Pythagorean Generalization of Testing the Equality of Two Symmetric Positive Definite Matrices

JIN SEO CHO
School of Economics
Yonsei University
50 Yonsei-ro, , Seodaemun-gu, Seoul, Korea
Email: jinseocho@yonsei.ac.kr

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

November 9, 2015

Abstract

We provide a new test for equality of two symmetric positive-definite matrices that leads to a convenient mechanism for testing specification using the information matrix equality and the sandwich asymptotic covariance matrix of the GMM estimator. The test relies on a new characterization of equality between two k dimensional symmetric positive-definite matrices A and B: the traces of AB^{-1} and BA^{-1} are equal to k if and only if $A = B$. Using this criterion, we introduce a class of omnibus test statistics for the equality and examine their null and local alternative approximations under some mild regularity conditions. Monte Carlo experiments are conducted to explore the performance characteristics of the test criteria and provide comparisons with existing tests under the null, local, and alternative. The tests are applied to the classic empirical models for voting turnout investigated by Wolfinger and Rosenstone (1980) and Nagler (1991, 1994). Our tests show that all classic models for the 1984 presidential voting turnout are misspecified in the sense that the information matrix equality fails.

Key Words: Matrix equality; Trace; Determinant; Arithmetic mean; Geometric mean; Harmonic mean; Information matrix; Sandwich covariance matrix; Eigenvalues; Bootstrap.

Subject Classification: C01, C12, C52, D72

Acknowledgements: The co-editor, Oliver Linton, Associated editor, and two anonymous referees provided helpful comments on the original version for which we are grateful. The authors acknowledge helpful discussions with Heejoon Han, Dukpa Kim, Eunyoung Kim, Joocheol Kim, Tae-Hwan Kim, Jin Lee, Nakyeong Lee, Tae Hwy Lee, Yoon-Jae Whang, Byung Sam Yoo, Jungmo Yoon, and other participants at the KEA conference (Korea University) and the seminar and conference participants at Yonsei University, the Korean Econometric Society, and the Western Economic Association International. Cho thanks the research grant support by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2015S1A5A2A01012140), and Phillips acknowledges research support from the NSF under grant No. SES 12-58258.

1 Introduction

Comparing covariance matrices and testing the equivalence of two symmetric positive-definite matrices have attracted substantial past attention in both multivariate analysis and econometrics. For example, the asymptotic distribution of the maximum likelihood (ML) estimator is characterized by the usual information matrix equality. On the other hand, the information matrix equality does not hold for the quasi-ML (QML) estimator. As another example, least squares (LS) and generalized method of moments (GMM) estimators have relatively simple covariance matrix structures except when heteroskedasticity, model misspecification, or autocorrelation is present. The simple covariance matrix structure is then delivered by the proportional equality of two symmetric positive-definite matrices (viz., $X'X$ and $X'\Sigma X$ in the usual regression notation).

These material econometric concerns have led to much literature on covariance matrix equality testing, with special attention being given to the information matrix equality (*e.g.*, White, 1982; 1994; Bera, 1986; Hall, 1987; Orme, 1988, 1990; Chesher and Spady, 1991; Horowitz 1994; Dhaene and Hoorelbeke, 2004; and Golden, Henley, White, and Kashner, 2013), although work is not limited to that setting alone (*e.g.*, Bera and Hall, 1991). Much of this past work arises from the desire for an omnibus test without level distortion and with high power. The problem in size control is simply stated. For two general $k \times k$ symmetric positive-definite matrices A and B say, testing *every* pair of corresponding elements in A and B generates enormous level distortions for the tests even with moderately sized k.

The goal of the current study is to develop simple and straightforward omnibus tests for the equality of two symmetric positive-definite matrices and to broaden their implications for applied research. The approach that we use here has an antecedent in Cho and White (2014; CW, henceforth). CW provided omnibus tests of matrix equality by using the fact that the conditions $\mathrm{tr}[BA^{-1}] = k$ and $\det[BA^{-1}] = 1$ are necessary and sufficient for $A = B$. Our starting point is to replace this condition with another, even simpler, characterization of equality that enables a new class of omnibus tests for equality that have little size distortion and comparable powers to other tests. The tests given in CW then become special cases of our approach. We also seek to clarify the interrelationships among the many tests that are now available and those that are developed in the current study. The paper therefore contributes by (i) introducing a class of easily implemented new tests that have good size and power properties, and (ii) providing a comprehensive study of the properties and performance characteristics of an extensive range of tests for covariance matrix equality.

These goals of the current study are achieved by extending the CW approach in a number of ways. First, we examine a number of omnibus test statistics for the equality of two symmetric positive-definite matrices by the Pythagorean characterization. We show that the simple dual conditions $\mathrm{tr}[AB^{-1}] = k$ and $\mathrm{tr}[BA^{-1}] = k$ are also necessary and sufficient for $A = B$. This characterization is made by noting that $k^{-1}\mathrm{tr}[BA^{-1}]$ and $k\mathrm{tr}[AB^{-1}]^{-1}$ are the arithmetic and harmonic means of the eigenvalues of BA^{-1}, respectively, and that these means are equal if and only if all eigenvalues are identical. Under the given conditions, all eigenvalues are unity, implying that $BA^{-1} = I$. Note that this characterization is additional to that given in CW, viz. $\mathrm{tr}[AB^{-1}] = k$ and $\det[BA^{-1}] = 1$ if and only if $A = B$, and a number of new testing factors can be obtained from this new characterization, that are also additional to those in CW. Further, an even wider range of test factors can be obtained by pairing the arithmetic and geometric means or combining all of the Pythagorean means. In addition, more tests are obtained by reversing the roles of A and B in the the relations $k^{-1}\mathrm{tr}[AB^{-1}]$, $\det[AB^{-1}]^{1/k}$, and $k\mathrm{tr}[BA^{-1}]^{-1}$. All these test factors form a class, and we can apply continuous distance functions to the class to yield omnibus test statistics. Within this general framework, the tests given in CW become special cases of those developed here.

Second, we examine the performance characteristics of the various tests under the null, local, and fixed alternative hypotheses, and derive their asymptotic approximations. This examination broadens the analysis commenced in CW as it transpires that the null and alternative approximations given in CW continue to apply for our test statistics, although the local alternative is generalized to include alternatives in other directions. For the alternative, we single out the factors leading to test consistency and analyze the power relationships of the tests. When the alternative hypothesis is partitioned into a number of explicit regions, the tests developed here are designed to test the dominant characteristic of each region.

Third, the practical applicability of our test statistics is extended to a wide range of estimation methodologies. CW specifically focus on the application of their tests in the context of QML estimation, and for practical implementation they use Horowitz's (1994) parametric bootstrap to test information matrix equality. Our approach achieves wider applicability through implementation in the GMM estimation context and by using the residual bootstrap. Specifically, we show that our test statistics are useful in practical work for testing the optimal weight matrix condition in GMM estimation. Extensive Monte Carlo simulations are conducted, and we evaluate the performance of our tests within both QML and GMM estimation contexts so that the most practically useful tests can be identified. The results show that test statistics constructed by Pythagorean means of the eigenvalues perform better overall than the other tests. The simulations also assist in confirming the relevance of the asymptotic theory and asymptotic comparisons in finite samples.

Finally, we apply our tests to empirical data on voting turnout. In the political economy and political science literature, an important research question involves identifying factors that determine presidential voting turnouts (*e.g.*, Wolfinger and Rosenstone, 1980; Feddersen and Pesendorfer, 1996; Nagler, 1991, 1994; Bénabou 2000; Besley and Case, 2003; Berry, DeMeritt, and Esarey, 2010, among many others). In particular, Wolfinger and Rosenstone's (1980) classic study examines the interaction effects between education and registry requirement to the voting turnout by estimating a probit model under ML. Nagler (1994) further extends their results by estimating a scobit model. We reexamine their models and empirically test whether their models are correctly specified for ML estimation. For this purpose, we use the 1984 presidential election data of the US that are provided by Altman and McDonald (2003).

The plan of this paper is as follows. Section 2 provides a fundamental result characterizing equality between two symmetric positive-definite matrices. Section 3 motivates and defines the test statistics employed, and develops asymptotic theory under the null and alternative hypotheses. Simulation results are reported in Section 4. We focus on linear normal and linear probit regression models and test the information matrix equality in these two frameworks. We also examine LS and two-stage least squares (TSLS) estimations and test their asymptotic covariance matrices. The empirical application is provided in Section 5. Mathematical proofs are collected in the Appendix.

Before proceeding, we provide some notation. A function mapping $f : \mathcal{X} \mapsto \mathcal{Y}$ is denoted by $f(\cdot)$, evaluated derivatives such as $f'(x)|_{x=x_*}$ are written simply as $f'(x_*)$, and $\partial_x f(x) := (\partial/\partial x)f(x)$, $\partial^2_{x,y} f(x,y)$ $:= (\partial^2/\partial x \partial y) f(x,y)$.

2 A Basic Lemma and Its Testing Implications

Our starting point is the following fundamental lemma that characterizes the equality of two symmetric positive-definite matrices.

Lemma 1. *Let A and B be real symmetric positive-definite $k \times k$ matrices with $k \in \mathbb{N}$. Then, $A = B$ if and only if*
(i) $\mathrm{tr}[D] = k$ *and* $\mathrm{tr}[D^{-1}] = k$, *where* $D := BA^{-1}$; *or*
(ii) $\det[D] = 1$ *and* $\mathrm{tr}[D^{-1}] = k$. □

To our knowledge and somewhat surprisingly given its simplicity, Lemma 1 is new to the literature and is proved in the Appendix. Briefly, part (*i*) follows because the arithmetic mean of positive numbers is identical to their harmonic mean, if and only if all of the positive numbers are identical. Since $k^{-1}\mathrm{tr}[D]$ is the arithmetic mean of the eigenvalues of D, and $k^{-1}\mathrm{tr}[D^{-1}]$ is the inverse of the harmonic mean of the same eigenvalues, we have $D = I$, if and only if all the eigenvalues are identical to unity, which implies that $A = B$. Notably, and most conveniently for practical work, the criteria in (i) and (ii) involve only the leading elementary symmetric functions of the matrices D and D^{-1}.

As pointed out by a reviewer, the characterization in Lemma 1(*i*) can be generalized by associating the eigenvalues of D with a strictly Schur-convex function of them (*e.g.*, Marshall, Olkin, and Arnold, 1979). Thus, if $f(\cdot)$ and λ are a Schur-convex function and the vector of the eigenvalues of D, respectively, we have the equivalence $D = I$ if and only if $\mathrm{tr}[D] = k$ and $f(\lambda) = f(\iota)$, where ι is the vector of ones. Note that $\mathrm{tr}[D^{-1}] = \sum_{i=1}^{k} \lambda_i^{-1}$, and $1/x$ is a Schur-convex function, so that $\mathrm{tr}[D^{-1}]$ is a convex function of the eigenvalues. This proves Lemma 1(*i*). In addition, we can apply D^{-1} to the criterion instead of D: if we let ρ be the vector of the eigenvalues of D^{-1}, $\mathrm{tr}[D^{-1}] = k$ and $\det[D] = \exp(-\sum_{i=1}^{k} \log(\rho_i))$. Here, we note that $-\log(x)$ is a Schur-convex function, so that $\det[D]$ is a strictly Schur-convex function of ρ. Therefore, $D^{-1} = I$ if and only if $\mathrm{tr}[D^{-1}] = k$ and $\det[D] = 1$, proving Lemma 1(*ii*). In another way, the characterization in Lemma 1(*i*) can also be associated with a convexity property of the trace operator. Note that $\phi(\cdot) := \mathrm{tr}[(\cdot)^{-1}] + \mathrm{tr}[\cdot]$ is a convex function on the space of $k \times k$ symmetric positive-definite matrices (*e.g.* , Bernstein, 2005, p. 283) and is also bounded from below by $2k$ (*e.g.*, Abadir and Magnus, 2005, p.338). The lower bound is achieved if and only if the argument of $\phi(\cdot)$ is I.

The characterization in Lemma 1 is different from that used in CW, in which the equality of two equal symmetric positive-definite matrices is characterized by both $\det[D]$ and $\mathrm{tr}[D]$. Note that $\det[D]^{1/k}$ is the geometric mean of the eigenvalues of D. Furthermore, the geometric mean of positive numbers is identical to the arithmetic mean, if and only if the positive numbers are identical. Using this simple fact, CW characterized two equal symmetric positive-definite matrices by the condition that $\det[D] = 1$ and $\mathrm{tr}[D] = k$. Lemma 1(*ii*) is then a corollary of Lemma 1(*i*) and the CW characterization.

Both Lemma 1 and the characterization in CW rely on fundamental properties of the Pythagorean (harmonic, geometric, and arithmetic) means of positive numbers:

$$\text{Harmonic Mean} \;\leq\; \text{Geometric mean} \;\leq\; \text{Arithmetic mean.} \tag{1}$$

All three means are identical if the positive numbers are identical. Lemma 1(*i*) is obtained by interrelating the harmonic mean with the arithmetic mean, and CW links the geometric mean to the arithmetic mean for their characterization. Lemma 1(*ii*) also associates the harmonic and geometric means for the equality.

There are solid grounds to use the trace and determinant-based test statistics for the equality of two symmetric positive-definite matrices. First, as mentioned above, these invariant polynomials are the leading elementary symmetric functions of the positive semi-definite matrices, and are simple and straightforward for practical implementation. Our Monte Carlo simulations also show that the test statistics defined below exhibit quality finite sample performance relative to other test statistics. Second, the theory of model selection information criteria has been developed by replacing Akaike's penalty term with the trace and/or determinant of the asymptotic covariance of an estimator (*e.g.*, Takeuchi, 1976; Bozdogan 2000), particularly when models are possibly misspecified. This motivates testing the equal covariance matrix hypothesis using the trace and determinant. Third, if the eigenvalues are explicitly involved with the test statistics (as distinct from implicitly via the elementary symmetric functions), it is challenging to obtain the null limit approximations of the test statistics. If we let $T_n := f(\widehat{\lambda}_n)$ for testing $D = I$, where $\widehat{\lambda}_n = \lambda(\widehat{D}_n)$ and $\widehat{D}_n$ is a consistent estimator for D, it is necessary to approximate T_n using the differential of $\lambda(\cdot)$ around I, which does not exist, *e.g.* Magnus (1985). Under the null $D = I$, $\lambda(D)$ is not simple, making it challenging to obtain the null limit approximation of T_n particularly when $\widehat{D}_n$ involves parameter estimation. James (1964) and Onatski, Moreira, and Hallin (2013) provide distributional properties of $\widehat{\lambda}_n$ for normally associated samples, although not for a general case that involves parameter estimation. Hence, testing equality of two symmetric positive definite matrices without explicitly involving the eigenvalues leads directly to the use of the elementary symmetric functions, thereby motivating the choice of $\mathrm{tr}[\widehat{D}_n]$ and $\det[\widehat{D}_n]$ as vehicles for testing $D = I$. Finally, as will be demonstrated below, our proposed tests are asymptotically equivalent to the likelihood ratio test statistic under a local alternative in a prototypical structural vector autoregressive model context, implying that these statistics are locally optimal and can therefore be expected to have good power properties.

We now exploit Lemma 1 to test the equality of two symmetric positive-definite matrices. Lemma 1(i) is our first focus. Let $\tau := k^{-1}\mathrm{tr}[D] - 1$, $\eta := k\mathrm{tr}[D^{-1}]^{-1} - 1$, and $\xi := k^{-1}\mathrm{tr}[D] - k\mathrm{tr}[D^{-1}]^{-1}$ for notational simplicity. Note that if any two of τ, η, and ξ equal to zero, the remaining one is also zero. Therefore, Lemma 1(i) holds if and only if any two of τ, η, and ξ equal to zero. This implies that the equality of two symmetric positive-definite matrices can be tested by testing one of the following hypotheses:

$$\begin{aligned}
&\mathcal{H}_0^{(1)}: \quad \tau = 0 \ \text{ and } \ \eta = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(1)}: \tau \neq 0 \ \text{ or } \ \eta \neq 0;\\
&\mathcal{H}_0^{(2)}: \quad \tau = 0 \ \text{ and } \ \xi = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(2)}: \tau \neq 0 \ \text{ or } \ \xi \neq 0;\\
&\mathcal{H}_0^{(3)}: \quad \eta = 0 \ \text{ and } \ \xi = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(3)}: \eta \neq 0 \ \text{ or } \ \xi \neq 0.
\end{aligned}$$

Similarly, we can exploit Lemma 1(ii) and for this, let $\delta := \det[D]^{1/k} - 1$ and $\gamma := \det[D]^{1/k} - k\mathrm{tr}[D^{-1}]^{-1}$. If any two of δ, η, and γ are zero, the remaining one is zero, so that Lemma 1(ii) holds if and only if any two of them are zero. Hence, we construct the corresponding hypotheses as

$$\begin{aligned}
\mathcal{H}_0^{(4)}: &\quad \delta = 0 \ \text{ and } \ \eta = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(4)}: \delta \neq 0 \ \text{ or } \ \eta \neq 0; \\
\mathcal{H}_0^{(5)}: &\quad \delta = 0 \ \text{ and } \ \gamma = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(5)}: \delta \neq 0 \ \text{ or } \ \gamma \neq 0; \\
\mathcal{H}_0^{(6)}: &\quad \eta = 0 \ \text{ and } \ \gamma = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(6)}: \eta \neq 0 \ \text{ or } \ \gamma \neq 0.
\end{aligned}$$

These hypotheses correspond to those considered in CW. They let $\sigma := k^{-1}\mathrm{tr}[D] - \det[D]^{1/k}$ and test whether any two of τ, δ, and σ are zero by considering the following hypotheses:

$$\begin{aligned}
\mathcal{H}_0^{(7)}: &\quad \tau = 0 \ \text{ and } \ \delta = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(7)}: \tau \neq 0 \ \text{ or } \ \delta \neq 0; \\
\mathcal{H}_0^{(8)}: &\quad \tau = 0 \ \text{ and } \ \sigma = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(8)}: \tau \neq 0 \ \text{ or } \ \sigma \neq 0; \\
\mathcal{H}_0^{(9)}: &\quad \delta = 0 \ \text{ and } \ \sigma = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(9)}: \delta \neq 0 \ \text{ or } \ \sigma \neq 0.
\end{aligned}$$

All these 9 hypothesis systems are equivalent systems of hypotheses to the simple null $\mathcal{H}_0 : A = B$ versus the alternative $\mathcal{H}_1 : A \neq B$.

Several remarks are warranted in terms of this testing methodology. First, note that the testing factors τ, δ, and η are invariant to linear transformations of the null $\mathcal{H}_0 : A = B$. Thus, for any invertible matrix H, if $\tau^\dagger$, $\delta^\dagger$, and $\eta^\dagger$ are computed using AH and BH, it easily follows that $\tau^\dagger = \tau$, $\delta^\dagger = \delta$, and $\eta^\dagger = \eta$ because $BH(AH)^{-1} = D$. Therefore, τ, δ, and η are invariant to linear transformations. The other factors ξ, γ, and σ share the same property. Second, as one of the reviewers pointed out, the roles of A and B can be reversed when computing the testing factors τ, δ and η. Furthermore, the prior hypotheses can be extended to involve τ, δ, and η at the same time, yielding

$$\begin{aligned}
\mathcal{H}_0^{(10)}: &\quad \tau = 0 \ \text{ and } \ \delta - \eta = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(10)}: \tau \neq 0 \ \text{ or } \ \delta - \eta \neq 0; \\
\mathcal{H}_0^{(11)}: &\quad \delta = 0 \ \text{ and } \ \tau - \eta = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(11)}: \delta \neq 0 \ \text{ or } \ \tau - \eta \neq 0; \\
\mathcal{H}_0^{(12)}: &\quad \eta = 0 \ \text{ and } \ \tau - \delta = 0 \quad \text{vs.} \quad \mathcal{H}_1^{(12)}: \eta \neq 0 \ \text{ or } \ \tau - \delta \neq 0.
\end{aligned}$$

The original null hypothesis $\mathcal{H}_0$ can also be written as a consequence of these hypotheses. We show below that test statistics that are obtained by reversing the roles of A and B or these additional hypotheses are asymptotically equivalent to the statistics that test the above hypotheses under the null and local alternative hypotheses. This feature implies that the testing factors τ, δ, and η also play fundamental roles when constructing omnibus test statistics for $A = B$.

We introduce testing environments by supposing that the previously defined matrices A and B are in fact parameterized as $A \equiv A(\theta_*)$ and $B \equiv B(\theta_*)$, respectively, where both $A(\cdot)$ and $B(\cdot)$ are defined on $\Theta \in \mathbb{R}^{\ell}$, and $\theta_* \in \Theta$ is an unknown parameter. This dependence is motivated by the fact that most covariance estimators are obtained as second-stage outputs after estimating the unknown parameter. For example, the asymptotic covariance matrix of the (Q)ML estimator is the sandwich covariance matrix that is often consistently obtained through the vehicle of heteroskedasticity autocorrelation consistent estimation:

$$\sqrt{n}(\widehat{\theta}_n - \theta_*) \overset{\text{A}}{\sim} N(0, A(\theta_*)^{-1}B(\theta_*)A(\theta_*)^{-1}).$$

If $A(\theta_*) = B(\theta_*)$, the asymptotic covariance of $\widehat{\theta}_n$ becomes identical to that of ML estimator, so that testing $A(\theta_*) = B(\theta_*)$ can be associated with model specification testing. When proceeding with this association, we further suppose that $A_n := A_n(\theta_*)$ and $B_n := B_n(\theta_*)$ estimate $A(\theta_*)$ and $B(\theta_*)$ consistently, where $A_n(\cdot)$ and $B_n(\cdot)$ are consistent for $A(\cdot)$ and $B(\cdot)$ uniformly on Θ and are uniformly positive definite almost surely on Θ for large enough n. That is, for any $x \in \mathbb{R}^k \setminus \{0\}$, $\mathbb{P}(\lim_{n\to\infty} \inf_{\theta\in\Theta} x'A_n(\theta)x > 0) = 1$ and $\mathbb{P}(\lim_{n\to\infty} \inf_{\theta\in\Theta} x'B_n(\theta)x > 0) = 1$. Therefore, $D_n := B_nA_n^{-1}$ and D_n^{-1} consistently estimate D and D^{-1}, respectively. Here, D is estimated by a two-step estimation procedure. Specifically, the unknown parameter θ_* is consistently estimated by an estimator $\widehat{\theta}_n$, so that $\widehat{A}_n := A_n(\widehat{\theta}_n)$ and $\widehat{B}_n := B_n(\widehat{\theta}_n)$ are consistent for $A(\theta_*)$ and $B(\theta_*)$, respectively. Therefore, $\widehat{D}_n := \widehat{B}_n\widehat{A}_n^{-1}$ and $\widehat{D}_n^{-1} = \widehat{A}_n\widehat{B}_n^{-1}$ are also consistent for D and D^{-1}, respectively. To reduce notational clutter, we simply indicate the influence of θ_* on these matrices by letting

$$A_* := A(\theta_*), \quad B_* := B(\theta_*), \quad D_* := B_*A_*^{-1}.$$

Similarly, let $\tau_* := k^{-1}\mathrm{tr}[D_*] - 1$, $\eta_* := k/\mathrm{tr}[D_*^{-1}] - 1$, $\delta_* := \det[D_*]^{1/k} - 1$, $\xi_* := \tau_* - \eta_*$, $\gamma_* := \delta_* - \eta_*$, and $\sigma_* := \tau_* - \delta_*$. When these matrices are estimated using $\widehat{A}_n$ and $\widehat{B}_n$, we denote the resulting statistics as $\widehat{\tau}_n := k^{-1}\mathrm{tr}[\widehat{D}_n] - 1$, $\widehat{\eta}_n := k/\mathrm{tr}[\widehat{D}_n^{-1}] - 1$, $\widehat{\delta}_n := \det[\widehat{D}_n]^{1/k} - 1$, $\widehat{\xi}_n := \widehat{\tau}_n - \widehat{\eta}_n$, $\widehat{\gamma}_n := \widehat{\delta}_n - \widehat{\eta}_n$, and $\widehat{\sigma}_n := \widehat{\tau}_n - \widehat{\delta}_n$. All these statistics, which form the base elements of the tests given below, are dependent upon $\widehat{\theta}_n$. For notational simplicity, we also let $\widetilde{D}_* := A_*B_*^{-1}$ and $\widetilde{D}_n := \widehat{A}_n\widehat{B}_n^{-1}$. Therefore, $\widetilde{D}_* = D_*^{-1}$ and $\widetilde{D}_n := \widehat{D}_n^{-1}$.

3 Test Statistics and Their Asymptotic Expansions

This section introduces the test statistics and examines their asymptotic expansions under the null, local, and fixed alternative hypotheses. We also supplement the test statistics considered in CW.

3.1 Definitions of Test Statistics and Asymptotic Approximations

Before defining the tests and examining their asymptotic approximations, we formally provide the following regularity conditions, some of which have already been mentioned.

Assumption A (Cho and White, 2014)**.** *(i)* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a complete probability space;*
(ii) $\Theta \subset \mathbb{R}^\ell$ *is a compact convex set with non-empty interior and* $\ell \in \mathbb{N}$;
(iii) a sequence of measurable mappings $\{\widehat{\theta}_n : \Omega \mapsto \Theta\}$ *is consistent for a unique* $\theta_* \in \mathrm{int}(\Theta)$*;*
(iv) $A : \Theta \mapsto \mathbb{R}^{k\times k}$ *and* $B : \Theta \mapsto \mathbb{R}^{k\times k}$ *are in* $\mathcal{C}^{(2)}(\Theta)$*, and* A_* *and* B_* *are positive definite;*
(v) $A_n(\cdot)$ *and* $B_n(\cdot)$ *are consistent for* $A(\cdot)$ *and* $B(\cdot)$*, respectively, uniformly on* Θ*, viz.,*

$$\sup_{\theta\in\Theta} \|A_n(\theta) - A(\theta)\|_\infty = o_{\mathbb{P}}(1) \quad \textit{and} \quad \sup_{\theta\in\Theta} \|B_n(\theta) - B(\theta)\|_\infty = o_{\mathbb{P}}(1),$$

where $\|\cdot\|_\infty$ *is the matrix maximum norm;*
(vi) $\sqrt{n}[(\widehat{\theta}_n - \theta_*)', \mathrm{vech}[A_n - A_*]', \mathrm{vech}[B_n - B_*]']' = O_{\mathbb{P}}(1)$*;*
(vii) for $j = 1, \ldots, \ell$*,* $\partial_j A_n(\cdot)$ *and* $\partial_j B_n(\cdot)$ *are consistent for* $\partial_j A(\cdot)$ *and* $\partial_j B(\cdot)$*, uniformly on* Θ*; and*
(viii) for $j = 1, \ldots, \ell$*,* $H_{j,n} = O_{\mathbb{P}}(n^{-1/2})$ *and* $G_{j,n} = O_{\mathbb{P}}(n^{-1/2})$*, where* $H_{j,n} := A_*^{-1}\partial_j(A_n - A_*)$ *and* $G_{j,n} := B_*^{-1}\partial_j(B_n - B_*)$*.* □

These conditions hold for most standard estimators based on (Q)MLE, LS, or (G)MM procedures when applied in standard environments. The same framework was employed in CW and facilitates comparison of our tests and findings with theirs under the same conditions.

Our omnibus tests are motivated by testing whether the critical quantities τ_*, δ_*, η_*, σ_* ξ_*, and γ_*, which we call the test base elements, equal zero. We first examine stochastic asymptotic representations of consistent estimates of these quantities. For this purpose, we let

$$L_n := P_n + \sum_{j=1}^{\ell} (\widehat{\theta}_{j,n} - \theta_{j,*}) R_{j,*}$$

for notational simplicity, where $P_n := W_n - U_n := B_*^{-1}(B_n - B_*) - A_*^{-1}(A_n - A_*)$, and for $j = 1, 2, \ldots, \ell$, $R_{j,*} := B_*^{-1}\partial_j B_* - A_*^{-1}\partial_j A_*$. Note under Assumption A we have $L_n = O_{\mathbb{P}}(n^{-1/2})$, $P_n = O_{\mathbb{P}}(n^{-1/2})$, and for $j = 1, 2, \ldots, \ell$, $R_{j,*} = O(1)$. These correspond with the definitions in CW. The following lemma provides the explicit asymptotic expansions.

Lemma 2. *Given Assumption A,*
(i) $\widehat{\tau}_n = \tau_* + k^{-1}\mathrm{tr}[L_n D_*'] + O_{\mathbb{P}}(n^{-1})$*;*

(ii) $\widehat{\delta}_n = \delta_* + k^{-1}\det[D_*]^{1/k}\mathrm{tr}[L_n] + O_{\mathbb{P}}(n^{-1})$;

(iii) $\widehat{\eta}_n = \eta_* + k^{-1}\mathrm{tr}[L_n\widetilde{D}_*{}']/(k^{-1}\mathrm{tr}[\widetilde{D}_*])^2 + O_{\mathbb{P}}(n^{-1})$;

(iv) $\widehat{\sigma}_n = \sigma_* + k^{-1}\mathrm{tr}[L_n D_*{}'] - k^{-1}\det[D_*]^{1/k}\mathrm{tr}[L_n] + O_{\mathbb{P}}(n^{-1})$;

(v) $\widehat{\xi}_n = \xi_* + k^{-1}\mathrm{tr}[L_n D_*{}'] - k^{-1}\mathrm{tr}[L_n\widetilde{D}_*{}']/(k^{-1}\mathrm{tr}[\widetilde{D}_*])^2 + O_{\mathbb{P}}(n^{-1})$; *and*

(vi) $\widehat{\gamma}_n = \gamma_* + k^{-1}\det[D_*]^{1/k}\mathrm{tr}[L_n] - k^{-1}\mathrm{tr}[L_n\widetilde{D}_*{}']/(k^{-1}\mathrm{tr}[\widetilde{D}_*])^2 + O_{\mathbb{P}}(n^{-1})$. □

Several remarks are in order. First, we do not prove Lemma 2 in the Appendix. Lemma 2(*i*, *ii*, and *iv*) are already established in lemma 4 of CW, and Lemma 2(*iii*) holds as a corollary of Lemma 3 below. The other results trivially follow from Lemma 2(*i*, *ii*, and *iii*). Second, some signs of the statistics are predetermined by the interrelationships between the Pythagorean means. That is, we know $\widehat{\tau}_n \geq \widehat{\delta}_n \geq \widehat{\eta}_n$, so that $\widehat{\sigma}_n$, $\widehat{\xi}_n$, and $\widehat{\gamma}_n$ are always greater than zero. Third, the asymptotic approximations of the statistics in Lemma 2 have different forms under $\mathcal{H}_0$ and $\mathcal{H}_1$. If $\mathcal{H}_0$ holds, $\tau_* = \delta_* = \eta_* = \sigma_* = \xi_* = \gamma_* = 0$, so that $\widehat{\tau}_n$, $\widehat{\delta}_n$, and $\widehat{\eta}_n$ are $O_{\mathbb{P}}(n^{-1/2})$, and $\widehat{\sigma}_n$, $\widehat{\xi}_n$, and $\widehat{\gamma}_n$ are $O_{\mathbb{P}}(n^{-1})$. On the other hand, $\widehat{\tau}_n$, $\widehat{\delta}_n$, $\widehat{\eta}_n$, $\widehat{\sigma}_n$, $\widehat{\xi}_n$, and $\widehat{\gamma}_n$ are $O_{\mathbb{P}}(1)$ under $\mathcal{H}_1$. These different forms make the test base element quantities useful in distinguishing $\mathcal{H}_0$ and $\mathcal{H}_1$.

We now define the first group of tests

$$\widehat{\mathfrak{B}}_n^{(1)} := \frac{nk}{4}\left(\widehat{\tau}_n^2 + \widehat{\delta}_n^2\right), \quad \widehat{\mathfrak{B}}_n^{(2)} := \frac{nk}{2}\left(\widehat{\tau}_n^2 + 2\widehat{\sigma}_n\right), \quad \text{and} \quad \widehat{\mathfrak{B}}_n^{(3)} := \frac{nk}{2}\left(\widehat{\delta}_n^2 + 2\widehat{\sigma}_n\right),$$

which modify the tests in CW. These tests exploit the discriminatory properties of the statistics $\widehat{\tau}_n$ and $\widehat{\delta}_n$, which embody elements of the Wald (1943) test principle. The coefficients of the statistics differ from those in CW: specifically, $\widehat{\mathfrak{B}}_n^{(1)}$ is (respectively, $\widehat{\mathfrak{B}}_n^{(2)}$ and $\widehat{\mathfrak{B}}_n^{(3)}$ are) defined by dividing the corresponding test in CW by $2k$ (respectively, 2). As detailed below, this modification is useful to achieve a direct comparison of the leading terms that are obtained as approximations of the tests under the alternative.

We define a second group of tests as follows:

$$\widehat{\mathfrak{D}}_n^{(1)} := \frac{nk}{4}\left(\widehat{\tau}_n^2 + \widehat{\eta}_n^2\right), \quad \widehat{\mathfrak{D}}_n^{(2)} := \frac{nk}{2}\left(\widehat{\tau}_n^2 + \widehat{\xi}_n\right), \quad \text{and} \quad \widehat{\mathfrak{D}}_n^{(3)} := \frac{nk}{2}\left(\widehat{\eta}_n^2 + \widehat{\xi}_n\right);$$

and

$$\widehat{\mathfrak{S}}_n^{(1)} := \frac{nk}{4}\left(\widehat{\delta}_n^2 + \widehat{\eta}_n^2\right), \quad \widehat{\mathfrak{S}}_n^{(2)} := \frac{nk}{2}\left(\widehat{\delta}_n^2 + 2\widehat{\gamma}_n\right), \quad \text{and} \quad \widehat{\mathfrak{S}}_n^{(3)} := \frac{nk}{2}\left(\widehat{\eta}_n^2 + 2\widehat{\gamma}_n\right).$$

Note that $\widehat{\mathfrak{D}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, and $\widehat{\mathfrak{D}}_n^{(3)}$ are defined by associating the arithmetic mean with the harmonic mean, whereas $\widehat{\mathfrak{S}}_n^{(1)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, and $\widehat{\mathfrak{S}}_n^{(3)}$ are defined by associating the geometric mean with the harmonic mean. As before, $\widehat{\tau}_n$, $\widehat{\delta}_n$, and $\widehat{\eta}_n$ are empowered with discriminatory capability.

In addition, we define other omnibus test statistics by reversing the roles of A_* and B_* or by combining all test base elements. For this purpose, we first let $\widetilde{\tau}_n := k^{-1}\mathrm{tr}[\widetilde{D}_n] - 1$, $\widetilde{\eta}_n := k/\mathrm{tr}[\widetilde{D}_n^{-1}] - 1$, $\widetilde{\delta}_n := \det[\widetilde{D}_n]^{1/k} - 1$, $\widetilde{\xi}_n := \widetilde{\tau}_n - \widetilde{\eta}_n$, $\widetilde{\gamma}_n := \widetilde{\delta}_n - \widetilde{\eta}_n$, and $\widetilde{\sigma}_n := \widetilde{\tau}_n - \widetilde{\delta}_n$. Note that the only difference between $\widetilde{D}_n$ and $\widehat{D}_n$ is in the fact that the roles of $\widehat{A}_n$ and $\widehat{B}_n$ is reversed. Using $\widetilde{D}_n$, additional test statistics are defined as before. Let $\widetilde{\mathfrak{B}}_n^{(j)}$, $\widetilde{\mathfrak{D}}_n^{(j)}$, and $\widetilde{\mathfrak{S}}_n^{(j)}$ be the test statistics corresponding to $\widehat{\mathfrak{B}}_n^{(j)}$, $\widehat{\mathfrak{D}}_n^{(j)}$, and $\widehat{\mathfrak{S}}_n^{(j)}$ $(j = 1, 2, 3)$, respectively, that are obtained by using $\widetilde{\tau}_n$, $\widetilde{\eta}_n$, and $\widetilde{\delta}_n$. Next, we also consider the following test statistics:

$$\widehat{\mathfrak{E}}_n^{(1)} := \frac{nk}{2}\left(\widehat{\tau}_n^2 + 2\widehat{\gamma}_n\right), \quad \widehat{\mathfrak{E}}_n^{(2)} := \frac{nk}{2}\left(\widehat{\delta}_n^2 + \widehat{\xi}_n\right), \quad \text{and} \quad \widehat{\mathfrak{E}}_n^{(3)} := \frac{nk}{2}\left(\widehat{\eta}_n^2 + 2\widehat{\sigma}_n\right)$$

to test $\mathcal{H}_0^{(10)}$, $\mathcal{H}_0^{(11)}$, and $\mathcal{H}_0^{(12)}$, respectively. The motivations of these tests are the same as for the earlier test statistics. Similarly, we also consider the test statistics that are obtained by reversing the roles of $\widehat{A}_n$ and $\widehat{B}_n$ and denote them as $\widetilde{\mathfrak{E}}_n^{(1)}$, $\widetilde{\mathfrak{E}}_n^{(2)}$, and $\widetilde{\mathfrak{E}}_n^{(3)}$ that correspond with $\widehat{\mathfrak{E}}_n^{(1)}$, $\widehat{\mathfrak{E}}_n^{(2)}$, and $\widehat{\mathfrak{E}}_n^{(3)}$, respectively. Indeed, these additional test statistics are linear combinations of the previous six test statistics, viz.,

$$\widehat{\mathfrak{E}}_n^{(1)} \equiv 2\widehat{\mathfrak{D}}_n^{(2)} - \widehat{\mathfrak{B}}_n^{(2)}, \qquad \widehat{\mathfrak{E}}_n^{(2)} \equiv \frac{1}{2}\widehat{\mathfrak{B}}_n^{(3)} + \frac{1}{2}\widehat{\mathfrak{S}}_n^{(2)}, \qquad \text{and} \qquad \widehat{\mathfrak{E}}_n^{(3)} \equiv 2\widehat{\mathfrak{D}}_n^{(3)} - \widehat{\mathfrak{S}}_n^{(3)}, \tag{2}$$

which implies that the asymptotic behaviors of $\widehat{\mathfrak{E}}_n^{(1)}$, $\widehat{\mathfrak{E}}_n^{(2)}$, and $\widehat{\mathfrak{E}}_n^{(3)}$ are determined by those of $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$, and there exists a similar association between $(\widetilde{\mathfrak{E}}_n^{(1)}, \widetilde{\mathfrak{E}}_n^{(2)}, \widetilde{\mathfrak{E}}_n^{(3)})$ with $(\widetilde{\mathfrak{B}}_n^{(i)}, \widetilde{\mathfrak{D}}_n^{(i)}, \widetilde{\mathfrak{S}}_n^{(i)})$ $(i = 2, 3)$. Below we show that these additional test statistics are equivalent to $(\widehat{\mathfrak{B}}_n^{(i)}, \widehat{\mathfrak{D}}_n^{(i)}, \widehat{\mathfrak{S}}_n^{(i)})$ $(i = 2, 3)$ under the null and local alternative using these identities.

Before moving to the next subsection, we make an additional remark. It is useful to think of the tests within the generic form as $T_n := f_n(\widehat{\tau}_n, \widehat{\delta}_n, \widehat{\eta}_n, \widehat{\sigma}_n, \widehat{\xi}_n, \widehat{\gamma}_n, \widetilde{\tau}_n, \widetilde{\delta}_n, \widetilde{\eta}_n, \widetilde{\sigma}_n, \widetilde{\xi}_n, \widetilde{\gamma}_n)$ with $f_n(\cdot)$ being a continuous distance function uniformly in n. Accordingly, a number of distance functions can be used to define tests, just as those employed in the current test statistics. For example, the uniform norm is popularly applied to continuous Gaussian processes in the field of specification testing for the conditional mean (*e.g.*, Cho and White, 2011; Baek, Cho, and Phillips, 2015). When the testing framework is embedded in this general setting, it is a natural question whether there exist functions capable of delivering optimal-powered tests, a challenging research topic that is beyond the scope of the current study. Instead, we examine below the relative performance of the tests in simulations under various alternatives and use these results to identify practically useful test statistics. Our tests also include the uniform-norm based test statistic.

3.2 Asymptotic Null Approximations of the Test Statistics

We now develop null approximations for each of the tests and start the development with corresponding null approximations of the test base elements. For notational simplicity, let

$$K_n := A_*^{-1}\{B_n - A_n + \sum_{j=1}^{\ell} \partial_j(B_* - A_*)(\widehat{\theta}_{j,n} - \theta_{j,*})\},$$

which follows by imposing the null $A_* = B_*$ on the linearization L_n. The following result is derived from Lemma 2.

Corollary 1. *Given Assumption A and $\mathcal{H}_0$,*
(i) $\widehat{\tau}_n = k^{-1}\mathrm{tr}[K_n] + O_{\mathbb{P}}(n^{-1})$;
(ii) $\widehat{\delta}_n = k^{-1}\mathrm{tr}[K_n] + O_{\mathbb{P}}(n^{-1})$;
(iii) $\widehat{\eta}_n = k^{-1}\mathrm{tr}[K_n] + O_{\mathbb{P}}(n^{-1})$;
(iv) $\widehat{\sigma}_n = O_{\mathbb{P}}(n^{-1})$;
(v) $\widehat{\xi}_n = O_{\mathbb{P}}(n^{-1})$*; and*
(vi) $\widehat{\gamma}_n = O_{\mathbb{P}}(n^{-1})$. □

Items (*i*), (*ii*), and (*iv*) of Corollary 1 are already available in CW.

The main implication of Corollary 1 is that $\widehat{\tau}_n$, $\widehat{\delta}_n$, and $\widehat{\eta}_n$ are asymptotically equivalent under the null, so that $\widehat{\sigma}_n$, $\widehat{\xi}_n$, and $\widehat{\gamma}_n$ have a convergence rate n^{-1} that is faster than $\widehat{\tau}_n$, $\widehat{\delta}_n$ and $\widehat{\eta}_n$. This aspect was noticed by CW in the case of $\widehat{\sigma}_n$. Corollary 1 extends their results by showing that the same properties apply for $\widehat{\xi}_n$ and $\widehat{\gamma}_n$. In consequence, the desired asymptotic null approximations involve study of higher order approximants for $\widehat{\tau}_n$, $\widehat{\delta}_n$, and $\widehat{\eta}_n$. Lemma 4 of CW provides these for $\widehat{\tau}_n$ and $\widehat{\delta}_n$, and we present them here for completeness to make this study self-contained:

$$\begin{aligned}\widehat{\tau}_n = \tau_* + k^{-1}\{\mathrm{tr}[L_n(I-U_n)D_*'] + [\mathrm{tr}[(J_{j,n} - P_n A_*^{-1}\partial_j A_*)D_*']]'(\widehat{\theta}_n - \theta_*)\} \\ + k^{-1}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2 \mathrm{tr}[D_*](\widehat{\theta}_n - \theta_*)/2 + o_{\mathbb{P}}(n^{-1}); \quad (3)\end{aligned}$$

$$\begin{aligned}\widehat{\delta}_n = \delta_* + k^{-1}\det[D_*]^{\frac{1}{k}}\{\mathrm{tr}[L_n] + \left(k^{-1}-1\right)\mathrm{tr}[L_n]^2/2 + (\mathrm{tr}[P_n]^2 + \mathrm{tr}[U_n^2] - \mathrm{tr}[W_n^2])/2\} \\ + k^{-1}\det[D_*]^{\frac{1}{k}}[\mathrm{tr}[P_n]\mathrm{tr}[R_{j,*}] + \mathrm{tr}[J_{j,n} + U_n A_*^{-1}\partial_j A_* - W_n B_*^{-1}\partial_j B_*]]'(\widehat{\theta}_n - \theta_*) \\ + k^{-1}\det[D_*]^{\frac{1}{k}-1}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\det[D_*](\widehat{\theta}_n - \theta_*)/2 + o_{\mathbb{P}}(n^{-1}), \quad (4)\end{aligned}$$

where $J_{j,n} := G_{j,n} - H_{j,n}$.

The following lemma gives the next-order expansion of $\widehat{\eta}_n$.

Lemma 3. *(i) Given Assumption A,*

$$\begin{aligned}\widehat{\eta}_n = \eta_* + k^{-1}\mathrm{tr}[L_n \widetilde{D}_*']/(k^{-1}\mathrm{tr}[\widetilde{D}_*])^2 \\ + (k^{-1}\mathrm{tr}[L_n \widetilde{D}_*'])^2/(k^{-1}\mathrm{tr}[\widetilde{D}_*])^3 - \{k^{-1}\mathrm{tr}[L_n W_n \widetilde{D}_*']\}/(k^{-1}\mathrm{tr}[\widetilde{D}_*])^2 \\ - \{k^{-1}[\mathrm{tr}[(-J_{j,n} + P_n B_*^{-1}\partial_j B_*)\widetilde{D}_*']]'(\widehat{\theta}_n - \theta_*)\}/(k^{-1}\mathrm{tr}[\widetilde{D}_*])^2 \\ - k^{-1}\{(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2 \mathrm{tr}[\widetilde{D}_*](\widehat{\theta}_n - \theta_*)\}/\{2(k^{-1}\mathrm{tr}[\widetilde{D}_*])^2\} + o_{\mathbb{P}}(n^{-1});\end{aligned}$$

(ii) If $\mathcal{H}_0$ holds, $\widehat{\eta}_n = \widehat{\eta}_n^\star + o_{\mathbb{P}}(n^{-1})$, where

$$\begin{aligned}\widehat{\eta}_n^\star \;:=&\; k^{-1}\mathrm{tr}[K_n] + (k^{-1}\mathrm{tr}[K_n])^2 - k^{-1}\mathrm{tr}[K_n W_n] \\ & -k^{-1}[\mathrm{tr}[(-J_{j,n} + M_n B_*^{-1}\partial_j B_*)]]'(\widehat{\theta}_n - \theta_*) - k^{-1}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2 \mathrm{tr}[\widetilde{D}_*](\widehat{\theta}_n - \theta_*)/2,\end{aligned}$$

and $M_n := B_*^{-1}(B_n - A_n)$. □

Lemma 3 is proved in the Appendix. Note that Lemma 2(*iii*) and Corollary 1(*iii*) follow from Lemma 3. Here, $\widehat{\eta}_n^\star$ of Lemma 3(*ii*) is the second-order expansion of $\widehat{\eta}_n$ under $\mathcal{H}_0$, and it is not hard to see that $\widehat{\eta}_n^\star = k^{-1}\mathrm{tr}[K_n] + O_{\mathbb{P}}(n^{-1})$ under Assumption A.

Lemma 3 can also be used to obtain the next-order asymptotic expansions of $\widehat{\xi}_n$ and $\widehat{\gamma}_n$ in Lemma 2, and these are provided in the following result.

Lemma 4. *Given Assumption A, we have:*

(i) if for all $d > 0$, $B_* \neq dA_*$,

(i.a) $\widehat{\xi}_n = \xi_* + k^{-1}\mathrm{tr}[L_n D_*'] - k^{-1}\mathrm{tr}[L_n \widetilde{D}_*']/(k^{-1}\mathrm{tr}[\widetilde{D}_*])^2 + O_{\mathbb{P}}(n^{-1})$;

(i.b) $\widehat{\gamma}_n = \gamma_* + k^{-1}\det[D_*]^{1/k}\mathrm{tr}[L_n] - k^{-1}\mathrm{tr}[L_n \widetilde{D}_*']/(k^{-1}\mathrm{tr}[\widetilde{D}_*])^2 + O_{\mathbb{P}}(n^{-1})$;

(ii) if for some $d_* > 0$, $B_* = d_* A_*$,

(ii.a) $\widehat{\xi}_n = d_*\left\{k^{-1}\mathrm{tr}[L_n^2] - k^{-2}\mathrm{tr}[L_n]^2\right\} + o_{\mathbb{P}}(n^{-1})$;

(ii.b) $\widehat{\gamma}_n = d_*\left\{k^{-1}\mathrm{tr}[L_n^2] - k^{-2}\mathrm{tr}[L_n]^2\right\}/2 + o_{\mathbb{P}}(n^{-1})$;

(iii) If in addition $\mathcal{H}_0$ holds,

(iii.a) $\widehat{\xi}_n = k^{-1}\mathrm{tr}[K_n^2] - k^{-2}\mathrm{tr}[K_n]^2 + o_{\mathbb{P}}(n^{-1})$;

(iii.b) $\widehat{\gamma}_n = \{k^{-1}\mathrm{tr}[K_n^2] - k^{-2}\mathrm{tr}[K_n]^2\}/2 + o_{\mathbb{P}}(n^{-1})$. □

We prove Lemma 4 by combining Lemma 3 and lemma 4 of CW. The point of Lemma 4(*i.a*) is that the asymptotic expansions for $\widehat{\xi}_n$ and $\widehat{\gamma}_n$ in Lemma 2 are useful when B_* is not proportional to A_*. If B_* is proportional to A_*, they are not usefully exploited in the approximations. Further higher-order expansions are needed, and they are given in Lemma 4(*ii*). CW's corollary 5 observes the same feature for $\widehat{\sigma}_n$. For convenience, we state their result here: if for all $d > 0$, $B_* \neq dA_*$, $\widehat{\sigma}_n = \sigma_* + k^{-1}\mathrm{tr}[(D_*{}' - \det[D_*]^{\frac{1}{k}} I)L_n] + o_{\mathbb{P}}(n^{-1/2})$; and if for some $d_* > 0$, $B_* = d_* A_*$, $\widehat{\sigma}_n = -d_*(2k^2)^{-1}\mathrm{tr}[L_n]^2 + d_*(2k)^{-1}\ \mathrm{tr}[L_n^2] + o_{\mathbb{P}}(n^{-1})$. Note the same property holds as those for $\widehat{\xi}_n$ and $\widehat{\gamma}_n$: the asymptotic expansion order of $\widehat{\sigma}_n$ depends on whether B_* is proportional to A_* or not.

Lemmas 3 and 4 now straightforwardly deliver the asymptotic null approximations of the tests. We collect these together in the following theorem which characterizes the relationships between the test statistics. In the statement of the result, and elsewhere in the paper we use $\mathrm{tr}\,[F]^2$ to represent $(\mathrm{tr}\,[F])^2$.

Theorem 1. *Given Assumption A and $\mathcal{H}_0$,*
(i) $\widehat{\mathfrak{B}}_n^{(1)} = \frac{n}{2k}\mathrm{tr}[K_n]^2 + o_{\mathbb{P}}(1)$, $\widehat{\mathfrak{B}}_n^{(2)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$, *and* $\widehat{\mathfrak{B}}_n^{(3)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$*;*
(ii) $\widehat{\mathfrak{D}}_n^{(1)} = \frac{n}{2k}\mathrm{tr}[K_n]^2 + o_{\mathbb{P}}(1)$, $\widehat{\mathfrak{D}}_n^{(2)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$, *and* $\widehat{\mathfrak{D}}_n^{(3)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$*; and*
(iii) $\widehat{\mathfrak{S}}_n^{(1)} = \frac{n}{2k}\mathrm{tr}[K_n]^2 + o_{\mathbb{P}}(1)$, $\widehat{\mathfrak{S}}_n^{(2)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$, *and* $\widehat{\mathfrak{S}}_n^{(3)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$. □

Theorem 1(*i*) corresponds to theorem 1 of CW. Applying their theorem yields Theorem 1(*i*) as a corollary. Theorems 1(*ii* and *iii*) also hold as corollaries of Corollary 1 and Lemma 4(*iii*). From Theorem 1, it follows that $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ are asymptotically equivalent under $\mathcal{H}_0$. Furthermore, $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{B}}_n^{(3)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, and $\widehat{\mathfrak{S}}_n^{(3)}$ are also asymptotically equivalent.

Theorem 1 has implications for the test statistics additional to the nine tests in Theorem 1. First, reversing the roles of $\widehat{A}_n$ and $\widehat{B}_n$ interchanges the positions of B_n and B_* in the definition of K_n with those of A_n and A_*, respectively, implying that $-K_n$ becomes the score for the test statistics using $\widetilde{D}_n$. Therefore, $\widetilde{\mathfrak{B}}_n^{(1)}$, $\widetilde{\mathfrak{D}}_n^{(1)}$, and $\widetilde{\mathfrak{S}}_n^{(1)}$ are approximated as $\frac{n}{2k}\mathrm{tr}[-K_n]^2 + o_{\mathbb{P}}(1)$, and $\widetilde{\mathfrak{B}}_n^{(i)}$, $\widetilde{\mathfrak{D}}_n^{(i)}$, and $\widetilde{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$ are approximated as $\frac{n}{2}\mathrm{tr}[(-K_n)^2] + o_{\mathbb{P}}(1)$ under $\mathcal{H}_0 : A_* = B_*$, so that the test statistics defined by $\widetilde{D}_n$ are also equivalent to $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{B}}_n^{(3)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, and $\widehat{\mathfrak{S}}_n^{(3)}$. Second, Theorem 1 and (2) imply that the asymptotic approximations of $\widehat{\mathfrak{E}}_n^{(1)}$, $\widehat{\mathfrak{E}}_n^{(2)}$, and $\widehat{\mathfrak{E}}_n^{(3)}$ are obtained as $\frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$ under $\mathcal{H}_0$. Therefore, $\widehat{\mathfrak{E}}_n^{(1)}$, $\widehat{\mathfrak{E}}_n^{(2)}$, and $\widehat{\mathfrak{E}}_n^{(3)}$ are asymptotically equivalent to $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{B}}_n^{(3)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, and $\widehat{\mathfrak{S}}_n^{(3)}$ under $\mathcal{H}_0$. Finally, the first and second implications imply that $\widetilde{\mathfrak{E}}_n^{(1)}$, $\widetilde{\mathfrak{E}}_n^{(2)}$, and $\widetilde{\mathfrak{E}}_n^{(3)}$ are also equivalent. We collect these results in the following corollary.

Corollary 2. *Given Assumption A and $\mathcal{H}_0$,*
(i) $\widetilde{\mathfrak{B}}_n^{(1)} = \frac{n}{2k}\mathrm{tr}[K_n]^2 + o_{\mathbb{P}}(1)$, $\widetilde{\mathfrak{D}}_n^{(1)} = \frac{n}{2k}\mathrm{tr}[K_n]^2 + o_{\mathbb{P}}(1)$, *and* $\widetilde{\mathfrak{S}}_n^{(1)} = \frac{n}{2k}\mathrm{tr}[K_n]^2 + o_{\mathbb{P}}(1)$*;*
(ii) $\widetilde{\mathfrak{B}}_n^{(i)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$, $\widetilde{\mathfrak{D}}_n^{(i)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$, *and* $\widetilde{\mathfrak{S}}_n^{(i)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$ $(i = 2, 3)$*; and*
(iii) $\widehat{\mathfrak{E}}_n^{(i)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$ *and* $\widetilde{\mathfrak{E}}_n^{(i)} = \frac{n}{2}\mathrm{tr}[K_n^2] + o_{\mathbb{P}}(1)$ $(i = 1, 2, 3)$. □

3.3 Asymptotic Alternative Approximations of the Test Statistics

We now examine asymptotic approximations of the tests under the alternative. As before, we first examine asymptotic approximations of the test base elements. For notational simplicity, we let $\widetilde{\tau}_* := k^{-1}\mathrm{tr}[\widetilde{D}_*] - 1$, $\widetilde{\eta}_* := k/\mathrm{tr}[\widetilde{D}_*^{-1}] - 1$, $\widetilde{\delta}_* := \det[\widetilde{D}_*]^{1/k} - 1$, $\widetilde{\xi}_* := \widetilde{\tau}_* - \widetilde{\eta}_*$, $\widetilde{\gamma}_* := \widetilde{\delta}_* - \widetilde{\eta}_*$, and $\widetilde{\sigma}_* := \widetilde{\tau}_* - \widetilde{\delta}_*$. Note that these are the ones that correspond to the previous testing quantities, but the roles of A_* and B_* are reversed, and there exist specific equivalent relationships: $\widetilde{\tau}_* = -\eta_*/(1+\eta_*)$, $\widetilde{\delta}_* = -\delta_*/(1+\delta_*)$ and $\widetilde{\eta}_* = -\tau_*/(1+\tau_*)$. The asymptotic approximations of the test base elements are easily obtained by combining Lemmas 3, 4, and lemma 4 of CW. The following corollary collects them together.

Corollary 3. *Given Assumption A,*

(i) if for all $d > 0$, $B_* \neq dA_*$,

(i.a) $\widehat{\mathfrak{B}}_n^{(1)} = \frac{nk}{2}(\frac{1}{2}\tau_*^2 + \frac{1}{2}\delta_*^2) + \frac{n}{2}\{\tau_*\mathrm{tr}[D_*'L_n] + \delta_*(\delta_*+1)\mathrm{tr}[L_n]\} + O_{\mathbb{P}}(1)$;

(i.b) $\widehat{\mathfrak{B}}_n^{(2)} = \frac{nk}{2}(\tau_*^2 + 2\sigma_*) + n\{(\tau_*+1)\mathrm{tr}[D_*'L_n] - (\delta_*+1)\mathrm{tr}[L_n]\} + O_{\mathbb{P}}(1)$;

(i.c) $\widehat{\mathfrak{B}}_n^{(3)} = \frac{nk}{2}(\delta_*^2 + 2\sigma_*) + n\{\mathrm{tr}[D_*'L_n] + (\delta_*^2 - 1)\mathrm{tr}[L_n]\} + O_{\mathbb{P}}(1)$;

(i.d) $\widehat{\mathfrak{D}}_n^{(1)} = \frac{nk}{2}(\frac{1}{2}\tau_*^2 + \frac{1}{2}\eta_*^2) + \frac{n}{2}\{\tau_*\mathrm{tr}[D_*'L_n] + (\eta_*^3/\widetilde{\tau}_*^2)\mathrm{tr}[\widetilde{D}_*'L_n]\} + O_{\mathbb{P}}(1)$;

(i.e) $\widehat{\mathfrak{D}}_n^{(2)} = \frac{nk}{2}(\tau_*^2 + \xi_*) + \frac{n}{2}\{(2\tau_*+1)\mathrm{tr}[D_*'L_n] - (\eta_*^2/\widetilde{\tau}_*^2)\mathrm{tr}[\widetilde{D}_*'L_n]\} + O_{\mathbb{P}}(1)$;

(i.f) $\widehat{\mathfrak{D}}_n^{(3)} = \frac{nk}{2}(\eta_*^2 + \xi_*) + \frac{n}{2}\{\mathrm{tr}[D_*'L_n] + (2\eta_* - 1)(\eta_*^2/\widetilde{\tau}_*^2)\mathrm{tr}[\widetilde{D}_*'L_n]\} + O_{\mathbb{P}}(1)$;

(i.g) $\widehat{\mathfrak{S}}_n^{(1)} = \frac{nk}{2}(\frac{1}{2}\delta_*^2 + \frac{1}{2}\eta_*^2) + \frac{n}{2}\{(\eta_*^3/\widetilde{\tau}_*^2)\mathrm{tr}[\widetilde{D}_*'L_n] + \delta_*(\delta_*+1)\mathrm{tr}[L_n]\} + O_{\mathbb{P}}(1)$;

(i.h) $\widehat{\mathfrak{S}}_n^{(2)} = \frac{nk}{2}(\delta_*^2 + 2\gamma_*) + n\{-(\eta_*^2/\widetilde{\tau}_*^2)\mathrm{tr}[\widetilde{D}_*'L_n] + (\delta_*+1)^2\mathrm{tr}[L_n]\} + O_{\mathbb{P}}(1)$;

(i.i) $\widehat{\mathfrak{S}}_n^{(3)} = \frac{nk}{2}(\eta_*^2 + 2\gamma_*) + n\{(\eta_* - 1)(\eta_*^2/\widetilde{\tau}_*^2)\mathrm{tr}[\widetilde{D}_*'L_n] + (\delta_*+1)\mathrm{tr}[L_n]\} + O_{\mathbb{P}}(1)$;

(ii) if for some $d_* > 0$, $B_* = d_*A_*$, for $j = 1,2,3$, $\widehat{\mathfrak{B}}_n^{(j)}$, $\widehat{\mathfrak{D}}_n^{(j)}$, and $\widehat{\mathfrak{S}}_n^{(j)}$ are equal to $\frac{nk}{2}(d_* - 1)^2 + nd_*(d_* - 1)\mathrm{tr}[L_n] + O_{\mathbb{P}}(1)$. □

Corollary 3 is implied by Lemmas 3, 4, and lemma 4 of CW so its proof is omitted.

Some remarks are warranted. First, the leading term of each test is the first term on the right sides of (*i*, *ii*), which in each case is $O(n)$. The second terms are $O_{\mathbb{P}}(n^{1/2})$ given that $L_n = O_{\mathbb{P}}(n^{-1/2})$. This divergence implies that all tests are consistent, and from this, the omnibus powers are derived. Second, if B_* is proportional to A_*, every test is equivalent even under the alternative. Third, if B_* is not proportional to A_*, the leading terms of the right sides in (*i.a*–*i.i*) are those that significantly affect the powers of the test statistics. The leading terms of all of our tests are affected by k in the same manner, so that their powers can be compared without considering the effects of k. The next-order terms are always determined as linear combinations of two of $\mathrm{tr}[D_*'L_n]$, $\mathrm{tr}[\widetilde{D}_n'L_n]$, and $\mathrm{tr}[L_n]$. As their coefficients are not always the same, the discriminatory powers of the test statistics are also affected by the next-order terms. Fourth, the asymptotic approximations in Corollary 3 can be properly exploited to obtain the asymptotic approximations of the additional test statistics. If the roles of A_* and B_* are reversed, the asymptotic approximations of the relevant test statistics are similarly obtained: the asymptotic approximations of $\widetilde{\mathfrak{B}}_n^{(i)}$, $\widetilde{\mathfrak{D}}_n^{(i)}$, and $\widetilde{\mathfrak{S}}_n^{(i)}$ ($i = 1,2,3$) are obtained by replacing τ_*, δ_*, η_*, σ_*, γ_*, ξ_*, $\widetilde{\tau}_*$ and L_n in Corollary 3(*i*) with $\widetilde{\tau}_*$, $\widetilde{\delta}_*$, $\widetilde{\eta}_*$, $\widetilde{\sigma}_*$, $\widetilde{\gamma}_*$, $\widetilde{\xi}_*$, τ_*, and $-L_n$, respectively. In particular, reversing the roles of A_* and B_* modifies L_n into $-L_n$ by the definition of L_n even under the alternative. By these replacements, for example, we obtain $\widetilde{\mathfrak{B}}_n^{(1)} = \frac{nk}{2}(\frac{1}{2}\widetilde{\tau}_*^2 + \frac{1}{2}\widetilde{\delta}_*^2) - \frac{n}{2}(\mathrm{tr}[\widetilde{\tau}_n \widetilde{D}_*{}' + \widetilde{\delta}_* \det[\widetilde{D}_*]^{1/k} I]L_n) + O_{\mathbb{P}}(1)$ if A_* is not proportional to B_*. If $B_* = d_* A_*$, for $j = 1,2,3$, $\widetilde{\mathfrak{B}}_n^{(j)}$, $\widetilde{\mathfrak{D}}_n^{(j)}$, and $\widetilde{\mathfrak{S}}_n^{(j)}$ are equal to $\frac{nk}{2}(d_*^{-1}-1)^2 - nd_*^{-1}(d_*^{-1}-1)\mathrm{tr}[L_n] + O_{\mathbb{P}}(1)$. This result is obtained simply by replacing notation in Corollary 3, so we do not repeat the statements. Fifth, the fourth remark applies even to the asymptotic approximations of $\widehat{\mathfrak{E}}_n^{(1)}$, $\widehat{\mathfrak{E}}_n^{(2)}$, $\widehat{\mathfrak{E}}_n^{(3)}$, $\widetilde{\mathfrak{E}}_n^{(1)}$, $\widetilde{\mathfrak{E}}_n^{(2)}$, and $\widetilde{\mathfrak{E}}_n^{(3)}$. Their asymptotic approximations are obtained by combining Corollary 3 and (2). We collect these together in the following corollary.

Corollary 4. *Given Assumption A,*

(i) if for all $d > 0$, $B_* \neq dA_*$,

(i.a) $\widehat{\mathfrak{E}}_n^{(1)} = \frac{nk}{2}(\tau_*^2 + 2\gamma_*) + n\{\tau_* \mathrm{tr}[D_*{}'L_n] - (\eta_*^2/\widetilde{\tau}_*^2)\mathrm{tr}[\widetilde{D}_*{}'L_n] + (\delta_* + 1)\mathrm{tr}[L_n]\} + O_{\mathbb{P}}(1)$*;*

(i.b) $\widehat{\mathfrak{E}}_n^{(2)} = \frac{nk}{2}(\delta_*^2 + \xi_*) + \frac{n}{2}\{\mathrm{tr}[D_*{}'L_n] - (\eta_*^2/\widetilde{\tau}_*^2)\mathrm{tr}[\widetilde{D}_*{}'L_n] + 2\delta_*(\delta_* + 1)\mathrm{tr}[L_n]\} + O_{\mathbb{P}}(1)$*;*

(i.c) $\widehat{\mathfrak{E}}_n^{(3)} = \frac{nk}{2}(\eta_*^2 + 2\sigma_*) + n\{\mathrm{tr}[D_*{}'L_n] + (\eta_*^3/\widetilde{\tau}_*^2)\mathrm{tr}[\widetilde{D}_*{}'L_n] - (\delta_* + 1)\mathrm{tr}[L_n]\} + O_{\mathbb{P}}(1)$*;*

(ii) if for some $d_* > 0$, $B_* = d_* A_*$, for $j = 1,2,3$, $\widehat{\mathfrak{E}}_n^{(j)} = \frac{nk}{2}(d_* - 1)^2 + nd_*(d_* - 1)\mathrm{tr}[L_n] + O_{\mathbb{P}}(1)$. □

We do not repeat the statements for $\widetilde{\mathfrak{E}}_n^{(1)}$, $\widetilde{\mathfrak{E}}_n^{(2)}$, and $\widetilde{\mathfrak{E}}_n^{(3)}$ to avoid repetition. Finally, as the leading terms of the tests are the main determinants of test power and their roles are significant in terms of their divergence speeds, we compare them and determine conditions for each term to be greater than the others. For this comparison we let μ_* be the maximum of the leading terms in Corollaries 3(*i.a*–*i.i*) and 4(*i.a*–*i.c*). The results are collected in the following theorem.

Theorem 2. *Given Assumption A and $\mathcal{H}_1$, if for all $d > 0$, $B_* \neq dA_*$,*

(i) the leading terms of $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ cannot be equal to μ_*;

(ii) the leading term of $\widehat{\mathfrak{B}}_n^{(2)}$ is equal to μ_* if and only if $\tau_*^2 > \max[\delta_*^2, \eta_*^2]$ and $\sigma_* > \gamma_*$;

(iii) the leading term of $\widehat{\mathfrak{B}}_n^{(3)}$ is equal to μ_* if and only if $\delta_*^2 > \max[\tau_*^2, \eta_*^2]$ and $\sigma_* > \gamma_*$;

(iv) if the leading term of $\widehat{\mathfrak{D}}_n^{(2)}$ is equal to μ_*, $\tau_*^2 > \max[\delta_*^2, \eta_*^2]$ and $\sigma_* = \gamma_*$, and if $\tau_*^2 > \max[\delta_*^2, \eta_*^2]$ and $\sigma_* = \gamma_*$, the leading terms of $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, and $\widehat{\mathfrak{E}}_n^{(1)}$ are equal to μ_*;

(v) if the leading term of $\widehat{\mathfrak{D}}_n^{(3)}$ is equal to μ_*, $\eta_*^2 > \max[\delta_*^2, \tau_*^2]$ and $\sigma_* = \gamma_*$, and if $\eta_*^2 > \max[\delta_*^2, \tau_*^2]$ and $\sigma_* = \gamma_*$, the leading terms of $\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(3)}$, and $\widehat{\mathfrak{E}}_n^{(3)}$ are equal to μ_*;

(vi) the leading term of $\widehat{\mathfrak{S}}_n^{(2)}$ *is equal to* μ_* *if and only if* $\delta_*^2 > \max[\eta_*^2, \tau_*^2]$ *and* $\gamma_* > \sigma_*$*;*

(vii) the leading term of $\widehat{\mathfrak{S}}_n^{(3)}$ *is equal to* μ_* *if and only if* $\eta_*^2 > \max[\tau_*^2, \delta_*^2]$ *and* $\gamma_* > \sigma_*$*;*

(viii) the leading term of $\widehat{\mathfrak{E}}_n^{(1)}$ *is equal to* μ_* *if and only if* $\tau_*^2 > \max[\delta_*^2, \eta_*^2]$ *and* $\gamma_* > \sigma_*$*;*

(ix) if the leading term of $\widehat{\mathfrak{E}}_n^{(2)}$ *is equal to* μ_**,* $\delta_*^2 > \max[\tau_*^2, \eta_*^2]$ *and* $\sigma_* = \gamma_*$*, and if* $\delta_*^2 > \max[\tau_*^2, \eta_*^2]$ *and* $\sigma_* = \gamma_*$*, the leading terms of* $\widehat{\mathfrak{B}}_n^{(3)}$*,* $\widehat{\mathfrak{E}}_n^{(2)}$*, and* $\widehat{\mathfrak{S}}_n^{(2)}$ *are equal to* μ_**; and*

(x) the leading term of $\widehat{\mathfrak{E}}_n^{(3)}$ *is equal to* μ_* *if and only if* $\eta_*^2 > \max[\tau_*^2, \delta_*^2]$ *and* $\sigma_* > \gamma_*$*.* □

Some remarks on these results are warranted. First, in view of Theorem 2(*i*), some caution is needed in testing $A_* = B_*$ using $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ since the leading terms in their expansions cannot be greater than the others, although their local alternative properties may be different. The other tests yield asymptotically better inferential discrimination in terms of the leading terms of the tests. Second, the conditions in Theorem 2 are summarized into two dimensional conditions, viz. the maximum condition between σ_* and γ_* and that among τ_*^2, δ_*^2, and η_*^2. Table 1 describes these relations in a table format and provides the test statistics with the maximum leading term under each condition. By Table 1, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(3)}$, and $\widehat{\mathfrak{E}}_n^{(2)}$ cannot have the maximum leading term alone, although the other tests can do so, as stated in Theorem 2. We also note that each cell of Table 1 is occupied by at least one of the nine test statistics, implying that the nine test statistics have the greatest leading term under each region, whose union is the alternative parameter space. If the roles of A_* and B_* are reversed, the alternative hypothesis is also partitioned in a different way to yield test statistics that have the maximum leading term under the region characterized by the maximum condition between $\widetilde{\sigma}_*$ and $\widetilde{\gamma}_*$ and that among $\widetilde{\tau}_*^2$, $\widetilde{\delta}_*^2$, and $\widetilde{\eta}_*^2$. Third, the conditions in Theorems 2(*i*–*x*) can be consistently selected by estimating τ_*, δ_*, σ_* and by comparing the conditions in Theorem 2. For example, if $\widehat{\tau}_n^2 > \widehat{\delta}_n^2$, $\widehat{\tau}_n^2 > \widehat{\eta}_n^2$, and $\widehat{\sigma}_n \geq \widehat{\gamma}_n$ and the sample size is reasonably large, testing the hypotheses by relying on $\widehat{\mathfrak{B}}_n^{(2)}$ can be better than the other tests, although it does not necessarily mean that $\widehat{\mathfrak{B}}_n^{(2)}$ is always more powerful than the other tests. This feature motivates testing the hypothesis by using the maximum of the test statistics. We examine below the performance of this extremum test by Monte Carlo experiments. Finally, if $k = 2$, it follows that $\sigma_* \geq \gamma_*$ from the fact that $(\delta_* + 1)^2 = (\tau_* + 1)(\eta_* + 1)$. Therefore, the leading terms of $\widehat{\mathfrak{S}}_n^{(2)}$, $\widehat{\mathfrak{S}}_n^{(3)}$, and $\widehat{\mathfrak{E}}_n^{(2)}$ cannot be the maximum leading term.

3.4 Asymptotic Local Alternative Approximations of the Test Statistics

We now examine asymptotic approximations of the tests under local alternatives. We consider the following local alternative: for some symmetric positive-definite $\bar{A}_*$ and $\bar{B}_*$ such that $\bar{A}_* \neq \bar{B}_*$,

$$\mathcal{H}_\ell : A_{*,n} = A_* + n^{-1/2}\bar{A}_*, \quad B_{*,n} = B_* + n^{-1/2}\bar{B}_*, \quad \text{and} \quad A_* = B_*.$$

Here, as the sample size n tends to infinity, $A_{*,n}$ and $B_{*,n}$ converge to A_* and B_*, respectively, at the rate $n^{-1/2}$. Note that $\mathcal{H}_\ell$ reduces to $\mathcal{H}_0$ if $\bar{A}_* = \bar{B}_*$. The local alternative differs from the null by requiring that $\bar{A}_* \neq \bar{B}_*$. This local alternative generalizes that used in CW, where it is assumed that $\bar{A}_* = 0$.

The following separate conditions are imposed for the local alternative approximations.

Assumption B (Local Alternative)**.** *(i)* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a complete probability space;*
(ii) $\Theta \subset \mathbb{R}^\ell$ *is a compact convex set with non-empty interior and* $k \in \mathbb{N}$*;*
(iii) a sequence of measurable mappings $\{\widehat{\theta}_n : \Omega \mapsto \Theta\}$ *is consistent for a unique* $\theta_* \in \mathrm{int}(\Theta)$*;*
(iv) $A : \Theta \mapsto \mathbb{R}^{k\times k}$ *and* $B : \Theta \mapsto \mathbb{R}^{k\times k}$ *are in* $\mathcal{C}^{(2)}(\Theta)$*, and* A_* *and* B_* *are symmetric and positive definite;*
(v) $\bar{A} : \Theta \mapsto \mathbb{R}^{k\times k}$ *and* $\bar{B} : \Theta \mapsto \mathbb{R}^{k\times k}$ *are in* $\mathcal{C}^{(1)}(\Theta)$ *and such that* $\bar{A}_* := \bar{A}(\theta_*)$ *and* $\bar{B}_* := \bar{B}(\theta_*)$ *are symmetric and positive definite, and* $\bar{A}_* \neq \bar{B}_*$*;*
(vi) $A_n(\cdot)$ *and* $B_n(\cdot)$ *are consistent for* $A(\cdot)$ *and* $B(\cdot)$*, respectively, uniformly on* Θ*;*
(vii) $\sqrt{n}[(\widehat{\theta}_n - \theta_*)', \mathrm{vech}[A_n - A_{*,n}]', \mathrm{vech}[B_n - B_{*,n}]']' = O_{\mathbb{P}}(1)$*;*
(viii) for $j = 1, \ldots, \ell$*,* $\partial_j A_n(\cdot)$ *and* $\partial_j B_n(\cdot)$ *are consistent for* $\partial_j A(\cdot)$ *and* $\partial_j B(\cdot)$*, uniformly on* Θ*; and*
(ix) for $j = 1, \ldots, \ell$*,* $H_{j,o,n} = O_{\mathbb{P}}(n^{-1/2})$ *and* $G_{j,o,n} = O_{\mathbb{P}}(n^{-1/2})$*, where* $H_{j,o,n} := A_*^{-1}\partial_j(A_n - A_{*,n})$ *and* $G_{j,o,n} := B_*^{-1}\partial_j(B_n - B_{*,n})$*.* □

The major differences between Assumptions A and B are in B(*v*, *vii*, and *ix*). The localizing parameters $\bar{A}_*$ and $\bar{B}_*$ are formally introduced in Assumption B(*v*), and the other two conditions modify the corresponding conditions in Assumption A to accommodate the presence of the localizing parameters.

Before examining the local asymptotic approximations, we provide notations relevant to the main claims of this section. We define

$$W_{o,n} := B_*^{-1}(B_n - B_{*,n}); \quad W_{a,n} := B_{*,n}^{-1}(B_n - B_{*,n});$$
$$U_{o,n} := A_*^{-1}(A_n - A_{*,n}); \quad U_{a,n} := A_{*,n}^{-1}(A_n - A_{*,n});$$
$$P_{o,n} := W_{o,n} - U_{o,n}; \quad P_{a,n} := W_{a,n} - U_{a,n};$$
$$M_{o,n} := B_*^{-1}(B_n - A_n - B_{*,n} + A_{*,n}); \quad M_{a,n} := B_{*,n}^{-1}(B_n - A_n - B_{*,n} + A_{*,n});$$
$$L_{o,n} := P_{o,n} + \sum_{j=1}^{\ell}(\widehat{\theta}_{j,n} - \theta_{j,*})R_{j,*}; \quad K_{o,n} := M_{o,n} + \sum_{j=1}^{\ell}(\widehat{\theta}_{j,n} - \theta_{j,*})S_{j,*};$$
$$L_{a,n} := P_{a,n} + \sum_{j=1}^{\ell}(\widehat{\theta}_{j,n} - \theta_{j,*})(B_{*,n}^{-1}\partial_j B_{*,n} - A_{*,n}^{-1}\partial_j A_{*,n});$$
$$K_{a,n} := M_{a,n} + \sum_{j=1}^{\ell}(\widehat{\theta}_{j,n} - \theta_{j,*})B_{*,n}^{-1}(\partial_j B_{*,n} - \partial_j A_{*,n}),$$

where $S_{j*} := A_*^{-1}(\partial_j B_* - \partial_j A_*)$. These statistics are defined to highlight the asymptotic roles of the localizing parameters. The statistics indexed by the subscript "o" correspond to those in previous sections, in which the localizing parameters are absent (zero). On the other hand, the statistics indexed by the subscript "a" are defined to explicitly consider the asymptotic effects of the locality parameters. Specifically, the inverse matrices in $W_{o,n}$, $U_{o,n}$, and $M_{o,n}$ are different from those in $W_{a,n}$, $U_{a,n}$, and $M_{a,n}$, respectively. If the localizing parameters are zero matrices in the inverse matrices, $W_{o,n}$, $U_{o,n}$, and $M_{o,n}$ are reduced versions of $W_{a,n}$, $U_{a,n}$, and $M_{a,n}$. Further note that $A_* = B_*$ under $\mathcal{H}_\ell$, so that $P_{o,n} = M_{o,n}$, $R_{j,*} = S_{j,*}$, and $L_{o,n} = K_{o,n}$. Using this fact, we let

$$\begin{aligned}\widehat{\tau}_{o,n} &:= k^{-1}\mathrm{tr}[K_{o,n}(I - U_{o,n})] \\ &\qquad + k^{-1}[\mathrm{tr}[J_{j,o,n} - M_{o,n}A_*^{-1}\partial_j A_*]]'(\widehat{\theta}_n - \theta_*) + (2k)^{-1}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2 \mathrm{tr}[D_*](\widehat{\theta}_n - \theta_*);\end{aligned}$$

$$\begin{aligned}\widehat{\delta}_{o,n} &:= k^{-1}\mathrm{tr}[K_{o,n}] + (2k)^{-1}\left(k^{-1} - 1\right)\mathrm{tr}[K_{o,n}]^2 + (2k)^{-1}(\mathrm{tr}[M_{o,n}]^2 + \mathrm{tr}[U_{o,n}^2] - \mathrm{tr}[W_{o,n}^2]) \\ &\qquad + k^{-1}[\mathrm{tr}[J_{j,o,n} + U_{o,n}A_*^{-1}\partial_j A_* - W_{o,n}A_*^{-1}\partial_j B_*]]'(\widehat{\theta}_n - \theta_*) \\ &\qquad + k^{-1}[\mathrm{tr}[M_{o,n}]\mathrm{tr}[S_{j,*}]]'(\widehat{\theta}_n - \theta_*) + (2k)^{-1}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2 \det[D_*](\widehat{\theta}_n - \theta_*);\end{aligned}$$

$$\begin{aligned}\widehat{\eta}_{o,n} &:= k^{-1}\mathrm{tr}[K_{o,n}] + (k^{-1}\mathrm{tr}[K_{o,n}])^2 - k^{-1}\mathrm{tr}[K_{o,n}W_{o,n}] \\ &\qquad - k^{-1}[\mathrm{tr}[-J_{j,o,n} + M_{o,n}B_*^{-1}\partial_j B_*]]'(\widehat{\theta}_n - \theta_*) - (2k)^{-1}(\widehat{\theta}_n - \theta_*)'\mathrm{tr}[\widetilde{D}_*](\widehat{\theta}_n - \theta_*);\end{aligned}$$

and define $\widehat{\sigma}_{o,n} := \widehat{\tau}_{o,n} - \widehat{\delta}_{o,n}$, $\widehat{\xi}_{o,n} := \widehat{\tau}_{o,n} - \widehat{\eta}_{o,n}$, and $\widehat{\gamma}_{o,n} := \widehat{\delta}_{o,n} - \widehat{\eta}_{o,n}$, where

$$\begin{aligned}J_{j,o,n} &:= G_{j,o,n} - H_{j,o,n} := B_*^{-1}\partial_j(B_n - B_{*,n}) - A_*^{-1}\partial_j(A_n - A_{*,n}); \qquad \text{and} \\ J_{j,a,n} &:= G_{j,a,n} - H_{j,a,n} := B_{*,n}^{-1}\partial_j(B_n - B_{*,n}) - A_{*,n}^{-1}\partial_j(A_n - A_{*,n}).\end{aligned}$$

These are the second-order approximations of the test base elements that are obtained by imposing $A_* = B_*$ and by letting the localizing parameters be zero in the inverse matrices. These definitions are obtained by reformulating (3), (4), and Lemma 3(*i*) to fit the current context. In particular, $\widehat{\sigma}_{o,n}$ is simplified to $\widehat{\sigma}_{o,n} := -\frac{1}{2k^2}\mathrm{tr}[K_{o,n}]^2 + \frac{1}{2k}\mathrm{tr}[K_{o,n}^2]$ after some tedious algebra. All these statistics are $O_{\mathbb{P}}(n^{-1})$.

Due to the effects of ignoring the asymptotic impact of $\bar{A}_*$ and $\bar{B}_*$, these statistics poorly approximate the test base elements under $\mathcal{H}_\ell$. Their differences from the second-order approximations of the tests are not asymptotically negligible and affect their asymptotic approximations under $\mathcal{H}_\ell$. The following lemma explicitly shows their differences.

Lemma 5. *Given Assumption B and $\mathcal{H}_\ell$,*

$$(i)\ \ \widehat{\tau}_n - \widehat{\tau}_{o,n} = n^{-1/2}k^{-1}\mathrm{tr}[V_*] - n^{-1/2}k^{-1}\mathrm{tr}[F_*W_{o,n} - C_*U_{o,n}] + n^{-1/2}k^{-1}\mathrm{tr}[K_{o,n}V_*] - (nk)^{-1}\mathrm{tr}[C_*V_*] + n^{-1/2}k^{-1}[\mathrm{tr}[Q_{j,*} - (F_*B_*^{-1}\partial_j B_* - C_*A_*^{-1}\partial_j A_*)]]'(\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}),$$

where $F_* := B_*^{-1}\bar{B}_*$, $C_* := A_*^{-1}\bar{A}_*$, $V_* := F_* - C_*$, and $Q_{j,*} := B_*^{-1}\partial_j\bar{B}_* - A_*^{-1}\partial_j\bar{A}_*$;

$$(ii)\ \ \widehat{\delta}_n - \widehat{\delta}_{o,n} = n^{-1/2}k^{-1}\mathrm{tr}[V_*] - n^{-1/2}k^{-1}\mathrm{tr}[F_*W_{o,n} - C_*U_{o,n}] + n^{-1/2}k^{-2}\mathrm{tr}[V_*]\mathrm{tr}[K_{o,n}] + (2nk^2)^{-1}\mathrm{tr}[V_*]^2 + (2nk)^{-1}(\mathrm{tr}[C_*^2] - \mathrm{tr}[F_*^2]) + n^{-1/2}k^{-1}[\mathrm{tr}[Q_{j,*} - (F_*B_*^{-1}\partial_j B_* - C_*A_*^{-1}\partial_j A_*)]]'(\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1});$$

$$(iii)\ \ \widehat{\eta}_n - \widehat{\eta}_{o,n} = n^{-1/2}k^{-1}\mathrm{tr}[V_*] - (nk)^{-1}\mathrm{tr}[F_*V_*] - n^{-1/2}k^{-1}\mathrm{tr}[K_{o,n}V_*] + 2(n^{1/2}k^2)^{-1}\mathrm{tr}[V_*]\mathrm{tr}[K_{o,n}] + (nk^2)^{-1}\mathrm{tr}[V_*]^2 - n^{-1/2}k^{-1}\mathrm{tr}[F_*W_{o,n} - \theta_*U_{o,n}] + n^{-1/2}k^{-1}[\mathrm{tr}[Q_{j,*} - (F_*B_*^{-1}\partial_j B_* - C_*A_*^{-1}\partial_j A_*)]]'(\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1});$$

$$(iv)\ \ \widehat{\sigma}_n - \widehat{\sigma}_{o,n} = (2k)^{-1}\mathrm{tr}[(K_{o,n} + n^{-1/2}V_*)^2] - (2k^2)^{-1}\mathrm{tr}[K_{o,n} + n^{-1/2}V_*]^2 + (2k^2)^{-1}\mathrm{tr}[K_{a,n}]^2 - (2k)^{-1}\mathrm{tr}[K_{o,n}^2] + o_{\mathbb{P}}(n^{-1});$$

$$(v)\ \ \widehat{\xi}_n - \widehat{\xi}_{o,n} = (nk)^{-1}\mathrm{tr}[V_*^2] - 2(n^{1/2}k^2)^{-1}\mathrm{tr}[V_*]\mathrm{tr}[K_{o,n}] + 2(n^{1/2}k)^{-1}\mathrm{tr}[K_{o,n}V_*] - (nk^2)^{-1}\mathrm{tr}[V_*]^2 + o_{\mathbb{P}}(n^{-1});$$

and

$$(vi)\ \ \widehat{\gamma}_n - \widehat{\gamma}_{o,n} = (2nk)^{-1}\mathrm{tr}[V_*^2] - (n^{1/2}k^2)^{-1}\mathrm{tr}[V_*]\mathrm{tr}[K_{o,n}] - (2nk^2)^{-1}\mathrm{tr}[V_*]^2 + (n^{1/2}k)^{-1}\mathrm{tr}[K_{o,n}V_*] + o_{\mathbb{P}}(n^{-1}). \qquad \square$$

Several remarks are warranted. First, note that if $\bar{A}_* = \bar{B}_* \equiv 0$, $F_* = C_* = V_* = 0$ and for each

$j = 1, 2, \ldots, \ell$, $Q_j = 0$, so that all the leading terms in Lemma 5(*i*–*vi*) are zero matrices. This implies that $\widehat{\tau}_n - \widehat{\tau}_{o,n} = o_{\mathbb{P}}(n^{-1})$, $\widehat{\delta}_n - \widehat{\delta}_{o,n} = o_{\mathbb{P}}(n^{-1})$, $\widehat{\eta}_n - \widehat{\eta}_{o,n} = o_{\mathbb{P}}(n^{-1})$, $\widehat{\sigma}_n - \widehat{\sigma}_{o,n} = o_{\mathbb{P}}(n^{-1})$, $\widehat{\xi}_n - \widehat{\xi}_{o,n} = o_{\mathbb{P}}(n^{-1})$, and $\widehat{\gamma}_n - \widehat{\gamma}_{o,n} = o_{\mathbb{P}}(n^{-1})$. Therefore, $\widehat{\tau}_{o,n}$, $\widehat{\delta}_{o,n}$, $\widehat{\eta}_{o,n}$, $\widehat{\sigma}_{o,n}$, $\widehat{\xi}_{o,n}$, and $\widehat{\gamma}_{o,n}$ are the second-order approximations under the condition that $\bar{A}_* = \bar{B}_* \equiv 0$. This is the desired feature of these definitions. Second, if $\bar{A}_* = 0$, Lemmas 5(*i*, *ii*, and *iv*) reduce to lemma 5 of CW. Thus, Lemma 5 generalizes those earlier results. Third, using these differences, the asymptotic approximations of the tests can also be derived, as they depend on the second-order approximations of the test base elements under $\mathcal{H}_\ell$. We provide them in the following theorem.

Theorem 3. *Given Assumption B and $\mathcal{H}_\ell$,*
(i) $\widehat{\mathfrak{B}}_n^{(1)} = \frac{1}{2k}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + o_{\mathbb{P}}(1)$*, and for* $i = 2, 3$, $\widehat{\mathfrak{B}}_n^{(i)} = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$*;*
(ii) $\widehat{\mathfrak{D}}_n^{(1)} = \frac{1}{2k}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + o_{\mathbb{P}}(1)$*, and for* $i = 2, 3$, $\widehat{\mathfrak{D}}_n^{(i)} = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$*; and*
(iii) $\widehat{\mathfrak{S}}_n^{(1)} = \frac{1}{2k}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + o_{\mathbb{P}}(1)$*, and for* $i = 2, 3$, $\widehat{\mathfrak{S}}_n^{(i)} = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$. □

Therefore, the asymptotic approximations of the tests are obtained by shifting the location parameter of $\mathrm{tr}[K_{o,n}]$ by $n^{-1/2}\mathrm{tr}[V_*]$, from which the local power of the tests is derived. We also note that if $\mathrm{tr}[V_*] = 0$, $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ do not have local power different from size. Similarly, if $\mathrm{tr}[V_*^2] = 0$, then $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{B}}_n^{(3)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, and $\widehat{\mathfrak{S}}_n^{(3)}$ have location parameters that are the same as those under the null hypothesis, and their local powers are correspondingly affected. Thus, $\mathrm{tr}[V_*] \neq 0$ and $\mathrm{tr}[V_*^2] \neq 0$ are necessary for these tests to have non-trivial local powers, respectively. We also note that even when $A_{n,*} = A_* + n^{-1/2}\bar{A}_* + o(n^{-1/2})$ or $B_{n,*} = B_* + n^{-1/2}\bar{B}_* + o(n^{-1/2})$, the results stated below still hold. For brevity, we omitted the $o(n^{-1/2})$ remainders from $A_{n,*}$ and $B_{n,*}$ in the local alternative hypothesis.

Theorem 3 implies several properties for the additional test statistics, and they extend the implications obtained under $\mathcal{H}_0$. First, we note that if the roles of A_* and B_* are reversed, the positions for $\bar{B}_*$, B_n, and $B_{*,n}$ in the definitions of V_* and $K_{o,n}$ are interchanged with those of $\bar{A}_*$, A_n, $A_{*,n}$, so that the score of the test statistics defined by $\widetilde{D}_n$ is $-(V_* + \sqrt{n}K_{o,n})$ under $\mathcal{H}_\ell$, implying that $\widetilde{\mathfrak{B}}_n^{(1)}$, $\widetilde{\mathfrak{D}}_n^{(1)}$, and $\widetilde{\mathfrak{S}}_n^{(1)}$ are approximated as $\frac{1}{2k}\mathrm{tr}[-(V_* + \sqrt{n}K_{o,n})]^2 + o_{\mathbb{P}}(1)$, whereas $\widetilde{\mathfrak{B}}_n^{(i)}$, $\widetilde{\mathfrak{D}}_n^{(i)}$, and $\widetilde{\mathfrak{S}}_n^{(i)}$ ($i = 2, 3$) are approximated as $\frac{1}{2}\mathrm{tr}[(-V_* - \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$ under $\mathcal{H}_\ell$. The test statistics defined by $\widehat{D}_n$ are, therefore, equivalent to those defined by $\widetilde{D}_n$. Second, Theorem 3 and (2) imply that the asymptotic approximations of $\widehat{\mathfrak{E}}_n^{(1)}$, $\widehat{\mathfrak{E}}_n^{(2)}$, and $\widehat{\mathfrak{E}}_n^{(3)}$ are obtained as $\frac{1}{2}\mathrm{tr}[(-V_* - \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$ under $\mathcal{H}_\ell$. Therefore, $\widehat{\mathfrak{E}}_n^{(1)}$, $\widehat{\mathfrak{E}}_n^{(2)}$, and $\widehat{\mathfrak{E}}_n^{(3)}$ are asymptotically equivalent to $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{B}}_n^{(3)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, and $\widehat{\mathfrak{S}}_n^{(3)}$ under $\mathcal{H}_\ell$. Finally, the first and second implications imply that $\widetilde{\mathfrak{E}}_n^{(1)}$, $\widetilde{\mathfrak{E}}_n^{(2)}$, and $\widetilde{\mathfrak{E}}_n^{(3)}$ are also asymptotically equivalent tests for the same reason. These three properties exactly extend the implications of Theorem 1. That is, if $V_* = 0$, the implications given below Theorem 1 are obtained. We collect these results in the following corollary.

Corollary 5. *Given Assumption B and $\mathcal{H}_\ell$,*

(i) $\widetilde{\mathfrak{B}}_n^{(1)} = \frac{1}{2k}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + o_{\mathbb{P}}(1)$, $\widetilde{\mathfrak{D}}_n^{(1)} = \frac{1}{2k}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + o_{\mathbb{P}}(1)$, *and* $\widetilde{\mathfrak{S}}_n^{(1)} = \frac{1}{2k}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + o_{\mathbb{P}}(1)$*;*

(ii) $\widetilde{\mathfrak{B}}_n^{(i)} = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$, $\widetilde{\mathfrak{D}}_n^{(i)} = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$, *and* $\widetilde{\mathfrak{S}}_n^{(i)} = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$ $(i = 2, 3)$*; and*

(iii) $\widehat{\mathfrak{E}}_n^{(i)} = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$ *and* $\widetilde{\mathfrak{E}}_n^{(i)} = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$ $(i = 1, 2, 3)$. □

Before moving to the next section, some additional remarks can be usefully included on these results. First, Mauchly (1940), Muirhead (1982), and Anderson (2003) test the sphericity condition: for some d_*, $B_* = d_* A_*$, and theorem 7 of CW implies that the likelihood ratio (LR) test that is obtained under the same distributional condition as above is locally equivalent to the difference between $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ with $i = 2, 3$ and those with $i = 1$. We examine specific examples of testing the sphericity condition in the GMM estimation context in Section 4.1.2. Second, the local asymptotic approximations are equivalent to that of the likelihood ratio test under certain conditions. Nagao (1967), Nagarsenker and Pillai (1973), Muirhead (1982), and Anderson (2003) examine the LR statistic that tests the equality of a covariance matrix to a certain matrix. In the next subsection, we extend this to the structural vector autoregressive (SVAR) model context and show that the tests are locally optimal.

3.5 Local Optimality of the Test Statistics

In this subsection, we examine the LR test statistic using the full-information maximum likelihood (FIML) estimation in the SVAR model context that tests over-identification and serves as a prototypical example of testing two equal symmetric positive-definite matrices. The main motivation of this examination is in the fact that the LR test statistic turns out locally most powerful test for many outstanding examples by the fact that the FIML estimation is obtained using a correctly specified distributional model assumption, so that we can compare the local power properties of the tests in the current study with that of the LR test statistic.

Specifically, we suppose the AB-SVAR model in Amisano and Giannini (1997) that is a synthetic generalization of popular SVAR models: for $k \times k$ invertible matrices $H_{n,*}$ and $N_{n,*}$,

$$H_{n,*}\phi_*(L)Y_{n,t} = H_{n,*}U_{n,t}, \qquad H_{n,*}U_{n,t} = N_{n,*}W_{n,t}$$

such that $E[W_{n,t}] = 0$ and $E[W_{n,t}W_{n,t}'] = I$, where $U_{n,t} \sim$ IID $N(0, B_{n,*})$, $\phi_*(L) := I - \phi_{1,*}L - \phi_{2,*}L^2 - \ldots - \phi_{p,*}L^p$, and L is the lag operator. The structural parameter $H_{n,*}$ and $N_{n,*}$ are estimated by maximizing the following log-likelihood function:

$$\mathcal{L}_n(A) := -\frac{nk}{2}\log(2\pi) - \frac{n}{2}\log(\det(A)) - \frac{n}{2}\mathrm{tr}(A^{-1}\widehat{B}_n)$$

with respect to H and N such that $A = H^{-1}NN'H'^{-1}$, where $\widehat{B}_n := \frac{1}{n}\sum_{t=1}^{n}\widehat{U}_n\widehat{U}_t'$, $\widehat{U}_t := \widehat{\phi}_n(L)Y_t$, and $\widehat{\phi}_n(L)$ is the LS estimator obtained by regressing Y_t against $(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p})$. Note that $\mathcal{L}_n(\cdot)$ is under-identified: there are $2k^2$ unknowns in $H_{n,*}$ and $N_{n,*}$, whereas there are $k(k+1)/2$ first-order equations, so that there are $2k^2 - k(k+1)/2$ free parameters. Thus, the structural parameters $H_{n,*}$ and $N_{n,*}$ are estimated by re-parameterizing them as function values of another parameter ψ that satisfies order and rank conditions (*e.g.*, Sargan, 1988), so that $(H_{n,*}, N_{n,*})$ can be represented as $(H_n(\psi_*), N_n(\psi_*))$ for some $H_n(\cdot)$ and $N_n(\cdot)$, and ψ_* is instead estimated by maximizing the log-likelihood function with respect to ψ. We denote the SVAR estimator obtained in this way as $(\widehat{H}_n, \widehat{N}_n) := (H_n(\widehat{\psi}_n), N_n(\widehat{\psi}_n))$ and $\widehat{A}_n := \widehat{H}_n^{-1}\widehat{N}_n\widehat{N}_n'\widehat{H}_n'^{-1}$ to estimate $A_{n,*} := H_{n,*}^{-1}N_{n,*}N_{n,*}'H_{n,*}'^{-1}$, where $\widehat{\psi}_n$ is the argument maximizing the log-likelihood function with respect to ψ.

Model over-identification is often tested using the LR test statistic. If the model is exactly identified, the log-likelihood is

$$\mathcal{L}_n(\widehat{B}_n) = -\frac{nk}{2}\log(2\pi) - \frac{n}{2}\log(\det(\widehat{B}_n)) - \frac{nk}{2},$$

so that the LR test statistic for over-identification is obtained as

$$\mathfrak{LR}_n := 2(\mathcal{L}_n(\widehat{B}_n) - \mathcal{L}_n(\widehat{A}_n)) = nk(\widehat{\tau}_n - \widehat{\lambda}_n),$$

where $\widehat{\lambda}_n = \frac{1}{k}\log(\det[\widehat{D}_n])$. That is, the LR test statistic measures the distance between $\widehat{\tau}_n$ and $\widehat{\lambda}_n$ that is a function of $\det(\widehat{D}_n)$, were $\widehat{D}_n = \widehat{B}_n\widehat{A}_n^{-1}$ as before, and it statistically tests the identity between $A_{n,*}$ and $B_{n,*}$.

We now provide the following set of assumptions to formalize our conditions and manage our discussions.

Assumption C (SVAR)**.** *(i) For every $n \in \mathbb{N}$, $\phi_*(L)Y_{n,t} = U_{n,t} \sim$ IID $N(0, B_{n,*})$ such that the roots of $\det(\phi_*(L)) = 0$ lie outside of the unit circle, and $B_{n,*} = B_* + n^{-1/2}\bar{B}_*$, where B_* and $\bar{B}_*$ are symmetric positive-definite matrices in $\mathbb{R}^{k\times k}$;*
(ii) For every $n \in \mathbb{N}$, $H_{n,*}$ and $N_{n,*}$ are invertible such that $A_{n,*} := H_{n,*}^{-1}N_{n,*}N_{n,*}'H_{n,*}'^{-1} = A_* + n^{-1/2}\bar{A}_*$ with $\bar{A}_* \neq \bar{B}_*$, where A_* and $\bar{A}_*$ are symmetric and positive definite;
(iii) $A : \Theta \mapsto \mathbb{R}^{k\times k}$ is in $\mathcal{C}^{(2)}(\Theta)$ and $A_* := A(\theta_*)$, where $\Theta(\in \mathbb{R}^{\ell})$ is a compact and convex parameter

space of $\theta := (\psi', \mathrm{vec}(\phi_1, \dots, \phi_p)')'$ *that contains* $\theta_* := (\psi_*', \mathrm{vec}(\phi_{1*}, \dots, \phi_{p*})')'$ *as an interior element;*

(iv) $\bar{A} : \Theta \mapsto \mathbb{R}^{k\times k}$ *and* $\bar{B} : \Theta \mapsto \mathbb{R}^{k\times k}$ *are in* $\mathcal{C}^{(1)}(\Theta)$, $\bar{A}_* := \bar{A}(\theta_*)$, *and* $\bar{B}_* := \bar{B}(\theta_*)$;

(v) $H_n : \Theta \mapsto \mathbb{R}^{k\times k}$ *and* $N_n : \Theta \mapsto \mathbb{R}^{k\times k}$ *are in* $\mathcal{C}^{(2)}(\Theta)$, $A_n(\cdot) := H_n(\cdot)^{-1} N_n(\cdot) N_n(\cdot)' H_n(\cdot)'^{-1}$ *is consistent for* $A(\cdot)$ *uniformly on* Θ, $H_{n,*} = H_n(\theta_*)$, *and* $N_{n,*} = N_n(\theta_*)$;

(vi) $\sqrt{n}[(\widehat{\theta}_n - \theta_*)', \mathrm{vech}[A_n - A_{*,n}]']' = O_{\mathbb{P}}(1)$;

(vii) for $j = 1, \dots, \ell$, $\partial_j A_n(\cdot)$ *is consistent for* $\partial_j A(\cdot)$ *uniformly on* Θ*; and*

(viii) for $j = 1, \dots, \ell$, $H_{j,o,n} := A_*^{-1}\partial_j(A_n - A_{*,n}) = O_{\mathbb{P}}(n^{-1/2})$. □

A number of relevant remarks are warranted on Assumption C. First, Assumption C is obtained from Assumption B by accommodating the SVAR features, so that the definitions of $\widehat{\phi}_n(L)$, $B_{n,*}$, and $\widehat{B}_n$ are given by this, and the conditions in Assumption B for $\widehat{\phi}_n(L)$, $B_{n,*}$, and $\widehat{B}_n$ are easily affirmed by their definitions. On the other hand, $A_{n,*}$ and $\widehat{A}_n$ cannot be uniformly defined. It is differently defined, depending on the order and rank conditions. Due to this, Assumption C focuses on the regularity conditions for $\widehat{A}_n(\cdot)$ and $A_{n,*}$. Second, we suppose that the roots of $\det(\phi_*(L)) = 0$ lie outside the unit circle for the stationarity of Y_t. Third, even when $A_{n,*} = A_* + n^{-1/2}\bar{A}_* + o(n^{-1/2})$ or $B_{n,*} = B_* + n^{-1/2}\bar{B}_* + o(n^{-1/2})$, the results stated below still hold. For brevity, we omitted the $o(n^{-1/2})$ remainders from $A_{n,*}$ and $B_{n,*}$. Fourth, we do not highlight the roles of order and rank conditions given by $H_n(\cdot)$ and $N_n(\cdot)$ as our interests are more given to testing for over-identification using $A_{n,*}$. We instead directly impose the regularity conditions for $A(\cdot)$ and $\bar{A}(\cdot)$ in Assumption C(*iii* and *iv*). Fifth, although $A(\cdot)$ and $\bar{A}(\cdot)$ are functions of only ψ, we treat them as functions of θ to comply with the theory in the previous section. By the same reason, $\bar{B}(\cdot)$ is also treated as a function of θ, although it is a function of only $\mathrm{vec}(\phi_1, \dots, \phi_p)$. Sixth, most SVAR models suppose locally identified SVAR model, and linear re-parameterizations of ψ with deterministic coefficients are assumed for $H_n(\cdot)$ and $N_n(\cdot)$. For such a case, Assumptions C(*v* and *vi*) trivially hold, although $H_n(\cdot)$ and $N_n(\cdot)$ do not have to be necessarily linear transformations of ψ. Finally, Assumption C is obtained by modifying Assumption B to fit the features of the SVAR model, so that the consequences in Theorem 3 and Corollary 5 are also valid for $\widehat{A}_n$ and $\widehat{B}_n$ of this subsection.

We first examine the local limit approximation of the ingredient of the LR test in the following lemma.

Lemma 6. *Given Assumption C and* $\mathcal{H}_\ell$,

$$\begin{aligned}\widehat{\lambda}_n - \widehat{\delta}_{o,n} = \frac{1}{\sqrt{n}k}\mathrm{tr}[V_*] - \frac{1}{\sqrt{n}k}\mathrm{tr}[F_* W_{o,n} - C_* U_{o,n}] + \frac{1}{2nk}(\mathrm{tr}[C_*^2] - \mathrm{tr}[F_*^2]) - \frac{1}{2k^2}\mathrm{tr}[K_{o,n}]^2 \\ + \frac{1}{\sqrt{n}k}[\mathrm{tr}[Q_{j,*} - (F_* B_*^{-1}\partial_j B_* - C_* A_*^{-1}\partial_j A_*)]]'(\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}). \qquad \square\end{aligned}$$

Note that the leading term of $\widehat{\lambda}_n$ is $\widehat{\delta}_{o,n}$, so that $\widehat{\delta}_n$ and $\widehat{\lambda}_n$ posse a similar asymptotic behavior, although their difference is not exactly zero:

$$\widehat{\delta}_n - \widehat{\lambda}_n = \frac{1}{2k^2}\mathrm{tr}[K_{o,n} + n^{-1/2}V_*]^2 + o_{\mathbb{P}}(n^{-1}).$$

Therefore, $\widehat{\delta}_n$ is asymptotically always greater than $\widehat{\lambda}_n$, and their difference is $O_{\mathbb{P}}(n^{-1})$. The following theorems derives the local limit approximation of the LR test statistic.

Theorem 4. *Given Assumptions C and $\mathcal{H}_\ell$, $\mathfrak{LR}_n = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$.* □

Note that the LR test statistic is asymptotically equivalent to the consequences in Theorem 3 and Corollary 5, implying that $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, $\widehat{\mathfrak{S}}_n^{(i)}$, $\widetilde{\mathfrak{B}}_n^{(i)}$, $\widetilde{\mathfrak{D}}_n^{(i)}$, $\widetilde{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$, and $\widehat{\mathfrak{E}}_n^{(i)}$, $\widetilde{\mathfrak{E}}_n^{(i)}$ $(i = 1, 2, 3)$ are asymptotically equivalent to the LR test statistic under the local alternative. Therefore, if the LR test statistic turns out locally optimal as is for the standard case that assumes a correct distributional condition along with minor regularity conditions, the test statistics of this study must be also locally optimal. In another way, the local optimality of testing over-identification is warranted if the trace and determinant functions are selected for the Schur-convex function in the SVAR context.

4 Monte Carlo Experiments

This section reports Monte Carlo experiments examining the performance of the tests analyzed in the previous section. We first consider applications of our procedures to information matrix testing as in CW and additionally apply residual bootstrapping to the linear LS and TSLS estimations (*e.g.* Efron and Tibshirani, 1993). Residual bootstrapping can also be applied to other GMM estimators.

4.1 Testing Information Matrix Equality

4.1.1 Linear Regression

We examine the finite sample properties of the tests by estimating the unknown parameters that are present in linear Gaussian regression models. Specifically, we start by assuming the model $y_t = X_t{}'\beta + u_t$ with $u_t|X_t \sim$ IID $N(0, \sigma^2)$, $X_t = (1, x_t)'$, and where the unknown parameters β and σ^2 are estimated by ML. We consider three DGP classes. First, we let

- Null: $y_t = x_t + u_t$, $u_t|X_t \sim$ IID $N(0, 1)$, and $x_t \sim$ IID $N(0, 1)$.

This DGP is correctly specified by the model, and the information matrix equality holds in ML estimation. Second, we consider three different DGPs for examining power properties. These are

- ALT1: $y_t = x_t + u_t$, $u_t|x_t \sim$ IND $N(0, \exp(x_t))$, and $x_t \sim$ IID $N(1, 1)$;
- ALT2: $y_t = x_t + u_t$, $u_t|x_t \sim$ IND *MN*$(-1, 1; 1, 1; 1/[1 + \exp(x_t)])$, and $x_t \sim$ IID $N(0, 1)$; and
- ALT3: $y_t = x_t + \frac{1}{2}x_t^2 + u_t$, $u_t|x_t \sim$ IID $N(0, 1)$, and $x_t \sim$ IID $N(0, 1)$.

Here, $Z \sim$ *MN*$(a, b; c, d; p)$ denotes a finite mixture of normal distributions: $Z \sim N(a, b)$ with probability p, and $Z \sim N(c, d)$ with probability $1 - p$. The first alternative exhibits conditional heteroskedasticity. Although the conditional mean is correctly specified by the model, the error distribution is misspecified by the presence of the conditional heteroskedasticity. The next alternative has a PDF with two peaks and dispersed distributions, and the final DGP has misspecification in the conditional mean, and this affects the asymptotic distribution of the ML estimator. Under these three DGPs, the model is misspecified, so that the information matrix equality does not hold. We use these alternatives for the powers of the tests. Third, we consider another three DGPs for examining the local power properties. These are

- LOC1: $y_t = x_t + u_t$, $u_t|x_t \sim$ IND $N(0, \exp(2n^{-1/2}x_t))$, and $x_t \sim$ IID $N(1, 1)$;
- LOC2: $y_t = x_t + u_t$, $u_t|x_t \sim$ IND *MN*$(0, 1; 1, 1; 1/[1 + n^{-1/2}\exp(x_t)])$, and $x_t \sim$ IID $N(0, 1)$; and
- LOC3: $y_t = x_t + 5n^{-1/2}x_t^2 + u_t$, $u_t|x_t \sim$ IID $N(0, 1)$, and $x_t \sim$ IID $N(0, 1)$.

These DGPs are obtained by modifying the DGPs in the second group. Note that as the sample size tends to infinity, they approach the first DGP at the rate $n^{-1/2}$. If the sample size is finite and small, they are also similarly distributed to the DGPs in the second group.

There is a caveat for the local DGPs. The distribution of x_t in LOC1 is different from the others. The non-zero mean condition of x_t is required for satisfying Assumption B. If $\mathbb{E}[x_t] = 0$ or $\mathbb{E}[x_t^3] = 0$ as in the other DGPs, Assumption B(v) does not hold, and it approaches the first DGP at the rate $n^{-1/4}$. Although its local power is not negligible, the theory in the previous section is not applicable for this case. Further higher-order approximations are required for the local asymptotic approximations of the tests. We thus let $x_t \sim N(1, 1)$ for ALT1 and LOC1 DGPs so that $\mathbb{E}[x_t] \neq 0$ and $\mathbb{E}[x_t^3] \neq 0$, thereby ensuring the relevance of the theory in the previous section.

Testing is implemented by the following two-step approach. First, we work on the given model assumptions and estimate $\widehat{D}_n := \widehat{B}_n\widehat{A}_n^{-1}$ by letting $\widehat{A}_n$ be the consistent negative Hessian matrix and $\widehat{B}_n$ the covariance matrix of the scores, viz.,

$$\widehat{A}_n := \frac{1}{n}\sum_{t=1}^{n}\begin{bmatrix} \frac{1}{\widehat{\sigma}_n^2}X_tX_t' & \frac{1}{\widehat{\sigma}_n^4}\widehat{u}_tX_t \\ \frac{1}{\widehat{\sigma}_n^4}\widehat{u}_tX_t' & \frac{1}{2\widehat{\sigma}_n^6}\widehat{u}_n^2 \end{bmatrix} \text{ and } \widehat{B}_n := \frac{1}{n}\sum_{t=1}^{n}\begin{bmatrix} \frac{1}{\widehat{\sigma}_n^4}\widehat{u}_t^2X_tX_t' & \frac{1}{2\widehat{\sigma}_n^6}(\widehat{u}_t-\widehat{\sigma}_n^2)\widehat{u}_tX_t \\ \frac{1}{2\widehat{\sigma}_n^6}(\widehat{u}_t-\widehat{\sigma}_n^2)\widehat{u}_tX_t' & \frac{1}{4\widehat{\sigma}_n^8}(\widehat{u}_t^2-\widehat{\sigma}_n^2)^2 \end{bmatrix},$$

where $\widehat{u}_t := y_t - X_t'\widehat{\beta}_n$, and $(\widehat{\beta}_n', \widehat{\sigma}_n^2)'$ is the ML estimator. These matrix estimators are used to construct the test statistics. Next, the parametric bootstrap is applied to our tests. As Horowitz (1994) points out, the parametric bootstrap helps to prevent the level distortion problem that can be enormous for testing the information matrix equality (*e.g.*, Taylor, 1987; Orme, 1990; Chesher and Spady, 1991; White, 1982). For applying the parametric bootstrap, n independent samples are drawn from $N(0, \widehat{\sigma}_n^2)$, and we let the samples be $\{u_t^b : t = 1, 2, \ldots, n\}$ and also $y_t^b := X_t'\widehat{\beta}_n + u_t^b$. Using $\{(X_t, y_t^b) : t = 1, 2, \ldots, n\}$, we compute the test statistics and iterate this process many times to obtain the critical values of the test statistics. CW provide a more detailed procedure for the parametric bootstrap.[1]

In addition to the tests of the current study, we extend the analysis to include other test statistics. Note that the basis for the 24 test statistics ($\widehat{\mathfrak{B}}_n^{(j)}$, $\widehat{\mathfrak{D}}_n^{(j)}$, $\widehat{\mathfrak{S}}_n^{(j)}$, $\widehat{\mathfrak{E}}_n^{(j)}$, $\widetilde{\mathfrak{B}}_n^{(j)}$, $\widetilde{\mathfrak{D}}_n^{(j)}$, $\widetilde{\mathfrak{S}}_n^{(j)}$, $\widetilde{\mathfrak{E}}_n^{(j)}$ with $j = 1, 2, 3$) can be thought of as a squared Gaussian stochastic process defined on a finite dimensional index set. Therefore, their extremum can also be used as a test statistic that is least favorable to the null hypothesis, as has been typically exploited in the context of the tests using infinite dimensional Gaussian stochastic processes (*e.g.*, Cho and White, 2011; Baek, Cho, and Phillips, 2015). We therefore consider the following tests:

$$\widehat{\mathfrak{M}}_n^{(1)} := \max[\widehat{\mathfrak{B}}_n^{(1)}, \widehat{\mathfrak{D}}_n^{(1)}, \widehat{\mathfrak{S}}_n^{(1)}, \widehat{\mathfrak{E}}_n^{(1)}, \widetilde{\mathfrak{B}}_n^{(1)}, \widetilde{\mathfrak{D}}_n^{(1)}, \widetilde{\mathfrak{S}}_n^{(1)}, \widetilde{\mathfrak{E}}_n^{(1)}],$$

$$\widehat{\mathfrak{M}}_n^{(2)} := \max_{j\in\{2,3\}}[\widehat{\mathfrak{B}}_n^{(j)}, \widehat{\mathfrak{D}}_n^{(j)}, \widehat{\mathfrak{S}}_n^{(j)}, \widehat{\mathfrak{E}}_n^{(j)}, \widetilde{\mathfrak{B}}_n^{(j)}, \widetilde{\mathfrak{D}}_n^{(j)}, \widetilde{\mathfrak{S}}_n^{(j)}, \widetilde{\mathfrak{E}}_n^{(j)}].$$

These two tests are separately examined as they have different asymptotic null and local approximations as given in Theorems 1, 3, and the remarks following them. We also let $\widehat{\mathfrak{M}}_n^{(3)} := \max[\widehat{\mathfrak{M}}_n^{(1)}, \widehat{\mathfrak{M}}_n^{(2)}]$. For our final comparisons, we consider two other popular statistics for testing the information matrix equality. Jarque and Bera's (1987) test is used for testing the normality assumption. We denote this as $\widehat{\mathfrak{J}}_n$ and use it to test the error normality assumption. Chesher's (1983) and Lancaster's (1984) information matrix test is also considered and is denoted $\widehat{\mathfrak{I}}_n$. We also apply the parametric bootstrap to $\widehat{\mathfrak{J}}_n$ and $\widehat{\mathfrak{I}}_n$.

The null simulation results are contained in Table 2. The bootstrap repetition is 500, and we report

[1] The following URL provides the GAUSS program codes for testing the information matrix equality of the linear normal, linear exponential, linear Weibull, linear probit, linear logit, linear gompit, linear scobit, and linear tobit models: http://web.yonsei.ac.kr/jinseocho/research.htm.

the empirical rejection rates in Table 2 by repeating independent experiments 5,000 times. The level of significance is 5%. All test statistics including $\widehat{\mathfrak{J}}_n$ and $\widehat{\mathfrak{I}}_n$ have empirical rejection rates very close to the nominal level. This result is not limited to the large sample size case alone. Even when the sample size is as small as 50, the empirical rejection rates are close to the nominal level. This aspect implies that the researcher can control the type I error without considering limitations on the sample size.

The power simulation results are contained in Tables 3 to 5. The bootstrap repetition number is 500, and we replicate independent experiments 2,000 times. The empirical rejection rates are contained in the tables using a significance level of 5%.

We summarize the power simulation results as follows.

1. All tests are consistent. As n tends to infinity, the empirical rejection rates approach unity.

2. The tests of this study are more powerful than $\widehat{\mathfrak{J}}_n$ or $\widehat{\mathfrak{I}}_n$. Nevertheless, it does not imply that they are always more powerful than $\widehat{\mathfrak{J}}_n$ or $\widehat{\mathfrak{I}}_n$. It is generally hard to say in advance which test is most powerful. But the simulations show that the tests of this study are more often powerful than $\widehat{\mathfrak{J}}_n$ or $\widehat{\mathfrak{I}}_n$ for the alternatives considered here.

3. When comparing only our tests, we note that $\widehat{\mathfrak{E}}_n$-indexed tests overall show relatively higher powers than the others when n is large. For ALT1 and ALT3, the $\widehat{\mathfrak{E}}_n$-indexed tests show higher or at least equivalent power than the others. On the other hand, the $\widehat{\mathfrak{E}}_n$-indexed tests are not the most powerful for ALT2 when n is small. As n increases, however, the power of the $\widehat{\mathfrak{E}}_n$-indexed tests approaches that of the most powerful test.

4. Although not the most powerful, the $\widehat{\mathfrak{M}}_n$-indexed tests also possess respectable powers. In particular, the power of $\widehat{\mathfrak{M}}_n^{(3)}$ is always determined by $\widehat{\mathfrak{M}}_n^{(2)}$ and $\widehat{\mathfrak{M}}_n^{(2)}$ always dominates $\widehat{\mathfrak{M}}_n^{(1)}$.

5. Restricting attention to the first nine tests, $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$ become more powerful than $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$, as n increases. To illustrate, we focus on the $\widehat{\mathfrak{B}}_n$-indexed tests in Table 4. When n is small, the most powerful test among the three tests is $\widehat{\mathfrak{B}}_n^{(1)}$. As n increases, however, $\widehat{\mathfrak{B}}_n^{(3)}$ becomes most powerful, and $\widehat{\mathfrak{B}}_n^{(2)}$ becomes more powerful than $\widehat{\mathfrak{B}}_n^{(1)}$. For $n \geq 1,200$, this power ordering is maintained. This tendency occurs not only for the $\widehat{\mathfrak{B}}_n$-indexed tests. The tests indexed by $\widehat{\mathfrak{D}}_n$ and $\widehat{\mathfrak{S}}_n$ in Tables 3, 4, and 5 exhibit similar tendencies. These findings corroborate Theorem 2(*i*). The leading terms of $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ cannot be greater than those of the other tests. Accordingly, they are unlikely to be the most powerful test when n is moderately large. Although the most powerful

test in finite samples is also determined by factors other than the leading terms, Tables 3 to 5 show the general tendency for $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$ to be more powerful than $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$.

The local power simulation results are contained in Tables 6 to 8. The bootstrap repetition number is 500, and 3,000 independent replications were conducted. As before, the level of significance is 5%.

We summarize the local power simulation results as follows.

1. For every local DGP, all tests converge to stable empirical rejection rates, as n increases. The limits of the empirical rejection rates are between 5% and 100%. This aspect corroborates the convergence rate $n^{-1/2}$ as the determining rate for stable local distributions of the tests.

2. Overall, the test statistics of this study exhibit higher local powers than the Jarque-Bera (1987) test and Chesher's (1983) and Lancaster's (1984) information matrix test; and the locally most powerful test is one of $\widehat{\mathfrak{E}}_n$-indexed tests. Nevertheless, these results do not imply that our tests are necessarily more powerful. Earlier simulation findings that are not reported here show that the local power of the Jarque-Bera (1987) test is higher than the others for mixture of normals or errors that follows t-distributions. The Jarque-Bera (1987) test is designed to test for Pearson family distributions such as the t-distribution, and mixtures of normal distributions are better approximated by Pearson family distributions than those with (normal) misspecification. Accordingly, the local powers of $\widehat{\mathfrak{J}}_n$ can be higher.

3. As for the fixed alternative case, the $\widehat{\mathfrak{M}}_n$-indexed tests also possess respectable local power, and $\widehat{\mathfrak{M}}_n^{(2)}$ always dominates $\widehat{\mathfrak{M}}_n^{(1)}$.

4. When comparing the first nine test statistics only, we note similar local power patterns among the tests. For every DGP, $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ (resp. $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ with $i = 2, 3$) have similar local powers among themselves. To illustrate, consider Table 6. Note that the empirical rejection rates of $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ converge to certain numbers around 5%, whereas those of $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{B}}_n^{(3)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, and $\widehat{\mathfrak{S}}_n^{(3)}$ converge to around 11.50%. This observation is not limited to Table 6 but applies also to Tables 7–8. Theorem 3 is therefore supported by this simulation finding. Note that the local approximations of $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ are equivalent, and so are those of $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{B}}_n^{(3)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, and $\widehat{\mathfrak{S}}_n^{(3)}$.

5. Again, when restricting attention to the first nine test statistics, the local power patterns can be different from the fixed alternative power patterns. $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ can be locally more powerful than

$\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$. For instance, observe Table 8 and note that the empirical rejection rates of $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ are generally higher than those of $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{B}}_n^{(3)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, and $\widehat{\mathfrak{S}}_n^{(3)}$.

In addition to the reported test statistics, the simulation results obtained by reversing the roles of A_* and B_* are more or less similar to the reported simulation results. For brevity, we do not report them here.

4.1.2 Probit

For the next experiment we use a probit specification and conduct the same simulations as in the previous subsection. The probit model specifies the conditional mean of a limited dependent variable y_t as $\mathbb{E}[y_t|X_t] = \Phi(X_t{}'\beta)$, where Φ is the standard normal CDF and $X_t := (1, x_t)'$.

We examine the following three DGP classes for our experiments. First,

- Null: $\mathbb{E}[y_t|x_t] = \text{Probit}(1 + x_t)$ and $x_t \sim$ IID $N(0, 1)$,

where 'Probit(x)' means $\Phi(x)$. The model is correctly specified for this DGP, and we use this DGP to examine the asymptotic null behavior of the tests. Second, we examine the following DGPs for the power properties of the tests:

- ALT1: $\mathbb{E}[y_t|x_t] = \text{Probit}(1 + x_t - x_t^4)$ and $x_t \sim$ IID $N(0, 1)$; and
- ALT2: $\mathbb{E}[y_t|x_t] = \text{Logit}[-\pi^2(1 + x_t)/6]$ and $x_t \sim$ IID $N(0, 1)$,

where 'Logit(x)' denotes $\{1 + \exp(-x)\}^{-1}$. ALT1 contains a nonlinear component $-x_t^4$, so that the linear probit model is misspecified. ALT2 is a linear logit process, and although no nonlinear component is involved, the linear probit model is functionally misspecified. Third, the local power properties of the tests are examined by means of the following DGPs:

- LOC1: $\mathbb{E}[y_t|x_t] = \text{Probit}(1 + x_t - n^{-1/2}x_t^4)$ and $x_t \sim$ IID $N(0, 1)$; and
- LOC2: $\mathbb{E}[y_t|x_t] = (1 - n^{-1/2})\text{Probit}(1 + x_t) + n^{-1/2}\text{Logit}[-\pi^2(1 + x_t)/6]$ and $x_t \sim$ IID $N(0, 1)$.

LOC1 and LOC2 are considered as local processes of ALT1 and ALT2, respectively. We also apply the same parametric bootstrap as before and compare Chesher's (1983) and Lancaster's (1984) information matrix test to our tests.

The null simulation results are contained in Table 9. The bootstrap repetitions are 500, and we report the empirical rejection rates in Table 9 using 5,000 replications, as before. The nominal level of significance is

5%. From these simulation results, the empirical nominal levels of all tests are very accurate, just as for the normal linear model case.

The power simulation results are given in Tables 10 and 11. The bootstrap repetitions were 500 and we used 2,000 replications, as before. The same nominal 5% significance level was used. We summarize the power simulation results as follows.

1. All tests including the information matrix test are consistent, and the overall performance of our tests is better than the information matrix test, although this outcome does not imply it is necessarily so.

2. As before, the most powerful test statistic is one of the $\widehat{\mathfrak{E}}_n$-indexed tests; and the $\widehat{\mathfrak{M}}_n$-indexed tests also possess respectable powers. Further, $\widehat{\mathfrak{M}}_n^{(2)}$ always dominates $\widehat{\mathfrak{M}}_n^{(1)}$.

3. The rejection rates of the tests are DGP-dependent. For ALT1, the empirical rejection rates of all the tests approach unity very quickly as n increases. On the other hand, the convergence rates are very slow for ALT2. This is due to the fact that the logit and probit probability functions are very similar to each other. Unless n is very large, it is hard to distinguish them.

4. When restricting to the first nine test statistics, it is hard to corroborate Theorem 2(*i*) using ALT1 as the tests approach unity very quickly. On the other hand, ALT2 shows that the rejection rates of $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$ are higher than $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$. These results affirm Theorem 2(*i*), just as for the linear normal model case.

The local power simulation results are contained in Tables 12 and 13. The bootstrap repetitions are 500, and 3,000 replications were used, as before. Again, the nominal significance level is 5%. We summarize the local power simulation results as follows.

1. Overall, just as before, one of the $\widehat{\mathfrak{E}}_n$-indexed tests is most powerful; and, again, the $\widehat{\mathfrak{M}}_n$-indexed tests also possess respectable local powers.

2. When restricting to the first nine test statistics and the information matrix test, there is a different power relationship among the tests. For LOC1, the locally most powerful tests are $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$, and the next is the information matrix test. $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ have the lowest local power. For LOC2, the locally most powerful tests are $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$, overall. Next are $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$, and the information matrix test has the lowest local power.

3. The simulations affirm Theorem 3. For both LOC1 and LOC2, $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ ($i = 2, 3$) show more or less similar empirical rejection rates, and $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ also show similar rejection rates. This outcome is predicted by Theorem 3.

In sum, using simulations based upon the parametric bootstrap, the results generally affirm the theory provided in the previous section. On the whole, we find that the performances of the $\widehat{\mathfrak{E}}$-indexed tests show greater discriminatory power in revealing (local) alternative hypotheses than the other tests.

4.2 Testing Optimal GMM Weight Matrix

In this subsection, we consider applying the residual bootstrap to GMM estimation for testing the equality of two symmetric positive-definite matrices. Suppose that a researcher wishes to estimate the unknown parameter using an asymptotically optimal weight matrix in order to ensure that the GMM estimator is asymptotically efficient. We therefore propose a test to determine whether the weight matrix selected by the researcher is asymptotically optimal. If the selected weight matrix is optimal, the sandwich asymptotic covariance matrix of the GMM estimator simplifies, and we can test this feature directly using the testing methodology of the present study. Specifically, when $\widehat{\theta}_n$ is the GMM estimator obtained by minimizing $g_n(\cdot)'W_n g_n(\cdot)$ under the standard model and the DGP assumption (*e.g.*, Hamilton, 1994, chap. 14), it is well known that $\sqrt{n}(\widehat{\theta}_n - \theta_*) \overset{\text{A}}{\sim} N[0, (G'WG)^{-1}(G'WSWG)(G'WG)^{-1}]$, where θ_* is the probability limit of the GMM estimator, W is the probability limit of W_n, S is the asymptotic covariance matrix of $\sqrt{n} g_n(\theta_*)$, and G is the probability limit of $\nabla_\theta g_n(\theta_*)$. The researcher wishes to select W_n to estimate S^{-1} asymptotically, so that $\sqrt{n}(\widehat{\theta}_n - \theta_*) \overset{\text{A}}{\sim} N[0, (G'S^{-1}G)^{-1}]$.

In this general set up, we further suppose that the researcher assumes conditional homoskedasticity and proceed with simulations on this basis: viz., for some $\sigma_*^2 > 0$ and positive-definite matrix Q, $S = \sigma_*^2 Q$. The conditional homoskedasticity assumption is violated if the error distribution is not conditionally homoskedastic, the model assumption is incorrect, or the error is serially correlated. We therefore test the proportionality condition between the two matrices and use the testing outcome as a diagnostic for conditional heteroskedasticity, model misspecification, or autocorrelation. This hypothesis is a specific example of the sphericity condition that was examined by Mauchly (1940).

4.2.1 Ordinary Least Squares Estimation

As the first example of GMM estimation, we estimate the unknown parameter by LS. Specifically, when a linear model is given as $y_t = {X_t}'\beta + u_t$, the linear coefficient and residuals are estimated by $\widehat{\beta}_n :=$

$(X'X)^{-1}X'Y$ and $\widehat{u}_t := y_t - X_t'\widehat{\beta}_n$, respectively. The conditionally homoskedastic variance σ_*^2 is estimated by $\widehat{\sigma}_n^2 := n^{-1}\sum_{t=1}^n \widehat{u}_t^2$. With these components in hand, we apply the following residual bootstrap technique.

1. Step 1: Let $\widehat{A}_n := \widehat{\sigma}_n^2(n^{-1}\sum X_t X_t')$ and $\widehat{B}_n$ be a heteroskedasticity autocorrelation consistent (HAC) covariance estimator of $n^{-1/2}\sum u_t X_t$. Using $\widehat{A}_n$ and $\widehat{B}_n$, we compute the test statistics.

2. Step 2: We randomly draw $\{\widehat{u}_t^b : t = 1, 2, \ldots, n\}$ from $\{\widehat{u}_t : t = 1, 2, \ldots, n\}$ with replacement and let $\widehat{B}_n^b$ be the corresponding HAC estimator (constructed in the same fashion as $\widehat{B}_n$) for the asymptotic covariance of $n^{-1/2}\sum \widehat{u}_t^b X_t$. Using $\widehat{A}_n$ and $\widehat{B}_n^b$, we compute the test statistics.

3. Step 3: Replicate Step 2 many times and compute the percentage of bootstrapped test statistics greater than the tests. If this percentage is less than the significance level, we reject the null.

Note that if the independent draws of Step 2 $\{\widehat{u}_t^b\}$ are conditionally homoskedastic given $\{X_1, \ldots, X_n\}$, the sphericity condition holds between the probability limits $\widehat{A}_n$ and $\widehat{B}_n^b$, and the null distribution of the test statistic is accordingly obtained. For our Monte Carlo simulations, the commonly used Newey and West (1987) HAC estimator is used in $\widehat{B}_n$ and $\widehat{B}_n^b$.

Our Monte Carlo experiments are conducted by generating null, local, and alternative DGPs. The following is the null DGP.

- Null: $y_t = \frac{1}{2}x_t + u_t$ and $(x_t, u_t)' \sim$ IID $N(0, I_2)$.

Note that the classical linear model condition is exactly satisfied. The conditional mean equation is correctly specified by the linear model, the error is conditionally homoskedastic, and the error term is serially uncorrelated. Therefore, the asymptotic limits of $\widehat{A}_n$ and $\widehat{B}_n$ are identical. Next, we consider three alternative DGPs.

- ALT1: $y_t = \frac{1}{2}x_t + (1 + \exp(x_t))u_t$ and $(x_t, u_t)' \sim$ IID $N(0, I_2)$;
- ALT2: $y_t = x_t^4 + u_t$ and $(x_t, u_t)' \sim$ IID $N(0, I_2)$; and
- ALT3: $y_t = \frac{1}{2}x_t + u_t$, $u_t = \frac{1}{2}u_{t-1} + \varepsilon_t$, and $(x_t, \varepsilon_t)' \sim$ IID $N(0, I_2)$.

Note that ALT1, ALT2, and ALT3 possess conditional heteroskedasticity, model misspecification, and serially correlated errors, respectively, so that $\widehat{A}_n$ and $\widehat{B}_n$ have different asymptotic limits. Finally, we modify the given alternative DGPs into the following local alternative DGPs.

- LOC1: $y_t = \frac{1}{2}x_t + (1 + n^{-1/2}\exp(x_t))u_t$ and $(x_t, u_t)' \sim$ IID $N(0, I_2)$;

- LOC2: $y_t = \frac{1}{2}n^{-1/2}x_t^4 + u_t$ and $(x_t, u_t)' \sim$ IID $N(0, I_2)$; and
- LOC3: $y_t = \frac{1}{2}x_t + u_t$, $u_t = \frac{1}{2}n^{-1/2}u_{t-1} + \varepsilon_t$, and $(x_t, \varepsilon_t)' \sim$ IID $N(0, I_2)$.

We report the simulation findings in Tables 14 to 20 and summarize them as follows. First, the null simulation results are given in Table 14. As n increases, the empirical rejection rates converge to the nominal level. If n is small, say 50, the empirical rejection rates differ from those using the parametric bootstrap. Overall, the rejection rates are conservative and undervalued the nominal level for the first nine test statistics. On the other hand, the $\widehat{\mathfrak{M}}_n$-indexed tests are liberal and overvalue the nominal values. This distortionary behavior disappears quickly as n increases to 100.

The power and local power simulation results are reported in Tables 15 to 17 and Tables 18 to 20, respectively. We summarize the power simulation results as follows.

1. As n increases, the empirical rejection rates converge to unity, implying that all tests of this study are consistent against conditional heteroskedasticity, model misspecification, or serially correlated error distribution.

2. For every alternative DGP, the most powerful test belongs to the test statistics indexed by $\widehat{\mathfrak{E}}_n$. This dominating feature is the same as that found earlier.

3. The $\widehat{\mathfrak{M}}_n$-indexed tests also possess respectable power except against model misspecification. As before, the power of $\widehat{\mathfrak{M}}_n^{(3)}$ is determined by that of $\widehat{\mathfrak{M}}_n^{(2)}$.

4. When restricting to the first nine test statistics, $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$ each shows higher power than $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$. This result affirms Theorem 2(*i*).

These simulation findings show that the $\widehat{\mathfrak{E}}_n$-indexed tests generally have better performance than the other tests, and we therefore recommend these tests for practical applications. The local power simulation results are similarly confirmatory and these are reported in Tables 18 to 20, which we summarize as follows.

1. As n increases, the empirical rejection rates converge to levels greater than 5%. The test statistics possess nontrivial local powers.

2. As before, the $\widehat{\mathfrak{E}}_n$-indexed test statistics are locally more powerful than other test statistics, overall, and the $\widehat{\mathfrak{M}}_n$-indexed tests also possess respectable local power.

3. When restricting to the first nine test statistics, $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$ exhibit higher powers than $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$. This also confirms the power relationship among the test statistics.

4. $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$ possess similar local powers. On the other hand, $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ appear to have different local powers. For example, for LOC1, the local powers of $\widehat{\mathfrak{B}}_n^{(1)}$ and $\widehat{\mathfrak{D}}_n^{(1)}$ are approximately similar, but the power of $\widehat{\mathfrak{S}}_n^{(1)}$ is quite different. This result differs from that obtained with the parametric bootstrap.

These simulations confirm theory, showing the tests of this study to be consistent and to have nontrivial local discriminatory power against conditional heteroskedasticity, model misspecification, and autocorrelation.

4.2.2 Two-Stage Least Squares Estimation (TSLS)

As a second example of GMM estimation, we consider TSLS estimation. Specifically, for the linear model $y_t = X_t{}'\beta + u_t$, where $X_t := (1, x_t)'$, we obtain the TSLS estimator $\widetilde{\beta}_n := (X'PX)^{-1}(X'PY)$ and residual $\widetilde{u}_t := y_t - X_t{}'\widetilde{\beta}_n$, where $Z_t := (1, z_t)'$ is the instrument variable and $P := Z(Z'Z)^{-1}Z'$. We also estimate the conditionally homoskedastic variance σ_*^2 by $\widetilde{\sigma}_n^2 := n^{-1}\sum_{t=1}^n \widetilde{u}_t^2$ and apply the residual bootstrap in Section 4.2.1. The only difference from the LS case is that $\widehat{A}_n$ is obtained as $\widetilde{\sigma}_n^2(n^{-1}\sum_{t=1}^n Z_t Z_t{}')$, and $\widehat{B}_n$ and $\widehat{B}_n^b$ estimate the asymptotic covariance matrices of $n^{-1/2}\sum_{t=1}^n u_t Z_t$ and $n^{-1/2}\sum_{t=1}^n \widetilde{u}_t^b Z_t$, respectively. The other aspects are the same.

As before, our Monte Carlo experiments are conducted under null, local, and alternative DGPs. The following is the null DGP.

- Null: $y_t = \frac{1}{2}x_t + u_t$, $x_t := u_t + z_t$, and $(z_t, u_t)' \sim$ IID $N(0, I_2)$.

Note that the TSLS estimator is consistent for the parameter vector $(0, \frac{1}{2})'$ because the model is correctly specified. Furthermore, the error is conditionally homoskedastic and serially uncorrelated. Therefore, the asymptotic limits of $\widehat{A}_n$ and $\widehat{B}_n$ are identical. Next, we consider three alternative DGPs.

- ALT1: $y_t = \frac{1}{2}x_t + (1 + \exp(x_t))u_t$, $x_t := u_t + z_t$, and $(z_t, u_t)' \sim$ IID $N(0, I_2)$;
- ALT2: $y_t = 1 + \frac{1}{2}x_t^4 + u_t$, $x_t := u_t + z_t$, and $(z_t, u_t)' \sim$ IID $N(0, I_2)$; and
- ALT3: $y_t = \frac{1}{2} + u_t$, $x_t := z_t + \varepsilon_t$, $u_t := \frac{1}{2}u_{t-1} + \frac{1}{2}u_{t-2} + \varepsilon_t + \frac{1}{2}\varepsilon_{t-1}$, and $(z_t, \varepsilon_t)' \sim$ IID $N(0, I_2)$.

Note that ALT1, ALT2, and ALT3 possess conditional heteroskedasticity, model misspecification, and serially correlated error distribution, respectively, so that estimated $\widehat{A}_n$ and $\widehat{B}_n$ have different asymptotic limits. Finally, we modify the given alternative DGPs into the following local alternative DGPs.

- LOC1: $y_t = \frac{1}{2}x_t + (1 + n^{-1/2}\exp(x_t))u_t$, $x_t := u_t + z_t$, and $(z_t, u_t)' \sim$ IID $N(0, I_2)$;

• LOC2: $y_t = 1 + \frac{1}{2}n^{-1/2}x_t^4 + u_t$, $x_t := u_t + z_t$, and $(z_t, u_t)' \sim$ IID $N(0, I_2)$; and

• LOC3: $y_t = \frac{1}{2} + u_t$, $x_t := z_t + \varepsilon_t$, $u_t := \rho_n(u_{t-1} + u_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$, $\rho_n := \frac{1}{2}n^{-1/2}$, and $(z_t, \varepsilon_t)' \sim$ IID $N(0, I_2)$.

The simulation results are given in Tables 21 to 27 and we summarize as follows. First, the null simulation results are reported in Table 21. As n increases, the empirical rejection rates converge to the nominal level. Just as for the LS estimation case, if n is small, say 50, the empirical rejection rates undervalue the nominal level for the first nine test statistics, and the $\widehat{\mathfrak{M}}_n$-indexed tests overvalue the nominal level. These discrepancies quickly disappear as n increases, but $\widehat{\mathfrak{D}}_n^{(2)}$ needs more observations to reduce size distortion than the others.

The power and local power simulation results are reported in Tables 22 to 24 and Tables 25 to 27, respectively. We summarize the power results as follows.

1. As before, when n increases, the empirical rejection rates converge to unity.

2. For every alternative DGP, the most powerful test is one of the $\widehat{\mathfrak{E}}_n$-indexed test statistics.

3. The $\widehat{\mathfrak{M}}_n$-indexed tests also possess respectable power except for model misspecification, just as for the LS estimation case. As before, the empirical power of $\widehat{\mathfrak{M}}_n^{(1)}$ is always dominated by $\widehat{\mathfrak{M}}_n^{(2)}$.

4. When restricting to the first nine test statistics, the powers of $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$ are generally higher than those of $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$. This result also affirms Theorem 2(*i*).

Note that these results are identical to those obtained from the LS case. The local power simulation results are summarized as follows.

1. As n increases, the empirical rejection rates converge to levels greater than 5%. This shows that the tests have nontrivial local powers.

2. As before, the $\widehat{\mathfrak{E}}_n$-indexed test statistics are overall more powerful than other test statistics, and the $\widehat{\mathfrak{M}}_n$-indexed tests also possess respectable powers.

3. When restricting to the first nine test statistics, $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$ show higher powers than $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$.

4. As before, $\widehat{\mathfrak{B}}_n^{(i)}$, $\widehat{\mathfrak{D}}_n^{(i)}$, and $\widehat{\mathfrak{S}}_n^{(i)}$ $(i = 2, 3)$ exhibit approximately similar local powers. On the other hand, $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ appear to have quite different local powers.

All these results mirror those found in the LS estimation case.

5 Empirical Applications

In the political economy and political science literature, a longstanding research question is to explain voting turnout. For example, in the seminal study of Wolfinger and Rosenstone (1980) on voting behavior, the authors estimated a voting turnout model using 1972 presidential election data. In later work in economics, Feddersen and Pesendorfer (1996) provided an economic model for voting turnout based on the asymmetric information of voters. In addition to these contributions, many research papers have attempted to explain voting turnout using empirical analysis and economic theory (*e.g.*, Nagler, 1991, 1994; Bénabou 2000; Besley and Case, 2003; Berry, DeMeritt, and Esarey, 2010, among others).

Wolfinger and Rosenstone's (1980) empirical model has been particularly influential in this literature. Using the level of education, (*Education*), the squared level of education (*Education*^{2}), age (*Age*), squared age (*Age*^{2}), a dummy for the South (*South*), a dummy for the presence of a gubernatorial election in the state (*Gubernatorial Election*), and the number of days before the election that registration closes (*Closing Date*), they estimate a linear probit model and find that it is the registration requirement in voting law that most severely affects the least educated group. Here, *Closing date* is used as a measure for the voting law requirement. Specifically, their model estimation shows that if *Closing Date* were hypothetically set to zero, the average voting turnout shows the greatest increase for the least educated group, whereas the increase is least for the most educated group turnout. Wolfinger and Rosenstone (1980) explain this in terms of the cost of voting: more educated people pay a lower cost for understanding the implications of complex and abstract political issues. This finding is now regarded as a stylized fact in the political economic and science literature.

Nagler (1991) points out that Wolfinger and Rosenstone's (1980) central empirical result is an artifact of the probit model methodology. The probit probability is most highly affected if the explanatory variable is around zero. In other words, the first-order derivative is greater at this level than at any other level of the explanatory variable. Therefore, if *Education* is near zero, the predicted probability level increase is greater than any group with higher education. So, the empirical finding of Wolfinger and Rosenstone (1980) result is simply an artifact of how the probit model manages the impact of data rather than a meaningful empirical finding. The same claim holds for a logit specification. To remedy this problem, Nagler (1991) estimates another probit model with two additional explanatory variables: *Closing Date* $\times$ *Education* and *Closing Date* $\times$ *Education*^{2}. He includes them to capture the interactive effects of *Closing Date* and *Education* to the turnout. From estimation of this model using 1972 and 1984 presidential election data, Nagler (1991) rejects Wolfinger and Rosenstone's (1980) empirical result.

Nagler (1994) attempts to improve these findings. Instead of specifying a probit model, he specifies a scobit model that assumes

$$\mathbb{P}(y_t = 1|X_t) = 1 - \frac{1}{(1 + \exp({X_t}'\beta_*))^{\alpha_*}},$$

where y_t is a dummy for voting, and X_t is a vector of explanatory variables. This specification follows a Burr type 10 distribution, for which the logit distribution is a special case obtained by setting $\alpha_* = 1$. If $\alpha_* \neq 1$, the distribution is skewed and so the model is also called the skewed-logit model. Using estimates of this model, Nagler (1991) modifies the earlier claim and finds that, under the skewed probability model, the interactive terms are not significant for the 1984 presidential election data and confirms Wolfinger and Rosenstone's (1980) earlier finding that the least educated group **is** more severely affected by the voting law requirement. He also argues that the scobit model is particularly useful in allowing for misspecification in logit or probit models.

Nagler's (1994) empirical conclusion may still be misleading and for the same reason. As Nagler (1991) points out, the probit model advocated by Wolfinger and Rosenstone (1980) can yield biased empirical findings because it is misspecified. In the same way, if the scobit model is misspecified, a similar critique applies. Although use of the skewness measure (α_*) generalizes the logit model, the model can easily be misspecified by other factors. As Berry, DeMeritt, and Esarey (2010) discuss, among the various features of probit and logit models for modeling voting turnout, use of the correct model assumptions and specification for the data is critical in reaching the conclusion of Nagler (1994).

Against this background, our tests of model specification provide a relevant new methodology to assess whether these empirical models of voting turnout are correctly specified or not. Using 1984 US presidential election data from Altman and McDonald (2003), we estimate the same model considered by Wolfinger and Rosenstone (1980) and Nagler (1991, 1994). The results are given in Table 28. Probit models without and with interactive terms are estimated by following Wolfinger and Rosenstone (1980) and Nagler (1991), respectively. The same model is also estimated by Berry, DeMeritt, and Esarey (2010) using the same data. Logit and scobit models without and with interactive terms are also estimated by following Nagler (1994). All the estimated parameters are similar to those in the literature. The only difference is that the p -values of the t-test statistics are computed by robust standard errors using the method in White (1980). These are provided in parentheses. The same table also provides the test results and their p -values. All the specification test statistics test the validity of the information matrix equality, and these are computed using the methodology of Section 4.1.2.

The findings in Table 28 can be summarized as follows. First, all empirical models for voting turnout appear to be misspecified, even though they have significantly dominated the empirical literature for some time. All the tests $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{B}}_n^{(3)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(1)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, $\widehat{\mathfrak{S}}_n^{(3)}$, $\widehat{\mathfrak{E}}_n^{(1)}$, $\widehat{\mathfrak{E}}_n^{(2)}$, $\widehat{\mathfrak{E}}_n^{(3)}$, $\widehat{\mathfrak{M}}_n^{(1)}$, $\widehat{\mathfrak{M}}_n^{(2)}$, and $\widehat{\mathfrak{M}}_n^{(3)}$ reject the information matrix equality. All of the test p-values are almost identical to zero, which implies that the conditional distributions of voting turnout that are assumed in these models are all misspecified. The scobit models appear to do best and have greater log-likelihood values than the probit and logit models. But the allowance for a skewed distribution is not enough to eliminate model misspecification. Second, the interactive terms in the scobit model are not statistically significant. The p-values of the interactive terms are 0.2283 and 0.6378, and this finding corresponds with Nagler (1994), although correct specification was assumed in reaching that conclusion in Nagler's work. Nevertheless, the outcome might be different if the correct model specification was used for voting turnout. In other words, the empirical findings obtained using the scobit specification may well be as misleading as for those from the probit and logit models.

Inferences drawn from these models in empirical work are inevitably approximate and the quality of the approximation depends on the scope of the model, its capability in testing the validity of a theory, and on the relevance of the model to the data. If the empirical model has a sufficiently flexible form that enables adequate estimation of the core part of the relevant theory, we may be able to exploit the model scope to test the theory within the framework of quasi-maximum likelihood estimation, as pointed out by Berry, DeMeritt, and Esarey (2010). The tests given here help to point to weaknesses in specification that may be repaired by the use of more flexible models with greater scope for empirical relevance.

6 Conclusion

The information matrix equality is a fundamental feature of correct specification in likelihood based econometric work, and the GMM estimator has a simple asymptotic covariance matrix if the models are correct and errors are conditionally homoskedastic and serially uncorrelated. We provide a new methodology for testing such equalities in empirical applications. Our approach is embedded in the general framework of testing the equality of two symmetric positive-definite matrices. The new approach improves earlier analytic attempts to control size in information matrix equality testing and includes testing optimal weight matrix conditions in GMM estimation, delivering a class of test procedures that are easily implemented in practical work. The test mechanism extends earlier test statistics developed in the literature by exploiting a simple characterization of equality between two k dimensional symmetric positive-definite matrices A and B involving only the traces of the two matrices AB^{-1} and BA^{-1}. This characterization leads to a group of new omnibus test statistics for testing the equality of covariance matrices.

Asymptotic theory for these tests under null, local, and alternatives are obtained under mild regularity conditions that support wide use of these procedures in empirical work. Simulation evidence affirms that good size control is obtained and test power in testing specification or optimal GMM estimation against various alternatives is generally strong, but power can be dominated in some cases by specific testing procedures such as specification testing based on direct tests for Gaussianity. The methods of specification testing based on the information matrix equality are well illustrated in the commonly occurring cases of logit and probit models, and testing the optimal GMM estimation is also illustrated using the least squares and two-stage least squares estimation. Empirical application of the logit and probit methods to voting turnout models show that classic models used in this literature all seem to suffer from specification failure, putting some of the empirical conclusions in the literature about voting turnout behavior at risk.

7 Technical Appendix and Proofs

7.1 Preliminary Lemmas

Before proving the main claims in the paper, we provide the following preliminary lemmas.

Lemma A1. *Given Assumption A, if for some* $d_* > 0$, $B_* = d_* A_*$,

$$\nabla_\theta^2 \mathrm{tr}[D_*] + d_*^2 \nabla_\theta^2 \mathrm{tr}[D_*^{-1}] = 2d_*[\mathrm{tr}[R_{j,*}R_{i,*}]]. \qquad \square$$

Lemma A2. *Given Assumption B and* $\mathcal{H}_\ell$,

(i) $A_{*,n}^{-1} = A_*^{-1} - n^{-1/2}C_* A_*^{-1} + n^{-1}C_*^2 A_*^{-1} + O(n^{-3/2})$;

(ii) $B_{*,n}^{-1} = B_*^{-1} - n^{-1/2}F_* B_*^{-1} + n^{-1}F_*^2 B_*^{-1} + O(n^{-3/2})$;

(iii) $U_{a,n} = U_{o,n} - n^{-1/2}C_* U_{o,n} + O_{\mathbb{P}}(n^{-3/2})$;

(iv) $W_{a,n} = W_{o,n} - n^{-1/2}F_* W_{o,n} + O_{\mathbb{P}}(n^{-3/2})$;

(v) $A_{*,n}^{-1}B_{*,n} = I + n^{-1/2}V_* - n^{-1}C_* V_* + O(n^{-3/2})$;

(vi) $B_{*,n}^{-1}A_{*,n} = I - n^{-1/2}V_* + n^{-1}F_* V_* + O(n^{-3/2})$;

(vii) $P_{a,n} = P_{o,n} - n^{-1/2}(F_* W_{o,n} - C_* U_{o,n}) + O_{\mathbb{P}}(n^{-3/2})$;

(viii) $B_{*,n}^{-1}\partial_j B_{*,n} = B_*^{-1}\partial_j B_* + n^{-1/2}(B_*^{-1}\partial_j \bar{B}_* - F_* B_*^{-1}\partial_j B_*) + O(n^{-1})$;

(ix) $A_{*,n}^{-1}\partial_j A_{*,n} = A_*^{-1}\partial_j A_* + n^{-1/2}(A_*^{-1}\partial_j \bar{A}_* - C_* A_*^{-1}\partial_j A_*) + O(n^{-1})$;

(x) $R_{j,a,*,n} = B_*^{-1}\partial_j B_* - A_*^{-1}\partial_j A_* + n^{-1/2}(Q_{j,*} - (F_* B_*^{-1}\partial_j B_* - C_* A_*^{-1}\partial_j A_*)) + O(n^{-1})$, *where* $R_{j,a,*,n} := B_{n,*}^{-1}\partial_j B_{n,*} - A_{n,*}^{-1}\partial_j A_{n,*}$;

(xi) $L_{a,n} = L_{o,n} - n^{-1/2}\{(F_*W_{o,n} - C_*U_{o,n}) - \sum_{j=1}^{\ell}(\widehat{\theta}_{j,n} - \theta_{j,*})(Q_{j,*} - (F_*B_*^{-1}\partial_j B_* - C_*A_*^{-1}\partial_j A_*))\} + O_{\mathbb{P}}(n^{-3/2})$*; and*

(xii) $L_{a,n}A_{*,n}^{-1}B_{*,n} = L_{o,n} - n^{-1/2}(F_*W_{o,n} - C_*U_{o,n}) + n^{-1/2}\sum_{j=1}^{\ell}(\widehat{\theta}_{j,n} - \theta_{j,*})(Q_{j,*} - (F_*B_*^{-1}\partial_j B_* - C_*A_*^{-1}\partial_j A_*)) + n^{-1/2}L_{o,n}(F_* - C_*) - n^{-1/2}(F_*W_{o,n} - C_*U_{o,n}) + O_{\mathbb{P}}(n^{-3/2})$. □

Lemma A3. *Given Assumption B and* $\mathcal{H}_\ell$*,*

(i) $k^{-1}\text{tr}[B_{*,n}^{-1}A_{*,n}] = 1 - n^{-1/2}k^{-1}\text{tr}[V_*] + n^{-1}k^{-1}\text{tr}[F_*V_*] + O(n^{-3/2})$*;*

(ii) $(k^{-1}\text{tr}[B_{*,n}^{-1}A_{*,n}])^2 = 1 - 2n^{-1/2}k^{-1}\text{tr}[V_*] + 2(nk)^{-1}\text{tr}[F_*V_*] + n^{-1}k^{-2}\text{tr}[V_*]^2 + O(n^{-3/2})$*;*

(iii) $(k^{-1}\text{tr}[B_{*,n}^{-1}A_{*,n}])^{-1} = 1 + n^{-1/2}k^{-1}\text{tr}[V_*] - n^{-1}k^{-1}\text{tr}[F_*V_*] + n^{-1}k^{-2}\text{tr}[V_*]^2 + O(n^{-3/2})$*; and*

(iv) $(k^{-1}\text{tr}[B_{*,n}^{-1}A_{*,n}])^{-2} = 1 + 2n^{-1/2}k^{-1}\text{tr}[V_*] - 2(nk)^{-1}\text{tr}[F_*V_*] + 3n^{-1}k^{-2}\text{tr}[V_*]^2 + O(n^{-3/2})$. □

Lemma A4. *Given Assumption B and* $\mathcal{H}_\ell$ *,*

(i) $\det[A_{*,n}] = \det[A_*]\{1 + n^{-1/2}\text{tr}[C_*] + \frac{1}{2n}(\text{tr}[C_*]^2 - \text{tr}[C_*^2])\} + O(n^{-3/2})$*;*

(ii) $\det[B_{*,n}] = \det[B_*]\{1 + n^{-1/2}\text{tr}[F_*] + \frac{1}{2n}(\text{tr}[F_*]^2 - \text{tr}[F_*^2])\} + O(n^{-3/2})$*;*

(iii) $\det[A_{*,n}]^{-1} = \det[A_*]^{-1}\{1 - n^{-1/2}\text{tr}[C_*] + \frac{1}{2n}(\text{tr}[C_*]^2 + \text{tr}[C_*^2])\} + O(n^{-3/2})$*;*

(iv) $\det[D_{*,n}] = 1 + n^{-1/2}\text{tr}[V_*] + \frac{1}{2n}(\text{tr}[V_*]^2 + \text{tr}[C_*^2] - \text{tr}[F_*^2]) + O(n^{-3/2})$*;*

(v) $\det[D_{*,n}]^{1/k} = 1 + \frac{1}{\sqrt{nk}}\text{tr}[V_*] + \frac{1}{2nk}(\text{tr}[C_*^2] - \text{tr}[F_*^2]) + \frac{1}{2nk^2}\text{tr}[V_*]^2 + O(n^{-3/2})$*; and*

(vi) $\det[D_{*,n}]^{1/k}\text{tr}[L_{a,n}] = \frac{1}{k}\text{tr}[K_{o,n}] + \frac{1}{\sqrt{nk^2}}\text{tr}[V_*]\text{tr}[K_{o,n}] - \frac{1}{\sqrt{nk}}\text{tr}[F_*W_{o,n} - C_*U_{o,n}] + \frac{1}{\sqrt{nk}}[\text{tr}[Q_{j,*} - (F_*B_*^{-1}\partial_j B_* - C_*A_*^{-1}\partial_j A_*)]]'(\widehat{\theta}_n - \theta_*) + O(n^{-3/2})$. □

Lemma A5. *Given Assumption B and* $\mathcal{H}_\ell$*,*

(i) $\widehat{\tau}_{o,n} = k^{-1}\text{tr}[K_{o,n}] + O_{\mathbb{P}}(n^{-1})$*;*

(ii) $\widehat{\delta}_{o,n} = k^{-1}\text{tr}[K_{o,n}] + O_{\mathbb{P}}(n^{-1})$*;*

(iii) $\widehat{\eta}_{o,n} = k^{-1}\text{tr}[K_{o,n}] + O_{\mathbb{P}}(n^{-1})$*;*

(iv) $\widehat{\xi}_{o,n} = k^{-1}\text{tr}[K_{o,n}^2] - (k^{-1}\text{tr}[K_{o,n}])^2 + o_{\mathbb{P}}(n^{-1})$*; and*

(v) $\widehat{\gamma}_{o,n} = (2k)^{-1}\text{tr}[K_{o,n}^2] - (2k^2)^{-1}\text{tr}[K_{o,n}]^2 + o_{\mathbb{P}}(n^{-1})$. □

Before proving the preliminary lemmas, we note that $P_{o,n} = M_{o,n}$ and $R_{j,*} = S_{j,*}$ under $\mathcal{H}_\ell$, so that $L_{o,n} = K_{o,n}$.

Proof of Lemma A1: By lemma A5(*i*) of CW and the fact that $A_*^{-1}B_* = d_*I$,

$$\begin{aligned}\partial_{ji}^2\text{tr}[D_*] &= \text{tr}[A_*^{-1}B_*\{(B_*^{-1}\partial_{ji}^2 B_* - A_*^{-1}\partial_{ji}^2 A_*) - (R_{j,*}A_*^{-1}\partial_i A_* + R_{i,*}A_*^{-1}\partial_j A_*)\}] \\ &= d_*\text{tr}[(B_*^{-1}\partial_{ji}^2 B_* - A_*^{-1}\partial_{ji}^2 A_*) - (R_{j,*}A_*^{-1}\partial_i A_* + R_{i,*}A_*^{-1}\partial_j A_*)].\end{aligned}$$

The asymptotic expansion of $\partial^2_{ji}\mathrm{tr}[D_*^{-1}]$ is also obtained by simply interchanging the roles of A_* and B_*:

$$\partial^2_{ji}\mathrm{tr}[D_*^{-1}] = \mathrm{tr}[B_*^{-1}A_*\{(A_*^{-1}\partial^2_{ji}A_* - B_*^{-1}\partial^2_{ji}B_*) + (R_{j,*}B_*^{-1}\partial_i B_* + R_{i,*}B_*^{-1}\partial_j B_*)\}]$$
$$= d_*^{-1}\mathrm{tr}[(A_*^{-1}\partial^2_{ji}A_* - B_*^{-1}\partial^2_{ji}B_*) + (R_{j,*}B_*^{-1}\partial_i B_* + R_{i,*}B_*^{-1}\partial_j B_*)].$$

Therefore, $\partial^2_{ji}\mathrm{tr}[D_*] + d_*^2\partial^2_{ji}\mathrm{tr}[D_*^{-1}] = 2d_*\mathrm{tr}[R_{j,*}R_{i,*}]$ by noting that $R_{i,*} := B_*^{-1}\partial_i B_* - A_*^{-1}\partial_i A_*$ and $R_{j,*} := B_*^{-1}\partial_j B_* - A_*^{-1}\partial_j A_*$. ■

Proof of Lemma A2: (*i*) Note that $A_{*,n}^{-1} = [I - n^{-1/2}A_*^{-1}(-\bar{A}_*)]^{-1}A_*^{-1}$. For large enough n, $[I - n^{-1/2}A_*^{-1}(-\bar{A}_*)]^{-1} = I - n^{-1/2}A_*^{-1}\bar{A}_* + n^{-1}A_*^{-1}\bar{A}_*A_*^{-1}\bar{A}_* + ...$, which implies that

$$A_{*,n}^{-1} = [I - n^{-1/2}A_*^{-1}(-\bar{A}_*)]^{-1}A_*^{-1}$$
$$= A_*^{-1} - n^{-1/2}A_*^{-1}\bar{A}_*A_*^{-1} + n^{-1}A_*^{-1}\bar{A}_*A_*^{-1}\bar{A}_*A_*^{-1} + O(n^{-3/2})$$
$$= A_*^{-1} - n^{-1/2}C_*A_*^{-1} + n^{-1}C_*^2A_*^{-1} + O(n^{-3/2}).$$

(*ii*) This follows from Lemma A2(*i*) and the symmetric structure between $A_{*,n}$ and $B_{*,n}$.

(*iii*) Note that $U_{a,n} = A_{*,n}^{-1}(A_n - A_{*,n}) = A_*^{-1}(A_n - A_{*,n}) - n^{-1/2}C_*A_*^{-1}(A_n - A_{*,n}) + O_{\mathbb{P}}(n^{-3/2})$ by Lemma A2(*i*). Here, the right side is $U_{o,n} - n^{-1/2}C_*U_{o,n} + O_{\mathbb{P}}(n^{-3/2})$ by the definition of $U_{o,n}$.

(*iv*) This follows from Lemma A2(*iii*) and the symmetric structure between $A_{*,n}$ and $B_{*,n}$.

(*v*) Note that

$$A_{*,n}^{-1}B_{*,n} = (A_*^{-1} - n^{-1/2}C_*A_*^{-1} + n^{-1}C_*^2A_*^{-1} + O(n^{-3/2}))(B_* + n^{-1/2}\bar{B}_*)$$
$$= I + n^{-1/2}(A_*^{-1}\bar{B}_* - A_*^{-1}\bar{A}_*) + n^{-1}(C_*^2A_*^{-1}B_* - C_*A_*^{-1}\bar{B}_*) + O(n^{-3/2})$$
$$= I + n^{-1/2}(B_*^{-1}\bar{B}_* - A_*^{-1}\bar{A}_*) - n^{-1}C_*(B_*^{-1}\bar{B}_* - C_*) + O(n^{-3/2})$$

by Lemma A2(*i*) and the definition of $B_{*,n}$. We now note that $V_* := B_*^{-1}\bar{B}_* - A_*^{-1}\bar{A}_* = B_*^{-1}\bar{B}_* - C_*$.

(*vi*) This follows from Lemma A2(*v*) and the symmetric structure between $A_{*,n}$ and $B_{*,n}$.

(*vii*) This follows from the definition of $P_{a,n} := W_{a,n} - U_{a,n}$ and Lemmas A2(*iii* and *iv*).

(*viii*) By Lemma A2(*ii*),

$$B_{*,n}^{-1}\partial_j B_{*,n} = (B_*^{-1} - n^{-1/2}F_*B_*^{-1} + n^{-1}F_*^2B_*^{-1} + O(n^{-3/2}))\partial_j(B_* + n^{-1/2}\bar{B}_*)$$
$$= B_*^{-1}\partial_j B_* + n^{-1/2}(B_*^{-1}\partial_j\bar{B}_* - F_*B_*^{-1}\partial_j B_*) + O(n^{-1}).$$

(*ix*) We can apply the proof of Lemma A2(*viii*).

(*x*) Apply Lemmas A2(*viii* and *ix*) to obtain the desired result.

(*xi*) By Lemmas A2(*vii* and *ix*),

$$L_{a,n} = P_{o,n} + \sum_{j=1}^{\ell} (\widehat{\theta}_{j,n} - \theta_{j,*})(B_*^{-1}\partial_j B_* - A_*^{-1}\partial_j A_*) - n^{-1/2}(F_* W_{o,n} - C_* U_{o,n})$$
$$+ n^{-1/2} \sum_{j=1}^{\ell} (\widehat{\theta}_{j,n} - \theta_{j,*})(Q_{j,*} - (F_* B_*^{-1}\partial_j B_* - C_* A_*^{-1}\partial_j A_*)) + O_{\mathbb{P}}(n^{-3/2}).$$

The desired result follows from the definition of $L_{o,n}$.

(*xii*) We combine Lemmas A2(*v* and *xi*) and collect the terms according to their convergence rates. This completes the proof. ■

Proof of Lemma A3: (*i*) This immediately follows from Lemma A2(*vi*).

(*ii*) This immediately follows from Lemma A2(*vi*).

(*iii*) Taylor expansion of $1/x$ at $x = 1$ gives $1/x = 1 - (x-1) + (x-1)^2 - (x-1)^3 + ...$ We now let x be $k^{-1}\mathrm{tr}[D_{*,n}^{-1}]$ and use Lemma A3(*i*). If the terms are rearranged according to their convergence rates, the desired result follows.

(*iv*) This immediately follows from Lemma A3(*iii*). ■

Proof of Lemma A4: (*i*) By the proof of lemma A2 (*i*) of CW,

$$\det[A_n] - \det[A_*] = \det[A_*]\mathrm{tr}[A_*^{-1}(A_n - A_*)]$$
$$+ \frac{1}{2}\det[A_*]\{\mathrm{tr}[A_*^{-1}(A_n - A_*)]^2 - \mathrm{tr}[A_*^{-1}(A_n - A_*)A_*^{-1}(A_n - A_*)]\} + O_{\mathbb{P}}(n^{-3/2}).$$

We now simply let A_n be $A_{*,n}$ and note that $C_* = A_*^{-1}(A_{*,n} - A_*) = A_*^{-1}\bar{A}_*$ under $\mathcal{H}_\ell$. This yields the desired result.

(*ii*) This immediately follows from Lemma A4(*i*) and the symmetric structure between $A_{*,n}$ and $B_{*,n}$.

(*iii*) Lemma A2(*iii*) of CW shows that $\det[A_n]^{-1} - \det[A_*]^{-1} = -\det[A_*]^{-1}(\mathrm{tr}[U_n] + \frac{1}{2}\mathrm{tr}[U_n]^2 - \frac{1}{2}\mathrm{tr}[U_n^2]) + O_{\mathbb{P}}(n^{-1})$. Under $\mathcal{H}_\ell$, $U_n = C_*$. If we further let their A_n be $A_{*,n}$, then

$$\det[A_{*,n}]^{-1} - \det[A_*]^{-1} = -\det[A_*]^{-1}\{\mathrm{tr}[A_*^{-1}(A_{*,n} - A_*)]$$
$$+ \frac{1}{2}\mathrm{tr}[A_*^{-1}(A_{*,n} - A_*)]^2 - \frac{1}{2}\mathrm{tr}[A_*^{-1}(A_{*,n} - A_*)^2]\} + O_{\mathbb{P}}(n^{-3/2}).$$

The desired result follows by noting that $C_* = A_*^{-1}(A_{*,n} - A_*) = A_*^{-1}\bar{A}_*$.

(*iv*) Note that

$$\det[D_{*,n}] = \det[A_{*,n}]^{-1}\det[B_{*,n}]$$
$$= \left\{1 + \frac{1}{\sqrt{n}}\mathrm{tr}[F_*] + \frac{1}{2n}(\mathrm{tr}[F_*]^2 - \mathrm{tr}[F_*^2])\right\}\left\{1 - \frac{1}{\sqrt{n}}\mathrm{tr}[C_*] + \frac{1}{2n}(\mathrm{tr}[C_*]^2 + \mathrm{tr}[C_*^2])\right\} + O(n^{-3/2}),$$

where the second equality follows from Lemmas A4(*ii* and *iii*) and the fact that $\det[D_*] = 1$ under $\mathcal{H}_\ell$. Thus,

$$\det[D_{*,n}] = 1 + \frac{1}{\sqrt{n}}\mathrm{tr}[F_* - C_*] + \frac{1}{2n}(\mathrm{tr}[F_* - C_*]^2 + \mathrm{tr}[C_*^2] - \mathrm{tr}[F_*^2]) + O(n^{-3/2}).$$

We further note that $V_* := F_* - C_*$ to yield the result.

(*v*) Taylor expansion applied to $\det[D_{*,n}]^{1/k}$ gives

$$\det[D_{*,n}]^{1/k} = \det[D_*]^{1/k} + \frac{1}{k}\det[D_{*,n}]^{1/k-1}\{\det[D_{*,n}] - \det[D_*]\}$$
$$+ \frac{1}{2k}\left(\frac{1}{k} - 1\right)\{\det[D_{*,n}] - \det[D_*]\}^2 + \dots. \quad (5)$$

Lemma A4(*iv*) implies that $\det[D_{*,n}] - \det[D_*] = \frac{1}{\sqrt{n}}\mathrm{tr}[V_*] + \frac{1}{2n}(\mathrm{tr}[V_*]^2 + \mathrm{tr}[C_*^2] - \mathrm{tr}[F_*^2]) + O(n^{-3/2})$ by noting that $\det[D_*] = 1$ under $\mathcal{H}_\ell$. We now substitute this into (5) and arrange the terms according to their convergence rates. This yields the desired result.

(*vi*) To show this, we combine Lemmas A2(*xi*) and A3(*v*) and rearrange the terms according to their convergence rates. This completes the proof. ■

As Lemma A5 is immediately obtained by applying Corollary 1 and Lemma 4(*ii*), we omit the proof.

7.2 Proofs of the Main Results

Proof of Lemma 1: (*i*) If $A = B$, then clearly $\mathrm{tr}[D] = \mathrm{tr}[A^{-1}B] = \mathrm{tr}[I] = k$ and $\mathrm{tr}[D^{-1}] = \mathrm{tr}[B^{-1}A] = \mathrm{tr}[I] = k$. For the converse, note that $k^{-1}\sum_{j=1}^{k}\lambda_j = 1$, where λ_j is the j-th largest eigenvalue of D and so $\mathrm{tr}[D] = k$. In addition, $k^{-1}\mathrm{tr}[D^{-1}] = 1$ implies that $k^{-1}\sum_{j=1}^{k}\lambda_j^{-1} = 1$, so that the harmonic mean of the eigenvalues of D is 1. That is, the arithmetic mean of the eigenvalues is identical to the harmonic mean. Therefore, for some λ, $\lambda = \lambda_1 = \ldots = \lambda_k$. The given condition also implies that $\lambda = 1$. If we now let C be the orthonormal matrix of the eigenvectors of $A^{-1/2}BA^{-1/2}$, $A^{-1/2}BA^{-1/2} = CIC' = I$. Therefore,

$A^{-1/2}BA^{-1/2} = I$ implies $A^{1/2}A^{-1/2}BA^{-1/2}A^{1/2} = A^{1/2}A^{1/2}$, which simplifies to $B = A$.

(*ii*) We can combine Lemma 1(*i*) with lemma 1 of CW. ■

Proofs of Lemma 2 follow from lemma 4 of CW and Lemma 3 in our study. We thus omit its proof. Furthermore, Corollary 1 follows from Lemma 2. We now prove Lemma 3.

Proof of Lemma 3: (*i*) Lemma 4(*i*) of CW gives the expansion of $\mathrm{tr}[\widehat{B}_n\widehat{A}_n^{-1}]$. We apply this expansion to expand $k^{-1}\mathrm{tr}[\widehat{D}_n^{-1}]$ by simply interchanging the roles of A_n and B_n. That is,

$$\begin{aligned}\frac{1}{k}\mathrm{tr}[\widehat{D}_n^{-1}] - \frac{1}{k}\mathrm{tr}[D_*^{-1}] = &-\frac{1}{k}\mathrm{tr}[L_nB_*^{-1}A_*] + \frac{1}{k}\mathrm{tr}[L_nW_nB_*^{-1}A_*]\\ &+ \frac{1}{k}[\mathrm{tr}[(-J_{j,n} + P_nB_*^{-1}\partial_jB_*)B_*^{-1}A_*]]'(\widehat{\theta}_n - \theta_*)\\ &+ \frac{1}{2k}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\mathrm{tr}[D_*^{-1}](\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}). \quad (6)\end{aligned}$$

We also note that by Taylor expansion of $\frac{1}{x}$ yields that $\frac{1}{x} - \frac{1}{x_0} = -\frac{1}{x_0^2}(x - x_0) + \frac{1}{x_0^3}(x - x_0)^2 + \mathcal{R}$, where $\mathcal{R}$ is the remainder. We now let x and x_0 be $\frac{1}{k}\mathrm{tr}[\widehat{D}_n^{-1}]$ and $\frac{1}{k}\mathrm{tr}[D_*^{-1}]$, respectively and also note that $\widehat{\eta}_n = k/\mathrm{tr}[\widehat{D}_n^{-1}] - 1$ and $\eta_* = k/\mathrm{tr}[D_*^{-1}] - 1$. We finally arrange the terms according to their convergence rates to obtain the desired result.

(*ii*) If $\mathcal{H}_0$ further holds, $\eta_* = 0$, $B_*^{-1}A_* = I$, $L_n = K_n$, $k^{-1}\mathrm{tr}[D_*^{-1}] = 1$, and $P_n = M_n$. If all these equalities are applied to (6), the asymptotic expansion of $\widehat{\eta}_n$ reduces to the desired expansion. ■

Proof of Lemma 4: (*i*) Lemmas 4(*i.a* and *i.b*) immediately follow from Lemma 2(*i*, *ii*, and *iii*).

(*ii*) (*ii.a*) From the fact that $B_* = d_*A_*$, it follows that $\mathrm{tr}[D_*^{-1}] = k/d_*$, $D_* = d_*I$, and $D_*^{-1} = d_*^{-1}I$. We now substitute these into $\widehat{\eta}_n$ in Lemma 3 and obtain

$$\begin{aligned}\widehat{\eta}_n = d_* - 1 + d_*k^{-1}\mathrm{tr}[L_n] + d_*(k^{-1}\mathrm{tr}[L_n])^2 - d_*k^{-1}[\mathrm{tr}[(-J_{j,n} + P_nB_*^{-1}\partial_jB_*)]]'(\widehat{\theta}_n - \theta_*)\\ - \frac{d_*}{k}\mathrm{tr}[L_nW_n] - \frac{d_*^2}{2k}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\mathrm{tr}[D_*^{-1}](\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}). \quad (7)\end{aligned}$$

In the same way, we substitute $\mathrm{tr}[D_*^{-1}] = k/d_*$, $D_* = d_*I$, and $D_*^{-1} = d_*^{-1}I$ into (3) and obtain

$$\begin{aligned}\widehat{\tau}_n = d_* - 1 + d_*k^{-1}\mathrm{tr}[L_n] - d_*k^{-1}\mathrm{tr}[L_nU_n]\\ + \frac{d_*}{k}[\mathrm{tr}[J_{j,n} - P_nA_*^{-1}\partial_jA_*]]'(\widehat{\theta}_n - \theta_*) + \frac{1}{2k}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\mathrm{tr}[D_*](\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}). \quad (8)\end{aligned}$$

Therefore, the asymptotic expansion of $\widehat{\xi}_n$ is obtained as

$$\widehat{\xi}_n := \widehat{\tau}_n - \widehat{\eta}_n = d_* k^{-1}\mathrm{tr}[L_n P_n] + d_* k^{-1}[\mathrm{tr}[P_n R_{j,*}]]'(\widehat{\theta}_n - \theta_*) - d_* k^{-2}\mathrm{tr}[L_n]^2 + \frac{1}{2k}(\widehat{\theta}_n - \theta_*)' \left\{\nabla^2_\theta \mathrm{tr}[D_*] + d_*^2 \nabla^2_\theta \mathrm{tr}[D_*^{-1}]\right\}(\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}). \quad (9)$$

Here, the definition of $P_n := W_n - U_n$ is used to simplify the expression. Given this, note that Lemma A1 implies that $\nabla^2_\theta \mathrm{tr}[D_*] + d_*^2 \nabla^2_\theta \mathrm{tr}[D_*^{-1} = 2d_* \mathrm{tr}[R_{j,*} R_{i,*}]$. Therefore,

$$\widehat{\xi}_n = d_* k^{-1}\mathrm{tr}[L_n P_n] + d_* k^{-1}[\mathrm{tr}[P_n R_{j,*}]]'(\widehat{\theta}_n - \theta_*) - d_* k^{-2}\mathrm{tr}[L_n]^2 + d_* k^{-1}(\widehat{\theta}_n - \theta_*)'[\mathrm{tr}[R_{j,*} R_{i,*}]](\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}).$$

We recall the definition of $L_n := P_n + \sum_{j=1}^{\ell}(\widehat{\theta}_{j,n} - \theta_{j,*})R_{j,*}$, and note that this implies

$$\widehat{\xi}_n = d_* k^{-1}\mathrm{tr}[P_n^2] + 2d_* k^{-1}[\mathrm{tr}[P_n R_{j,*}]]'(\widehat{\theta}_n - \theta_*) - d_* k^{-2}\mathrm{tr}[L_n]^2 + d_* k^{-1}(\widehat{\theta}_n - \theta_*)'[\mathrm{tr}[R_{j,*} R_{i,*}]](\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}), \quad (10)$$

that is also equal to $d_* k^{-1}\mathrm{tr}[L_n^2] - d_* k^{-2}\mathrm{tr}[L_n]^2 + o_{\mathbb{P}}(n^{-1})$.

(*ii.b*) Note that corollary 5(*ii*) of CW shows that $\widehat{\sigma}_n = -\frac{d_*}{2k^2}\mathrm{tr}[L_n]^2 + \frac{d_*}{2k}\mathrm{tr}[L_n^2] + o_{\mathbb{P}}(n^{-1})$. Also, $\widehat{\gamma}_n := \widehat{\xi}_n - \widehat{\sigma}_n$. Thus, $\widehat{\gamma}_n = 2^{-1} d_* \{k^{-1}\mathrm{tr}[L_n^2] - (k^{-1}\mathrm{tr}[L_n])^2\} + o_{\mathbb{P}}(n^{-1})$ using (*ii.a*).

(*iii*) Given Lemma 4(*ii*), we let $d_* = 1$ and $L_n = K_n$ to complete the proof. ■

Theorem 1(*i*) follows as a corollary of theorem 1 of CW, and Theorems 1(*ii* and *iii*) also follow as corollaries of Corollary 1 and Lemma 4(*iii*). Corollary 3 is implied by Lemmas 3, 4, and lemma 5 of CW. We now prove Theorem 2.

Proof of Theorem 2: The claim structures given for the statistics in Corollaries 3(*i.a*– *i.i*) and 4(*i.a*– *i.c*) are symmetric and similar. We therefore prove only the claim on $\widehat{\mathfrak{B}}_n^{(1)}$ in (*i*), $\widehat{\mathfrak{B}}_n^{(2)}$ in (*ii*), and $\widehat{\mathfrak{D}}_n^{(2)}$ in (*iv*) to save the space. The others are proved in a similar fashion.

(*i*) For $\widehat{\mathfrak{B}}_n^{(1)}$ to have the greatest leading term, it has to be greater than those $\widehat{\mathfrak{B}}_n^{(2)}$ and $\widehat{\mathfrak{B}}_n^{(3)}$. This implies that $\frac{1}{2}\delta_*^2 \geq \frac{1}{2}\tau_*^2 + 2(\tau_* - \gamma_*)$ and $\frac{1}{2}\tau_*^2 \geq \frac{1}{2}\delta_*^2 + 2(\tau_* - \gamma_*)$. These two inequalities hold only when $\tau_* = \gamma_*$, so that all eigenvalues of D_* are identical. This implies that B_* is proportional to A_* and contradicts the assumption of Theorem 2. This proving methodology also applies to $\widehat{\mathfrak{D}}_n^{(1)}$ and $\widehat{\mathfrak{S}}_n^{(1)}$.

(*ii*) For $\widehat{\mathfrak{B}}_n^{(2)}$ to have the greatest leading term, it has to be greater than those of the other tests. By (*i*), we compare the leading term of $\widehat{\mathfrak{B}}_n^{(2)}$ with those of other test statistics than $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$:

$$\tau_*^2 > \delta_*^2, \tag{11}$$

$$\sigma_* > \gamma_*, \tag{12}$$

$$\tau_*^2 + \sigma_* > \eta_*^2, \tag{13}$$

$$\tau_*^2 + 2\sigma_* > \delta_*^2 + 2\gamma_*, \tag{14}$$

$$\tau_*^2 + 2\sigma_* > \eta_*^2 + 2\gamma_*, \tag{15}$$

$$\sigma_* > \gamma_*, \tag{16}$$

$$\tau_*^2 + \sigma_* > \delta_*^2 + \gamma_* \tag{17}$$

$$\tau_*^2 > \eta_*^2. \tag{18}$$

Each inequality is obtained by letting the leading term of Corollary 3(*i.b*) be greater than the leading terms of Corollaries 3(*i.c*, *i.e*, *i.f*, *i.h*, *i.i*), 4(*i.a*, *i.b*, *i.c*), and the fact that $\xi_* \equiv \sigma_* + \gamma_*$. These 8 inequalities are necessary for the desired condition.

Given this, note that (11), (12), and (18) are the conditions for $\widehat{\mathfrak{B}}_n^{(2)}$ to have the greatest leading term that are given by Theorem 2(*ii*). This proves sufficiency. For necessity, note that (12) is identical to (16); (11) and (12) imply (14) and (17); (12) and (18) imply (13) and (15). The same proving methodology applies to the proof of (*iii*, *vi*, *vii*, *viii*, and *x*).

(*iii*) Given the conditions, we note that μ_* is equal to $\frac{nk}{2}(\tau_*^2 + 2\gamma_*)$ that is achieved by $\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, and $\widehat{\mathfrak{E}}_n^{(1)}$. This proves the sufficiency. For necessity, we compare the leading term of $\widehat{\mathfrak{D}}_n^{(2)}$ with those of other test statistics than $\widehat{\mathfrak{B}}_n^{(1)}$, $\widehat{\mathfrak{D}}_n^{(1)}$, and $\widehat{\mathfrak{S}}_n^{(1)}$ as before:

$$\gamma_* > \sigma_*, \tag{19}$$

$$\tau_*^2 + \gamma_* > \delta_*^2 + \sigma_*, \tag{20}$$

$$\tau_*^2 > \eta_*^2, \tag{21}$$

$$\tau_*^2 + \sigma_* > \delta_*^2 + \gamma_*, \tag{22}$$

$$\tau_*^2 + \sigma_* > \eta_*^2 + \gamma_*, \tag{23}$$

$$\tau_*^2 + \sigma_* > \tau_*^2 + \gamma_*, \tag{24}$$

$$\tau_*^2 > \delta_*^2 \tag{25}$$

$$\tau_*^2 + \gamma_* > \eta_*^2 + \sigma_*. \tag{26}$$

Each inequality is obtained by letting the leading term of Corollary 3(*i.e*) be greater than the leading terms of Corollaries 3(*i.b*, *i.c*, *i.f*, *i.h*, *i.i*), 4(*i.a*, *i.b*, *i.c*). These 8 inequalities are necessary for the desired condition.

Given this, note that (19) and (24) are contradictory, implying that μ_* cannot be uniquely maximized by the leading term of $\widehat{\mathfrak{D}}_n^{(2)}$. We therefore let $\gamma_* = \sigma_*$ and allow for the existence of multiple maximizers. Then, (19)–(26) reduce to the necessary conditions. The same proving methodology applies to the proof of (*iv*, *v*, and *ix*), and this completes the proof. ■

Proof of Lemma 5: (*i*) We apply lemma 4(*i*) of CW and obtain the following expansion for $\widehat{\tau}_n$:

$$\begin{aligned}\widehat{\tau}_n = \tau_{*,n} + \frac{1}{k}\mathrm{tr}[L_{a,n}A_{*,n}^{-1}B_{*,n}] + \frac{1}{k}[\mathrm{tr}[(J_{j,a,n} - P_{a,n}A_*^{-1}\partial_j A_*)A_*^{-1}B_*]]'(\widehat{\theta}_n - \theta_*)\\ - \frac{1}{k}\mathrm{tr}[L_{a,n}U_{a,n}A_{*,n}^{-1}B_{*,n}] + \frac{1}{2k}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\mathrm{tr}[D_*](\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}),\end{aligned}$$

where $\tau_{*,n} := k^{-1}\mathrm{tr}[B_{*,n}A_{*,n}^{-1}] - 1$. We now use Lemma A2(*ii*, *v*, *vii*, and *xii*) for the first three terms and obtain

$$\begin{aligned}\widehat{\tau}_n = \frac{1}{n^{1/2}k}\mathrm{tr}[V_*] - \frac{1}{nk}\mathrm{tr}[C_*V_*] - \frac{1}{n^{1/2}k}\mathrm{tr}[(F_*W_{o,n} - C_*U_{o,n})]\\ + \frac{1}{n^{1/2}k}\mathrm{tr}[L_{o,n}V_*] + \frac{1}{n^{1/2}k}[\mathrm{tr}[(Q_{j,*} - (F_*B_*^{-1}\partial_j B_* - C_*A_*^{-1}\partial_j A_*))]]'(\widehat{\theta}_n - \theta_*)\\ + \frac{1}{k}\mathrm{tr}[L_{o,n}] + \frac{1}{k}[\mathrm{tr}[(J_{j,o,n} - P_{o,n}A_*^{-1}\partial_j A_*)A_*^{-1}B_*]]'(\widehat{\theta}_n - \theta_*)\\ - \frac{1}{k}\mathrm{tr}[L_{o,n}U_{o,n}] + \frac{1}{2k}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\mathrm{tr}[D_*](\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}).\end{aligned} \quad (27)$$

Note that $P_{o,n} = M_{o,n}$, $L_{o,n} = K_{o,n}$ under $\mathcal{H}_\ell$ and also that

$$\widehat{\tau}_{o,n} := \frac{1}{k}\mathrm{tr}[K_{o,n}(I - U_{o,n})] + \frac{1}{k}[\mathrm{tr}[J_{j,o,n} - M_{o,n}A_*^{-1}\partial_j A_*]]'(\widehat{\theta}_n - \theta_*) + \frac{1}{2k}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\mathrm{tr}[D_*](\widehat{\theta}_n - \theta_*).$$

This represents the last second to the last fourth terms in (27). Substituting $\widehat{\tau}_{o,n}$ into these terms completes the proof.

(ii) We apply Lemma 4(ii) of CW and obtain the following expansion for $\widehat{\delta}_n$:

$$\begin{aligned}\widehat{\delta}_n = \delta_{*,n} &+ \frac{1}{k}\det[D_{*,n}]^{-\frac{1}{k}}\mathrm{tr}[L_{a,n}] + \frac{1}{2k}\det[D_{*,n}]^{-\frac{1}{k}}\left\{\left(\frac{1}{k}-1\right)\mathrm{tr}[L_{a,n}]^2 - \mathrm{tr}[W_{a,n}^2]\right\}\\
&+ \frac{1}{k}\det[D_{*,n}]^{-\frac{1}{k}}\left\{\frac{1}{2}(\mathrm{tr}[P_{a,n}]^2 + \mathrm{tr}[U_{a,n}^2]) + \mathrm{tr}[P_{a,n}][\mathrm{tr}[R_{j,a,*,n}]]'(\widehat{\theta}_n - \theta_*)\right\}\\
&+ \frac{1}{k}\det[D_{*,n}]^{-\frac{1}{k}}[\mathrm{tr}[J_{j,o,n} + U_{a,n}A_{*,n}^{-1}\partial_j A_{*,n} - W_{a,n}B_{*,n}^{-1}\partial_j B_{*,n}]]'(\widehat{\theta}_n - \theta_*)\\
&+ \frac{1}{2k}\det[D_{*,n}]^{\frac{1}{k}-1}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\det[D_{*,n}](\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}),\end{aligned}$$

where $\delta_{*,n} := \det[B_{*,n}A_{*,n}^{-1}]^{1/k} - 1$. We note that Lemma A3($v$) implies that $\delta_{*,n} = \frac{1}{\sqrt{n}k}\mathrm{tr}[V_*] + \frac{1}{2nk}(\mathrm{tr}[C_*^2] - \mathrm{tr}[F_*^2]) + \frac{1}{2nk^2}\mathrm{tr}[V_*]^2 + O(n^{-3/2})$, and the asymptotic expansion of $\det[D_{*,n}]^{1/k}\mathrm{tr}[L_{a,n}]$ is given by Lemma A3(vi). If we collect all these terms,

$$\begin{aligned}\widehat{\delta}_n = &\frac{1}{\sqrt{n}k}\mathrm{tr}[V_*] + \frac{1}{2nk}(\mathrm{tr}[C_*^2] - \mathrm{tr}[F_*^2]) + \frac{1}{2nk^2}\mathrm{tr}[V_*]^2 + \frac{1}{k}\mathrm{tr}[K_{o,n}] + \frac{1}{\sqrt{n}k^2}\mathrm{tr}[V_*]\mathrm{tr}[K_{o,n}]\\
&- \frac{1}{\sqrt{n}k}\mathrm{tr}[F_*W_{o,n}] + \frac{1}{\sqrt{n}k}\mathrm{tr}[C_*U_{o,n}] + \frac{1}{2k}(\mathrm{tr}[M_{o,n}]^2 + \mathrm{tr}[U_{o,n}^2]) + \frac{1}{2k}\left(\frac{1}{k}-1\right)\mathrm{tr}[K_{o,n}]^2\\
&- \frac{1}{2k}\mathrm{tr}[W_{o,n}^2] + \frac{1}{\sqrt{n}k}[\mathrm{tr}[Q_{j,*} - (F_*B_*^{-1}\partial_j B_* - C_*A_*^{-1}\partial_j A_*)]]'(\widehat{\theta}_n - \theta_*)\\
&+ \frac{1}{k}[\mathrm{tr}[U_{o,n}A_*^{-1}\partial_j A_* - W_{o,n}B_*^{-1}\partial_j B_*]]'(\widehat{\theta}_n - \theta_*) + \frac{1}{k}\mathrm{tr}[M_{o,n}][\mathrm{tr}[S_{j,*}]]'(\widehat{\theta}_n - \theta_*)\\
&+ \frac{1}{k}[\mathrm{tr}[J_{j,o,n}]]'(\widehat{\theta}_n - \theta_*) + \frac{1}{2k}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\det[D_*](\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}).\end{aligned} \tag{28}$$

This equation is derived by using the fact that $L_{o,n} = K_{o,n}$, $R_{j,*} = S_{j,*}$, and $P_{o,n} = M_{o,n}$ under $\mathcal{H}_\ell$. We now note the definition of $\widehat{\delta}_{o,n}$:

$$\begin{aligned}\widehat{\delta}_{o,n} := &\frac{1}{k}\mathrm{tr}[K_{o,n}] + \frac{1}{2k}\left(\frac{1}{k}-1\right)\mathrm{tr}[K_{o,n}]^2 + \frac{1}{2k}(\mathrm{tr}[M_{o,n}]^2 + \mathrm{tr}[U_{o,n}^2] - \mathrm{tr}[W_{o,n}^2])\\
&+ \frac{1}{k}[\mathrm{tr}[J_{j,o,n} + U_{o,n}A_*^{-1}\partial_j A_* - W_{o,n}A_*^{-1}\partial_j B_*]]'(\widehat{\theta}_n - \theta_*)\\
&+ \frac{1}{k}[\mathrm{tr}[M_{o,n}]\mathrm{tr}[S_{j,*}]]'(\widehat{\theta}_n - \theta_*) + \frac{1}{2k}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\det[D_*](\widehat{\theta}_n - \theta_*).\end{aligned}$$

If the right-side terms of (28) that correspond to the definition of $\widehat{\delta}_{o,n}$ are collected into $\widehat{\delta}_{o,n}$, the desired result follows.

(*iii*) Note that Lemma 3(*i*) is simplified into

$$\widehat{\eta}_n = \eta_{*,n} + k^{-1}\mathrm{tr}[L_{a,n}B_{*,n}^{-1}A_{*,n}]/(k^{-1}\mathrm{tr}[D_{*,n}^{-1}])^2 \\ + (k^{-1}\mathrm{tr}[L_{o,n}])^2 - k^{-1}\mathrm{tr}[L_{o,n}W_{o,n}] - k^{-1}[\mathrm{tr}[(-J_{j,n} + P_{o,n}B_*^{-1}\partial_j B_*)B_*^{-1}A_*]]'(\widehat{\theta}_n - \theta_*) \\ - (2k)^{-1}(\widehat{\theta}_n - \theta_*)'\nabla_\theta^2\mathrm{tr}[D_*^{-1}](\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}) \quad (29)$$

under $\mathcal{H}_\ell$, where $\eta_{*,n} := (k^{-1}\mathrm{tr}[B_{*,n}^{-1}A_{*,n}])^{-1} - 1$. Given this, we further note that

$$k^{-1}\mathrm{tr}[L_{a,n}B_{*,n}^{-1}A_{*,n}]/(k^{-1}\mathrm{tr}[D_{*,n}])^2 = k^{-1}\mathrm{tr}[L_{o,n}] - n^{-1/2}k^{-1}\mathrm{tr}[F_*W_{o,n} - C_*U_{o,n}] \\ + n^{-1/2}k^{-1}\sum_{j=1}^{\ell}(\widehat{\theta}_{j,n} - \theta_{j,*})[Q_{j,*} - (F_*B_*^{-1}\partial_j B_* - C_*A_*^{-1}\partial_j A_*)] \\ - n^{-1/2}k^{-1}\mathrm{tr}[L_{o,n}V_*] + 2n^{-1/2}k^{-2}\mathrm{tr}[V_*]\mathrm{tr}[L_{o,n}] + o_{\mathbb{P}}(n^{-1}) \quad (30)$$

using Lemmas A1(*vi*, *xi*) and A2(*iv*). If we substitute (30) into (29) and use Lemma A1(*vi*), the the desired result is obtained.

(*iv*) We now use Lemmas 5(*i* and *ii*) and compute $\widehat{\sigma}_n$ by its definition. That is,

$$\widehat{\sigma}_n := \widehat{\tau}_n - \widehat{\delta}_n = \widehat{\sigma}_{o,n} + \frac{1}{2k}\left\{\frac{1}{n}\mathrm{tr}[V_*^2] + \frac{2}{\sqrt{n}}\mathrm{tr}[K_{o,n}V_*] + \mathrm{tr}[K_{o,n}^2]\right\} - \frac{1}{2k}\mathrm{tr}[K_{o,n}^2] \\ - \frac{1}{2k^2}\left\{\frac{1}{n}\mathrm{tr}[V_*]^2 + \frac{2}{\sqrt{n}}\mathrm{tr}[V_*]\mathrm{tr}[K_{o,n}] + \mathrm{tr}[K_{o,n}]^2\right\} + \frac{1}{2k^2}\mathrm{tr}[K_{o,n}]^2 + o_{\mathbb{P}}(n^{-1}).$$

Note further that $\frac{1}{n}\mathrm{tr}[V_*^2] + \frac{2}{\sqrt{n}}\mathrm{tr}[K_{o,n}V_*] + \mathrm{tr}[K_{o,n}^2] = \mathrm{tr}[(K_{o,n} + n^{-1/2}V_*)^2]$ and $\frac{1}{n}\mathrm{tr}[V_*]^2 + \frac{2}{\sqrt{n}}\mathrm{tr}[V_*]$ $\mathrm{tr}[K_{o,n}] + \mathrm{tr}[K_{o,n}]^2 = \mathrm{tr}[K_{o,n} + n^{-1/2}V_*]^2$. Using these facts, we obtain that

$$\widehat{\sigma}_n = \widehat{\sigma}_{o,n} + \frac{1}{2k}\{\mathrm{tr}[(K_{o,n} + n^{-1/2}V_*)^2] - \mathrm{tr}[K_{o,n}^2]\} - \frac{1}{2k^2}\{\mathrm{tr}[K_{o,n} + n^{-1/2}V_*]^2 - \mathrm{tr}[K_{o,n}]^2\} + o_{\mathbb{P}}(n^{-1}).$$

This is the desired result.

(*v*) Note that $\widehat{\xi}_n \equiv \widehat{\tau}_n - \widehat{\eta}_n$ and that the asymptotic approximations of $\widehat{\tau}_n$ and $\widehat{\eta}_n$ are provided in Lemmas 5(*i* and *ii*).

$$\widehat{\xi}_n = \widehat{\tau}_{o,n} - \widehat{\eta}_{o,n} + (nk)^{-1}\mathrm{tr}[C_*^2 - 2C_*F_* + F_*^2] + 2n^{-1/2}k^{-1}\mathrm{tr}[K_{o,n}V_*] \\ - n^{-1}k^{-2}\mathrm{tr}[V_*]^2 - 2n^{-1/2}k^{-2}\mathrm{tr}[V_*]\mathrm{tr}[K_{o,n}] + o_{\mathbb{P}}(n^{-1}). \quad (31)$$

Note that $\mathrm{tr}[C_*^2 - 2C_*F_* + F_*^2] = \mathrm{tr}[(F_* - C_*)^2] = \mathrm{tr}[V_*^2]$. The desired result follows from this.

(*vi*) Note that $\widehat{\gamma}_n \equiv \widehat{\xi}_n - \widehat{\sigma}_n$. Furthermore, the asymptotic approximations of $\widehat{\xi}_n$ and $\widehat{\sigma}_n$ are provided in Lemmas 5(*i* and *iii*). From these, it follows that

$$\widehat{\gamma}_n = \widehat{\xi}_{o,n} - \widehat{\sigma}_{o,n} + (2nk)^{-1}\mathrm{tr}[V_*^2] - n^{-1/2}k^{-1}\mathrm{tr}[V_*]\mathrm{tr}[K_{o,n}] \\ - (2nk^2)^{-1}\mathrm{tr}[V_*]^2 + n^{-1/2}k^{-1}\mathrm{tr}[K_{o,n}V_*] + o_{\mathbb{P}}(n^{-1}).$$

We finally note that $\widehat{\xi}_{o,n} - \widehat{\sigma}_{o,n} = \widehat{\gamma}_{o,n}$ to complete the proof. ■

Proof of Theorem 3: (*i*) Note that $\widehat{\mathfrak{B}}_n^{(1)} := \frac{nk}{4}(\widehat{\tau}_n^2 + \widehat{\delta}_n^2)$. Furthermore, the asymptotic approximations of $\widehat{\tau}_n$ and $\widehat{\delta}_n$ are provided in Lemmas 5(*i* and *ii*). Therefore,

$$\begin{aligned}\widehat{\mathfrak{B}}_n^{(1)} &= \frac{nk}{4}\left\{\widehat{\tau}_{o,n}^2 + \widehat{\delta}_{o,n}^2 + \frac{2}{\sqrt{n}k}(\widehat{\tau}_{o,n} + \widehat{\delta}_{o,n})\mathrm{tr}[V_*] + \frac{2}{nk^2}\mathrm{tr}[V_*]^2\right\} + o_{\mathbb{P}}(1) \\ &= \frac{nk}{4}\left\{\frac{2}{k^2}\mathrm{tr}[K_{o,n}]^2 + \frac{4}{\sqrt{n}k^2}\mathrm{tr}[K_{o,n}]\mathrm{tr}[V_*] + \frac{2}{nk^2}\mathrm{tr}[V_*]^2\right\} + o_{\mathbb{P}}(1) \\ &= \frac{1}{2k}\left\{n\mathrm{tr}[K_{o,n}]^2 + 2\sqrt{n}\mathrm{tr}[K_{o,n}]\mathrm{tr}[V_*] + \mathrm{tr}[V_*]^2\right\} + o_{\mathbb{P}}(1) \\ &= \frac{1}{2k}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + o_{\mathbb{P}}(1),\end{aligned}$$

where the third to last equality holds by the definitions of $\widehat{\tau}_{o,n}$ and $\widehat{\delta}_{o,n}$. This shows the asymptotic approximation of $\widehat{\mathfrak{B}}_n^{(1)}$. We next note that $\widehat{\mathfrak{B}}_n^{(2)} := \frac{nk}{2}(\widehat{\tau}_n^2 + 2\widehat{\sigma}_n)$. Therefore,

$$\widehat{\mathfrak{B}}_n^{(2)} = \frac{nk}{2}\widehat{\tau}_{o,n}^2 + nk\widehat{\sigma}_{o,n} + \sqrt{n}\mathrm{tr}[V_*]\widehat{\tau}_{o,n} + \frac{1}{2k}\mathrm{tr}[V_*]^2 + \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] \\ - \frac{1}{2k}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + \frac{n}{2k}\mathrm{tr}[K_{o,n}]^2 - \frac{n}{2}\mathrm{tr}[K_{o,n}^2] + o_{\mathbb{P}}(1) = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1),$$

where the last equality holds by virtue of the definitions of $\widehat{\tau}_{o,n}$ and $\widehat{\sigma}_{o,n}$.

Finally, the structure of $\widehat{\mathfrak{B}}_n^{(3)}$ is symmetric to that of $\widehat{\mathfrak{B}}_n^{(2)}$. In the same way, it follows that $\widehat{\mathfrak{B}}_n^{(3)} = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)$.

(*ii*) From Lemmas 5(*i* and *iii*), it follows that

$$\begin{aligned}\widehat{\tau}_n^2 &= (\widehat{\tau}_{o,n} + n^{-1/2}k^{-1}\mathrm{tr}[V_*])^2 + o_{\mathbb{P}}(n^{-1}) = (k^{-1}\mathrm{tr}[K_{o,n}] + n^{-1/2}k^{-1}\mathrm{tr}[V_*] + O_{\mathbb{P}}(n^{-1}))^2 + o_{\mathbb{P}}(n^{-1}) \\ &= (k^{-1}\mathrm{tr}[K_{o,n} + n^{-1/2}V_*]))^2 + o_{\mathbb{P}}(n^{-1}), \qquad (32)\end{aligned}$$

and

$$\widehat{\eta}_n^2 = (\widehat{\eta}_{o,n} + n^{-1/2}k^{-1}\mathrm{tr}[V_*])^2 + o_{\mathbb{P}}(n^{-1}) = (k^{-1}\mathrm{tr}[K_{o,n}] + n^{-1/2}k^{-1}\mathrm{tr}[V_*] + O_{\mathbb{P}}(n^{-1}))^2 + o_{\mathbb{P}}(n^{-1})$$
$$= (k^{-1}\mathrm{tr}[K_{o,n} + n^{-1/2}V_*]))^2 + o_{\mathbb{P}}(n^{-1}), \tag{33}$$

where the second equality holds by Lemmas A5(*i* and *iii*). Given that $\widehat{\mathfrak{D}}_n^{(1)} := \frac{nk}{4}(\widehat{\tau}_n^2 + \widehat{\eta}_n^2)$, (32) and (33) imply that $\widehat{\mathfrak{D}}_n^{(1)} = \frac{1}{2k}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + o_{\mathbb{P}}(n^{-1})$, as desired.

From the definition $\widehat{\mathfrak{D}}_n^{(2)} := \frac{nk}{2}(\widehat{\tau}_n^2 + \widehat{\xi}_n)$, if we combine this with Lemma A5(*iv*) and (32), it follows that

$$\widehat{\mathfrak{D}}_n^{(2)} = \frac{k}{2}\{k^{-2}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + k^{-1}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2]$$
$$- k^{-2}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2\} + o_{\mathbb{P}}(n^{-1}) = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(n^{-1}).$$

This is the desired result for $\widehat{\mathfrak{D}}_n^{(2)}$.

Finally, from the fact that (32) has the same asymptotic approximation as that of (33), the asymptotic approximation of $\widehat{\mathfrak{D}}_n^{(3)}$ is identical to that of $\widehat{\mathfrak{D}}_n^{(2)}$.

(*iii*) From Lemma 5(*ii*),

$$\widehat{\delta}_n^2 = (\widehat{\delta}_{o,n} + n^{-1/2}k^{-1}\mathrm{tr}[V_*])^2 + O_{\mathbb{P}}(n^{-3/2})$$
$$= (k^{-1}\mathrm{tr}[K_{o,n}] + n^{-1/2}k^{-1}\mathrm{tr}[V_*] + O_{\mathbb{P}}(n^{-1}))^2 + O_{\mathbb{P}}(n^{-3/2})$$
$$= (k^{-1}\mathrm{tr}[K_{o,n} + n^{-1/2}V_*]))^2 + O_{\mathbb{P}}(n^{-3/2}), \tag{34}$$

where the second equality holds by Lemma A5(*ii*). Given that $\widehat{\mathfrak{S}}_n^{(1)} := \frac{nk}{4}(\widehat{\delta}_n^2 + \widehat{\eta}_n^2)$, (32) and (34) imply that $\widehat{\mathfrak{S}}_n^{(1)} = \mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + o_{\mathbb{P}}(n^{-1})$, as desired.

From the definition of $\widehat{\mathfrak{S}}_n^{(2)} := \frac{nk}{2}(\widehat{\delta}_n^2 + 2\widehat{\gamma}_n)$, it follows that

$$\widehat{\mathfrak{S}}_n^{(2)} = \frac{k}{2}\{k^{-2}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2 + k^{-1}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2]$$
$$- k^{-2}\mathrm{tr}[V_* + \sqrt{n}K_{o,n}]^2\} + o_{\mathbb{P}}(n^{-1}) = \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(n^{-1})$$

by using Lemma A5(*v*) and (33). This is the desired approximation for $\widehat{\mathfrak{S}}_n^{(2)}$. Finally, (34) has the same asymptotic approximation as that of (33), and this implies that the asymptotic expansion of $\widehat{\mathfrak{S}}_n^{(3)}$ is identical to that of $\widehat{\mathfrak{S}}_n^{(2)}$. ■

Proof of Lemma 6: From the definition of $\widehat{\lambda}_n$, if we approximate the log function around unity, it follows that

$$\widehat{\lambda}_n = \widehat{\delta}_n - \frac{1}{2}\widehat{\delta}_n^2 + o_{\mathbb{P}}(n^{-1}).$$

By Lemma 5(*ii*), $\widehat{\delta}_n = n^{-1/2}k^{-1}\mathrm{tr}[V_*] + k^{-1}\mathrm{tr}[K_{o,n}] + O_{\mathbb{P}}(n^{-1})$, so that $\frac{1}{2}\widehat{\delta}_n^2 = (2nk^2)^{-1}\mathrm{tr}[V_*]^2 + n^{-1/2}k^{-2}\mathrm{tr}[K_{o.n}]\mathrm{tr}[V_*] + (2k^2)^{-1}\mathrm{tr}[K_{o,n}]^2 + o_{\mathbb{P}}(n^{-1})$. Therefore, if we combine this result with Lemma 5(*ii*),

$$\begin{aligned}\widehat{\lambda}_n = \widehat{\delta}_{o,n} + \frac{1}{\sqrt{n}k}\mathrm{tr}[V_*] - \frac{1}{\sqrt{n}k}\mathrm{tr}[F_*W_{o,n} - C_*U_{o,n}] + \frac{1}{2nk}(\mathrm{tr}[C_*^2] - \mathrm{tr}[F_*^2]) - \frac{1}{2k^2}\mathrm{tr}[K_{o,n}]^2 \\ + \frac{1}{\sqrt{n}k}[\mathrm{tr}[Q_{j,*} - (F_*B_*^{-1}\partial_j B_* - C_*A_*^{-1}\partial_j A_*)]]'(\widehat{\theta}_n - \theta_*) + o_{\mathbb{P}}(n^{-1}),\end{aligned}$$

as desired. ■

Proof of Theorem 4: Using the equality that $\mathfrak{LR}_n = nk(\widehat{\tau}_n - \widehat{\lambda}_n)$ and Lemmas 5(*i*) and 6, we obtain that

$$\begin{aligned}\mathfrak{LR}_n &= nk(\widehat{\tau}_{o,n} - \widehat{\delta}_{n,o}) + \sqrt{n}\mathrm{tr}[K_{o,n}V_*] - \mathrm{tr}[C_*V_*] - \frac{1}{2}(\mathrm{tr}[C_*^2] - \mathrm{tr}[F_*^2]) + \frac{n}{2k}\mathrm{tr}[K_{o,n}]^2 + o_{\mathbb{P}}(1) \\ &= \frac{n}{2}\mathrm{tr}[K_{o,n}^2] + \sqrt{n}\mathrm{tr}[K_{o,n}V_*] - \mathrm{tr}[C_*V_*] - \frac{1}{2}(\mathrm{tr}[C_*^2] - \mathrm{tr}[F_*^2]) + o_{\mathbb{P}}(1) \\ &= \frac{n}{2}\mathrm{tr}[K_{o,n}^2] + \sqrt{n}\mathrm{tr}[K_{o,n}V_*] + \frac{1}{2}\mathrm{tr}[V_*^2] + o_{\mathbb{P}}(1) \\ &= \frac{1}{2}\mathrm{tr}[(V_* + \sqrt{n}K_{o,n})^2] + o_{\mathbb{P}}(1)\end{aligned}$$

using the fact that $\widehat{\sigma}_{o,n} := \widehat{\tau}_{o,n} - \widehat{\delta}_{o,n} = -(2k^2)^{-1}\mathrm{tr}[K_{o,n}]^2 + (2k)^{-1}\mathrm{tr}[K_{o,n}^2]$, where the second last holds by the definition of $V_* := F_* - C_*$. Note that the final right side is the desired result. ■

References

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

ALTMAN, M. AND MCDONALD, M. (2003): "Replication with Attnetion to Accuracy," *Political Analysis*, 11, 302–307.

AMISANO, G. AND GIANNINI, C. (1997): *Topics in Structural VAR Ecnometrics*. Berlin/Heidelberg: Springer-Verlag, 2nd edition.

ANDERSON, T. (2003): *An Introduction to Multivariate Statistical Analysis*. Hoboken, NJ: John Wiley & Sons, Inc.

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

BÉNABOU, R. (2000): "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review*, 90, 96–129.

BERA, A. (1986): "Model Specification Test Through Eigenvalues," Department of Economics, University of Illinois, mimeo.

BERA, A. AND HALL, A. (1991): "An Eigenvalue Based Test for Heteroskedasticity," Department of Economics, University of Illinois, mimeo.

BERNSTEIN, D. (2005): *Matrix Mathematics*. Princeton, NJ: Princeton University Press.

BERRY, W., DEMERITT, J., AND ESARLEY, J. (2010): "Testing for Interaction in Binary Logit and Probit Models: Is a Product Terms Essential?," *American Journal of Political Science*, 54, 248–266.

BESLEY, T. AND CASE, A. (2003): "Political Institutions and Policy Choices: Evidence from the United States," *Journal of Economics Literature*, 41, 7–73.

BOZDOGAN, H. (2000): "Akaike's Information Criterion and Recent Developments in Information Complexity," *Journal of Mathematical Psychology*, 44, 62–91.

CHESHER, A. (1983): "Information Matrix Test," *Economics Letters*, 13, 45–48.

CHESHER, A. AND SPADY, R. (1991): "Asymptotic Expansions of the Information Matrix Test Statistic," *Econometrica*, 59, 787–815.

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. (2014): "Testing the Equality of Two Positive-Definite Matrices with Application to Information Matrix Testing," *Advances in Econometrics: Essays in Honor of Peter C. B. Phillips*. Vol. 33. Eds. Yoosoon Chang, Thomas B. Fomby, and Joon Y. Park, 491–556. West Yorkshire, UK: Emerald Group Publishing Limited.

DHAENE, G. AND HOORELBEKE, D. (2004): "The Information Matrix Test with Bootstrap-Based Covariance Matrix Estimation," *Economics Letters*, 82, 341–347.

EFRON, B. AND TIBSHIRANI, R. (1993): *Introduction to the Bootstrap*, Monographs on Statistics & Applied Probability 57. New York: Chapman and Hall/CRC.

FEDDERSEN, T. AND PESENDORFER, W. (1996): "The Swing Voter's Curse," *American Economic Review*, 86, 408–424.

GOLDEN, R., HENLEY, S., WHITE, H., AND KASHNER, T. (2013): "New Directions in Information Matrix Testing: Eigenspectrum Tests," in *Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions Essays in Honour of Halbert L. White, Jr.*, eds. Norman Rasmus Swanson and Xiaohong Chen. New York: Springer, pp. 145–177.

HALL, A. (1987): "The Information Matrix Test for the Linear Model," *Review of Economic Studies*, 54, 257–263.

HAMILTON, J. (1994): *Time Series Analysis*. NJ: Princeton University Proess.

HOROWITZ, J. (1994): "Bootstrap-based Critical Values for the Information Matrix Test," *Journal of Econometrics*, 61, 395–411.

JAMES, A. T. (1964): "Distributions of Matrix Variates and Latent Roots Derived from Normal Samples," *Annals of Mathematical Statistics*, 35, 475–501.

JARQUE, C. AND BERA, A. (1987): "A Test for Normality of Observations and Regression Residuals," *International Statistical Review*, 55, 163–218.

LANCASTER, T. (1984): "The Covariance Matrix of the Information Matrix Test," *Econometrica*, 52, 1052–1053.

MAGNUS, J. (1985): "On Differentiating Eigenvalues and Eigenvectors," *Econometric Theory*, 1, 179–191.

MAGNUS, J. AND NEUDECKER, H. (1999): *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Chichester: John Wiley & Sons.

MARSHALL, A.W., OLKIN, I., AND ARNOLD, B. C. (1979): *Inequalities: Theory of Majorization and Its Applications*. Springer Series in Statistics. New York: Springer.

MAUCHLY, J. (1940): "Significance Test for Sphericity of a Normal n-variable Distribution," *Annals of Mathematical Statistics*, 11, 204–209.

MUIRHEAD, R. (1982): *Aspects of Multivariate Statistical Theory*. New York: John Wiley & Sons, Inc.

NAGAO, H. (1967): "Monotonicity of the Modified Likelihood Ratio Test for a Covariance Matrix," *Journal of Science of the Hiroshima University. series A-I*, 31, 147–150.

NAGARSENKER, B. AND PILLAI, K. (1973): "The Distribution of the Sphericity Test Criterion," *Journal of Multivariate Analysis*, 3, 226–235.

NAGLER, J. (1994): "Scobit: An Alternative Estimator to Logit and Probit," *American Journal of Political Science*, 38, 230–255.

NAGLER, J. (1991): "The Effect of Registration Laws and Education on U.S. Voter Turnout," *The American Political Science Review*, 85, 1393–1405.

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

ONATSKI, A., MOREIRA, M., AND HALLIN, M. (2013): "Asymptotic Power of Sphericity Tests for High-Dimensional Data," *Annals of Statistics*, 41, 1204–1231.

ORME, C. (1988): "The Calculation of the Information Matrix Test for Binary Data Models," *The Manchester School*, 56, 370–376.

ORME, C. (1990): "The Small-Sample Performance of the Information-Matrix Test," *Journal of Econometrics*, 46, 309–331.

SARGAN, J.D. (1988): *Lectures on Advanced Econometric Theory*. Oxford: Basil Blackwell.

TAYLOR, L. (1987): "The Size Bias of White's Information Matrix Test," *Economics Letters*, 24, 63–67.

TAKEUCHI, K. (1976): 'Distribution of Information Statistics and a Criterion of Model Fitting," *Suri-Kagaku (Mathematical Sciences)*, 153, 12–18 (in Japanese).

WALD, A. (1943): "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large," *Transactions of the 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*, 50, 1–26.

WHITE, H. (1994): *Estimation, Inference and Specification Analysis*. New York: Cambridge University Press.

WILKS, S. (1935): "On the Independence of k sets of Normally Distributed Statistical Variables," *Econometrica*, 3, 309–326.

WOLFINGER, R. AND ROSENSTONE, S. (1980): *Who Votes?* New Haven: Yale University Press.

	$\tau_*^2 > \max[\delta_*^2, \eta_*^2]$	$\delta_*^2 > \max[\tau_*^2, \eta_*^2]$	$\eta_*^2 > \max[\tau_*^2, \delta_*^2]$
$\sigma_* > \gamma_*$	$\widehat{\mathfrak{B}}_n^{(2)}$	$\widehat{\mathfrak{B}}_n^{(3)}$	$\widehat{\mathfrak{E}}_n^{(3)}$
$\sigma_* = \gamma_*$	$\widehat{\mathfrak{B}}_n^{(2)}$, $\widehat{\mathfrak{D}}_n^{(2)}$, $\widehat{\mathfrak{E}}_n^{(1)}$	$\widehat{\mathfrak{B}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(2)}$, $\widehat{\mathfrak{E}}_n^{(2)}$	$\widehat{\mathfrak{D}}_n^{(3)}$, $\widehat{\mathfrak{S}}_n^{(3)}$, $\widehat{\mathfrak{E}}_n^{(3)}$
$\sigma_* < \gamma_*$	$\widehat{\mathfrak{E}}_n^{(1)}$	$\widehat{\mathfrak{S}}_n^{(2)}$	$\widehat{\mathfrak{S}}_n^{(3)}$

Table 1: TEST STATISTICS WITH THE GREATEST LEADING TERMS UNDER THE ALTERNATIVE. The Test statistics in each cell indicates those with the greatest leading term under the alternative hypothesis and the condition for each cell.

Statistics $\setminus$ n	50	100	200	300	400	500
$\widehat{\mathfrak{B}}_n^{(1)}$	5.42	4.98	4.68	4.82	4.96	5.36
$\widehat{\mathfrak{B}}_n^{(2)}$	5.66	5.24	4.60	4.48	4.94	5.74
$\widehat{\mathfrak{B}}_n^{(3)}$	5.50	5.46	4.60	4.42	4.88	5.72
$\widehat{\mathfrak{D}}_n^{(1)}$	5.54	4.98	4.30	4.98	4.86	5.66
$\widehat{\mathfrak{D}}_n^{(2)}$	5.90	5.20	4.62	4.38	4.90	5.66
$\widehat{\mathfrak{D}}_n^{(3)}$	5.60	5.50	4.90	4.44	4.82	5.74
$\widehat{\mathfrak{S}}_n^{(1)}$	5.14	4.96	4.52	5.18	4.68	5.62
$\widehat{\mathfrak{S}}_n^{(2)}$	5.66	5.22	4.74	4.38	4.94	5.58
$\widehat{\mathfrak{S}}_n^{(3)}$	5.48	5.34	4.90	4.44	4.94	5.58
$\widehat{\mathfrak{E}}_n^{(1)}$	5.86	5.24	4.66	4.30	4.92	5.42
$\widehat{\mathfrak{E}}_n^{(2)}$	5.62	5.26	4.56	4.32	4.74	5.78
$\widehat{\mathfrak{E}}_n^{(3)}$	5.64	5.28	4.78	4.54	4.76	5.76
$\widehat{\mathfrak{M}}_n^{(1)}$	5.12	4.92	4.60	5.10	4.68	5.64
$\widehat{\mathfrak{M}}_n^{(2)}$	5.38	5.40	4.78	4.44	4.70	5.68
$\widehat{\mathfrak{M}}_n^{(3)}$	5.38	5.40	4.78	4.44	4.70	5.68
$\mathfrak{J}_n$	5.18	5.18	5.02	4.42	5.08	5.58
$\mathfrak{I}_n$	4.88	5.04	4.98	5.14	5.08	5.14

Table 2: EMPIRICAL LEVELS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 5,000. Bootstrap Repetitions: 500. Model: $y_t = X_t{}'\beta + u_t$, $X_t = (1, x_t)'$ and $u_t \sim N(0, \sigma^2)$. DGP: $y_t = x_t + u_t$, $u_t|x_t \sim N(0,1)$, and $x_t \sim N(0,1)$.

Statistics \ n	50	100	150	200	250	300
$\widehat{\mathfrak{B}}_n^{(1)}$	52.20	80.85	93.10	97.15	98.90	99.55
$\widehat{\mathfrak{B}}_n^{(2)}$	85.10	99.55	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{B}}_n^{(3)}$	87.30	99.90	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(1)}$	70.55	93.05	99.25	99.50	99.95	100.0
$\widehat{\mathfrak{D}}_n^{(2)}$	86.05	99.75	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(3)}$	89.60	99.90	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(1)}$	36.05	61.95	83.85	91.95	97.20	99.35
$\widehat{\mathfrak{S}}_n^{(2)}$	89.05	99.95	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(3)}$	90.15	99.95	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(1)}$	86.70	99.90	99.95	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(2)}$	88.55	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(3)}$	88.80	99.90	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(1)}$	59.85	88.15	98.05	99.05	99.90	100.0
$\widehat{\mathfrak{M}}_n^{(2)}$	88.30	99.90	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(3)}$	88.30	99.90	100.0	100.0	100.0	100.0
$\mathfrak{I}_n$	49.35	79.00	92.30	96.10	98.85	99.50
$\mathfrak{I}_n$	19.45	87.90	100.0	100.0	100.0	100.0

Table 3: EMPIRICAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model: $y_t = X_t{}'\beta + u_t$, $X_t = (1, x_t)'$, and $u_t \sim N(0, \sigma^2)$. DGP: $y_t = x_t + u_t$, $u_t|x_t \sim N(0, \exp(x_t))$, and $x_t \sim N(1, 1)$.

Statistics \ n	100	500	800	1,200	1,600	2,000
$\widehat{\mathfrak{B}}_n^{(1)}$	18.90	64.45	82.15	93.85	98.45	99.50
$\widehat{\mathfrak{B}}_n^{(2)}$	3.90	46.25	84.15	97.30	99.95	100.0
$\widehat{\mathfrak{B}}_n^{(3)}$	6.25	54.90	87.75	98.10	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(1)}$	22.80	72.45	87.60	96.50	99.30	99.85
$\widehat{\mathfrak{D}}_n^{(2)}$	3.85	48.55	85.45	97.45	99.95	100.0
$\widehat{\mathfrak{D}}_n^{(3)}$	10.00	62.10	90.70	98.55	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(1)}$	27.00	77.05	90.55	97.50	99.65	99.90
$\widehat{\mathfrak{S}}_n^{(2)}$	7.90	58.35	89.30	98.35	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(3)}$	10.75	63.40	91.50	98.60	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(1)}$	4.30	50.50	86.35	97.70	99.95	100.0
$\widehat{\mathfrak{E}}_n^{(2)}$	5.35	44.75	79.70	95.85	99.90	100.0
$\widehat{\mathfrak{E}}_n^{(3)}$	9.25	60.70	90.25	98.50	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(1)}$	25.65	76.40	90.30	97.45	99.65	99.90
$\widehat{\mathfrak{M}}_n^{(2)}$	12.65	60.50	88.95	98.00	99.95	100.0
$\widehat{\mathfrak{M}}_n^{(3)}$	12.65	60.50	88.95	98.00	99.95	100.0
$\mathfrak{I}_n$	1.60	9.35	20.10	38.40	52.65	65.85
$\mathfrak{I}_n$	15.45	50.65	66.25	82.45	90.10	95.25

Table 4: EMPIRICAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model: $y_t = X_t{}'\beta + u_t$, $X_t = (1, x_t)'$, and $u_t \sim N(0, \sigma^2)$. DGP: $y_t = x_t + u_t$, $u_t|x_t$ IID $\sim$ *MN*$(-1, 1; 1, 1; 1/[1 + \exp(x_t)])$, and $x_t \sim N(0, 1)$.

Statistics \ n	200	500	800	1,200	1,600	2,000
$\widehat{\mathfrak{B}}_n^{(1)}$	39.40	70.30	85.20	93.55	96.80	98.80
$\widehat{\mathfrak{B}}_n^{(2)}$	42.85	71.70	87.00	94.75	98.25	99.20
$\widehat{\mathfrak{B}}_n^{(3)}$	38.75	68.80	85.35	94.05	97.70	99.05
$\widehat{\mathfrak{D}}_n^{(1)}$	33.45	65.90	81.05	91.65	96.00	98.30
$\widehat{\mathfrak{D}}_n^{(2)}$	43.90	72.30	86.80	94.70	98.15	99.20
$\widehat{\mathfrak{D}}_n^{(3)}$	33.45	64.40	82.70	92.45	97.10	98.70
$\widehat{\mathfrak{S}}_n^{(1)}$	12.85	46.80	70.15	84.30	92.10	95.75
$\widehat{\mathfrak{S}}_n^{(2)}$	39.50	69.20	85.15	93.85	97.85	99.00
$\widehat{\mathfrak{S}}_n^{(3)}$	32.80	64.15	82.50	92.15	97.00	98.45
$\widehat{\mathfrak{E}}_n^{(1)}$	44.50	72.85	87.00	94.80	98.20	99.30
$\widehat{\mathfrak{E}}_n^{(2)}$	38.70	67.25	84.05	93.30	97.35	98.55
$\widehat{\mathfrak{E}}_n^{(3)}$	33.50	64.50	82.65	92.80	97.20	98.75
$\widehat{\mathfrak{M}}_n^{(1)}$	22.65	58.85	78.05	89.90	95.45	98.00
$\widehat{\mathfrak{M}}_n^{(2)}$	29.85	62.95	81.30	91.85	96.70	98.35
$\widehat{\mathfrak{M}}_n^{(3)}$	29.85	62.95	81.30	91.85	96.70	98.35
$\mathfrak{I}_n$	16.40	36.90	51.55	64.55	73.30	78.45
$\mathfrak{I}_n$	20.60	33.05	45.20	59.95	73.65	84.45

Table 5: EMPIRICAL GLOBAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model: $y_t = X_t{}'\beta + u_t$, $X_t = (1, x_t)'$, and $u_t \sim N(0, \sigma^2)$. DGP: $y_t = x_t + \frac{1}{2}x_t^2 + u_t$, $u_t|x_t \sim N(0,1)$, and $x_t \sim N(0,1)$.

Statistics \ n	100	200	300	400	500	1,000
$\widehat{\mathfrak{B}}_n^{(1)}$	7.30	6.70	6.53	5.47	5.33	5.63
$\widehat{\mathfrak{B}}_n^{(2)}$	13.07	13.00	12.13	12.47	11.23	11.87
$\widehat{\mathfrak{B}}_n^{(3)}$	13.80	13.13	12.53	13.00	11.33	11.97
$\widehat{\mathfrak{D}}_n^{(1)}$	6.83	6.47	6.27	5.23	4.83	5.57
$\widehat{\mathfrak{D}}_n^{(2)}$	13.23	13.00	12.30	12.43	11.37	11.83
$\widehat{\mathfrak{D}}_n^{(3)}$	13.50	13.10	12.87	12.77	11.40	11.97
$\widehat{\mathfrak{S}}_n^{(1)}$	4.83	5.60	5.13	5.00	4.40	5.50
$\widehat{\mathfrak{S}}_n^{(2)}$	13.83	13.27	12.57	12.73	11.40	12.07
$\widehat{\mathfrak{S}}_n^{(3)}$	12.90	13.23	12.77	12.80	11.40	12.10
$\widehat{\mathfrak{E}}_n^{(1)}$	13.03	13.03	12.27	12.87	11.27	11.77
$\widehat{\mathfrak{E}}_n^{(2)}$	13.87	13.70	12.97	13.43	11.73	12.13
$\widehat{\mathfrak{E}}_n^{(3)}$	13.37	12.83	13.07	12.90	11.27	12.03
$\widehat{\mathfrak{M}}_n^{(1)}$	5.77	6.23	6.17	5.47	4.83	5.63
$\widehat{\mathfrak{M}}_n^{(2)}$	12.77	12.87	12.87	12.90	11.53	12.40
$\widehat{\mathfrak{M}}_n^{(3)}$	12.77	12.87	12.87	12.90	11.53	12.40
$\mathfrak{I}_n$	6.87	6.27	6.37	6.00	5.67	5.13
$\mathfrak{I}_n$	6.17	8.73	9.30	10.17	10.03	10.27

Table 6: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 3,000. Bootstrap Repetitions: 500. Model: $y_t = X_t{}'\beta + u_t$, $X_t = (1, x_t)'$, and $u_t \sim N(0, \sigma^2)$. DGP: $y_t = x_t + u_t$, $u_t|x_t \sim N(0, \exp(2n^{-1/2}x_t))$, and $x_t \sim N(1,1)$.

Statistics \ n	1,000	5,000	10,000	15,000	20,000	30,000
$\widehat{\mathfrak{B}}_n^{(1)}$	7.30	8.93	9.13	8.57	10.10	9.60
$\widehat{\mathfrak{B}}_n^{(2)}$	10.83	14.03	14.70	16.03	17.00	18.37
$\widehat{\mathfrak{B}}_n^{(3)}$	10.90	13.87	14.43	15.73	16.97	18.23
$\widehat{\mathfrak{D}}_n^{(1)}$	6.40	8.63	8.73	8.57	9.90	9.30
$\widehat{\mathfrak{D}}_n^{(2)}$	10.87	14.10	14.57	15.87	17.03	18.23
$\widehat{\mathfrak{D}}_n^{(3)}$	10.60	13.57	14.00	15.23	16.70	17.87
$\widehat{\mathfrak{S}}_n^{(1)}$	5.63	7.90	8.30	8.27	9.37	9.03
$\widehat{\mathfrak{S}}_n^{(2)}$	10.90	13.80	14.40	15.60	16.83	18.10
$\widehat{\mathfrak{S}}_n^{(3)}$	10.40	13.57	13.90	15.30	16.57	17.87
$\widehat{\mathfrak{E}}_n^{(1)}$	10.73	13.90	14.60	16.00	17.10	18.27
$\widehat{\mathfrak{E}}_n^{(2)}$	10.97	14.03	14.40	15.53	16.67	17.73
$\widehat{\mathfrak{E}}_n^{(3)}$	10.63	13.60	14.10	15.60	16.80	17.90
$\widehat{\mathfrak{M}}_n^{(1)}$	6.03	8.13	8.53	8.40	9.57	9.30
$\widehat{\mathfrak{M}}_n^{(2)}$	10.30	13.80	14.03	15.23	16.40	17.50
$\widehat{\mathfrak{M}}_n^{(3)}$	10.30	13.80	14.03	15.23	16.40	17.50
$\mathfrak{I}_n$	8.13	8.53	8.30	8.37	9.53	9.63
$\mathfrak{I}_n$	6.23	8.67	11.47	12.67	14.37	15.83

Table 7: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 3,000. Bootstrap Repetitions: 500. Model: $y_t = X_t{}'\beta + u_t$, $X_t = (1, x_t)'$, and $u_t \sim N(0, \sigma^2)$. DGP: $y_t = x_t + u_t$, $u_t|x_t \sim$ IND *MN*$(0, 1; 1, 1; 1/[1 + n^{-1/2}\exp(x_t)])$, and $x_t \sim$ IID $N(0, 1)$.

Statistics \ n	3,000	4,000	5,000	10,000	20,000	30,000
$\widehat{\mathfrak{B}}_n^{(1)}$	21.73	18.03	15.43	11.70	7.57	8.13
$\widehat{\mathfrak{B}}_n^{(2)}$	20.43	17.30	14.47	9.67	7.47	7.03
$\widehat{\mathfrak{B}}_n^{(3)}$	19.67	16.43	13.97	9.57	7.37	6.93
$\widehat{\mathfrak{D}}_n^{(1)}$	20.03	17.30	15.13	11.33	7.43	7.90
$\widehat{\mathfrak{D}}_n^{(2)}$	20.43	17.10	14.43	9.63	7.47	7.03
$\widehat{\mathfrak{D}}_n^{(3)}$	18.87	15.77	13.57	9.50	7.20	7.10
$\widehat{\mathfrak{S}}_n^{(1)}$	19.03	16.57	14.43	11.20	7.30	7.70
$\widehat{\mathfrak{S}}_n^{(2)}$	19.60	16.40	14.03	9.77	7.40	7.07
$\widehat{\mathfrak{S}}_n^{(3)}$	18.97	15.80	13.30	9.50	7.23	7.17
$\widehat{\mathfrak{E}}_n^{(1)}$	20.43	17.03	14.40	9.77	7.57	7.10
$\widehat{\mathfrak{E}}_n^{(2)}$	17.57	14.60	12.43	8.67	6.77	6.37
$\widehat{\mathfrak{E}}_n^{(3)}$	19.03	15.93	13.60	9.57	7.17	7.07
$\widehat{\mathfrak{M}}_n^{(1)}$	19.47	16.73	14.63	11.13	7.30	7.83
$\widehat{\mathfrak{M}}_n^{(2)}$	17.00	14.03	12.20	8.53	6.73	6.47
$\widehat{\mathfrak{M}}_n^{(3)}$	17.00	14.03	12.20	8.53	6.73	6.47
$\mathfrak{I}_n$	5.20	5.60	5.53	5.23	4.97	5.17
$\mathfrak{I}_n$	11.60	10.63	9.10	7.70	6.47	6.27

Table 8: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 3,000. Bootstrap Repetitions: 500. Model: $y_t = X_t{}'\beta + u_t$, $X_t = (1, x_t)'$, and $u_t \sim N(0, \sigma^2)$. DGP: $y_t = x_t + 5n^{-1/2}x_t^2 + u_t$, $u_t|x_t \sim N(0, 1)$, and $x_t \sim N(0, 1)$.

Statistics $\backslash$ n	50	100	200	300	400	500
$\widehat{\mathfrak{B}}_n^{(1)}$	2.32	4.06	4.60	4.82	4.84	4.80
$\widehat{\mathfrak{B}}_n^{(2)}$	3.34	4.10	4.28	5.02	4.84	4.44
$\widehat{\mathfrak{B}}_n^{(3)}$	2.54	3.82	4.30	4.80	4.74	4.54
$\widehat{\mathfrak{D}}_n^{(1)}$	1.94	3.86	4.42	4.80	4.56	4.76
$\widehat{\mathfrak{D}}_n^{(2)}$	3.40	4.14	4.26	5.00	4.86	4.48
$\widehat{\mathfrak{D}}_n^{(3)}$	1.88	3.66	4.20	4.66	4.52	4.70
$\widehat{\mathfrak{S}}_n^{(1)}$	1.28	3.50	4.26	4.98	4.64	4.84
$\widehat{\mathfrak{S}}_n^{(2)}$	2.40	3.78	4.36	4.92	4.76	4.60
$\widehat{\mathfrak{S}}_n^{(3)}$	1.78	3.64	4.22	4.70	4.50	4.72
$\widehat{\mathfrak{E}}_n^{(1)}$	3.44	4.20	4.28	5.00	4.88	4.46
$\widehat{\mathfrak{E}}_n^{(2)}$	2.96	3.80	4.42	4.90	4.82	4.64
$\widehat{\mathfrak{E}}_n^{(3)}$	2.00	3.62	4.20	4.72	4.58	4.68
$\widehat{\mathfrak{M}}_n^{(1)}$	1.20	3.56	4.20	4.88	4.50	4.74
$\widehat{\mathfrak{M}}_n^{(2)}$	1.40	3.54	4.30	4.78	4.56	4.68
$\widehat{\mathfrak{M}}_n^{(3)}$	1.40	3.54	4.30	4.78	4.56	4.68
$\mathfrak{I}_n$	4.26	4.44	4.70	5.42	4.30	5.08

Table 9: EMPIRICAL LEVELS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 5,000. Bootstrap Repetitions: 500. Model for $\mathbb{E}[y_t|x_t]$: Probit$(X_t{}'\beta)$ and $X_t = (1, x_t)'$. DGP: $\mathbb{E}[y_t|x_t] =$ Probit$(1 + x_t)$ and $x_t \sim N(0, 1)$.

Statistics $\backslash$ n	50	100	150	200	250	300
$\widehat{\mathfrak{B}}_n^{(1)}$	44.60	73.90	88.15	93.65	97.30	98.90
$\widehat{\mathfrak{B}}_n^{(2)}$	77.45	97.80	99.80	99.95	99.95	100.0
$\widehat{\mathfrak{B}}_n^{(3)}$	74.60	97.80	99.80	99.95	99.95	100.0
$\widehat{\mathfrak{D}}_n^{(1)}$	37.75	69.20	85.55	92.15	96.00	98.00
$\widehat{\mathfrak{D}}_n^{(2)}$	77.50	97.80	99.80	99.95	99.95	100.0
$\widehat{\mathfrak{D}}_n^{(3)}$	70.85	97.75	99.80	99.95	99.95	100.0
$\widehat{\mathfrak{S}}_n^{(1)}$	21.80	49.35	66.50	77.00	84.90	90.55
$\widehat{\mathfrak{S}}_n^{(2)}$	74.45	97.75	99.80	99.95	99.95	100.0
$\widehat{\mathfrak{S}}_n^{(3)}$	70.60	97.75	99.80	99.95	99.95	100.0
$\widehat{\mathfrak{E}}_n^{(1)}$	77.60	97.80	99.80	99.95	99.95	100.0
$\widehat{\mathfrak{E}}_n^{(2)}$	77.35	98.05	99.80	99.95	99.95	100.0
$\widehat{\mathfrak{E}}_n^{(3)}$	71.05	97.80	99.80	99.95	99.95	100.0
$\widehat{\mathfrak{M}}_n^{(1)}$	31.00	69.45	86.35	93.10	96.85	98.55
$\widehat{\mathfrak{M}}_n^{(2)}$	69.45	97.85	99.75	99.95	99.95	100.0
$\widehat{\mathfrak{M}}_n^{(3)}$	69.45	97.85	99.75	99.95	99.95	100.0
$\mathfrak{I}_n$	96.00	100.0	100.0	100.0	100.0	100.0

Table 10: EMPIRICAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model for $\mathbb{E}[y_t|x_t]$: Probit$(X_t{}'\beta)$ and $X_t = (1, x_t)'$. DGP: $\mathbb{E}[y_t|x_t] =$ Probit$(1 + x_t - x_t^4)$ and $x_t \sim N(0, 1)$.

Statistics $\setminus n$	1,000	2,000	3,000	4,000	5,000	10,000
$\widehat{\mathfrak{B}}_n^{(1)}$	8.75	11.75	12.95	15.60	18.55	17.83
$\widehat{\mathfrak{B}}_n^{(2)}$	9.20	12.90	14.45	17.20	21.45	22.00
$\widehat{\mathfrak{B}}_n^{(3)}$	8.70	12.10	13.75	16.60	20.25	21.25
$\widehat{\mathfrak{D}}_n^{(1)}$	8.05	10.75	12.35	14.05	17.05	16.17
$\widehat{\mathfrak{D}}_n^{(2)}$	9.20	12.85	14.50	17.25	21.45	22.00
$\widehat{\mathfrak{D}}_n^{(3)}$	8.00	11.05	12.95	15.40	19.15	19.00
$\widehat{\mathfrak{S}}_n^{(1)}$	6.55	9.50	11.10	13.00	15.10	14.83
$\widehat{\mathfrak{S}}_n^{(2)}$	8.65	12.00	13.85	16.60	20.25	21.25
$\widehat{\mathfrak{S}}_n^{(3)}$	7.95	11.10	12.95	15.40	19.15	19.00
$\widehat{\mathfrak{E}}_n^{(1)}$	9.20	12.85	14.50	17.25	21.55	22.00
$\widehat{\mathfrak{E}}_n^{(2)}$	8.60	12.10	14.05	16.90	20.35	21.08
$\widehat{\mathfrak{E}}_n^{(3)}$	8.15	11.05	12.95	15.40	19.15	18.92
$\widehat{\mathfrak{M}}_n^{(1)}$	7.25	9.70	10.80	12.75	15.45	15.17
$\widehat{\mathfrak{M}}_n^{(2)}$	7.70	10.50	12.60	14.85	19.20	19.42
$\widehat{\mathfrak{M}}_n^{(3)}$	7.70	10.50	12.60	14.85	19.20	19.42
$\mathfrak{I}_n$	2.75	2.90	2.40	2.25	2.85	2.75

Table 11: EMPIRICAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model for $\mathbb{E}[y_t|x_t]$: Probit$(X_t{}'\beta)$ and $X_t = (1, x_t)'$. DGP: $\mathbb{E}[y_t|x_t] =$ Logit$[-\pi^2(1 + x_t)/6]$ and $x_t \sim N(0, 1)$.

Statistics $\setminus n$	500	1,000	1,500	2,000	2,500	3,000
$\widehat{\mathfrak{B}}_n^{(1)}$	63.00	53.13	45.50	44.18	30.90	29.40
$\widehat{\mathfrak{B}}_n^{(2)}$	65.70	57.47	51.47	54.98	41.17	39.80
$\widehat{\mathfrak{B}}_n^{(3)}$	65.40	56.97	51.37	54.37	41.00	39.50
$\widehat{\mathfrak{D}}_n^{(1)}$	62.53	52.40	45.07	43.63	30.43	28.80
$\widehat{\mathfrak{D}}_n^{(2)}$	65.67	57.47	51.43	54.80	41.13	39.80
$\widehat{\mathfrak{D}}_n^{(3)}$	65.07	56.47	51.30	54.06	40.57	39.37
$\widehat{\mathfrak{S}}_n^{(1)}$	61.83	51.57	44.07	42.31	29.53	28.23
$\widehat{\mathfrak{S}}_n^{(2)}$	65.43	56.93	51.37	54.37	40.93	39.53
$\widehat{\mathfrak{S}}_n^{(3)}$	65.07	56.47	51.30	54.03	40.57	39.37
$\widehat{\mathfrak{E}}_n^{(1)}$	65.70	57.47	51.43	54.80	41.10	39.80
$\widehat{\mathfrak{E}}_n^{(2)}$	66.23	58.20	52.77	57.60	43.37	42.47
$\widehat{\mathfrak{E}}_n^{(3)}$	65.07	56.47	51.30	54.06	40.57	39.37
$\widehat{\mathfrak{M}}_n^{(1)}$	62.57	52.33	44.83	43.69	30.80	28.90
$\widehat{\mathfrak{M}}_n^{(2)}$	65.43	57.47	52.03	55.94	42.80	41.80
$\widehat{\mathfrak{M}}_n^{(3)}$	65.43	57.47	52.03	55.94	42.80	41.80
$\mathfrak{I}_n$	22.57	24.33	25.73	35.11	30.83	30.70

Table 12: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 3,000. Bootstrap Repetitions: 500. Model for $\mathbb{E}[y_t|x_t]$: Probit$(X_t{}'\beta)$ and $X_t = (1, x_t)'$. DGP: $\mathbb{E}[y_t|x_t] =$ Probit$(1 + x_t - n^{-1/2}x_t^4)$ and $x_t \sim$ IID $N(0, 1)$.

Statistics $\setminus$ n	2,500	3,000	3,500	4,000	4,500	5,000
$\widehat{\mathfrak{B}}_n^{(1)}$	72.03	75.45	77.67	78.77	80.37	82.00
$\widehat{\mathfrak{B}}_n^{(2)}$	75.80	78.60	81.23	82.33	83.37	84.90
$\widehat{\mathfrak{B}}_n^{(3)}$	75.30	78.20	80.63	81.83	83.00	84.33
$\widehat{\mathfrak{D}}_n^{(1)}$	70.83	74.85	76.77	77.83	79.50	81.37
$\widehat{\mathfrak{D}}_n^{(2)}$	75.80	78.60	81.23	82.33	83.37	84.90
$\widehat{\mathfrak{D}}_n^{(3)}$	74.63	77.55	80.17	81.37	82.67	83.87
$\widehat{\mathfrak{S}}_n^{(1)}$	69.43	73.60	75.43	76.67	78.87	80.30
$\widehat{\mathfrak{S}}_n^{(2)}$	75.27	78.10	80.63	81.83	83.03	84.33
$\widehat{\mathfrak{S}}_n^{(3)}$	74.63	77.55	80.17	81.37	82.67	83.87
$\widehat{\mathfrak{E}}_n^{(1)}$	75.80	78.60	81.23	82.33	83.37	84.90
$\widehat{\mathfrak{E}}_n^{(2)}$	75.23	78.20	80.60	81.90	83.10	84.17
$\widehat{\mathfrak{E}}_n^{(3)}$	74.63	77.55	80.17	81.37	82.67	83.87
$\widehat{\mathfrak{M}}_n^{(1)}$	70.10	74.20	75.83	77.43	79.23	80.53
$\widehat{\mathfrak{M}}_n^{(2)}$	74.20	77.30	79.80	81.03	82.50	83.57
$\widehat{\mathfrak{M}}_n^{(3)}$	74.20	77.30	79.80	81.03	82.50	83.57
$\mathfrak{I}_n$	16.07	20.35	24.87	27.60	30.20	31.53

Table 13: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 3,000. Bootstrap Repetitions: 500. Model for $\mathbb{E}[y_t|x_t]$: Probit$(X_t{}'\beta)$ and $X_t = (1, x_t)'$. DGP: $\mathbb{E}[y_t|x_t] = (1 - n^{-1/2})$Probit$(1 + x_t) + n^{-1/2}$Logit$[-\pi^2(1 + x_t)/6]$ and $x_t \sim$ IID $N(0, 1)$.

Statistics $\setminus$ n	50	100	200	300	400	500
$\widehat{\mathfrak{B}}_n^{(1)}$	2.54	3.50	4.62	4.82	4.94	4.94
$\widehat{\mathfrak{B}}_n^{(2)}$	3.02	3.56	4.18	4.30	4.44	4.36
$\widehat{\mathfrak{B}}_n^{(3)}$	3.84	4.18	4.26	4.62	4.58	4.76
$\widehat{\mathfrak{D}}_n^{(1)}$	3.08	4.02	4.98	5.08	5.12	4.96
$\widehat{\mathfrak{D}}_n^{(2)}$	2.88	3.42	4.08	4.32	4.40	4.32
$\widehat{\mathfrak{D}}_n^{(3)}$	4.92	4.80	4.50	4.86	4.66	5.02
$\widehat{\mathfrak{S}}_n^{(1)}$	4.70	5.10	5.38	5.12	5.56	5.28
$\widehat{\mathfrak{S}}_n^{(2)}$	3.28	4.12	4.20	4.54	4.60	4.62
$\widehat{\mathfrak{S}}_n^{(3)}$	4.90	4.76	4.48	4.88	4.68	5.02
$\widehat{\mathfrak{E}}_n^{(1)}$	2.90	3.32	4.02	4.40	4.44	4.42
$\widehat{\mathfrak{E}}_n^{(2)}$	4.30	4.30	4.54	4.86	4.60	4.70
$\widehat{\mathfrak{E}}_n^{(3)}$	5.02	4.78	4.42	4.94	4.64	5.06
$\widehat{\mathfrak{M}}_n^{(1)}$	7.84	6.76	6.32	5.42	5.96	5.52
$\widehat{\mathfrak{M}}_n^{(2)}$	7.64	6.12	5.12	5.52	5.02	5.32
$\widehat{\mathfrak{M}}_n^{(3)}$	7.64	6.12	5.12	5.52	5.02	5.32

Table 14: EMPIRICAL LEVELS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 5,000. Bootstrap Repetitions: 500. Model for LS: $X_t{}'\beta_*$ and $X_t = (1, x_t)'$. DGP: $y_t = \frac{1}{2}x_t + u_t$ and $(x_t, u_t)' \sim$ IID $N(0, I_2)$.

Statistics \ n	50	100	150	200	250	300
$\widehat{\mathfrak{B}}_n^{(1)}$	40.87	66.00	81.80	91.60	96.65	97.05
$\widehat{\mathfrak{B}}_n^{(2)}$	81.53	98.27	99.85	100.0	100.0	100.0
$\widehat{\mathfrak{B}}_n^{(3)}$	90.10	99.30	99.95	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(1)}$	63.23	91.53	99.15	99.95	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(2)}$	75.97	97.27	99.85	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(3)}$	93.23	99.70	99.95	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(1)}$	29.83	49.40	67.65	79.80	88.65	91.30
$\widehat{\mathfrak{S}}_n^{(2)}$	83.77	98.93	99.95	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(3)}$	93.03	99.67	99.95	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(1)}$	65.40	93.13	99.75	99.90	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(2)}$	89.83	99.30	99.95	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(3)}$	93.53	99.73	99.95	100.0	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(1)}$	60.53	87.93	98.15	99.55	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(2)}$	84.70	98.87	99.95	99.90	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(3)}$	84.70	98.87	99.95	99.90	100.0	100.0

Table 15: EMPIRICAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model for LS: ${X_t}'\beta_*$ and $X_t = (1, x_t)'$. DGP: $y_t = \frac{1}{2}x_t + (1 + \exp(x_t))u_t$ and $(x_t, u_t)' \sim$ IID $N(0, I_2)$.

Statistics \ n	50	100	150	200	250	300
$\widehat{\mathfrak{B}}_n^{(1)}$	86.30	98.10	99.65	100.0	100.0	100.0
$\widehat{\mathfrak{B}}_n^{(2)}$	91.60	99.45	99.90	100.0	100.0	100.0
$\widehat{\mathfrak{B}}_n^{(3)}$	90.10	99.30	99.80	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(1)}$	86.00	97.90	99.70	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(2)}$	92.00	99.50	99.90	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(3)}$	86.50	98.70	99.80	100.0	99.95	100.0
$\widehat{\mathfrak{S}}_n^{(1)}$	27.55	39.45	53.35	69.90	82.60	87.45
$\widehat{\mathfrak{S}}_n^{(2)}$	91.05	99.50	99.90	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(3)}$	86.70	98.65	99.80	100.0	99.95	100.0
$\widehat{\mathfrak{E}}_n^{(1)}$	92.20	99.60	99.90	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(2)}$	90.10	99.05	99.90	99.95	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(3)}$	86.35	98.70	99.80	100.0	99.95	100.0
$\widehat{\mathfrak{M}}_n^{(1)}$	50.50	88.20	98.70	99.90	99.90	100.0
$\widehat{\mathfrak{M}}_n^{(2)}$	68.30	96.60	99.55	99.90	99.95	100.0
$\widehat{\mathfrak{M}}_n^{(3)}$	68.30	96.60	99.55	99.90	99.95	100.0

Table 16: EMPIRICAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model for LS: ${X_t}'\beta_*$ and $X_t = (1, x_t)'$. DGP: $y_t = x_t^4 + u_t$ and $(x_t, u_t)' \sim$ IID $N(0, I_2)$.

Statistics $\backslash$ n	50	100	150	200	250	300
$\widehat{\mathfrak{B}}_n^{(1)}$	99.05	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{B}}_n^{(2)}$	99.85	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{B}}_n^{(3)}$	99.60	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(1)}$	99.15	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(2)}$	99.85	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(3)}$	99.50	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(1)}$	61.85	97.25	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(2)}$	99.75	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(3)}$	99.55	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(1)}$	99.80	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(2)}$	99.80	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(3)}$	99.45	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(1)}$	95.85	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(2)}$	99.05	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(3)}$	99.05	100.0	100.0	100.0	100.0	100.0

Table 17: EMPIRICAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model for LS: $X_t{}'\beta_*$ and $X_t = (1, x_t)'$. DGP: $y_t = \frac{1}{2}x_t + u_t$, $u_t = \frac{1}{2}u_{t-1} + \varepsilon_t$, and $(x_t, \varepsilon_t)' \sim$ IID $N(0, I_2)$.

Statistics $\backslash$ n	100	200	400	600	800	1,000
$\widehat{\mathfrak{B}}_n^{(1)}$	30.73	38.43	43.07	48.10	47.37	47.13
$\widehat{\mathfrak{B}}_n^{(2)}$	62.17	78.67	85.20	90.00	89.07	90.73
$\widehat{\mathfrak{B}}_n^{(3)}$	63.70	78.97	84.57	89.53	88.50	89.80
$\widehat{\mathfrak{D}}_n^{(1)}$	32.70	37.37	39.63	43.20	41.40	41.60
$\widehat{\mathfrak{D}}_n^{(2)}$	61.07	78.40	85.03	89.90	89.03	90.67
$\widehat{\mathfrak{D}}_n^{(3)}$	63.37	77.97	83.77	88.73	88.03	88.93
$\widehat{\mathfrak{S}}_n^{(1)}$	7.40	12.63	17.50	24.17	24.33	28.03
$\widehat{\mathfrak{S}}_n^{(2)}$	62.27	78.37	84.03	89.33	88.50	89.77
$\widehat{\mathfrak{S}}_n^{(3)}$	62.60	77.63	83.60	88.77	87.97	88.93
$\widehat{\mathfrak{E}}_n^{(1)}$	59.50	78.00	84.90	89.90	88.97	90.63
$\widehat{\mathfrak{E}}_n^{(2)}$	66.37	80.93	86.20	91.00	89.87	91.00
$\widehat{\mathfrak{E}}_n^{(3)}$	64.23	78.43	83.97	88.80	88.07	89.03
$\widehat{\mathfrak{M}}_n^{(1)}$	24.93	31.77	36.60	42.40	41.07	42.30
$\widehat{\mathfrak{M}}_n^{(2)}$	59.93	77.70	83.57	89.23	88.50	89.80
$\widehat{\mathfrak{M}}_n^{(3)}$	59.93	77.70	83.57	89.23	88.50	89.80

Table 18: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 3,000. Bootstrap Repetitions: 500. Model for LS: $X_t{}'\beta_*$ and $X_t = (1, x_t)'$. DGP: $y_t = \frac{1}{2}x_t + (1 + n^{-1/2}\exp(x_t))u_t$ and $(x_t, u_t)' \sim$ IID $N(0, I_2)$.

Statistics \ n	100	200	400	600	800	1,000
$\widehat{\mathfrak{B}}_n^{(1)}$	66.67	75.73	78.77	78.27	73.50	70.23
$\widehat{\mathfrak{B}}_n^{(2)}$	74.83	83.37	86.07	86.50	83.13	82.37
$\widehat{\mathfrak{B}}_n^{(3)}$	73.00	81.93	84.73	85.63	82.00	81.73
$\widehat{\mathfrak{D}}_n^{(1)}$	64.30	73.00	75.67	75.57	70.70	68.10
$\widehat{\mathfrak{D}}_n^{(2)}$	75.20	83.33	86.07	86.57	83.20	82.33
$\widehat{\mathfrak{D}}_n^{(3)}$	69.57	80.20	83.33	84.27	80.53	80.17
$\widehat{\mathfrak{S}}_n^{(1)}$	36.93	59.17	65.07	67.77	63.97	63.23
$\widehat{\mathfrak{S}}_n^{(2)}$	73.17	82.07	85.00	85.70	82.00	81.73
$\widehat{\mathfrak{S}}_n^{(3)}$	69.43	80.27	83.47	84.20	80.60	80.10
$\widehat{\mathfrak{E}}_n^{(1)}$	75.37	83.43	86.20	86.50	83.27	82.40
$\widehat{\mathfrak{E}}_n^{(2)}$	72.57	81.53	83.93	84.27	80.27	80.53
$\widehat{\mathfrak{E}}_n^{(3)}$	69.70	80.13	83.27	84.20	80.53	80.23
$\widehat{\mathfrak{M}}_n^{(1)}$	54.10	69.07	73.30	74.43	69.33	66.77
$\widehat{\mathfrak{M}}_n^{(2)}$	64.77	77.80	81.97	82.43	77.77	78.13
$\widehat{\mathfrak{M}}_n^{(3)}$	64.77	77.80	81.97	82.43	77.77	78.13

Table 19: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 3,000. Bootstrap Repetitions: 500. Model for LS: $X_t{}'\beta_*$ and $X_t = (1, x_t)'$. DGP: $y_t = \frac{1}{2}n^{-1/2}x_t^4 + u_t$ and $(x_t, u_t)' \sim$ IID $N(0, I_2)$.

Statistics \ n	400	800	1,200	1,600	2,000	2,400
$\widehat{\mathfrak{B}}_n^{(1)}$	8.50	8.17	8.27	8.43	8.77	8.47
$\widehat{\mathfrak{B}}_n^{(2)}$	9.63	9.00	9.97	8.57	9.20	8.60
$\widehat{\mathfrak{B}}_n^{(3)}$	9.07	8.53	9.47	8.40	8.80	8.20
$\widehat{\mathfrak{D}}_n^{(1)}$	8.07	7.77	8.33	8.13	8.33	8.30
$\widehat{\mathfrak{D}}_n^{(2)}$	9.77	9.00	9.93	8.63	9.13	8.53
$\widehat{\mathfrak{D}}_n^{(3)}$	8.83	8.13	8.87	8.07	8.37	7.50
$\widehat{\mathfrak{S}}_n^{(1)}$	7.67	7.47	7.93	7.90	7.90	8.10
$\widehat{\mathfrak{S}}_n^{(2)}$	9.13	8.67	9.53	8.47	8.80	8.20
$\widehat{\mathfrak{S}}_n^{(3)}$	8.90	8.17	8.87	8.03	8.37	7.50
$\widehat{\mathfrak{E}}_n^{(1)}$	9.60	9.13	9.93	8.63	9.17	8.53
$\widehat{\mathfrak{E}}_n^{(2)}$	8.77	8.10	8.43	7.80	8.33	7.73
$\widehat{\mathfrak{E}}_n^{(3)}$	8.80	8.03	8.80	8.07	8.37	7.50
$\widehat{\mathfrak{M}}_n^{(1)}$	5.67	6.53	7.07	7.13	7.37	7.50
$\widehat{\mathfrak{M}}_n^{(2)}$	7.90	7.30	7.33	7.10	7.77	6.93
$\widehat{\mathfrak{M}}_n^{(3)}$	7.90	7.30	7.33	7.10	7.77	6.93

Table 20: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model for LS: $X_t{}'\beta_*$ and $X_t = (1, x_t)'$. DGP: $y_t = \frac{1}{2}x_t + u_t$, $u_t = \frac{1}{2}n^{-1/2}u_{t-1} + \varepsilon_t$, and $(x_t, \varepsilon_t)' \sim$ IID $N(0, I_2)$.

Statistics $\backslash$ n	50	100	200	300	400	500
$\widehat{\mathfrak{B}}_n^{(1)}$	2.78	4.16	3.90	4.50	4.58	4.88
$\widehat{\mathfrak{B}}_n^{(2)}$	3.00	3.98	4.00	3.76	4.74	5.26
$\widehat{\mathfrak{B}}_n^{(3)}$	3.36	4.40	4.28	4.22	4.80	5.44
$\widehat{\mathfrak{D}}_n^{(1)}$	3.26	4.74	4.32	4.80	4.96	4.98
$\widehat{\mathfrak{D}}_n^{(2)}$	2.84	3.80	3.98	3.72	4.82	5.28
$\widehat{\mathfrak{D}}_n^{(3)}$	4.22	5.04	4.58	4.76	4.98	5.52
$\widehat{\mathfrak{S}}_n^{(1)}$	4.42	5.86	4.62	5.08	4.98	5.08
$\widehat{\mathfrak{S}}_n^{(2)}$	3.06	4.22	4.26	4.14	4.80	5.44
$\widehat{\mathfrak{S}}_n^{(3)}$	4.26	5.26	4.52	4.82	4.98	5.54
$\widehat{\mathfrak{E}}_n^{(1)}$	2.64	3.82	4.02	3.74	4.84	5.32
$\widehat{\mathfrak{E}}_n^{(2)}$	3.64	4.46	4.52	4.36	4.84	5.28
$\widehat{\mathfrak{E}}_n^{(3)}$	4.22	5.06	4.54	4.82	4.94	5.58
$\widehat{\mathfrak{M}}_n^{(1)}$	6.68	7.66	5.54	6.10	5.48	5.60
$\widehat{\mathfrak{M}}_n^{(2)}$	6.24	6.50	5.06	5.22	5.34	5.62
$\widehat{\mathfrak{M}}_n^{(3)}$	6.24	6.50	5.06	5.22	5.34	5.62

Table 21: EMPIRICAL LEVELS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 5,000. Bootstrap Repetitions: 500. Model for TSLS: ${X_t}'\beta_*$ with $X_t = (1, x_t)'$ and IVs: $(1, z_t)'$. DGP: $y_t = \frac{1}{2}x_t + u_t$, $x_t := u_t + z_t$, and $(z_t, u_t)' \sim$ IID $N(0, I_2)$.

Statistics $\backslash$ n	100	500	1,000	1,500	2,000	2,500
$\widehat{\mathfrak{B}}_n^{(1)}$	28.45	61.70	79.80	87.65	92.45	94.65
$\widehat{\mathfrak{B}}_n^{(2)}$	50.35	89.35	95.75	97.75	98.90	98.95
$\widehat{\mathfrak{B}}_n^{(3)}$	65.75	94.30	97.80	99.05	99.45	99.55
$\widehat{\mathfrak{D}}_n^{(1)}$	41.35	83.85	94.15	96.90	98.45	98.70
$\widehat{\mathfrak{D}}_n^{(2)}$	44.20	85.95	94.00	97.20	98.05	98.55
$\widehat{\mathfrak{D}}_n^{(3)}$	70.00	95.60	98.10	99.25	99.55	99.60
$\widehat{\mathfrak{S}}_n^{(1)}$	24.45	46.80	62.00	71.40	79.25	81.85
$\widehat{\mathfrak{S}}_n^{(2)}$	52.20	91.65	96.35	98.35	99.15	99.20
$\widehat{\mathfrak{S}}_n^{(3)}$	69.00	95.15	97.85	99.05	99.50	99.50
$\widehat{\mathfrak{E}}_n^{(1)}$	37.00	77.85	90.05	94.60	95.90	97.30
$\widehat{\mathfrak{E}}_n^{(2)}$	61.40	93.65	97.35	98.65	99.35	99.50
$\widehat{\mathfrak{E}}_n^{(3)}$	70.15	95.70	98.05	99.20	99.50	99.60
$\widehat{\mathfrak{M}}_n^{(1)}$	44.50	78.50	90.10	95.20	96.65	98.35
$\widehat{\mathfrak{M}}_n^{(2)}$	57.85	88.75	94.65	97.55	98.45	99.15
$\widehat{\mathfrak{M}}_n^{(3)}$	57.85	88.75	94.65	97.55	98.45	99.15

Table 22: EMPIRICAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model for TSLS: ${X_t}'\beta_*$ with $X_t = (1, x_t)'$ and IVs: $(1, z_t)'$. DGP: $y_t = \frac{1}{2}x_t + (1 + \exp(x_t))u_t$, $x_t := u_t + z_t$, and $(z_t, u_t)' \sim$ IID $N(0, I_2)$.

Statistics \ n	100	200	400	600	800	1,000
$\widehat{\mathfrak{B}}_n^{(1)}$	65.95	89.15	97.30	99.60	99.95	100.0
$\widehat{\mathfrak{B}}_n^{(2)}$	63.75	84.95	95.10	98.85	99.55	100.0
$\widehat{\mathfrak{B}}_n^{(3)}$	59.40	81.55	94.00	98.30	99.40	100.0
$\widehat{\mathfrak{D}}_n^{(1)}$	63.00	86.50	95.90	99.35	99.75	100.0
$\widehat{\mathfrak{D}}_n^{(2)}$	64.45	85.75	95.25	99.00	99.60	100.0
$\widehat{\mathfrak{D}}_n^{(3)}$	49.50	72.10	90.50	96.55	98.70	99.45
$\widehat{\mathfrak{S}}_n^{(1)}$	31.30	56.45	79.90	91.90	96.65	97.70
$\widehat{\mathfrak{S}}_n^{(2)}$	64.70	85.75	95.15	98.85	99.65	100.0
$\widehat{\mathfrak{S}}_n^{(3)}$	50.45	73.50	90.95	96.75	98.80	99.60
$\widehat{\mathfrak{E}}_n^{(1)}$	66.05	86.75	95.80	99.10	99.70	100.0
$\widehat{\mathfrak{E}}_n^{(2)}$	59.00	80.15	93.20	97.80	99.05	99.90
$\widehat{\mathfrak{E}}_n^{(3)}$	48.60	71.05	90.20	96.45	98.45	99.30
$\widehat{\mathfrak{M}}_n^{(1)}$	22.60	51.70	81.70	93.75	97.65	98.80
$\widehat{\mathfrak{M}}_n^{(2)}$	24.90	51.50	80.90	92.95	96.80	98.55
$\widehat{\mathfrak{M}}_n^{(3)}$	24.90	51.50	80.90	92.95	96.80	98.55

Table 23: EMPIRICAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model for TSLS: ${X_t}'\beta_*$ with $X_t = (1, x_t)'$ and IVs: $(1, z_t)'$. DGP: $y_t = 1 + \frac{1}{2}x_t^4 + u_t$, $x_t := u_t + z_t$ and $(z_t, u_t)' \sim$ IID $N(0, I_2)$.

Statistics \ n	50	100	150	200	250	300
$\widehat{\mathfrak{B}}_n^{(1)}$	98.10	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{B}}_n^{(2)}$	99.70	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{B}}_n^{(3)}$	99.50	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(1)}$	98.50	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(2)}$	99.70	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{D}}_n^{(3)}$	99.15	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(1)}$	59.55	95.95	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(2)}$	99.60	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{S}}_n^{(3)}$	99.10	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(1)}$	99.70	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(2)}$	99.35	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{E}}_n^{(3)}$	99.20	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(1)}$	94.10	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(2)}$	98.70	100.0	100.0	100.0	100.0	100.0
$\widehat{\mathfrak{M}}_n^{(3)}$	98.70	100.0	100.0	100.0	100.0	100.0

Table 24: EMPIRICAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 2,000. Bootstrap Repetitions: 500. Model for TSLS: ${X_t}'\beta_*$ with $X_t = (1, x_t)'$ and IVs: $(1, z_t)'$. DGP: $y_t = \frac{1}{2} + u_t$, $x_t := z_t + \varepsilon_t$, $u_t = \frac{1}{2}u_{t-1} + \frac{1}{2}u_{t-2} + \varepsilon_t + \frac{1}{2}\varepsilon_{t-1}$, and $(z_t, \varepsilon_t)' \sim$ IID $N(0, I_2)$.

Statistics $\backslash$ n	500	600	700	800	900	1,000
$\widehat{\mathfrak{B}}_n^{(1)}$	35.20	38.53	38.63	38.73	42.23	42.40
$\widehat{\mathfrak{B}}_n^{(2)}$	69.33	73.70	73.67	76.20	80.40	81.93
$\widehat{\mathfrak{B}}_n^{(3)}$	70.60	74.00	74.50	76.50	80.07	81.80
$\widehat{\mathfrak{D}}_n^{(1)}$	34.70	37.40	37.53	37.17	39.90	39.53
$\widehat{\mathfrak{D}}_n^{(2)}$	68.97	73.27	73.27	75.93	80.33	81.70
$\widehat{\mathfrak{D}}_n^{(3)}$	71.37	74.40	74.93	76.93	79.87	81.73
$\widehat{\mathfrak{S}}_n^{(1)}$	8.17	10.47	11.43	13.27	15.90	16.30
$\widehat{\mathfrak{S}}_n^{(2)}$	69.47	73.13	73.77	76.13	79.77	81.67
$\widehat{\mathfrak{S}}_n^{(3)}$	71.23	74.13	74.73	76.80	79.70	81.50
$\widehat{\mathfrak{E}}_n^{(1)}$	68.37	72.90	73.00	75.43	80.00	81.47
$\widehat{\mathfrak{E}}_n^{(2)}$	72.50	76.27	76.53	78.80	82.07	83.67
$\widehat{\mathfrak{E}}_n^{(3)}$	71.57	74.63	75.07	76.97	80.17	82.13
$\widehat{\mathfrak{M}}_n^{(1)}$	25.73	29.30	29.80	31.87	34.80	35.53
$\widehat{\mathfrak{M}}_n^{(2)}$	68.50	72.93	73.63	75.97	79.40	81.73
$\widehat{\mathfrak{M}}_n^{(3)}$	68.50	72.93	73.63	75.97	79.40	81.73

Table 25: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 3,000. Bootstrap Repetitions: 500. Model for TSLS: ${X_t}'\beta_*$ with $X_t = (1, x_t)'$ and IVs: $(1, z_t)'$. DGP: $y_t = \frac{1}{2}x_t + (1 + n^{-1/2}\exp(x_t))u_t$, $x_t := u_t + z_t$, and $(z_t, u_t)' \sim$ IID $N(0, I_2)$.

Statistics $\backslash$ n	500	600	700	800	900	1,000
$\widehat{\mathfrak{B}}_n^{(1)}$	78.50	81.23	78.90	80.93	81.73	81.40
$\widehat{\mathfrak{B}}_n^{(2)}$	96.63	97.97	97.83	98.80	99.03	99.33
$\widehat{\mathfrak{B}}_n^{(3)}$	96.57	97.70	97.77	98.57	99.03	99.30
$\widehat{\mathfrak{D}}_n^{(1)}$	78.40	80.87	78.90	80.23	81.13	81.07
$\widehat{\mathfrak{D}}_n^{(2)}$	96.67	97.93	97.90	98.80	99.03	99.37
$\widehat{\mathfrak{D}}_n^{(3)}$	96.23	97.60	97.53	98.40	98.97	99.07
$\widehat{\mathfrak{S}}_n^{(1)}$	45.97	48.67	46.37	47.30	49.23	49.80
$\widehat{\mathfrak{S}}_n^{(2)}$	96.57	97.70	97.77	98.63	99.00	99.33
$\widehat{\mathfrak{S}}_n^{(3)}$	96.13	97.60	97.53	98.40	98.93	99.07
$\widehat{\mathfrak{E}}_n^{(1)}$	96.57	97.97	98.00	98.83	99.07	99.37
$\widehat{\mathfrak{E}}_n^{(2)}$	96.70	97.87	97.87	98.63	99.10	99.33
$\widehat{\mathfrak{E}}_n^{(3)}$	96.20	97.53	97.50	98.37	98.97	99.07
$\widehat{\mathfrak{M}}_n^{(1)}$	72.57	75.87	74.23	76.80	78.80	78.60
$\widehat{\mathfrak{M}}_n^{(2)}$	95.33	97.13	97.33	98.17	98.83	99.03
$\widehat{\mathfrak{M}}_n^{(3)}$	95.33	97.13	97.33	98.17	98.83	99.03

Table 26: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 3,000. Bootstrap Repetitions: 500. Model for TSLS: $(1, z_t)'$. DGP: $y_t = 1 + \frac{1}{2}n^{-1/2}x_t^4 + u_t$, $x_t := u_t + z_t$, and $(z_t, u_t)' \sim$ IID $N(0, I_2)$.

Statistics $\backslash$ n	700	800	900	1,000	1,100	1,200
$\widehat{\mathfrak{B}}_n^{(1)}$	8.80	9.33	8.97	9.07	8.20	8.17
$\widehat{\mathfrak{B}}_n^{(2)}$	8.87	10.20	8.70	9.50	9.70	9.10
$\widehat{\mathfrak{B}}_n^{(3)}$	8.47	9.43	8.30	8.80	9.40	9.00
$\widehat{\mathfrak{D}}_n^{(1)}$	8.50	8.87	8.47	8.33	8.00	7.80
$\widehat{\mathfrak{D}}_n^{(2)}$	8.77	10.27	8.80	9.57	9.67	9.10
$\widehat{\mathfrak{D}}_n^{(3)}$	8.10	9.00	7.63	8.10	9.00	8.23
$\widehat{\mathfrak{S}}_n^{(1)}$	8.20	8.50	8.07	7.93	7.70	7.60
$\widehat{\mathfrak{S}}_n^{(2)}$	8.47	9.53	8.30	8.83	9.40	9.10
$\widehat{\mathfrak{S}}_n^{(3)}$	8.07	9.13	7.67	8.10	8.93	8.30
$\widehat{\mathfrak{E}}_n^{(1)}$	8.90	10.23	8.83	9.60	9.70	9.10
$\widehat{\mathfrak{E}}_n^{(2)}$	8.43	9.33	7.53	8.27	9.17	8.60
$\widehat{\mathfrak{E}}_n^{(3)}$	8.17	9.00	7.60	8.10	8.87	8.23
$\widehat{\mathfrak{M}}_n^{(1)}$	6.50	7.73	6.77	7.43	6.50	6.80
$\widehat{\mathfrak{M}}_n^{(2)}$	7.03	7.90	6.63	7.00	8.03	7.67
$\widehat{\mathfrak{M}}_n^{(3)}$	7.03	7.90	6.63	7.00	8.03	7.67

Table 27: EMPIRICAL LOCAL POWERS OF THE TEST STATISTICS (LEVEL OF SIGNIFICANCE: 5%). Repetitions: 3,000. Bootstrap Repetitions: 500. Model for TSLS: ${X_t}'\beta_*$ with $X_t = (1, x_t)'$ and IVs: $(1, z_t)'$. DGP: $y_t = \frac{1}{2} + u_t$, $x_t := z_t + \varepsilon_t$, $u_t = \frac{1}{2}n^{-1/2}u_{t-1} + \frac{1}{2}n^{-1/2}u_{t-2} + \varepsilon_t + \frac{1}{2}n^{-1/2}\varepsilon_{t-1}$, and $(z_t, \varepsilon_t)' \sim$ IID $N(0, I_2)$.

Statistics \ Models	Probit Models		Logit Models		Scobit Models	
	w/ Products	w/o Products	w/ Products	w/o Products	w/ Products	w/o Products
Constant	-2.7431	-2.5229	-4.4129	-4.0727	-5.3465	-4.4062
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Closing Date	0.0006	-0.0078	-0.0001	-0.0132	-0.0024	-0.0217
	(0.8685)	(0.0000)	(0.9832)	(0.0000)	(0.7997)	(0.0000)
Education	0.2645	0.1818	0.3585	0.2426	0.3494	0.2094
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0041)	(0.0000)
$Education^2$	0.0050	0.0123	0.0192	0.0282	0.0663	0.0711
	(0.2433)	(0.0000)	(0.0133)	(0.0000)	(0.0000)	(0.0000)
Age	0.0696	0.0697	0.1141	0.1142	0.1837	0.1813
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Age^2	-0.0005	-0.0005	-0.0008	-0.0008	-0.0012	-0.0012
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
South	-0.1154	-0.1159	-0.1897	-0.1904	-0.2975	-0.2956
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Gubernatorial Election	0.0034	0.0034	0.0048	0.0052	-0.0014	-0.0000
	(0.7670)	(0.7666)	(0.8012)	(0.7853)	(0.9637)	(0.9998)
Closing Date $\times$ *Education*	-0.0031		-0.0044		-0.0052	
	(0.0399)		(0.0956)		(0.2283)	
Closing Date $\times$ $Education^2$	0.0002		0.0003		0.0002	
	(0.0075)		(0.2219)		(0.6378)	
α_*					0.4105	0.4194
					(0.0000)	(0.0000)
Log-Likelihood	-55,815.28	-55,818.03	-55,774.55	-55,777.67	-55,725.09	-55,730.63
$\widehat{\mathfrak{B}}_n^{(1)}$	589.90	272.79	509.42	224.26	812.99	651.59
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{B}}_n^{(2)}$	3,007.43	1,875.64	2,849.08	1,803.54	5,673.81	5,337.46
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{B}}_n^{(3)}$	2,925.71	1,834.28	2,775.50	1,766.55	5,490.45	5,162.59
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{D}}_n^{(1)}$	553.60	254.71	476.66	208.09	737.84	582.65
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{D}}_n^{(2)}$	2,951.45	1,831.22	2,801.19	1,765.28	5,482.02	5,118.08
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{D}}_n^{(3)}$	2,797.14	1,753.71	2,662.08	1,695.95	5,148.35	4,805.33
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{S}}_n^{(1)}$	512.75	234.03	439.87	189.60	646.16	495.21
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{S}}_n^{(2)}$	2,813.76	1,745.44	2,679.71	1,690.03	5,106.86	4,723.83
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{S}}_n^{(3)}$	2,741.16	1,709.28	2,614.19	1,657.70	4,956.56	4,585.95
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{E}}_n^{(1)}$	2,895.48	1,786.79	2,753.29	1,727.02	5,290.23	4,898.70
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{E}}_n^{(2)}$	2,869.73	1,789.13	2,727.61	1,728.29	5,298.66	4,898.70
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{E}}_n^{(3)}$	2,853.11	1,798.13	2,709.98	1,734.21	5,340.14	5,024.71
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{M}}_n^{(1)}$	3,007.43	1,875.64	2,849.08	1,803.54	5,673.81	5,337.46
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{M}}_n^{(2)}$	589.90	272.79	509.42	224.26	812.99	651.59
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{\mathfrak{M}}_n^{(3)}$	3,007.43	1,875.64	2,849.08	1,803.54	5,673.81	5,337.46
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Table 28: EMPIRICAL MODEL ESTIMATIONS AND INFERENCES OF THE TEST STATISTICS. The figures in parentheses stand for the p-values. The p-values of the parameter estimates are computed by White's (1980) heteroskedasticity consistent standard errors, and the p-values of the test statistics are obtained by implementing the parametric bootstrap. The sample size is 99,676.