Comprehensively Testing Linearity Hypothesis Using the Smooth Transition Autoregressive Model Applied to Macroeconomic Variables

DAKYUNG SEONG

School of Economics, University of Sydney, Camperdown NSW 2006, Australia

Email: sdkseong@gmail.com

JIN SEO CHO

School of Humanities and Social Sciences, Beijing Institute of Technology, Haidian, Beijing 100081, China
School of Economics, Yonsei University, 50 Yonsei-ro, Seodaemun, Seoul, Korea

Email: jinseocho@yonsei.ac.kr

TIMO TERÄSVIRTA

CREATES, Department of Economics and Business Economics, Aarhus University,
C.A.S.E., Humbold-Universität zu Berlin

Email: tterasvirta@econ.au.dk

This version: October 2020

Abstract

This paper examines the null limit distribution of the quasi-likelihood ratio (QLR) statistic that tests linearity condition using the smooth transition autoregressive (STAR) model. We explicitly show that the QLR test statistic weakly converges to a functional of a Gaussian process under the null of linearity by resolving the issue of twofold identification meaning that Davies's (1977, 1987) identification problem arises in two different ways under the null. We illustrate our theory using the exponential STAR and logistic STAR models and also conduct Monte Carlo simulations. Finally, we test for neglected nonlinearity in the multiplier effect of US government spending, and growth rates of US unemployment. These empirical examples also demonstrate that the QLR test statistic is useful for detecting the nonlinear structure among economic variables and complements the linearity test of the Lagrange multiplier test statistic in Teräsvirta (1994).

Key Words: QLR test statistic, STAR model, linearity test, Gaussian process, null limit distribution, nonstandard testing problem.

Subject Classification: C12, C18, C46, C52.

Acknowledgements: Part of the work for this paper was done when the third author was visiting the School of Economics and Finance, Queensland University of Technology, whose kind hospitality is gratefully acknowledged. The authors have benefitted from discussions with Otilia Boldea, Tae-Hwan Kim and Byungsam Yoo. Responsibility for any errors and shortcomings in this work remains ours.

1 Introduction

The smooth transition autoregressive (STAR) model is popularly used in empirical data analysis. For example, Teräsvirta (1994) examines the nonlinearity structure of the German industrial production. As another example, Auerbach and Gorodnichenko (2012b, AG hereafter) estimate the vector smooth transition autoregressive model to identify the nonlinear fiscal multiplier effect.

It is also popular to test linearity of time series against the STAR model as a first step of building nonlinear STAR models. Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994) and Granger and Teräsvirta (1993) among others suggested to test linearity using the Lagrange multiplier (LM) test statistic based upon the STAR model, and their LM test statistics are popularly used for empirical applications.

Nevertheless, the LM test statistic does not comprehensively test for the nonlinearity entailed by the STAR model. As we detail below, the STAR model violates the linearity condition in two different ways, and the LM statistic tests for only one of the two violations. In other words, the STAR model is not identified when the null hypothesis of linearity holds.

The main goal of the current study is to develop a testing procedure that tests for the alternative nonlinearity in two different ways and combines the testing results into a single test statistic, thereby demonstrating its proper empirical applications along with our theory on the test statistic.

The identification problem considered here was first studied by Davies (1977, 1987). The problem arises when the model to be tested, in our case a linear model, can be defined by more than one set of parameter restrictions on the alternative. When one such set of restrictions is selected, some of the parameters of the alternative model are not identified when the null hypothesis holds, that is, the data provide no information about their values. This being the case, it is also possible, however, to choose a different set of restrictions to define the null model using some or all of the parameters that were unidentified in the previous case. This implies that a different set of parameters, including the one or ones that defined the previous null model, are now unidentified when this null hypothesis holds. The main consequence of this lack of identification is that the parameters of the alternative model cannot be estimated consistently when the null hypothesis is valid, and the standard asymptotic inference becomes invalid. This problem appears when linearity is tested against nonlinear models such as smooth transition models, threshold autoregressive or switching regression models, hidden Markov models or artificial neural network models, to name just a few.

This goal is achieved by generalising the results in the previous literature to the STAR case. This literature examines testing linearity using the artificial neural network (ANN) framework. Cho, Ishida, and White (2011, 2014), Cho and Ishida (2012), White and Cho (2012), and Baek, Cho, and Phillips (2015), among others, study testing for neglected nonlinearity using analytic functions and note that the null of linearity can arise in two or three different ways, each of which carries its own identification problem. They call this feature the twofold or trifold identification problem and propose a quasi-likelihood ratio (QLR) test statistic to resolve the issues associated with the identification problem. We transform their approach to the STAR framework and develop a testing procedure that is readily available for applications. In essence, two views exist to look at the QLR test statistics in terms of the linear model hypothesis, and the methodologies supporting the two views produce two different null weak limits. We below examine how the weak limits are interrelated to the null limit distribution of the QLR test statistic testing the linearity hypothesis.

The second goal of this study is to shed light on the empirical usefulness of the QLR test statistic. In order to do this, we examine popular nonlinearity assumptions imposed in applied macroeconomic literature and revisit two published empirical studies. We first re-examine the macroeconomic data in AG who examine the nonlinearity structure of government multiplier effect. Using the vector smooth-transition autoregression (VSTAR) model, they relax the linear model assumption among the government spending, tax revenue, and gross domestic product posited as by Blanchard and Perotti's (2002) vector autoregression (VAR) model and, using a nonlinear model, estimate the response of certain economic variables to a \$1 increase of government spending. Reconsidering this model is motivated by the debate in the literature, in which the new Keynesian empirical models generally agree on the stable fiscal multiplier effect (*e.g.*, Ramey, 2011a; Coenen et al., 2012; Ramey and Zubairy, 2018; Christiano, Eichenbaum, and Rebelo, 2011), whereas traditional Keynesian models argues for state-dependent multiplier effect. The latter approach leads to characterizing the nonlinear interrelationship by a Vector STAR (VSTAR) model. In particular, the strong multiplier effect estimated by AG has been doubted in the new Keynesian literature. This contradiction makes it interesting to see what using the QLR test can contribute to this debate.

We apply our QLR test statistic to the AG dataset to detect the associated nonlinearity. For this purpose, we reformulate the VAR and VSTAR models into vector error correction (VEC) and vector smooth-transition error correction (VSTEC) models, respectively and apply the QLR test statistic, so that we can exploit the stationarity assumption of data needed for the application of the QLR test statistic and the long-run relationship among the variables can be properly detected. As it turns out, the QLR test statistic favors nonlinearity against linearity, and further, despite their differences, the VSTEC model estimates the impulse response functions similar to those in AG, but our estimated multiplier effects are not so strong as what AG obtained. Through this empirical example, we demonstrate the usefulness of the QLR test statistic.

We use a univariate economic time series to discuss the usefulness of the QLR test statistic relative to the LM test statistic frequently applied in the literature. We provide evidence that the QLR and LM test statistics are complementary to each other. We extend the quarterly US unemployment rate series that has been previously studied by van Dijk, Teräsvirta, and Franses (2002). They tested linearity by the LM statistic, and in this study we illustrate the use of the QLR test statistic alongside the LM statistic and find nonlinear features in the series that could not have been found by the LM or the QLR statistic alone.

The plan of this paper is as follows. In Section 2, we derive the null limit distribution of the QLR test statistic by resolving the twofold identification problem. We do this by generalizing the approach developed for the artificial neural network model. In Section 3, we apply our theory to the ESTAR and LSTAR models and demonstrate its relevance. In this section we also report results on Monte Carlo simulations. In particular, we demonstrate how to apply Hansen's (1996) weighted bootstrap to the QLR test statistic. Section 4 contains applications of the QLR test statistic to the multiplier effect of US government spending and the US unemployment rate. The QLR test statistic is exploited to detect the nonlinearity structure among economic variables, and the performances of the QLR and LM statistics are compared with each other. The detailed proofs of our claims can be found in the Appendix.

Before proceeding, we provide some notation. A function mapping $f : \mathcal{X} \mapsto \mathcal{Y}$ is denoted by $f(\cdot)$, evaluated derivatives such as $f'(x)|_{x=x_*}$ are written simply as $f'(x_*)$. We also let "$a_n \Rightarrow a$""$a_n \stackrel{\text{a.s.}}{\to} a$" indicate "$a_n$ weakly converges to a" and "a_n almost surely converges to a," respectively. The latter is occasionally denoted as $\lim_{n\to\infty} a_n \stackrel{\text{a.s.}}{=} a$.

2 Testing Linearity Using the STAR Model

2.1 Preliminaries

As already mentioned, the main goal of the current study is to develop a testing procedure for testing linearity against a nonlinear alternative in two different ways such that one ends up with a single (QLR) test statistic. This goal is achieved by applying the results in the previous literature to the STAR model. This literature examines testing linearity using the ANN framework. Cho, Ishida, and White (2011, 2014), Cho and Ishida (2012), White and Cho (2012), and Baek, Cho, and Phillips (2015), among others, study testing for neglected nonlinearity using analytic functions and note that the null of linearity can arise in two or three different ways in their model framework as it does in the STAR model. They call this feature the twofold or trifold identification problem and propose a quasi-likelihood ratio (QLR) test statistic to resolve the issues arising from this ambiguity.

The standard single-hidden layer (univariate) ANN model has the following form

$$y_t = \pi_0 + \widetilde{z}_t'\pi + \sum_{j=1}^{q} \theta_j f(z_t'\gamma_j) + \varepsilon_t \tag{1}$$

where $z_t := (1, \widetilde{z}_t')'$ with $\widetilde{z}_t := (y_{t-1}, ..., y_{t-p})'$, $f(0) =$ constant, and $\pi_0, \pi, \theta_j, \gamma_j$, $j = 1, ..., q$, are parameters. In many applications, $\pi = 0$. The ANN model thus contains a linear combination of continuous and bounded functions (a hidden layer), typically logistic functions, although other bounded functions are possible. Nowadays, ANN models are often more complicated than this one containing more than one hidden layer, but (1) serves as a benchmark against which a STAR model may be compared. The twofold identification problem appears here: the ANN model becomes linear by assuming either $\theta_j = 0$ or $\gamma_j = 0$, $j = 1, ..., q$.

A standard STAR model, see for example, Teräsvirta (1994), Granger and Teräsvirta (1993), and van Dijk, Teräsvirta, and Franses (2002), can be written as follows:

$$y_t = z_t'\pi + z_t'\theta f(z_t'\alpha; \gamma) + \varepsilon_t$$

where $f(z_t'\alpha;\gamma)$ is another analytic bounded function, $f(z_t'\alpha;0) = 0$, $\gamma > 0$ is a scalar parameter, and $\alpha = (0,...,0,1,0,...,0)'$ is a known vector. It becomes linear by either setting $\theta = 0$ or $\gamma = 0$, so the twofold identification problem is present. The main difference between these two models is that the single hidden-layer ANN model contains a linear combination of several transitions that are themselves functions of linear combinations of elements of z_t, whereas in the standard STAR model a linear combination of these elements is multiplied by a single nonlinear function.[1] This transition function is typically driven by only one variable which is most often an element of $\widetilde{z}_t$.

Due to these differences, the analysis of the QLR test statistic needs to be generalized in order to make the QLR test statistic applicable in the STAR framework. As an example, Cho, Ishida, and White (2011, 2014) characterize the null limit distribution of the QLR test statistic as a functional of a univariate Gaussian process. This limit distribution cannot, however, be simply applied to STAR models, because, as it turns out, for this purpose a multidimensional Gaussian process is called for. This means that we have to generalize the approach based upon the ANN model to fit the complexity of the STAR model.

2.2 Motivation of Testing Linearity Using the STAR Model

The following STAR model of order p is popularly specified as a prediction model of a time-series data y_t (*e.g.*, Teräsvirta, 1994; Granger and Teräsvirta, 1993):

$$\mathcal{M}_0 := \{h_0(\,\cdot\,;\pi,\theta,\gamma) : (\pi,\theta,\gamma) \in \Pi \times \Theta \times \Gamma\},$$

where $h_0(z_t;\pi,\theta,\gamma) := z_t'\pi + f(\widetilde{z}_t'\alpha,\gamma)(z_t'\theta)$, $z_t := (1,\widetilde{z}_t')'$ is a $(p+1)\times 1$ vector of regressors with a transition variable $\widetilde{z}_t'\alpha$. Here, $\widetilde{z}_t := (y_{t-1}, y_{t-2}, \ldots, y_{t-p})'$, and $\alpha = (0,\ldots,1,0,\ldots,0)'$ denotes a selection vector chosen by the researcher. The other parameter vectors $\pi := (\pi_0,\pi_1,\ldots,\pi_p)'$ and $\theta := (\theta_0,\theta_1,\ldots,\theta_p)'$ are the mean transition parameters, and γ is used to describe the smooth transition from one extreme regime to the other. Symbols Π, Θ, and Γ denote the parameter spaces of π, θ, and γ, respectively. The transition function $f(\cdot,\gamma)$ is a nonlinear, continuously differentiable, and uniformly bounded function. Here, we observe that the empirical researcher often flexibly modifies $\mathcal{M}_0$ by removing the constant from z_t or adding other exogenous variables to z_t. We also note that the STAR model is a special case of the original STAR model in the literature, in which $f(\widetilde{z}_t'\alpha - c,\gamma)$ is specified for $\mathcal{M}_0$ instead of $f(\widetilde{z}_t'\alpha,\gamma)$ by employing the additional parameter c. We set $c = 0$ in $\mathcal{M}_0$ as in the regular exponential autoregressive model in Haggan and Ozaki (1981) because the essential property in testing the linearity is that $f(\widetilde{z}_t'\alpha,\cdot)$ is an analytic function. As we detail below, if c is estimated along with the other parameters π and θ, the inferential procedure becomes more complicated than the one of the current study, and this limits its applicability due to its complexity.

The most popular STAR models are the exponential smooth transition autoregressive (ESTAR) and logistic smooth transition autoregressive (LSTAR) models. They are characterized by the exponential and logistic cumulative distri-

[1] STAR models can also contain more than one additive transition, but this seems to be uncommon in applications.

bution functions, respectively, and each of them displays different nonlinear patterns:

$$f_E(\widetilde{z}_t'\alpha, \gamma) := 1 - \exp(-\gamma(\widetilde{z}_t'\alpha)^2) \quad \text{and} \quad f_L(\widetilde{z}_t'\alpha, \gamma) := \{1 + \exp(-\gamma\widetilde{z}_t'\alpha)\}^{-1},$$

where $\gamma > 0$ are the nonlinear functional forms exhibited by the ESTAR and LSTAR models, respectively. It is seen from these expressions that the STAR model has a continuum of regimes defined by transition functions obtaining values between 0 to 1. This aspect makes the model appealing for empirical analysis because the presence of multiple regimes are often structural and attributed to the behaviour of economic agents. For more discussion on the STAR model the reader is referred to van Dijk, Teräsvirta, and Franses (2002), Teräsvirta (1994), Granger and Teräsvirta (1993), and Teräsvirta, Tjøstheim, and Granger (2010), among others.

This study focuses on testing linearity against STAR. The ESTAR and LSTAR models are specified by transforming the exponential function that is analytic, so that it is generically comprehensively revealing for model misspecification as pointed out by Stinchcombe and White (1998). Therefore, the estimated parameters in the transition function become statistically significant such that the nonlinear component necessarily reduces the mean squared error of the model, even when the assumed STAR model is misspecified. This fact implies that if the linear model is misspecified, the mean square error obtained from the STAR models becomes smaller than that from the linear model, motivating testing linearity hypothesis by comparing the estimated mean squared errors from the STAR and the linear model nested in the STAR. The QLR test statistic is often motivated this way. This process delivers an omnibus testing procedure for nonlinearity.

Similar arguments can be found in the previous literature. First, Cho, Ishida, and White (2011, 2014) and Baek, Cho, and Phillips (2015) examine testing linearity using both analytic functions and power transformations. They test linearity using the QLR test statistic and demonstrate usefulness of the test by Monte Carlo experiments. We take advantage of this literature and apply the QLR statistic to testing linearity against STAR. We note, however, that in the previous literature the QLR statistic is applied to testing linearity against artificial neural network models. In the STAR case, the nonlinear functions are different from what they are when the alternative is an artificial neural network. Because of this, the QLR test statistic against STAR exhibits power patterns different from those in Cho, Ishida, and White (2011, 2014) and Baek, Cho, and Phillips (2015). Deriving the null limit distribution of the QLR test based against STAR leads to generalizing the corresponding derivations in Cho, Ishida, and White (2011, 2014) and Baek, Cho, and Phillips (2015).

Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994), Granger and Teräsvirta (1993), and Teräsvirta, Tjøstheim, and Granger (2010), among others, examine the LM statistic of testing linearity against STAR. As we discuss below, the LM test is defined to test one of the two hypotheses that characterize the linearity condition using the STAR model, whereas the QLR test statistic is defined to handle the two hypotheses at the same time. This aspect of the QLR test statistic extends the testing scope aimed by the LM test statistic, and below we illustrate how the QLR and LM test statistics can complement each other using empirical examples.

2.3 DGP and QLR Test Statistic

In order to proceed, we make the following assumptions:

Assumption 1. $\{(y_t, \widetilde{z}_t')' \in \mathbb{R}^{1+p} : t = 1, 2, \ldots\}$ $(p \in \mathbb{N})$ *is a strictly stationary and absolutely regular process defined on the complete probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, with* $\mathbb{E}[|y_t|] < \infty$ *and mixing coefficient* β_τ *such that for some* $\rho > 1$*,* $\sum_{\tau=1}^{\infty} \tau^{1/(\rho-1)} \beta_\tau < \infty$. □

Here, the mixing coefficient is defined as $\beta_\tau := \sup_{s \in \mathbb{N}} \mathbb{E}[\sup_{A \in \mathcal{F}_{s+\tau}^{\infty}} |\mathbb{P}(A|\mathcal{F}_{-\infty}^{s}) - \mathbb{P}(A)|]$, where $\mathcal{F}_\tau^s$ is the σ-field generated by $(y_t, \ldots, y_{t+s})$. Many time series models satisfy this condition, and the autoregressive process is one of them. It is general enough to cover the stationary time series we are interested in.

We impose the following regular STAR model condition:

Assumption 2. *Let* $f(\widetilde{z}_t'\alpha, \cdot) : \Gamma \mapsto [0, 1]$ *be a non-polynomial analytic function with probability 1. Let* $\Pi \in \mathbb{R}^{p+1}$*,* $\Theta \in \mathbb{R}^{p+1}$*, and* $\Gamma \in \mathbb{R}$ *be non-empty convex and compact sets such that* $0 \in \Gamma$*. Let* $h(z_t; \pi, \theta, \gamma) := z_t'\pi + \{f(\widetilde{z}_t'\alpha, \gamma) - f(\widetilde{z}_t'\alpha, 0)\}(z_t'\theta)$*, and let* $\mathcal{M} := \{h(\,\cdot\,; \pi, \theta, \gamma) : (\pi, \theta, \gamma) \in \Pi \times \Theta \times \Gamma\}$ *be the model specified for* $\mathbb{E}[y_t|z_t]$. □

Note that $\mathcal{M}$ differs from $\mathcal{M}_0$. The transition function is centered at $f(\widetilde{z}_t'\alpha, 0)$ for analytical convenience. As $f(\widetilde{z}_t'\alpha, 0)$ is constant, the nonlinearity feature of the STAR model is not modified by the centering. For example, we have $f_E(\widetilde{z}_t'\alpha, 0) = 0$ and $f_L(\widetilde{z}_t'\alpha, 0) = 1/2$, and so it will be centered to have value zero. Furthermore, the centering further reduces the dimension of the identification problem as detailed below.

As already mentioned, the STAR model is different from the artificial neural network model. This means that the null limit distribution of the QLR test be derived under regularity conditions that are different from those needed for the ANN model. The parameters to be estimated are π, θ, and γ, as α is defined by the researcher.

Using Assumption 2, the linearity hypothesis and the alternative are specified as follows:

$$\mathcal{H}_0 : \exists \pi \in \mathbb{R}^{p+1} \text{ such that } \mathbb{P}(\mathbb{E}[y_t|z_t] = z_t'\pi) = 1; \text{ vs. } \mathcal{H}_1 : \forall \pi \in \mathbb{R}^{p+1}, \mathbb{P}(\mathbb{E}[y_t|z_t] = z_t'\pi) < 1.$$

These hypotheses are the same as the ones in Cho, Ishida, and White (2011, 2014) and Baek, Cho, and Phillips (2015). As in the previous literature, the focus is on developing an omnibus test statistic, but now against STAR, and we use the QLR test statistic as a vehicle for reaching this goal. The QLR test statistic is formally defined as

$$QLR_n := n\left(1 - \frac{\widehat{\sigma}_{n,A}^2}{\widehat{\sigma}_{n,0}^2}\right),$$

where

$$\widehat{\sigma}_{n,0}^2 := \min_{\pi} \frac{1}{n}\sum_{t=1}^{n}(y_t - z_t'\pi)^2, \qquad \widehat{\sigma}_{n,A}^2 := \min_{\pi,\theta,\gamma} \frac{1}{n}\sum_{t=1}^{n}\{y_t - z_t'\pi - f_t(\gamma)(z_t'\theta)\}^2,$$

and $f_t(\gamma) := f(\widetilde{z}_t'\alpha, \gamma) - f(\widetilde{z}_t'\alpha, 0)$. We also let the nonlinear least squares (NLS) estimator $(\widehat{\pi}_n, \widehat{\theta}_n, \widehat{\gamma}_n)$ be obtained by minimizing the squared errors with respect to (π, θ, γ), and $(\pi_*, \theta_*, \gamma_*)$ denotes the probability limits of the NLS estimator. That is, $(\pi_*, \theta_*, \gamma_*) := \arg\min_{\pi,\theta,\gamma} \mathbb{E}[\{y_t - z_t'\pi - f_t(\gamma)(z_t'\theta)\}^2]$ is the pseudo-true parameter.

The main reason to proceed with the QLR statistic is that the linearity hypothesis contains a twofold identification problem, and this statistic is able to handle both parts of it. To be more specific, the twofold identification problem can be described by the parameter set $(\pi_*, \theta_*, \gamma_*)$. If $\mathbb{E}[y_t|z_t]$ is linear with respect to z_t with coefficient π_*, we can generate a linear function from $h(\cdot; \pi_*, \theta_*, \gamma_*)$ in two different ways by letting $\theta_* = 0$ or $\gamma_* = 0$. Because of this, $(\pi_*, \theta_*, \gamma_*)$ is not uniquely determined, and the identification problem can be classified into cases depending on the value of θ_* and γ_*. If $\theta_* = 0$, $h(\cdot; \pi_*, 0, \gamma_*) = z_t'\pi_*$, so that γ_* is not identified. We call this problem type I identification problem, under which $(\pi_*, \theta_*, \gamma_*)$ becomes any element in $\{(\pi, \theta, \gamma) \in \Pi \times \Theta \times \Gamma : \pi = \pi_*, \theta = 0\}$. If we had employed $f(\widetilde{z}_t'\alpha - c, \gamma)$ instead of $f(\widetilde{z}_t'\alpha, \gamma)$ for $\mathcal{M}$ according to the original STAR model in the literature, both γ_* and c_* are not identified under $\theta_* = 0$, leading to a more complicated identification problem. We fix our interest to the current derivative model $\mathcal{M}$. Alternatively, if $\gamma_* = 0$, $h(\cdot; \pi_*, \theta_*, 0) = z_t'\pi_*$, so that θ_* is not identified, leading to another identification problem. We call it type II identification problem, under which $(\pi_*, \theta_*, \gamma_*)$ becomes any element in $\{(\pi, \theta, \gamma) \in \Pi \times \Theta \times \Gamma : \pi = \pi_*, \gamma = 0\}$. On the other hand, if the transition function is not centered at $f(\widetilde{z}_t'\alpha, 0)$, letting $\gamma_* = 0$ leads to $h_0(z_t; \pi_*, \theta_*, 0) = z_t'(\pi_* + f(\widetilde{z}_t'\alpha, 0)\theta_*)$. This implies that the type II identification problem becomes more complicated as π_* and θ_* are not separately identified. Here, the centering process is a device to make this complication a relatively simple identification problem. In addition to this, the null limit distribution is not modified by this centering, mainly due to the invariance principle. Note that π in $\mathcal{M}_0$ is reparameterized to $\pi - f(\widetilde{z}_t'\alpha, 0)\theta$ in $\mathcal{M}$, so that the QLR test statistic obtained by this reparameterization becomes identical to that before the reparameterization. Without this reparameterization, the null model investigation has to be separately conducted by discerning the parameters with type II identification problem as in Cho and Phillips (2018). So, we avoid the involved complication by the centering and obtain the null limit distribution of the QLR test statistic efficiently. We also observe that this centering process is often applied in the literature to derive the null limit distribution of the LM test statistic. For example, Teräsvirta (1994) shows that if z_t contains the constant, the null limit distribution of the LM test statistic is not affected by this centering because the centering parameter is merged with other linear components while applying Taylor expansions.

As described above, the null holds for the following two sub-hypotheses: $\mathcal{H}_{01} : \theta_* = 0$ and $\mathcal{H}_{02} : \gamma_* = 0$. The limit distribution of the QLR test statistic can be derived both under $\mathcal{H}_{01}$ and $\mathcal{H}_{02}$, leading to different null limit distributions even for the same QLR test statistic. We call these derivations type I and type II analysis, respectively and show below that they yield two different null weak limits under both hypotheses. The null hypothesis of linearity against STAR has to be properly tested by tackling both $\mathcal{H}_{01}$ and $\mathcal{H}_{02}$ simultaneously, and we shall demonstrate that the QLR test statistic has the capability of doing so. For this purpose, we derive its null limit distribution from the two null weak limits obtained under $\mathcal{H}_{01}$ and $\mathcal{H}_{02}$ in the vein of the approaches in Cho and White (2007), Cho, Ishida, and White (2011, 2014), Cho and Ishida (2012), Baek, Cho, and Phillips (2015), and Cho, and Phillips (2015). In essence, two views exist to look at the QLR test statistics in terms of the linear model hypothesis, and the methodologies supporting the two views produce two different null weak limits. We will examine how the weak limits are interrelated to the null limit distribution of the QLR test statistic testing the linearity hypothesis.

Our view to test the linearity hypothesis by accommodating types I and II analyses is different from other tests in prior literature. For example, the aforementioned LM test statistic does not accommodate the twofold identification problem. Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994), and Granger and Teräsvirta (1993) study the null limit distribution of the LM statistic for testing linearity using $\mathcal{H}_{02}$. The main argument for the LM test is that its asymptotic null distribution is chi-squared, which makes the test easily applicable. As another example, Cheng (2015) assumes the model in (1) by letting $\pi_0 = 0$ and $\pi = 0$ to develop a Wald statistic testing whether some or all of θ_{j*}'s are equal to zero or not in the vein of type I analysis.

2.4 The Null Limit Distribution of the QLR Test

We now derive the null limit distribution of the QLR test following the approach in Cho, Ishida, and White (2011, 2014) and Baek, Cho, and Phillips (2015) and highlight the difference between the STAR-based approach and the ANN-based one.

We first study the limit distributions of the QLR test under $\mathcal{H}_{01}$ and $\mathcal{H}_{02}$ separately, combine them and, finally, obtain the limit distribution under $\mathcal{H}_0$. For this purpose, we let our objective function or quasi-likelihood (QL) function be

$$\mathcal{L}_n(\pi, \theta, \gamma) := -\sum_{t=1}^{n} \{y_t - z_t'\pi - f_t(\gamma)(z_t'\theta)\}^2.$$

2.4.1 Type I Analysis: Testing $\mathcal{H}_{01} : \theta_* = 0$

In this subsection, we discuss the limit distribution of the QLR test under $\mathcal{H}_{01} : \theta_* = 0$. The problem is that γ_* is not identified under this hypothesis. We obtain the NLS estimator by maximizing the QL function with respect to γ in the final stage for the purpose of testing $\mathcal{H}_{01}$:

$$\mathcal{L}_n^{(1)} := \max_{\gamma} \max_{\theta} \max_{\pi} -\sum_{t=1}^{n} \{y_t - z_t'\pi - f_t(\gamma)(z_t'\theta)\}^2$$

and let $QLR_n^{(1)}$ be the QLR statistic obtained by this optimization process. That is,

$$\mathcal{L}_n^{(1)} := \max_{\gamma \in \Gamma} \{-u'Mu + u'MF(\gamma)Z[Z'F(\gamma)MF(\gamma)Z]^{-1}Z'F(\gamma)Mu\},$$

where $u := [u_1, u_2, \ldots, u_n]'$, $u_t := y_t - \mathbb{E}[y_t|z_t]$, $Z := [Z_1, Z_2, \ldots, Z_n]'$, $M := I - Z(Z'Z)^{-1}Z'$, $F(\gamma) := \mathrm{diag}[f_1(\gamma), f_2(\gamma), \ldots, f_n(\gamma)]$, and

$$QLR_n^{(1)} := \max_{\gamma \in \Gamma} \frac{1}{\widehat{\sigma}_{n,0}^2} u'MF(\gamma)Z[Z'F(\gamma)MF(\gamma)Z]^{-1}Z'F(\gamma)Mu \quad (2)$$

under $\mathcal{H}_{01}$ using the fact that $y_t = \mathbb{E}[y_t|z_t] + u_t = z_t'\pi_* + u_t$. Note that the numerator on the right-hand side of (2) is identical to $n(\widehat{\sigma}_{n,0}^2 - \widehat{\sigma}_{n,A}^2)$ under $\mathcal{H}_{01} : \theta_* = 0$, so that the definition of the QLR test statistic accords with $QLR_n^{(1)}$. Furthermore, we cannot let $\gamma = 0$ when the objective function in (2) is derived. Note that if $\gamma = 0$, the alternative model reduces to the linear model, so that the QLR test statistic cannot test the null model by letting $\gamma = 0$. We will therefore examine the null limit distribution by supposing $\gamma \neq 0$.

We now derive the limit distribution of $QLR_n^{(1)}$ under $\mathcal{H}_{01}$ similarly to Cho, Ishida, and White (2011, 2014) and Baek, Cho, and Phillips (2015). For this purpose and to guarantee regular behaviour of the null limit distribution, we impose the following conditions:

Assumption 3. *(i)* $\mathbb{E}[u_t|z_t, u_{t-1}, z_{t-1}, \ldots] = 0$*; and (ii)* $\mathbb{E}[u_t^2|z_t, u_{t-1}, z_{t-1}, \ldots] = \sigma_*^2$. □

Assumption 4. $\sup_{\gamma\in\Gamma} |(\partial/\partial\gamma) f_t(\gamma)| \leq m_t$. □

Assumption 5. *There exists a sequence of stationary ergodic random variables m_t such that for $i = 1, 2, \ldots, p$, $|\widetilde{z}_{t,i}| \leq m_t$, $|u_t| \leq m_t$, $|y_t| \leq m_t$, and for some $\omega \geq 2(\rho - 1)$, $\mathbb{E}[m_t^{6+3\omega}] < \infty$, where ρ is given in Assumption 1, and $z_{t,i}$ is the i-th row element of z_t.* □

Assumption 6. *For each $\gamma \neq 0$, $V_1(\gamma)$ and $V_2(\gamma)$ are positive definite, where for each γ, $V_1(\gamma) := \mathbb{E}[u_t^2 \widetilde{r}_t(\gamma)\widetilde{r}_t(\gamma)']$ and $V_2(\gamma) := \mathbb{E}[\widetilde{r}_t(\gamma)\widetilde{r}_t(\gamma)']$ with $\widetilde{r}_t(\gamma) := (f_t(\gamma)z_t', z_t')'$.* □

Assumption 3(*i*) implies that the model in Assumption 2 is not dynamically misspecified, and Assumption 3(*ii*) implies that the error is conditionally homoskedastic. Here, conditional homoskedasticity is not needed in proving the main theorems in this study, but it will be below additionally assumed whenever this facilitates understanding the theoretical results. Assumption 4 plays an integral role in applying the tightness condition in Doukhan, Massart, and Rio (1995) to the QLR test statistic. Here, Assumption 4 can be easily verified for the ESTAR and LSTAR models by noting that $|(\partial/\partial\gamma) f_E(\widetilde{z}_t'\alpha, \gamma)| = (1 - f_E(\widetilde{z}_t'\alpha, \gamma))(\widetilde{z}_t'\alpha)^2 \leq (\widetilde{z}_t'\alpha)^2$ and $|(\partial/\partial\gamma) f_L(\widetilde{z}_t'\alpha, \gamma)| = f_L(\widetilde{z}_t'\alpha, \gamma)(1 - f_L(\widetilde{z}_t'\alpha, \gamma))|(\widetilde{z}_t'\alpha)| \leq |(\widetilde{z}_t'\alpha)|$, so that we can let m_t in Assumption 4 be $(\widetilde{z}_t'\alpha)^2$ and $|(\widetilde{z}_t'\alpha)|$, respectively. The moment condition in Assumption 5 is stronger than those in Cho, Ishida, and White (2011, 2014), and it also implies that $\mathbb{E}[u_t^6]$ and $\mathbb{E}[y_t^6]$ are finite. The multiplicative component $f_t(\gamma)z_t'\theta$ in the STAR model makes the stronger moment condition necessary in deriving the regular null limit distribution of the QLR test statistic. Assumption 6 is imposed for the invertibility of the limit covariance matrix. This makes our test statistic non-degenerate.

Given these assumptions, we have the following lemma.

Lemma 1. *Given Assumptions 1, 2, 3(i), 4, 5, 6, and $\mathcal{H}_{01}$, (i) $\widehat{\sigma}_{n,0}^2 \overset{\text{a.s.}}{\to} \sigma_*^2 := \mathbb{E}[u_t^2]$; (ii) $\{n^{-1/2}Z'F(\cdot)Mu, \widehat{\sigma}_{n,0}^2 n^{-1}Z'F(\cdot)M\,F(\cdot)Z\} \Rightarrow \{\mathcal{Z}_1(\cdot), A_1(\cdot,\cdot)\}$ on $\Gamma(\epsilon)$ and $\Gamma(\epsilon)\times\Gamma(\epsilon)$, respectively, where $\Gamma(\epsilon) := \{\gamma \in \Gamma : |\gamma| \geq \epsilon\}$, $\mathcal{Z}_1(\cdot)$ is a continuous Gaussian process with $\mathbb{E}[\mathcal{Z}_1(\gamma)] = 0$, and for each γ and $\widetilde{\gamma}$, $\mathbb{E}[\mathcal{Z}_1(\gamma)\mathcal{Z}_1(\widetilde{\gamma})'] = B_1(\gamma,\widetilde{\gamma})$ such that $B_1(\gamma,\widetilde{\gamma}) := \mathbb{E}[u_t^2 f_t^*(\gamma) f_t^*(\widetilde{\gamma})']$ and $A_1(\gamma,\widetilde{\gamma}) := \sigma_*^2\mathbb{E}[f_t^*(\gamma)f_t^*(\widetilde{\gamma})']$ with $f_t^*(\gamma) = f_t(\gamma)z_t - \mathbb{E}[f_t(\gamma)z_t z_t']\,\mathbb{E}[z_t z_t']^{-1}z_t$; (iii) if, in addition, Assumption 3(ii) holds, $B_1(\gamma,\widetilde{\gamma}) = A_1(\gamma,\widetilde{\gamma})$.* □

There is a caveat to Lemma 1. It is clear from (2) that the null limit distribution of $QLR_n^{(1)}$ is determined by the limit behaviour under $\mathcal{H}_{01}$ of both $n^{-1/2}Z'F(\cdot)Mu$ and $n^{-1}Z'F(\cdot)MF(\cdot)Z$. Furthermore, $\lim_{\gamma\to 0} Z'F(\gamma)Mu \overset{\text{a.s.}}{=} Z'F(0)Mu = 0$ and $\lim_{\gamma\to 0} Z'F(\gamma)MF(\gamma)Z \overset{\text{a.s.}}{=} Z'F(0)MF(0)Z = 0$ by the definition of $f_t(\cdot)$. This implies that it is hard to obtain the limit distribution of $QLR_n^{(1)}$ around $\gamma = 0$. We therefore assume for the moment that 0 is not included in Γ by considering $\Gamma(\epsilon)$ instead of Γ and accommodate this effect by restricting the QLR test statistic to

$$QLR_n^{(1)}(\epsilon) := \max_{\gamma\in\Gamma(\epsilon)} \frac{1}{\widehat{\sigma}_{n,0}^2} u'MF(\gamma)Z[Z'F(\gamma)MF(\gamma)Z]^{-1}Z'F(\gamma)Mu.$$

We relax this restriction when the limit distribution is below examined under $\mathcal{H}_0$.

Lemma 1 plays a central role in deriving the null limit distribution of $QLR_n^{(1)}(\epsilon)$ and corresponds to lemma 1 of Cho, Ishida, and White (2011). Despite being similar, the two lemmas are not identical. Note that $\mathcal{Z}_1(\cdot)$ is mapped to $\mathbb{R}^{p+1}$, whereas their lemma obtains a univariate Gaussian process. The multidimensional Gaussian process $\mathcal{Z}_1(\cdot)$ distinguishes the STAR model-based testing from the ANN-based approach. The STAR model has a different null limit distribution by this, and the QLR test based upon the STAR model has power over alternatives in different directions from those of the ANN-based approach.

Theorem 1. *Given Assumptions 1, 2, 3(i), 4, 5, 6, and $\mathcal{H}_{01}$, for each $\epsilon > 0$, (i) $QLR_n^{(1)}(\epsilon) \Rightarrow \sup_{\gamma\in\Gamma(\epsilon)} \mathcal{G}_1(\gamma)'\mathcal{G}_1(\gamma)$, where $\mathcal{G}_1(\cdot)$ is a Gaussian process such that for each γ, $\mathbb{E}[\mathcal{G}_1(\gamma)] = 0$ and $\mathbb{E}[\mathcal{G}_1(\gamma)\mathcal{G}_1(\widetilde{\gamma})'] = A_1(\gamma,\gamma)^{-1/2}B_1(\gamma,\widetilde{\gamma})$ $A_1(\widetilde{\gamma},\widetilde{\gamma})^{-1/2}$; (ii) if, in addition, Assumption 3 (ii) holds, $\mathbb{E}[\mathcal{G}_1(\gamma)\mathcal{G}_1(\widetilde{\gamma})'] = A_1(\gamma,\gamma)^{-1/2}A_1(\gamma,\widetilde{\gamma})\ A_1(\widetilde{\gamma},\widetilde{\gamma})^{-1/2}$.* $\square$

As continuous mapping makes proving Theorem 1 trivial, no proof is given.

Theorem 1 implies that $QLR_n^{(1)}(\epsilon)$ does not asymptotically follow a chi-squared distribution under $\mathcal{H}_{01}$ as does the LM statistic in Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994), and Granger and Teräsvirta (1993). The difficulty here is that the null limit distribution contains the unidentified nuisance parameter γ. We can overcome this obstacle by applying Hansen's (1996) weighted bootstrap as in Cho, Cheong, and White (2011) and Cho, Ishida, and White (2011, 2014) to the QLR test statistic.

2.4.2 Type II Analysis: Testing $\mathcal{H}_{02} : \gamma_* = 0$

In this subsection, we study the limit distribution under $\mathcal{H}_{02} : \gamma_* = 0$. This is the null hypothesis used in deriving the LM statistic in Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994), and Granger and Teräsvirta (1993). As we know, θ_* is not identified under $\mathcal{H}_{02}$. We therefore maximize the QL function with respect to θ at the final stage: $\mathcal{L}_n^{(2)} := \sup_\theta \sup_\gamma \sup_\pi -\sum_{t=1}^n \{y_t - z_t'\pi - f_t(\gamma)(z_t'\theta)\}^2$, and denote the QLR test defined by this maximization process by $QLR_n^{(2)}$.

Several remarks are in order. First, maximizing the QL with respect to π is relatively simple due to linearity. We let the concentrated QL (CQL) function be $\mathcal{L}_n^{(2)}(\gamma,\theta) := \sup_\pi \mathcal{L}_n(\pi,\theta,\gamma) = -[y - F(\gamma)Z\theta]'M[y - F(\gamma)Z\theta]$, where $y := [y_1, y_2, \ldots, y_n]$. Here, we cannot let $\theta = 0$. If so, the alternative model reduces to a linear model, so that the

QLR test statistic cannot compare the alternative and null models. Second, $\mathcal{L}_n^{(2)}(\cdot)$ is not linear with respect to γ, so that the next stage CQL function with respect to γ cannot be analytically derived. We approximate the CQL function with respect to γ around $\gamma_* = 0$ and capture its limit behaviour under $\mathcal{H}_{02}$. The first-order derivative of $\mathcal{L}_n^{(2)}(\gamma, \theta)$ with respect to γ is

$$(d/d\gamma)\, \mathcal{L}_n^{(2)}(\gamma, \theta) = 2[y - F(\gamma) Z\theta]' M (\partial F(\gamma)/\partial\gamma) Z\theta,$$

where $(\partial F(\gamma)/\partial\gamma) := (\partial/\partial\gamma)\,(f(\widetilde{z}_1'\alpha, \gamma), ..., f(\widetilde{z}_n'\alpha, \gamma))$. For the LSTAR model, $\partial f_L(\widetilde{z}_t'\alpha, \gamma)/\partial\gamma = f_L(\widetilde{z}_t'\alpha, \gamma)(1 - f_L(\widetilde{z}_t'\alpha, \gamma))\widetilde{z}_t'\alpha$ and $\partial F(0)/\partial\gamma = (1/4)(\widetilde{z}_1'\alpha, ..., \widetilde{z}_n'\alpha)'$, whereas for ESTAR, it follows that $\partial f_E(\widetilde{z}_t'\alpha, \gamma)/\partial\gamma = (\widetilde{z}_t'\alpha)^2(1 - f_E(\widetilde{z}_t'\alpha, \gamma))$, so $\partial F(0)/\partial\gamma = ((\widetilde{z}_1'\alpha)^2, ..., (\widetilde{z}_n'\alpha)^2)'$, implying that we can approximate the CQL function by a second-order approximation. Nevertheless, as Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994), and Cho, Ishida, and White (2011, 2014) point out, the first-order derivative is often zero for many other models. For example, if we had let $\mathcal{M}_A := \{\pi y_{t-1} + \theta\{1 + \exp(\gamma y_{t-1})\}^{-1} : (\pi, \theta, \gamma) \in \Pi \times \Theta \times \Gamma\}$, the first-order derivative is zero when $\gamma_* = 0$. Due to this fact, a higher-order approximation is needed. Cho, Ishida, and White (2014) adopt a sixth-order Taylor expansion, whereas Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994), and Cho, Ishida, and White (2011) use fourth-order Taylor expansions to obtain the null limit distributions of the LM or QLR test statistics under $\mathcal{H}_{02}$. The order of expansion is determined by the functional form of $f(\widetilde{z}_t'\alpha, \cdot)$.

As our model does not have a fixed form of the STAR model, we fix our model scope by letting κ ($\kappa \in \mathbb{N}$) be the smallest order such that the κ-th order partial derivative with respect to γ is different from zero at $\gamma = 0$, so that for all $j < \kappa$, $(\partial^j/\partial\gamma^j)\mathcal{L}_n^{(2)}(0, \cdot) \equiv 0$. For example, $\kappa = 3$ for $\mathcal{M}_A$. Then, the CQL function is expanded as

$$\mathcal{L}_n^{(2)}(\gamma, \theta) = \mathcal{L}_n^{(2)}(0, \theta) + \frac{1}{\kappa!}\frac{\partial^\kappa}{\partial\gamma^\kappa}\mathcal{L}_n^{(2)}(0, \theta)\gamma^\kappa + \ldots + \frac{1}{(2\kappa)!}\frac{\partial^{2\kappa}}{\partial\gamma^{2\kappa}}\mathcal{L}_n^{(2)}(0, \theta)\gamma^{2\kappa} + o_{\mathbb{P}}(\gamma^{2\kappa}). \tag{3}$$

Note that for $j = 1, 2, \ldots, \kappa - 1$, $(\partial^j/\partial\gamma^j)\mathcal{L}_n^{(2)}(0, \theta) = 0$ by the definition of κ. If $\kappa = 1$, the first-order derivative differs from zero, so that none of the derivatives is zero, meaning that $j = 0$. The partial derivatives in (3) are given in the following lemma:

Lemma 2. *Given Assumption 2, the definition of κ, and $\mathcal{H}_{02}$, for each $\theta \neq 0$,*

$$\frac{\partial^j}{\partial\gamma^j}\mathcal{L}_n^{(2)}(0, \theta) = \begin{cases} 2\theta' Z' H_j(0) M u, & \textit{for } \kappa \leq j < 2\kappa\,; \\ 2\theta' Z' H_{2\kappa}(0) M u - \binom{2\kappa}{\kappa}\theta' Z' H_\kappa(0) M H_\kappa(0) Z\theta, & \textit{for } j = 2\kappa, \end{cases}$$

where $H_j(\gamma) := (\partial^j/\partial\gamma^j) F(\gamma)$. □

Using Lemma 2 we can specifically write (3) as

$$\mathcal{L}_n^{(2)}(\gamma, \theta) - \mathcal{L}_n^{(2)}(0, \theta) = \sum_{j=\kappa}^{2\kappa}\frac{2}{j!}\{\theta' Z' H_j(0) M u\}\gamma^j - \frac{1}{(2\kappa)!}\binom{2\kappa}{\kappa}\theta' Z' H_\kappa(0) M H_\kappa(0) Z\theta\gamma^{2\kappa} + o_{\mathbb{P}}(\gamma^{2\kappa}). \tag{4}$$

To reduce notational clutter, we further let $G_j := [g_{j,1}, g_{j,2}, \ldots, g_{j,n}]' := MH_j(0)Z$, where $g_{j,t} := h_{j,t}(0)z_t - Z'H_j(0)Z\,(Z'Z)^{-1}Z'z_t$ and $\varsigma_n := n^{1/2\kappa}\gamma$ with $h_{j,t}(0)$ being the t-th diagonal element of $H_j(0)$. Then, (4) is written as

$$\mathcal{L}_n^{(2)}(\gamma,\theta) - \mathcal{L}_n^{(2)}(0,\theta) = \sum_{j=\kappa}^{2\kappa} \frac{2}{j!\,n^{j/2\kappa}}\{\theta' G_j' u\}\varsigma_n^j - \frac{1}{(2\kappa)!\,n}\binom{2\kappa}{\kappa}\{\theta' G_\kappa' G_\kappa \theta\}\varsigma_n^{2\kappa} + o_{\mathbb{P}}(\gamma^{2\kappa}). \tag{5}$$

We note that if $j = \kappa$, $n^{-j/2\kappa}G_j'u = O_{\mathbb{P}}(1)$ by applying the central limit theorem. Furthermore, for $j = \kappa + 1, \ldots, 2\kappa - 1$, $n^{-j/(2\kappa)}(\partial^j/\partial\gamma^j)\mathcal{L}_n^{(2)}(\gamma,\theta) = o_{\mathbb{P}}(1)$ and $\theta' G_{2\kappa} u = o_{\mathbb{P}}(n)$ by the ergodic theorem, so that they become asymptotically negligible, implying that the smallest j-th component greater than κ and surviving at the limit becomes the second-final term in the right side of (5) that is obtained by letting $j = 2\kappa$. Note that $n^{-1}G_\kappa' G_\kappa = O_{\mathbb{P}}(1)$, if the ergodic theorem applies. Furthermore, the terms with $j > 2\kappa$ belong to $o_{\mathbb{P}}(\gamma^{2\kappa})$ by Taylor's theorem, so that they are asymptotically negligible under the null at any rate. Due to this fact, $\mathcal{L}_n^{(2)}(\cdot,\theta)$ is approximated by the 2κ-th degree polynomial function in (5), and we provide the following condition for the asymptotic analysis of the polynomial function:

Assumption 7. *For each $j = \kappa, \kappa+1, \ldots, 2\kappa$ and $i = 0, 1, \ldots, p$, (i) $\mathbb{E}[|u_t|^8] < \infty$, $\mathbb{E}[|h_{j,t}(0)|^8] < \infty$, and $\mathbb{E}[|z_{t,i}|^4] < \infty$; or (ii) $\mathbb{E}[|u_t|^4] < \infty$, $\mathbb{E}[|h_{j,t}(0)|^8] < \infty$, and $\mathbb{E}[|z_{t,i}|^8] < \infty$.* □

Assumption 7 is provided to apply the CLT to $n^{-1/2}G_j'u$ for $j = \kappa, \kappa+1, \ldots, 2\kappa$. Note that $G_j'u = \sum_{t=1}^n (u_t g_{j,t})$, and $\mathbb{E}[(u_t g_{j,t})^2] < \infty$ by the moment conditions in Assumption 7 and Cauchy-Schwarz inequality as shown below in the proof of Lemma 4, so that for $j = \kappa+1, \ldots, 2\kappa-1$, $n^{-j/2\kappa}G_j'u = o_{\mathbb{P}}(1)$ and $\overline{QLR}_n^{(2)}(\theta)$ is derived from (5). Although the QLR test statistic is approximated by the 2κ-th degree polynomial function, the moment conditions in Assumption 7 are sufficient to apply the CLT to the first term in (5).

We establish the following lemma by collecting the terms asymptotically surviving under the null:

Lemma 3. *Given Assumptions 1, 2, 7, and $\mathcal{H}_{02}$, $QLR_n^{(2)} = \sup_\theta \overline{QLR}_n^{(2)}(\theta) + o_{\mathbb{P}}(n)$, where for given $\theta \neq 0$,*

$$\overline{QLR}_n^{(2)}(\theta) := \sup_{\varsigma_n} \frac{1}{\widehat{\sigma}_{n,0}^2}\left\{\frac{2}{\kappa! n^{1/2}}\{\theta' G_\kappa' u\}\varsigma_n^\kappa - \frac{1}{(2\kappa)! n}\binom{2\kappa}{\kappa}\theta' G_\kappa' G_\kappa \theta \varsigma_n^{2\kappa}\right\}$$

and $\widehat{\varsigma}_n^\kappa(\theta)$ denotes the value of ς_n^κ that maximizes the given objective function, so that

$$\widehat{\varsigma}_n^\kappa(\theta) = \begin{cases} W_n(\theta), & \textit{if } \kappa \textit{ is odd;} \\ \max[0, W_n(\theta)], & \textit{if } \kappa \textit{ is even,} \end{cases}$$

where $W_n(\theta) := \kappa! n^{1/2}\{\theta' G_\kappa' u\}/\{\theta' G_\kappa' G_\kappa \theta\}$. □

Lemma 3 implies that the functional form of $\overline{QLR}_n^{(2)}(\cdot)$ depends on κ: for each $\theta \neq 0$,

$$\overline{QLR}_n^{(2)}(\theta) = \begin{cases} \frac{1}{\widehat{\sigma}_{n,0}^2}\frac{(\theta' G_\kappa' u)^2}{\theta' G_\kappa' G_\kappa \theta}, & \text{if } \kappa \text{ is odd;} \\ \frac{1}{\widehat{\sigma}_{n,0}^2}\max\left[0, \frac{(\theta' G_\kappa' u)^2}{\theta' G_\kappa' G_\kappa \theta}\right], & \text{if } \kappa \text{ is even.} \end{cases}$$

If θ is a scalar as in the previous literature, θ cancels out, so maximization with respect to θ does not matter any longer, see Cho, Ishida, and White (2011, 2014). This implies that $QLR_n^{(2)}$ and $\overline{QLR}_n^{(2)}(\cdot)$ are asymptotically equivalent under $\mathcal{H}_{02}$. On the other hand, if θ is a vector, the asymptotic null distribution of the test statistic has to be determined by further maximizing $\overline{QLR}_n^{(2)}(\cdot)$ with respect to θ.

We now derive the regular asymptotic distribution of QLR test statistic under $\mathcal{H}_{02}$. The following additional condition is sufficient for doing it:

Assumption 8. *$V_3(0)$ and $V_4(0)$ are positive definite, where for each γ, $V_3(\gamma) := \mathbb{E}[u_t^2 \bar{r}_t(\gamma)\bar{r}_t(\gamma)']$ and $V_4(\gamma) := \mathbb{E}[\bar{r}_t(\gamma)\ \bar{r}_t(\gamma)']$ with $\bar{r}_t(\gamma) := (h_{t,\kappa}(\gamma)z_t', z_t')'$.* □

We note that the nuisance parameter γ does not play a significant role in Assumption 8 as it does in the previous case, because $\overline{QLR}_n(\cdot)$ has already concentrated the QL function with respect to γ. Given these regularity conditions, the key limit results of the components that constitute $\overline{QLR}_n^{(2)}(\cdot)$ appear in the following lemma:

Lemma 4. *Given Assumptions 1, 2, 3(i), 4, 7, 8, and $\mathcal{H}_{02}$, (i) $n^{-1/2}G_\kappa' u \Rightarrow \mathcal{Z}_2$, where $\mathbb{E}[\mathcal{Z}_2] = 0$ and $\mathbb{E}[\mathcal{Z}_2\mathcal{Z}_2'] = \mathbb{E}[u_t^2 g_{t,\kappa} g_{t,\kappa}']$; (ii) $n^{-1}G_\kappa' G_\kappa \overset{\text{a.s.}}{\to} A_2$, where $A_2 := \mathbb{E}[g_{t,\kappa} g_{t,\kappa}']$; and (iii) if, additionally, Assumption 3(iii) holds, $\mathbb{E}[u_t^2 g_{t,\kappa} g_{t,\kappa}'] = \sigma_*^2 \mathbb{E}[g_{t,\kappa} g_{t,\kappa}']$.* □

Using this lemma, the following theorem describes the limit distribution of $QLR_n^{(2)}$ under $\mathcal{H}_{02}$.

Theorem 2. *Given Assumptions 1, 2, 3(i), 4, 7, 8, and $\mathcal{H}_{02}$, (i)*

$$QLR_n^{(2)} \Rightarrow \begin{cases} \max_{\theta\in\Theta} \mathcal{G}_2(\theta)^2, & \textit{if } \kappa \textit{ is odd;} \\ \max_{\theta\in\Theta} \max[0, \mathcal{G}_2(\theta)]^2, & \textit{if } \kappa \textit{ is even,} \end{cases}$$

where $\mathcal{G}_2(\cdot)$ is a Gaussian process such that for each θ, $\mathbb{E}[\mathcal{G}_2(\theta)] = 0$ and

$$\mathbb{E}[\mathcal{G}_2(\theta)\mathcal{G}_2(\widetilde{\theta})] = \frac{B_2(\theta, \widetilde{\theta})}{A_2(\theta,\theta)^{1/2} A_2(\widetilde{\theta},\widetilde{\theta})^{1/2}},$$

where $B_2(\theta,\widetilde{\theta}) := \theta'\mathbb{E}[u_t^2 g_{t,\kappa} g_{t,\kappa}']\widetilde{\theta}$ and $A_2(\theta,\widetilde{\theta}) := \sigma_*^2\theta'\mathbb{E}[g_{t,\kappa} g_{t,\kappa}']\widetilde{\theta}$; (ii) if, additionally, Assumption 3(iii) holds,

$$\mathbb{E}[\mathcal{G}_2(\theta)\mathcal{G}_2(\widetilde{\theta})] = \frac{A_2(\theta, \widetilde{\theta})}{A_2(\theta,\theta)^{1/2} A_2(\widetilde{\theta},\widetilde{\theta})^{1/2}}.$$ □

As Theorem 2 trivially follows from Lemma 4 and continuous mapping, its proof is omitted.

Several remarks are in order. First, the covariance kernel of $\mathcal{G}_2(\cdot)$ is bilinear with respect to θ and $\widetilde{\theta}$. This implies that $\mathcal{G}_2(\theta)$ is a linear Gaussian process with respect to θ. Therefore, if $z \sim N(0, \mathbb{E}[u_t^2 g_{t,\kappa} g_{t,\kappa}'])$, $z'\theta$ as a function of θ is distributionally equivalent to $\mathcal{G}_2(\cdot)$. This fact relates the null limit distribution to the chi-squared distribution. Corollary 1 of Cho and White (2018) shows that $\max_{\theta\in\Theta}\mathcal{G}_2(\theta)^2 \overset{\text{d}}{=} \mathcal{X}_{p+1}^2$ if $\mathcal{G}_2(\cdot)$ is a linear Gaussian process and $\mathbb{E}[u_t^2 g_{t,\kappa} g_{t,\kappa}'] = \sigma_*^2\mathbb{E}[g_{t,\kappa} g_{t,\kappa}']$, where $\mathcal{X}_{p+1}^2$ is a chi-squared distribution with $p+1$ degrees of freedom . Second, the chi-squared null limit distributions of the LM test statistics in Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994), and Granger and Teräsvirta (1993) follow from the fact that the LM test statistic is equivalent to the QLR test statistic under $\mathcal{H}_{02}$. Third, the process to obtain the limit distribution of the QLR test under $\mathcal{H}_{02}$ of here is simpler than that of Cho, Ishida, and White (2011) and Cho and Phillips (2018) in which they examine two other identification problems. In our context, $f_t(\cdot)$ is defined by centering $f(\cdot,\gamma)$ at $f(\cdot,0)$, as π_* and θ_* are not separately identified. Finally, if $\theta = 0$, $\mathcal{G}_2(\theta)$ is not well defined as the weak limit in Theorem 2 is obtained by assuming that $\theta \neq 0$. Notwithstanding this feature, the null limit distribution of the QLR test statistic is well represented by Theorem 2 as the QLR test statistic obtain the alternative model by letting $\theta \neq 0$.

2.4.3 Null Limit Distribution of the QLR Test Statistic under $\mathcal{H}_0$

In this subsection, we derive the limit distribution of the QLR test under $\mathcal{H}_0$ by examining the relationship between $QLR_n^{(1)}$ and $QLR_n^{(2)}$. Specifically, we show that $QLR_n^{(1)} \geq QLR_n^{(2)}$, which means the limit distribution under $\mathcal{H}_0$ equals that of $QLR_n^{(1)}$. Although this idea is the same as the one in Cho, Ishida, and White (2011, 2014), their approach cannot be applied in the current context. This is because the associated Gaussian process $\mathcal{G}_1(\cdot)$ is multidimensional.

The following lemma generalizes the approach in Cho, Ishida, and White (2011, 2014) to STAR models.

Lemma 5. *Let $n(\gamma) := Z'F(\gamma)Mu$ and $D(\gamma) := Z'F(\gamma)MF(\gamma)Z'$ with $n^{(j)}(\gamma) := (\partial^j/\partial\gamma^j)n(\gamma)$, and $D^{(j)}(\gamma) := (\partial^j/\partial\gamma^j)D(\gamma)$. Under Assumptions 1, 2 and 3, (i) for $j < \kappa$, $\lim_{\gamma\to 0} n^{(j)}(\gamma) \overset{\text{a.s.}}{=} 0$ and $\lim_{\gamma\to 0} D^{(j)}(\gamma) \overset{\text{a.s.}}{=} 0$; (ii) $\lim_{\gamma\to 0} n^{(\kappa)}(\gamma) \overset{\text{a.s.}}{=} G_\kappa' u$; and (iii) $\lim_{\gamma\to 0} D^{(\kappa)}(\gamma) \overset{\text{a.s.}}{=} G_\kappa' G_\kappa$.* □

The limit obtained by letting $\gamma \to 0$ under $\mathcal{H}_{01}$ can be compared with that obtained under $\mathcal{H}_{02}$. More specifically, using Lemma 5 and L'Hôpital's rule, we obtain $\lim_{\gamma\to 0} n(\gamma)'D(\gamma)^{-1}n(\gamma) \overset{\text{a.s.}}{=} \lim_{\gamma\to 0} n^{(\kappa)}(\gamma)'D^{(\kappa)}(\gamma)^{-1}n^{(\kappa)}(\gamma) \overset{\text{a.s.}}{=} u'G_\kappa(G_\kappa'G_\kappa)^{-1}G_\kappa'u$. From this, it follows that $QLR_n^{(1)} \geq \sup_\theta \overline{QLR}_n^{(2)}(\theta)$ because $QLR_n^{(1)} = \sup_{\gamma\in\Gamma} \frac{1}{\widehat{\sigma}_{n,0}^2} n(\gamma)'$ $D(\gamma)^{-1}\, n(\gamma) \geq \lim_{\gamma\to 0} \frac{1}{\widehat{\sigma}_{n,0}^2} n(\gamma)'D(\gamma)^{-1}n(\gamma) \overset{\text{a.s.}}{=} \frac{1}{\widehat{\sigma}_{n,0}^2} u'G_\kappa(G_\kappa'G_\kappa)^{-1}G_\kappa'u$. Furthermore, $\overline{QLR}_n^{(2)}(\theta)$ is asymptotically equal to $\frac{1}{\widehat{\sigma}_{n,0}^2} u'G_\kappa\theta(\theta'G_\kappa'G_\kappa\theta)^{-1}\,\theta'G_\kappa'u$. Thus, it follows that $QLR_n^{(1)} \geq \sup_\theta \overline{QLR}_n^{(2)}(\theta) + o_{\mathbb{P}}(1)$, if

$$G_\kappa(G_\kappa'G_\kappa)^{-1}G_\kappa' - G_\kappa\theta(\theta'G_\kappa'G_\kappa\theta)^{-1}\theta'G_\kappa' \tag{6}$$

is positive semidefinite irrespective of θ. To show this we first note that the two terms in (6) are idempotent and symmetric matrices. Therefore, we may apply Exercise 8.58 in Abadir and Magnus (2005, p. 233). Then,

$$\{G_\kappa(G_\kappa'G_\kappa)^{-1}G_\kappa'\}\{G_\kappa\theta(\theta'G_\kappa'G_\kappa\theta)^{-1}\theta'G_\kappa'\} = G_\kappa\theta(\theta'G_\kappa'G_\kappa\theta)^{-1}\theta'G_\kappa'$$

so that (6) is positive semidefinite. This implies

$$QLR_n = \max[QLR_n^{(1)}, QLR_n^{(2)}] + o_{\mathbb{P}}(1) = \max\left[QLR_n^{(1)}, \sup_\theta \overline{QLR}_n^{(2)}(\theta)\right] + o_{\mathbb{P}}(1) = QLR_n^{(1)} + o_{\mathbb{P}}(1).$$

Given that $\Gamma(\epsilon)$ was considered in Theorem 1 to remove $\gamma = 0$ from Γ so that the QLR test statistic is computed by comparing the alternative model with the null model, if ϵ is selected as small as possible to have $QLR_n = QLR_n(\epsilon) + o_{\mathbb{P}}(1)$ and we can let $\gamma \to 0$ as posited in Lemma 5, the null limit distribution of the QLR test statistic is characterized by the Gaussian process provided in Theorem 1. Therefore, if the conditions in Theorems 1 and 2 hold simultaneously, the null limit distribution of the QLR test statistic is derived by combining Theorems 1 and 2. For this purpose, we combine Assumptions 6 and 8 into a new assumption as follows:

Assumption 9. *For each $\gamma \neq 0$, $V_5(\gamma)$ and $V_6(\gamma)$ are positive definite, where for each γ, $V_5(\gamma) := \mathbb{E}[u_t^2 \ddot{r}_t(\gamma)\ddot{r}_t(\gamma)']$ and $V_6(\gamma) := \mathbb{E}[\ddot{r}_t(\gamma)\ddot{r}_t(\gamma)']$ with $\ddot{r}_t(\gamma) := (h_{t,\kappa}(0)z_t', f_t(\gamma)z_t', z_t')'$.* □

The following theorem now yields the limit distribution of the QLR test under $\mathcal{H}_0$.

Theorem 3. *Given Assumptions 1, 2, 3(i), 4, 5, 7, 9, and $\mathcal{H}_0$, (i) $QLR_n \Rightarrow \sup_{\gamma \in \Gamma} \mathcal{G}_1(\gamma)'\mathcal{G}_1(\gamma)$, where $\mathcal{G}_1(\cdot)$ is a Gaussian process such that for each γ, $\mathbb{E}[\mathcal{G}_1(\gamma)] = 0$ with $\mathbb{E}[\mathcal{G}_1(\gamma)\mathcal{G}_1(\widetilde{\gamma})'] = A_1(\gamma,\gamma)^{-1/2}B_1(\gamma,\widetilde{\gamma})A_1(\widetilde{\gamma},\widetilde{\gamma})^{-1/2}$; (ii) if Assumption 3(ii) also holds, $\mathbb{E}[\mathcal{G}_1(\gamma)\mathcal{G}_1(\widetilde{\gamma})'] = A_1(\gamma,\gamma)^{-1/2}A_1(\gamma,\widetilde{\gamma})A_1(\widetilde{\gamma},\widetilde{\gamma})^{-1/2}$.* □

Theorem 3 immediately follows from Theorems 1 and 2 and from our earlier argument that $QLR_n = QLR_n^{(1)} + o_{\mathbb{P}}(1)$, which is why we do not prove it in the Appendix. Note that the consequence of Theorem 3 is the same as that of Theorem 1, although the null hypothesis is extended to $\mathcal{H}_0$ from $\mathcal{H}_{01}$ by enlarging the parameter space from $\Gamma(\epsilon)$ to Γ with sufficiently small ϵ. We here again emphasize that the Gaussian process $\mathcal{G}_1(\cdot)$ is obtained by supposing that $\gamma \neq 0$. Otherwise, a meaningful QLR test statistic is not properly defined.

The given null limit distribution in Theorem 3 is derived as in Cho, Ishida, and White (2011, 2014) and Baek, Cho, and Phillips (2015). Nevertheless, our proofs generalize theirs due to the existence of the multidimensional Gaussian process. Furthermore, this null limit distribution extends the scope of the LM test statistics in Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994), and Granger and Teräsvirta (1993) who only test $\mathcal{H}_{02}$.

3 Monte Carlo Experiments

In this section, we illustrate testing linearity using the ESTAR and LSTAR models and simulate the QLR test statistic to support the statistical theory in Section 2. Hansen's (1996) weighted bootstrap is also applied to enhance the applicability of our methodology.

3.1 Illustration Using the ESTAR Model

To simplify our illustration, we assume that for all $t = 1, 2, \ldots$, $u_t \sim$ IID $N(0, \sigma_*^2)$ and $y_t = \pi_* y_{t-1} + u_t$ with $\pi_* = 0.5$. Under this DGP, we specify the following first-order ESTAR model:

$$\mathcal{M}_{ESTAR} := \{\pi y_{t-1} + \theta y_{t-1}\{1 - \exp[-\gamma y_{t-1}^2]\} : \pi \in \Pi,\ \theta \in \Theta,\ \text{and } \gamma \in \Gamma\}.$$

The model does not contain an intercept, and the transition variable is y_{t-1}. The nonlinear function $f_t(\gamma) = 1 - \exp(-\gamma y_{t-1}^2)$ is defined on Γ which is compact and convex, and the exponential function is analytic. This means that the QLR test statistic is generically comprehensively revealing. To identify the model it is assumed that $\gamma_* > 0$. In our model set-up, we allow 0 to be included in Γ. The nonlinear function $f_t(\cdot)$ satisfies $f_t(0) = 0$. Given this model, the following hypotheses are of interest:

$$\mathcal{H}_0' : \exists \pi \in \mathbb{R}, \mathbb{P}(\mathbb{E}[y_t|y_{t-1}] = \pi y_{t-1}) = 1; \text{ vs. } \mathcal{H}_1' : \forall \pi \in \mathbb{R}, \mathbb{P}(\mathbb{E}[y_t|y_{t-1}] = \pi y_{t-1}) < 1,$$

Two parameter restrictions make $\mathcal{H}_0'$ valid: either $\theta_* = 0$ or $\gamma_* = 0$. The sub-hypotheses are thus $\mathcal{H}_{01}' : \theta_* = 0$ and $\mathcal{H}_{02}' : \gamma_* = 0$.

We first examine the null distribution of the QLR test under $\mathcal{H}_{01}'$. By Theorem 1, the limit null distribution of this test statistic is given as

$$QLR_n^{(1)} = \sup_{\gamma \in \Gamma} \frac{1}{\widehat{\sigma}_{n,0}^2} \frac{(u'MF(\gamma)Z)^2}{Z'F(\gamma)MF(\gamma)Z} \Rightarrow \sup_{\gamma \in \Gamma} \widetilde{\mathcal{G}}_1(\gamma)^2$$

where $\widetilde{\mathcal{G}}_1(\cdot)$ is a mean-zero Gaussian process with the covariance structure

$$\widetilde{\rho}_1(\gamma, \widetilde{\gamma}) = \frac{\widetilde{k}_1(\gamma, \widetilde{\gamma})}{c_1(\gamma, \gamma)^{1/2} \widetilde{c}_1(\widetilde{\gamma}, \widetilde{\gamma})^{1/2}} \tag{7}$$

with $\widetilde{k}_1(\gamma, \widetilde{\gamma}) = \widetilde{c}_1(\gamma, \widetilde{\gamma}) = \sigma_*^2 \mathbb{E}[y_t^2 \exp(-(\gamma + \widetilde{\gamma}) y_t^2)] - \sigma_*^2 \mathbb{E}[y_t^2 \exp(-\gamma y_t^2)] \mathbb{E}[y_t^2]^{-1} \mathbb{E}[y_t^2 \exp(-\widetilde{\gamma} y_t^2)]$. Furthermore, under $\mathcal{H}_{01}'$, y_t is normally distributed with $\mathbb{E}[y_t] = 0$ and $\mathrm{var}[y_t] = \sigma_y^2 := \sigma_*^2/(1 - \pi_*^2)$, so that y_t^2 follows the gamma distribution with shape parameter $1/2$ and scale parameter $2\sigma_*^2/(1 - \pi_*^2)$. Define

$$\widetilde{m}(\gamma) := \left(1 + \frac{2\sigma_*^2}{1 - \pi_*^2} \gamma\right)^{-\frac{1}{2}}, \quad \text{and} \quad \widetilde{h}(\gamma, \widetilde{\gamma}) := \frac{1}{\sigma_y^2} \left(\left[\frac{1 + 2\sigma_y^2(\gamma + \widetilde{\gamma})}{(1 + 2\sigma_y^2 \gamma)(1 + 2\sigma_y^2 \widetilde{\gamma})} \right]^{-3/2} - 1 \right).$$

Note that $\widetilde{m}(\gamma) = \mathbb{E}[\exp(-\gamma y_t^2)]$, so that $\mathbb{E}[y_t^2 \exp(-\gamma y_t^2)] = -\widetilde{m}'(\gamma)$. As a result, (7) is further simplified to $\widetilde{k}_1(\gamma, \widetilde{\gamma}) = \sigma_*^2 \widetilde{m}'(\gamma) \widetilde{m}'(\widetilde{\gamma}) \widetilde{h}(\gamma, \widetilde{\gamma})$, and

$$\widetilde{\rho}_1(\gamma, \widetilde{\gamma}) = \frac{\widetilde{k}_1(\gamma, \widetilde{\gamma})}{\widetilde{c}_1(\gamma, \gamma)^{1/2} \widetilde{c}_1(\widetilde{\gamma}, \widetilde{\gamma})^{1/2}} = \frac{\widetilde{h}(\gamma, \widetilde{\gamma})}{\widetilde{h}(\gamma, \gamma)^{1/2} \widetilde{h}(\widetilde{\gamma}, \widetilde{\gamma})^{1/2}}.$$

We next examine the limit distribution of the QLR test statistic under $\mathcal{H}_{02}'$: $\gamma_* = 0$. The first-order derivative $(\partial/\partial\gamma) f_t(\gamma) = y_{t-1}^2 \exp(-\gamma y_{t-1}^2)$, which is different from zero even when $\gamma = 0$, so that in this case $\kappa = 1$. Thus, we can apply the second-order Taylor expansion to obtain the limit distribution of the QLR test statistic under $\mathcal{H}_{02}'$. As a result,

$$\overline{QLR}_n^{(2)}(\theta) = \frac{1}{\widehat{\sigma}_{n,0}^2} \frac{(\theta' G_\kappa' u)^2}{\theta' G_\kappa' G_\kappa \theta}, \tag{8}$$

where

$$\theta' G_\kappa' u = \theta \left[\sum y_{t-1}^3 u_t - \frac{\sum y_{t-1}^4 \sum y_{t-1} u_t}{\sum y_{t-1}^2} \right] \quad \text{and} \quad \theta' G_\kappa' G_\kappa \theta = \theta^2 \left[\sum y_{t-1}^6 - \frac{(\sum y_{t-1}^4)^2}{\sum y_{t-1}^2} \right].$$

In (8), θ is a scalar, so that cancels out, and it follows that $QLR_n^{(2)} \Rightarrow \widetilde{\mathcal{G}}_2^2$, where $\widetilde{\mathcal{G}}_2 \sim N(0,1)$.

These two separate results can be combined, which means that we can examine the limit distribution of the QLR test under $\mathcal{H}_0'$. We have $QLR_n \Rightarrow \sup_{\gamma \in \Gamma} \widetilde{\mathcal{G}}(\gamma)^2$, where

$$\widetilde{\mathcal{G}}(\gamma) = \begin{cases} \widetilde{\mathcal{G}}_1(\gamma), & \text{if } \gamma \neq 0; \\ \widetilde{\mathcal{G}}_2 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad \mathbb{E}[\widetilde{\mathcal{G}}(\gamma)\widetilde{\mathcal{G}}(\widetilde{\gamma})] = \begin{cases} \widetilde{\rho}_1(\gamma, \widetilde{\gamma}), & \text{if } \gamma \neq 0, \widetilde{\gamma} \neq 0; \\ 1 & \text{if } \gamma = 0, \widetilde{\gamma} = 0; \\ \widetilde{\rho}_3(\gamma). & \text{if } \gamma \neq 0, \widetilde{\gamma} = 0, \end{cases}$$

with

$$\widetilde{\rho}_3(\gamma) := \mathbb{E}[\widetilde{\mathcal{G}}_1(\gamma)\widetilde{\mathcal{G}}_2] = \frac{\sqrt{6}\sigma_y^2 \gamma}{\widetilde{h}(\gamma,\gamma)^{1/2}(1+2\sigma_y^2\gamma)} \quad \text{such that} \quad \widetilde{\rho}_3(\gamma)^2 = \lim_{\widetilde{\gamma} \to 0} \widetilde{\rho}_1(\gamma, \widetilde{\gamma})^2 = \left(\frac{\sqrt{6}\sigma_y^2 \gamma}{\widetilde{h}(\gamma,\gamma)^{1/2}(1+2\sigma_y^2\gamma)} \right)^2.$$

Thus, we conclude that $QLR_n \Rightarrow \sup_\gamma \widetilde{\mathcal{G}}(\gamma)^2$, which agrees with Theorem 3.

The null limit distribution can be approximated numerically by simulating a distributionally equivalent Gaussian process. To do this we present the following lemma:

Lemma 6. *If $\{z_k : k = 0, 1, 2, \ldots\}$ is an IID sequence of standard normal random variables, $\widetilde{\mathcal{G}}(\cdot) \overset{\mathrm{d}}{=} \overline{\mathcal{G}}(\cdot)$, where for each $\gamma \in \Gamma := \{\gamma \in \mathbb{R} : \gamma \geq 0\}$, $\overline{\mathcal{G}}(\gamma) := \sum_{k=1}^{\infty} c(\gamma) \cdot a(\gamma)^k [(-1)^k \binom{-3/2}{k}]^{1/2} z_k$, $c(\gamma) := \{\sum_{k=1}^{\infty} (-1)^k a(\gamma)^{2k} \binom{-3/2}{k} \}^{-1/2}$, and $a(\gamma) := 2\sigma_y^2\gamma/(1+2\sigma_y^2\gamma)$.* □

Note that the term $(-1)^k \binom{-3/2}{k}$ in Lemma 6 is always positive irrespective of k, and for any γ,

$$\lim_{k\to\infty} \mathrm{var}\left[a(\gamma)^k \left((-1)^k \binom{-3/2}{k} \right)^{1/2} z_k \right] = \lim_{k\to\infty} a(\gamma)^{2k} (-1)^k \binom{-3/2}{k} = 0 \tag{9}$$

and $\widetilde{h}(\gamma,\gamma) = \sum_{k=1}^{\infty} a(\gamma)^{2k}(-1)^k \binom{-3/2}{k}$. Using these facts Lemma 6 shows that for any $\gamma, \widetilde{\gamma} \neq 0$, $\mathbb{E}[\overline{\mathcal{G}}(\gamma)\overline{\mathcal{G}}(\widetilde{\gamma})] = \widetilde{\rho}_1(\gamma, \widetilde{\gamma})$. Here, the non-negative parameter space condition for Γ is necessary for $\overline{\mathcal{G}}(\cdot)$ to be properly defined on Γ. Without this condition, $\overline{\mathcal{G}}(\gamma)$ cannot be properly generated. We note that $\lim_{\gamma \downarrow 0} \overline{\mathcal{G}}(\gamma) \overset{\text{a.s.}}{=} z_1$, so that if we let $z_1 = \overline{\mathcal{G}}_2$,

$$\mathbb{E}[\overline{\mathcal{G}}(\gamma)\overline{\mathcal{G}}_2] = \frac{\sqrt{6}\sigma_y^2\gamma}{\widetilde{h}(\gamma,\gamma)^{1/2}(1+2\sigma_y^2\gamma)} = \widetilde{\rho}_3(\gamma).$$

It follows that the distribution of $\widetilde{\mathcal{G}}(\cdot)$ can be simulated by iteratively generating $\overline{\mathcal{G}}(\cdot)$. In practice,

$$\overline{\mathcal{G}}(\gamma; K) := \frac{\sum_{k=1}^{K} a(\gamma)^k [(-1)^k \binom{-3/2}{k}]^{1/2} z_k}{\sqrt{\sum_{k=1}^{K} a(\gamma)^{2k} (-1)^k \binom{-3/2}{k}}}$$

is generated by choosing K to be sufficiently large . By (9), if this is the case, the difference between the distributions of $\overline{\mathcal{G}}(\cdot)$ and $\overline{\mathcal{G}}(\cdot; K)$ becomes negligible.

We now examine the empirical distributions of the QLR statistic under several different environments. First, we consider four different parameter spaces: $\Gamma_1 = [0, 2]$, $\Gamma_2 = [0, 3]$, $\Gamma_3 = [0, 4]$, and $\Gamma_4 = [0, 5]$. They are selected to examine how the null limit distribution of the QLR test is influenced by the choice of Γ. We obtain the limit distribution by simulating $\sup_{\gamma \in \Gamma} \bar{\mathcal{G}}(\gamma; K)^2$ 5,000 times with $K = 2,000$, where Γ is in turn Γ_1, Γ_2, Γ_3, and Γ_4. Second, we study how the empirical distribution of the QLR test statistic changes with the sample size. We consider the sample sizes $n = 100$, $1,000$, $2,000$, and $5,000$.

Figure 1 summarizes the simulation results and shows that the empirical distribution approaches the null limit distribution under different parameter space conditions. We also provide the estimates of the probability density functions next to the empirical distributions. For every parameter space considered, the empirical rejection rates of the QLR test statistics are most accurate when $n = 2,000$. The empirical rejection rates are closer to the nominal levels when the parameter space is small. This result is significant when $n = 100$: the empirical rejection rates for $\Gamma = \Gamma_1$ are closer to the nominal ones than when $\Gamma = \Gamma_4$. Nonetheless, this difference becomes negligible as the sample size increases. The empirical rejection rates obtained using $n = 2,000$ are already satisfactorily close to the nominal levels, and this result is more or less similar to that from 5,000 observations. This suggests that the theory in Section 2 is effective for the ESTAR model. Considering even larger parameter spaces for γ yielded similar results, so they are not reported here.

3.2 Illustration Using the LSTAR Model

As another illustration, we consider testing against the first-order LSTAR model. We assume that the data-generating process is $y_t = \pi_* y_{t-1} + u_t$ with $\pi_* = 0.5$ and

$$u_t = \begin{cases} \ell_t, & \text{w.p. } 1 - \pi_*^2 \, ; \\ 0, & \text{w.p. } \pi_*^2 \end{cases}$$

where $\{\ell_t\}_{t=1}^n$ follows the Laplace distribution with mean 0 and variance 2. Under this assumption, y_t follows the same distribution as ℓ_t that makes the algebra associated with the LSTAR model straightforward. For example, the covariance kernel of the Gaussian process associated with the null limit distribution of the QLR test statistic is analytically obtained thanks to this distributional assumption. This DGP is a variation of the exponential autoregressive model in Lawrence and Lewis (1980). Their exponential distribution is replaced by the Laplace distribution to allow

y_t to obtain negative values.

Given this DGP, the first-order LSTAR model for $\mathbb{E}[y_t|y_{t-1}, y_{t-2}, \ldots]$ is defined as follows:

$$\mathcal{M}^0_{LSTAR} := \{\pi y_{t-1} + \theta y_{t-1}\{1 + \exp(-\gamma y_{t-1})\}^{-1} : \pi \in \Pi,\ \theta \in \Theta,\ \text{and } \gamma \in \Gamma := [0, \bar{\gamma}]\}.$$

The nonlinear logistic function $\{1 + \exp(-\gamma y_{t-1})\}^{-1}$ contains an exponential function. It is therefore analytic, and this fact delivers a consistent power for the QLR test statistic. Note, however, that for $\gamma = 0$ the value of the logistic function equals 1/2. This difficulty is avoided by subtracting 1/2 from the logistic function when carrying out the test, viz.,

$$\mathcal{M}_{LSTAR} := \{\pi y_{t-1} + \theta y_{t-1}\{[1 + \exp(-\gamma y_{t-1})]^{-1} - 1/2\} : \pi \in \Pi,\ \theta \in \Theta,\ \text{and } \gamma \in \Gamma := [0, \bar{\gamma}]\}.$$

By the invariance principle, this shift does not affect the null limit distribution of the QLR test statistic. We here let $\gamma \geq 0$ so that transition function is bounded, which modifies the limit space of ς_n into $\mathbb{R}^+$. The null and the alternative hypotheses are identical to those in the ESTAR case.

Before proceeding, note that

$$\{1 + \exp(-\gamma y_{t-1})\}^{-1} - \frac{1}{2} = \frac{1}{2} \tanh\left(\frac{\gamma y_{t-1}}{2}\right)$$

Using the hyperbolic tangent function as in Bacon and Watts (1971) makes it easy to find a Gaussian process that is distributionally equivalent to the Gaussian process obtained under the null.

Using this fact, the limit distribution of QLR test statistic under $\mathcal{H}'_{01}$ is derived as in before. By Theorem 1,

$$QLR_n^{(1)} = \sup_{\gamma \in \Gamma} \frac{1}{\widehat{\sigma}^2_{n,0}} \frac{(u'MF(\gamma)Z)^2}{Z'F(\gamma)MF(\gamma)Z} \Rightarrow \sup_{\gamma \in \Gamma} \ddot{\mathcal{G}}_1(\gamma)^2$$

where $\ddot{\mathcal{G}}_1(\cdot)$ is a mean-zero Gaussian process with the covariance structure

$$\ddot{\rho}_1(\gamma, \widetilde{\gamma}) := \frac{\ddot{k}_1(\gamma, \widetilde{\gamma})}{\ddot{c}_1(\gamma, \gamma)^{1/2} \ddot{c}_1(\widetilde{\gamma}, \widetilde{\gamma})^{1/2}}.$$

The function $\ddot{k}_1(\gamma, \tilde{\gamma})$ is equivalent to $\ddot{c}_1(\gamma, \tilde{\gamma})$ by the conditional homoskedasticity condition, and for each nonzero γ and $\widetilde{\gamma}$, we now obtain that

$$\ddot{k}_1(\gamma, \tilde{\gamma}) = \frac{1}{4}\mathbb{E}\left[y_t^2 \tanh\left(\frac{\gamma y_t}{2}\right) \tanh\left(\frac{\widetilde{\gamma} y_t}{2}\right)\right] - \frac{1}{4}\mathbb{E}\left[y_t^2 \tanh\left(\frac{\gamma y_t}{2}\right)\right] \mathbb{E}[y_t^2]^{-1} \mathbb{E}\left[y_t^2 \tanh\left(\frac{\widetilde{\gamma} y_t}{2}\right)\right].$$

In the proof of Lemma 7 given in the Appendix, we further show that $\ddot{k}_1(\gamma, \widetilde{\gamma}) = \sum_{n=1}^{\infty} b_n(\gamma) b_n(\widetilde{\gamma})$, where $b_1(\gamma) :=$

$\frac{1}{\sqrt{2}}(1-2a(\gamma))$ with $a(\gamma) := \sum_{k=1}^{\infty}(-1)^{k-1}/(1+\gamma k)^3$ and for $n = 2, 3, \ldots$,

$$b_n(\gamma) := \sqrt{n(n+1)}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}(\gamma k)^{n-1}}{(1+\gamma k)^{n+2}}.$$

Next, we derive the limit distribution of the QLR test statistic under $\mathcal{H}_{02}'$. For $\gamma = 0$, $(\partial/\partial\gamma)f_t(\gamma) = y_{t-1}\exp(\gamma y_{t-1})/[1+\exp(-\gamma y_{t-1})]^2 \neq 0$, implying that κ is unity as for the ESTAR case, so that we can apply a second-order Taylor expansion to obtain the limit distribution of the QLR test statistic under $\mathcal{H}_{02}'$:

$$\overline{QLR}_n^{(2)}(\theta) = \frac{1}{\widehat{\sigma}_{n,0}^2}\frac{(\theta' G_\kappa' u)^2}{\theta' G_\kappa' G_\kappa \theta},$$

where, similarly to the ESTAR case,

$$\theta' G_\kappa' u = \frac{\theta}{4}\left[\sum y_{t-1}^2 u_t - \frac{\theta}{4}\frac{\sum y_{t-1}^3 \sum y_{t-1}u_t}{\sum y_{t-1}^2}\right] \quad \text{and} \quad \theta' G_\kappa' G_\kappa \theta = \frac{\theta^2}{16}\left[\sum y_{t-1}^4 - \frac{(\sum y_{t-1}^3)^2}{\sum y_{t-1}^2}\right].$$

From this it follows that $QLR_n^{(2)} \Rightarrow \ddot{\mathcal{G}}_2^2$, where $\ddot{\mathcal{G}}_2 \sim N(0,1)$.

Therefore, we conclude that $QLR_n \Rightarrow \sup_\gamma \ddot{\mathcal{G}}(\gamma)^2$, where

$$\ddot{\mathcal{G}}(\gamma) := \begin{cases} \ddot{\mathcal{G}}_1(\gamma), & \text{if } \gamma \neq 0; \\ \ddot{\mathcal{G}}_2, & \text{otherwise.} \end{cases}$$

The limit variance of $\ddot{\mathcal{G}}(\gamma)$ is given as

$$\ddot{\rho}(\gamma,\widetilde{\gamma}) := \mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}(\widetilde{\gamma})] = \begin{cases} \ddot{\rho}_1(\gamma,\widetilde{\gamma}), & \text{if } \gamma \neq 0, \widetilde{\gamma} \neq 0; \\ 1 & \text{if } \gamma = 0, \widetilde{\gamma} = 0; \\ \ddot{\rho}_3(\gamma), & \text{if } \gamma \neq 0, \widetilde{\gamma} = 0, \end{cases} \quad \text{where} \quad \ddot{\rho}_3(\gamma) := \mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}_2] = \frac{\ddot{r}_1(\gamma)}{\ddot{k}_1(\gamma,\gamma)^{1/2}\ddot{q}^{1/2}}$$

with $\ddot{r}_1(\gamma) := \frac{1}{2}\mathbb{E}\left[y_{t-1}^3 \tanh\left(\frac{\gamma y_{t-1}}{2}\right)\right]$ and $\ddot{q} := \mathbb{E}[y_t^4] - \mathbb{E}[y_t^3]^2/\mathbb{E}[y_t^2]$. From this it follows that $QLR_n \Rightarrow \sup_{\gamma\in\Gamma}\ddot{\mathcal{G}}(\gamma)^2$. Furthermore, $\mathbb{E}[y_t^3] = 0$ and $\mathbb{E}[y_t^4] = 24$ given our DGP, so that

$$\ddot{\rho}_3(\gamma) = \frac{\mathbb{E}[y_t^3\tanh(\gamma y_t/2)]}{4\sqrt{6}\ddot{k}_1(\gamma,\gamma)^{1/2}}. \tag{10}$$

Here, we note that

$$\mathbb{E}[y_t^3\tanh(\gamma y_t/2)] = \frac{1}{8\gamma^4}\left[48\gamma^4 + P_G\left(3, 1+\frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right] \tag{11}$$

by some tedious algebra assisted by Mathematica, where $P_G(n,x)$ is the polygamma function: $P_G(n,x) := d^{n+1}/d$

$x^{n+1}\log(\Gamma(x))$. Inserting (11) into (10) yields

$$\ddot{\rho}_3(\gamma) = \frac{1}{32\sqrt{6}\gamma^4 \ddot{k}_1(\gamma,\gamma)^{1/2}}\left[48\gamma^4 + P_G\left(3, 1+\frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right]. \tag{12}$$

In addition, we show in the supplementary lemma (Lemma 8) given in the Appendix that applying L'Hôpital's rule iteratively yields that

$$\lim_{\widetilde{\gamma}\downarrow 0}\ddot{\rho}_1(\gamma,\widetilde{\gamma})^2 = \left(\frac{1}{32\sqrt{6}\gamma^4 \ddot{k}_1(\gamma,\gamma)^{1/2}}\left[48\gamma^4 + P_G\left(3, 1+\frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right]\right)^2. \tag{13}$$

This fact implies that $\mathrm{plim}_{\gamma\downarrow 0}\ddot{\mathcal{G}}_1(\gamma)^2 = \ddot{\mathcal{G}}_2^2$. That is, the weak limit of the QLR test statistic under $\mathcal{H}_{02}'$ can be obtained from $\ddot{\mathcal{G}}_1(\cdot)^2$ by letting γ converging to zero, so that $QLR_n \Rightarrow \sup_{\gamma\in\Gamma}\ddot{\mathcal{G}}_1(\gamma)^2$ under $\mathcal{H}_0'$.

Next, we derive another Gaussian process that is distributionally equivalent to $\ddot{\mathcal{G}}(\cdot)$ and conduct Monte Carlo simulations using it. The process is presented in the following lemma.

Lemma 7. *If $\{z_k\}_{k=1}^{\infty}$ is an IID sequences of standard normal random variables, then for each γ and $\widetilde{\gamma} \in \Gamma := \{\gamma \in \mathbb{R} : \gamma \geq 0\}$, $\ddot{\mathcal{G}}(\cdot) \overset{\mathrm{d}}{=} \dot{\mathcal{G}}(\cdot)$, where $\ddot{\mathcal{Z}}_1(\gamma) := \sum_{n=1}^{\infty} b_n(\gamma) z_n$ and $\dot{\mathcal{G}}(\gamma) := \left(\sum_{n=1}^{\infty} b_n(\gamma)^2\right)^{-1/2}\ddot{\mathcal{Z}}_1(\gamma)$.* □

We prove Lemma 7 by showing that the Gaussian process $\dot{\mathcal{G}}(\cdot)$ given in Lemma 7 has the same covariance structure as $\ddot{\mathcal{G}}(\cdot)$, and for this purpose, we focus on proving that for all $\gamma, \widetilde{\gamma} \geq 0$, $\mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}(\widetilde{\gamma})] = \mathbb{E}[\dot{\mathcal{G}}(\gamma)\dot{\mathcal{G}}(\widetilde{\gamma})]$ in the Appendix. If $\gamma, \widetilde{\gamma} > 0$, the desired equality trivially follows from the definition of $\dot{\mathcal{G}}(\cdot)$. On the other hand, applying L'Hôpital's rule iterative shows that $\mathrm{plim}_{\gamma\downarrow 0}\dot{\mathcal{G}}(\gamma) = \frac{\sqrt{3}}{2}z_1 + \frac{1}{2}z_2 \sim N(0,1)$, so that if we let $\dot{\mathcal{G}}_2 := \lim_{\gamma\downarrow 0}\dot{\mathcal{G}}(\gamma)$, then for $\gamma \neq 0$, $\mathbb{E}[\dot{\mathcal{G}}(\gamma)\dot{\mathcal{G}}_2] = \left[\sqrt{3}b_1(\gamma) + b_2(\gamma)\right]/\{2\ddot{k}_1(\gamma,\gamma)^{1/2}\}$. We show in the proof of Lemma 7 that the term on the right side is identical to $\ddot{\rho}_3(\gamma)$ in (12), so that the covariance kernel of $\dot{\mathcal{G}}(\cdot)$ is identical to $\ddot{\rho}(\cdot,\cdot)$. This fact implies that $\ddot{\mathcal{G}}(\cdot)$ has the same distribution as $\dot{\mathcal{G}}(\cdot)$, and $\dot{\mathcal{G}}_2$ can be regarded as the weak limit obtained under $\mathcal{H}_{02}'$.

Lemma 7 can be used to obtain the approximate null limit distribution of the QLR test statistic. We cannot generate $\dot{\mathcal{G}}(\cdot)$ using the infinite number of $b_n(\cdot)$, but we can simulate the following process to approximate the distribution of $\dot{\mathcal{G}}(\cdot)$:

$$\dot{\mathcal{G}}(\gamma; K) := \left(\sum_{n=1}^{K} b_{K,n}(\gamma)^2\right)^{-1/2}\sum_{n=1}^{K} b_{K,n}(\gamma) z_k,$$

where for $n = 2, 3, \ldots,$

$$b_{K,1}(\gamma) := \frac{1}{\sqrt{2}}(1 - 2a_K(\gamma)), \quad a_K(\gamma) := \sum_{k=1}^{K}\frac{(-1)^{k-1}}{(1+\gamma k)^3} \quad \text{and} \quad b_{K,n}(\gamma) := \sqrt{n(n+1)}\sum_{k=1}^{K}\frac{(-1)^{k-1}(\gamma k)^{n-1}}{(1+\gamma k)^3}.$$

If K is sufficiently large, the distribution of $\dot{\mathcal{G}}(\cdot; K)$ is close to that of $\dot{\mathcal{G}}(\cdot)$ as can be easily affirmed by simulations.

We conduct Monte Carlo Simulations for the LSTAR case as in the ESTAR case. The only aspect different from the ESTAR case is that the DGP is the one defined in the beginning of this section. Simulation results are summarized into Figure 2. We use the same parameter spaces $\Gamma = \Gamma_i$, $i = 1, \ldots, 4$, as before, and we can see that the empirical distribution and PDF estimate of the QLR test approach the null limit distribution and its PDF that are obtained using $\dot{\mathcal{G}}(\cdot; K)$ with $K = 2,500$. This shows that the theory in Section 2 is also valid for the LSTAR model. When the parameter space Γ for γ becomes even larger, we obtain similar results. To save space, they are not reported.

3.3 Application of the Weighted Bootstrap

The standard approach to obtaining the null limit distribution of the QLR test is not applicable for empirical analysis because it requires knowledge of the error distribution. Without this information it is not possible in practice to obtain a distributionally equivalent Gaussian process. Hansen's (1996) weighted bootstrap is useful for this case. We apply it to our models as in Cho and White (2010), Cho, Ishida, and White (2011, 2014), and Cho, Cheong, and White (2011).

Although the relevant weighted bootstrap is available in Cho, Cheong, and White (2011), we provide here a version adapted to the structure of the STAR model. We consider the previously studied ESTAR and LSTAR models and proceed as follows. First, we compute the following score for each grid point of $\gamma \in \Gamma$:

$$\widetilde{W}_n(\gamma) := \frac{1}{n}\sum_{t=1}^{n} \widetilde{u}_{n,t}^2 f_t^2(\gamma) z_t z_t' - \frac{1}{n}\sum_{t=1}^{n} \widetilde{u}_{n,t}^2 f_t(\gamma) z_t z_t' \left[\frac{1}{n}\sum_{t=1}^{n} \widetilde{u}_{n,t}^2 z_t z_t'\right]^{-1} \frac{1}{n}\sum_{t=1}^{n} \widetilde{u}_{n,t}^2 f_t(\gamma) z_t z_t',$$

$$\widetilde{d}_{n,t}(\gamma) := z_t f_t(\gamma)\widetilde{u}_{n,t} - \frac{1}{n}\sum_{t=1}^{n} \widetilde{u}_{n,t}^2 f_t(\gamma) z_t z_t' \left[\frac{1}{n}\sum_{t=1}^{n} \widetilde{u}_{n,t}^2 z_t z_t'\right]^{-1} z_t \widetilde{u}_{n,t},$$

where $\widetilde{u}_{n,t} := y_t - y_{t-1}\widetilde{\theta}_n$, and $\widetilde{\theta}_n$ is the least squares estimator of θ from the null model. Here, $f_t(\gamma) = 1 - \exp(-\gamma y_{t-1}^2)$ for ESTAR and $f_t(\gamma) = \{1 + \exp(\gamma y_{t-1})\}^{-1} - 1/2$ for the LSTAR model. Second, given these functions, we construct the following score function and pseudo-QLR test statistic:

$$\overline{QLR}_{b,n} := \sup_{\gamma\in\Gamma} \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \widetilde{s}_{n,t}(\gamma)\right)' \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \widetilde{s}_{n,t}(\gamma)\right) \text{ and } \widetilde{s}_{n,t}(\gamma) := \{\widetilde{W}_n(\gamma)\}^{-1/2}\widetilde{d}_{n,t}(\gamma) z_{b,t},$$

where $z_{b,t} \sim$ IID $(0, 1)$ with respect to b and t, $b = 1, 2, \ldots, B$, and B is the number of bootstrap replications. For example, we can resample $z_{b,t}$ from the standard normal distribution. For possible two-point distributions, see Davidson et al. (2007). Third, we estimate the empirical p-value by $\widehat{p}_n := B^{-1}\sum_{b=1}^{B} \mathbb{I}[QLR_n < \overline{QLR}_{b,n}]$, where $\mathbb{I}[\cdot]$ is the indicator function. We set $B = 300$ to obtain $\widehat{p}_n^{(i)}$ with $i = 1, 2, \ldots, 2,000$. Finally, for a specified nominal value of α, we compute $\frac{1}{2000}\sum_{i=1}^{2000} \mathbb{I}[\widehat{p}_n^{(i)} < \alpha]$. When the null hypothesis holds, this proportion converges to α.

The intuition of the weighted bootstrap is straightforward. Note that if the null hypothesis is valid, the QLR test statistic is bounded in probability, and its null limit distribution can be revealed by the covariance structure of $\widetilde{s}_{n,t}(\cdot)$ asymptotically. That is, for each γ and $\widetilde{\gamma}$, $\mathbb{E}[\widetilde{s}_{n,t}(\gamma)\widetilde{s}_{n,t}(\widetilde{\gamma})']$ converges to $\mathbb{E}[\mathcal{G}_1(\gamma)\mathcal{G}_1(\widetilde{\gamma})']$ from the fact that $z_{b,t}$ is independent of $\widetilde{d}_{n,t}(\cdot)$ such that its population mean is zero and variance is unity. This means that the $\mathbb{E}[\widetilde{s}_{n,t}(\gamma)\widetilde{s}_{n,t}(\widetilde{\gamma})']$ is asymptotically equivalent to $\{\widetilde{W}_n(\gamma)\}^{-1/2}\mathbb{E}[\widetilde{d}_{n,t}(\gamma)\widetilde{d}_{n,t}(\widetilde{\gamma})']\{\widetilde{W}_n(\widetilde{\gamma})\}^{-1/2}$, which is asymptotically equivalent to $\mathbb{E}[\mathcal{G}_1(\gamma)\mathcal{G}_1(\widetilde{\gamma})']$ as given in Theorem 3. Therefore, the null limit distribution can be asymptotically revealed by the resampling distribution of $\overline{QLR}_{b,n}$. On the contrary, if the alternative hypothesis is valid, the QLR test statistic is not bounded in probability, but $\overline{QLR}_{b,n}$ is bounded in probability from the fact that $z_{b,t}$ is distributed around zero, so that the chance for the QLR test statistic to be bounded by the critical value obtained by the resampling distribution of $\overline{QLR}_{b,n}$ gets smaller, as n increases. This aspect implies that if $n^{-1/2}\sum_{t=1}^{n}\widetilde{d}_{n,t}(\cdot)$ weakly converges to $\mathcal{G}_1(\cdot)$ and $\widetilde{W}_n(\cdot)$ uniformly converges to $A_1(\cdot)$ under the null, the weighted bootstrap is asymptotically valid.

The results are displayed in the percentile-percentile (PP) plots for the ESTAR and LSTAR models in Figures 3 (ESTAR) and 4 (LSTAR). The horizonal unit interval stands for α, and the vertical unit interval is the space of p-values. As a function of α, the aforementioned proportion should converge to the 45-degree line under the null hypothesis. As before, the four parameter spaces are considered: $\Gamma = \Gamma_i$, $i = 1, ..., 4$. The results are summarized as follows. First, as a function of α, the proportion $\frac{1}{2000}\sum_{i=1}^{2000}\mathbb{I}[\widehat{p}_n^{(i)} < \alpha]$ does converge to the 45-degree line. Second, the empirical rejection rates estimated by the weighted bootstrap are closest to the nominal levels when the parameter space is small. Although the overall finite sample level distortions are smaller for the ESTAR model than the LSTAR model, the empirical rejection rate is close to the nominal significance level if α is close to zero. Finally, as the size of the parameter space increases, more observations are needed to better approximate the 45-degree line in the PP plots. We have conducted simulations using even larger parameter spaces and obtained similar results. We omit reporting them for brevity.

4 Empirical Application

In this section, we illustrate use of the QLR test statistic. The purpose of this section is twofold. We first aim to demonstrate how to exploit the QLR test statistic for empirical economic analysis, and for this purpose, we examine a proper inference on fiscal policy effect to other macroeconomic variables. The next purpose is in comparing the results using the QLR test statistic with those from the LM test statistic proposed by Luukkonen, Saikkonen, and Teräsvirta (1988) and applied by Granger and Teräsvirta (1993) and Teräsvirta (1994), among others. Through this comparison, we desire to demonstrate that the QLR and LM tests can complement each other when the p-values of the QLR and LM tests are computed by the weighted bootstrap and the methodology for the F-test statistic in Teräsvirta (1994), respectively. Two empirical examples are provided. The first example is devoted to the first goal, and the last example is devoted to the second goal.

4.1 Example 1: State-Dependent Multiplier Effect of Government Spending

As our first example, we demonstrate use of the QLR test statistic to draw a proper inference on fiscal policy effects on other macroeconomic variables. For this purpose, we revisit the empirical study in AG that examines the multiplier effect of government spending. We specifically investigate potential nonlinearity of government spending, *i.e.*, how the aggregate output and its components respond nonlinearly to the change of government spending.

Nonlinearities in the effects of government spending have recently received increasing attention in the empirical literature. For example, Ramey (2011a) and Parker (2011) describe the debates on the state-dependent effect of the fiscal multiplier. Specifically, the literature on the new Keynesian empirical models generally agrees on a stable fiscal multiplier effect, although its size and sign depend on the model assumptions imposed under the new Keynesian framework. For example, Ramey (2011a), Coenen et al. (2012) and Ramey and Zubairy (2018), among others, provide empirical evidence on the stable fiscal multiplier effect. Christiano, Eichenbaum, and Rebelo (2011) provide theoretical and empirical evidence of stable but larger fiscal multiplier effects when the nominal interest rate is close to the zero lower bound. In contrast to these studies, the literature on traditional Keynesian models argues that the fiscal multiplier has economic state-dependent features. According to the traditional Keynesian model framework, for example, a positive shock of government spending is expected to result in direct employment of many idle resources, leading to positive responses in the aggregate output and its components. Hence, the multiplier effect of government spending tends to be larger than normal during recession, which implies that the effect is state-dependent. This makes it different from the new Keynesian model framework in which the multiplier effects of government spending on other macroeconomic variables remain stable.

The smooth transition model is typically employed to estimate the state-dependent multiplier effect. AG empirically apply it in order to capture the behaviour of the multiplier effect, influencing other studies with similar motivations (*e.g.*, Fazzari, Morley, and Panovska, 2014; Owyang, Ramey, and Zubairy, 2013a; and Auerbach and Gorodnichenko, 2012a). In particular, AG capture the nonlinear transition effect by employing the LSTAR model, and this idea has been followed in the other studies. STAR models or the threshold autoregressive model have then used as alternative models.

Notwithstanding its empirical popularity and significant policy implications, the STAR model in the literature could be applied in a more systematic fashion. Note that the nonlinearity plays an integral role in AG when the authors study state-dependent response of the aggregate output to the government spending. This suggests that the goal of the aforementioned studies can be achieved by considering a potential nonlinear relationship between the variables using proper statistical inference. However, in the literature simply the coefficients estimated from the linear and nonlinear models are compared with each other and differences in them are used as a means of validating the state-dependent multiplier effect.

In this section, we instead apply our QLR test statistic to study the nonlinear assumption adopted by AG and explore nonlinearity implied by the model. Note that if linearity is rejected by the QLR test statistic, this does not necessarily imply that the nonlinearity as defined by the model correctly describes the state-dependent fiscal multiplier effect. In fact, our test statistic is a diagnostic tool, so that even if linearity is rejected against the STAR model, the latter may not correctly describe the time-varying structure of the government spending multiplier. For this reason we also apply other test statistics than the QLR test and study validity of the model given the assumptions made by AG. In case the STAR model adequately describes the nonlinearity in the data, we exploit the model estimates to draw conclusions about the state-dependent multiplier effect on the macroeconomic variables of our interest.

The plan of our empirical analysis is as follows. We first note that AG introduce the following VSTAR model:

$$z_t = (1 - f_t(\gamma_*))\Pi_{E*}(L)z_{t-1} + f_t(\gamma_*)\Pi_{R*}(L)z_{t-1} + u_t \tag{14}$$

such that $u_t|(z_t, z_{t-1}, \ldots) \sim N(0, \Omega_t)$, where for some Ω_{E*} and Ω_{R*}, $\Omega_t := (1 - f_t(\gamma_*))\Omega_{E*} + f_t(\gamma_*)\Omega_{R*}$, and $z_t := (g_t, \tau_t, y_t)'$ with g_t, τ_t, and y_t being log real government spending, log real government net tax receipts, and log real gross domestic product (GDP) deflated by 2012 GDP deflator, respectively. Here, $\Pi_{E*}(L)$ and $\Pi_{R*}(L)$ are lag operators employed to capture the lag coefficients under economic expansion and recession periods, respectively. AG use the logistic function for the transition function $f_t(\gamma)$ by letting the transition variable be a seven-quarter moving average of the output growth rate centered at the sample average 0.80. The VSTAR model is motivated from the VAR model in Blanchard and Perotti (2002): $z_t = \Pi_*(L)z_{t-1} + u_t$ by supposing the two lag operators $\Pi_{E*}(L)$ and $\Pi_{R*}(L)$ under the expansive and recessive economic periods.

Despite the appealing features of the STAR model detailed in AG, nonstationarity of their time series makes it difficult to estimate the associated parameters by the standard approach. The maximum likelihood estimator may often represent a local optimum, which makes estimation of the parameters of the VSTAR model more complicated than it would be if the variables were stationary. In addition, their estimation method is technically complex as it employs the Markov-Chain Monte Carlo method.

We avoid this difficulty by converting the model in (14) into a vector error correction (VEC) form. Doing this to the VAR form in Blanchard and Perotti (2002) yields the following model:

$$\Delta z_t = \Phi_* z_{t-1} + \widetilde{\Pi}_*(L)\Delta z_{t-1} + u_t. \tag{15}$$

The first term on the right-hand side of (15) captures the long-run relationship between the variables. It is assumed invariant to the economic states, which means that we may rewrite (15) as $\Delta z_t = \alpha_* \beta_*' z_{t-1} + \widetilde{\Pi}_*(L)\Delta z_{t-1} + u_t$, where α_* and β_*' are $3 \times m$ and $m \times 3$ matrices, respectively, and m is the cointegration rank. This conversion allows us to rewrite (14) into the following vector smooth-transition error-correction (VSTEC) model:

$$\Delta z_t = (1 - f_t(\gamma_*))[\alpha_{E*} z_{t-1}^* + \widetilde{\Pi}_{E*}(L)\Delta z_{t-1}] + f_t(\gamma_*)[\alpha_{R*} z_{t-1}^* + \widetilde{\Pi}_{R*}(L)\Delta z_{t-1}] + u_t \tag{16}$$

by letting $z_{t-1}^* := \beta_*' z_{t-1}$, which is an $m \times 1$ vector of stationary variables (*e.g.*, Rothman, van Dijk, and Franses, 2001; Hubrich and Teräsvirta, 2013). Note that the cointegration vector β_* is invariant to the economic state, although we allow α_* to have different values for different states. In addition to this framework, we also modify the assumption of conditional heteroskedasticity made by AG by introducing the error covariance matrix $\Omega_t = D_t P D_t$, where D_t is the diagonal matrix of conditionally heteroskedastic standard deviations, and P is the corresponding correlation matrix assumed constant over time, see Eklund and Teräsvirta (2007). This decomposition allows us to test homoskedasticity before assuming heteroskedastic errors. We specifically assume that, under the alternative, $D_t := (1 - f_t(\gamma_*))D_{E*} + f_t(\gamma_*)D_{R*}$, where D_{E*} and D_{R*} have diagonal elements σ^2_{i*} and $\sigma^2_{i*} + \lambda_{i*}$, respectively for $i = 1, 2$, and 3, so that D_t becomes a diagonal matrix with $\sigma^2_{i*} + \lambda_{i*} f_t(\gamma_*)$. Under the null hypothesis of conditional homoskedasticity, $D_{E*} = D_{R*}$, and we test this hypothesis by the LM test statistic provided by Eklund and Teräsvirta (2007). Admittedly, there is no theoretical reason for believing that correlations remain constant over time while variances exhibits conditional heteroskedasticity. Nevertheless, we believe that the conditional error variances are likely to dominate the conditional heteroskedastic correlations for many empirical data exhibiting conditional heteroskedasticity. Using this belief and the fact that the number of parameters of here is fewer than that in AG, we test for conditional heteroskedasticity by employing this specification test.

To the best of our knowledge, VSTEC model has been rarely applied in macroeconomic studies on the effects of fiscal policy, although it has been popular in applications to macroeconomic models investigating monetary and financial policies (*e.g.*, Rothman, van Dijk, and Franses, 2001 among others). This aspect in another sense motivates to compare the empirical outputs obtained by the VSTAR and VSTEC models.

Using the data analyzed by AG, we estimate the parameters in (16).[2] Their data consist of US quarterly macroeconomic variables and the sample runs from 1947Q1 to 2008Q4. In order to estimate the parameters of the VEC model, we first specify the cointegration rank m and estimate β_* following Johansen (1995). As the maximum likelihood estimator for β_*, denoted by $\widehat{\beta}_n$, is super-consistent, we simply replace z^*_{t-1} in (16) with $\widehat{z}_{t-1} := \widehat{\beta}'_n z_{t-1}$ and estimate the other parameters by nonlinear least squares.

Second, we test linearity against the VSTAR model using the QLR test statistic and employ other diagnostic test statistics to assess the adequacy of the model. In order to do this we modify the VSTEC model in (16) such that testing it becomes possible. Note that (16) can be rewritten as

$$\Delta z_t = \Psi_{1*}(L) w_{t-1} + f_t(\gamma_*)\Psi_{2*}(L) w_{t-1} + u_t,$$

where $u_t = (u_{1t}, u_{2t}, u_{3t})'$, $w_{t-1} := [z^{*\prime}_{t-1}, \Delta z'_{t-1}]'$, $\Psi_{1*}(L) := [\alpha_{R*}, \widetilde{\Pi}_{R*}(L)]$, and $\Psi_{2*}(L) := [\alpha_{D*}, \widetilde{\Pi}_{D*}(L)]$ with $\alpha_{D*} := \alpha_{R*} - \alpha_{E*}$ and $\widetilde{\Pi}_{D*}(L) := \widetilde{\Pi}_{R*}(L) - \widetilde{\Pi}_{E*}(L)$. Given the normality assumption for u_t, each equation can be marginalized as follows: for $j = 1, 2, 3$,

$$\Delta z_{jt} = \theta'_j \Delta z_{-jt} + \xi_{j1*}(L)' w_{t-1} + f_t(\gamma_*)\xi_{j2*}(L)' w_{t-1} + \epsilon_{jt}, \tag{17}$$

where for $j = 1, 2, 3$, $\theta'_{j*} := \mathbb{E}[u_{jt} u_{-jt}]\mathbb{E}[u_{-jt} u'^{-1}_{-jt}]$, and, further,

$$\xi_{j1*}(L)' := \psi_{j1*}(L)' - \theta'_{j*}\psi_{-j1*}(L)' \quad \text{and} \quad \xi_{j2*}(L)' := \psi_{j2*}(L)' - \theta'_{j*}\psi_{-j2*}(L)'.$$

[2] When estimating this model, we also include the intercept term.

Here, u_{jt} and u_{-jt} denote the j-th row element of u_t and the 2×1 vector obtained by removing u_{jt} from u_t, respectively. Furthermore, for each $i = 1$ and 2, $\psi_{ji*}(L)'$ and $\psi_{-ji*}(L)'$ are the j-th row vector of $\Psi_{i*}(L)$ and $2 \times (m+3)$ matrix obtained by removing the j-th row from $\Psi_{i*}(L)$, respectively. We use (17) as our baseline model for testing for nonlinearity, where we replace w_{t-1} in (17) with $\widehat{w}_{t-1} := [\widehat{z}'_{t-1}, \Delta z'_{t-1}]'$ to again exploit the super-consistency of $\widehat{\beta}_n$. We apply the QLR test statistic for each $j = 1, 2, 3$. In addition to the QLR test statistic, we test for serial correlation and conditional heteroskedasticity in the errors. Note that serially uncorrelated errors are a necessary condition for the adequacy of the model. If the errors are serially correlated, the nonlinear model is misspecified in the sense that the dynamic structure of the model is incorrect. By testing for serial correlation in the error using the multivariate Ljung-Box test statistic, we aim at detecting model misspecification. In addition, as detailed above, Eklund and Teräsvirta's (2007) test statistic is employed to test for conditional heteroskedasticity. If it turns out to be conditionally homoskedastic, we do not have to specify a parametric model for error covariances as AG do.

Finally, based on the estimated nonlinearity associated with z_t, we examine the impact of government spending by separately estimating the impulse response functions of the variables in z_t under the expansion and recession states. This is done in order to compare our empirical findings based upon the VSTEC model with those given by AG. They estimate the corresponding impulse response functions by estimating the VSTAR model directly. In order to generate the impulse response functions, we use the VAR structure given in Blanchard and Perotti (2002) by ordering the variables in z_t as above and relating it to the Cholesky decomposition. AG also adopted this ordering. More specifically, after estimating the expansion and recession regime parameters in (16), we convert them into the VAR parameters by applying the Cholesky decomposition. Using this structure we generate the impulse response functions of the variables in z_t under expansion and recession states in response to a \$1 increase of government spending. The results are compared with the corresponding ones in AG.

We now report our empirical findings. In order to estimate the cointegration rank m, we apply Johansen's (1988, 1991) trace testing procedure with lag equal to 3 selected by AIC and BIC[3] and cannot reject the hypothesis that $m = 2$ at the 5% significance level. Using this rank, we estimate the cointegration coefficient β_* to obtain $\widehat{z}_{t-1}$.

Next, we apply several diagnostic testing procedures to validate our model assumptions. As the first diagnostic test statistic, we apply the LM test statistic in Eklund and Teräsvirta (2007) to test for conditional heteroskedasticity. The p-value of the test can be found in Table 1, indicating that the null of conditional homoskedasticity is not rejected at the 5% level of significance.[4] Next we test for serial correlation in the errors using the multivariate Ljung-Box test statistic and fail to reject the null of no serial correlation of the errors at 5% level of significance as reported in Table 1. Our final diagnostic test is for the nonlinearity assumption made for the VSTEC model. As explained above, we replace w_{t-1} in (17) with $\widehat{w}_{t-1} := [\widehat{z}'_{t-1}, \Delta z'_{t-1}]'$ and test the linearity hypothesis by the QLR test statistic based upon LSTAR model. We report the p-values of the QLR test statistics in Table 1. They are obtained by applying the

[3] The lag order is also identical to that selected by AG.

[4] In addition to Eklund and Teräsvirta's (2007) LM test statistic, we also applied Breusch-Pagan's conditional heteroskedasticity and multivariate Ljung-Box test statistics to detect conditional heteroskedasticity but failed to detect. Testing results are available from the authors upon request.

weighted bootstrap with 20,000 replications, and they strongly reject the linearity hypothesis. This implies that the linear VEC model is not adequate and that the VSTEC model can better capture the dynamic interrelationship among the variables.

We compare our model estimates with those obtained by AG. After converting the VSTEC parameters in expansion and recession states into VAR parameters by applying the Cholesky decomposition, we estimate the impulse response functions of the three variables in response to a \$1 increase of government spending. The results appear in Figure 5. The impulse responses are plotted following the figure structure in AG. That is, we show how g_t, τ_t, and y_t respond to the government spending over 10 quarters under expansion and recession states and depict their 90% confidence bands obtained from 50,000 Monte Carlo replications. In addition, we plot the impulse response functions obtained from the linear VEC model. The impulse response functions of g_t, τ_t, and y_t are shown from the top, to the middle and bottom rows in the figure, respectively, and they correspond to those in Figure 2 of AG.

Our empirical findings are as follows. First, the overall shapes of the functions agree with those in AG. Specifically, the response of government purchases under the expansion state achieves its maximum after a short delay, as argued in AG. Furthermore, the immediate effect of the increase in government spending on y_t is less than unity, irrespective of the economic states. The estimated immediate effect is about 0.114 for both expansion and recession states, which is slightly greater than that obtained from the linear VEC model, 0.106. This also agrees with AG. Second, the details of the impulse response functions, nevertheless, are not exactly the same as those in AG. The bottom row of Figure 5 shows the divergent response of y_t to the government spending in both recession and expansion states. This also agrees with AG, but our maximum government spending multiplier is not as large as theirs. It reaches only around 0.2 in the recession state, which is about one tenth of that in AG.

As to the numerical results on the impulse response functions, we report the estimated spending multipliers in Table 2. They are estimated following AG. That is, we estimate them by the maximum output after increasing the government spending, $\max_{h=1,\ldots,H} Y_h$, and the sum of outputs relative to the sum of government purchases over H quarters after increasing the government spending, $\sum_{h=1}^{H} Y_h / \sum_{h=1}^{H} G_h$, where we let H be 8 and 20 quarters. To do this, we convert the estimated percentage changes of government spending and output into their dollar equivalents over the same sample period and report the point estimates along with standard errors. Table 2 reports the estimates in detail, and the figures in parentheses are their standard errors. Compared to the estimated multipliers in Table 1 of AG, most of our estimates have larger standard errors, especially under the expansion and recession states, and thus we cannot rule out the possibility that the fiscal multipliers are not different from zero or that they are of smaller magnitude than those in AG if they are nonzero. The large multiplier effect estimated by AG for the recession state has been doubted in a number of studies (*e.g.*, Owyang et al., 2013a; Ramey and Zubairy, 2018), whereas our model estimates are free from this criticism. Specifically, our fiscal multiplier estimates for the shorter horizon are larger than those for the long horizon, which is in line with Owyang et al. (2013a) and Ramey and Zubariy (2018), and exhibiting smaller standard errors for the short than the long horizon. This also suggests that the multiplier effect does not last long.

4.2 Example 2: US Unemployment Rates

We next compare the QLR and LM test statistics. We desire to demonstrate specifically how these two different test statistics can supplement each other through the empirical examples in this and next subsections.

Before examining the empirical examples, we briefly review the model framework for the LM test statistics. The following auxiliary model is first estimated for the LM test statistics:

$$\mathcal{M}_{AUX} := \{h_{AUX}(\,\cdot\,;\alpha_0,\alpha_1,\alpha_2,\alpha_3,\alpha_4) : (\alpha_0',\alpha_1',\alpha_1',\alpha_1',\alpha_1')' \in \Theta\},$$

where $h_{AUX}(z_t;\alpha_0,\alpha_1,\alpha_2,\alpha_3,\alpha_4) := \alpha_0' z_t + \alpha_1'(\widetilde{z}_t t_t) + \alpha_2'(\widetilde{z}_t t_t^2) + \alpha_3'(\widetilde{z}_t t_t^3) + \alpha_4'(\widetilde{z}_t t_t^4)$, and t_t is the transition variable, viz., $\widetilde{z}_t'\alpha$. This auxiliary model is obtained by applying a fourth-order Taylor expansion to the analytic function as an intermediate step to compute the LM test statistics conveniently. Although this auxiliary model is different from the STAR model, testing the coefficients of nonlinear components by the LM test statistics turns out to be equivalent to computing the LM test statistics that test the STAR model assumption under $\mathcal{H}_{02}$. Luukkonen, Saikkonen, and Teräsvirta (1988) and Teräsvirta (1994) provide detailed rationales of this equivalence.

The following four sets of hypotheses are considered as common hypotheses of the three empirical examples:

$$\mathcal{H}_{0,1} : \alpha_{1*} = \alpha_{2*} = \alpha_{3*} = 0 | \alpha_{4*} = 0; \text{ vs. } \mathcal{H}_{1,1} : \alpha_{1*} \neq 0, \alpha_{2*} \neq 0, \text{ or } \alpha_{3*} \neq 0 | \alpha_{4*} = 0.$$

$$\mathcal{H}_{0,2} : \alpha_{1*} = \alpha_{2*} = \alpha_{3*} = \alpha_{4*} = 0; \text{ vs. } \mathcal{H}_{1,2} : \alpha_{1*} \neq 0, \alpha_{2*} \neq 0, \alpha_{3*} \neq 0, \text{ or } \alpha_{4*} \neq 0.$$

$$\mathcal{H}_{0,3} : \alpha_{1*} = \alpha_{3*} = 0; \text{ vs. } \mathcal{H}_{1,3} : \alpha_{1*} \neq 0 \text{ or } \alpha_{3*} \neq 0.$$

$$\mathcal{H}_{0,4} : \alpha_{2*} = \alpha_{4*} = 0; \text{ vs. } \mathcal{H}_{1,4} : \alpha_{2*} \neq 0 \text{ or } \alpha_{4*} \neq 0.$$

These hypotheses are specified by following Teräsvirta (1994) and Escribano and Jordà (1999). We denote the LM test statistics testing $\mathcal{H}_{0,i}$ as $LM_{i,n}$, $i = 1, ..., 4$. $LM_{1,n}$ and $LM_{2,n}$ are general tests against STAR. On the other hand, $LM_{3,n}$ and $LM_{4,n}$ are tests against the LSTAR and ESTAR models, respectively. The QLR statistic against ESTAR is denoted by QLR_n^E the one against LSTAR is called QLR_n^L.

We now consider the performance of the tests when applied to the monthly US unemployment rate. van Dijk, Teräsvirta, and Franses (2002) tested linearity of this series using observations from June 1968 to December 1999. We use the same dataset and compare their LM test statistics with our QLR test statistics. We also extend the series to August 2015 and perform the same tests.[5]

van Dijk, Teräsvirta, and Franses (2002) point out that the US unemployment rate is a persistent series with an

[5]The data set used by van Dijk, Teräsvirta, and Franses (2002) is available at `<http://swopec.hhs.se/hastef/abs/hastef0380.htm>` that was originally retrieved from the Bureau of Labor Statistics.

asymmetric adjustment process and strong seasonality. They specify a STAR model with monthly dummy variables mainly because first differences of the seasonally unadjusted unemployment rate of males aged 20 and over is used for Δy_t. They test linearity against STAR assuming that the transition variable is a lagged twelve-month difference of the unemployment rate. The alternative (STAR) model has the following form (the lag length has been determined by AIC):

$$\begin{aligned}\Delta y_t = \pi_0 + \pi_1 y_{t-1} + \sum_{p=1}^{15} \pi_{p+2} \Delta y_{t-p} + \sum_{k=1}^{11} \pi_{17+k} d_{t,k} \\ + \left[\theta_0 + \theta_1 y_{t-1} + \sum_{p=1}^{15} \theta_{p+2} \Delta y_{t-p} + \sum_{k=1}^{11} \theta_{17+k} d_{t,k} \right] f(\Delta_{12} y_{t-d}; \gamma) + u_t,\end{aligned}$$

where y_t is the monthly US unemployment rate in question, Δy_t is the first difference of y_t, $f(\cdot, \cdot)$ is a nonlinear transition function, $\Delta_{12} y_t$ is the twelve-month difference of y_t, $d_{t,k}$ is the dummy for month k, and $u_t \sim$ IID $(0, \sigma^2)$. The twelve-month difference $\Delta_{12} y_{t-d}$ is not included as an explanatory variable in the null (linear) model. The theory in Section 2 can nonetheless be used without modification because a null model including $\Delta_{12} y_{t-d}$ can be thought of having a zero coefficient for this variable. Following van Dijk, Teräsvirta, and Franses (2002), we test linearity by using $\Delta_{12} y_{t-d}$, $d = 1, 2, \ldots, 6$, as the transition variable.

Our test results using the same series as van Dijk, Teräsvirta, and Franses (2002) can be found in the top panel of Table 3. Both the LM tests and QLR_n^L reject linearity when $d = 2$, and, besides, $LM_{3,n}$ that has power against LSTAR yields $p = 0.037$ for $d = 2$. The p-values of QLR_n^L, however, lie at or below 0.05 for all six lags, suggesting that at least in this particular case this QLR test is more powerful than the LM tests. The smallest p-value is even here attained for $d = 2$. The results from QLR_n^E are quite different and reject the null for $d = 1$ and 2, but not for other lags. This makes sense as this statistic is designed for ESTAR alternatives, and asymmetry in the unemployment rate is best described by an LSTAR model.

The bottom panel of Table 3 contains the results from the series extended to August 2015.[6] Now there seems to be plenty of evidence of asymmetry: all p-values of $LM_{1,n}$ are rather small. $LM_{3,n}$ also has small values for the first three lags, as has $LM_{2,n}$. The p-values from QLR_n^L are smallest of all, which is in line with the results in the top panel. Even QLR_n^E rejects the null of linearity at the 5% level for $d = 1, 2, 3, 4, 5$. This outcome may be expected because the QLR statistics are omnibus tests and as such respond to any deviation from the null hypothesis as the sample size increases. Note, however, that even $LM_{4,n}$ now yields two p-values $(d = 2, 3)$ that lie below 0.05, although the test does not have the omnibus property. The behaviour of the unemployment rate during and after the financial crisis (a quick upswing and slow decrease) has probably contributed to theses results.

[6]The recent observations of the monthly US unemployment rate are available at <http://beta.bls.gov/dataViewer/view/timeseries/LNU04000025>.

5 Conclusion

The current study examines the null limit distribution of the QLR test statistic for neglected nonlinearity using the STAR model. The QLR test statistic contains a twofold identification problem under the null, and we explicitly examine how the twofold identification problem affects the null limit distribution of the QLR test statistic. We show that the QLR test statistic is shown to converge to a functional of a multidimensional Gaussian process under the null of linearity by extending the testing scope of the LM test statistic in Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994) and Granger and Teräsvirta (1993).

We further illustrate our theory on the QLR test statistic to ESTAR and LSTAR models and affirm our theory by obtaining the null limit critical values and conducting Monte Carlo simulations. Finally, three empirical examples are revisited to empirically demonstrate use of the QLR test statistic. We test for neglected nonlinearity in the multiplier effect of US government spending and the growth rates of US unemployment using the QLR test statistic by revisiting the empirical data examined by Auerbach and Gordonichenko (2012b) and van Dijk, Teräsvirta, and Franses (2002), respectively. Through these examinations, the QLR test statistic turns out useful for detecting the nonlinear structure among the economic variables and complements the linearity test of the Lagrange multiplier test statistic in Teräsvirta (1994).

6 Appendix

Proof of Lemma 1: (*i*) Given Assumptions 1, 2, 3, and 5, it is trivial to show that $\widehat{\sigma}^2_{n,0} \stackrel{\text{a.s.}}{\to} \sigma^2_*$ by the ergodic theorem.

(*ii*) The null limit distribution of $QLR_n^{(1)}$ is determined by the two terms in (2): $Z'F(\cdot)Mu$ and $Z'F(\cdot)MF(\cdot)Z$. We examine their null limit behaviour one by one and combine the limit results using the converging-together lemma in Billingsley (1999, p. 39).

(a) We show the weak convergence part of $n^{-1/2}Z'F(\cdot)Mu$. Using the definition of $M := I - Z(Z'Z)^{-1}Z'$ we have $Z'F(\gamma)Mu = Z'F(\gamma)u - Z'F(\gamma)Z(Z'Z)^{-1}Z'u$, and we now examine the components on the right-hand side of this equation separately. For each $\gamma \in \Gamma$, we define $\widehat{f}_{n,t}(\gamma) := f_t(\gamma)u_t z_t - (\sum_{t=1}^n f_t(\gamma)z_t z_t')\left(\sum_{t=1}^n z_t z_t'\right)^{-1}\sum_{t=1}^n z_t u_t$, $\widetilde{f}_{n,t}(\gamma) := f_t(\gamma)u_t z_t - \mathbb{E}[f_t(\gamma)z_t z_t']\mathbb{E}[z_t z_t']^{-1}\sum_{t=1}^n z_t u_t$ and show that

$$\sup_{\gamma\in\Gamma(\epsilon)} \left\| n^{-1/2}\sum_{t=1}^n \left[\widehat{f}_{n,t}(\gamma) - \widetilde{f}_{n,t}(\gamma)\right] \right\|_\infty = o_{\mathbb{P}(1)}, \tag{18}$$

where $\Gamma(\epsilon) := \{\gamma \in \Gamma : |\gamma| \geq \epsilon\}$ and $\|\cdot\|_\infty$ is the uniform matrix norm. We have

$$\begin{aligned}
\sup_{\gamma\in\Gamma(\epsilon)} &\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left[\widehat{f}_{n,t}(\gamma) - \widetilde{f}_{n,t}(\gamma) \right] \right\|_\infty \\
&\leq \sup_{\gamma\in\Gamma(\epsilon)} \left\| \left(\frac{1}{n} \sum_{t=1}^{n} f_t(\gamma) z_t z_t' \right) \left\{ \left(\frac{1}{n} \sum_{t=1}^{n} z_t z_t' \right)^{-1} - \mathbb{E}[z_t z_t']^{-1} \right\} n^{-1/2} \sum_{t=1}^{n} z_t u_t \right\|_\infty \\
&\quad + \sup_{\gamma\in\Gamma(\epsilon)} \left\| \left\{ \left(n^{-1} \sum_{t=1}^{n} f_t(\gamma) z_t z_t' \right) - \mathbb{E}[f_t(\gamma) z_t z_t'] \right\} \mathbb{E}[z_t z_t']^{-1} n^{-1/2} \sum_{t=1}^{n} z_t u_t \right\|_\infty . \quad (19)
\end{aligned}$$

We show that each term on the right-hand side of (19) is $o_{\mathbb{P}}(1)$. Now, $\{z_t u_t, \mathcal{F}_t\}$ is a martingale difference sequence, where $\mathcal{F}_t$ is the smallest sigma-field generated by $\{z_t u_t, z_{t-1} u_{t-1}, \ldots\}$. Therefore, $\mathbb{E}[z_t u_t | \mathcal{F}_{t-1}] = 0$, $\mathbb{E}[|Z_{t,j} u_t|^2] = \mathbb{E}[u_t^4]^{1/2} \mathbb{E}[|Z_{t,j}|^4]^{1/2} \leq \mathbb{E}[m_t^4]^{1/2} \mathbb{E}[Z_{t,j}^4]^{1/2} < \infty$, and $\mathbb{E}[u_t^2 z_t z_t']$ is positive definite. Thus, $n^{-1/2} \sum_{t=1}^{n} z_t u_t$ is asymptotically normal. Next, we note that $n^{-1/2} \sum_{t=1}^{n} f_t(\gamma) u_t z_t$ is also asymptotically normal. This follows from the fact that $\{f_t(\gamma) u_t z_t, \mathcal{F}_t\}$ is a martingale difference sequence, and for each j, $|f_t(\gamma) u_t z_{t,j}|^2 \leq m_t^6$, and $\mathbb{E}[m_t^6] < \infty$ by Assumptions 4 and 5. Furthermore, $\sup_{\gamma\in\Gamma} \left\| n^{-1} \sum_{t=1}^{n} f_t(\gamma) z_t z_t' - \mathbb{E}[f_t(\gamma) z_t z_t'] \right\|_\infty = o_{\mathbb{P}}(1)$ by Ranga Rao's (1962) uniform law of large numbers. Thus,

$$\sup_{\gamma\in\Gamma(\epsilon)} \left\| \left\{ n^{-1} \sum_{t=1}^{n} f_t(\gamma) z_t z_t' - \mathbb{E}[f_t(\gamma) z_t z_t'] \right\} \mathbb{E}[z_t z_t']^{-1} n^{-1/2} \sum_{t=1}^{n} z_t u_t \right\|_\infty = o_{\mathbb{P}}(1). \quad (20)$$

This shows that the second term of (19) is $o_{\mathbb{P}}(1)$. We now demonstrate that the first term of (19) is also $o_{\mathbb{P}}(1)$. By Assumption 4 and the ergodic theorem, we note that $\|n^{-1} \sum_{t=1}^{n} z_t z_t' - \mathbb{E}[z_t z_t']\|_\infty = o_{\mathbb{P}}(1)$, and $|\sum_{t=1}^{n} f_t(\gamma) z_{t,j} z_{t,i}| \leq \sum_{t=1}^{n} m_t^3 = O_{\mathbb{P}}(n)$, so that (20) follows, leading to (18). Therefore, $n^{-1/2} Z' F(\gamma) M u \overset{\text{A}}{\sim} N[0, B_1(\gamma,\gamma)]$ by noting that $\mathbb{E}[\widetilde{f}_{n,t}(\gamma) \widetilde{f}_{n,t}(\gamma)'] = B_1(\gamma,\gamma)$. Using the same methodology, we can show that for each $\gamma, \widetilde{\gamma} \in \Gamma(\epsilon)$,

$$\frac{1}{\sqrt{n}} \begin{bmatrix} Z' F(\gamma) M u \\ Z' F(\widetilde{\gamma}) M u \end{bmatrix} \overset{\text{A}}{\sim} N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} B_1(\gamma,\gamma) & B_1(\gamma,\widetilde{\gamma}) \\ B_1(\widetilde{\gamma},\gamma) & B_1(\widetilde{\gamma},\widetilde{\gamma}) \end{bmatrix} \right].$$

Finally, we have to show that $\{\widetilde{f}_{n,t}(\cdot)\}$ is tight. First note that by Assumptions 1, 2, and 4, it follows that $|f_t(\gamma) z_{t,j} u_t - f_t(\widetilde{\gamma}) z_{t,j} u_t| \leq m_t |z_{t,j} u_t| \; |\gamma - \widetilde{\gamma}|$ for each j. From this we obtain that $\sup_{|\gamma-\widetilde{\gamma}|<\eta} |f_t(\gamma) z_{t,j} u_t - f_t(\widetilde{\gamma}) z_{t,j} u_t|^{2+\omega} \leq m_t^{2+\omega} |z_{t,j} u_t|^{2+\omega} \eta^{2+\omega} \leq m_t^{6+3\omega} \eta^{2+\omega}$, so that $\mathbb{E}[\sup_{|\gamma-\widetilde{\gamma}|<\eta} |f_t(\gamma) z_{t,j} u_t - f_t(\widetilde{\gamma}) z_{t,j} u_t|^{2+\omega}]^{\frac{1}{2+\omega}} \leq \mathbb{E}[m_t^{6+3\omega}]^{\frac{1}{2+\omega}} \eta$ for each j. This implies that $\{n^{-1/2} f_t(\cdot) z_{t,j} u_t\}$ is tight because Ossiander's $L^{2+\omega}$ entropy is finite.

Next, for some $c > 0$, $\|\mathbb{E}[f_t(\gamma) z_t z_t'] \mathbb{E}[z_t \, z_t']^{-1} z_t u_t - \mathbb{E}[f_t(\widetilde{\gamma}) z_t z_t'] \mathbb{E}[z_t z_t']^{-1} z_t u_t\|_\infty = \|\mathbb{E}[\{f_t(\gamma) - f_t(\widetilde{\gamma})\} z_t z_t'] \mathbb{E}[z_t \, z_t']^{-1} z_t u_t\|_\infty \leq c m_t^2 \left\| \mathbb{E}[z_t z_t']^{-1} \right\|_\infty \|\mathbb{E}[\{f_t(\gamma) - f_t(\widetilde{\gamma})\} z_t z_t']\|_\infty$ by the property of the uniform norm and Assumption 5. Also note that $\|\mathbb{E}[f_t(\gamma) z_t z_t' - f_t(\widetilde{\gamma}) z_t z_t']\|_\infty \leq \|\mathbb{E}[\{f_t(\gamma) - f_t(\widetilde{\gamma})\} z_t z_t']\|_1$ and by Assumption 4, for each $i, j =$

$1, 2, \ldots, m+1$, $|z_{t,j} z_{t,i}[f_t(\gamma) - f_t(\widetilde{\gamma})]| \leq m_t^3 |\gamma - \widetilde{\gamma}|$, where $\|[g_{i,j}]\|_1 := \sum_i \sum_j |g_{i,j}|$. Therefore,

$$\begin{aligned}
&\left\|\mathbb{E}[f_t(\gamma) z_t z_t'] \mathbb{E}[z_t z_t']^{-1} z_t u_t - \mathbb{E}[f_t(\widetilde{\gamma}) z_t z_t'] \mathbb{E}[z_t z_t']^{-1} z_t u_t\right\|_\infty \\
&\leq c m_t^2 \left\|\mathbb{E}[z_t z_t']^{-1}\right\|_\infty \left\|\mathbb{E}[\{f_t(\gamma) - f_t(\widetilde{\gamma})\} z_t z_t']\right\|_\infty \leq c(m+1)^2 m_t^2 \left\|\mathbb{E}[z_t z_t']^{-1}\right\|_\infty \mathbb{E}[m_t^3]|\gamma - \widetilde{\gamma}|. \qquad (21)
\end{aligned}$$

This inequality (21) implies that $\{n^{-1/2}\mathbb{E}[f_t(\cdot) z_t z_t'] \mathbb{E}[z_t z_t']^{-1} z_t u_t\}$ is also tight. Hence, it follows that for some $b < \infty$, $\mathbb{E}[\sup_{|\gamma - \widetilde{\gamma}| < \eta} |\widetilde{f}_t(\gamma) - \widetilde{f}_t(\widetilde{\gamma})|^{2+\omega}] \leq b \cdot \eta$. That is, $\{n^{-1/2} \sum_{t=1}^n \widetilde{f}_{n,t}(\cdot)\}$ is tight. From this and the fact that the finite-dimensional multivariate CLT holds, the weak convergence of $\{n^{-1/2} \sum_{t=1}^n \widetilde{f}_{n,t}(\cdot)\}$ is established.

(b) Next, we examine the limit behaviour of $n^{-1} Z' F(\cdot) F(\cdot) Z$. For this purpose, we note that $n^{-1} Z' F(\gamma) F(\gamma) Z = \frac{1}{n} \sum_{t=1}^n f_t(\gamma)^2 z_t z_t' - \left\{n^{-1} \sum_{t=1}^n f_t(\gamma) z_t z_t'\right\} \left\{n^{-1} \sum_{t=1}^n z_t z_t'\right\}^{-1} \left\{n^{-1} \sum_{t=1}^n f_t(\gamma) z_t z_t'\right\}$ and, given Assumptions 1, 2, 3, 4, and 6, $\sup_{\gamma \in \Gamma(\epsilon)} \|n^{-1} \sum_{t=1}^n f_t(\gamma)^2 z_t z_t' - \mathbb{E}[f_t(\gamma)^2 z_t z_t']\| \overset{\text{a.s.}}{\to} 0$ and $\sup_{\gamma \in \Gamma(\epsilon)} \|n^{-1} \sum_{t=1}^n f_t(\gamma) z_t z_t' - \mathbb{E}[f_t(\gamma) z_t z_t']\| \overset{\text{a.s.}}{\to} 0$ by Ranga Rao's (1962) uniform law of large numbers. Therefore, given $\|n^{-1} \sum_{t=1}^n z_t z_t' - \mathbb{E}[z_t z_t']\|_\infty = o_{\mathbb{P}}(1)$, it follows that $\sup_{\gamma \in \Gamma(\epsilon)} |n^{-1} Z' F(\gamma) M F(\gamma) Z - \{\mathbb{E}[f_t(\gamma)^2 z_t z_t'] - \mathbb{E}[f_t(\gamma) z_t z_t'] \mathbb{E}[z_t z_t']^{-1} \mathbb{E}[f_t(\gamma) z_t z_t']\}| = o_{\mathbb{P}}(1)$. Applying the converging-together lemma yields the desired result.

(iii) This result trivially follows from the fact that $\mathbb{E}[u_t^2 | z_t] = \sigma_*^2$. ■

Proof of Lemma 2: Given Assumption 2, $\mathcal{H}_{02}$, and the definition of $H_j(\gamma)$, the j-th order derivative of $\mathcal{L}_n^{(2)}(\,\cdot\,;\theta)$ is obtained as

$$\begin{aligned}
\frac{\partial^j}{\partial \gamma^j} \mathcal{L}_n^{(2)}(\gamma; \theta) &= -\sum_{k=0}^{j} \binom{j}{k} \left\{\frac{\partial^k}{\partial \gamma^k}(y - F(\gamma) Z\theta)'\right\} M \left\{\frac{\partial^{j-k}}{\partial \gamma^{j-k}}(y - F(\gamma) Z\theta)\right\} \\
&= 2\theta' Z' H_j(\gamma) M u - \sum_{k=1}^{j-1} \binom{j}{k} \theta' Z' H_j(\gamma) M H_{j-k}(\gamma) Z\theta \qquad (22)
\end{aligned}$$

by iteratively applying the general Leibniz rule. We now evaluate this derivative at $\gamma = 0$. Note that $H_j(0) = 0$ if $j < \kappa$ by the definition of κ. This implies that $(\partial^j / \partial \gamma^j) \mathcal{L}_n^{(2)}(0; \theta) = 0$ for $j = 1, 2, \ldots, \kappa - 1$. This also implies that $\binom{j}{k} \theta' Z' H_j(0) M H_{j-k}(0) Z\theta = 0$ for $j = \kappa$, $\kappa + 1$, ..., $2\kappa - 1$. Therefore, $\frac{\partial^j}{\partial \gamma^j} \mathcal{L}_n^{(2)}(0; \theta) = 2\theta' Z' H_j(0) M u$. Finally, we examine the case in which $j = 2\kappa$. For each $j < 2\kappa$, $H_j(0) = 0$ and $H_\kappa(0) \neq 0$, so that the summand of the second term in the right side of (22) is different from zero only when $j = 2\kappa$ and $k = \kappa$: $\frac{\partial^{2\kappa}}{\partial \gamma^{2\kappa}} \mathcal{L}_n^{(2)}(0; \theta) = 2\theta' Z' H_{2\kappa}(\gamma) M u - \binom{2\kappa}{\kappa} \theta' Z' H_\kappa(\gamma) M H_\kappa(\gamma) Z\theta$. This completes the proof. ■

Proof of Lemma 3: Given Assumptions 1, 2, 7, and $\mathcal{H}_{02}$, we note that

$$QLR_n^{(2)} := \sup_\theta \overline{QLR}_n^{(2)}(\theta) = \sup_\theta \sup_\varsigma \frac{1}{\widehat{\sigma}_{n,0}^2} \left[\frac{2\{\theta' G_\kappa' u\} \varsigma^\kappa}{\kappa! \sqrt{n}} - \frac{1}{(2\kappa)! n} \left\{\binom{2\kappa}{\kappa} \theta' G_\kappa' G_\kappa \theta\right\} \varsigma^{2\kappa}\right] + o_{\mathbb{P}}(n). \qquad (23)$$

Then, the FOC with respect to ς implies that

$$\widehat{\varsigma}_n^{\kappa}(\theta) = \begin{cases} W_n(\theta), & \text{if } \kappa \text{ is odd;} \\ \max[0, W_n(\theta)] & \text{if } \kappa \text{ is even} \end{cases}$$

by noting that $\widehat{\varsigma}_n^{\kappa}(\theta)$ cannot be negative. If we plug $\widehat{\varsigma}_n^{\kappa}(\theta)$ back into the right side of (23), the desired result follows. ■

Proof of Lemma 4: Before proving Lemma 4, we first show that for each j, $Z'H_j(0)Mu = O_{\mathbb{P}}(n^{1/2})$, so that $j = \kappa+1, \ldots, 2\kappa-1$, $Z'H_j(0)Mu = o_{\mathbb{P}}(n^{j/2\kappa})$. Note that for $j = \kappa+1, \ldots, 2\kappa$, $Z'H_jMu = \sum_{t=1}^n z_t h_{t,j}(0)u_t - \sum_{t=1}^n z_t h_{t,j}(0)z_t' \left(\sum_{t=1}^n z_t z_t'\right)^{-1} \sum_{t=1}^n z_t u_t$. First, we apply the ergodic theorem to $n^{-1}\sum_t z_t h_{t,j}(0)z_t'$ and $n^{-1}\sum_t z_t$ z_t', respectively. Second, given Assumptions 1, 2, 3, 7, and 8, following the proof of Lemma 1, we have that $n^{-1/2}\sum_t z_t u_t$ is asymptotically normal. Furthermore, for all $j = \kappa+1, \ldots, 2\kappa$, we show that $n^{-1/2}\sum_t z_t h_{t,j}(0)u_t$ is asymptotically normal. To do this, first note that $\{z_t h_{t,j}(0)u_t, \mathcal{F}_t\}$ is a martingale difference sequence, so that for each j, $\mathbb{E}[z_t h_{t,j}(0)u_t|\mathcal{F}_{t-1}] = 0$. Next, we prove that for each j, $\mathbb{E}[z_{t,i}^2 h_{t,j}^2(0)u_t^2] < \infty$. First note that using the moment conditions in Assumption 7, $\mathbb{E}[|z_{t,i}^2 h_{t,j}^2(0)u_t^2|] \leq \mathbb{E}[|u_t|^4]^{1/2}\mathbb{E}[|h_{t,j}^2(0)z_{t,i}^2|^2]^{1/2} \leq \mathbb{E}[|u_t|^4]^{1/2}\mathbb{E}[|h_{t,j}(0)|^8]^{1/4}\mathbb{E}[|z_{t,i}|^8$ $]^{1/4} < \infty$ by the Cauchy-Schwarz's inequality. For the same reason, $\mathbb{E}[|z_{t,i}^2 h_{t,j}^2(0)u_t^2|] \leq \mathbb{E}[|u_t h_{t,j}(0)|^4]^{1/2}\mathbb{E}[|z_{t,i}|^4$ $]^{1/2} \leq \mathbb{E}[|u_t|^8]^{1/2}\mathbb{E}[|h_{t,j}(0)|^8]^{1/2}\mathbb{E}[|z_{t,i}|^4]^{1/2} < \infty$. By Assumption 8, $\mathbb{E}[u_t^2 z_t h_{t,j}(0)^2 z_t']$ is positive definite. It then follows by Theorem 5.25 of White (2001) that $n^{-1/2}\sum_t z_t h_{t,j}(0)u_t$ is asymptotically normal. Thus, $Z'H_j(0)Mu = O_{\mathbb{P}}(n^{1/2})$.

We now consider the statements (*i*)–(*iii*).

(*i*) First, we show that $\theta' Z'H_\kappa(0)Mu = O_{\mathbb{P}}(n^{1/2})$. By the definition of M,

$$Z'H_\kappa(0)Mu = \sum_{t=1}^n z_t h_{t,\kappa}(0)u_t - \sum_{t=1}^n z_t h_{t,\kappa}(0)z_t' \left(\sum_{t=1}^n z_t z_t'\right)^{-1} \sum_{t=1}^n z_t u_t. \tag{24}$$

We examine all sums on the right-hand side of (24). First, $h_{t,\kappa}(0)$ is a function of z_t, which implies that, given the moment condition in Assumption 7, $n^{-1}\sum z_t h_{t,\kappa}(0)z_t'$ obeys the ergodic theorem. Second, similarly under Assumptions 1, 2, 3, 7, 8, and $\mathcal{H}_{02}$, $n^{-1}\sum z_t z_t'$ also obeys the ergodic theorem. Third, given the assumptions and the proof of Lemma 1, we have already proved that $n^{-1/2}\sum z_t u_t$ is asymptotically normally distributed. Finally, $n^{-1/2}\sum z_t h_{t,\kappa}(0)u_t$ is asymptotically normal, and the proof is similar to that of the asymptotic normality of $n^{-1/2}\sum_t z_t h_{t,j}(0)u_t$ ($j = \kappa+1, \ldots, 2\kappa$). All these facts imply that $Z'H_\kappa(0)Mu = O_{\mathbb{P}}(n^{1/2})$.

(*ii*) $n^{-1}G_\kappa' G_\kappa \overset{\text{a.s.}}{\to} A_2$ by the ergodic theorem.

(*iii*) Note that

$$Z'H_\kappa(0)MH_\kappa(0)Z = \sum_{t=1}^n z_t h_{t,\kappa}(0)^2 z_t' - \sum_{t=1}^n z_t h_{t,\kappa}(0)z_t' \left(\sum_{t=1}^n z_t z_t'\right)^{-1} \sum_{t=1}^n z_t h_{t,\kappa}(0)z_t'. \tag{25}$$

The limit of (25) is revealed by applying the ergodic theorem to each term on the right-hand side of this expression.

Consequently, $n^{-1}Z'H_\kappa(0)MH_\kappa(0)Z \overset{\text{a.s.}}{\to} \mathbb{E}[g_{t,\kappa}g_{t,\kappa}']$, where $\mathbb{E}[g_{t,\kappa}g_{t,\kappa}'] := \mathbb{E}[z_t H_{2\kappa}(0)^2 z_t'] - \mathbb{E}[z_t H_{2\kappa}(0) z_t']\mathbb{E}[z_t z_t']^{-1}$ $\mathbb{E}[z_t H_{2\kappa}(0)\ z_t']$. This completes the proof. ■

Proof of Lemma 7: The distributional equivalence between $\dot{\mathcal{G}}(\cdot)$ and $\ddot{\mathcal{G}}(\cdot)$ can be established by showing that for all γ, $\widetilde{\gamma} \geq 0$, $\mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}(\widetilde{\gamma})] = \mathbb{E}[\dot{\mathcal{G}}(\gamma)\dot{\mathcal{G}}(\widetilde{\gamma})]$. We will proceed in three steps. First, we derive the functional form of $\ddot{\rho}(\gamma, \widetilde{\gamma})$. We show that if $\gamma, \widetilde{\gamma} > 0$, then $\ddot{k}_1(\gamma, \widetilde{\gamma}) = \sum_{n=1}^{\infty} b_n(\gamma) b_n(\widetilde{\gamma})$. This in turn implies that for $\gamma, \widetilde{\gamma} > 0$,

$$\ddot{\rho}(\gamma, \widetilde{\gamma}) = \frac{\sum_{n=1}^{\infty} b_n(\gamma) b_n(\widetilde{\gamma})}{\ddot{k}_1(\gamma, \gamma)^{1/2} \ddot{k}_1(\widetilde{\gamma}, \widetilde{\gamma})^{1/2}}.$$

It follows that the specific functional form of $\ddot{\rho}(\gamma, \widetilde{\gamma})$ can be obtained from this result and (12).

Second, similarly for all $\gamma, \widetilde{\gamma} \geq 0$, we derive the functional form of $\dot{\rho}(\gamma, \widetilde{\gamma})$ and compare it to $\ddot{\rho}(\gamma, \widetilde{\gamma})$. To do all this, we first note that for all $\gamma, \widetilde{\gamma} > 0$,

$$\begin{aligned}
\ddot{k}_1(\gamma, \tilde{\gamma}) &= \frac{1}{4}\mathbb{E}\left[y_t^2 \tanh\left(\frac{\gamma y_t}{2}\right) \tanh\left(\frac{\widetilde{\gamma} y_t}{2}\right)\right] - \frac{1}{4}\mathbb{E}\left[y_t^2 \tanh\left(\frac{\gamma y_t}{2}\right)\right] \mathbb{E}[y_t^2]^{-1} \mathbb{E}\left[y_t^2 \tanh\left(\frac{\widetilde{\gamma} y_t}{2}\right)\right] \\
&= \frac{1}{4}\mathbb{E}\left[y_t^2 \tanh\left(\frac{\gamma y_t}{2}\right) \tanh\left(\frac{\widetilde{\gamma} y_t}{2}\right)\right]. \qquad (26)
\end{aligned}$$

This follows from that fact that for any $x \in \mathbb{R}$, $\tanh(x) = -\tanh(-x)$ and that y_t follows the Laplace distribution with mean zero and variance 2, so that $\mathbb{E}\left[y_t^2 \tanh\left(\gamma y_t/2\right)\right] = 0$. Given this, we can apply the Dirichlet series to $\tanh(\cdot)$ to obtain the functional form of $\ddot{k}_1(\cdot, \cdot)$. Thus, for any $x \in \mathbb{R}$, $\tanh(x) = \mathrm{sgn}(x)(1 - 2\sum_{k=0}^{\infty}(-1)^k \exp(-2|x|(k+1)))$ and, furthermore, that $\mathbb{E}\left[s_t^2 \exp(-s_t \gamma k)\right] = 2/(1+\gamma k)^3$ and $\mathbb{E}[s_t^2] = 2$, where $s_t := |y_t|$ follows the exponential distribution with mean 1 and variance 2. Applying these to (26) yields

$$\begin{aligned}
\ddot{k}_1(\gamma, \widetilde{\gamma}) &= \mathbb{E}\left[\frac{y_t^2}{4} \tanh\left(\frac{\gamma y_t}{2}\right) \tanh\left(\frac{\tilde{\gamma} y_t}{2}\right)\right] \\
&= \mathbb{E}\left[\frac{s_t^2}{4}\right] - \mathbb{E}\left[\frac{s_t^2}{2}\sum_{k=1}^{\infty}(-1)^{k-1}\exp(-s\gamma k)\right] - \mathbb{E}\left[\frac{s_t^2}{2}\sum_{k=1}^{\infty}(-1)^{k-1}\exp(-s\widetilde{\gamma} k)\right] \\
&\quad + \mathbb{E}\left[s_t^2 \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(-1)^{k+j-2}\exp(-s_t(\gamma k + \widetilde{\gamma} j))\right] \\
&= \frac{1}{2} - \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\gamma k)^3} - \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\widetilde{\gamma} k)^3} + \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(-1)^{j+k-2}\frac{2}{(1+\gamma k + \widetilde{\gamma} j)^3}.
\end{aligned}$$

Next, for $|x| < 1$ we have

$$\frac{1}{(1-x)^3} = \sum_{n=1}^{\infty}\frac{n(n+1)}{2}x^{n-1}, \quad \text{so that} \quad \frac{1}{(1+\gamma k + \widetilde{\gamma} j)^3} = \frac{1}{(1+\gamma k)^3(1+\widetilde{\gamma} j)^3\left(1 - \frac{\gamma k}{1+\gamma k}\frac{\widetilde{\gamma} j}{1+\widetilde{\gamma} j}\right)^3},$$

where

$$\frac{1}{\left(1-\frac{\gamma k}{1+\gamma k}\frac{\widetilde{\gamma} j}{1+\widetilde{\gamma} j}\right)^3}=\sum_{n=1}^{\infty}\frac{n(n+1)}{2}\left(\frac{\gamma k}{1+\gamma k}\frac{\widetilde{\gamma} j}{1+\widetilde{\gamma} j}\right)^{n-1}.$$

Therefore,

$$\ddot{k}_1(\gamma,\widetilde{\gamma})=\frac{1}{2}-\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\gamma k)^3}-\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\widetilde{\gamma} k)^3}+\sum_{n=1}^{\infty}\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(-1)^{j+k-2}n(n+1)\frac{(\gamma k)^{n-1}}{(1+\gamma k)^{n+2}}\frac{(\widetilde{\gamma} j)^{n-1}}{(1+\widetilde{\gamma} j)^{n+2}}.$$

Furthermore,

$$\begin{aligned}
&\sum_{n=1}^{\infty}\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(-1)^{j+k-2}n(n+1)\frac{(\gamma k)^{n-1}}{(1+\gamma k)^{n+2}}\frac{(\widetilde{\gamma} j)^{n-1}}{(1+\widetilde{\gamma} j)^{n+2}}\\
&\quad=2\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\gamma k)^3}\sum_{j=1}^{\infty}\frac{(-1)^{j-1}}{(1+\widetilde{\gamma} j)^3}+\sum_{n=2}^{\infty}n(n+1)\sum_{k=1}^{\infty}\frac{(-1)^{k-1}(\gamma k)^{n-1}}{(1+\gamma k)^{n+2}}\sum_{j=1}^{\infty}\frac{(-1)^{j-1}(\widetilde{\gamma} j)^{n-1}}{(1+\widetilde{\gamma} j)^{n+2}}\\
&\quad=2a(\gamma)a(\widetilde{\gamma})+\sum_{n=2}^{\infty}b_n(\gamma)b_n(\widetilde{\gamma}),
\end{aligned}$$

where for $n=2,3,\ldots$,

$$a(\gamma):=\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\gamma k)^3}\quad\text{and}\quad b_n(\gamma):=\sqrt{n(n+1)}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}(\gamma k)^{n-1}}{(1+\gamma k)^{n+2}}.$$

In particular, $b_1(\gamma):=2^{-1/2}(1-2a(\gamma))$, so that

$$\begin{aligned}
\ddot{k}_1(\gamma,\widetilde{\gamma})&=\frac{1}{2}-a(\gamma)-a(\widetilde{\gamma})+2a(\gamma)a(\widetilde{\gamma})+\sum_{n=2}^{\infty}b_n(\gamma)b_n(\widetilde{\gamma})\\
&=\frac{1}{2}(1-2a(\gamma))(1-2a(\widetilde{\gamma}))+\sum_{n=2}^{\infty}b_n(\gamma)b_n(\widetilde{\gamma})=\sum_{n=1}^{\infty}b_n(\gamma)b_n(\widetilde{\gamma}).
\end{aligned}$$

Then, for each $\gamma,\widetilde{\gamma}>0$,

$$\ddot{\rho}_1(\gamma,\widetilde{\gamma}):=\mathbb{E}[\ddot{\mathcal{G}}_1(\gamma)\ddot{\mathcal{G}}_1(\widetilde{\gamma})]=\frac{\sum_{n=1}^{\infty}b_n(\gamma)b_n(\widetilde{\gamma})}{\ddot{k}_1(\gamma,\gamma)^{1/2}\ddot{k}_1(\widetilde{\gamma},\widetilde{\gamma})^{1/2}}.$$

In addition, for $\gamma>0$, we examine $\ddot{\rho}_3(\gamma):=\mathbb{E}[\ddot{\mathcal{G}}_1(\gamma)\ddot{\mathcal{G}}_2]$. Note that from (12),

$$\ddot{\rho}_3(\gamma)=\frac{\mathbb{E}[y_t^3\tanh(\gamma y_t/2)]}{4\sqrt{6}\ddot{k}_1(\gamma,\gamma)^{1/2}}=\frac{1}{32\sqrt{6}\gamma^4\ddot{k}_1(\gamma,\gamma)^{1/2}}\left[48\gamma^4+P_G\left(3,1+\frac{1}{2\gamma}\right)-P_G\left(3,\frac{1+\gamma}{2\gamma}\right)\right]$$

as affirmed by Mathematica. It follows that the specific functional form of $\ddot{\rho}(\gamma,\widetilde{\gamma})$ is given as

$$\ddot{\rho}(\gamma,\widetilde{\gamma})=\begin{cases}\frac{\ddot{k}_1(\gamma,\widetilde{\gamma})}{\ddot{k}_1(\gamma,\gamma)^{1/2}\ddot{k}_1(\widetilde{\gamma},\widetilde{\gamma})^{1/2}}, & \text{if } \gamma>0 \text{ and } \widetilde{\gamma}>0;\\ 1, & \text{if } \gamma=0 \text{ and } \widetilde{\gamma}=0;\\ \frac{48\gamma^4+P_G(3,1+\frac{1}{2\gamma})-P_G(3,\frac{1+\gamma}{2\gamma})}{32\sqrt{6}\gamma^4\ddot{k}_1(\gamma,\gamma)^{1/2}}, & \text{if } \gamma>0 \text{ and } \widetilde{\gamma}=0,\end{cases}. \tag{27}$$

Third, we examine the covariance kernel of $\dot{\mathcal{G}}(\cdot)$, viz., $\dot{\rho}(\cdot,\cdot)$. If we let $\gamma, \widetilde{\gamma} > 0$,

$$\dot{\rho}(\gamma,\widetilde{\gamma}) := \mathbb{E}[\dot{\mathcal{G}}(\gamma)\cdot\mathcal{G}(\widetilde{\gamma})] = \frac{\sum_{n=1}^{\infty} b_n(\gamma)b_n(\widetilde{\gamma})}{\ddot{k}_1(\gamma,\gamma)^{1/2}\ddot{k}_1(\widetilde{\gamma},\widetilde{\gamma})^{1/2}} = \frac{\ddot{k}_1(\gamma,\widetilde{\gamma})}{\ddot{k}_1(\gamma,\gamma)^{1/2}\ddot{k}_1(\widetilde{\gamma},\widetilde{\gamma})^{1/2}} = \ddot{\rho}_1(\gamma,\widetilde{\gamma}).$$

Furthermore, by some tedious algebra,

$$\mathrm{plim}_{\gamma\downarrow 0}\ddot{\mathcal{Z}}_1^2(\gamma) = 0, \quad \mathrm{plim}_{\gamma\downarrow 0}\frac{\partial}{\partial\gamma}\ddot{\mathcal{Z}}_1^2(\gamma) = 0, \quad \mathrm{plim}_{\gamma\downarrow 0}\frac{\partial^2}{\partial\gamma^2}\ddot{\mathcal{Z}}_1^2(\gamma) = \frac{1}{8}\{3\sqrt{2}Z_1 + \sqrt{6}Z_2\}^2,$$

$$\mathrm{plim}_{\gamma\downarrow 0}\ddot{k}_1(\gamma,\gamma) = 0, \quad \mathrm{plim}_{\gamma\downarrow 0}\frac{\partial}{\partial\gamma}\ddot{k}_1(\gamma,\gamma) = 0, \quad \text{and} \quad \mathrm{plim}_{\gamma\downarrow 0}\frac{\partial^2}{\partial\gamma^2}\ddot{k}_1(\gamma,\gamma) = 3$$

so that $\mathrm{plim}_{\gamma\downarrow 0}\dot{\mathcal{G}}^2(\gamma) = \left(\frac{\sqrt{3}}{2}z_1 + \frac{1}{2}z_2\right)^2$, which implies $\dot{\mathcal{G}}_2 := \mathrm{plim}_{\gamma\downarrow 0}\dot{\mathcal{G}}(\gamma) = \frac{\sqrt{3}}{2}z_1 + \frac{1}{2}z_2 \sim N(0,1)$. Consequently, if $\gamma > 0$,

$$\begin{aligned}
\mathbb{E}[\dot{\mathcal{G}}(\gamma)\dot{\mathcal{G}}_2] &= \ddot{k}_1(\gamma,\gamma)^{-1/2}\mathbb{E}\left[\ddot{\mathcal{Z}}_1(\gamma)\left(\frac{\sqrt{3}}{2}z_1 + \frac{1}{2}z_2\right)\right] = \ddot{k}_1(\gamma,\gamma)^{-1/2}\left[\frac{\sqrt{3}}{2}b_1(\gamma) + \frac{1}{2}b_2(\gamma)\right] \\
&= \frac{1}{32\sqrt{6}\gamma^4\ddot{k}_1(\gamma,\gamma)^{1/2}}\left[48\gamma^4 + P_G\left(3, 1+\frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right].
\end{aligned} \tag{28}$$

The last equality follows from

$$b_1(\gamma) = \frac{1}{8\sqrt{2}\gamma^3}\left[8\gamma^3 - P_G\left(2, 1+\frac{1}{2\gamma}\right) + P_G\left(2, \frac{1+\gamma}{2\gamma}\right)\right],$$

$$b_2(\gamma) = \frac{1}{16\sqrt{6}\gamma^4}\left[6\gamma P_G\left(2, \frac{1}{2\gamma}\right) - 6\gamma P_G\left(2, \frac{1+\gamma}{2\gamma}\right) + P_G\left(3, \frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right],$$

$$P_G\left(2, \frac{1}{2\gamma}\right) - P_G\left(2, 1+\frac{1}{2\gamma}\right) = -16\gamma^3, \quad \text{and} \quad P_G\left(3, \frac{1}{2\gamma}\right) - P_G\left(3, 1+\frac{1}{2\gamma}\right) = 96\gamma^4,$$

as obtained by Mathematica. Equation (28) then leads to the following functional form for $\dot{\rho}(\gamma,\widetilde{\gamma}) := \mathbb{E}[\dot{\mathcal{G}}(\gamma)\dot{\mathcal{G}}(\widetilde{\gamma})]$:

$$\dot{\rho}(\gamma,\widetilde{\gamma}) = \begin{cases} \frac{\ddot{k}_1(\gamma,\widetilde{\gamma})}{\ddot{k}_1(\gamma,\gamma)^{1/2}\ddot{k}_1(\widetilde{\gamma},\widetilde{\gamma})^{1/2}}, & \text{if } \gamma > 0 \text{ and } \widetilde{\gamma} > 0; \\ 1, & \text{if } \gamma = 0 \text{ and } \widetilde{\gamma} = 0; \\ \frac{48\gamma^4 + P_G(3, 1+\frac{1}{2\gamma}) - P_G(3, \frac{1+\gamma}{2\gamma})}{32\sqrt{6}\gamma^4\ddot{k}_1(\gamma,\gamma)^{1/2}}, & \text{if } \gamma > 0 \text{ and } \widetilde{\gamma} = 0, \end{cases}$$

which is identical to the functional form of $\ddot{\rho}(\cdot,\cdot)$ in (27). This allows the conclusion that $\ddot{\mathcal{G}}(\cdot)$ has the same distribution as $\dot{\mathcal{G}}(\cdot)$. ■

In the following, we provide additional supplementary claim in (13) that is given in the following lemma:

Lemma 8. *Given the DGP and Model conditions in Section 3.2,*

$$\lim_{\tilde{\gamma}\downarrow 0}\ddot{\rho}_1(\gamma,\tilde{\gamma})^2 = \left(\frac{1}{32\sqrt{6}\gamma^4\ddot{k}_1(\gamma,\gamma)^{1/2}}\left[48\gamma^4 + P_G\left(3, 1+\frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right]\right)^2. \qquad \square$$

Lemma 8 implies that $\text{plim}_{\gamma\downarrow 0}\ddot{\mathcal{G}}_1(\gamma)^2 = \ddot{\mathcal{G}}_2^2$, so that $\sup_{\gamma\in\Gamma}\ddot{\mathcal{G}}_1(\gamma)^2 \geq \ddot{\mathcal{G}}_2^2$ and $QLR_n \Rightarrow \sup_{\gamma\in\Gamma}\ddot{\mathcal{G}}_1(\gamma)^2$.

Proof of Lemma 8: From the definition of $\ddot{\rho}_1(\gamma,\tilde{\gamma})$, we note that

$$\ddot{\rho}_1(\gamma,\tilde{\gamma})^2 := \frac{\ddot{k}_1(\gamma,\tilde{\gamma})^2}{\ddot{k}_1(\gamma,\gamma)\ddot{k}_1(\tilde{\gamma},\tilde{\gamma})}.$$

Furthermore, we have $\text{plim}_{\tilde{\gamma}\downarrow 0}\ddot{k}_1(\gamma,\tilde{\gamma})^2 = 0$, $\text{plim}_{\tilde{\gamma}\downarrow 0}\frac{\partial}{\partial\tilde{\gamma}}\ddot{k}_1(\gamma,\tilde{\gamma})^2 = 0$, $\text{plim}_{\tilde{\gamma}\downarrow 0}\ddot{k}_1(\tilde{\gamma},\tilde{\gamma}) = 0$, $\text{plim}_{\tilde{\gamma}\downarrow 0}(\partial/\partial\tilde{\gamma})\ddot{k}_1(\tilde{\gamma},\tilde{\gamma}) = 0$, $\text{plim}_{\tilde{\gamma}\downarrow 0}(\partial^2/\partial\tilde{\gamma}^2)\ddot{k}_1(\tilde{\gamma},\tilde{\gamma}) = 3$, and

$$\text{plim}_{\tilde{\gamma}\downarrow 0}\frac{\partial^2}{\partial\tilde{\gamma}^2}\ddot{k}_1(\gamma,\tilde{\gamma})^2 = \left(\frac{48\gamma^4 + P_G(3, 1+\frac{1}{2\gamma}) - P_G(3, \frac{1+\gamma}{2\gamma})}{32\sqrt{2}\gamma^4}\right)^2$$

by some tedious algebra using Mathematica. This property implies that

$$\begin{aligned}\lim_{\tilde{\gamma}\downarrow 0}\ddot{\rho}_1(\gamma,\tilde{\gamma})^2 &= \frac{1}{3\ddot{k}_1(\gamma,\gamma)}\left(\frac{48\gamma^4 + P_G(3, 1+\frac{1}{2\gamma}) - P_G(3, \frac{1+\gamma}{2\gamma})}{32\sqrt{2}\gamma^4}\right)^2 \\ &= \left(\frac{1}{32\sqrt{6}\gamma^4\ddot{k}_1(\gamma,\gamma)^{1/2}}\left[48\gamma^4 + P_G\left(3, 1+\frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right]\right)^2.\end{aligned}$$

This completes the proof. ■

References

ABADIR, K. AND MAGNUS, J. (2005): *Matrix Algebra*. NY: Cambridge University Press.

AUERBACH, A. J. AND GORODNICHENKO, Y. (2012a): "Fiscal Multipliers in Recession and Expansion," in *Fiscal Policy after the Financial Crisis*. Eds. Alberto Alesina and Francesco Giavazzi, 63–98. Chicago, IL: University of Chicago Press.

AUERBACH, A. J. AND GORODNICHENKO, Y. (2012b): "Measuring the Output Responses to Fiscal Policy," *American Economic journal: Economic Policy*, 4, 1–27.

AUERBACH, A. J. AND GORODNICHENKO, Y. (2012c): "Replication Data for: Measuring the Output Responses to Fiscal Policy," Nashville, TN: American Economic Association. Ann Arbor, MI: Inter-university Consortium for Political and Social Research, 2019-10-13.

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

BACON, D. W. AND WATTS, D.G. (1971): "Estimating the Transition between Two Intersecting Straight Lines," *Biometrika*, 58, 525–534.

BILLINGSLEY, P. (1999): *Convergence of Probability Measures*. New York: Wiley.

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

CHAN, K. S., AND TONG, H. (1986): "On Estimation Thresholds in Autoregressive Models," *Journal of Time Series Analysis*, 7, 179–190.

CHENG, X. (2015): "Robust Inference in Nonlinear Models with Mixed Identification Strength," *Journal of Econometrics*, 189, 207–228.

CHO, J. S., CHEONG, T. U., AND WHITE, H. (2011): "Experience with the Weighted Bootstrap in Testing for Unobserved Heterogeneity in Exponential and Weibull Duration Models," *Journal of Economic Theory and Econometrics*, 22:2, 60–91.

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. (2014): "Testing for Neglected Nonlinearity Using Twofold Unidentified Models under the Null and Hexic Expansions," *Essays in Nonlinear Time Series Econometrics*. Eds. Niels Haldrup, Mika Meitz, and Pentti Saikkonen, 3–27. Oxford: Oxford University Press.

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

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

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

CHO, J.S. AND WHITE, H. (2018): "Directionally Differentiable Econometric Models," *Econometric Theory*, 34, 1101–1131.

CHRISTIANO, L., EICHENBAUM, M., AND REBELO, S. (2011): "When is the Government Spending Multiplier Large?" *Journal of Political Economic*, 119, 78–121.

COENEN, G., ERCEG, C. J., FREEDMAN, C., FURCERI, D., KUMHOF, M., LALONDE, R., LAXTON, D., LIND'E, J., MOUROUGANE, A., MUIR, D., MURSULA, S., DE RESENDE, C., ROBERTS, J., ROGER, W., SNUDDEN, S., TRABANDT, M., AND IN'T VELD, J. (2012): "Effects of Fiscal Stimulus in Structural Models," *American Economic Journal: Macroeconomics*, 4, 22–68.

DAVIDSON, J., MONTICINI, A., AND PEEL, D. (2007): "Implementing the Wild Bootstrap Using a Two-Point Distribution," *Economics Letters*, 96, 309–315.

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.

DESCHAMPS, P. (2008): "Comparing Smooth Transition and Markov Switching Autoregressive Models of US Unemployment," *Journal of Applied Econometrics*, 23, 435–462.

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.

EKLUND, B. AND TERÄSVIRTA, T. (2007): "Testing Consistancy of the Error Covariance Matrix in Vector Models," *Journal of Econometrics*, 140, 753–780.

ESCRIBANO, Á AND JORDÁ, O. (1999): "Testing Nonlinearity: Decision Rules for Selecting between Logistic and Expnential STAR Models," *Spanish Economic Review*, 3, 193–209.

FAZZRI, S. M., MORLEY, S. C., AND PANOVSKA, I. (2014): "State-Dependent Effects of Fiscal Policy," *Studies in Nonlinear Dynamics & Econometrics*, 19, 285–315.

GRANGER, C. W. J., AND T. TERÄSVIRTA (1993): *Modelling Non-Linear Economic Relationships*. Oxford: Oxford University Press.

HAGGAN, V. AND T. OZAKI (1981): "Modelling Non-Linear Random Vibrations Using an Amplitude-Dependent Time Series Model," *Biometrika*, 68, 189–196.

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

HUBRICH, K. AND TERÄSVIRTA, T. (2013): "Thresholds and Smooth Transitions in Vector Autoregressive Models," *VAR Models in Macroeconomics–New Developments and Applications: Essays in Honor of Christopher A. Sims*. Eds. Thomas B. Fomby, Lutz Kilian, and Anthony Murphy. *Advances in Econometrics*. Vol. 32, 273–326. Bingley, U.K.: Emerald.

JOHANSEN, S. (1988): "Statistical analysis of cointegration vectors," *Journal of economic dynamics and control*, 12, 231–254.

JOHANSEN, S. (1991): "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models," *Econometrica*, 59, 1551–1580.

JOHANSEN, S. (1995): *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford, U.K.: Oxford University Press.

LAWRANCE, A. J., AND LEWIS, P. A. W. (1980): "The Exponential Autoregressive-Moving Average EARMA (p, q) Process," *Journal of the Royal Statistical Society. Series B*, 47, 150–161.

LÜTKEPOHL, H., TERÄSVIRTA, T., AND WOLTERS, J. (1999): "Investigating Stability and Linearity of a German M1 Money Demand Function," *Journal of Applied Econometrics*, 14, 511-525.

LUUKKONEN, R., SAIKKONEN, P., AND TERÄSVIRTA, T., (1988): "Testing Linearity Against Transition Autoregressive Models," *Biometrika*, 75, 491–499.

OWYANG, M. T., RAMEY, V. A., AND ZUBAIRY, S. (2013): "Are Government Spending Multipliers Greater during Periods of Slack? Evidence from Twentieth-Century Historical Data," *American Economic Review*, 103, 129–134.

PARKER, J. A. (2011): "On Measuring the Effects of Fiscal Policy in Recessions," *Journal of Economic Literature*, 49, 703–718.

RAMEY, V. A. (2011a): "Can Government Purchases Stimulate the Economy?" *Journal of Economic Literature*, 49, 673–685.

RAMEY, V. A. (2011b): "Identifying Government Spending Shocks: It's all in the Timing," *Quarterly Journal of Economics*, 126, 1–50.

RAMEY, V. A. AND ZUBAIRY, S. (2018): "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data," *Journal of Political Economy*, 126, 850–901.

RANGA RAO, R. (1962): "Relations Between Weak and Uniform Convergence of Measures with Applications," *Annals of Mathematical Statistics*, 33, 659–680.

ROTHMAN, P., VAN DIJIK, D., AND FRANSES, P. H. (2001): “Multivariate STAR Analysis of Money-Output Relationship,” *Macroeconomic Dynamics*, 5, 506–532.

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

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

TERÄSVIRTA, T., TJØSTHEIM, D., AND GRANGER, C.W.J. (2010): *Modelling Nonlinear Economic Time Series*. Oxford: Oxford University Press.

VAN DIJK, D., TERÄSVIRTA, T., AND FRANSES, P. H. (2002): “Smooth Transition Autoregressive Models–a Survey of Recent Developments,” *Econometric Reviews*, 21, 1–47.

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.

QLR tests \ variables	Δg_t	$\Delta \tau_t$	Δy_t
QLR_n	$\mathbf{0.000}$	$\mathbf{0.004}$	$\mathbf{0.000}$
Eklund and Teräsvirta's (2007) conditional heteroskedasticity test			0.158
Multivariate Ljung-Box test			0.175

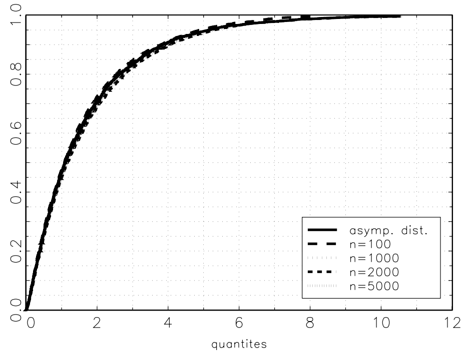
Table 1: p-VALUES OF THE DIAGNOSTIC TEST STATISTICS. Notes: The figures in QLR_n row show the p-values of the QLR test statistics for linearity based upon the LSTAR VSTEC model, and they are obtained using 20,000 bootstrap replications. The variables in the first row denote the dependent variables in the marginal models of (17). In addition, we repot the p-values of the multivariate Ljung-Box test statistic and Eklund and Teräsvirta (2007) test statistic for conditional heteroskedasticity. Boldface p-values indicate significance levels less than or equal to 0.05.

	$\max_{h=1,\ldots,H} Y_h$		$\sum_{h=1}^{H} Y_h / \sum_{h=1}^{H} G_h$	
	$H = 8$	$H = 20$	$H = 8$	$H = 20$
Linear	1.299	0.664	1.299	0.926
	(0.762)	(0.849)	(1.530)	(1.550)
Expansive	0.666	0.223	0.666	0.272
	(0.440)	(0.848)	(1.906)	(1.617)
Recessive	0.961	0.587	1.125	1.001
	(0.248)	(0.331)	(0.525)	(0.569)

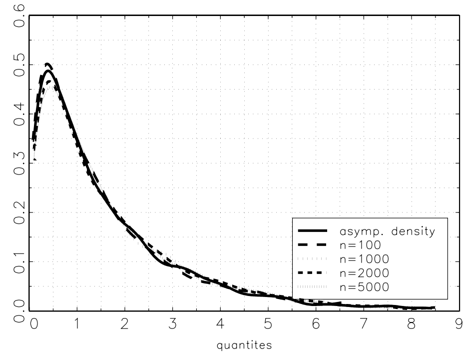
Table 2: ESTIMATED GOVERNMENT SPENDING MULTIPLIERS UNDER EACH LINEAR, RECESSIVE, AND EXPANSIVE STATES. This table shows the estimated output multipliers for a \$1 increase in government spending, measured by maximum output increase and total output increase relative to total government spending over 8 and 20 quarters. The figures in parentheses are the standard errors of the estimates.

Periods	Transition Variable	$LM_{1,n}$	$LM_{2,n}$	$LM_{3,n}$	$LM_{4,n}$	QLR_n^L	QLR_n^E
1968.06∼1999.12	$\Delta_{12} y_{t-1}$	0.150	0.532	0.412	0.895	**0.000**	**0.045**
	$\Delta_{12} y_{t-2}$	**0.037**	0.093	0.057	0.195	**0.000**	**0.028**
	$\Delta_{12} y_{t-3}$	0.162	0.326	0.163	0.555	**0.012**	0.054
	$\Delta_{12} y_{t-4}$	0.665	0.745	0.546	0.619	**0.014**	0.098
	$\Delta_{12} y_{t-5}$	0.662	0.886	0.954	0.830	**0.003**	0.099
	$\Delta_{12} y_{t-6}$	0.588	0.306	0.121	0.234	**0.003**	0.157
1968.06∼2015.08	$\Delta_{12} y_{t-1}$	**0.000**	**0.000**	**0.000**	0.098	**0.000**	**0.000**
	$\Delta_{12} y_{t-2}$	**0.000**	**0.000**	**0.000**	**0.016**	**0.000**	**0.000**
	$\Delta_{12} y_{t-3}$	**0.001**	**0.000**	**0.008**	**0.045**	**0.000**	**0.014**
	$\Delta_{12} y_{t-4}$	**0.008**	**0.012**	0.070	0.111	**0.000**	**0.009**
	$\Delta_{12} y_{t-5}$	**0.038**	0.237	0.274	0.861	**0.000**	**0.049**
	$\Delta_{12} y_{t-6}$	**0.003**	0.068	**0.017**	0.582	**0.000**	0.350

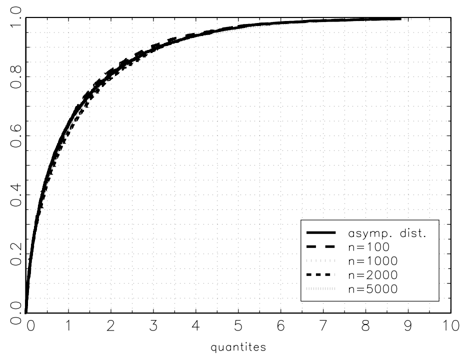
Table 3: LINEARITY TESTS FOR THE MONTHLY US UNEMPLOYMENT RATE. Notes: The p-values of the linearity tests for the first differenced monthly US unemployment rate are provided. The p-values in the top panel are obtained using observations from 1968.06 to 1999.12, and the p-values of the bottom panel are obtained using observations from 1968.06 to 2015.08. The null linear model is given as AR(15) by AIC, and the twelve-month differences are considered as a transition variable. Boldface p-values indicate significance levels less than or equal to 0.05.



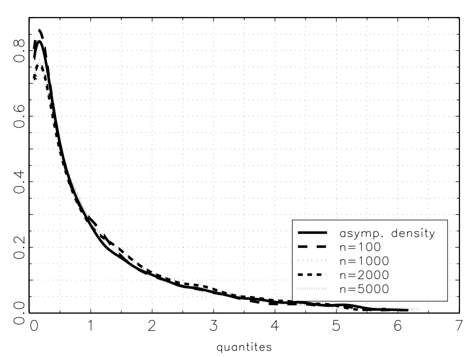
$\Gamma_1 = [0, 2]$



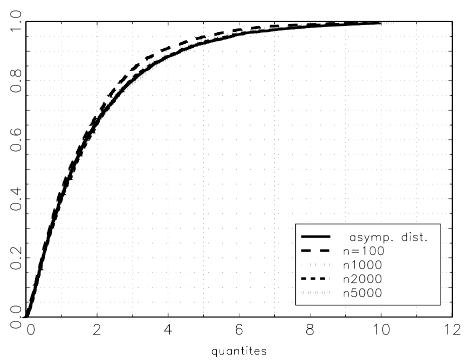
$\Gamma_1 = [0, 2]$



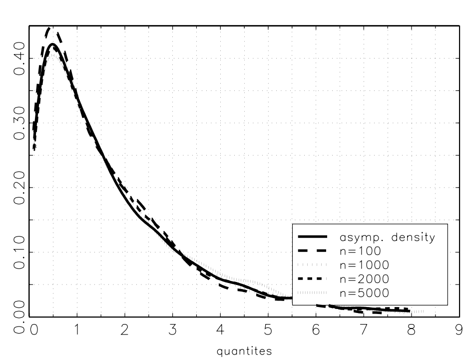
$\Gamma_2 = [0, 3]$



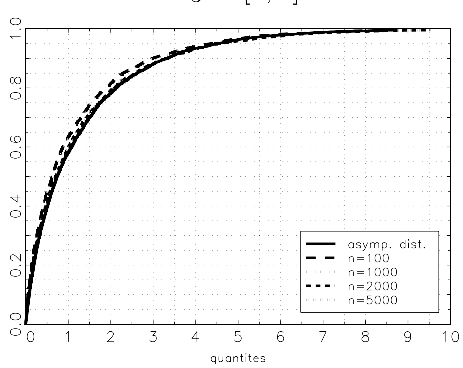
$\Gamma_2 = [0, 3]$



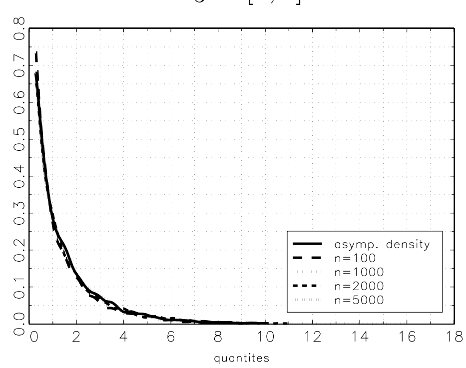
$\Gamma_3 = [0, 4]$



$\Gamma_3 = [0, 4]$

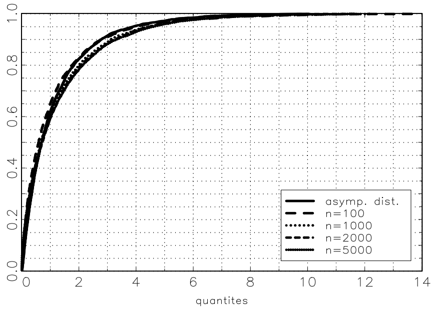


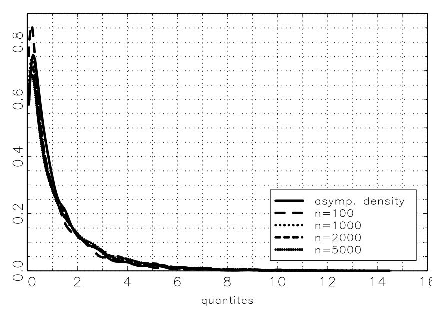
$\Gamma_4 = [0, 5]$



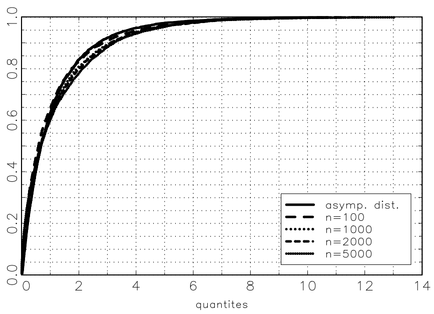
$\Gamma_4 = [0, 5]$

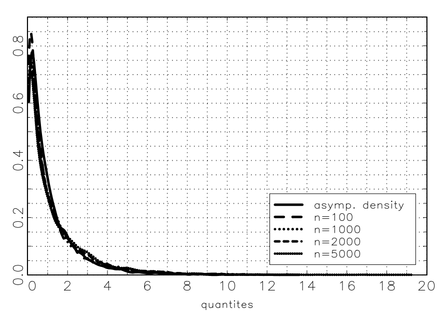
Figure 1: EMPIRICAL NULL DISTRIBUTIONS OF THE QLR STATISTIC AND ITS NULL LIMIT DISTRIBUTION (ESTAR MODEL CASE). Notes: (*i*) Number of Iterations: 5,000; (*ii*) DGP: $y_t = 0.5y_{t-1} + u_t$ and $u_t \sim$ IID $N(0, 1)$; (*iii*) Model: $y_t = \pi y_{t-1} + \theta y_{t-1}\{1 - \exp(-\gamma y_{t-1}^2)\} + u_t$ and $u_t \sim$ IID $N(0, 1)$; and (*iv*) $\Gamma_1 = [0, 2]$, $\Gamma_2 = [0, 3]$, $\Gamma_3 = [0, 4]$, $\Gamma_4 = [0, 5]$.



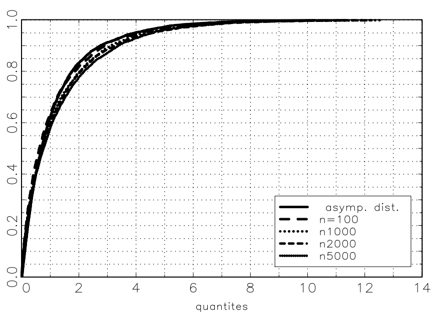


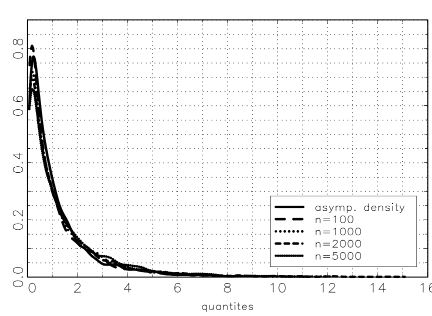
$\Gamma_1 = [0, 2]$ $\Gamma_1 = [0, 2]$



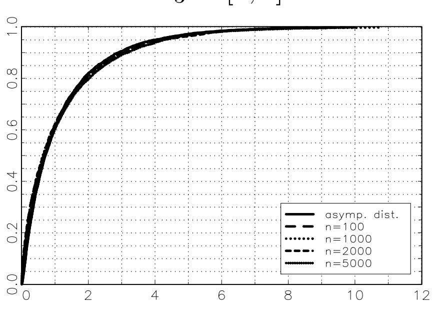


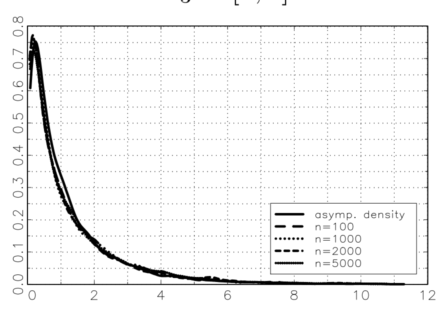
$\Gamma_2 = [0, 3]$ $\Gamma_2 = [0, 3]$





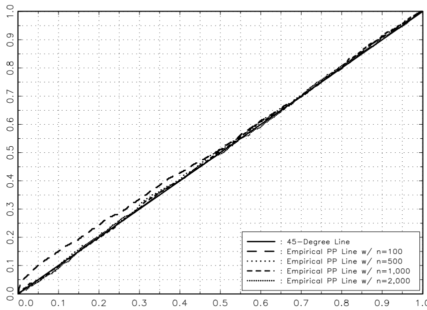
$\Gamma_3 = [0, 4]$ $\Gamma_3 = [0, 4]$



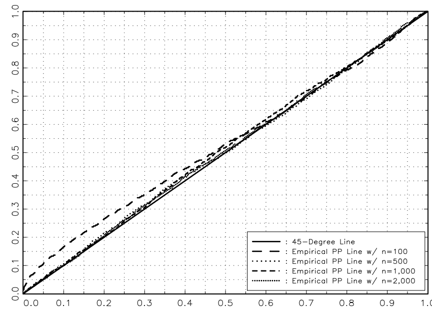


$\Gamma_4 = [0, 5]$ $\Gamma_4 = [0, 5]$

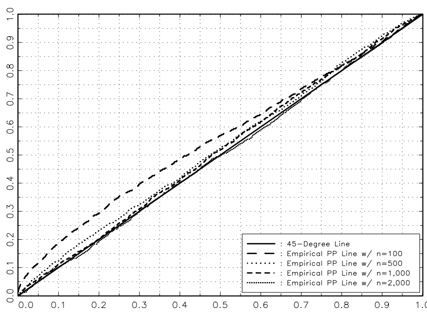
Figure 2: EMPIRICAL NULL DISTRIBUTIONS OF THE QLR STATISTIC AND ITS NULL LIMIT DISTRIBUTION (LSTAR MODEL CASE). Notes: (*i*) Number of Iterations: 5,000; (*ii*) DGP: $y_t = 0.5 y_{t-1} + u_t$ and $u_t = i_t \ell_t$, where $\{i_t\}$ is an IID sequence in which $\mathbf{P}\{i_t = 1\} = 1 - 0.5^2$ and $\{\ell_t\} \sim$ Laplace$(0, 2)$; (*iii*) Model: $y_t = \pi y_{t-1} + \theta y_{t-1}\{(1+\exp(-\gamma y_{t-1}))^{-1}\} + u_t$ and $u_t = i_t \ell_t$, where $\{i_t\}$ is an IID sequence in which $\mathbf{P}\{i_t = 1\} = 1 - 0.5^2$ and $\{\ell_t\} \sim$ Laplace$(0, 2)$; and (*iv*) $\Gamma_1 = [0, 2]$, $\Gamma_2 = [0, 3]$, $\Gamma_3 = [0, 4]$, $\Gamma_4 = [0, 5]$.



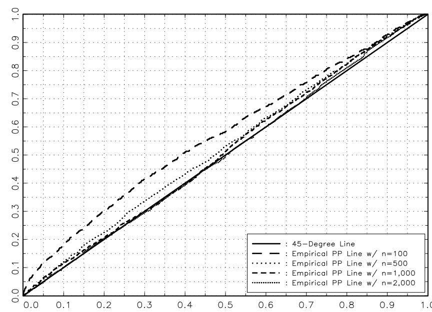
$\Gamma_1 = [0, 2]$



$\Gamma_2 = [0, 3]$

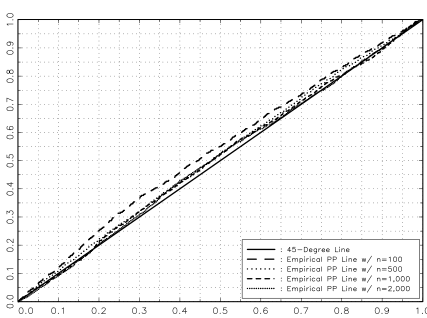


$\Gamma_3 = [0, 4]$

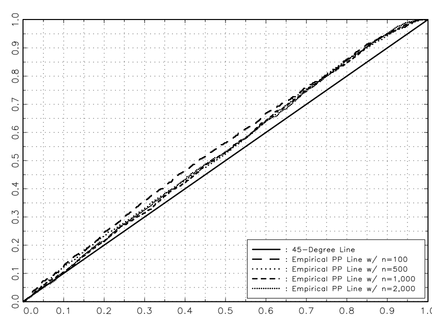


$\Gamma_4 = [0, 5]$

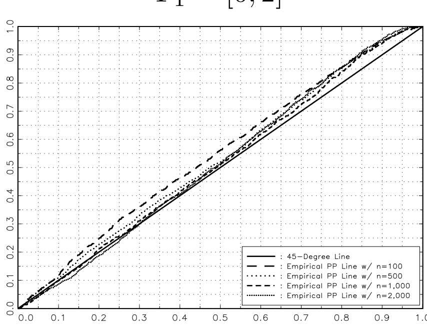
Figure 3: PP PLOTS OF THE QLR STATISTIC USING THE WEIGHTED BOOTSTRAP (ESTAR MODEL CASE). Notes: (*i*) Number of Iterations: 2,000, Bootstrap Iterations: 300; (*ii*) DGP: $y_t = 0.5y_{t-1} + u_t$ and $u_t \sim$ IID $N(0, 1)$; (*iii*) Model: $y_t = \pi y_{t-1} + \theta y_{t-1}\{1 - \exp(-\gamma y_{t-1}^2)\}u_t$ and $u_t \sim$ IID $N(0, 1)$; and (*iv*) $\Gamma_1 = [0, 2]$, $\Gamma_2 = [0, 3]$, $\Gamma_3 = [0, 4]$, $\Gamma_4 = [0, 5]$.



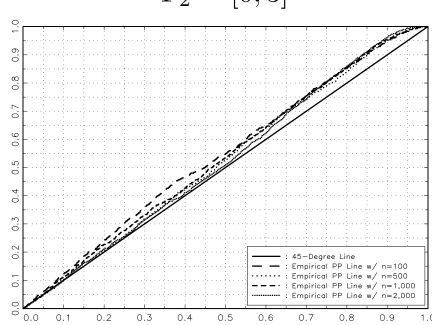
$\Gamma_1 = [0, 2]$



$\Gamma_2 = [0, 3]$



$\Gamma_3 = [0, 4]$



$\Gamma_4 = [0, 5]$

Figure 4: PP PLOTS OF THE QLR STATISTIC USING THE WEIGHTED BOOTSTRAP (LSTAR MODEL CASE). Notes: (*i*) Number of Iterations: 2,000, Bootstrap Iterations: 300; (*ii*) DGP: $y_t = 0.5y_{t-1} + u_t$ and $u_t = i_t\ell_t$, where $\{i_t\}$ is an IID sequence in which $\mathbf{P}\{i_t = 1\} = 1 - 0.25$ and $\{\ell_t\} \sim$ Laplace$(0, 2)$; (*iii*) Model: $y_t = \pi y_{t-1} + \theta y_{t-1}\{(1+\exp(-\gamma y_{t-1}))^{-1} - 1/2\} + u_t$ and $u_t = i_t\ell_t$, where $\{i_t\}$ is an IID sequence in which $\mathbf{P}\{i_t = 1\} = 1 - 0.5^2$ and $\{\ell_t\} \sim$ Laplace$(0, 2)$; and (*iv*) $\Gamma_1 = [0, 2]$, $\Gamma_2 = [0, 3]$, $\Gamma_3 = [0, 4]$, $\Gamma_4 = [0, 5]$.

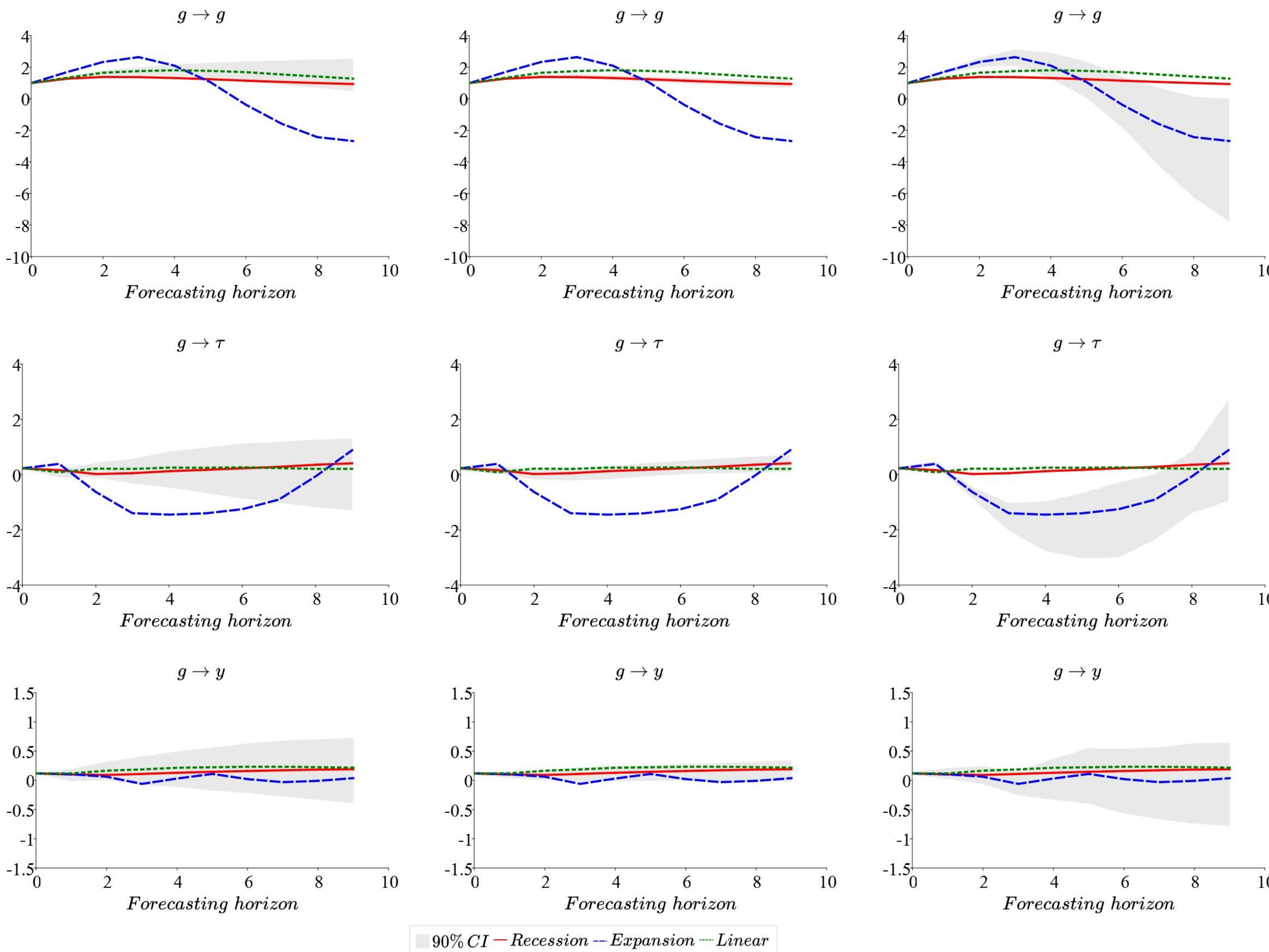


Figure 5: IMPULSE RESPONSES IN LINEAR, EXPANSION, AND RECESSION STATES Notes: The first, second, and third columns show the impulse and response functions of g_t, τ_t, and y_t in response to a \$1 increase of government spending obtained by linear model, recession and expansion state parameters, respectively. Shaded region is the 90 percent confidence band. Long dashed and solid lines show the responses in expansion and recession. Dotted line shows the responses using the linear model.