Testing for the Sandwich-Form Covariance Matrix of the Quasi-Maximum Likelihood Estimator

LIJUAN HUO

School of Humanities and Social Sciences, Beijing Institute of Technology, Haidian, Beijing, 100081, China

ljhuo@bit.edu.cn

JIN SEO CHO

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

jinseocho@yonsei.ac.kr

This version: February 2020

Abstract

This study tests for the sandwich-form asymptotic covariance matrices entailed by conditional heteroskedastic and/or autocorrelated regression errors or conditionally uncorrelated homoskedastic errors. In doing so, we enable the empirical researcher to estimate the asymptotic covariance matrix of the quasi-maximum likelihood estimator by a proper estimator. Accordingly, we provide test methodologies by extending the approaches in Cho and White (2014) and Cho and Phillips (2018a) to detect the influence of heteroskedastic and/or autocorrelated regression errors on the asymptotic covariance matrix. In particular, we establish a sequential testing procedure to achieve our goal. We affirm the theory on our test statistics through simulation and apply the test statistics to energy price growth rate data for illustrative purposes; here, we also apply our test methodology to test the fully correct model hypothesis.

Key Words: Information matrix equality; sandwich-form covariance matrix; heteroskedasticity-consistent covariance matrix estimator; heteroskedasticity and autocorrelation-consistent covariance matrix estimator.

JEL Codes: C12, C22, O47, Q47.

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 (IM) equality holds so that the standard t-test statistic can be used to infer unknown parameters.

Nevertheless, the asymptotic covariance matrix displays a sandwich form in a general environment, allowing the standard t-ratio to be redefined by the heteroskedasticity-consistent (HC) or heteroskedasticity and autocorrelation-consistent (HAC) covariance matrix estimator. As conditionally heteroskedastic and/or autocorrelated error properties are only required for a sandwich-form asymptotic covariance matrix, the failure of the IM equality entails regression errors influenced by either conditional heteroskedasticity or autocorrelation, necessitating the estimation of the covariance matrix by the HC or HAC estimator. Furthermore, if the empirical researcher knows how the covariance matrix is affected by conditional heteroskedastic and/or autocorrelated errors, a more pertinent estimator can be applied to estimate the asymptotic covariance matrix. If the regression errors exhibit conditional homoskedastic and MDS properties, it is redundant to estimate the covariance matrix by the HC or HAC estimator.

This study primarily seeks to provide a test methodology to detect a sandwich-form asymptotic covariance matrix directly and further identify how conditional heteroskedasticity and/or autocorrelation affect the asymptotic covariance matrix. This differs from the conventional approach to testing a sandwich-form asymptotic covariance matrix by testing conditionally heteroskedastic and/or autocorrelated errors indirectly. As they are only necessary conditions, a sandwich-form asymptotic covariance matrix is not necessarily obtained by conditionally heteroskedastic and/or autocorrelated errors. Another goal is to provide an empirical analysis using the proposed test methodology in classical regression and time-series data analysis settings. This serves the dual purpose of illustrating our method and reinforcing the inferential process on regression analysis.

We apply the methodologies proposed by Cho and White (2014, CW) and Cho and Phillips (2018a, CP) to the quasi-maximum likelihood (QML) estimator, which nests the LS estimator as a special case. CW and CP 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 test methodologies. We apply their methodologies to test for a sandwich-form asymptotic covariance matrix entailed by conditional heteroskedastic and/or autocorrelated regression errors. Specifically, we first generalize the maximum test statistic in CP in our context so that our test statistics have greater power than the maximum test statistic in CP. Next, we separately construct two maximum test statistics for the sandwich-form asymptotic covariance matrix entailed by conditional heteroskedastic and autocorrelated errors. Specifically, the first maximum test statistic is designed to test the covariance matrix entailed by conditionally homoskedastic MDS errors, whereas the second tests the covariance matrix affected by conditionally heteroskedastic errors against the matrix affected by serially correlated errors. Although failure to reject the null hypotheses using the first and second maximum tests do not necessarily imply conditionally homoskedastic MDS and uncorrelated errors, respectively, our statistics are omnibus test statistics. Therefore, rejecting the null hypothesis implies a consistent detection of the sandwich-form covariance matrix entailed by conditional heteroskedastic and/or autocorrelated errors.

Furthermore, we develop a sequential testing procedure (STP) by implementing our maximum test statistics systematically. Although the two statistics are separately designed to test for the sandwich-form covariance matrices entailed by conditional heteroskedastic and autocorrelated errors, respectively, the empirical researcher can control the overall type-I error of the two test statistics by applying them in a proper order. Specifically, we first apply the second maximum statistic and test for the sandwich-form covariance matrix affected by autocorrelated errors. If we cannot reject the null, we next apply the first maximum statistic to test for the sandwich-form covariance matrix entailed by conditionally heteroskedastic errors. As discussed below, this test plan can consistently identify how conditional heteroskedasticity and/or autocorrelation affects the asymptotic covariance matrix.

The test methodology in this study is especially useful if the maximum test statistics defined herein are combined with resampling bootstraps. The residual bootstrap and wild bootstrap methods of Efron and Tibsharani (1988) and Wu (1986) are particularly useful for enhancing the utility of the maximum test statistics. Although this study provides asymptotic null limit distributions for the maximum test statistics, they may lead to relatively large level 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 confirm IM equality, leading to large degrees of freedom. We overcome the level distortion problem by applying the bootstrap methods to the maximum test statistics. Indeed, the resampling methods are efficient remedies for the level distortion problem. We provide simulation results to affirm our theory on the maximum test statistics and STP.

For the second goal of illustrating the test methodology, we revisit empirical studies on energy price growth rates. Estimating the model using autoregressive (AR) and generalized autoregressive conditional heteroskedastic (GARCH) models for the conditional mean and variance of the energy price growth rate, respectively, is standard. 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 if the skewed t-distribution is correct for the standardized error, their model must be entirely correct, and their QML estimator becomes a maximum likelihood (ML) estimator, leading to IM equality for their estimator. Conversely, if only the distribution of the standardized error is misspecified, their QML estimator consistently estimates the conditional mean and variance equations, while possessing a sandwich-form asymptotic covariance matrix for the QML estimator.

This empirical outline implies that we can apply these test methodologies for purposes other than the goals of this study. In addition to applying the STP to the AR and GARCH models, we can further test the fully correct model hypothesis by testing the IM equality of the (Q)ML estimator. Although failure to reject IM equality does not mean that the model is entirely correct, rejecting the IM equality means that the full model is indeed misspecified. Therefore, we can associate this rejection with the maximum test results to detect how the model is misspecified. We detail our illustration using the models and data in Bai and Lam (2019).

The empirical applicability of the test methodology is potentially huge, given that estimating a sandwich-form asymptotic covariance matrix is necessary for many cross-sectional and time-series data analyses. Our test methodology can be applied even when the error term is present is a multiplicative form, although we focus on an additive error case in the current study.

The remainder of the paper is structured as follows. Section 2 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 an empirical example. 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 the goals. 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 by 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 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 $\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_*) \overset{A}{\sim} N(\mathbf{0}, \mathbf{A}_*^{-1})$; otherwise, it generally holds that

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

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, $\mathbf{A}_* := -\mathbb{E}[\nabla_{\boldsymbol{\theta}}^2 \ell_t(\boldsymbol{\theta}_*)]$ 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 $m(\mathbf{X}_t, \boldsymbol{\theta}_*) = \mathbf{X}_t' \boldsymbol{\theta}_*$, $\mathbf{A}_* = \sigma_*^{-2} \mathbb{E}[\mathbf{X}_t \mathbf{X}_t']$ and $\mathbf{B}_* = \sigma_*^{-4} \mathbb{E}[U_t^2 \mathbf{X}_t \mathbf{X}_t']$, 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 IM 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 IM inequality holds. If U_t is autocorrelated, the asymptotic covariance matrix $\mathbf{B}_*$ in (1) is modified to

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

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

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

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 IM equality holds even with conditionally heteroskedastic and autocorrelated regression errors.

This aspect motivates to directly 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:

$$\dot{\mathcal{H}}_0 : \mathbf{A}_* = \mathbf{B}_* \quad \text{vs.} \quad \dot{\mathcal{H}}_1 : \mathbf{A}_* \neq \mathbf{B}_*; \quad \text{and} \quad \ddot{\mathcal{H}}_0 : \mathbf{B}_* = \mathbf{C}_* \quad \text{versus} \quad \ddot{\mathcal{H}}_1 : \mathbf{B}_* \neq \mathbf{C}_*.$$

$\dot{\mathcal{H}}_0$ and $\dot{\mathcal{H}}_1$ are provided to test for the sandwich-form asymptotic covariance matrix entailed by conditional heteroskedastic errors, whereas $\ddot{\mathcal{H}}_0$ and $\ddot{\mathcal{H}}_1$ 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 $\ddot{\mathcal{H}}_0$ 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 CW and CP. They provide simple necessary and sufficient conditions for the equality of two symmetric positive-definite matrices. For example, when we test $\dot{\mathcal{H}}_0$ against $\dot{\mathcal{H}}_1$, any two of the following conditions hold:

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

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:

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

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 test 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 $\widehat{\boldsymbol{\xi}}_n := (\widehat{\boldsymbol{\theta}}'_n, \widehat{\sigma}^2_n)'$: $\widehat{\boldsymbol{\xi}}_n := \arg\max_{\boldsymbol{\xi} \in \Xi} \bar{L}_n(\boldsymbol{\xi})$, where $\bar{L}_n(\cdot) := n^{-1} \sum_{t=1}^n \ell_t(\cdot)$, and $\ell_t(\cdot)$ is the individual quasi-log-likelihood function defined on $\Xi := \Theta \times [c, d]$ (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., $\widehat{\mathbf{A}}_n \overset{\mathbb{P}}{\to} \mathbf{A}_*$, $\widehat{\mathbf{B}}_n \overset{\mathbb{P}}{\to} \mathbf{B}_*$, and $\widehat{\mathbf{C}}_n \overset{\mathbb{P}}{\to} \mathbf{C}_*$. For example, we can let $\widehat{\mathbf{A}}_n := -n^{-1} \sum_{t=1}^n \nabla^2_{\boldsymbol{\theta}} \ell_t(\widehat{\boldsymbol{\xi}}_n)$, $\widehat{\mathbf{B}}_n := n^{-1} \sum_{t=1}^n \widehat{\boldsymbol{s}}_t \widehat{\boldsymbol{s}}'_t$, and

$$\widehat{\mathbf{C}}_n := \frac{1}{n} \sum_{t=1}^n \widehat{\boldsymbol{s}}_t \widehat{\boldsymbol{s}}'_t + \frac{1}{n} \sum_{k=1}^{\ell} \omega_{\ell k} \sum_{t=k+1}^n \left(\widehat{\boldsymbol{s}}_{t-k} \widehat{\boldsymbol{s}}'_t + \widehat{\boldsymbol{s}}_t \widehat{\boldsymbol{s}}'_{t-k} \right),$$

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 with $\ell = O(n^{1/5})$. 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 CW and CP. For example, if $\dot{\mathcal{H}}_0$ is tested against $\dot{\mathcal{H}}_1$, we let $\widehat{\mathbf{D}}_n := \widehat{\mathbf{A}}_n \widehat{\mathbf{B}}_n^{-1}$ and define the following statistics:

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

These statistics are introduced to consistently estimate

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

respectively.

The motivations of these statistics stem from the fact that $d^{-1}\text{tr}[\widehat{\mathbf{D}}]$, $\det[\widehat{\mathbf{D}}]^{\frac{1}{d}}$, and $d\text{tr}[\widehat{\mathbf{D}}_n^{-1}]^{-1}$ 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 $\dot{\mathcal{H}}_0$ against $\dot{\mathcal{H}}_1$ and define the following test statistics:

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

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 $\ddot{\mathcal{H}}_0$ against $\ddot{\mathcal{H}}_1$, 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 $\widehat{\mathfrak{B}}_{1,n} := (nd/2)(\widehat{\tau}_n^2 + 2\widehat{\zeta}_n)$, $\widehat{\mathfrak{B}}_{2,n} := (nd/2)(\widehat{\delta}_n^2 + 2\widehat{\zeta}_n)$, $\widehat{\mathfrak{S}}_{1,n} := (nd/2)(\widehat{\delta}_n^2 + 2\widehat{\gamma}_n)$, $\widehat{\mathfrak{S}}_{2,n} := (nd/2)(\widehat{\eta}_n^2 + 2\widehat{\gamma}_n)$, $\widehat{\mathfrak{E}}_{1,n} := (nd/2)(\widehat{\tau}_n^2 + 2\widehat{\gamma}_n)$, and $\widehat{\mathfrak{E}}_{2,n} := (nd/2)(\widehat{\eta}_n^2 + 2\widehat{\zeta}_n)$. 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, CP 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 CP. 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 CP 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, CP 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 $\dot{\mathcal{H}}_0$ against $\dot{\mathcal{H}}_1$ and $\ddot{\mathcal{H}}_0$ against $\ddot{\mathcal{H}}_1$, 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 CP. 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 $\dot{\mathcal{H}}_0$ or $\ddot{\mathcal{H}}_0$. For example, we let $\mathbf{P}^{(1)}(\boldsymbol{\xi}_*) = \mathbf{A}_*$ and $\mathbf{P}^{(2)}(\boldsymbol{\xi}_*) = \mathbf{B}_*$ under $\dot{\mathcal{H}}_0$ and $\ddot{\mathcal{H}}_0$, respectively, which are the left-hand side matrices in $\dot{\mathcal{H}}_0$ and $\ddot{\mathcal{H}}_0$, respectively. Likewise, for $i = 1$ and 2, 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 matrices in $\dot{\mathcal{H}}_0$ and $\ddot{\mathcal{H}}_0$, respectively. 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 $\dot{\mathcal{H}}_0$.

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

(iii) The sequence of $\{\widehat{\boldsymbol{\xi}}_n := (\widehat{\boldsymbol{\theta}}_n', \widehat{\sigma}_n^2)' : \Omega \mapsto \Xi\}$ *converges to* $\boldsymbol{\xi}_* := (\boldsymbol{\theta}_*', \sigma_*^2)' \in \Xi$ *in probability, where* $\boldsymbol{\xi}_*$ *is unique and interior to* Ξ, $\widehat{\boldsymbol{\xi}}_n := \arg\max_{\boldsymbol{\xi} \in \Xi} \bar{L}_n(\boldsymbol{\xi})$ *with* $\bar{L}_n(\boldsymbol{\xi}) := n^{-1} \sum_{t=1}^n \ell_t(\boldsymbol{\xi})$, *and* $\ell_t(\cdot) := \ell(\mathbf{Z}_t, \cdot)$ *is the quasi-log-likelihood function assuming a correctly specified model* $\mathcal{M} := \{m(\cdot, \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d\}$ *for* $E[Y_t | \mathbf{X}_t]$ *and a constant conditional variance such that* $\ell_t(\cdot)$ *is continuous on* Ξ *almost surely* $-\mathbb{P}$;

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

(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* $[\mathbf{Z}_{\boldsymbol{\xi}}', \text{vech}(\mathbf{Z}_{\mathbf{P}^{(i)}})', \text{vech}(\mathbf{Z}_{\mathbf{Q}^{(i)}})']' \sim N(\mathbf{0}, \boldsymbol{\Sigma}_*^{(i)})$, *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* $\mathbf{P}^{(i)} : \Xi \mapsto \mathbb{R}^{d \times d}$ *and* $\mathbf{Q}^{(i)} : \Xi \mapsto \mathbb{R}^{d \times d}$ *are in* $\mathcal{C}^{(2)}(\Xi)$; *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* $\bar{\mathbf{P}}^{(i)} : \Xi \mapsto \mathbb{R}^{d \times d}$ *and* $\bar{\mathbf{Q}}^{(i)} : \Xi \mapsto \mathbb{R}^{d \times d}$ *are in* $\mathcal{C}^{(1)}(\Xi)$;

(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* Ξ *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 ($\dot{\mathcal{H}}_0$ and $\ddot{\mathcal{H}}_0$), alternative hypotheses ($\dot{\mathcal{H}}_1$ and $\ddot{\mathcal{H}}_1$), and following local alternative hypothe-

ses:

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

- $\mathcal{H}_\ell^{(2)}: \mathbf{B}_* = \mathbf{C}_*$ with $\mathbf{B}_{*n} = \mathbf{B}_* + n^{-1/2}\bar{\mathbf{B}}_* + o(n^{-1/2})$ 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 $\dot{\mathcal{H}}_0$ and $\ddot{\mathcal{H}}_0$, 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.

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, $\widehat{\mathfrak{M}}_n^{(i)} \Rightarrow (\mathcal{Z}^{(i)} + \ddot{\mathbf{V}}_*^{(i)})'\Omega_*^{(i)}(\mathcal{Z}^{(i)} + \ddot{\mathbf{V}}_*^{(i)})$ under $\mathcal{H}_\ell^{(i)}$, where*

$$\mathcal{Z}^{(i)} := \left[\begin{array}{c} \text{vec}(\mathbf{Z}_{\mathbf{Q}^{(i)}} - \mathbf{Z}_{\mathbf{P}^{(i)}}) \\ \text{vec}(\mathbf{Z}_{\boldsymbol{\theta}}' \otimes \mathbf{I}_d) \end{array} \right], \quad \Omega_*^{(i)} := \left[\begin{array}{cc} (\mathbf{P}_*^{(i)})^{-1} \otimes (\mathbf{P}_*^{(i)})^{-1} & (\mathbf{P}_*^{(i)})^{-1}\mathbf{R}_*^{(i)\prime} \otimes (\mathbf{P}_*^{(i)})^{-1} \\ (\mathbf{P}_*^{(i)})^{-1} \otimes \mathbf{R}_*^{(i)}(\mathbf{P}_*^{(i)})^{-1} & (\mathbf{P}_*^{(i)})^{-1}\mathbf{R}_*^{(i)\prime} \otimes \mathbf{R}_*^{(i)}(\mathbf{P}_*^{(i)})^{-1} \end{array} \right]$$

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

The local alternative limit distribution in Theorem 1 is obtained by deriving the local alternative weak limit of each maximand as in CP. 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 $(\mathcal{Z}^{(i)} + \mathbf{V}_*^{(i)})'\Omega_*^{(i)}(\mathcal{Z}^{(i)} + \mathbf{V}_*^{(i)})$ for $i = 1$ and 2, so that their maximum also weakly converges to $(\mathcal{Z}^{(i)} + \mathbf{V}_*^{(i)})'\Omega_*^{(i)}(\mathcal{Z}^{(i)} + \mathbf{V}_*^{(i)})$ 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.

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

Corollary 1 simply follows by letting $\mathbf{V}_*^{(i)} = \mathbf{0}$ in Theorem 1. 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. The finite sample level distortion can be substantially large if the asymptotic critical values apply, because the degrees of freedom in the null limit distribution can be large. Note that $\mathcal{Z}^{(i)}$ 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 substantially differ from the null limit distributions. Instead, the finite sample null distributions can be easily generated using

the 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 the 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 $\mu_*^{(i)} := (d/2)\max[\dot{\mathfrak{B}}_{1,*}^{(i)}, \dot{\mathfrak{E}}_{1,*}^{(i)}, \dot{\mathfrak{E}}_{2,*}^{(i)}, \ddot{\mathfrak{B}}_{1,*}^{(i)}, \ddot{\mathfrak{E}}_{1,*}^{(i)}, \ddot{\mathfrak{E}}_{2,*}^{(i)}]$, $\dot{\mathfrak{B}}_{1,*}^{(i)} := (\dot{\tau}_*^{(i)})^2 + 2\dot{\zeta}_*$, $\dot{\mathfrak{E}}_{1,*}^{(i)} := (\dot{\tau}_*^{(i)})^2 + 2\dot{\gamma}_*$, $\dot{\mathfrak{E}}_{2,*}^{(i)} := (\dot{\eta}_*^{(i)})^2 + 2\dot{\zeta}_*$, $\ddot{\mathfrak{B}}_{1,*}^{(i)} := (\ddot{\tau}_*^{(i)})^2 + 2\ddot{\zeta}_*^{(j)}$, $\ddot{\mathfrak{E}}_{1,*}^{(i)} := (\ddot{\tau}_*^{(i)})^2 + 2\ddot{\gamma}_*$, 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 CP if the power direction of their maximum test statistic differs from $\mu_*^{(i)}$. 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 $\dot{\mathfrak{S}}_{1,*}^{(i)} \leq \dot{\mathfrak{E}}_{1,*}^{(i)}$. Furthermore, $\dot{\mathfrak{S}}_{2,*}^{(i)} \leq \max[\ddot{\mathfrak{E}}_{1,*}^{(i)}, \ddot{\mathfrak{B}}_{1,*}^{(i)}]$ by the definitions of $\ddot{\mathfrak{E}}_{1,*}^{(i)}$ and $\ddot{\mathfrak{B}}_{1,*}^{(i)}$, implying that $\dot{\mu}_*^{(i)} \leq \mu_*^{(i)}$. From this feature, for each $i = 1$ and 2, the leading term of $\widehat{\mathfrak{M}}_n^{(i)}$ becomes greater than that in CP. 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/or autocorrelated errors, we recommend testing autocorrelation first and then conditional heteroskedasticity. One of the assumptions retained to test $\dot{\mathcal{H}}_0$ against $\dot{\mathcal{H}}_1$ is that $\{U_t, \mathcal{F}_t\}$ is an MDS. If $\{U_t, \mathcal{F}_t\}$ is not an MDS, it is difficult to detect how the asymptotic covariance matrix is formed by conditionally heteroskedastic and autocorrelated errors. We therefore recommend testing for the sandwich-form asymptotic covariance matrix sequentially, viz., first testing for the asymptotic covariance matrix influenced by autocorrelated errors and then testing for the asymptotic covariance matrix influenced by heteroskedastic regression errors. Specifically, we provide the following STP by applying that in Cho and Phillips (2018b):

- Step 1: Test $\ddot{\mathcal{H}}_0$ against $\ddot{\mathcal{H}}_1$ using $\widehat{\mathfrak{M}}_n^{(2)}$. If $\ddot{\mathcal{H}}_0$ is rejected, stop the STP. Otherwise, move to the next step.
- Step 2: Test $\dot{\mathcal{H}}_0$ against $\dot{\mathcal{H}}_1$ using $\widehat{\mathfrak{M}}_n^{(1)}$. If $\dot{\mathcal{H}}_0$ 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 $\ddot{\mathcal{H}}_1$, 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 $\dot{\mathcal{H}}_1$ that also belongs to $\ddot{\mathcal{H}}_0$, 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 $\dot{\mathcal{H}}_0$, the empirical researcher can detect the correct hypothesis with an asymptotic probability converging to $(1 - 2\alpha) \times 100\%$. $\dot{\mathcal{H}}_0$ can be selected only when neither $\ddot{\mathcal{H}}_0$ nor $\dot{\mathcal{H}}_0$ 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 $\dot{\mathcal{H}}_0$ against $\dot{\mathcal{H}}_1$ and $\ddot{\mathcal{H}}_0$ against $\ddot{\mathcal{H}}_1$ 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 $h_t = \kappa_* + \phi_* h_{t-1} + \gamma_* U_{t-1}^2$ and $\varepsilon_t \sim$ 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 level property of $\widehat{\mathfrak{M}}_n^{(1)}$ under $\dot{\mathcal{H}}_0$. 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 $\ddot{\mathcal{H}}_0$. 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 $\dot{\mathcal{H}}_1$. 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 power performance of $\widehat{\mathfrak{M}}_n^{(2)}$ under $\ddot{\mathcal{H}}_1$.

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 slightly mentioned 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 $\dot{\mathcal{H}}_0$ against $\dot{\mathcal{H}}_1$ by $\widehat{\mathfrak{M}}_n^{(1)}$, whereas wild bootstrap resampling is useful for testing $\ddot{\mathcal{H}}_0$ against $\ddot{\mathcal{H}}_1$ by $\widehat{\mathfrak{M}}_n^{(2)}$. The residual bootstrap method is 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

$$\widehat{\sigma}_n^2 := \frac{1}{n} \sum_{t=1}^n \widehat{U}_t^2, \quad \widehat{\mathbf{A}}_n := \frac{1}{n} \sum_{t=1}^n \mathbf{X}_t \mathbf{X}_t', \quad \text{and} \quad \widehat{\mathbf{B}}_n := \frac{1}{n} \sum_{t=1}^n \widehat{U}_t^2 \mathbf{X}_t \mathbf{X}_t'; \quad (3)$$

- 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 $\widehat{\sigma}_n^{2b} := n^{-1} \sum_{t=1}^n \widehat{U}_t^{b2}$ 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 $\dot{\mathcal{H}}_0$; otherwise, do not reject $\dot{\mathcal{H}}_0$.

On the contrary, the wild bootstrap method is implemented in the following plan to test $\ddot{\mathcal{H}}_0$ against $\ddot{\mathcal{H}}_1$:

- 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

$$\widehat{\mathbf{C}}_n := \widehat{\mathbf{B}}_n + \frac{1}{n} \sum_{k=1}^{\ell} \omega_{\ell k} \sum_{t=k+1}^n (\widehat{U}_{t-k} \widehat{U}_t \mathbf{X}_{t-k} \mathbf{X}_t' + \widehat{U}_t \widehat{U}_{t-k} \mathbf{X}_t \mathbf{X}_{t-k}'). \quad (4)$$

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

$$\ddot{\mathbf{B}}_n := \frac{1}{n} \sum_{t=1}^n W_t^{b2} \mathbf{X}_t \mathbf{X}_t' \text{ and } \ddot{\mathbf{C}}_n := \ddot{\mathbf{B}}_n + \frac{1}{n} \sum_{k=1}^{\ell} \sum_{t=k+1}^n \omega_{\ell k} (W_{t-k} W_t \mathbf{X}_{t-k} \mathbf{X}_t' + W_t W_{t-k} \mathbf{X}_t \mathbf{X}_{t-k}');$$

- 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 $\ddot{\mathcal{H}}_0$; otherwise, do not reject $\ddot{\mathcal{H}}_0$.

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 Andrews (1991). 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 methods with $B = 1,000$. Table 1 reports the simulation results.[1] Figures in parentheses are the empirical rejection rates obtained by the maximum test statistics in CP that are computed by letting $\widehat{\mathbf{A}}_n^{-1}\widehat{\mathbf{B}}_n$ and $\widehat{\mathbf{C}}_n^{-1}\widehat{\mathbf{B}}_n$ be $\widehat{\mathbf{D}}_n$ in CP to test $\dot{\mathcal{H}}_0$ and $\dot{\mathcal{H}}_1$, respectively. 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 $\dot{\mathcal{H}}_0$. 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 $\ddot{\mathcal{H}}_0$. Both DGP1 and DGP2 can be used to generate observations belonging to $\ddot{\mathcal{H}}_0$. 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 $\ddot{\mathcal{H}}_0$. 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 a 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 $\dot{\mathcal{H}}_1$. Both DGP2 and DGP3 can be used to generate observations belonging to $\dot{\mathcal{H}}_1$. 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 $\ddot{\mathcal{H}}_1$. 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 $\dot{\mathcal{H}}_1$.

(e) Finally, when comparing the empirical rejection rates of the maximum test statistics of here with those in CP, the empirical powers of our test statistics are always greater than those in CP under DGPs 2 and 3. This is mainly because our maximum test statistics are defined by including the test bases whose power directions produce

---

[1]The following URL provides the GAUSS, MATLAB, and R codes to compute the maximum test statistics and their p-values for linear models: `http://web.yonsei.ac.kr/jinseocho/cospimt.htm`.

higher powers than those in CP. If their maximum test statistics are also defined by the same test bases with the greatest power direction, the power properties of their maximum tests are expected more or less similar to ours.

These simulation results show that the theories on $\widehat{\mathfrak{M}}_n^{(1)}$ and $\widehat{\mathfrak{M}}_n^{(2)}$ are affirmed when testing $\dot{\mathcal{H}}_0$ and $\ddot{\mathcal{H}}_0$ against $\dot{\mathcal{H}}_1$ and $\ddot{\mathcal{H}}_1$, 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, $\dot{\mathcal{H}}_0$ should be selected with an asymptotic probability converging to $(1 - 2\alpha) \times 100\%$ under DGP1. On the contrary, $\ddot{\mathcal{H}}_0$ should be selected with an asymptotic probability converging to $(1 - \alpha) \times 100\%$ under DGP2. Finally, $\ddot{\mathcal{H}}_1$ 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 using the maximum test statistics of here and CP. We denote the precision rates obtained by CP in parentheses. We summarize the simulation results as follows:

(a) Under DGP1, $\dot{\mathcal{H}}_0$ is estimated with a proportion close to 90% as the sample size n increases. This aspect affirms that the STP estimates $\dot{\mathcal{H}}_0$ with an asymptotic probability converging to $(1 - 2\alpha) \times 100\%$. On the contrary, both $\ddot{\mathcal{H}}_0$ and $\ddot{\mathcal{H}}_1$ 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 $\dot{\mathcal{H}}_0$ with a high precision rate.

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

(c) Under DGP3, $\ddot{\mathcal{H}}_1$ 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 $\ddot{\mathcal{H}}_0$ with an asymptotic probability converging to unity, so that $\ddot{\mathcal{H}}_1$ can be consistently selected. When $n = 2,000$, $\ddot{\mathcal{H}}_1$ is selected with 86.10%, whereas if the sample size n further increases, $\ddot{\mathcal{H}}_1$ must be selected with a probability converging to 100%.

(d) When comparing with STP implemented by the maximum test statistics in CP, we observe that the finite sample precision rates of the STP using our maximum test statistics are closer to the hypothetical rates than their tests. This feature is well expected from the fact that our maximum tests display higher powers than theirs under DGPs 2 an 3.

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

Before moving to the next section, we note that further extension of the current research can be made to have a better powered maximum test statistic. Li, Yu, and Zeng (2018) provide a Bayesian perspective to test the equality of two matrices using the power enhancement test statistic in Fan, Liao, and Yao (2015). Although our test statistic

is not built on the Bayesian principle, it may be feasible to combine the power enhancing component to enable the maximum test statistics to have greater powers than those obtained here. For this goal, it is necessary to remedy the level distortions of the maximum test statistics other than bootstrapping, because bootstrapping can nullify the use of the power enhancement component. As this remedy is beyond scope of the current study, we leave it as our future research.

4   Empirical Applications

In this section, we apply the matrix equality testing to empirical data. We 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).

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, the researcher estimates the conditional heteroskedastic behavior based on a GARCH(1,1) model and attempts to estimate the parameters using ML 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) for the others. In addition, they adopt a GARCH(1,1) model based on ML 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 $\widehat{\mathfrak{M}}_n^{(2)}$ assisted by the residual and wild bootstraps, respectively with $B = 100,000$. 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 $\ddot{\mathcal{H}}_0 : \mathbf{B}_* = \mathbf{C}_*$ using $\widehat{\mathbf{B}}_n$ and $\widehat{\mathbf{C}}_n$ in (3) and (4), 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 3 reports the p-values estimated using the wild bootstrap method. 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 $\ddot{\mathcal{H}}_0$ cannot be rejected at the 5% level of significance, most AR models with lags greater than unity turn out not to reject $\ddot{\mathcal{H}}_0$. 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 $p = 2, 3, \ldots, 10$, AR(p) is not correctly specified for MB as implied by $\widehat{\mathfrak{M}}_n^{(2)}$. Although not reported, we increased the order of AR(p) to 15 in a further empirical analysis for MB and also found that the regression errors are still autocorrelated. 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 as the p-value rises above 5% again for $p = 10$. 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 3, 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) We also computed the maximum test statistics in CP and contained their p-values in the parentheses next to the p-values of the maximum tests. As a result, we could see that the inferential results are more or less similar to each other.

(e) Finally, we test the correct model assumption using the IM equality. White (1982) provides the IM test for the model specification of ML estimation. Given that the AR-GARCH(1,1) model specified by Bai and Lam (2019) is estimated using ML estimation assuming a skewed t-distribution for the standardized error, we apply White's

(1982) IM 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 ML 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. This test motivation is also addressed by Golden, Henley, White, and Kashner (2013, 2016) by generalizing the IM test in White (1980). We obtain the p-values of the maximum test statistic using the parametric bootstrap method and report them in Table 4, which is obtained by letting $B = 3,000$. The desired IM 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 if 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 4, the IM 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 3, implying that the IM equality is unlikely to hold for ML estimation for AFEICPS. Finally, the IM 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 3 already showed that AR-GARCH(1,1) is dynamically misspecified, implying that the model is misspecified. Given that the IM 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.

5   Concluding Remarks

The current study tests for the sandwich-form covariance matrix entailed by conditional heteroskedasticity and/or autocorrelation in the regression error. Given that conditional heteroskedastic and/or autocorrelated errors produce different sandwich forms for the covariance matrix of the LS estimator, if the precise form is known for the covariance matrix, the empirical researcher can estimate it by a pertinent estimator to the covariance matrix. Accordingly, we extend the test methodologies in CW and CP to construct maximum test statistics via the standard, HC, and HAC covariance matrix estimators of the QML estimator, which nests the LS estimator as a special case. Furthermore, we provide the STP to identify how the sandwich-form covariance matrix is affected by the conditionally heteroskedastic and/or autocorrelated regression errors. The null, alternative, and local alternative limit distributions of the maximum test statistics are provided, and we affirm the theories on the maximum test statistics through a simulation. We also apply our test principle to test the correct model hypothesis of the empirical models for many growth rates of energy prices provided by Bai and Lam (2019); we cannot find evidence that their model assumptions are incorrect for most

growth rates, although some are misspecified.

Appendix: Proofs

Proof of Theorem 1: For each $i = 1$ and 2, theorem 2 of CP 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 $\mathbf{K}_{o,n}^{(i)} := \mathbf{M}_{o,n}^{(i)} + \sum_{j=1}^{d}(\widehat{\theta}_{j,n} - \theta_{j*})\mathbf{S}_{j,*}^{(i)}$, $\mathbf{M}_{o,n}^{(i)} := (\mathbf{Q}_*^{(i)})^{-1}(\mathbf{Q}_n^{(i)}(\boldsymbol{\xi}_*) - \mathbf{P}_n^{(i)}(\boldsymbol{\xi}_*) - \mathbf{Q}_{*n}^{(i)} + \mathbf{P}_{*n}^{(i)})$, and $\mathbf{S}_{j*}^{(i)} := (\mathbf{P}_*^{(i)})^{-1}[\partial/(\partial\theta_j)\mathbf{Q}^{(i)}(\boldsymbol{\xi}_*) - \partial/(\partial\theta_j)\mathbf{P}_*^{(i)}(\boldsymbol{\xi}_*)]$. Furthermore, the symmetry between $\mathbf{P}_*^{(i)}$ and $\mathbf{Q}_*^{(i)}$ implies that for $j = 1, 2$, 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 $\widetilde{\mathbf{V}}_*^{(i)} := (\mathbf{P}_*^{(i)})^{-1}\bar{\mathbf{P}}_*^{(i)} - (\mathbf{Q}_*^{(i)})^{-1}\bar{\mathbf{Q}}_*^{(i)}$, $\widetilde{\mathbf{K}}_{o,n}^{(i)} := \widetilde{\mathbf{M}}_{o,n}^{(i)} + \sum_{j=1}^{d}(\widehat{\theta}_{j,n} - \theta_{j*})\widetilde{\mathbf{S}}_{j,*}^{(i)}$, $\widetilde{\mathbf{M}}_{o,n}^{(i)} := (\mathbf{P}_*^{(i)})^{-1}(\mathbf{P}_n^{(i)}(\boldsymbol{\xi}_*) - \mathbf{Q}_n^{(i)}(\boldsymbol{\xi}_*) - \mathbf{P}_{*n}^{(i)} + \mathbf{Q}_{*n}^{(i)})$, and $\widetilde{\mathbf{S}}_{j*}^{(i)} := (\mathbf{Q}_*^{(i)})^{-1}[\partial/(\partial\theta_j)\mathbf{P}_*^{(i)}(\boldsymbol{\xi}_*) - \partial/(\partial\theta_j)\mathbf{Q}^{(i)}(\boldsymbol{\xi}_*)]$. That is, $\mathbf{V}_*^{(i)} = -\widetilde{\mathbf{V}}_*^{(i)}$ and $\mathbf{K}_{o,n}^{(i)} = -\widetilde{\mathbf{K}}_{o,n}^{(i)}$, 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

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

In addition, Corollary 1 of CP 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 $\widehat{\mathfrak{M}}_n^{(i)} \Rightarrow (\boldsymbol{\mathcal{Z}}^{(i)} + \ddot{\mathbf{V}}_*^{(i)})^\prime\boldsymbol{\Omega}_*^{(i)}(\boldsymbol{\mathcal{Z}}^{(i)} + \ddot{\mathbf{V}}_*^{(i)})$, 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 $\dot{\mathfrak{S}}_{2,*}^{(i)} \leq \max[\dot{\mathfrak{E}}_{1,*}^{(i)}, \ddot{\mathfrak{B}}_{1,*}^{(i)}]$ and $\ddot{\mathfrak{S}}_{2,*}^{(i)} \leq \max[\ddot{\mathfrak{E}}_{1,*}^{(i)}, \dot{\mathfrak{B}}_{1,*}^{(i)}]$. 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. (2018a): "Pythagorean Generalization of Testing the Equality of Two Symmetric Positive Definite Matrices," *Journal of Econometrics*, 202, 45–56.

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

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.

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

DURBIN, 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.

FAN, J., LIAO, Y., AND YAO, J. (2015): "Power Enhancement in High-Dimensional Cross-Sectional Tests," *Econometrica*, 83, 1497–1541.

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.

GOLDEN, R. M., HENLEY, S. S., WHITE, H., AND T. M. KASHNER (2013): "New Directions in Information Matrix Testing: Eigenspectrum Tests," *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis: Festschcrify Hal White Conference*, Chen, X., Swanson, N. R., Eds. 145-178. New York: Springer.

GOLDEN, R. M., HENLEY, S. S., WHITE, H., AND T. M. KASHNER (2013): "Generalized Information Matrix Tests for Detecting Model Misspecification," *Econometrics*, 4, 46.

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.

LI, Y., YU, J., AND ZENG, T. (2018): "Specification Tests Bases on MCMC Output," *Journal of Econometrics*, 207, 237–260.

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.

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.

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.

Table 1: Empirical rejection rates of the maximum test statistics (in percent)

DGP1: $\dot{\mathcal{H}}_0$ or $\ddot{\mathcal{H}}_0$							
test statistics	levels\n	200	400	600	800	1,000	1,200
$\widehat{\mathfrak{M}}_n^{(1)}$	1%	0.93 (0.83)	0.90 (1.00)	0.70 (1.30)	1.17 (1.13)	0.77 (1.23)	1.03 (1.03)
	5%	4.70 (4.33)	4.63 (4.97)	4.77 (5.57)	4.47 (5.40)	4.30 (5.20)	4.80 (4.57)
	10%	9.23 (9.30)	8.83 (9.63)	9.73 (10.63)	9.53 (10.17)	9.23 (9.13)	9.97 (9.57)
$\widehat{\mathfrak{M}}_n^{(2)}$	1%	0.27 (0.10)	0.60 (0.17)	0.87 (0.27)	0.67 (0.43)	0.60 (0.33)	1.23 (0.73)
	5%	2.83 (0.93)	4.07 (2.23)	3.97 (3.07)	3.83 (3.33)	4.10 (3.90)	6.00 (4.60)
	10%	6.10 (4.03)	8.83 (8.03)	8.60 (7.83)	8.37 (7.13)	7.63 (8.47)	11.40 (9.70)
DGP2: $\dot{\mathcal{H}}_1$ or $\ddot{\mathcal{H}}_0$							
test statistics	levels\n	200	400	600	800	1,000	1,500
$\widehat{\mathfrak{M}}_n^{(1)}$	1%	13.00 (7.20)	35.93 (27.43)	57.00 (50.43)	73.97 (67.37)	84.93 (80.13)	96.83 (96.00)
	5%	28.63 (21.20)	56.33 (48.90)	76.13 (71.77)	87.03 (85.20)	94.43 (92.27)	98.97 (99.13)
	10%	39.33 (33.13)	66.30 (61.07)	83.10 (80.80)	91.93 (91.27)	96.70 (95.73)	99.43 (99.60)
$\widehat{\mathfrak{M}}_n^{(2)}$	1%	0.33 (0.00)	0.57 (0.17)	0.97 (0.10)	0.87 (0.27)	0.67 (0.30)	1.20 (0.50)
	5%	3.20 (1.00)	5.20 (2.13)	5.03 (3.23)	5.23 (3.57)	5.00 (4.40)	5.63 (4.83)
	10%	8.07 (5.10)	11.37 (7.87)	11.03 (8.67)	11.20 (10.40)	11.57 (10.37)	11.73 (11.37)
DGP3: $\dot{\mathcal{H}}_1$ or $\ddot{\mathcal{H}}_1$							
test statistics	levels\n	200	400	600	800	1,000	2,000
$\widehat{\mathfrak{M}}_n^{(1)}$	1%	14.70 (7.27)	38.90 (28.43)	58.87 (50.73)	75.70 (70.23)	84.80 (82.03)	99.33 (99.17)
	5%	29.10 (22.57)	59.27 (52.10)	77.57 (72.97)	88.87 (86.60)	93.93 (93.00)	99.87 (99.90)
	10%	39.80 (33.13)	68.10 (64.70)	84.13 (81.57)	93.33 (91.70)	96.33 (95.87)	100.0 (100.0)
$\widehat{\mathfrak{M}}_n^{(2)}$	1%	5.17 (0.53)	18.70 (2.57)	24.53 (7.40)	30.07 (13.20)	37.27 (18.13)	67.73 (45.63)
	5%	18.13 (10.07)	40.27 (22.03)	48.63 (31.93)	54.07 (41.87)	60.87 (49.13)	86.10 (76.90)
	10%	28.83 (23.20)	51.97 (42.43)	59.17 (49.00)	65.53 (56.87)	70.27 (65.03)	91.43 (87.30)

Table 2: Empirical precision rates using the STP (in percent)

	DGP1: $\dot{\mathcal{H}}_0$ or $\ddot{\mathcal{H}}_0$					
hypotheses $\backslash n$	200	400	600	800	1,000	1,200
$\dot{\mathcal{H}}_0$	92.57 (94.80)	91.40 (92.90)	91.40 (91.60)	91.93 (91.43)	91.80 (91.10)	89.57 (91.13)
$\ddot{\mathcal{H}}_0$	4.60 (4.27)	4.53 (4.87)	4.63 (5.33)	4.23 (5.23)	4.10 (5.00)	4.43 (4.27)
$\ddot{\mathcal{H}}_1$	2.83 (0.93)	4.07 (2.23)	3.97 (3.07)	3.83 (3.33)	4.10 (3.90)	6.00 (4.60)
Sum	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)

	DGP2: $\dot{\mathcal{H}}_1$ or $\ddot{\mathcal{H}}_0$					
hypotheses $\backslash n$	200	400	600	800	1,000	1,500
$\dot{\mathcal{H}}_0$	69.17 (77.90)	41.67 (49.83)	23.00 (27.20)	12.37 (14.33)	5.30 (7.37)	0.97 (0.83)
$\ddot{\mathcal{H}}_0$	27.63 (21.10)	53.13 (48.03)	71.97 (69.57)	82.40 (82.10)	89.70 (88.23)	93.40 (94.33)
$\ddot{\mathcal{H}}_1$	3.20 (1.00)	5.20 (2.13)	5.03 (3.23)	5.23 (3.57)	5.00 (4.40)	5.63 (4.83)
Sum	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)

	DGP3: $\dot{\mathcal{H}}_1$ or $\ddot{\mathcal{H}}_1$					
hypotheses $\backslash n$	200	400	600	800	1,000	2,000
$\dot{\mathcal{H}}_0$	57.80 (68.97)	24.87 (35.53)	10.60 (17.00)	4.83 (6.67)	1.87 (3.10)	0.00 (0.00)
$\ddot{\mathcal{H}}_0$	24.07 (20.97)	34.87 (42.43)	40.77 (51.07)	41.10 (51.47)	37.27 (42.77)	13.90 (23.10)
$\ddot{\mathcal{H}}_1$	18.13 (10.07)	40.27 (22.03)	48.63 (31.93)	54.07 (41.87)	60.87 (49.13)	86.10 (76.90)
Sum	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)

Table 3: p-values of testing for the influence of conditional heteroskedasticity and autocorrelation (in percent)

	BLPG		BRENT		MB		PAFEI	
lags	$\dot{\mathcal{H}}_0$ vs $\dot{\mathcal{H}}_1$	$\ddot{\mathcal{H}}_0$ vs $\ddot{\mathcal{H}}_1$	$\dot{\mathcal{H}}_0$ vs $\dot{\mathcal{H}}_1$	$\ddot{\mathcal{H}}_0$ vs $\ddot{\mathcal{H}}_1$	$\dot{\mathcal{H}}_0$ vs $\dot{\mathcal{H}}_1$	$\ddot{\mathcal{H}}_0$ vs $\ddot{\mathcal{H}}_1$	$\dot{\mathcal{H}}_0$ vs $\dot{\mathcal{H}}_1$	$\ddot{\mathcal{H}}_0$ vs $\ddot{\mathcal{H}}_1$
1	0.00 (0.19)	26.67 (30.55)	0.03 (0.32)	96.67 (96.59)	0.01 (0.42)	6.31 (11.14)	0.86 (2.34)	68.25 (66.08)
2	0.01 (0.38)	81.82 (83.09)	0.11 (0.54)	40.23 (48.04)	0.04 (1.40)	4.49 (10.61)	0.59 (2.83)	80.64 (85.51)
3	0.00 (0.30)	86.05 (89.05)	0.10 (0.89)	30.20 (53.27)	0.07 (2.09)	4.18 (8.62)	0.09 (2.42)	91.11 (92.59)
4	0.01 (0.96)	94.97 (95.72)	0.02 (1.03)	6.06 (14.08)	0.13 (3.20)	2.51 (3.57)	0.02 (1.72)	98.84 (99.15)
5	0.03 (2.02)	96.62 (97.29)	0.05 (1.76)	11.88 (23.73)	0.01 (1.80)	1.85 (3.15)	0.10 (3.00)	96.76 (98.48)
6	0.04 (1.99)	99.55 (99.70)	0.02 (2.07)	12.89 (27.93)	0.04 (2.74)	4.15 (6.75)	0.10 (4.10)	94.15 (97.08)
7	0.09 (2.71)	80.45 (85.07)	0.01 (1.99)	10.60 (20.02)	0.03 (3.32)	3.26 (5.78)	0.05 (3.58)	71.65 (87.16)
8	0.38 (5.72)	43.99 (47.65)	0.01 (3.09)	12.80 (21.59)	0.04 (3.21)	2.37 (5.73)	0.01 (2.52)	80.64 (93.64)
9	1.13 (9.78)	43.57 (51.13)	0.01 (3.32)	17.04 (25.86)	0.08 (5.19)	2.88 (6.45)	0.03 (4.91)	71.58 (85.69)
10	1.49 (9.12)	44.59 (53.04)	0.03 (4.30)	16.40 (24.28)	0.18 (7.54)	3.95 (7.04)	0.06 (6.13)	66.72 (80.25)

	CPS		AFEIUS		AFEICPS			
lags	$\dot{\mathcal{H}}_0$ vs $\dot{\mathcal{H}}_1$	$\ddot{\mathcal{H}}_0$ vs $\ddot{\mathcal{H}}_1$	$\dot{\mathcal{H}}_0$ vs $\dot{\mathcal{H}}_1$	$\ddot{\mathcal{H}}_0$ vs $\ddot{\mathcal{H}}_1$	$\dot{\mathcal{H}}_0$ vs $\dot{\mathcal{H}}_1$	$\ddot{\mathcal{H}}_0$ vs $\ddot{\mathcal{H}}_1$		
1	1.48 (3.17)	57.92 (63.10)	2.50 (3.75)	6.51 (12.71)	0.03 (3.49)	0.15 (1.13)		
2	1.41 (4.56)	10.47 (26.05)	4.06 (8.17)	15.34 (11.39)	0.06 (10.69)	1.06 (0.83)		
3	1.47 (6.29)	23.08 (52.21)	2.91 (10.14)	15.47 (7.95)	0.06 (14.17)	5.01 (4.19)		
4	2.79 (9.53)	29.61 (45.07)	3.32 (18.41)	14.77 (8.82)	0.22 (24.46)	9.37 (10.22)		
5	4.64 (15.00)	16.01 (30.49)	4.38 (21.28)	13.72 (8.37)	0.28 (29.46)	8.91 (10.95)		
6	4.59 (18.38)	31.65 (46.15)	8.55 (37.94)	13.65 (10.79)	0.37 (36.90)	6.09 (8.72)		
7	1.74 (14.00)	21.51 (39.63)	12.34 (44.86)	13.24 (11.74)	0.40 (43.58)	8.92 (13.42)		
8	0.30 (9.74)	30.13 (45.37)	21.39 (62.33)	27.07 (21.71)	0.81 (51.53)	14.74 (22.99)		
9	0.26 (11.15)	39.55 (55.95)	29.71 (76.30)	43.42 (34.91)	1.11 (54.03)	4.19 (6.53)		
10	0.54 (11.50)	42.43 (55.26)	41.11 (76.14)	39.27 (20.86)	1.30 (56.68)	8.90 (15.06)		

Table 4: p-value of the information matrix equality test (in percent)

	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