the corollary in Billingsley (1995, p. 211). ■

Lemma A6. *We suppose that* $\{(U_t,\mathbf{D}_t')'\}$ *is a strictly stationary process. If for each* $j=1,2,\ldots,k$, $\mathbb{E}[D_{t,j}^4]<\infty$ *and* $\mathbb{E}[U_t^4]<\infty$*, then for each* $\gamma\in\mathbf{\Gamma}$ *with* $\inf\mathbf{\Gamma}>-1/2$,

(i) $n^{-1}\sum\mathbf{G}_{n,t}(\gamma)\mathbf{G}_{n,t}(\gamma)'\overset{a.s.}{\to}\widetilde{\mathbf{A}}(\gamma)$*; and*

(ii) $n^{-1}\sum U_t^2\mathbf{G}_{n,t}(\gamma)\mathbf{G}_{n,t}(\gamma)'\overset{a.s.}{\to}\widetilde{\mathbf{B}}(\gamma)$. □

Proof of Lemma A6: (*i* and *ii*) We let $\xi_{n,t}$ of Lemma A5 be $s_{n,t}^{2\gamma}/n$, $s_{n,t}^{\gamma+1}\log(s_{n,t})/n$, $s_{n,t}^{\gamma}\log(s_{n,t})/n$, $s_{n,t}^{\gamma}/n$, $s_{n,t}^{\gamma+1}/n$, $s_{n,t}^{2}\log^2(s_{n,t})/n$, $s_{n,t}\log^2(s_{n,t})/n$, $s_{n,t}\log(s_{n,t})/n$, $s_{n,t}^{2}\log(s_{n,t})/n$, $\log^2(s_{n,t})/n$, $\log(s_{n,t})/n$,

$s_{n,t}\log(s_{n,t})/n$, $s_{n,t}/n$, or $s_{n,t}^2/n$. Then, Lemma A4 implies that $\sum \xi_{n,t}$ converges to $1/(2\gamma+1)$, $-1/(\gamma+2)^2$, $-1/(\gamma+1)^2$, $1/(\gamma+1)$, $1/(\gamma+2)$, $2/27$, $1/4$, $-1/4$, $-1/9$, 2, -1, $-1/4$, $1/2$, or $1/3$, respectively. We let these limits be denoted by ξ_o. Lemma A5 implies that $\sum \xi_{n,t}\mathbf{D}_t$, $\sum \xi_{n,t}U_t^2$, and $\sum \xi_{n,t}U_t^2\mathbf{D}_t$ almost surely converge to $\xi_o\mathbb{E}[\mathbf{D}_t]$, $\xi_o\mathbb{E}[U_t^2]$ and $\xi_o\mathbb{E}[U_t^2\mathbf{D}_t]$, respectively. Finally, $n^{-1}\sum \mathbf{D}_t\mathbf{D}_t' \stackrel{\text{a.s.}}{\to} \mathbb{E}[\mathbf{D}_t\mathbf{D}_t']$ and $n^{-1}\sum U_t^2\mathbf{D}_t\mathbf{D}_t' \stackrel{\text{a.s.}}{\to} \mathbb{E}[U_t^2\mathbf{D}_t\mathbf{D}_t']$ by the ET and that $\mathbb{E}[D_{t,j}^4] < \infty$ and $\mathbb{E}[U_t^4] < \infty$. These limit results are sufficient for the desired results. ■

Lemma A7. *Given the definition of* $s_{n,t} := (t/n)$, *if for each* $j = 1, 2, \ldots, k$, $\mathbb{E}[|D_{t,j}|] < \infty$ *and* $\mathbf{\Gamma}$ *is a compact and convex subset in* $\mathbb{R}$ *such that* $\inf \mathbf{\Gamma} > -1$,

(i) $\sup_{\gamma\in\mathbf{\Gamma}} |n^{-1}\sum s_{n,t}^{\gamma} - \frac{1}{\gamma+1}| \to 0$*; and*

(ii) $\sup_{\gamma\in\mathbf{\Gamma}} |n^{-1}\sum s_{n,t}^{\gamma}D_{t,j} - \frac{1}{\gamma+1}\mathbb{E}[D_{t,j}]| \stackrel{\text{a.s.}}{\to} 0$. □

Proof of Lemma A7: (*i*) Lemma A4(*i* and *iv*) implies the desired result.

(*ii*) For each γ, Lemma A6(*i*) implies that $n^{-1}\sum s_{n,t}^{\gamma}D_{t,j} \stackrel{\text{a.s.}}{\to} \frac{1}{\gamma+1}\mathbb{E}[D_{t,j}]$. To show the desired result, we show the stochastic equicontinuity of $\{n^{-1}\sum s_{n,t}^{(\cdot)}D_{t,j} : \mathbf{\Gamma} \mapsto \mathbb{R}\}$. We note that

$$\left|\frac{1}{n}\sum s_{n,t}^{\gamma}D_{t,j} - \frac{1}{n}\sum s_{n,t}^{\gamma'}D_{t,j}\right| \le \frac{1}{n}\sum |s_{n,t}|^{\gamma_o}\cdot|\log(s_{n,t})|\cdot|D_{t,j}|\cdot|\gamma-\gamma'|,$$

where $\gamma_o := \inf \mathbf{\Gamma}$. This implies that for any $\epsilon > 0$,

$$\begin{aligned}\limsup_{n\to\infty} P\left(\sup_{|\gamma-\gamma'|<\delta}\left|\frac{1}{n}\sum s_{n,t}^{\gamma}D_{t,j} - \frac{1}{n}\sum s_{n,t}^{\gamma'}D_{t,j}\right| > \epsilon\right)&\\ \le \limsup_{n\to\infty} P\left(\frac{1}{n}\sum |s_{n,t}|^{\gamma_o}\cdot|\log(s_{n,t})|\cdot|D_{t,j}|\cdot\delta > \epsilon\right).&\end{aligned}$$

Therefore, if δ is sufficiently small, the right side can be made smaller than ϵ by using Fatou's lemma since $n^{-1}\sum |s_{n,t}|^{\gamma_o}\cdot|\log(s_{n,t})| \to 1/(\gamma_o+2)$, implying that $n^{-1}\sum |s_{n,t}|^{\gamma_o}\cdot|\log(s_{n,t})|\cdot|D_{t,j}| \stackrel{\text{a.s.}}{\to} [|D_{t,j}|]1/(\gamma_o+2)$ by Lemma A5. The desired result follows. ■

Proof of Lemma 8: We first note that Lemmas A3 and A4 show that $(n^{1+2\gamma})^{-1}\sum t^{2\gamma} = n^{-1}\sum s_{n,t}^{2\gamma} \to \frac{1}{2\gamma+1}$, $(n^{2+\gamma}\log(n))^{-1}\sum t^{1+\gamma}\log(t) \to \frac{1}{\gamma+2}$, $(n^{1+\gamma}\log(n))^{-1}\sum t^{\gamma}\log(t) \to \frac{1}{\gamma+1}$, $(n^{1+\gamma})^{-1}\sum t^{\gamma} \to \frac{1}{\gamma+1}$, $(n^{2+\gamma})^{-1}\sum t^{\gamma+1} \to \frac{1}{\gamma+2}$, $(n^3\log^2(n))^{-1}\sum t^2\log^2(t) \to \frac{1}{3}$, $(n^2\log^2(n))^{-1}\sum t\log^2(t) \to \frac{1}{2}$, $(n^2\log(n))^{-1}\sum t\log(t) \to \frac{1}{2}$, $(n^3\log(n))^{-1}\sum t^2\log(t) \to \frac{1}{3}$, $(n\log^2(n))^{-1}\sum\log^2(t) \to 1$, $(n\log(n))^{-1}\sum\log(t) \to 1$, $n^{-2}\sum t \to \frac{1}{2}$, and $n^{-3}\sum t^2 \to \frac{1}{3}$.

We also note that $n^{-1}\sum \mathbf{D}_t\mathbf{D}_t' \stackrel{\text{a.s.}}{\to} \mathbb{E}[\mathbf{D}_t\mathbf{D}_t']$ by ergodicity and $\mathbb{E}[D_{t,j}^2] < \infty$. If we further let $\xi_{n,t}$ of Lemma A5 be $t^{\gamma}/n^{1+\gamma}$, $t\log(t)/(n^2\log(n))$, $\log(t)/(n\log(n))$, $1/n$, or t/n^2, then $\sum \xi_{n,t}$ converges to $1/(\gamma+1)$, $\frac{1}{2}$, 1, 1, or, $\frac{1}{2}$, respectively. These facts and Lemma A5 imply that $\sum \xi_{n,t}\mathbf{D}_t$ almost surely converges to $\frac{1}{\gamma+1}\mathbb{E}[\mathbf{D}_t]$, $\frac{1}{2}\mathbb{E}[\mathbf{D}_t]$, $\mathbb{E}[\mathbf{D}_t]$, $\mathbb{E}[\mathbf{D}_t]$, or $\frac{1}{2}\mathbb{E}[\mathbf{D}_t]$, respectively. This completes the proof. ■

Proof of Theorem 5: (*i*) The proof is the same as the proof of Theorem 4(*i*), the the result follows simply by replacing $\mathbf{X}(\gamma)$ of Theorem 4(*i*) with $\mathbf{T}(\gamma)$.

(*ii*) We note that the QLR test statistic under $\widetilde{\mathcal{H}}_0$ is equal to

$$\sup_{\gamma\in\mathbf{\Gamma}} \frac{\{\mathbf{T}(\gamma)'\mathbf{M}\mathbf{U}\}^2}{\{\hat{\sigma}^2_{n,0}\mathbf{T}(\gamma)'\mathbf{M}\mathbf{T}(\gamma)\}} \tag{27}$$

by (*i*). In particular, if we let $\widetilde{\mathbf{L}}_1 := [\log(s_{n,1}),\ldots,\log(s_{n,1})]'$ and $\widetilde{\mathbf{C}}_1 := [s_{n,1}\log(s_{n,1}),\ldots,s_{n,n}\log(s_{n,n})]'$, the QLR test is equal to

$$\frac{\{\widetilde{\mathbf{L}}_1'\mathbf{M}\mathbf{U}\}^2}{\hat{\sigma}^2_{n,0}\{\widetilde{\mathbf{L}}_1'\mathbf{M}\widetilde{\mathbf{L}}_1\}} \quad \text{and} \quad \frac{\{\widetilde{\mathbf{C}}_1'\mathbf{M}\mathbf{U}\}^2}{\hat{\sigma}^2_{n,0}\{\widetilde{\mathbf{C}}_1'\mathbf{M}\widetilde{\mathbf{C}}_1\}} \tag{28}$$

under $\widetilde{\mathcal{H}}_0''$ and $\widetilde{\mathcal{H}}_0'''$, respectively. We separate the proof into three parts: (*a*), (*b*), and (*c*). In (*a*) and (*b*) we examine the denominators and the numerators of the statistics in (27) and (28), respectively, so that the asymptotic null behavior of the QLR test can be revealed by joint convergence. In (*c*) we derive the covariance structure given in the theorem.

(*a*) We examine the denominators of the statistics in (27) and (28). It is elementary to show that $\widehat{\sigma}^2_{n,0} \stackrel{\text{a.s.}}{\to} \sigma^2_*$ under $\widetilde{\mathcal{H}}_0$. Next note that Lemma A6(*i*) implies that $n^{-1}\widetilde{\mathbf{L}}'\mathbf{M}\widetilde{\mathbf{L}} \stackrel{\text{a.s.}}{\to} 2-\widetilde{\mathbf{A}}_{2,1}'\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{2,1}$ and $n^{-1}\widetilde{\mathbf{C}}'\mathbf{M}\widetilde{\mathbf{C}} \stackrel{\text{a.s.}}{\to} 2/27-\widetilde{\mathbf{A}}_{3,1}'\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{3,1}$, where

$$\widetilde{\mathbf{A}}_{2,1} := \begin{bmatrix} -1 \\ -\frac{1}{4} \\ -\mathbb{E}[\mathbf{D}_t] \end{bmatrix}, \quad \widetilde{\mathbf{A}}_{3,1} := \begin{bmatrix} -\frac{1}{4} \\ -\frac{1}{9} \\ -\frac{1}{4}\mathbb{E}[\mathbf{D}_t] \end{bmatrix}, \quad \text{and} \quad \widetilde{\mathbf{A}}_{1,1} := \begin{bmatrix} 1 & \frac{1}{2} & \mathbb{E}[\mathbf{D}_t'] \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{2}\mathbb{E}[\mathbf{D}_t'] \\ \mathbb{E}[\mathbf{D}_t] & \frac{1}{2}\mathbb{E}[\mathbf{D}_t] & \mathbb{E}[\mathbf{D}_t\mathbf{D}_t'] \end{bmatrix}.$$

We finally examine the denominator of $\left\{n^{-1/2}\mathbf{T}(\cdot)'\mathbf{M}\mathbf{U}\right\}/\{\hat{\sigma}^2_{n,0}n^{-1}\mathbf{T}(\cdot)'\mathbf{M}\mathbf{T}(\cdot)\}^{1/2}$. Observer that

$$n^{-1}\mathbf{T}(\gamma)'\mathbf{M}\mathbf{T}(\gamma) = n^{-1}\mathbf{T}(\gamma)'\mathbf{T}(\gamma) - n^{-1}\mathbf{T}(\gamma)'\mathbf{Z}(n^{-1}\mathbf{Z}'\mathbf{Z})^{-1}n^{-1}\mathbf{Z}'\mathbf{T}(\gamma),$$

and Lemma A6(*i*) implies that $n^{-1}\mathbf{T}(\gamma)'\mathbf{T}(\gamma)$, $n^{-1}\mathbf{Z}'\mathbf{T}(\gamma)$, and $n^{-1}\mathbf{Z}'\mathbf{Z}$ almost surely converges to $\widetilde{\mathbf{A}}_{4,4}(\gamma) := \frac{1}{2\gamma+1}$, $\widetilde{\mathbf{A}}_{4,1}(\gamma)$, and $\widetilde{\mathbf{A}}_{1,1}$, respectively, where $\widetilde{\mathbf{A}}_{4,1}(\gamma) := [\frac{1}{\gamma+1}, \frac{1}{\gamma+2}, \frac{1}{\gamma+1}\mathbb{E}[\mathbf{D}_t']]'$. Furthermore, Lemma A6(*i* and *ii*) implies that

$$\sup_{\gamma\in\mathbf{\Gamma}} |n^{-1}\mathbf{T}(\gamma)'\mathbf{T}(\gamma) - 1/(2\gamma+1)| \stackrel{\text{a.s.}}{\to} 0 \quad \text{and} \quad \sup_{\gamma\in\mathbf{\Gamma}} \|n^{-1}\mathbf{Z}'\mathbf{T}(\gamma) - \widetilde{\mathbf{A}}_{4,1}(\gamma)\|_\infty \stackrel{\text{a.s.}}{\to} 0.$$

Therefore,

$$\sup_{\gamma\in\boldsymbol{\Gamma}}\left|n^{-1}\mathbf{T}(\gamma)'\mathbf{M}\mathbf{T}(\gamma)-\{1/(2\gamma+1)-\widetilde{\mathbf{A}}_{4,1}(\gamma)'\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma)\}\right|\stackrel{\text{a.s.}}{\to}0, \tag{29}$$

since

$$\begin{aligned}
&\sup_{\gamma\in\boldsymbol{\Gamma}}\left|n^{-1}\mathbf{T}(\gamma)'\mathbf{M}\mathbf{T}(\gamma)-\{1/(2\gamma+1)-\widetilde{\mathbf{A}}_{4,1}(\gamma)'\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma)\}\right|\\
&\quad\le\sup_{\gamma\in\boldsymbol{\Gamma}}|n^{-1}\mathbf{T}(\gamma)'\mathbf{T}(\gamma)-1/(2\gamma+1)|+\sup_{\gamma\in\boldsymbol{\Gamma}}\left|\{n^{-1}\mathbf{T}(\gamma)'\mathbf{Z}-\widetilde{\mathbf{A}}_{4,1}(\gamma)\}'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{T}(\gamma)\right|\\
&\qquad+\sup_{\gamma\in\boldsymbol{\Gamma}}\left|\{\widetilde{\mathbf{A}}_{4,1}(\gamma)'^{-1}\mathbf{Z}'\mathbf{Z})^{-1}-\widetilde{\mathbf{A}}_{1,1}^{-1}\}\{n^{-1}\mathbf{Z}'\mathbf{T}(\gamma)\}\right|\\
&\qquad\quad+\sup_{\gamma\in\boldsymbol{\Gamma}}\left|\{\widetilde{\mathbf{A}}_{4,1}(\gamma)'\{\widetilde{\mathbf{A}}_{1,1}^{-1}\}\{n^{-1}\mathbf{T}(\gamma)'\mathbf{Z}-\widetilde{\mathbf{A}}_{4,1}(\gamma)\}\right|,
\end{aligned}$$

and each element on the right side almost surely converges to zero. This shows that $n^{-1}\mathbf{T}(\cdot)'\mathbf{M}\mathbf{T}(\cdot)$ obeys the ULLN. We further note that

$$\widetilde{\mathbf{A}}_{4,4}(\gamma)-\widetilde{\mathbf{A}}_{4,1}(\gamma)'\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma)=\frac{\sigma_*^2\gamma^2(\gamma-1)^2}{(\gamma+1)^2(\gamma+2)^2(2\gamma+1)}$$

by using the definition of $\widetilde{\mathbf{A}}_{4,4}(\gamma)$, $\widetilde{\mathbf{A}}_{4,1}(\gamma)$, and $\widetilde{\mathbf{A}}_{1,1}$. For notational simplicity, let the right side be $\sigma^2(\gamma,\gamma)$. If we combine all these limit results, it follows that

$$\begin{aligned}
&\left\{\sup_{\gamma\in\boldsymbol{\Gamma}}|n^{-1}\widehat{\sigma}_{n,0}^2\mathbf{T}(\gamma)'\mathbf{M}\mathbf{T}(\gamma)-\sigma^2(\gamma,\gamma)|,n^{-1}\widehat{\sigma}_{n,0}^2\widetilde{\mathbf{L}}'\mathbf{M}\widetilde{\mathbf{L}},n^{-1}\widehat{\sigma}_{n,0}^2\widetilde{\mathbf{C}}'\mathbf{M}\widetilde{\mathbf{C}},\right\}\\
&\qquad\stackrel{\text{a.s.}}{\to}\left\{0,\sigma_*^2(2-\widetilde{\mathbf{A}}_{2,1}'\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{2,1}),\sigma_*^2(2/27-\widetilde{\mathbf{A}}_{3,1}'\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{3,1})\right\}.
\end{aligned}\tag{30}$$

(*b*) We next examine the numerators of the statistics in (27) and (28). We first show that for each γ, $\{n^{-1/2}\mathbf{T}(\gamma)'\mathbf{M}\mathbf{U},n^{-1/2}\widetilde{\mathbf{L}}_1'\mathbf{M}\mathbf{U},n^{-1/2}\widetilde{\mathbf{C}}_1'\mathbf{M}\mathbf{U}\}$ weakly converges to a multivariate normal variate. We note that

$$n^{-1/2}\mathbf{T}(\gamma)'\mathbf{M}\mathbf{U}=n^{-1/2}\mathbf{T}(\gamma)'\mathbf{U}-(n^{-1}\mathbf{T}(\gamma)'\mathbf{Z})(n^{-1}\mathbf{Z}'\mathbf{Z})^{-1}(n^{-1/2}\mathbf{Z}'\mathbf{U}),$$

$$n^{-1/2}\widetilde{\mathbf{C}}_1'\mathbf{M}\mathbf{U}=n^{-1/2}\widetilde{\mathbf{C}}_1'\mathbf{U}-(n^{-1}\widetilde{\mathbf{C}}_1'\mathbf{Z})(n^{-1}\mathbf{Z}'\mathbf{Z})^{-1}(n^{-1/2}\mathbf{Z}'\mathbf{U}),\quad\text{and}$$

$$n^{-1/2}\widetilde{\mathbf{L}}_1'\mathbf{M}\mathbf{U}=n^{-1/2}\widetilde{\mathbf{L}}_1'\mathbf{U}-(n^{-1}\widetilde{\mathbf{L}}_1'\mathbf{Z})(n^{-1}\mathbf{Z}'\mathbf{Z})^{-1}(n^{-1/2}\mathbf{Z}'\mathbf{U}),$$

and (30) implies that for each γ, $\{n^{-1}\mathbf{T}(\gamma)'\mathbf{Z},n^{-1}\widetilde{\mathbf{L}}_1'\mathbf{Z},n^{-1}\widetilde{\mathbf{C}}_1'\mathbf{Z},n^{-1}\mathbf{Z}'\mathbf{Z}\}$ has its own almost sure limit. Furthermore, for each $\gamma\in\boldsymbol{\Gamma}\ \setminus\{0,1\}$, $\{U_t\mathbf{G}_{n,t}(\gamma),\mathcal{F}_t\}$ is an MDS and we can apply McLeish's (1974) CLT. Assumption 7 implies that $n^{-1}\sum\mathbb{E}[U_t^2\mathbf{G}_t(\gamma)\ \mathbf{G}_t(\gamma)']$ is uniformly positive definite with respect to

n. Thus, for each γ, $n^{-1/2}\sum U_t \mathbf{G}_t(\gamma) \overset{\text{A}}{\sim} N(\mathbf{0}, \widetilde{\mathbf{B}}(\gamma))$. We also note that for each $\gamma \in \mathbf{\Gamma}$, $\sum U_t \mathbf{G}_t(\gamma) = [\mathbf{T}(\gamma)'\mathbf{U}, \widetilde{\mathbf{C}}_1'\mathbf{U}, \widetilde{\mathbf{L}}_1'\mathbf{U}, (\mathbf{Z}'\mathbf{U})']'$, so that $\{n^{-1/2}\mathbf{T}(\gamma)'\mathbf{MU}, n^{-1/2}\widetilde{\mathbf{L}}_1'\mathbf{MU}, n^{-1/2}\widetilde{\mathbf{C}}_1'\mathbf{MU}\}$ weakly converges to a multivariate normal vector by joint convergence. We denote this weak limit by $\{\widetilde{\mathcal{G}}(\gamma), \widetilde{\mathcal{G}}_0, \widetilde{\mathcal{G}}_1\}$.

Similarly, we have finite dimensional convergence of the vectors $\{n^{-1/2}\mathbf{T}(\cdot)'\mathbf{MU}\}$. So we concentrate on showing that $\{n^{-1/2}\mathbf{T}(\cdot)'\mathbf{MU}\}$ is tight. As we have already shown in (*a*) that $\sup_{\gamma\in\mathbf{\Gamma}} |n^{-1}\mathbf{Z}'\mathbf{T}(\gamma) - \widetilde{\mathbf{A}}_{4,1}(\gamma)| \overset{\text{a.s.}}{\to} 0$ and $n^{-1}\mathbf{Z}'\mathbf{Z} \overset{\text{a.s.}}{\to} \widetilde{\mathbf{A}}_{1,1}$, if $\{n^{-1/2}\mathbf{T}(\cdot)'\mathbf{U}\}$ is tight, then $\{n^{-1/2}\mathbf{T}(\cdot)'\mathbf{MU}\}$ weakly converges to a Gaussian process. Without loss of generality, we let $\gamma' > \gamma$. Then, for some $\bar{\gamma}$ between γ and γ', $s_{n,t}^{\gamma} - s_{n,t}^{\gamma'} = s_{n,t}^{\bar{\gamma}} \log(s_{n,t}) \cdot (\gamma - \gamma') \leq s_{n,t}^{\gamma_o} |\log(s_{n,t})| \cdot |\gamma - \gamma'|$, where $\gamma_o := \inf_{\gamma} \mathbf{\Gamma}$, so that for any $\epsilon > 0$,

$$\limsup_{n\to\infty} P\left(\sup_{|\gamma-\gamma'|<\delta}\left|\frac{1}{\sqrt{n}}\sum s_{n,t}^{\gamma}U_t - \frac{1}{\sqrt{n}}\sum s_{n,t}^{\gamma'}U_t\right| > \epsilon\right)$$
$$\leq \limsup_{n\to\infty} P\left(\left|\frac{1}{\sqrt{n}}\sum |s_{n,t}|^{\gamma_o}\cdot|\log(s_{n,t})|\cdot U_t\right|\cdot\delta > \epsilon\right).$$

We further note that $n^{-1/2}\sum |s_{n,t}|^{\gamma_o}|\log(s_{n,t})|U_t \overset{\text{A}}{\sim} N(0, 2\sigma_*^2/(1+2\gamma_o)^3)$. Thus, if δ is sufficiently small, the right side can be made as small as desired. Hence, the random process sequence $\{n^{-1/2}\mathbf{T}(\cdot)'\mathbf{U}\}$ is tight, so that

$$\{n^{-1/2}\mathbf{T}(\cdot)'\mathbf{MU}, n^{-1/2}\widetilde{\mathbf{L}}'\mathbf{MU}, n^{-1/2}\widetilde{\mathbf{C}}'\mathbf{MU}\} \Rightarrow \{\widetilde{\mathcal{G}}(\cdot), \widetilde{\mathcal{G}}_0, \widetilde{\mathcal{G}}_1\}.$$

(*c*) Finally, we derive the covariance structure of the power Gaussian process. We first examine the limit covariance structure of the numerator in $\{\mathbf{T}(\cdot)'\mathbf{MU}\}/\{\hat{\sigma}_{n,0}^2\mathbf{T}(\cdot)'\mathbf{MT}(\cdot)\}^{1/2}$. Note that $\mathbf{T}(\gamma)'\mathbf{MU} = \mathbf{T}(\gamma)'\mathbf{U} - (\mathbf{T}(\gamma)'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{U})$, so that

$$\begin{aligned}\mathbf{T}(\gamma)'\mathbf{MUU}'\mathbf{MT}(\gamma') =& (\mathbf{T}(\gamma)'\mathbf{U})(\mathbf{U}'\mathbf{T}(\gamma')) \\ &- (\mathbf{T}(\gamma)'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\{(\mathbf{Z}'\mathbf{U})(\mathbf{U}'\mathbf{T}(\gamma'))\} - \left\{(\mathbf{T}(\gamma)'\mathbf{U})\left(\mathbf{U}'\mathbf{Z}\right)\right\}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{T}(\gamma') \\ &+ (\mathbf{T}(\gamma)'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\{(\mathbf{Z}'\mathbf{U})(\mathbf{U}'\mathbf{Z})\}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{T}(\gamma')).\end{aligned}$$

Lemma A6 shows that $n^{-1}\mathbf{T}(\gamma)'\mathbf{Z} \overset{\text{a.s.}}{\to} \widetilde{\mathbf{A}}_{4,1}(\gamma)'$ and $n^{-1}\mathbf{Z}'\mathbf{Z} \overset{\text{a.s.}}{\to} \widetilde{\mathbf{A}}_{1,1}$, respectively. This implies that

$$\begin{aligned}&n^{-1}\mathbf{T}(\gamma)'\mathbf{MUU}'\mathbf{MT}(\gamma') = n^{-1}(\mathbf{T}(\gamma)'\mathbf{U})(\mathbf{U}'\mathbf{T}(\gamma')) \\ &- n^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma)'\widetilde{\mathbf{A}}_{1,1}^{-1}\{(\mathbf{Z}'\mathbf{U})(\mathbf{U}'\mathbf{T}(\gamma'))\} - n^{-1}\left\{(\mathbf{T}(\gamma)'\mathbf{U})\left(\mathbf{U}'\mathbf{Z}\right)\right\}\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma') \\ &+ n^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma)'\widetilde{\mathbf{A}}_{1,1}^{-1}\{(\mathbf{Z}'\mathbf{U})(\mathbf{U}'\mathbf{Z})\}\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma') + o_p(1).\end{aligned} \tag{31}$$

To find the covariance structure of the limit process of $n^{-1/2}\mathbf{T}(\cdot)'\mathbf{MU}$, we consider the limit expectations of the terms on the right side of (31). First,

$$n^{-1}\mathbb{E}\left[\mathbf{T}(\gamma)'\mathbf{U}\mathbf{U}'\mathbf{T}(\gamma')\right] = n^{-1}\sum s_{n,t}^{\gamma+\gamma'}\mathbb{E}[U_t^2] \rightarrow \frac{\sigma_*^2}{\gamma+\gamma'+1}, \tag{32}$$

using Lemma A4(*i*) and the fact that $\{U_t, \mathcal{F}_t\}$ is an MDS. Second,

$$n^{-1}\mathbb{E}\left[(\mathbf{Z}'\mathbf{U})(\mathbf{U}'\mathbf{T}(\gamma')\right] = n^{-1}\sum s_{n,t}^{\gamma'}\mathbb{E}[U_t^2\mathbf{Z}_{n,t}] \rightarrow \widetilde{\mathbf{B}}_{4,1}(\gamma') := \left[\frac{\sigma_*^2}{\gamma'+1}, \frac{\sigma_*^2}{\gamma'+2}, \frac{1}{\gamma'+1}\mathbb{E}[U_t^2\mathbf{D}_t']\right]',$$

and so

$$\widetilde{\mathbf{A}}_{4,1}(\gamma)'\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{B}}_{4,1}(\gamma') = \frac{\sigma_*^2(4\gamma\gamma'+2\gamma+2\gamma'+4)}{(\gamma+1)(\gamma+2)(\gamma'+1)(\gamma'+2)}, \tag{33}$$

which is symmetric between γ and γ', thereby giving the limit of the expectation of the second and third terms of (31). Next observe that

$$n^{-1}\mathbb{E}[(\mathbf{Z}'\mathbf{U})(\mathbf{U}'\mathbf{Z})] = n^{-1}\sum \mathbb{E}[U_t^2\mathbf{Z}_{n,t}\mathbf{Z}_{n,t}'] \rightarrow \widetilde{\mathbf{B}}_{1,1} := \begin{bmatrix} \sigma_*^2 & \frac{1}{2}\sigma_*^2 & \mathbb{E}[U_t^2\mathbf{D}_t'] \\ \frac{1}{2}\sigma_*^2 & \frac{1}{3}\sigma_*^2 & \frac{1}{2}\mathbb{E}[U_t^2\mathbf{D}_t'] \\ \mathbb{E}[U_t^2\mathbf{D}_t] & \frac{1}{2}\mathbb{E}[U_t^2\mathbf{D}_t] & \mathbb{E}[U_t^2\mathbf{D}_t\mathbf{D}_t'] \end{bmatrix}$$

using Lemma A4(*ii*) and the fact that $\{U_t, \mathcal{F}_t\}$ is an MDS. Then,

$$\begin{aligned} & n^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma)'\widetilde{\mathbf{A}}_{1,1}^{-1}\mathbb{E}\{(\mathbf{Z}'\mathbf{U})(\mathbf{U}'\mathbf{Z})\}\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma') \\ = \; & \widetilde{\mathbf{A}}_{4,1}(\gamma)'\widetilde{\mathbf{A}}_{1,1}^{-1}\left\{\mathbb{E}[n^{-1}\sum U_t^2\mathbf{Z}_{n,t}\mathbf{Z}_{n,t}']\right\}\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma') \\ \rightarrow \; & \widetilde{\mathbf{A}}_{4,1}(\gamma)'\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{B}}_{1,1}\widetilde{\mathbf{A}}_{1,1}^{-1}\widetilde{\mathbf{A}}_{4,1}(\gamma') = \frac{\sigma_*^2(4\gamma\gamma'+2\gamma+2\gamma'+4)}{(\gamma+1)(\gamma+2)(\gamma'+1)(\gamma'+2)}. \end{aligned} \tag{34}$$

We combine all the limit results in (32), (33), and (34) to obtain the following limiting covariance kernel of the process $n^{-1/2}\mathbf{T}(\cdot)'\mathbf{M}\mathbf{U}$

$$\sigma(\gamma,\gamma') := \frac{\sigma_*^2\gamma\gamma'(\gamma-1)(\gamma'-1)}{(\gamma+1)(\gamma+2)(\gamma'+1)(\gamma'+2)(\gamma+\gamma'+1)}.$$

The limit behavior of the denominator of $\{\mathbf{T}(\cdot)'\mathbf{M}\mathbf{U}\}/\{\hat{\sigma}_{n,0}^2\mathbf{T}(\cdot)'\mathbf{M}\mathbf{T}(\cdot)\}^{1/2}$ is already given in (*a*). That is, $\hat{\sigma}_{n,0}^2 n^{-1}\mathbf{T}(\cdot)'\mathbf{M}\mathbf{T}(\cdot)$ almost surely converges to $\sigma^2(\cdot,\cdot)$ uniformly on $\mathbf{\Gamma}$. Therefore, using the definition

$$c(\gamma,\gamma') := \gamma\gamma'(\gamma-1)(\gamma'-1)/|\gamma\gamma'(\gamma-1)(\gamma'-1)|,$$

the covariance kernel of the limit $\widetilde{\mathcal{Z}}(\gamma)$ of the process $\{\mathbf{T}(\cdot)'\mathbf{MU}\}/\{\hat{\sigma}^2_{n,0}\mathbf{T}(\cdot)'\mathbf{MT}(\cdot)\}^{1/2}$ is given by

$$\tilde{\kappa}\left(\gamma,\gamma'\right) = \mathbb{E}[\widetilde{\mathcal{Z}}(\gamma)\widetilde{\mathcal{Z}}(\gamma')] = \frac{\sigma(\gamma,\gamma')}{\sqrt{\sigma^2(\gamma,\gamma)}\sqrt{\sigma^2(\gamma',\gamma')}} = c(\gamma,\gamma')\frac{(1+2\gamma)^{1/2}(1+2\gamma')^{1/2}}{(1+\gamma+\gamma')},$$

as stated. This completes the proof. ■

References

ANDREWS, D. (1992): "Generic Uniform Convergence," *Econometric Theory*, 8, 241–257.

BIERENS, H. (1990): "A Consistent Conditional Moment Test of Functional Form," *Econometrica*, 58, 1443–1458.

BILLINGSLEY, P. (1995): *Probability and Measure*. New York: John Wiley & Sons.

BOX, G. AND COX, D. (1964): "An Analysis of Transformations," *Journal of the Royal Statistical Society, Series B*, 26, 211-252.

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

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

CHO, J.S., ISHIDA, I., AND WHITE, H. (2013): "Testing for Neglected Nonlinearity Using Twofold Unidentified Models under the Null and Hexic Expansions," *Essays on Nonlinear Time Series Econometrics: A Festschrift in Honor of Timo Teräsvirta*. Eds. Niels Haldrup, Mika Meitz, and Pentti Saikkonen. Oxford: Oxford University Press, forthcoming.

CHO, J.S. AND WHITE, H. (2010): "Testing for Unobserved Heterogeneity in Exponential and Weibull Duration Models," *Journal of Econometrics*, 157 458–480.

DAVIES, R. (1977): "Hypothesis Testing When a Nuisance Parameter is Present only under the Alternative," *Biometrika*, 64, 247-254.

DAVIES, R. (1987): "Hypothesis Testing When a Nuisance Parameter is Present only under the Alternative," *Biometrika*, 74, 33-43.

DOUKHAN, P., MASSART, P., AND RIO, E. (1995): “Invariance Principles for Absolutely Regular Empirical Processes,” *Annales de l’Institut Henri Poincaré, Probabilites et Statistiques*, 31 393–427.

HANSEN, B. (1996): “Inference When a Nuisance Parameter is Not Identified under the Null Hypothesis,” *Econometrica*, 64, 413-430.

MCLEISH, D. (1974): “Dependent Central Limit Theorems and Invariance Principles,” *Annals of Probability*, 2, 620–628.

PARK, J.Y. AND PHILLIPS, P.C.B. (1999): “Asymptotics for Nonlinear transformations of Integrated Time Series,” *Econometric Theory*, 15, 269–298.

PHILLIPS, P.C.B. (2007): “Regression with Slowly Varying Regressors and Nonlinear Trends,” *Econometric Theory*, 23, 557–614.

RAMSEY, J.B. (1969): “Tests for Specification Errors in Classical Linear Least Squares Regression Analysis,” *Journal of Royal Statistical Society B*, 31, 350–371.

SAKIA, R. (1992): “The Box-Cox transformation Technique: A Review,” *The Statistician*, 41, 169–178.

SHI, X. AND PHILLIPS, P.C.B. (2012): “Nonlinear Cointegrating Regression under Weak Identification,” *Econometric Theory*, 28, 509–547.

STINCHCOMBE, M. AND WHITE, H. (1998): “Consistent Specification Testing with Nuisance Parameters Present Only under the Alternative,” *Econometric Theory*, 14, 295–324.

TUKEY, J. (1957): “On the Comparative Anatomy of Transformations,” *Annals of Mathematical Statistics*, 28, 602–632.

TUKEY, J. (1977): *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

WHITE, H. (2001): *Asymptotic Theory for Econometricians*. Orlando, FL: Academic Press.

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

WU, C.-F. (1981): “Asymptotic Theory of Nonlinear Least Squares Estimation,” *Annals of Statistics*, 9, 501–513.

Levels \ $\mathbf{\Gamma}$	$[-0.20, 1.50]$	$[-0.10, 1.50]$	$[0.00, 1.50]$	$[0.10, 1.50]$
10%	3.7186	3.6326	3.4669	3.4098
5%	4.9641	4.9065	4.7112	4.6196
1%	7.9861	7.9549	7.7336	7.6404
Levels \ $\mathbf{\Gamma}$	$[-0.20, 2.50]$	$[-0.10, 2.50]$	$[0.00, 2.50]$	$[0.10, 2.50]$
10%	3.9051	3.7636	3.6795	3.6334
5%	5.1772	5.0426	4.9499	4.8933
1%	8.2651	8.0254	8.0236	7.8715

Table 1: ASYMPTOTIC CRITICAL VALUES OF THE QLR TEST STATISTIC. This table contains the asymptotic critical values obtained by generating the approximated power Gaussian process 100,000 times. We used a grid search method to obtain the maximum of the squared Gaussian process. The grid distance is 0.01, and we let m be 500.

$\mathbf{\Gamma}$	Levels $\setminus n$	50	100	200	300	400	500
$[-0.20, 1.50]$	1%	0.59	0.73	0.80	0.78	0.96	0.84*
	5%	3.84	3.92	4.29	3.83	4.24	4.23*
	10%	8.22	8.27	8.86	8.01	8.73	8.68*
$[-0.10, 1.50]$	1%	0.66	0.66	0.55	0.73	0.79	0.82
	5%	4.21	3.94	3.95	4.34	4.26	4.63
	10%	8.63	8.22	8.11	8.92	8.99	9.52
$[0.00, 1.50]$	1%	0.83	0.86	0.99	0.91	0.85	0.91
	5%	4.73	4.75	5.19	4.77	4.87	5.15
	10%	9.52	9.67	10.22	9.87	9.94	10.52
$[0.10, 1.50]$	1%	0.86	0.93	1.10	0.78	1.17	0.94
	5%	4.88	5.03	5.45	4.75	5.43	5.03
	10%	9.75	10.03	10.10	9.80	10.32	10.15
$[-0.20, 2.50]$	1%	0.72	0.68	0.75	0.94	0.81	0.97†
	5%	4.28	4.36	4.14	4.62	4.21	4.55†
	10%	8.39	8.58	8.32	9.09	8.72	8.72†
$[-0.10, 2.50]$	1%	1.09	0.95	0.83	0.87	0.96	0.83
	5%	4.99	4.31	4.61	4.66	4.58	4.38
	10%	9.59	8.90	9.34	9.67	9.58	9.44
$[0.00, 2.50]$	1%	0.61	1.01	0.88	0.89	0.70	1.00
	5%	4.47	4.66	4.66	4.38	4.73	4.66
	10%	9.31	9.54	9.52	8.98	9.09	9.60
$[0.10, 2.50]$	1%	0.72	0.89	0.95	0.79	0.99	1.11
	5%	4.57	4.69	4.74	4.58	5.07	4.77
	10%	9.61	9.62	9.50	9.63	9.90	9.59

Table 2: LEVELS OF THE QLR TEST STATISTICS. Number of Repetitions: 10,000. MODEL: $Y_t = \alpha + \xi X_t + \beta X_t^{\gamma} + U_t$. DGP: $Y_t = \alpha_* + \xi_* X_t + U_t$, $X_t := \exp(-\lambda_* H_t)$, $U_t \sim$ IID $N(0, \sigma_*^2)$, $H_t \sim$ IID Exp(λ_*) such that U_t is independent of H_t and $(\alpha_*, \xi_*, \sigma_*^2, \lambda_*) = (1, 1, 1, 1)$. Notes. $*$: for $n = 50,000$, the empirical rejection rates are 0.87, 4.89, and 10.04 when the levels of significance are 1%, 5%, and 10%, respectively; †: for $n = 50,000$, the empirical rejection rates are 0.82, 4.67, and 9.23 when the levels of significance are 1%, 5%, and 10%, respectively.

Figure 1: A TYPICAL SAMPLE PATH OF $\mathcal{Z}(\cdot)$. A typical sample path of $\mathcal{Z}(\cdot)$ is provided, which is discontinuous at $\gamma = 0$ and 1 with probability 1. Furthermore, $\lim_{\gamma\uparrow 0} |\mathcal{Z}(\gamma)| = \lim_{\gamma\downarrow 0} |\mathcal{Z}(\gamma)|$ and $\lim_{\gamma\uparrow 1} |\mathcal{Z}(\gamma)| = \lim_{\gamma\downarrow 1} |\mathcal{Z}(\gamma)|$.

$\mathbf{\Gamma} = [-0.1, 1.5]$ $\quad$ $\mathbf{\Gamma} = [-0.1, 2.5]$

Figure 2: ASYMPTOTIC AND EMPIRICAL NULL DISTRIBUTIONS OF THE QLR STATISTICS WITH $\mathbf{\Gamma} = [-0.1, 1.5]$ AND $\mathbf{\Gamma} = [-0.1, 2.5]$. Number of Repetitions: 10,000. MODEL: $Y_t = \alpha + \xi X_t + \beta X_t^{\gamma} + U_t$. DGP: $Y_t = \alpha_* + \xi_* X_t + U_t$, $X_t := \exp(-\lambda_* H_t)$, $U_t \sim$ IID $N(0, \sigma_*^2)$, $H_t \sim$ IID Exp(λ_*) such that U_t is independent of H_t and $(\alpha_*, \xi_*, \sigma_*^2, \lambda_*) = (1, 1, 1, 1)$. Similar figures are obtained for $\mathbf{\Gamma} = [-0.2, 1.5]$, $\mathbf{\Gamma} = [-0.2, 2.5]$, $\mathbf{\Gamma} = [0.0, 1.5]$, $\mathbf{\Gamma} = [0.0, 2.5]$, $\mathbf{\Gamma} = [0.1, 1.5]$, and $\mathbf{\Gamma} = [0.1, 2.5]$.

$\mathbf{\Gamma}$	Levels \ n	50	100	200	300	400	500
$[-0.20, 1.50]$	1%	1.08	1.28	1.01	0.95	1.04	0.95
	5%	5.77	5.56	4.78	5.09	4.89	5.01
	10%	11.32	10.80	9.78	9.88	9.50	9.58
$[-0.10, 1.50]$	1%	1.25	0.97	1.11	1.11	0.95	1.07
	5%	5.59	5.22	4.84	5.21	5.08	4.75
	10%	10.98	9.97	9.77	10.39	9.98	9.50
$[0.00, 1.50]$	1%	1.11	1.20	1.05	1.05	1.14	1.12
	5%	5.96	5.23	5.42	5.00	5.46	5.11
	10%	11.43	10.72	10.83	10.27	10.40	10.39
$[0.10, 1.50]$	1%	1.37	1.18	1.15	1.30	0.92	1.16
	5%	6.38	5.92	5.12	5.58	5.09	5.15
	10%	11.97	11.41	10.60	10.53	10.38	10.25
$[-0.20, 2.50]$	1%	1.04	1.11	0.84	1.20	1.23	0.91
	5%	5.72	5.17	4.71	4.84	4.92	4.86
	10%	10.76	10.63	9.56	9.64	9.95	9.72
$[-0.10, 2.50]$	1%	1.24	1.26	0.93	0.96	1.05	1.01
	5%	6.09	5.88	4.94	4.87	5.07	4.82
	10%	11.83	11.27	9.83	9.81	9.98	9.72
$[0.00, 2.50]$	1%	1.37	1.15	1.25	1.03	1.19	1.02
	5%	6.17	5.55	5.17	5.19	5.38	4.95
	10%	11.56	11.04	10.48	10.47	10.45	10.28
$[0.10, 2.50]$	1%	1.25	1.07	1.08	1.17	1.12	1.09
	5%	5.87	5.53	5.62	5.40	4.98	5.30
	10%	11.68	10.97	10.94	10.42	9.75	10.22

Table 3: EMPIRICAL P-VALUES OF THE QLR STATISTICS OBTAINED BY THE WEIGHTED BOOTSTRAP. Number of Repetitions: 10,000. MODEL: $Y_t = \alpha + \xi X_t + \pi Z_t + \beta X_t^{\gamma} + U_t$. DGP: $Y_t = \alpha_* + \xi_* X_t + \pi_* Z_t + \cos(Z_t) U_t$, $X_t := \exp(-\lambda_* H_t)$, $(Z_t, U_t)' \sim$ IID $N(\mathbf{0}, \sigma_*^2 \mathbf{I}_2)$, $H_t \sim$ IID Exp(λ_*) such that U_t is independent of H_t and $(\alpha_*, \xi_*, \pi_*, \sigma_*^2, \lambda_*) = (1, 1, 1, 1, 1)$.

$\mathbf{\Gamma} = [-0.1, 1.5]$ $\qquad$ $\mathbf{\Gamma} = [-0.1, 2.5]$

Figure 3: EMPIRICAL P-P PLOTS OF THE QLR STATISTICS OBTAINED BY THE WEIGHTED BOOTSTRAP WITH $\mathbf{\Gamma} = [-0.1, 1.5]$ AND $\mathbf{\Gamma} = [-0.1, 2.5]$. Number of Repetitions: 10,000. MODEL: $Y_t = \alpha + \xi X_t + \pi Z_t + \beta X_t^{\gamma} + U_t$. DGP: $Y_t = \alpha_* + \xi_* X_t + \pi_* Z_t + \cos(Z_t) U_t$, $X_t := \exp(-\lambda_* H_t)$, $(Z_t, U_t)' \sim$ IID $N(\mathbf{0}, \sigma_*^2 \mathbf{I}_2)$, $H_t \sim$ IID Exp(λ_*) such that U_t is independent of H_t and $(\alpha_*, \xi_*, \pi_*, \sigma_*^2, \lambda_*) = (1, 1, 1, 1, 1)$. Similar figures are obtained for $\mathbf{\Gamma} = [-0.2, 1.5]$, $\mathbf{\Gamma} = [-0.2, 2.5]$, $\mathbf{\Gamma} = [0.0, 1.5]$, $\mathbf{\Gamma} = [0.0, 2.5]$, $\mathbf{\Gamma} = [0.1, 1.5]$, and $\mathbf{\Gamma} = [0.1, 2.5]$.

$\mathbf{\Gamma} = [-0.1, 1.5]$ $\qquad$ $\mathbf{\Gamma} = [-0.1, 2.5]$

Figure 4: ASYMPTOTIC AND EMPIRICAL NULL DISTRIBUTIONS OF THE QLR STATISTICS WITH $\mathbf{\Gamma} = [-0.1, 1.5]$ AND $\mathbf{\Gamma} = [-0.1, 2.5]$. Number of Repetitions: 10,000. MODEL: $Y_t = \alpha + \eta D_t + \xi t + \beta t^{\gamma} + U_t$. DGP: $Y_t = \alpha_* + \eta_* D_t + \xi_* t + U_t$, $D_t := \rho_* D_{t-1} + V_t$, and $(U_t, V_t)' \sim$ IID $N(\mathbf{0}, \sigma_*^2 \mathbf{I}_2)$ such that $(\alpha_*, \eta_*, \xi_*, \sigma_*^2, \rho_*) = (1, 1, 0, 1, 0.5)$. Similar figures are obtained for $\mathbf{\Gamma} = [-0.2, 1.5]$, $\mathbf{\Gamma} = [-0.2, 2.5]$, $\mathbf{\Gamma} = [0.0, 1.5]$, $\mathbf{\Gamma} = [0.0, 2.5]$, $\mathbf{\Gamma} = [0.1, 1.5]$, and $\mathbf{\Gamma} = [0.1, 2.5]$.

$\mathbf{\Gamma}$	Levels $\setminus$ n	50	100	200	300	400	500
$[-0.20, 1.50]$	1%	0.89	0.81	0.65	0.85	0.80	0.87*
	5%	4.31	4.18	4.08	4.32	4.61	4.37*
	10%	8.82	8.32	8.23	8.85	8.96	8.55*
$[-0.10, 1.50]$	1%	0.63	0.66	0.67	0.90	0.76	0.76
	5%	3.99	4.18	3.92	4.17	4.13	4.07
	10%	8.01	8.42	8.55	8.60	8.70	8.53
$[0.00, 1.50]$	1%	0.78	0.76	1.03	1.01	0.87	0.98
	5%	4.71	4.96	4.78	4.82	4.89	4.72
	10%	9.54	9.72	9.65	9.59	9.53	9.35
$[0.10, 1.50]$	1%	0.86	0.82	0.93	0.97	1.14	1.03
	5%	5.04	4.91	5.20	5.20	5.16	5.36
	10%	10.04	9.82	10.12	9.86	9.86	10.09
$[-0.20, 2.50]$	1%	0.98	0.72	0.81	0.70	0.97	0.79†
	5%	4.61	4.10	4.12	4.30	4.31	4.28†
	10%	9.13	8.82	8.46	8.56	8.44	8.53†
$[-0.10, 2.50]$	1%	0.96	0.81	1.02	0.98	0.77	0.88
	5%	4.53	4.34	4.75	4.70	4.27	4.80
	10%	9.29	9.03	9.32	8.87	8.79	9.48
$[0.00, 2.50]$	1%	1.03	1.18	1.17	1.15	0.92	1.10
	5%	5.34	5.93	5.56	5.85	5.29	5.37
	10%	10.93	11.73	11.02	11.03	10.46	10.99
$[0.10, 2.50]$	1%	0.90	1.09	1.04	0.92	0.97	1.05
	5%	5.06	5.13	4.68	4.55	5.02	4.99
	10%	10.23	10.39	9.46	9.50	9.94	9.82

Table 4: LEVELS OF THE QLR TEST STATISTICS. Number of Repetitions: 10,000. MODEL: $Y_t = \alpha + \eta D_t + \xi t + \beta t^\gamma + U_t$. DGP: $Y_t = \alpha_* + \eta_* D_t + \xi_* t + U_t$, $D_t := \rho_* D_{t-1} + V_t$, and $(U_t, V_t)' \sim$ IID $N(\mathbf{0}, \sigma_*^2 \mathbf{I}_2)$ such that $(\alpha_*, \eta_*, \xi_*, \sigma_*^2, \rho_*) = (1, 1, 0, 1, 0.5)$. Notes. $*$: for $n = 50,000$, the empirical rejection rates are 0.79, 4.11, and 8.53 when the levels of significance are 1%, 5%, and 10%, respectively; †: for $n = 50,000$, the empirical rejection rates are 0.91, 4.64, and 9.03 when the levels of significance are 1%, 5%, and 10%, respectively.

$\mathbf{\Gamma} = [-0.1, 1.5]$ $\qquad$ $\mathbf{\Gamma} = [-0.1, 2.5]$

Figure 5: ASYMPTOTIC AND EMPIRICAL NULL DISTRIBUTIONS OF THE QLR STATISTICS WITH $\mathbf{\Gamma} = [-0.1, 1.5]$ AND $\mathbf{\Gamma} = [-0.1, 2.5]$. Number of Repetitions: 10,000. MODEL: $Y_t = \alpha + \eta D_t + \xi t + \beta t^\gamma + U_t$. DGP: $Y_t = \alpha_* + \eta_* D_t + \xi_* t + \cos(D_t) U_t$, $D_t := \rho_* D_{t-1} + V_t$, and $(U_t, V_t)' \sim$ IID $N(\mathbf{0}, \sigma_*^2 \mathbf{I}_2)$ such that $(\alpha_*, \eta_*, \xi_*, \sigma_*^2, \rho_*) = (1, 1, 0, 1, 0.5)$. Similar figures are obtained for $\mathbf{\Gamma} = [-0.2, 1.5]$, $\mathbf{\Gamma} = [-0.2, 2.5]$, $\mathbf{\Gamma} = [0.0, 1.5]$, $\mathbf{\Gamma} = [0.0, 2.5]$, $\mathbf{\Gamma} = [0.1, 1.5]$, and $\mathbf{\Gamma} = [0.1, 2.5]$.

$\mathbf{\Gamma} = [-0.1, 1.5]$ $\qquad$ $\mathbf{\Gamma} = [-0.1, 2.5]$

Figure 6: EMPIRICAL P-P PLOTS OF THE QLR STATISTICS OBTAINED BY THE WEIGHTED BOOTSTRAP WITH $\mathbf{\Gamma} = [-0.1, 1.5]$ AND $\mathbf{\Gamma} = [-0.1, 2.5]$. Number of Repetitions: 10,000. MODEL: $Y_t = \alpha + \eta D_t + \xi t + \beta t^\gamma + U_t$. DGP: $Y_t = \alpha_* + \eta_* D_t + \xi_* t + \cos(D_t) U_t$, $D_t := \rho_* D_{t-1} + V_t$, and $(U_t, V_t)' \sim$ IID $N(\mathbf{0}, \sigma_*^2 \mathbf{I}_2)$ such that $(\alpha_*, \eta_*, \xi_*, \sigma_*^2, \rho_*) = (1, 1, 0, 1, 0.5)$. Similar figures are obtained for $\mathbf{\Gamma} = [-0.2, 1.5]$, $\mathbf{\Gamma} = [-0.2, 2.5]$, $\mathbf{\Gamma} = [0.0, 1.5]$, $\mathbf{\Gamma} = [0.0, 2.5]$, $\mathbf{\Gamma} = [0.1, 1.5]$, and $\mathbf{\Gamma} = [0.1, 2.5]$.

$\mathbf{\Gamma}$	Levels $\setminus$ n	50	100	200	300	400	500
$[-0.20, 1.50]$	1%	1.11	0.83	0.95	0.90	0.92	0.81^*
	5%	4.95	4.58	4.83	4.40	4.49	4.37^*
	10%	9.51	8.69	9.31	8.40	8.71	8.62^*
$[-0.10, 1.50]$	1%	1.04	0.88	0.74	1.02	0.92	0.77
	5%	5.12	4.66	4.34	4.76	4.56	4.50
	10%	9.64	9.35	8.73	9.22	9.48	8.84
$[0.00, 1.50]$	1%	1.10	1.10	1.10	0.98	1.14	1.10
	5%	5.42	5.11	4.88	4.91	5.30	4.82
	10%	10.43	10.01	9.55	9.69	10.14	10.05
$[0.10, 1.50]$	1%	1.21	1.10	1.17	1.21	1.10	1.13
	5%	5.13	4.95	5.25	5.30	5.07	5.52
	10%	10.94	10.05	10.37	10.50	9.98	10.55
$[-0.20, 2.50]$	1%	0.95	0.76	0.73	0.76	0.83	$0.80^\dagger$
	5%	4.68	4.35	3.78	4.24	4.23	$4.27^\dagger$
	10%	9.16	8.85	8.48	8.58	8.27	$8.70^\dagger$
$[-0.10, 2.50]$	1%	1.12	0.90	0.85	1.03	0.77	0.99
	5%	5.02	4.65	4.50	4.50	4.61	4.55
	10%	10.32	9.47	9.41	9.16	8.86	9.11
$[0.00, 2.50]$	1%	0.87	0.95	0.88	0.91	0.86	1.01
	5%	5.13	4.84	4.87	4.96	4.75	5.17
	10%	10.30	10.07	10.06	10.21	9.80	10.23
$[0.10, 2.50]$	1%	1.11	0.99	1.09	1.09	1.03	1.09
	5%	5.24	5.06	5.33	5.33	5.19	5.57
	10%	10.63	9.86	10.05	10.17	9.73	10.32

Table 5: LEVELS OF THE QLR TEST STATISTICS. Number of Repetitions: 10,000. MODEL: $Y_t = \alpha + \eta D_t + \xi t + \beta t^\gamma + U_t$. DGP: $Y_t = \alpha_* + \eta_* D_t + \xi_* t + \cos(D_t)U_t$, $D_t := \rho_* D_{t-1} + V_t$, and $(U_t, V_t)' \sim$ IID $N(\mathbf{0}, \sigma_*^2 \mathbf{I}_2)$ such that $(\alpha_*, \eta_*, \xi_*, \sigma_*^2, \rho_*) = (1, 1, 0, 1, 0.5)$. Notes. $*$: for $n = 50,000$, the empirical rejection rates are 0.79%, 4.58%, and 8.72% when the levels of significance are 1%, 5%, and 10%, respectively; †: for $n = 50,000$, the empirical rejection rates are 0.71%, 4.25%, and 8.69% when the levels of significance are 1%, 5%, and 10%, respectively.

$\mathbf{\Gamma}$	Levels $\backslash$ n	50	100	200	300	400	500
	1%	1.40	1.03	1.17	0.99	0.92	0.92
$[-0.20, 1.50]$	5%	6.00	5.43	5.42	4.86	4.78	4.71
	10%	11.98	10.46	10.45	9.87	9.59	9.52
	1%	1.30	1.01	0.96	1.14	1.07	1.01
$[-0.10, 1.50]$	5%	6.35	5.66	4.94	5.45	5.19	4.91
	10%	11.71	10.85	9.80	10.16	10.35	9.68
	1%	1.34	1.22	1.26	1.05	1.29	1.14
$[0.00, 1.50]$	5%	6.24	5.65	5.26	5.18	5.55	4.94
	10%	11.83	11.09	10.10	10.28	10.65	10.27
	1%	1.25	1.05	1.20	1.23	1.10	1.04
$[0.10, 1.50]$	5%	6.00	5.47	5.39	5.60	5.21	5.68
	10%	11.88	10.62	10.49	10.88	10.23	10.60
	1%	1.31	1.03	1.00	0.97	0.98	0.97
$[-0.20, 2.50]$	5%	6.03	5.34	4.67	4.90	4.76	4.87
	10%	11.43	10.62	9.58	9.84	9.40	9.72
	1%	1.32	1.08	0.92	1.12	0.81	1.02
$[-0.10, 2.50]$	5%	6.06	5.35	5.00	4.87	4.87	4.92
	10%	12.12	10.72	10.13	9.91	9.47	9.60
	1%	1.04	1.15	1.01	1.00	0.97	1.13
$[0.00, 2.50]$	5%	6.22	5.49	5.08	5.39	4.99	5.46
	10%	11.79	11.16	10.84	10.68	10.21	10.80
	1%	1.38	1.24	1.04	1.01	1.03	1.03
$[0.10, 2.50]$	5%	6.32	5.52	5.19	4.99	5.01	4.99
	10%	11.93	10.98	10.22	10.02	10.18	10.23

Table 6: EMPIRICAL P-VALUES OF THE QLR STATISTICS OBTAINED BY THE WEIGHTED BOOTSTRAP. Number of Repetitions: 10,000. MODEL: $Y_t = \alpha + \eta D_t + \xi t + \beta t^{\gamma} + U_t$. DGP: $Y_t = \alpha_* + \eta_* D_t + \xi_* t + \cos(D_t) U_t$, $D_t := \rho_* D_{t-1} + V_t$, and $(U_t, V_t)' \sim$ IID $N(\mathbf{0}, \sigma_*^2 \mathbf{I}_2)$ such that $(\alpha_*, \eta_*, \xi_*, \sigma_*^2, \rho_*) = (1, 1, 0, 1, 0.5)$.