Testing for the Sandwich-Form Covariance Matrix Applied to Quasi-Maximum Likelihood Estimation Using Economic and Energy Price Growth Rates

LIJUAN HUO

Department of Economics, Beijing Institute of Technology, Haidian, Beijing, 100081, China

hlj8913@msn.com

JIN SEO CHO

Department of Economics, Beijing Institute of Technology, Haidian, Beijing, 100081, China
and
School of Economics, Yonsei University, Seodaemun, Seoul, 03722, Korea

jinseocho@yonsei.ac.kr

September 12, 2019

Abstract

This study aims to directly test for the sandwich-form asymptotic covariance matrix entailed by conditional heteroskedasticity and autocorrelation in the regression error. Given that none of the conditional heteroskedastic or autocorrelated regression errors yield the sandwich-form asymptotic covariance matrix for the least squares estimator, it is not necessary to estimate the asymptotic covariance matrix using the heteroskedasticity-consistent (HC) or heteroskedasticity and autocorrelation-consistent (HAC) covariance matrix estimator. Because of this fact, we first examine testing for the sandwich-form asymptotic covariance matrix before applying the HC or HAC covariance matrix estimator. For this goal, we apply the testing methodologies proposed by Cho and White (2015) and Cho and Phillips (2018) to fit the context of this study by extending the scope of their maximum test statistic to have greater power and further establishing a methodology to sequentially detect the influence of heteroskedastic and autocorrelated regression errors on the asymptotic covariance matrix. We affirm the theory on the test statistics of this study through a simulation and further apply our test statistics to economic and energy price growth rate data for illustrative purposes.

Key Words: Information matrix equality; sandwich-form covariance matrix; heteroskedasticity-consistent covariance matrix estimator; heteroskedasticity and autocorrelation-consistent covariance matrix estimator; economic growth rate; energy price growth rate.

JEL Codes: C12, C22, O47, G17, Q47.

Acknowledgements: The authors benefitted from discussions with the seminar participants at BIT. The responsibility for any errors and shortcomings in this work remains ours. Cho is grateful for the research grant by the Ministry of Education in the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5A2A01035256). Huo is grateful for the research grant support by National Natural Science Foundation of China (No.7183009) and MOE Project of Humanities and Sciences (No. 17YJC790057).

1 Introduction

It is critical to estimate the correct form of the asymptotic covariance matrix for the least squares (LS) estimator. For example, if the regression error exhibits conditional homoskedastic and martingale difference sequence (MDS) properties, the information matrix equality holds, so that the standard t-test statistic can be used to infer the unknown parameters. Unless otherwise imposed, the asymptotic covariance matrix is generally known to have a sandwich form, letting the standard t-ratio be redefined by the heteroskedasticity-consistent (HC) or heteroskedasticity and autocorrelation-consistent (HAC) covariance matrix estimator.

Nevertheless, conditionally heteroskedastic and/or autocorrelated errors do not necessarily yield the sandwich-form asymptotic covariance matrix. Conditionally heteroskedastic and/or autocorrelated error properties are the only necessary conditions to yield a sandwich-form asymptotic covariance matrix, implying that if the information matrix equality holds even under conditionally heteroskedastic and/or autocorrelated error properties, we do not have to estimate the asymptotic covariance matrix using the HC or HAC covariance matrix estimator. Given that the HC or HAC covariance matrix estimator is not the most efficient estimator under the information matrix equality, the empirical researcher should use the standard t-test statistic to infer the unknown parameters.

The main goal of this study is therefore to provide a testing methodology for a sandwich-form asymptotic covariance matrix instead of testing heteroskedastic and/or autocorrelated regression errors indirectly. Another goal of this study is to provide an empirical data analysis using the proposed testing methodology in classical regression and time-series data analysis settings. This serves the dual purpose of illustrating our method, and, as it turns out, reinforcing the inferencing process on regression analysis.

To achieve this goal, we apply the methodologies proposed by Cho and White (2015) and Cho and Phillips (2018). In the earlier literature, they provide simple necessary and sufficient conditions for the equality of two positive-definite matrices using the trace and determinant of a matrix ratio and further extend them to develop testing methodologies. In the current study, we apply their methodologies to test for a sandwich-form asymptotic covariance matrix entailed by conditional heteroskedastic or autocorrelated regression errors. Specifically, we first generalize the maximum test statistic in Cho and Phillips (2018) in our contexts, so that our test statistics have greater power than the maximum test statistic in Cho and Phillips (2018), where we separately test for the sandwich-form asymptotic covariance matrix entailed by conditional heteroskedastic and/or autocorrelated errors. Although our maximum test statistics are not the most powerful test statistics for any sandwich-form asymptotic covariance matrix, they are omnibus test statistics that have consistent power against any sandwich form of the asymptotic covariance matrix.

As a further step to achieve these goals, we develop a sequential testing procedure (STP) by implementing our maximum test statistics in a systematic way. Given that our maximum test statistics are separately constructed to test for the sandwich-form asymptotic covariance matrices entailed by conditional heteroskedasticity and/or autocorrelated errors, we further examine a suitable order to apply the test statistics, so that the empirical researcher can control the overall type-I error of applying these two test statistics. Using this STP, (s)he can then consistently estimate the status of the asymptotic covariance matrix, so that (s)he can be assured of using a suitable version of the t-ratio to infer the unknown parameters.

The testing methodology in this study is highly utilized if the maximum test statistics are combined with the resampling bootstraps. In particular, the residual bootstrap and wild bootstrap methods of Efron and Tibsharani (1988) and Wu (1986) are useful for raising the utility of the maximum test statistics. Although the asymptotic null limit distributions are provided for the maximum test statistics in this study, they may lead to relatively large size distortions when the sample size is relatively small. This fact is mainly due to the large degrees of freedom of the maximum test statistics. Many components in the two matrices have to be compared to claim the information matrix equality, leading to the large degrees of freedom. We overcome this size distortion problem by applying the bootstrap methods to the maximum test statistics. Indeed, the resampling methods turn out to be efficient remedies for the size distortion problem, controlling the type-I errors successfully. Simulation results are also provided as evidence to affirm our theory on the maximum test statistics and STP.

The empirical applicability of the maximum test statistics is potentially huge given that many cross-sectional and time-series data analyses may have to estimate a sandwich-form asymptotic covariance matrix. For the second goal of this study to illustrate the use of the testing methodology, we revisit empirical studies of economic growth and finance. The standard t-test statistic is typically computed by assuming conditional homoskedasticity for cross-sectional data, and this practice is popularly found in empirical studies. For example, Mankiw, Romer, and Weil (1992) estimate Solow's (1957) economic growth model and test the conditional convergence hypothesis using the standard t-ratio statistic. Using one of our maximum test statistics, we test the sandwich-form asymptotic covariance matrix implied by the conditionally heteroskedastic regression error and investigate whether their inference results are modified by the maximum test statistic. As another illustration, we examine a time-series model for the growth rate of energy prices. In the finance literature, it is standard to estimate conditional mean and variance models simultaneously. It is also usual to estimate the model using autoregressive (AR) and generalized autoregressive conditional heteroskedastic (GARCH) models for the conditional mean and variance, respectively. For example, Bai and Lam (2019) estimate the growth rates of energy prices using AR and GARCH models after supposing a skewed t-distribution for the standardized error, so that their model can be estimated using maximum likelihood estimation. Furthermore, they deepen their investigation of the marginal distribution for each energy variable by further estimating the dependence structure between the growth rates of energy prices through copula estimation. Using our maximum test statistics, we revisit their data analysis and examine whether their model assumption can be reinforced by the maximum test statistics.

The plan of the remainder of this paper is as follows. Section 2 specifically discusses the motivation of this study and provides the theories and definitions of our maximum test statistics along with their null, asymptotic, and local alternative limit behaviors. We also develop the STP to estimate the status of the asymptotic covariance matrix. Section 3 provides the simulation evidence of our theory and Section 4 illustrates the use of our maximum test statistics using two empirical examples. We provide concluding remarks in Section 5 and collect the mathematical proofs in the Appendix.

2 Testing for the Influence of Conditional Heteroskedasticity and Autocorrelation

In this section, we discuss the motivation of our study and the maximum test statistics used to achieve its goal. We also provide the null, alternative, and local alternative limit behaviors of the test statistics, so that they can be systematically structured to estimate the status of the asymptotic covariance matrix using an STP.

2.1 Motivation and Hypotheses

Suppose that the empirical researcher wishes to estimate a parametric model for a conditional equation. That is, for a stationary process $(Y_t, \mathbf{X}_t) \in \mathbb{R}^{1+k}$, the researcher specifies a parametric model as follows:

$$\mathcal{M} := \{m(\cdot, \boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^d\}$$

to estimate the conditional mean of Y_t on $\mathbf{X}_t$ such that $\mathcal{M}$ satisfies the standard regularity conditions for the LS estimation (e.g., White and Domowitz, 1984). We further suppose that $\mathbf{X}_t$ may possess lagged dependent Y_t and that $\mathcal{M}$ is correctly specified for $\mathbb{E}[Y_t|\mathbf{X}_t]$, viz., for some $\boldsymbol{\theta}_* \in \boldsymbol{\Theta}$, $m(\mathbf{X}_t, \boldsymbol{\theta}_*) = \mathbb{E}[Y_t|\mathbf{X}_t]$ with probability 1.

The conditional mean equation can be consistently estimated using quasi-maximum likelihood (QML) estimation. If the model for the conditional mean equation is correct and the assumed distribution for the QML estimation belongs to the linear exponential family, Gourieroux, Monfort, and Trognon (1984) show that the conditional mean equation can be consistently estimated using the QML estimator that asymptotically follows a normal distribution.

To motivate the main goal of this study, we consider the following model environment. If the model $\mathcal{M}$ is estimated using QML estimation, we may suppose that the QML estimator denoted by $\widehat{\boldsymbol{\theta}}_n$ converges to $\boldsymbol{\theta}_* \in \boldsymbol{\Theta}$ by supposing regularity conditions (e.g., White, 1982). We denote the error by U_t, viz.,

$$U_t := Y_t - m(\mathbf{X}_t, \boldsymbol{\theta}_*).$$

As established in the literature, the asymptotic behavior of the QML estimator is critically dependent upon the properties of U_t and $\mathcal{M}$. For example, if U_t is conditionally heteroskedastic on $\mathbf{X}_t$, the asymptotic variance of $\widehat{\boldsymbol{\theta}}_n$ is different from that obtained by assuming conditional homoskedasticity (e.g., White, 1980). This fact implies that the inferencing procedure on $\boldsymbol{\theta}_*$ has to be differently conducted depending on the properties of U_t and $\mathcal{M}$. We can classify the properties into the following two categories. First, we let $\mathcal{F}_t$ be the smallest sigma field generated by $\{\mathbf{X}_t, U_{t-1}, \mathbf{X}_{t-1}, U_{t-2}, \ldots\}$ and suppose that $\{U_t, \mathcal{F}_t\}$ forms an MDS. With conditional homoskedasticity, viz., for some finite $\sigma_*^2 \in \mathbb{R}^+$, $\mathbb{E}[U_t^2|\mathcal{F}_t] = \sigma_*^2$, it is straightforward to obtain that

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_*) \overset{\text{A}}{\sim} N(\mathbf{0}, \mathbf{A}_*^{-1});$$

unless otherwise imposed, it follows that

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_*) \overset{\text{A}}{\sim} N(\mathbf{0}, \mathbf{A}_*^{-1}\mathbf{B}_*\mathbf{A}_*^{-1}), \tag{1}$$

where $\mathbf{A}_*$ and $\mathbf{B}_*$ are the limit of the Hessian matrix of the quasi-log-likelihood function and asymptotic covariance matrix of the scores evaluated at $\boldsymbol{\theta}_*$. More specifically, if we let $\ell_t(\cdot)$ be the quasi-likelihood function of individual observations, $\mathbf{A}_* := -\mathbb{E}[\nabla^2_{\boldsymbol{\theta}}\ell_t(\boldsymbol{\theta}_*)]$ and $\mathbf{B}_* := \mathbb{E}[\boldsymbol{s}_t\boldsymbol{s}_t']$, where $\boldsymbol{s}_t$ denotes the score at $\boldsymbol{\theta}_*$, namely $\boldsymbol{s}_t := \nabla_{\boldsymbol{\theta}}\ell_t(\boldsymbol{\theta}_*)$. For example, if the normal distribution is assumed for the quasi-likelihood function and $m(\mathbf{X}_t, \boldsymbol{\theta}_*) = \mathbf{X}_t'\boldsymbol{\theta}_*$,

$$\mathbf{A}_* = \sigma_*^{-2}\mathbb{E}[\mathbf{X}_t\mathbf{X}_t'] \quad \text{and} \quad \mathbf{B}_* = \sigma_*^{-4}\mathbb{E}[U_t^2\mathbf{X}_t\mathbf{X}_t'],$$

so that the asymptotic covariance matrix of $\widehat{\boldsymbol{\theta}}_n$ is $\mathbf{A}_*^{-1} = \sigma_*^2\mathbb{E}[\mathbf{X}_t\mathbf{X}_t']^{-1}$ under the conditional homoskedastic error assumption. On the contrary, the asymptotic covariance matrix becomes

$$\mathbf{A}_*^{-1}\mathbf{B}_*\mathbf{A}_*^{-1} = \mathbb{E}[\mathbf{X}_t\mathbf{X}_t']^{-1}\mathbb{E}[U_t^2\mathbf{X}_t\mathbf{X}_t']\mathbb{E}[\mathbf{X}_t\mathbf{X}_t']^{-1}$$

under the conditional heteroskedastic error assumption. From this aspect, it is now clear that if the information matrix equality does not hold for the asymptotic covariance matrix of the QML estimator, namely $\mathbf{A}_* \neq \mathbf{B}_*$, it must follow that U_t exhibits conditional heteroskedasticity under the maintained regularity conditions. Therefore, the inference on $\boldsymbol{\theta}_*$ has to be conducted using the asymptotic distribution in (1).

In the literature, a number of different test statistics are popularly used to test for conditional heteroskedasticity. Among others, Godfrey (1978b), White (1980), Breusch and Pagan (1979), and Engle's (1982) conditional heteroskedasticity test statistics are popularly used for empirical studies under various data assumptions. These test statistics are obtained by assuming a particular form of conditional heteroskedasticity; however, they are defined to have consistent power against a broad class of conditional heteroskedasticity even though they lack omnibus power against any form of heteroskedasticity. It is also standard to estimate the asymptotic covariance matrix using the HC covariance matrix estimator if the conditional homoskedasticity hypothesis is rejected by the test statistics.

Second, if $\{U_t, \mathcal{F}_t\}$ does not form an MDS, another form of information matrix inequality holds. If U_t is autocorrelated, the asymptotic covariance matrix $\mathbf{B}_*$ in (1) is modified to

$$\mathbf{C}_* = \mathbf{B}_* + \lim_{n\to\infty}\frac{1}{n}\sum_{\substack{t=1\\t\neq\tau}}^{n}\sum_{\tau=1}^{n}\mathbb{E}[\boldsymbol{s}_t\boldsymbol{s}_\tau'],$$

so that the limit distribution of the QML estimator is accordingly modified as follows:

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_*) \overset{\text{A}}{\sim} N(\mathbf{0}, \mathbf{A}_*^{-1}\mathbf{C}_*\mathbf{A}_*^{-1}). \tag{2}$$

The second term on the right-hand side of $\mathbf{C}_*$ vanishes to zero if U_t is an MDS, which does not necessarily hold if U_t

is autocorrelated. From this aspect, we can conclude that if $\mathbf{B}_* \neq \mathbf{C}_*$, U_t must be autocorrelated and the inference on $\boldsymbol{\theta}_*$ has to be conducted using the asymptotic distribution in (2), thus allowing a suitable test statistic to be defined by estimating $\mathbf{C}_*$ consistently.

In the literature, a number of test statistics are available to test for autocorrelation. For example, Durbin and Watson (1950, 1951), Breusch (1978), Godfrey (1978a), and Ljung and Box (1978) are popularly used for autocorrelation as well as developed by supposing a particular form of autocorrelation as for the test statistics for conditional heteroskedasticity. In addition, there are non-parametric test statistics for autocorrelation. To mention a few, Robinson (1991), Skaug and Tjøstheim (1996), Hong and White (2005), and Cho and White (2011) provide non-parametric test statistics for autocorrelation. When the autocorrelation hypothesis is rejected by these test statistics, it is standard to estimate the asymptotic covariance matrix using the HAC covariance matrix estimator.

As we see from the test statistics in the literature, testing procedures are developed to indirectly test for conditional heteroskedasticity and/or autocorrelation in U_t. Nevertheless, any form of conditional heteroskedasticity or autocorrelation does not necessarily lead to the sandwich-form asymptotic covariance matrix. Conditional heteroskedasticity and autocorrelation are just two of the necessary conditions for the sandwich-form asymptotic covariance matrix. That is, the presence of conditional heteroskedasticity and/or autocorrelation in U_t does not necessarily lead to the sandwich-form asymptotic covariance matrix, so that the empirical researcher may not have to estimate the asymptotic covariance matrix using the HC or HAC covariance matrix estimator if the information matrix equality holds even with conditionally heteroskedastic and autocorrelated regression errors.

This aspect motivates us to test for the sandwich-form asymptotic covariance matrix hypothesis instead of indirectly testing for conditional heteroskedastic and/or autocorrelated regression errors. For this goal, we view the following as the main hypotheses of interest in the current study:

$$\mathcal{H}_0^{(1)} : \mathbf{A}_* = \mathbf{B}_* \quad \text{versus} \quad \mathcal{H}_1^{(1)} : \mathbf{A}_* \neq \mathbf{B}_*; \quad \text{and}$$

$$\mathcal{H}_0^{(2)} : \mathbf{B}_* = \mathbf{C}_* \quad \text{versus} \quad \mathcal{H}_1^{(2)} : \mathbf{B}_* \neq \mathbf{C}_*.$$

$\mathcal{H}_0^{(1)}$ and $\mathcal{H}_1^{(1)}$ are provided to test for the sandwich-form asymptotic covariance matrix entailed by conditional heteroskedastic errors, whereas $\mathcal{H}_0^{(2)}$ and $\mathcal{H}_1^{(2)}$ are investigated to test for the sandwich-form covariance matrix entailed by autocorrelated errors. Each hypothesis characterizes the status of the asymptotic covariance matrix, and a suitable test statistic can be applied according to the status. For example, if $\mathcal{H}_0^{(2)}$ turns out to be a suitable status for the asymptotic covariance matrix, we can estimate the asymptotic covariance matrix using the HC covariance matrix estimator.

2.2 Maximum Test Statistics

The test statistics we examine in this study are developed from the methodologies presented by Cho and White (2014) and Cho and Phillips (2018). They provide simple necessary and sufficient conditions for the equality of two symmetric positive-definite matrices. For example, when we test $\mathcal{H}_0^{(1)}$ against $\mathcal{H}_1^{(1)}$, any two of the following conditions hold:

$$\mathrm{tr}[\mathbf{D}_*] = d; \quad \det[\mathbf{D}_*] = 1; \quad \text{and} \quad \mathrm{tr}[\mathbf{D}_*^{-1}] = d,$$

if and only if $\mathbf{A}_* = \mathbf{B}_*$, where $\mathbf{D}_* := \mathbf{A}_*\mathbf{B}_*^{-1}$. It is also possible to reverse the role of $\mathbf{A}_*$ with that of $\mathbf{B}_*$ to define $\mathbf{D}_*$, so that the same equality holds between $\mathbf{A}_*$ and $\mathbf{B}_*$ if and only if any two of the following conditions hold:

$$\mathrm{tr}[\widetilde{\mathbf{D}}_*] = d; \quad \det[\widetilde{\mathbf{D}}_*^{-1}] = 1; \quad \text{and} \quad \mathrm{tr}[\widetilde{\mathbf{D}}_*^{-1}] = d,$$

where $\widetilde{\mathbf{D}}_* := \mathbf{B}_*\mathbf{A}_*^{-1}$. Using these equivalent conditions, these authors define a number of auxiliary omnibus test statistics to test the equality of two symmetric positive-definite matrices, recommending the most powerful test statistic as the maximum of the auxiliary test statistics.

We apply their testing methodology to our testing problem. For this purpose, we define $\boldsymbol{\xi} := (\boldsymbol{\theta}', \sigma^2)'$ and suppose that $\boldsymbol{\xi}_* := (\boldsymbol{\theta}'_*, \sigma_*^2)'$ is consistently estimated using the QML estimator $\widehat{\boldsymbol{\xi}}_n := (\widehat{\boldsymbol{\theta}}'_n, \widehat{\sigma}_n^2)'$:

$$\widehat{\boldsymbol{\xi}}_n := \arg\max_{\boldsymbol{\xi}\in\boldsymbol{\Xi}} \bar{L}_n(\boldsymbol{\xi}), \quad \text{where} \quad \bar{L}_n(\cdot) := \frac{1}{n}\sum_{t=1}^{n} \ell_t(\cdot),$$

and $\ell_t(\cdot)$ is the individual quasi-log-likelihood function defined on $\boldsymbol{\Xi} := \boldsymbol{\Theta} \times [c, d]$ (c and $d \in \mathbb{R}^+$). Here, n denotes the sample size. A particular form of quasi-log-likelihood function is assumed, so that it can consistently estimate $\boldsymbol{\theta}_*$ and the unconditional variance of U_t denoted by σ_*^2. For example, the normal likelihood function belongs to this case. We further suppose that $\mathbf{A}_*$ and the associated covariance matrices can be consistently estimated using $\widehat{\mathbf{A}}_n$, $\widehat{\mathbf{B}}_n$, and $\widehat{\mathbf{C}}_n$, respectively, viz.,

$$\widehat{\mathbf{A}}_n \xrightarrow{\mathbb{P}} \mathbf{A}_*, \quad \widehat{\mathbf{B}}_n \xrightarrow{\mathbb{P}} \mathbf{B}_*, \quad \text{and} \quad \widehat{\mathbf{C}}_n \xrightarrow{\mathbb{P}} \mathbf{C}_*.$$

For example, we can let

$$\widehat{\mathbf{A}}_n := -\frac{1}{n}\sum_{t=1}^{n} \nabla_{\boldsymbol{\theta}}^2 \ell_t(\widehat{\boldsymbol{\xi}}_n), \quad \widehat{\mathbf{B}}_n := \frac{1}{n}\sum_{t=1}^{T} \widehat{\boldsymbol{s}}_t\widehat{\boldsymbol{s}}'_t, \quad \text{and} \quad \widehat{\mathbf{C}}_n := \frac{1}{n}\sum_{t=1}^{T} \widehat{\boldsymbol{s}}_t\widehat{\boldsymbol{s}}'_t + \frac{1}{n}\sum_{k=1}^{\ell} \omega_{\ell k} \sum_{t=k+1}^{n} \left(\widehat{\boldsymbol{s}}_{t-k}\widehat{\boldsymbol{s}}'_t + \widehat{\boldsymbol{s}}_t\widehat{\boldsymbol{s}}'_{t-k}\right),$$

where $\widehat{\boldsymbol{s}}_t := \nabla_{\boldsymbol{\theta}} \ell_t(\widehat{\boldsymbol{\xi}}_n)$ and $\omega_{\ell k}$ and ℓ are the kernel and bandwidth used to estimate $\mathbf{C}_*$ consistently; moreover, $\widehat{\mathbf{B}}_n$ is White's (1980) HC covariance matrix estimator, and a number of different kernels and bandwidths can be employed for $\widehat{\mathbf{C}}_n$. For example, we can let $\omega_{\ell k} := 1 - k/(1+\ell)$ and $\ell = O(n^{1/3})$ to employ Newey and West's (1987) HAC covariance matrix estimator. As another example, Andrews (1991) recommends using the quadratic spectral kernel:

$$\omega_{\ell k} := \frac{25}{12\pi^2 x^2}\left\{\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5)\right\} \tag{3}$$

with $x := k/(1+\ell)$ and $\ell = O(n^{1/5})$. Indeed, there are many other kernels and bandwidths for the HAC estimation of $\mathbf{C}_*$ (e.g., Gallant, 1987; Ng and Perron, 1996).

Using these estimators, we define the auxiliary test statistics as in Cho and White (2014) and Cho and Phillips

(2018). For example, if $\mathcal{H}_0^{(1)}$ is tested against $\mathcal{H}_1^{(1)}$, we let $\widehat{\mathbf{D}}_n := \widehat{\mathbf{A}}_n\widehat{\mathbf{B}}_n^{-1}$ and define the following statistics:

$$\widehat{\tau}_n := \frac{1}{d}\mathrm{tr}[\widehat{\mathbf{D}}_n] - 1, \quad \widehat{\eta}_n := \frac{d}{\mathrm{tr}[\widehat{\mathbf{D}}_n^{-1}]} - 1, \quad \widehat{\delta}_n := \det[\widehat{\mathbf{D}}_n]^{\frac{1}{d}} - 1, \quad \widehat{\gamma}_n := \widehat{\delta}_n - \widehat{\eta}_n, \quad \text{and} \quad \widehat{\zeta}_n := \widehat{\tau}_n - \widehat{\delta}_n.$$

These statistics are introduced to consistently estimate

$$\dot{\tau}_* := \frac{1}{d}\mathrm{tr}[\mathbf{D}_*] - 1, \quad \dot{\eta}_* := \frac{d}{\mathrm{tr}[\mathbf{D}_*^{-1}]} - 1, \quad \dot{\delta}_* := \det[\mathbf{D}_*]^{\frac{1}{d}} - 1, \quad \dot{\gamma}_* := \dot{\delta}_* - \dot{\eta}_*, \quad \text{and} \quad \dot{\zeta}_* := \dot{\tau}_* - \dot{\delta}_*,$$

respectively.

The motivations of these statistics stem from the fact that $d^{-1}\mathrm{tr}[\widehat{\mathbf{D}}]$, $\det[\widehat{\mathbf{D}}]^{\frac{1}{d}}$, and $d\mathrm{tr}[\widehat{\mathbf{D}}_n^{-1}]^{-1}$ estimate the arithmetic, geometric, and harmonic means of the eigenvalues of $\mathbf{D}_*$. If any two of these are equal to unity, $\mathbf{D}_*$ must be $\mathbf{I}_d$, and its converse is also true because any two of the means are equal to each other if and only if all the eigenvalues are identical. Therefore, if any two of the above five auxiliary statistics estimate zero consistently, this implies that $\mathbf{D}_* = \mathbf{I}_d$.

In addition to these auxiliary test statistics, we let $\widetilde{\mathbf{D}}_n := \widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n^{-1}$ to test $\mathcal{H}_0^{(1)}$ against $\mathcal{H}_1^{(1)}$ and define the following test statistics:

$$\widetilde{\tau}_n := \frac{1}{d}\mathrm{tr}[\widetilde{\mathbf{D}}_n] - 1, \quad \widetilde{\eta}_n := \frac{d}{\mathrm{tr}[\widetilde{\mathbf{D}}_n^{-1}]} - 1, \quad \widetilde{\delta}_n := \det[\widetilde{\mathbf{D}}_n]^{\frac{1}{d}} - 1, \quad \widetilde{\gamma}_n := \widetilde{\delta}_n - \widetilde{\eta}_n, \quad \text{and} \quad \widetilde{\zeta}_n := \widetilde{\tau}_n - \widetilde{\delta}_n,$$

where the roles of $\mathbf{A}_*$ and $\mathbf{B}_*$ are reversed from those used to define $\mathbf{D}_*$. As before, these additional statistics are introduced to estimate

$$\ddot{\tau}_* := \frac{1}{d}\mathrm{tr}[\mathbf{D}_*^{-1}] - 1, \quad \ddot{\eta}_* := \frac{d}{\mathrm{tr}[\mathbf{D}_*]} - 1, \quad \ddot{\delta}_* := \det[\mathbf{D}_*^{-1}]^{\frac{1}{d}} - 1, \quad \ddot{\gamma}_* := \ddot{\delta}_* - \ddot{\eta}_*, \quad \text{and} \quad \ddot{\zeta}_* := \ddot{\tau}_* - \ddot{\delta}_*,$$

consistently, respectively.

In a different environment, $\widehat{\mathbf{D}}_n$ and $\widetilde{\mathbf{D}}_n$ can be redefined to test the hypotheses. If the researcher wishes to test $\mathcal{H}_0^{(2)}$ against $\mathcal{H}_1^{(2)}$, we can let $\widehat{\mathbf{D}}_n := \widehat{\mathbf{B}}_n\widehat{\mathbf{C}}_n^{-1}$ and $\widetilde{\mathbf{D}}_n := \widehat{\mathbf{C}}_n\widehat{\mathbf{B}}_n^{-1}$ and obtain the statistics accordingly using the new $\widehat{\mathbf{D}}_n$ and $\widetilde{\mathbf{D}}_n$.

The maximum test statistic in this study is defined using the auxiliary statistics. Specifically, we let our test statistic be defined as follows:

$$\widehat{\mathfrak{M}}_n := \max_{j=1,2}\left[\widehat{\mathfrak{B}}_{j,n}, \widehat{\mathfrak{S}}_{j,n}, \widehat{\mathfrak{E}}_{j,n}, \widetilde{\mathfrak{B}}_{j,n}, \widetilde{\mathfrak{S}}_{j,n}, \widetilde{\mathfrak{E}}_{j,n}\right],$$

where

$$\widehat{\mathfrak{B}}_{1,n} := \frac{nd}{2}(\widehat{\tau}_n^2 + 2\widehat{\zeta}_n), \quad \widehat{\mathfrak{B}}_{2,n} := \frac{nd}{2}(\widehat{\delta}_n^2 + 2\widehat{\zeta}_n), \quad \widehat{\mathfrak{S}}_{1,n} := \frac{nd}{2}(\widehat{\delta}_n^2 + 2\widehat{\gamma}_n), \quad \widehat{\mathfrak{S}}_{2,n} := \frac{nd}{2}(\widehat{\eta}_n^2 + 2\widehat{\gamma}_n),$$

$$\widehat{\mathfrak{E}}_{1,n} := \frac{nd}{2}(\widehat{\tau}_n^2 + 2\widehat{\gamma}_n), \quad \text{and} \quad \widehat{\mathfrak{E}}_{2,n} := \frac{nd}{2}(\widehat{\eta}_n^2 + 2\widehat{\zeta}_n).$$

For $j = 1$ and 2, we also let $\widetilde{\mathfrak{B}}_{j,n}$, $\widetilde{\mathfrak{S}}_{j,n}$, and $\widetilde{\mathfrak{E}}_{j,n}$ be similarly defined using $\widetilde{\tau}_n$, $\widetilde{\eta}_n$, $\widetilde{\delta}_n$, $\widetilde{\gamma}_n$, and $\widetilde{\zeta}_n$.

The auxiliary test statistics constituting $\widehat{\mathfrak{M}}_n$ are defined by applying Wald's (1943) test principle, and $\widehat{\mathfrak{M}}_n$ is designed to have the maximum power out of the 12 auxiliary test statistics. For example, $\widehat{\tau}_n$ estimates the distance between the arithmetic average of the eigenvalues of $\widehat{\mathbf{D}}_n$ and unity, and $\widehat{\zeta}_n$ estimates the distance between $\widehat{\tau}_n$ and $\widehat{\delta}_n$ that estimates the distance between the geometric average of the eigenvalues of $\widehat{\mathbf{D}}_n$ and unity, so that $\widehat{\mathfrak{B}}_{1,n}$ now tests whether the arithmetic average of the eigenvalues of $\mathbf{D}_*$ is unity and also equal to the geometric average of the same eigenvalues. Likewise, for j = 1 and 2, the auxiliary test statistics $\widehat{\mathfrak{B}}_{j,n}$, $\widehat{\mathfrak{S}}_{j,n}$, $\widehat{\mathfrak{E}}_{j,n}$, $\widetilde{\mathfrak{B}}_{j,n}$, $\widetilde{\mathfrak{S}}_{j,n}$, and $\widetilde{\mathfrak{E}}_{j,n}$ test whether any two of the arithmetic, geometric, and harmonic averages of $\mathbf{D}_*$ and $\mathbf{D}_*^{-1}$ are equal to unity using Wald's (1943) test principle. Furthermore, Cho and Phillips (2018) show that the null limit behaviors of the auxiliary statistics are equivalent, so that the null limit behavior of $\widehat{\mathfrak{M}}_n$ turns out to be identical to that of each auxiliary test statistic.

The maximum test statistic $\widehat{\mathfrak{M}}_n$ is different from that in Cho and Phillips (2018). Although our test statistic is constructed by following their approach, their maximum test statistic is defined as the maximum of only $\widehat{\mathfrak{B}}_{j,n}$, $\widehat{\mathfrak{S}}_{j,n}$, and $\widehat{\mathfrak{E}}_{j,n}$ with j = 1 and 2. They do not explicitly examine the role of $\widetilde{\mathbf{D}}_n$ when defining their maximum test statistic. Accommodating the roles of both $\widehat{\mathbf{D}}_n$ and $\widetilde{\mathbf{D}}_n$, we define $\widehat{\mathfrak{M}}_n$ and expect more powerful behavior from $\widehat{\mathfrak{M}}_n$ than that in Cho and Phillips (2018) in addition to being an omnibus test statistic.

When defining the maximum test statistic $\widehat{\mathfrak{M}}_n$, not every auxiliary test statistic is used to define the maximum test statistic. For example, the researcher may define another auxiliary test statistic by applying Wald's test principle to $\widehat{\tau}_n$ and the distance between $\widehat{\tau}_n$ and $\widehat{\eta}_n$, so that it can contribute to $\widehat{\mathfrak{M}}_n$. Nevertheless, Cho and Phillips (2018, remarks below theorem 3) show that the other auxiliary test statistics not considered here have asymptotic power inferior to the auxiliary test statistics contributing to $\widehat{\mathfrak{M}}_n$. We therefore focus only on $\widehat{\mathfrak{B}}_{j,n}$, $\widehat{\mathfrak{S}}_{j,n}$, $\widehat{\mathfrak{E}}_{j,n}$, $\widetilde{\mathfrak{B}}_{j,n}$, $\widetilde{\mathfrak{S}}_{j,n}$, and $\widetilde{\mathfrak{E}}_{j,n}$ to define the maximum test statistics.

Hereafter, we denote the maximum test statistics and their auxiliary test statistics using a superscript to represent the hypothesis tested by the maximum test statistics and avoid any confusion. That is, $\widehat{\mathfrak{M}}_n^{(1)}$ and $\widehat{\mathfrak{M}}_n^{(2)}$ denote the maximum test statistics designed to test $\mathcal{H}_0^{(1)}$ against $\mathcal{H}_1^{(1)}$ and $\mathcal{H}_0^{(2)}$ against $\mathcal{H}_1^{(2)}$, respectively.

Before examining the limit behaviors of the maximum test statistics, we formally state the regularity conditions for the maximum test statistics by applying those in Cho and Phillips (2018). For an efficient provision of the regularity conditions, for each i = 1 and 2, we let $\mathbf{P}^{(i)}(\cdot)$, $\bar{\mathbf{P}}^{(i)}(\cdot)$, and $\mathbf{P}_n^{(i)}(\cdot)$ denote the functions corresponding to the left-hand side matrix in $\mathcal{H}_0^{(i)}$. For example, we let $\mathbf{P}^{(1)}(\boldsymbol{\xi}_*) = \mathbf{A}_*$ under $\mathcal{H}_0^{(1)}$, which is the left-hand side matrix in $\mathcal{H}_0^{(1)}$. Likewise, we let $\mathbf{Q}^{(i)}(\cdot)$, $\bar{\mathbf{Q}}^{(i)}(\cdot)$, and $\mathbf{Q}_n^{(i)}(\cdot)$ denote the functions corresponding to the right-hand side matrix in $\mathcal{H}_0^{(i)}$. For each i = 1 and 2, we further let $\widehat{\mathbf{P}}_n^{(i)} := \mathbf{P}_n(\widehat{\boldsymbol{\xi}}_n)$ and $\widehat{\mathbf{Q}}_n^{(i)} := \mathbf{Q}_n(\widehat{\boldsymbol{\xi}}_n)$, so that, for example, $\widehat{\mathbf{P}}_n^{(1)} = \widehat{\mathbf{A}}_n$ and $\widehat{\mathbf{Q}}_n^{(1)} = \widehat{\mathbf{B}}_n$ under $\mathcal{H}_0^{(1)}$.

Assumption 1. *(i)* $\{\mathbf{Z}_t := (Y_t, \mathbf{X}_t')' \in \mathbb{R}^{(1+k)} : t = 1, 2, \ldots\}$ *is a strictly stationary process defined on a complete probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *with* $k \in \mathbb{N}$*;*

(ii) $\boldsymbol{\Xi} := \boldsymbol{\Theta} \times [b, c]$ *is a compact and convex set with a non-empty interior, and* $\boldsymbol{\Theta} \subset \mathbb{R}^d$ *with* $d \in \mathbb{N}$*;*

(iii) The sequence of $\{\widehat{\boldsymbol{\xi}}_n := (\widehat{\boldsymbol{\theta}}_n', \widehat{\sigma}_n^2)' : \Omega \mapsto \boldsymbol{\Xi}\}$ *converges to* $\boldsymbol{\xi}_* := (\boldsymbol{\theta}_*', \sigma_*^2)' \in \boldsymbol{\Xi}$ *in probability, where* $\boldsymbol{\xi}_*$ *is unique and interior to* $\boldsymbol{\Xi}$*,*

$$\widehat{\boldsymbol{\xi}}_n := \arg\max_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} \bar{L}_n(\boldsymbol{\xi}) \quad \textit{with} \quad \bar{L}_n(\boldsymbol{\xi}) := \frac{1}{n} \sum_{t=1}^{n} \ell_t(\boldsymbol{\xi}),$$

and $\ell_t(\cdot) := \ell(\mathbf{Z}_t, \cdot)$ *is the quasi-log-likelihood function assuming a correctly specified model* $\mathcal{M} := \{m(\cdot, \boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^d\}$ *for* $E[Y_t|\mathbf{X}_t]$ *and a constant conditional variance such that* $\ell_t(\cdot)$ *is continuous on* $\boldsymbol{\Xi}$ *almost surely* $-\mathbb{P}$*;*

(iv) For each $i = 1$ *and* 2*, there are* $\mathbf{P}_n^{(i)}(\cdot) : \boldsymbol{\Xi} \mapsto \mathbb{R}^{d\times d}$ *and* $\mathbf{Q}_n^{(i)}(\cdot) : \boldsymbol{\Xi} \mapsto \mathbb{R}^{d\times d}$ *consistent for* $\mathbf{P}^{(i)}(\cdot)$ *and* $\mathbf{Q}^{(i)}(\cdot)$ *uniformly on* $\boldsymbol{\Xi}$*;*

(v) For each $i = 1$ *and* 2 *and for some positive-definite* $\boldsymbol{\Sigma}_*^{(i)}$*,* $\sqrt{n}[(\widehat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}_*), (\widehat{\mathbf{P}}_n^{(i)} - \mathbf{P}_*^{(i)}), (\widehat{\mathbf{Q}}_n^{(i)} - \mathbf{Q}_*^{(i)})] \Rightarrow (\mathbf{Z}_{\boldsymbol{\xi}}, \bar{\mathbf{P}}_*^{(i)} + \mathbf{Z}_{\mathbf{Q}^{(i)}}, \bar{\mathbf{Q}}_*^{(i)} + \mathbf{Z}_{\mathbf{Q}^{(i)}})$ *such that* $[\mathbf{Z}_{\boldsymbol{\xi}}', \mathrm{vech}(\mathbf{Z}_{\mathbf{P}^{(i)}})', \mathrm{vech}(\mathbf{Z}_{\mathbf{Q}^{(i)}})']' \sim N(\mathbf{0}, \boldsymbol{\Sigma}_*^{(i)})$*, where*

(v.i) $\widehat{\mathbf{P}}_n^{(i)} := \mathbf{P}_n^{(i)}(\widehat{\boldsymbol{\xi}}_n)$ *and* $\widehat{\mathbf{Q}}_n^{(i)} := \mathbf{Q}_n^{(i)}(\widehat{\boldsymbol{\xi}}_n)$*;*

(v.ii) $\mathbf{P}_*^{(i)} := \mathbf{P}^{(i)}(\boldsymbol{\xi}_*)$ *and* $\mathbf{Q}_*^{(i)} := \mathbf{Q}^{(i)}(\boldsymbol{\xi}_*)$ *are symmetric and positive definite such that* $\mathbf{P}^{(i)} : \boldsymbol{\Xi} \mapsto \mathbb{R}^{d\times d}$ *and* $\mathbf{Q}^{(i)} : \boldsymbol{\Xi} \mapsto \mathbb{R}^{d\times d}$ *are in* $\mathcal{C}^{(2)}(\boldsymbol{\Xi})$*; and*

(v.iii) $\bar{\mathbf{P}}_*^{(i)} := \bar{\mathbf{P}}^{(i)}(\boldsymbol{\xi}_*)$ *and* $\bar{\mathbf{Q}}_*^{(i)} := \bar{\mathbf{Q}}^{(i)}(\boldsymbol{\xi}_*)$ *are symmetric and positive semi-definite such that* $\bar{\mathbf{P}}_*^{(i)} \neq \bar{\mathbf{Q}}_*^{(i)}$*, and* $\bar{\mathbf{P}}^{(i)} : \boldsymbol{\Xi} \mapsto \mathbb{R}^{d\times d}$ *and* $\bar{\mathbf{Q}}^{(i)} : \boldsymbol{\Xi} \mapsto \mathbb{R}^{d\times d}$ *are in* $\mathcal{C}^{(1)}(\boldsymbol{\Xi})$*;*

(vi) For each $i = 1, 2$ *and* $j = 1, 2, \ldots, d$*,* $(\partial/\partial\theta_j)\mathbf{P}_n^{(i)}(\cdot)$ *and* $(\partial/\partial\theta_j)\mathbf{Q}_n^{(i)}(\cdot)$ *are consistent for* $(\partial/\partial\theta_j)\mathbf{P}^{(i)}(\cdot)$ *and* $(\partial/\partial\theta_j)\mathbf{Q}^{(i)}(\cdot)$ *uniformly on* $\boldsymbol{\Xi}$ *in probability; and*

(vii) For each $i = 1, 2$ *and* $j = 1, 2, \ldots, d$*,* $(\mathbf{P}_*^{(i)})^{-1}(\partial/\partial\theta_j)(\mathbf{P}_n^{(i)}(\boldsymbol{\xi}_*) - \mathbf{P}_{*n}^{(i)})$ *and* $(\mathbf{Q}_*^{(i)})^{-1}(\partial/\partial\theta_j)$ $(\mathbf{Q}_n^{(i)}(\boldsymbol{\xi}_*) - \mathbf{Q}_{*n}^{(i)})$ *are* $O_{\mathbb{P}}(n^{-1/2})$*.* □

Here, Assumption 1(v) is imposed to apply a multivariate central limit theorem to $\widehat{\boldsymbol{\xi}}_n$, $\widehat{\mathbf{P}}_n^{(i)}$, and $\widehat{\mathbf{Q}}_n^{(i)}$, where the last two statistics are involved in the local-to-zero parameters $\bar{\mathbf{P}}_*^{(i)}$ and $\bar{\mathbf{Q}}_*^{(i)}$, respectively. As we show in the next subsection, these local-to-zero parameters are used to examine the limit behaviors of the maximum test statistics under the local alternative. If both $\bar{\mathbf{P}}_*^{(i)}$ and $\bar{\mathbf{Q}}_*^{(i)}$ are zero, Assumption 1(vii) can be used to derive the limit behaviors of the maximum test statistics under the null and alternative hypotheses.

2.3 Limit Behaviors of the Test Statistics

In this section, we examine the limit distributions of the maximum test statistics discussed in the previous sections under the null hypotheses ($\mathcal{H}_0^{(1)}$ and $\mathcal{H}_0^{(2)}$), alternative hypotheses ($\mathcal{H}_1^{(1)}$ and $\mathcal{H}_1^{(2)}$), and following local alternative hypotheses:

- $\mathcal{H}_\ell^{(1)} : \mathbf{A}_* = \mathbf{B}_*$ such that $\mathbf{A}_{*n} = \mathbf{A}_* + n^{-1/2}\bar{\mathbf{A}}_* + o(n^{-1/2})$ and $\mathbf{B}_{*n} = \mathbf{B}_* + n^{-1/2}\bar{\mathbf{B}}_* + o(n^{-1/2})$; and
- $\mathcal{H}_\ell^{(2)} : \mathbf{B}_* = \mathbf{C}_*$ such that $\mathbf{B}_{*n} = \mathbf{B}_* + n^{-1/2}\bar{\mathbf{B}}_* + o(n^{-1/2})$ and $\mathbf{C}_{*n} = \mathbf{C}_* + n^{-1/2}\bar{\mathbf{C}}_* + o(n^{-1/2})$,

where $\bar{\mathbf{A}}_*$, $\bar{\mathbf{B}}_*$, and $\bar{\mathbf{C}}_*$ are positive semi-definite matrices. $\mathcal{H}_\ell^{(1)}$ and $\mathcal{H}_\ell^{(2)}$ capture the local alternative hypotheses to their corresponding null hypotheses. For example, $\mathbf{A}_{*n}$ differs from $\mathbf{B}_{*n}$ under $\mathcal{H}_\ell^{(1)}$, but their limits are identical at the convergence rate $n^{-1/2}$. If $\bar{\mathbf{A}}_* = \bar{\mathbf{B}}_* = \bar{\mathbf{C}}_* = \mathbf{0}$, $\mathcal{H}_\ell^{(1)}$ and $\mathcal{H}_\ell^{(2)}$ reduce to the null hypotheses $\mathcal{H}_0^{(1)}$ and $\mathcal{H}_0^{(2)}$, respectively.

We first obtain the local alternative limit distributions of the maximum test statistics. If we let the local-to-zero parameters be zero, the local alternative limit distributions of the maximum test statistics become the null limit distributions of the same test statistic. We therefore first obtain the local limit distributions of the maximum test statistics and next obtain the null limit distributions from these as by-products.

We now formally provide the local alternative limit distributions of the maximum test statistics in the following theorem:

Theorem 1. *For $i = 1$ and 2, if Assumption 1 holds,*

$$\widehat{\mathfrak{M}}_n^{(i)} \Rightarrow (\mathcal{Z}^{(i)} + \ddot{\mathbf{V}}_*^{(i)})'\mathbf{\Omega}_*^{(i)}(\mathcal{Z}^{(i)} + \ddot{\mathbf{V}}_*^{(i)})$$

under $\mathcal{H}_\ell^{(i)}$, where

$$\mathcal{Z}^{(i)} := \begin{bmatrix} \mathrm{vec}(\mathbf{Z}_{\mathbf{Q}^{(i)}} - \mathbf{Z}_{\mathbf{P}^{(i)}}) \\ \mathrm{vec}(\mathbf{Z}_{\boldsymbol{\theta}}' \otimes \mathbf{I}_d) \end{bmatrix}, \quad \mathbf{\Omega}_*^{(i)} := \begin{bmatrix} (\mathbf{P}_*^{(i)})^{-1} \otimes (\mathbf{P}_*^{(i)})^{-1} & (\mathbf{P}_*^{(i)})^{-1}\mathbf{R}_*^{(i)\prime} \otimes (\mathbf{P}_*^{(i)})^{-1} \\ (\mathbf{P}_*^{(i)})^{-1} \otimes \mathbf{R}_*^{(i)}(\mathbf{P}_*^{(i)})^{-1} & (\mathbf{P}_*^{(i)})^{-1}\mathbf{R}_*^{(i)\prime} \otimes \mathbf{R}_*^{(i)}(\mathbf{P}_*^{(i)})^{-1} \end{bmatrix}$$

with $\mathbf{R}_*^{(i)} := [\nabla_{\boldsymbol{\theta}}(\mathbf{Q}_*^{(i)} - \mathbf{P}_*^{(i)})']'$, $\ddot{\mathbf{V}}_*^{(i)} := \mathbf{\Omega}_*^{(i)-1/2}\mathbf{V}_*^{(i)}$ with $\mathbf{V}_*^{(i)} := (\mathbf{Q}_*^{(i)})^{-1}\bar{\mathbf{Q}}_*^{(i)} - (\mathbf{P}_*^{(i)})^{-1}\bar{\mathbf{P}}_*^{(i)}$, and $\mathbf{Z}_{\boldsymbol{\xi}} := [\mathbf{Z}_{\boldsymbol{\theta}}', Z_{\sigma^2}]'$. □

The local alternative limit distribution in Theorem 1 is obtained by deriving the local alternative weak limit of each maximand as in Cho and Phillips (2018). More specifically, we show that for each $j = 1$ and 2, all $\widehat{\mathfrak{B}}_{j,n}$, $\widehat{\mathfrak{S}}_{j,n}$, $\widehat{\mathfrak{E}}_{j,n}$, $\widetilde{\mathfrak{B}}_{j,n}$, $\widetilde{\mathfrak{S}}_{j,n}$, and $\widetilde{\mathfrak{E}}_{j,n}$ weakly converge to the same weak limit given as $(\mathcal{Z}^{(i)} + \mathbf{V}_*^{(i)})'\mathbf{\Omega}_*^{(i)}(\mathcal{Z}^{(i)} + \mathbf{V}_*^{(i)})$ for $i = 1$ and 2, so that their maximum also weakly converges to $(\mathcal{Z}^{(i)} + \mathbf{V}_*^{(i)})'\mathbf{\Omega}_*^{(i)}(\mathcal{Z}^{(i)} + \mathbf{V}_*^{(i)})$ under the local alternative $\mathcal{H}_\ell^{(i)}$.

We next derive the null limit distributions of the maximum test statistics by letting the local-to-zero parameters be zero, as mentioned above. The following corollary formally states the null limit distributions of the maximum test statistics.

Corollary 1. *For $i = 1$ and 2, if Assumption 1 holds, $\widehat{\mathfrak{M}}_n^{(i)} \Rightarrow \mathcal{Z}^{(i)\prime}\mathbf{\Omega}_*^{(i)}\mathcal{Z}^{(i)}$ under $\mathcal{H}_0^{(i)}$.* □

If the null hypothesis is imposed, so that for $i = 1$ and 2, $\mathbf{V}_*^{(i)} = \mathbf{0}$, Corollary 1 follows. This fact implies that the null limit distributions of the maximum test statistics are obtained as non-central chi-squared distributions with a zero non-centrality parameter.

If the sample size n is relatively small, the null distributions of the maximum test statistics can be better approximated by applying the residual and wild bootstraps to $\widehat{\mathfrak{M}}_n^{(1)}$ and $\widehat{\mathfrak{M}}_n^{(2)}$, respectively than applying the null limit distributions in Corollary 1. The finite sample size distortion can be substantially large if the asymptotic critical values apply when the sample size n is relatively small, because the degrees of freedom in the null limit distribution can be large. $\mathcal{Z}^{(i)}$ is obtained by vectorizing $\mathbf{Z}_{\mathbf{Q}^{(i)}} - \mathbf{Z}_{\mathbf{P}^{(i)}}$ and $\mathbf{Z}_{\boldsymbol{\theta}}' \otimes \mathbf{I}_d$, so that relatively large degrees of freedom are obtained from the vectorization, leading to finite sample null distributions that can differ substantially from the null limit distributions. Instead, the finite sample null distributions of the maximum test statistics can be easily generated using resampling methods. In Section 3, we present the results of Monte Carlo simulations conducted to examine the performance of the maximum test statistics implemented using resampling methods.

We next provide the power properties of the maximum test statistics in the following theorem.

Theorem 2. *For $i = 1$ and 2, if Assumption 1 holds, $\widehat{\mathfrak{M}}_n^{(i)} = n\{\mu_*^{(i)} + O_{\mathbb{P}}(1)\}$ under the fixed alternative hypothesis $\mathcal{H}_1^{(i)}$, where*

$$\mu_*^{(i)} := \left(\frac{d}{2}\right) \max\left[\dot{\mathfrak{B}}_{1,*}^{(i)}, \dot{\mathfrak{E}}_{1,*}^{(i)}, \dot{\mathfrak{E}}_{2,*}^{(i)}, \ddot{\mathfrak{B}}_{1,*}^{(i)}, \ddot{\mathfrak{E}}_{1,*}^{(i)}, \ddot{\mathfrak{E}}_{2,*}^{(i)}\right],$$

$\dot{\mathfrak{B}}_{1,*}^{(i)} := (\dot{\tau}_*^{(i)})^2 + 2\dot{\zeta}_*$, $\dot{\mathfrak{E}}_{1,*}^{(i)} := (\dot{\tau}_*^{(i)})^2 + 2\dot{\gamma}_*$, $\dot{\mathfrak{E}}_{2,*}^{(i)} := (\dot{\eta}_*^{(i)})^2 + 2\dot{\zeta}_*$, $\ddot{\mathfrak{B}}_{1,*}^{(i)} := (\ddot{\tau}_*^{(i)})^2 + 2\ddot{\zeta}_*^{(j)}$, $\ddot{\mathfrak{E}}_{1,*}^{(i)} := (\ddot{\tau}_*^{(i)})^2 + 2\ddot{\gamma}_*$, *and* $\ddot{\mathfrak{E}}_{2,*}^{(i)} := (\ddot{\eta}_*^{(i)})^2 + 2\ddot{\zeta}_*$. □

That is, the limit distribution of the maximum test statistics diverges to positive infinity at the rate of n. Here, the first term $\mu_*^{(i)}$ denotes the location parameter leading to the divergence of the maximum test statistics. On the contrary, the remaining term $O_{\mathbb{P}}(n)$ determines the dispersion scale of the maximum test statistics under the alternative.

The asymptotic power of the maximum test statistics is expected to be greater than that in Cho and Phillips (2018). For each $i = 1$ and 2, applying their maximum test statistic yields the leading term $\dot{\mu}_*^{(i)} := \max_{j=1,2}[\dot{\mathfrak{B}}_{j,*}^{(i)}, \dot{\mathfrak{S}}_{j,*}^{(i)}, \dot{\mathfrak{E}}_{j,*}^{(i)}]$, where $\dot{\mathfrak{S}}_{1*}^{(i)} := (\dot{\delta}_*^{(i)})^2 + 2\dot{\gamma}_*^{(i)}$ and $\dot{\mathfrak{S}}_{2*}^{(i)} := (\dot{\eta}_*^{(i)})^2 + 2\dot{\zeta}_*^{(i)}$. Here, $\dot{\delta}_*^{(i)} \in [\dot{\eta}_*^{(i)}, \dot{\tau}_*^{(i)}]$, so that $\dot{\mathfrak{S}}_{1,*}^{(i)} \le \dot{\mathfrak{E}}_{1,*}^{(i)}$. Furthermore, $\dot{\mathfrak{S}}_{2,*}^{(i)} \le \max[\ddot{\mathfrak{E}}_{1,*}^{(i)}, \ddot{\mathfrak{B}}_{1,*}^{(i)}]$ by the definitions of $\ddot{\mathfrak{E}}_{1,*}^{(i)}$ and $\ddot{\mathfrak{B}}_{1,*}^{(i)}$, implying that $\dot{\mu}_*^{(i)} \le \mu_*^{(i)}$. From this feature, for each $i = 1$ and 2, the leading term of $\widehat{\mathfrak{M}}_n^{(i)}$ becomes greater than that in Cho and Phillips (2018). From this fact, greater power is expected from the maximum test statistics.

2.4 Plan to Test for the Influence of Conditional Heteroskedasticity and Autocorrelation

When testing for the sandwich-form asymptotic covariance matrix entailed by both conditional heteroskedastic and autocorrelated errors, we recommend testing autocorrelation first and then conditional heteroskedasticity. One of the assumptions retained to test $\mathcal{H}_0^{(1)}$ against $\mathcal{H}_1^{(1)}$ is that $\{U_t, \mathcal{F}_t\}$ is an MDS. If $\{U_t, \mathcal{F}_t\}$ is not an MDS, it is difficult to trust the test result. We therefore recommend testing for the sandwich-form asymptotic covariance matrix entailed by autocorrelated errors first and then testing for the asymptotic covariance matrix form influenced by heteroskedastic regression errors. Specifically, we provide the following STP:

- Step 1: Test $\mathcal{H}_0^{(2)}$ against $\mathcal{H}_1^{(2)}$ using $\widehat{\mathfrak{M}}_n^{(2)}$. If $\mathcal{H}_0^{(2)}$ is rejected, stop the STP. Otherwise, move to the next step.
- Step 2: Test $\mathcal{H}_0^{(1)}$ against $\mathcal{H}_1^{(1)}$ using $\widehat{\mathfrak{M}}_n^{(1)}$. If $\mathcal{H}_0^{(1)}$ is rejected, conclude that U_t is conditionally heteroskedastic. Otherwise, conclude that $\mathbf{A}_* = \mathbf{B}_*$.

Using the provided STP, the given hypotheses (i.e., the statuses of the asymptotic covariance matrix) are selected with different asymptotic probabilities. Under $\mathcal{H}_1^{(2)}$, the empirical researcher can detect the correct hypothesis with an asymptotic probability converging to 100% using the STP from the fact that $\widehat{\mathfrak{M}}_n^{(2)}$ is a consistent test statistic. On the contrary, under $\mathcal{H}_1^{(1)}$ that also belongs to $\mathcal{H}_0^{(2)}$, the empirical researcher can detect the correct hypothesis with an asymptotic probability converging to $(1 - \alpha) \times 100\%$ using the STP, where α denotes the level of significance selected by the researcher. In the first step, the empirical researcher can commit to an $\alpha \times 100\%$ type-I error while testing for the sandwich-form asymptotic covariance matrix entailed by autocorrelated errors. On the contrary, the sandwich-form asymptotic covariance matrix entailed by conditional heteroskedastic errors can be consistently detected by $\widehat{\mathfrak{M}}_n^{(1)}$ with an asymptotic probability converging to 100%. In total, the researcher therefore commits to an asymptotic probability converging to $(1-\alpha)\times 100\%$. Finally, under $\mathcal{H}_0^{(1)}$, the empirical researcher can detect the correct hypothesis with an asymptotic probability converging to $(1-2\alpha)\times 100\%$. $\mathcal{H}_0^{(1)}$ can be selected only when neither $\mathcal{H}_0^{(2)}$ nor $\mathcal{H}_0^{(1)}$ is rejected by the STP. As the testing procedure is conducted at $\alpha\times 100\%$ in each step, the overall asymptotic type-I error becomes $2\alpha\times 100\%$. In the next section, we show the Monte Carlo simulations and affirm this feature.

3 Simulations

In this section, we present the results of the Monte Carlo simulations conducted to affirm the theory in Section 2. Given that the test statistics can be sequentially applied, we first discuss testing $\mathcal{H}_0^{(1)}$ against $\mathcal{H}_1^{(1)}$ and $\mathcal{H}_0^{(2)}$ against $\mathcal{H}_1^{(2)}$ and then discuss the STP.

For the goal of this section, we consider the following simulation setup. We first suppose that time-series observations are generated according to the following formula:

$$Y_t = \beta_{0*} + \beta_{1*}Y_{t-1} + \beta_{2*}Y_{t-2} + U_t \quad \text{and} \quad U_t = \sqrt{h_t}\varepsilon_t,$$

such that

$$h_t = \kappa_* + \phi_* h_{t-1} + \gamma_* U_{t-1}^2 \quad \text{and} \quad \varepsilon_t \sim \text{ IID } N(0,1).$$

Different time-series observations can be generated by letting the parameters in the data-generating process (DGP) be specified differently. We consider the following three parameter specifications:

- DGP1: $(\beta_{0*},\beta_{1*},\beta_{2*},\kappa_*,\phi_*,\gamma_*) = (0.0, 0.5, 0.0, 1.0, 0.0, 0.0)$;
- DGP2: $(\beta_{0*},\beta_{1*},\beta_{2*},\kappa_*,\phi_*,\gamma_*) = (0.0, 0.5, 0.0, 1.0, 0.2, 0.2)$; and
- DGP3: $(\beta_{0*},\beta_{1*},\beta_{2*},\kappa_*,\phi_*,\gamma_*) = (0.0, 0.5, -0.2, 1.0, 0.2, 0.2)$.

DGP1 generates serially uncorrelated errors with conditionally homoskedastic variance from the fact that $h_t = 1$, so that we can use DGP1 to examine the size property of $\widehat{\mathfrak{M}}_n^{(1)}$ under $\mathcal{H}_0^{(1)}$. That is, if we let $\sigma_*^2 := \mathbb{E}[U_t^2]$, $\sigma_*^2\mathbb{E}[\mathbf{X}_t\mathbf{X}_t']$ becomes identical to the probability limit of the HC covariance matrix estimator for $\mathbb{E}[U_t^2\mathbf{X}_t\mathbf{X}_t']$. On the contrary, DGP2 generates conditionally heteroskedastic errors, although the errors form an MDS. Therefore, we can use DGP2 to examine the performance of $\widehat{\mathfrak{M}}_n^{(2)}$ under $\mathcal{H}_0^{(2)}$. In addition to this feature, the probability limit of the HC covariance matrix estimator becomes identical to the probability limit of the HAC covariance matrix estimator, whereas $\sigma_*^2\mathbb{E}[\mathbf{X}_t\mathbf{X}_t']$ becomes different from the probability limit of the HC covariance matrix estimator. Therefore, DGP2 can be used to investigate the performance of $\widehat{\mathfrak{M}}_n^{(1)}$ under $\mathcal{H}_1^{(1)}$. Finally, DGP3 generates autocorrelated and conditionally heteroskedastic errors from the fact that $\mathcal{M}$ is dynamically misspecified for $\mathbb{E}[Y_t|Y_{t-1},Y_{t-2},\ldots]$, so that the probability limits of the HC and HAC covariance matrix estimators differ. From this fact, DGP3 can be used to examine the performance of $\widehat{\mathfrak{M}}_n^{(2)}$ under $\mathcal{H}_1^{(2)}$.

Given these DGP conditions, we further suppose that the researcher specifies the following model:

$$\mathcal{M} := \{\theta_0 + \theta_1 Y_{t-1} : (\theta_0, \theta_1)' \in \mathbb{R}^2\}$$

and estimates the unknown coefficients in $\mathcal{M}$ using the LS estimator. This estimation procedure is equivalent to estimating the linear model by supposing that the error is normally distributed. For notational simplicity, we let $\mathbf{X}_t := (1, Y_{t-1})'$.

It is useful to apply resampling methods to test for the sandwich-form asymptotic covariance matrices, as discussed above. For our simulations, we apply the residual bootstrap in Efron and Tibsharani (1988) and the wild bootstrap in Wu (1986). Specifically, residual bootstrap resampling is useful for testing $\mathcal{H}_0^{(1)}$ against $\mathcal{H}_1^{(1)}$ by $\widehat{\mathfrak{M}}_n^{(1)}$, whereas wild bootstrap resampling is useful for testing $\mathcal{H}_0^{(2)}$ against $\mathcal{H}_1^{(2)}$ by $\widehat{\mathfrak{M}}_n^{(2)}$. The residual bootstrap resampling method can thus be applied in the following plan:

- Step 1: Estimate $\widehat{\boldsymbol{\theta}}_n$ using the LS method and obtain $\{\widehat{U}_t := Y_t - \mathbf{X}_t'\widehat{\boldsymbol{\theta}}_n : t = 1, 2, \ldots, n\}$ to estimate $\widehat{\sigma}_n^2 \widehat{\mathbf{A}}_n$ and $\widehat{\mathbf{B}}_n$. From these matrix estimators, compute $\widehat{\mathfrak{M}}_n^{(1)}$, where

$$\widehat{\sigma}_n^2 := \frac{1}{n}\sum_{t=1}^{n} \widehat{U}_t^2, \quad \widehat{\mathbf{A}}_n := \frac{1}{n}\sum_{t=1}^{n} \mathbf{X}_t\mathbf{X}_t', \quad \text{and} \quad \widehat{\mathbf{B}}_n := \frac{1}{n}\sum_{t=1}^{n} \widehat{U}_t^2 \mathbf{X}_t\mathbf{X}_t'; \tag{4}$$

- Step 2: Resample $\{\widehat{U}_t^b : t = 1, 2, \ldots, n\}$ and obtain $\widehat{\sigma}_n^{2b}\widehat{\mathbf{A}}_n$ and $\widehat{\mathbf{B}}_n^b$ to compute $\widehat{\mathfrak{M}}_n^{(1)b}$ by $\widehat{\sigma}_n^{2b} := n^{-1}\sum_{t=1}^n (\widehat{U}_t^b)^2$ and $\widehat{\mathbf{B}}_n^b := n^{-1}\sum_{t=1}^n \widehat{U}_t^{b2}\mathbf{X}_t\mathbf{X}_t'$;
- Step 3: Iterate Step 2 for $b = 1, 2, \ldots, B$ and estimate the bootstrapped p-value by $\widehat{p}_n^{(1)} := B^{-1}\sum_{b=1}^B \mathbf{1}\{\widehat{\mathfrak{M}}_n^{(1)} < \widehat{\mathfrak{M}}_n^{(1)b}\}$;
- Step 4: If $\widehat{p}_n^{(1)} < \alpha$, reject $\mathcal{H}_0^{(1)}$; otherwise, do not reject $\mathcal{H}_0^{(1)}$.

On the contrary, the wild bootstrap resampling method can be implemented in the following plan to test $\mathcal{H}_0^{(2)}$ against $\mathcal{H}_1^{(2)}$:

- Step 1: Estimate $\widehat{\boldsymbol{\theta}}_n$ using the LS estimation and obtain $\{\widehat{U}_t := Y_t - \mathbf{X}_t'\widehat{\boldsymbol{\theta}}_n : t = 1, 2, \ldots, n\}$ to estimate $\widehat{\mathbf{B}}_n$ and $\widehat{\mathbf{C}}_n$. From these matrix estimators, compute $\widehat{\mathfrak{M}}_n^{(2)}$, where $\widehat{\mathbf{B}}_n$ is the same estimator as $\widehat{\mathbf{B}}_n$ in the residual bootstrap, and

$$\widehat{\mathbf{C}}_n := \widehat{\mathbf{B}}_n + \frac{1}{n}\sum_{k=1}^{\ell}\omega_{\ell k}\sum_{t=k+1}^{n}(\widehat{U}_{t-k}\widehat{U}_t\mathbf{X}_{t-k}\mathbf{X}_t' + \widehat{U}_t\widehat{U}_{t-k}\mathbf{X}_t\mathbf{X}_{t-k}'). \tag{5}$$

Here, $\omega_{\ell k}$ and ℓ are the HAC kernel and bandwidth, respectively used to estimate $\mathbf{C}_*$ consistently;

- Step 2: Resample $V_t \sim$ IID $(0, 1)$ and define $\{W_t^b := V_t\widehat{U}_t : t = 1, 2, \ldots, n\}$ to compute $\ddot{\mathbf{B}}_n$ and $\ddot{\mathbf{C}}_n$ and obtain $\widehat{\mathfrak{M}}_n^{(2)b}$, where

$$\ddot{\mathbf{B}}_n := \frac{1}{n}\sum_{t=1}^{n}(W_t^b)^2\mathbf{X}_t\mathbf{X}_t' \quad \text{and} \quad \ddot{\mathbf{C}}_n := \ddot{\mathbf{B}}_n + \frac{1}{n}\sum_{k=1}^{\ell}\omega_{\ell k}\sum_{t=k+1}^{n}(W_{t-k}W_t\mathbf{X}_{t-k}\mathbf{X}_t' + W_tW_{t-k}\mathbf{X}_t\mathbf{X}_{t-k}');$$

- Step 3: Iterate Step 2 for $b = 1, 2, \ldots, B$ and estimate bootstrapped p-value by $\widehat{p}_n^{(2)} := B^{-1}\sum_{b=1}^{B} \mathbf{1}\{\widehat{\mathfrak{M}}_n^{(2)} < \widehat{\mathfrak{M}}_n^{(2)b}\}$;
- Step 4: If $\widehat{p}_n^{(2)} < \alpha$, reject $\mathcal{H}_0^{(2)}$; otherwise, do not reject $\mathcal{H}_0^{(2)}$.

Wild bootstrap resampling has many variations, as V_t can be drawn differently from the standard normal and many alternative estimators exist for $\mathbf{C}_*$. For example, the Rademacher distribution is also popularly used for V_t in the literature. For our simulations, we let $V_t \sim$ IID $N(0,1)$ and $\omega_{\ell k}$ be the quadratic kernel in (3). We further let $\ell = \lfloor n^{1/5} \rfloor - 1$.

We conduct Monte Carlo simulations using DGP1, -2, and -3 according to the above two resampling methods. Table 1 reports the simulation results. We summarize the simulation results as follows:

(a) The first panel of Table 1 reports the empirical rejection rates of $\widehat{\mathfrak{M}}_n^{(1)}$ under $\mathcal{H}_0^{(1)}$. We applied the residual bootstrap to evaluate $\widehat{\mathfrak{M}}_n^{(1)}$, which is computed using data observations following DGP1. As the sample size n increases, the empirical rejection rates approach nominal levels, implying that the resampling method enables us to control the type-I error successfully.

(b) The first two panels of Table 1 report the empirical rejection rates of $\widehat{\mathfrak{M}}_n^{(2)}$ under $\mathcal{H}_0^{(2)}$. Both DGP1 and DGP2 can be used to generate observations belonging to $\mathcal{H}_0^{(2)}$. Hence, we applied the wild bootstrap method to $\widehat{\mathfrak{M}}_n^{(2)}$. As the sample size n increases, the empirical rejection rates approach nominal levels, implying that the resampling method also enables us to control the type-I error successfully under $\mathcal{H}_0^{(2)}$. In particular, when the error term exhibits conditional homoskedasticity (i.e., under DGP1), the empirical rejection rates of $\widehat{\mathfrak{M}}_n^{(2)}$ are closer to nominal levels than those under DGP2. Furthermore, the empirical rejection rates are closer to nominal levels when the level of significance α is small. Although we do not report these here, our simulations show that a rather larger sample size n is required for the successful application of the wild bootstrap method when the level of significance α is relatively large.

(c) The second and third panels of Table 1 report the empirical rejection rates of $\widehat{\mathfrak{M}}_n^{(1)}$ under $\mathcal{H}_1^{(1)}$. Both DGP2 and DGP3 can be used to generate observations belonging to $\mathcal{H}_1^{(1)}$. The two panels of Table 1 show that the empirical rejection rates converge to unity as the sample size n increases. The empirical rejection rates are more or less similar between the two panels.

(d) The third panel of Table 1 reports the empirical rejection rates of $\widehat{\mathfrak{M}}_n^{(2)}$ under $\mathcal{H}_1^{(2)}$. As the sample size n increases, the empirical rejection rates converge to unity, whereas the rejection rates are lower than those of $\widehat{\mathfrak{M}}_n^{(1)}$ under $\mathcal{H}_1^{(1)}$.

These simulation results show that the theories on $\widehat{\mathfrak{M}}_n^{(1)}$ and $\widehat{\mathfrak{M}}_n^{(2)}$ are affirmed when testing $\mathcal{H}_0^{(1)}$ and $\mathcal{H}_0^{(2)}$ against $\mathcal{H}_1^{(1)}$ and $\mathcal{H}_1^{(2)}$, respectively, implying that we can consistently distinguish the statuses of the asymptotic covariance matrices generated by homoskedastic, heteroskedastic, and autocorrelated errors.

We next examine the performance of the STP using Monte Carlo simulations. Using DGP1, -2, and -3, we examine the asymptotic portion of the hypotheses selected by the STP. As discussed in Section 2.4, if the STP is successfully implemented, $\mathcal{H}_0^{(1)}$ should be selected with an asymptotic probability converging to $(1 - 2\alpha) \times 100\%$ under DGP1.

On the contrary, $\mathcal{H}_0^{(2)}$ should be selected with an asymptotic probability converging to $(1 - \alpha) \times 100\%$ under DGP2. Finally, $\mathcal{H}_1^{(2)}$ should be selected with an asymptotic probability converging to 100% under DGP3. We let the level of significance α be 0.05 to conduct the simulations. Table 2 reports the simulation results, which can be summarized as follows:

(a) Under DGP1, $\mathcal{H}_0^{(1)}$ is estimated with a proportion close to 90% as the sample size n increases. This aspect affirms that the STP estimates $\mathcal{H}_0^{(1)}$ with an asymptotic probability converging to $(1 - 2\alpha) \times 100\%$. On the contrary, both $\mathcal{H}_0^{(2)}$ and $\mathcal{H}_1^{(2)}$ are selected with asymptotic probabilities converging to $\alpha \times 100\%$. This feature is caused by the $\alpha \times 100\%$ type-I error, so that if a small α is selected, the researcher can estimate $\mathcal{H}_0^{(1)}$ with a high precision rate.

(b) Under DGP2, $\mathcal{H}_0^{(2)}$ is estimated with a proportion close to 95% as the sample size n increases. This aspect affirms that the STP estimates $\mathcal{H}_0^{(2)}$ with an asymptotic probability converging to $(1 - \alpha) \times 100\%$. On the contrary, $\mathcal{H}_0^{(1)}$ and $\mathcal{H}_1^{(2)}$ are selected with probabilities converging to 0% and $\alpha \times 100\%$, respectively. The portion of $\mathcal{H}_1^{(2)}$ is determined by the type-I error. As for DGP1, if the researcher selects α to be close to zero, (s)he can estimate $\mathcal{H}_0^{(2)}$ with a high precision rate.

(c) Under DGP3, $\mathcal{H}_1^{(2)}$ is estimated with a proportion close to 100% as the sample size n increases. Under the first step of the STP, the researcher can reject $\mathcal{H}_0^{(2)}$ with an asymptotic probability converging to unity, so that $\mathcal{H}_1^{(2)}$ can be consistently selected. When $n = 2,000$, $\mathcal{H}_1^{(2)}$ is selected with 86.10%, whereas if the sample size n further increases, $\mathcal{H}_1^{(2)}$ must be selected with a probability converging to 100%.

Our Monte Carlo simulations affirm the theories in Section 2.3, validating the use of the matrix equality testing.

4 Empirical Applications

In this section, we apply the matrix equality testing to empirical data. We examine two empirical models. First, we revisit the conditionally converging economic growth model using the empirical data provided by Mankiw, Romer, and Weil (1992) and then examine how a typical time-series data analysis can be combined with the matrix equality testing. For this purpose, we use the energy data provided by Bai and Lam (2019).

4.1 Conditionally Converging Economic Growth Model Estimation

In the economic growth literature, there are many theories on economic growth and many empirical studies have been conducted to compare economic growth rates between countries. In particular, the conditional convergence hypothesis is popularly examined in empirical studies, which states that different savings propensities between countries result in the same economic growth rate if the same technological possibilities and population growth rate are possessed by different countries. For example, Mankiw, Romer, and Weil (1992) empirically examine Solow's (1957) growth model and find evidence for conditional convergence using cross-sectional data consisting of three country groups: non-oil, intermediate, and OECD countries. Their economic growth rates are collected using GDP observations between 1960 and 1985.

In this section, we revisit the empirical inference in Mankiw, Romer, and Weil (1992) that test the conditional convergence hypothesis by assuming a conditional homoskedastic regression error. Specifically, we test for the sandwich-form asymptotic covariance matrix entailed by heteroskedastic regression errors using the maximum test statistic $\widehat{\mathfrak{M}}_n^{(1)}$ and examine how this additional testing result modifies the inference results in Mankiw, Romer, and Weil (1992). We further compare our inference results with Breusch and Pagan's (1979) and White's (1980) test statistics for heteroskedasticity.

Mankiw, Romer, and Weil (1992) estimate the following model to infer the conditional convergence hypothesis of economic growth:

$$\log(Y_{85,t}) = \beta_{0*} + \beta_{1*}\log(Y_{60,t}) + \beta_{2*}\log(s_t) + \beta_{3*}\log(n_t + g_t + \delta_t) + U_t,$$

where $Y_{85,t}$ and $Y_{60,t}$ represent the GDP per working-age person of the t-th country in 1985 and 1960, respectively and s_t, n_t, g_t, and δ_t denote the average share of real investment in real GDP, average growth rate of the working-age population, technology growth, and depreciation rate, respectively.[1] In the above model, $\log(s_t)$ and $\log(n_t + g_t + \delta_t)$ are controlled as the determinants of steady states. According to the Solow model, poor countries with a lower initial income level tend to grow faster than richer ones, which means that β_{1*} is negative. Conditional convergence exists if the estimate of β_{1*} is significantly negative. Following their inferencing plan, we reproduce their model estimation results and further apply the maximum test statistic to the estimated asymptotic covariance matrix of the estimated parameters.

Using the data sets provided by Mankiw, Romer, and Weil (1992), we obtain the model estimation results using the LS estimation and present them in Table 3. In the same table, we also report the p-values of the maximum test statistic and other conditional heteroskedasticity test statistics. Specifically, the top panel of Table 3 reports the LS estimates for the coefficients, standard errors, and White's HC standard errors for the three data sets of non-oil, intermediate, and OECD countries. The two standard errors are separately provided to compute the t-ratio using a suitable standard error. If the maximum test statistic $\widehat{\mathfrak{M}}_n^{(1)}$ rejects the null hypothesis $\mathcal{H}_0^{(1)}$, the t-ratio should be computed using the heteroskedasticity-robust standard error, so that it asymptotically follows a standard normal distribution under the null. On the contrary, Mankiw, Romer, and Weil (1992) assume a conditional homoskedastic regression error and infer the unknown parameters using standard t-test statistics only. The bottom panel also shows the p-values of the test statistics $\widehat{\mathfrak{M}}_n^{(1)}$, Breusch and Pagan's (1979), and White's (1980) test statistics for the three groups of countries. We summarize the key results of Table 3 as follows:

(a) For non-oil countries, the null hypothesis of no conditional heteroskedasticity $\mathcal{H}_0^{(1)}$ is rejected by the maximum test statistic $\widehat{\mathfrak{M}}_n^{(1)}$ at the 5% significance level. This aspect is also affirmed by Breusch and Pagan's (1979) and White's (1980) test statistics. Breusch and Pagan's (1979) test statistic rejects the conditional homoskedasticity hypothesis at the 5% level of significance, while White's (1980) test statistic also reveals a p-value close to 5%, that is 6.58%. This aspect implies that the t-test statistics should have been computed using the HC covariance

[1] Mankiw, Romer, and Weil (1992) assume that $g_t + \delta_t = 0.05$.

matrix estimator to infer the unknown coefficients suitably.

(b) For intermediate countries, we draw the same conclusion as for non-oil countries. The maximum test statistic $\widehat{\mathfrak{M}}_n^{(1)}$ shows a p-value less than 5% and the other two test statistics also have p-values less than 5%, implying that the t-ratios should have been computed using the HC covariance matrix estimator to infer the unknown coefficients as for non-oil countries.

(c) On the contrary, for OECD countries, the opposite result is obtained. The maximum test statistic $\widehat{\mathfrak{M}}_n^{(1)}$ has a p-value substantially greater than 5%, so that we cannot reject the null hypothesis $\mathcal{H}_0^{(1)}$. This conclusion is further reinforced by Breusch and Pagan's (1979) and White's (1980) test statistics that yield p-values greater than 5%. Therefore, the standard t-test statistic can be used to infer the unknown coefficients, implying that Mankiw, Romer, and Weil's (1992) inference on the unknown parameters is suitable.

(d) Based on the inference results of the maximum test statistic, we compute the t-test statistics for the unknown coefficient of $\log(Y_{60,t})$ using the HC covariance matrix estimator. For both non-oil and intermediate countries, they are negatively valued and turn out to be statistically significant as for the standard t-test statistics. Although different standard errors are used from those in Mankiw, Romer, and Weil (1992), there exist opposite relationships between economic growth and the initial value for both groups of countries. From this, the same conclusion is reached as by Mankiw, Romer, and Weil (1992).

(e) A rather large influence is made to the standard errors for the coefficients of $\log(n_t + g_t + \delta_t)$. The t-test statistic for the coefficient of $\log(n_t + g_t + \delta_t)$ using the standard error is $-0.4575 \div 0.3074 \cong 1.48$ and its p-value is about 0.14, which is insignificant. On the contrary, the t-test statistic using the HC covariance matrix estimator and its p-value are about -1.93 and 0.05, respectively, implying that the estimate of $\log(n_t + g_t + \delta_t)$ is significant at the 6% level of significance.

This result shows that the maximum test statistic $\widehat{\mathfrak{M}}_n^{(1)}$ provides strong evidence for the sandwich-form asymptotic covariance matrices entailed by conditionally heteroskedastic regression errors for non-oil and intermediate countries. Nevertheless, the t-test statistics modified by this inference do not modify the original inference on the conditional convergence hypothesis in Mankiw, Romer, and Weil (1992) for either group of countries.

4.2 Dynamic Model Estimation for Energy Prices

Classical time-series analysis attempts to specify a model without autocorrelation in the prediction errors and estimate a model with a high precision rate for the parameter estimates. For this purpose, the empirical researcher popularly estimates the conditional mean and variance models specified as AR and GARCH models, respectively.

In this section, we conduct an empirical practice using energy price growth rates. The typical model assumption for the energy price growth rate is an AR-GARCH(1,1) model, and it has been empirically reported that the growth rates of energy prices such as the liquefied petroleum gas (LPG) freight rate and growth rate of Brent crude oil price are well estimated using the AR-GARCH(1,1) model. The AR order is typically selected using the Bayesian information criterion and the autocorrelation in the residuals is tested using Ljung–Box's (1978) Q-test statistic. If the autocorrelation cannot be detected using the Q-test statistic, the researcher estimates the conditional heteroskedastic behavior based on a GARCH(1,1) model and attempts to estimate the parameters using maximum likelihood estimation after supposing a particular distribution for the standardized error.

Specifying a correct model without autocorrelation in the prediction error and for conditional variance is important for further examining the growth rates of energy prices. For example, Bai and Lam (2019) analyze the conditional dependence structure among the LPG freight rate, product price arbitrage, and crude oil price by analyzing them using a copula model estimation. Without having the correct models for the individual variables, estimating the next-stage models using copula estimation can lead to the failure to infer the economic variables correctly. This is mainly because the dependence structure between the individual variables cannot be correctly inferred without estimating the marginal distribution of each variable.

In this section, we revisit the data analysis to estimate the marginal distributions of the energy price growth rates in Bai and Lam (2019) by testing the conditional heteroskedasticity and autocorrelation assumption in the prediction error. Seven marginal distributions of energy variables are examined in Bai and Lam (2019): the growth rates of the Baltic LPG (BLPG) price, Brent crude oil price (BRENT), propane Mt Belvieu price (MB), propane Argus Far East index (PAFEI), propane CP swap month 1 price (CPS), AFEI-US price index (AFEIUS), and AFEI-CP swap price index (AFEICPS). They are weekly observed data from the second week of 2005 to the 35th week of 2016, and the total number of observations is 601. Table 1 in Bai and Lam (2019) provides the descriptive statistics of the variables.

Using the data sets provided by Bai and Lam (2019), we verify whether their model analysis is affirmed by the maximum test statistics. Bai and Lam (2019) employ Ljung–Box's (1978)) Q-test statistic and specify an AR(2) model for BLPG, whereas they select an AR(1) model for the other variables. In addition, they adopt a GARCH(1,1) model based on maximum likelihood estimation by assuming the skewed t-distribution for the standardized error distribution. Table 3 in Bai and Lam (2019) presents Engle's (1982) ARCH-LM test statistic to show that the standardized errors are homoskedastic.

We conduct our empirical analysis for the same data sets using $\widehat{\mathfrak{M}}_n^{(1)}$ and $\widehat{\mathfrak{M}}_n^{(2)}$. We summarize our empirical inference results as follows:

(a) We first test for the sandwich-form asymptotic covariance matrix entailed by autocorrelated errors using $\widehat{\mathfrak{M}}_n^{(2)}$. Specifically, for $p = 1, 2, \ldots, 10$, we estimate the AR(p) model using the LS estimation and test $\mathcal{H}_0^{(2)} : \mathbf{B}_* = \mathbf{C}_*$ using $\widehat{\mathbf{B}}_n$ and $\widehat{\mathbf{C}}_n$ in (4) and (5), respectively by applying the wild bootstrap method. Here, we let $\mathbf{X}_t := (1, Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p})'$. When applying the wild bootstrap method, we use the HAC estimation with the quadratic kernel and bandwidth $\ell = \lfloor n^{1/5} \rfloor - 1$ in Andrews (1991). Table 4 reports the p-values estimated using the wild bootstrap method and we denote p-values greater than 5% using boldface font. According to the maximum test statistics, the AR(1) model is correctly specified for BLPG, BRENT, PAFEI, CPS, AFEIUS, and AFEICPS. For these variables, if $\mathcal{H}_0^{(2)}$ cannot be rejected at the 5% level of significance, most AR models with lags greater than unity turn out not to reject $\mathcal{H}_0^{(2)}$. This aspect implies that the AR(2) model may not be the most parsimonious model for BLPG, although AR(2) is also correct for the conditional mean.

(b) Nevertheless, the AR(1) model is not correctly specified for MB and AFEICPS, although Bai and Lam (2019) select the AR(1) model using the Bayesian information criterion. For every $p = 1, 2, \ldots, 10$, AR(p) is not correctly specified as implied by the maximum test statistic $\widehat{\mathfrak{M}}_n^{(2)}$. Although not reported in Table 4, we increased the order of AR(p) to 15 in a further empirical analysis and also found that AR(p) generated autocorrelated errors. On the contrary, if $p \geq 3$, the AR(p) model is correctly specified for AFEICPS. Although the AR(9) model yields a p-value less than 5% for $\widehat{\mathfrak{M}}_n^{(2)}$, it does not show a consistent test result. For the AR(10) model, the p-value rises above 5% again. Although not reported here, this feature also holds for AR models with degrees greater than 10. These investigations allow us to conclude that the AR(1) model is dynamically misspecified for MB and AFEICPS, which could not be detected using Ljung–Box's (1978) Q-test statistic.

(c) We next test for conditional heteroskedasticity using $\widehat{\mathfrak{M}}_n^{(1)}$. As shown in Table 4, the p-values of the $\widehat{\mathfrak{M}}_n^{(1)}$ test statistics are less than 5% for BLPG, BRENT, MB, PAFEI, CPS, and AFEICPS, implying that conditional heteroskedastic aspects are persistently observed from these variables.

(d) Finally, we test the correct model assumption using the information matrix equality. White (1982) provides the information matrix test for the model specification of maximum likelihood estimation. Given that the AR-GARCH(1,1) model specified by Bai and Lam (2019) is estimated using maximum likelihood estimation assuming a skewed t-distribution for the standardized error, we apply White's (1982) information matrix equality testing to Bai and Lam's (2019) models. Specifically, if we let $\mathbf{A}_*$ be the limit matrix of the negative Hessian matrix and $\mathbf{B}_*$ be the asymptotic covariance matrix of the scores for maximum likelihood estimation, we can also test $\mathbf{A}_* = \mathbf{B}_*$ using the maximum test statistic of the current study to test distributional misspecification for standardized errors. We obtain the p-values of the maximum test statistic using the parametric bootstrap method and report them in Table 5. The desired information matrix equality is not sufficient but is necessary for the correct distributional specification. Hence, we cannot say whether the AR-GARCH(1,1) model with a skewed t-distribution is the correct assumption for the DGP, even though we cannot reject $\mathbf{A}_* = \mathbf{B}_*$. Nevertheless, we can be assured that the model is misspecified if the hypothesis $\mathbf{A}_* = \mathbf{B}_*$ is rejected. As shown in Table 5, the information matrix equality for the correct distributional assumption cannot be rejected for BLPG, BRENT, PAFEI, and CPS, implying that the skewed t-distribution fits the standardized error distribution sufficiently well. On the contrary, the skewed t-distribution is misspecified for AFEIUS. The p-value for AFEIUS is 0.25%. Although the AR-GARCH(1,1) model is correct for AFEIUS, the low p-value shows that the skewed t-distribution is misspecified for the standardized error distribution. Likewise, the skewed t-distribution is also misspecified for the standardized error distribution of AFEICPS. The AR-GARCH(1,1) model with an AR lag equal to unity is dynamically misspecified, as already seen in Table 4, implying that the information matrix equality is unlikely to hold for maximum likelihood estimation for AFEICPS. Finally, the information matrix equality test statistic does not reject the hypothesis that $\mathbf{A}_* = \mathbf{B}_*$ for MB, whereas it does not imply that the skewed t-distribution is correctly specified for the standardized error distribution of MB. Table 4 already showed that AR-GARCH(1,1) is dynamically misspecified, implying that the model is misspecified. Given that the information matrix equality is a necessary condition for the correct model specification, this result does not necessarily imply that the AR-GARCH(1,1) model with a skewed t-distribution is the correct model for MB.

From the empirical analysis in this section, we find that the AR-GARCH(1,1) model is correctly specified for the energy variables examined by Bai and Lam (2019). Nevertheless, some of them such as MB, AFEIUS, and AFEICPS are dynamically or distributionally misspecified.

5 Concluding Remarks

In the current study, we examine testing for the sandwich-form covariance matrix entailed by conditional heteroskedasticity and autocorrelation in the regression error. Given that none of the conditional heteroskedastic or autocorrelated errors yields the sandwich-form covariance matrix for the LS estimator, it is not necessary to estimate the covariance matrix by the HC or HAC covariance matrix estimator. Only if the covariance matrix exhibits the sandwich-form covariance matrix would it be desirable to estimate the asymptotic covariance matrix using the HC or HAC covariance matrix estimator given that it is not the most efficient estimator under the information matrix equality. Because of this fact, we first examine testing for the sandwich-form covariance matrix before applying the HC or HAC covariance matrix estimator. For this purpose, we apply the testing methodologies proposed by Cho and White (2015) and Cho and Phillips (2018) to fit the context of this study and define the maximum test statistics by extending the scope of their maximum test statistic. We derive the null, alternative, and local alternative limit distributions of the maximum test statistics and further affirm the theories on the maximum test statistics through a simulation. We also apply the maximum test statistics to empirical data sets popularly examined in the literature on economic growth and the growth rates of energy prices, affirming the empirical result in Mankiw, Romer, and Weil (1992) as well as confirming that the model assumptions for the growth rates of the energy prices in Bai and Lam (2019) are correct for many of the energy variables, although some of them are misspecified.

Appendix: Proofs

Proof of Theorem 1: For each $i = 1$ and 2, Theorem 2 of Cho and Phillips (2018) implies that for $j = 1$ and 2, the local alternative approximations of $\widehat{\mathfrak{B}}_{j,n}^{(i)}$, $\widehat{\mathfrak{S}}_{j,n}^{(i)}$, and $\widehat{\mathfrak{E}}_{j,n}^{(i)}$ are equivalent and obtained as $\frac{1}{2}\mathrm{tr}[(\mathbf{V}_*^{(i)} + \sqrt{n}\mathbf{K}_{o,n}^{(i)})^2] + o_{\mathbb{P}}(1)$, where

$$\mathbf{K}_{o,n}^{(i)} := \mathbf{M}_{o,n}^{(i)} + \sum_{j=1}^{d}(\widehat{\theta}_{j,n} - \theta_{j*})\mathbf{S}_{j,*}^{(i)},$$

$$\mathbf{M}_{o,n}^{(i)} := (\mathbf{Q}_*^{(i)})^{-1}(\mathbf{Q}_n^{(i)}(\boldsymbol{\xi}_*) - \mathbf{P}_n^{(i)}(\boldsymbol{\xi}_*) - \mathbf{Q}_{*n}^{(i)} + \mathbf{P}_{*n}^{(i)}) \quad \text{and} \quad \mathbf{S}_{j*}^{(i)} := (\mathbf{P}_*^{(i)})^{-1}\left(\frac{\partial}{\partial\theta_j}\mathbf{Q}^{(i)}(\boldsymbol{\xi}_*) - \frac{\partial}{\partial\theta_j}\mathbf{P}_*^{(i)}(\boldsymbol{\xi}_*)\right).$$

Further, the symmetry between $\mathbf{P}_*^{(i)}$ and $\mathbf{Q}_*^{(i)}$ implies that for $j = 1$ and 2, the local alternative approximations of $\widetilde{\mathfrak{B}}_{j,n}^{(i)}$, $\widetilde{\mathfrak{S}}_{j,n}^{(i)}$, and $\widetilde{\mathfrak{E}}_{j,n}^{(i)}$ are equivalently obtained as $\frac{1}{2}\mathrm{tr}[(\widetilde{\mathbf{V}}_*^{(i)} + \sqrt{n}\widetilde{\mathbf{K}}_{o,n}^{(i)})^2] + o_{\mathbb{P}}(1)$, where

$$\widetilde{\mathbf{V}}_*^{(i)} := (\mathbf{P}_*^{(i)})^{-1}\bar{\mathbf{P}}_*^{(i)} - (\mathbf{Q}_*^{(i)})^{-1}\bar{\mathbf{Q}}_*^{(i)}, \quad \widetilde{\mathbf{K}}_{o,n}^{(i)} := \widetilde{\mathbf{M}}_{o,n}^{(i)} + \sum_{j=1}^{d}(\widehat{\theta}_{j,n} - \theta_{j*})\widetilde{\mathbf{S}}_{j,*},$$

$$\widetilde{\mathbf{M}}_{o,n}^{(i)} := (\mathbf{P}_*^{(i)})^{-1}(\mathbf{P}_n^{(i)}(\boldsymbol{\xi}_*) - \mathbf{Q}_n^{(i)}(\boldsymbol{\xi}_*) - \mathbf{P}_{*n}^{(i)} + \mathbf{Q}_{*n}^{(i)}) \quad \text{and} \quad \widetilde{\mathbf{S}}_{j*}^{(i)} := (\mathbf{Q}_*^{(i)})^{-1}\left(\frac{\partial}{\partial\theta_j}\mathbf{P}_*^{(i)}(\boldsymbol{\xi}_*) - \frac{\partial}{\partial\theta_j}\mathbf{Q}^{(i)}(\boldsymbol{\xi}_*)\right).$$

That is, $\mathbf{V}_*^{(i)} = -\widetilde{\mathbf{V}}_*^{(i)}$ and $\mathbf{K}_{o,n}^{(i)} = -\mathbf{K}_{o,n}^{(i)}$, so that for $j = 1$ and 2, the local alternative approximations of $\widehat{\mathfrak{B}}_{j,n}^{(i)}$, $\widehat{\mathfrak{S}}_{j,n}^{(i)}$, and $\widehat{\mathfrak{E}}_{j,n}^{(i)}$ are equivalent to the local alternative approximations of $\widetilde{\mathfrak{B}}_{j,n}^{(i)}$, $\widetilde{\mathfrak{S}}_{j,n}^{(i)}$, and $\widetilde{\mathfrak{E}}_{j,n}^{(i)}$. Therefore, it now follows that

$$\widehat{\mathfrak{M}}_n^{(i)} := \max_{j=1,2}\Big[\widehat{\mathfrak{B}}_{j,n}^{(i)}, \widehat{\mathfrak{S}}_{j,n}^{(i)}, \widehat{\mathfrak{E}}_{j,n}^{(i)}, \widetilde{\mathfrak{B}}_{j,n}^{(i)}, \widetilde{\mathfrak{S}}_{j,n}^{(i)}, \widetilde{\mathfrak{E}}_{j,n}^{(i)}\Big] = \frac{1}{2}\mathrm{tr}[(\mathbf{V}_*^{(i)} + \sqrt{n}\mathbf{K}_{o,n}^{(i)})^2] + o_{\mathbb{P}}(1).$$

In addition, Corollary 1 of Cho and Phillips (2018) implies that $\frac{1}{2}\mathrm{tr}[(\sqrt{n}\mathbf{K}_{o,n}^{(i)})^2] \Rightarrow \boldsymbol{\mathcal{Z}}^{(i)\prime}\boldsymbol{\Omega}_*^{(i)}\boldsymbol{\mathcal{Z}}^{(i)}$, suggesting that $\frac{1}{2}\mathrm{tr}[(\mathbf{V}_*^{(i)} + \sqrt{n}\mathbf{K}_{o,n}^{(i)})^2] \Rightarrow (\boldsymbol{\mathcal{Z}}^{(i)\prime} + \mathbf{V}_*^{(i)\prime}\boldsymbol{\Omega}_*^{(i)-1/2})\boldsymbol{\Omega}_*^{(i)}(\boldsymbol{\mathcal{Z}}^{(i)} + \boldsymbol{\Omega}_*^{(i)-1/2}\mathbf{V}_*^{(i)})$. Note that $\ddot{\mathbf{V}}_*^{(i)} := \boldsymbol{\Omega}_*^{(i)-1/2}\mathbf{V}_*^{(i)}$, so that $\widehat{\mathfrak{M}}_n^{(i)} \Rightarrow (\boldsymbol{\mathcal{Z}}^{(i)} + \ddot{\mathbf{V}}_*^{(i)})'\boldsymbol{\Omega}_*^{(i)}(\boldsymbol{\mathcal{Z}}^{(i)} + \ddot{\mathbf{V}}_*^{(i)})$, as desired. ■

Proof of Corollary 1: For each $i = 1$ and 2, given the null hypothesis $\mathcal{H}_0^{(i)}$, $\bar{\mathbf{P}}_*^{(i)} = \bar{\mathbf{Q}}_*^{(i)} = \mathbf{0}$. Therefore, Theorem 1 implies that $\widehat{\mathfrak{M}}_n^{(i)} \Rightarrow \boldsymbol{\mathcal{Z}}^{(i)\prime}\boldsymbol{\Omega}_*^{(i)}\boldsymbol{\mathcal{Z}}^{(i)}$. ■

Proof of Theorem 2: For $i = 1$ and 2, we first note that $\widehat{\mathfrak{M}}_n^{(i)} = \max[\dot{\mathfrak{M}}_n^{(i)}, \ddot{\mathfrak{M}}_n^{(i)}]$, where $\dot{\mathfrak{M}}_n^{(i)} := \max_{j=1,2}[\widehat{\mathfrak{B}}_{j,n}^{(i)}, \widehat{\mathfrak{S}}_{j,n}^{(i)}, \widehat{\mathfrak{E}}_{j,n}^{(i)}]$ and $\ddot{\mathfrak{M}}_n^{(i)} := \max_{j=1,2}[\widetilde{\mathfrak{B}}_{j,n}^{(i)}, \widetilde{\mathfrak{S}}_{j,n}^{(i)}, \widetilde{\mathfrak{E}}_{j,n}^{(i)}]$. Here, the leading term of $\dot{\mathfrak{M}}_n^{(i)}$ is determined by

$$\dot{\mu}_*^{(i)} := \max_{j=1,2}[\dot{\mathfrak{B}}_{j,*}^{(i)}, \dot{\mathfrak{S}}_{j,*}^{(i)}, \dot{\mathfrak{E}}_{j,*}^{(i)}],$$

where $\dot{\mathfrak{B}}_{2*}^{(i)} := (\dot{\delta}_*^{(i)})^2 + 2\dot{\zeta}_*^{(i)}$, $\dot{\mathfrak{S}}_{1*}^{(i)} := (\dot{\delta}_*^{(i)})^2 + 2\dot{\gamma}_*^{(i)}$, and $\dot{\mathfrak{S}}_{2*}^{(i)} := (\dot{\eta}_*^{(i)})^2 + 2\dot{\zeta}_*^{(i)}$. Here, $(\dot{\delta}_*^{(i)})^2$ is dominated by $(\dot{\tau}_*^{(i)})^2$ or $(\dot{\eta}_*^{(i)})^2$ from the fact that $\dot{\delta}_*^{(i)} \in [\dot{\eta}_*^{(i)}, \dot{\tau}_*^{(i)}]$ and $\dot{\eta}_*^{(i)} \geq -1$. Therefore, it now follows that $\dot{\mu}_*^{(i)} = \max[\dot{\mathfrak{B}}_{1,*}^{(i)}, \dot{\mathfrak{S}}_{2,*}^{(i)}, \dot{\mathfrak{E}}_{1,*}^{(i)}, \dot{\mathfrak{E}}_{2,*}^{(i)}]$.

Likewise, if we let $\ddot{\mu}_*^{(i)}$ be the leading term of $\ddot{\mathfrak{M}}_n^{(i)}$, $\ddot{\mu}_*^{(i)} = \max[\ddot{\mathfrak{B}}_{1,*}^{(i)}, \ddot{\mathfrak{S}}_{2,*}^{(i)}, \ddot{\mathfrak{E}}_{1,*}^{(i)}, \ddot{\mathfrak{E}}_{2,*}^{(i)}]$, where $\ddot{\mathfrak{S}}_{2,*}^{(i)} := (\ddot{\eta}_*^{(i)})^2 + 2\ddot{\gamma}_*^{(i)}$. Therefore, it now follows that $\mu_*^{(i)} = \max[\dot{\mathfrak{B}}_{1,*}^{(i)}, \dot{\mathfrak{S}}_{2,*}^{(i)}, \dot{\mathfrak{E}}_{1,*}^{(i)}, \dot{\mathfrak{E}}_{2,*}^{(i)}, \ddot{\mathfrak{B}}_{1,*}^{(i)}, \ddot{\mathfrak{S}}_{2,*}^{(i)}, \ddot{\mathfrak{E}}_{1,*}^{(i)}, \ddot{\mathfrak{E}}_{2,*}^{(i)}]$, and further $\dot{\mathfrak{S}}_{2,*}^{(i)} \leq \max[\dot{\mathfrak{E}}_{1,*}^{(i)}, \dot{\mathfrak{B}}_{1,*}^{(i)}]$ and $\ddot{\mathfrak{S}}_{2,*}^{(i)} \leq \max[\ddot{\mathfrak{E}}_{1,*}^{(i)}, \ddot{\mathfrak{B}}_{1,*}^{(i)}]$. Therefore, we can further simplify $\mu_*^{(i)}$: $\mu_*^{(i)} = \max[\dot{\mathfrak{B}}_{1,*}^{(i)}, \dot{\mathfrak{E}}_{1,*}^{(i)}, \dot{\mathfrak{E}}_{2,*}^{(i)}, \ddot{\mathfrak{B}}_{1,*}^{(i)}, \ddot{\mathfrak{E}}_{1,*}^{(i)}, \ddot{\mathfrak{E}}_{2,*}^{(i)}]$. This completes the proof. ■

References

ANDREWS, D.W.K. (1991): "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, 59, 817–858.

BAI, X. AND J. S. L. LAM, (2019): "A Copula-GARCH Approach for Analyzing Dynamic Conditional Dependency Structure between Liquefied Pertroleum Gas Freight Rate, Product Price Arbitrage and Crude Oil Price," *Energy Economics*, 78, 412–427.

BREUSCH, T. S. (1978): "Testing for Autocorrelation in Dynamic Linear Models," *Australian Economic Papers*, 17, 334–355.

BREUSCH, T. S. AND PAGAN, A. R. (1979): "A Simple Test for Heteroscedasticity and Random Coefficient Variation," *Econometrica*, 47, 1287–1294.

CHO, J.S., KIM, T.-H., AND SHIN, Y. (2015): "Quantile Cointegration in the Autoregressive Distributed Lag Modeling Framework," *Journal of Econometrics*, 188, 281–300.

CHO, J.S. AND PHILLIPS, P. C. B. (2018): "Pythagorean Generalization of Testing the Equality of Two Symmetric Positive Definite Matrices," *Journal of Econometrics*, 202, 45–56.

CHO, J.S. AND WHITE, H. (2011): "Generalized Runs Tests for the IID Hypothesis," *Journal of Econometrics*, 162, 326–344.

CHO, J.S. AND WHITE, H. (2015): "Testing the Equality of Two Positive-Definite Matrices with Application to Information Matrix Testing," *Essays in Honor of Peter CB Phillips*, 491–556.

DURNIN, J. AND WATSON, G. S. (1950): "Testing for Serial Correlation in Least Squares Regression, I," *Biometrika*, 37, 409–428.

DURNIN, J. AND WATSON, G. S. (1951): "Testing for Serial Correlation in Least Squares Regression, II," *Biometrika*, 38, 159–179.

EFRON, B. AND TIBSHARANI, R. J. (1988): *An Introduction to the Bootstrap*. Chapman & Hall/CRC.

ENGLE, R. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1008.

GALLANT, R. (1987): *Nonlinear Statistical Models*, New York: John Wiley & Sons.

GODFREY, L. G. (1978a): "Testing against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables," *Econometrica*, 46, 1293–1301.

GODFREY, L. G. (1978b): "Testing for Multiplicative Heteroskedasticity," *Journal of Econometrics*, 8, 227–236.

GOURIEROUX, C., MONFORT, A., AND TROGNON, A. "Pseudo Maximum Likelihood Methods: Theory," *Econometrica*, 52, 681–700.

HONG, Y. AND WHITE, H. (2005): "Asymptotic Distribution Theory for Nonparametric Entropy Measures of Serial Dependence," *Econometrica*, 73, 837–901.

LJUNG, G. M. AND BOX, G. E. P. (1978): "On a Measure of a Lack of Fit in Time Series Models," *Biometrika*, 65, 297–303.

MANKIW, N., R0MER, D., AND WEIL, D. (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 407–437.

NEWEY, W. K. AND WEST, K. D. (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

NG, S. AND PERRON, P. (1996): “The Exaxt Error in Estimating the Spectral Density at the Origin,” *Journal of Time Series Analysis*, 17, 379–408.

ROBINSON, P. M. (1991): “Consistent Nonparametric Entropy-Based Testing,” *Review of Economic Studies*, 58, 437–453.

SKAUG, H. AND TJØSTHEIM, D. (1996): “Measures of Distance between Densities with Applications to Testing for Serial Independence,” *Time Series Analysis in Memory of E. J. Hannan*, ed. P. Robinson and M. Rosenblatt. New York: Springer-Verlag, 363-377.

SOLOW, R. (1957): “Technical Change and the Aggregate Production Function,” *Review of Economics and Statistics*, 39, 312–320.

WALD, A. (1943): “Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large,” *Transactions of American Mathematical Society*, 54, 426–482.

White, H. (1980): “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity,” *Econometrica*, 48, 817–838.

White, H. (1982): “Maximum Likelihood Estimation of Misspecified Models,,” *Econometrica*, 48, 1–25.

WHITE, H. AND DOMOWITZ, I. (1984): “Nonlinear Regression with Dependent Observations,,” *Econometrica*, 52, 143–161.

WU, C.F.J. (1986): “Jacknife Bootstrap and Other Resampling Methods in Regression Analysis,” *Annlas of Statistics*, 14, 1261–1295.

DGP1: $\mathcal{H}_0^{(1)}$ or $\mathcal{H}_0^{(2)}$							
Test Statistics	Levels$\backslash n$	200	400	600	800	1,000	1,200
	1%	0.93	0.90	0.70	1.17	0.77	1.03
$\widehat{\mathfrak{M}}_n^{(1)}$	5%	4.70	4.63	4.77	4.47	4.30	4.80
	10%	9.23	8.83	9.73	9.53	9.23	9.97
	1%	0.27	0.60	0.87	0.67	0.60	1.23
$\widehat{\mathfrak{M}}_n^{(2)}$	5%	2.83	4.07	3.97	3.83	4.10	6.00
	10%	6.10	8.83	8.60	8.37	7.63	11.40

DGP2: $\mathcal{H}_1^{(1)}$ or $\mathcal{H}_0^{(2)}$							
Test Statistics	Levels$\backslash n$	200	400	600	800	1,000	1,500
	1%	13.00	35.93	57.00	73.97	84.93	96.83
$\widehat{\mathfrak{M}}_n^{(1)}$	5%	28.63	56.33	76.13	87.03	94.43	98.97
	10%	39.33	66.30	83.10	91.93	96.70	99.43
	1%	0.33	0.57	0.97	0.87	0.67	1.20
$\widehat{\mathfrak{M}}_n^{(2)}$	5%	3.20	5.20	5.03	5.23	5.00	5.63
	10%	8.07	11.37	11.03	11.20	11.57	11.73

DGP3: $\mathcal{H}_1^{(1)}$ or $\mathcal{H}_1^{(2)}$							
Test Statistics	Levels$\backslash n$	200	400	600	800	1,000	2,000
	1%	14.70	38.90	58.87	75.70	84.80	99.33
$\widehat{\mathfrak{M}}_n^{(1)}$	5%	29.10	59.27	77.57	88.87	93.93	99.87
	10%	39.80	68.10	84.13	93.33	96.33	100.0
	1%	5.17	18.70	24.53	30.07	37.27	67.73
$\widehat{\mathfrak{M}}_n^{(2)}$	5%	18.13	40.27	48.63	54.07	60.87	86.10
	10%	28.83	51.97	59.17	65.53	70.27	91.43

Table 1: Empirical Rejection Rates of the Maximum Test Statistics under the Null and Alternative Hypotheses (in Percent). This table shows the rejection rates of the maximum test statistics $\widehat{\mathfrak{M}}_n^{(1)}$ and $\widehat{\mathfrak{M}}_n^{(2)}$ under $\mathcal{H}_0^{(1)} : \sigma_*^2 \mathbb{E}[\mathbf{X}_t \mathbf{X}_t'] = \mathbb{E}[U_t^2 \mathbf{X}_t \mathbf{X}_t']$, $\mathcal{H}_1^{(1)} : \sigma_*^2 \mathbb{E}[\mathbf{X}_t \mathbf{X}_t'] \neq \mathbb{E}[U_t^2 \mathbf{X}_t \mathbf{X}_t']$, $\mathcal{H}_0^{(2)} : \mathbb{E}[U_t^2 \mathbf{X}_t \mathbf{X}_t'] = \lim_{n\to\infty} n^{-1} \sum_{t=1}^n \sum_{\tau=1 \, t\neq\tau}^n \mathbb{E}[U_t U_\tau \mathbf{X}_t \mathbf{X}_\tau']$, and $\mathcal{H}_1^{(2)} : \mathbb{E}[U_t^2 \mathbf{X}_t \mathbf{X}_t'] \neq \lim_{n\to\infty} n^{-1} \sum_{t=1}^n \sum_{\tau=1 \, t\neq\tau}^n \mathbb{E}[U_t U_\tau \mathbf{X}_t \mathbf{X}_\tau']$. Here, the HC and HAC covariance matrix estimators are used to estimate the covariance matrices consistently. In particular, Andrews's (1991) quadratic kernel is employed to estimate the HAC covariance matrix estimator with the bandwidth $\lfloor n^{1/5} \rfloor - 1$.

			DGP1: $\mathcal{H}_0^{(1)}$ or $\mathcal{H}_0^{(2)}$			
Hypotheses $\backslash n$	200	400	600	800	1,000	1,200
$\mathcal{H}_0^{(1)}$	92.57	91.40	91.40	91.93	91.80	89.57
$\mathcal{H}_0^{(2)}$	4.60	4.53	4.63	4.23	4.10	4.43
$\mathcal{H}_1^{(2)}$	2.83	4.07	3.97	3.83	4.10	6.00
Sum	100.0	100.0	100.0	100.0	100.0	100.0
			DGP2: $\mathcal{H}_1^{(1)}$ or $\mathcal{H}_0^{(2)}$			
Hypotheses $\backslash n$	200	400	600	800	1,000	1,500
$\mathcal{H}_0^{(1)}$	69.17	41.67	23.00	12.37	5.30	0.97
$\mathcal{H}_0^{(2)}$	27.63	53.13	71.97	82.40	89.70	93.40
$\mathcal{H}_1^{(2)}$	3.20	5.20	5.03	5.23	5.00	5.63
Sum	100.0	100.0	100.0	100.0	100.0	100.0
			DGP3: $\mathcal{H}_1^{(1)}$ or $\mathcal{H}_1^{(2)}$			
Hypotheses $\backslash n$	200	400	600	800	1,000	2,000
$\mathcal{H}_0^{(1)}$	57.80	24.87	10.60	4.83	1.87	0.00
$\mathcal{H}_0^{(2)}$	24.07	34.87	40.77	41.10	37.27	13.90
$\mathcal{H}_1^{(2)}$	18.13	40.27	48.63	54.07	60.87	86.10
Sum	100.0	100.0	100.0	100.0	100.0	100.0

Table 2: EMPIRICAL REJECTION RATES OF THE TEST STATISTICS USING THE STP (IN PERCENT). This table shows the proportion of the hypotheses selected by the STP when the maximum test statistics $\widehat{\mathfrak{M}}_n^{(1)}$ and $\widehat{\mathfrak{M}}_n^{(2)}$ are applied to the data observations generated under DGP1, -2, and -3. As before, $\widehat{\mathfrak{M}}_n^{(1)}$ and $\widehat{\mathfrak{M}}_n^{(2)}$, respectively test $\mathcal{H}_0^{(1)} : \sigma_*^2\mathbb{E}[\mathbf{X}_t\mathbf{X}_t'] = \mathbb{E}[U_t^2\mathbf{X}_t\mathbf{X}_t']$ against $\mathcal{H}_1^{(1)} : \sigma_*^2\mathbb{E}[\mathbf{X}_t\mathbf{X}_t'] \neq \mathbb{E}[U_t^2\mathbf{X}_t\mathbf{X}_t']$; and $\mathcal{H}_0^{(2)} : \mathbb{E}[U_t^2\mathbf{X}_t\mathbf{X}_t'] = \lim_{n\to\infty} n^{-1}\sum_{t=1}^n\sum_{\tau=1\, t\neq\tau}^n \mathbb{E}[U_tU_\tau\mathbf{X}_t\mathbf{X}_\tau']$ against $\mathcal{H}_0^{(2)} : \mathbb{E}[U_t^2\mathbf{X}_t\mathbf{X}_t'] \neq \lim_{n\to\infty} n^{-1}\sum_{t=1}^n\sum_{\tau=1\, t\neq\tau}^n \mathbb{E}[U_tU_\tau\mathbf{X}_t\mathbf{X}_\tau']$. For the implementation of the STP, we first test $\mathcal{H}_0^{(2)}$ and then test $\mathcal{H}_0^{(1)}$ if $\mathcal{H}_0^{(2)}$ is rejected. We also let the level of significance α be 5%. Here, the HC and HAC covariance matrix estimators are used to estimate the covariance matrices consistently. In particular, Andrews's (1991) quadratic kernel is employed to estimate the HAC covariance matrix estimator with the bandwidth $\lfloor n^{1/5}\rfloor - 1$.

	Non-Oil	Intermediate	OECD
Constant	1.9194	2.2497	2.1404
Std. Error	0.8337	0.8547	1.1807
Heteroskedasticity-Robust Std. Error	0.8042	0.9246	1.1445
$\log(Y_{60,t})$	-0.1409	-0.2278	-0.3499
Std. Error	0.0520	0.0573	0.0657
Heteroskedasticity-Robust Std. Error	0.0482	0.0560	0.0649
$\log(s_t)$	0.6472	0.6459	0.3901
Std. Error	0.0867	0.1039	0.1761
Heteroskedasticity-Robust Std. Error	0.1010	0.1318	0.2299
$\log(n_t + g_t + \delta_t)$	-0.3023	-0.4575	-0.7662
Std. Error	0.3044	0.3074	0.3452
Heteroskedasticity-Robust Std. Error	0.2411	0.2363	0.2704
	p-values for Heteroskedasticity (in Percent)		
$\widehat{\mathfrak{M}}_n^{(1)}$	**0.49**	**2.48**	61.33
Breusch-Pagan's (1978) test	**2.74**	**3.06**	21.87
White's (1980) test	6.58	**1.37**	51.31

Table 3: p-VALUE OF TESTING FOR CONDITIONAL HETEROSKEDASTICITY BY THE MATRIX EQUALITY TESTING (IN PERCENT). The dependent variable of the model is the log of GDP per working-age person in 1985. This table shows the p-values of the matrix equality tests using the same data and models in Mankiw, Romer, and Weil (1992). The table also provides Breusch and Pagan's (1978) and White's (1980) conditional heteroskedasticity test statistics. The standard errors of the LS estimators are obtained by assuming conditional homoskedasticity and heteroskedasticity. The p-values denoted by the boldface font indicate that they are less than 5%. As a result, we obtain the same inference result as in Mankiw, Romer, and Weil (1992), although they do not accommodate conditional heteroskedasticity when testing zero coefficients using the t-test statistics.

	BLPG		BRENT		MB		PAFEI	
Lags	$\mathcal{H}_0^{(1)}$ vs. $\mathcal{H}_1^{(1)}$	$\mathcal{H}_0^{(2)}$ vs. $\mathcal{H}_1^{(2)}$	$\mathcal{H}_0^{(1)}$ vs. $\mathcal{H}_1^{(1)}$	$\mathcal{H}_0^{(2)}$ vs. $\mathcal{H}_1^{(2)}$	$\mathcal{H}_0^{(1)}$ vs. $\mathcal{H}_1^{(1)}$	$\mathcal{H}_0^{(2)}$ vs. $\mathcal{H}_1^{(2)}$	$\mathcal{H}_0^{(1)}$ vs. $\mathcal{H}_1^{(1)}$	$\mathcal{H}_0^{(2)}$ vs. $\mathcal{H}_1^{(2)}$
1	0.00	**26.67**	0.03	**96.67**	0.01	6.31	0.86	**68.25**
2	0.01	**81.82**	0.11	**40.23**	0.04	4.49	0.59	**80.64**
3	0.00	**86.05**	0.10	**30.20**	0.07	4.18	0.09	**91.11**
4	0.01	**94.97**	0.02	**6.06**	0.13	2.51	0.02	**98.84**
5	0.03	**96.62**	0.05	**11.88**	0.01	1.85	0.10	**96.76**
6	0.04	**99.55**	0.02	**12.89**	0.04	4.15	0.10	**94.15**
7	0.09	**80.45**	0.01	**10.60**	0.03	3.26	0.05	**71.65**
8	0.38	**43.99**	0.01	**12.80**	0.04	2.37	0.01	**80.64**
9	1.13	**43.57**	0.01	**17.04**	0.08	2.88	0.03	**71.58**
10	1.49	**44.59**	0.03	**16.40**	0.18	3.95	0.06	**66.72**

	CPS		AFEIUS		AFEICPS			
Lags	$\mathcal{H}_0^{(1)}$ vs. $\mathcal{H}_1^{(1)}$	$\mathcal{H}_0^{(2)}$ vs. $\mathcal{H}_1^{(2)}$	$\mathcal{H}_0^{(1)}$ vs. $\mathcal{H}_1^{(1)}$	$\mathcal{H}_0^{(2)}$ vs. $\mathcal{H}_1^{(2)}$	$\mathcal{H}_0^{(1)}$ vs. $\mathcal{H}_1^{(1)}$	$\mathcal{H}_0^{(2)}$ vs. $\mathcal{H}_1^{(2)}$		
1	1.48	**57.92**	2.50	**6.51**	0.03	0.15		
2	1.41	**10.47**	4.06	**15.34**	0.06	1.06		
3	1.47	**23.08**	2.91	**15.47**	0.06	**5.01**		
4	2.79	**29.61**	3.32	**14.77**	0.22	**9.37**		
5	4.64	**16.01**	4.38	**13.72**	0.28	**8.91**		
6	4.59	**31.65**	**8.55**	**13.65**	0.37	**6.09**		
7	1.74	**21.51**	**12.34**	**13.24**	0.40	**8.92**		
8	0.30	**30.13**	**21.39**	**27.07**	0.81	**14.74**		
9	0.26	**39.55**	**29.71**	**43.42**	1.11	4.19		
10	0.54	**42.43**	**41.11**	**39.27**	1.30	**8.90**		

Table 4: p-VALUES OF TESTING FOR THE INFLUENCE OF CONDITIONAL HETEROSKEDASTICITY AND AUTOCORRELATION (IN PERCENT). This table shows the p-values of the maximum test statistics $\widehat{\mathfrak{M}}_n^{(1)}$ and $\widehat{\mathfrak{M}}_n^{(2)}$ obtained by applying them to the models and energy data in Bai and Lam (2019). The p-values denoted by the boldface font indicate that they are greater than 5%. As a result, the AR(1) and AR(2) models are dynamically and correctly specified for BLPG, BRENT, PAFEI, CPS, and AFEIUS. On the contrary, the AR(1) model specified for MB and AFEICPS is misspecified.

	BLPG	BRENT	MB	PAFEI	CPS	AFEIUS	AFEICPS
Information matrix test statistic	31.80	22.90	63.05	83.55	56.90	0.25	0.35

Table 5: p-VALUE OF THE INFORMATION MATRIX EQUALITY TEST (IN PERCENT). This table shows the p-values of the maximum test statistic of the correct specification for the AR-GARCH(1,1) model estimated by assuming that the standardized error follows a skewed t-distribution. Using the maximum test statistic, we test the information matrix equality. As a result, the maximum test statistic cannot reject the hypothesis that the AR-GARCH(1,1) model driven by the skewed t-distribution is correctly specified for BLPG, BRENT, MB, PAFEI, and CPS, whereas the same model is misspecified for AFEIUS and AFEICPS.