

SECOND-ORDER REFINEMENTS FOR t -RATIOS WITH MANY INSTRUMENTS

YUKITOSHI MATSUSHITA AND TAISUKE OTSU

ABSTRACT. This paper studies second-order properties of the many instruments robust t -ratios based on the limited information maximum likelihood and Fuller estimators for instrumental variable regression models under the many instruments asymptotics, where the number of instruments may increase proportionally with the sample size n , and proposes second-order refinements to the t -ratios to improve the size and power properties. Based on asymptotic expansions of the null and non-null distributions of the t -ratios derived under the many instruments asymptotics, we show that the second order terms of those expansions may have non-trivial impacts on the size as well as the power properties. Furthermore, we propose adjusted t -ratios whose approximation errors for the null rejection probabilities are of order $O(n^{-1})$ in contrast to the ones for the unadjusted t -ratios of order $O(n^{-1/2})$, and show that these adjustments induce some desirable power properties in terms of the local maximinity.

1. INTRODUCTION

Instrumental variable regression is one of the most widely used methods in empirical economic analysis. Particularly in microeconomic applications, researchers often use many instrumental variables to improve efficiency of estimators and associated inference methods (e.g., Angrist and Krueger, 1991). However, in such cases, it has been found that approximate distributions of the estimators and statistics based on the conventional asymptotic theory can be inaccurate. For example, the two stage least squares (TSLS) estimator tends to have large bias. Although the limited information maximum likelihood (LIML) estimator is less biased, its distribution is often more dispersed than the limiting distribution based on the conventional asymptotics (see, e.g., Anderson, Kunitomo and Sawa, 1982, and Anderson, Kunitomo and Matsushita, 2010, 2011).

In order to give more accurate approximations under many instruments, Kunitomo (1980, 1982) and Morimune (1983) considered a limiting sequence where the number of instruments K is allowed to increase proportionally with the sample size n (called the large- K asymptotics), and derived the limiting distribution of the LIML estimator when the disturbances are normal and there is one endogenous regressor in the regression model. Bekker (1994) derived multivariate first-order approximations to the distributions of several estimators under the large- K asymptotics with the normal disturbances, while Hansen, Hausman and Newey (2008), van Hasselt (2010), and Anderson, Kunitomo and Matsushita (2010) extended those results to non-normal cases. Hansen, Hausman and Newey (2012) considered a more general model, where the reduced form may be nonlinear and the disturbances may be heteroskedastic, and suggested to use the

We are grateful to Naoto Kunitomo for helpful comments. Matsushita acknowledges financial support from the JSPS KAKENHI (18K01541).

t -ratios for the LIML or Fuller (1977) estimators using the many instruments and heteroskedasticity robust standard errors. It should be noted that existing works on the large- K asymptotics mostly focus on the *first-order* asymptotic properties of the estimators and test statistics and their higher-order properties are largely unexplored.¹

This paper studies second-order properties of the t -ratios based on the LIML and Fuller estimators for instrumental variable regression models under the large- K asymptotics, and proposes higher-order refinements to the t -ratios to improve the size and power properties. To explore the finite sample properties of the t -ratios with many instruments, asymptotic expansions of the null and non-null distributions of the large- K robust t -ratios associated with the LIML and Fuller estimators are derived under the large- K asymptotics. Moreover, to assess the effects of variance estimation, we derive asymptotic expansions of the LIML and Fuller estimators under the large- K asymptotics. Based on these asymptotic expansions, it is shown that the finite sample distributions of the large- K t -ratios can be quite different from those of the corresponding standardized estimators although they have the same asymptotic normal distribution under the large- K asymptotics. In fact, the absolute values of the second-order terms of the asymptotic expansions of the standardized LIML and Fuller estimators and their large- K t -ratios are the same but have opposite signs. Also the null distributions of the large- K t -ratios can be skewed and largely deviated from the standard normal distribution. For two-sided testing, although the second-order terms cancel out under the null distribution, we find that these second-order terms may have non-trivial impacts on the power properties. Based on these expansions, we propose adjusted t -ratios whose approximation errors for the null rejection probabilities are of order $O(n^{-1})$ in contrast to the ones for the unadjusted t -ratios of order $O(n^{-1/2})$. Furthermore, we show that these adjustments induce some desirable power properties in terms of the local maximinity. Finally, these findings are illustrated by some simulation studies.

This paper also contributes to the literature of the asymptotic higher-order expansion approach, which has been developed extensively to investigate the finite sample properties of econometric methods (see, e.g., Rothenberg, 1984, and Ullah, 2004, for an overview). For simultaneous equation models, it has been used to give more accurate approximations to distributions of estimators and test statistics, or to compare their higher-order properties, see, Anderson (1974), Sargan (1975), Phillips (1977), Rothenberg (1988), Fujikoshi *et al.* (1982), Morimune (1989), to name a few. Our main contribution in this context is that, to the best of our knowledge, this is the first paper which investigates higher-order properties of testing methods under the large- K asymptotics.

The paper is organized as follows. In Section 2, we introduce our setup and estimators (Section 2.1) and define the the large- K robust t -ratios (Section 2.2). Section 3 presents our main results: asymptotic expansions of the large- K t -ratios under the null hypothesis (Section 3.1), asymptotic expansions under the local alternatives and adjusted t -ratios (Section 3.2), and a discussion for

¹Exceptions are Kunitomo (1980) and Morimune (1983), which established asymptotic expansions for the distributions of the LIML and k -class estimators, respectively, for the case of the normal disturbances and one endogenous regressor.

the case of heteroskedastic and non-elliptically contoured disturbances (Section 3.3). Section 4 illustrates our findings by some simulation results.

2. SETUP, ESTIMATORS AND TEST STATISTICS

2.1. Setup and estimators. We first introduce our basic setup and some estimators. Consider a single structural equation

$$y_1 = Y_2\beta + Z_1\gamma + u, \quad (1)$$

where y_1 is an n -vector of dependent variables, Y_2 is an $n \times G_1$ matrix of endogenous regressors, Z_1 is an $n \times K_1$ matrix of exogenous regressors, u is an n -vector of error terms, and β and γ are G_1 and K_1 dimensional vectors of unknown parameters, respectively. We assume that (1) is the first equation in a simultaneous system of $G_1 + 1$ linear stochastic equations relating $G_1 + 1$ endogenous variables $Y = (y_1, Y_2)$, and $K_1 + K_2$ exogenous variables $Z = (Z_1, Z_2)$, where Z_2 is an $n \times K_2$ matrix of instrumental variables for (1). In this paper, we consider the case, where Z is nonrandom (which may be viewed as conditioning on Z). Let $K = K_1 + K_2$. The reduced form of Y is defined as

$$Y = Z\Pi + V = (Z_1, Z_2) \begin{pmatrix} \Pi_1 \\ \Pi_2 \end{pmatrix} + (v_1, V_2), \quad (2)$$

where $\Pi_1 = (\pi_{11}, \Pi_{12})$ and $\Pi_2 = (\pi_{21}, \Pi_{22})$ are $K_1 \times (1 + G_1)$ and $K_2 \times (1 + G_1)$ matrices, respectively, of the reduced form coefficients, and (v_1, V_2) is an $n \times (1 + G_1)$ matrix of disturbances. We assume that the rows of V are independently distributed, and each row of V has mean 0 and nonsingular covariance matrix $\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \Omega_{22} \end{pmatrix}$. In order to relate (1) and (2), we postmultiply (2) by $(1, -\beta)'$. Then it can be written as $u = v_1 - V_2\beta$, $\gamma = \pi_{11} - \Pi_{12}\beta$, and $\pi_{21} = \Pi_{22}\beta$. Assume that the matrix Π_{22} is of rank G_1 . The components of u are independently distributed with mean 0 and variance $\sigma^2 = \omega_{11} - 2\beta'\omega_{21} + \beta'\Omega_{22}\beta$. We mainly focus on the case where the error term u is homoskedastic, and discuss heteroskedastic errors in Section 3.3.

Let $P_F = F(F'F)^{-1}F'$ and $M_F = I - P_F$ for a full column rank matrix F . The k -class estimator is defined as

$$\begin{pmatrix} \hat{\beta}_k \\ \hat{\gamma}_k \end{pmatrix} = \begin{bmatrix} Y_2'Y_2 - kY_2'M_ZY_2 & Y_2'Z_1 \\ Z_1'Y_2 & Z_1'Z_1 \end{bmatrix}^{-1} \begin{pmatrix} Y_2'(I - kM_Z)y_1 \\ Z_1'y_1 \end{pmatrix}. \quad (3)$$

This estimator covers (i) OLS ($k = 0$), (ii) TSLS ($k = 1$), (iii) LIML ($k = \hat{\lambda}$), and (iv) Fuller (1977) ($k = \hat{\lambda} - a/(n - K)$ for some $a > 0$) as special cases, where $\hat{\lambda}$ is the smallest root of

$$\left| \begin{pmatrix} Y' \\ Z_1' \end{pmatrix} P_Z(Y, Z_1) - \lambda \begin{pmatrix} Y' \\ Z_1' \end{pmatrix} M_Z(Y, Z_1) \right| = 0.$$

Under the conventional asymptotics, where the number of instruments K is fixed, both the LIML and TSLS estimators are consistent and follow the same limiting normal distribution. However, it has been known that the exact distributions of these estimators can be quite different from the normal distribution. When K is large, the TSLS estimator can be severely biased (see, e.g., Anderson, Kunitomo and Matsushita, 2010). On the other hand, the distribution

of the LIML estimator is more dispersed than the limiting distribution under the conventional asymptotics. The Fuller estimator has moments of all orders, and is known to have good finite sample properties in some situations (e.g., Hahn, Hausman and Kuersteiner, 2004, and Hansen, Hausman and Newey, 2008).

Bekker (1994) pointed out that the large- K asymptotic theory, where K may grow proportionally to n , may be suited better to applications, even when the number of instruments is not large. Under the large- K asymptotics, the (first-order) asymptotic distributions of the LIML and TSLS estimators are rather different. The LIML estimator is consistent and asymptotically normal while the TSLS estimator loses consistency. Also the LIML estimator attains the asymptotic efficiency bound when the number of instruments is large (see, Kunitomo, 1982, Chioda and Jansson, 2009, and Anderson, Kunitomo and Matsushita, 2010).

In contrast, this paper is concerned with higher-order properties of the t -ratios for testing parameter hypotheses under the large- K asymptotics, which will be introduced in the next subsection.

2.2. Many instruments robust t -statistics. Let ι be a $(G_1 + K_1)$ -vector of zeros, apart from its j -th element which is unity. We are interested in testing the null hypothesis

$$H_0 : \iota' \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = 0,$$

i.e., the j -th coefficient in (1) is zero, against the one-sided or two-sided alternative hypothesis.

Let $D_2 = \begin{pmatrix} \Pi_{12} & I_{K_1} \\ \Pi_{22} & 0 \end{pmatrix}$. Throughout the paper, we assume

$$\frac{1}{n} D_2' Z' Z D_2 = Q + O(n^{-1}) \quad \text{for some positive definite } Q, \quad (4)$$

and $K = K_n$ satisfies

$$\frac{K}{n} = c + O(n^{-1}) \quad \text{for some } c \in [0, 1). \quad (5)$$

Note that the number of instruments K can grow at either the same rate ($c > 0$) as the sample size or at slower rate ($c = 0$), where the latter includes the conventional fixed- K asymptotics as a special case.

Under the sequence (5) and certain regularity conditions, the results in Anderson, Kunitomo and Matsushita (2010) can be adapted to derive the limiting distributions of the LIML estimator $(\hat{\beta}'_{LI}, \hat{\gamma}'_{LI})'$ and Fuller estimator $(\hat{\beta}'_F, \hat{\gamma}'_F)'$ as

$$\sqrt{n} \begin{pmatrix} \hat{\beta}_{LI} - \beta \\ \hat{\gamma}_{LI} - \gamma \end{pmatrix} \xrightarrow{d} N(0, \Psi^*), \quad \sqrt{n} \begin{pmatrix} \hat{\beta}_F - \beta \\ \hat{\gamma}_F - \gamma \end{pmatrix} \xrightarrow{d} N(0, \Psi^*),$$

where

$$\Psi^* = \sigma^2 Q^{-1} + \frac{c}{1-c} Q^{-1} \left[\begin{pmatrix} \Omega_{22} \sigma^2 & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^4 \right] Q^{-1} + Q^{-1} [(\Xi_3 + \Xi_3') + \eta \Xi_4] Q^{-1}. \quad (6)$$

For Ψ^* , we use the following notation

$$\begin{aligned} q_2 &= \frac{1}{\sigma^2}(\omega'_{21} - \beta'\Omega_{22}, 0)', & \Xi_3 &= \frac{1}{1-c} \lim_{n \rightarrow \infty} D_2' \frac{1}{n} \sum_{i=1}^n z_i(p_{ii} - c)E[u_i^2 w'_{2i}], \\ \eta &= \frac{1}{(1-c)^2} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (p_{ii} - c)^2, & \Xi_4 &= E[u_i^2 w_{2i} w'_{2i}] - \sigma^2 E[w_{2i} w'_{2i}], \end{aligned}$$

where $w_{2i} = (v'_{2i}, 0) - u_i q_2$ and $p_{ii} = z'_i(Z'Z)^{-1}z_i$. Compared to the conventional variance formula $\sigma^2 Q^{-1}$, there are two additional terms in Ψ^* due to the large- K asymptotics, which vanish when the number of instruments grows at a slower rate (i.e., $c = 0$).

The focus of this paper is to investigate higher-order properties of the t -tests for H_0 under the large- K asymptotics (5). To this end, we introduce an additional assumption, which greatly simplifies our higher-order analysis in the following section. Let $EC(\Omega)$ be the class of elliptically contoured distributions.² This class contains many important distributions including the multivariate normal and t distributions and the uniform distribution on the sphere. Hereafter, we assume

$$\text{the distribution of the rows of } V \text{ belongs to } EC(\Omega). \quad (7)$$

Under (7), the asymptotic variance Ψ^* simplifies to (see, Anderson, Kunitomo and Matsushita, 2010)

$$\Psi = \sigma^2 Q^{-1} + \left(\frac{c}{1-c} + \eta\kappa \right) Q^{-1} \left[\begin{pmatrix} \Omega_{22}\sigma^2 & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^4 \right] Q^{-1}, \quad (8)$$

where $\kappa = (E[u_i^4]/\sigma^4 - 3)/3$. Note that Ψ is identical to Bekker's (1994) variance when the error terms are normally distributed. By taking the sample counterparts, a consistent estimator of Ψ is given by

$$\begin{aligned} \hat{\Psi} &= \hat{\sigma}^2 \hat{Q}^{-1} \\ &+ \left(\frac{K}{n-K} + \hat{\eta}\hat{\kappa} \right) \hat{Q}^{-1} \begin{pmatrix} \frac{1}{n-K} Y_2' M_Z Y_2 \hat{\sigma}^2 - \frac{1}{(n-K)^2} Y_2' M_Z Y \hat{b} \hat{b}' Y' M_Z Y_2 & 0 \\ 0 & 0 \end{pmatrix} \hat{Q}^{-1}, \quad (9) \end{aligned}$$

where

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{n-K} \hat{b}' Y' M_Z Y \hat{b}, & \hat{b} &= (1, -\hat{\beta})', & \hat{Q} &= \frac{1}{n} \begin{pmatrix} Y_2' P_Z Y_2 - \lambda Y_2' M_Z Y_2 & Y_2' Z_1 \\ Z_1' Y_2 & Z_1' Z_1 \end{pmatrix}, \\ \hat{\eta} &= \left(\frac{n}{n-K} \right)^2 \frac{1}{n} \sum_{i=1}^n \left(p_{ii} - \frac{K}{n} \right)^2, & \hat{\kappa} &= \frac{1}{3} \left\{ \frac{1}{\hat{\sigma}^4} \frac{1}{n} \sum_{i=1}^n (y_{1i} - y'_{2i} \hat{\beta} - z'_{1i} \hat{\gamma})^4 - 3 \right\}, \end{aligned}$$

and $(\hat{\beta}, \hat{\gamma}, \lambda) = (\hat{\beta}_{LI}, \hat{\gamma}_{LI}, \hat{\lambda})$ for the LIML estimator, or $(\hat{\beta}_F, \hat{\gamma}_F, \hat{\lambda} - a/(n-k))$ for the Fuller estimator.

Based on this variance estimator, the large- K t -ratio for testing H_0 is given by

$$t_K = \frac{1}{\sqrt{\hat{\Psi}_j}} l' \sqrt{n} \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix}, \quad (10)$$

²If the characteristic function of a random vector X has the form of $\phi(\sum_{i=1}^n t' \Omega t)$ for some function $\phi(\cdot)$ and positive definite Ω , then we say that X is distributed according to an elliptically contoured distribution. See Anderson (2003, Section 2.7).

where $\hat{\Psi}_j$ is the j -th diagonal element of $\hat{\Psi}$. In the next section, we study higher order properties of the t -test based on this statistic under the large- K asymptotics.

3. MAIN RESULTS

3.1. Asymptotic expansions under null hypothesis. Let t_K^{LI} and t_K^F be the large- K t -ratios defined in (10) based on the LIML and Fuller estimators, respectively. We first present asymptotic expansions of the null distributions of these t -ratios under the large- K asymptotics in (5). Let $\Phi(\cdot)$ and $\phi(\cdot)$ be the standard normal distribution and density functions, respectively, and Ψ_j be the j -th diagonal element of Ψ .

Theorem 1. *Consider the setup of Section 2. Suppose that (4) and (7) hold true, and the rows of V have finite 8th moments. Then the asymptotic expansions of the null distributions of the large- K t -ratios t_K^{LI} and t_K^F under the large- K asymptotics (5) are given by*

$$\begin{aligned} P\{t_K^{LI} \leq \tau\} &= \Phi(\tau) - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\Psi_j}} \tau^2 (\iota' \Psi q_2) \phi(\tau) + O(n^{-1}), \\ P\{t_K^F \leq \tau\} &= \Phi(\tau) - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\Psi_j}} \left[\tau^2 \iota' \Psi + \frac{a}{1-c} \sigma^2 \iota' Q^{-1} \right] q_2 \phi(\tau) + O(n^{-1}), \end{aligned}$$

for each $\tau \in \mathbb{R}$, respectively.

This theorem says that (i) the approximation errors in rejection probability of both tests are of order $O(n^{-1/2})$, (ii) the approximation errors become larger as the degree of endogeneity $q_2 = E[v_{2i}u_i]/\sigma^2$ increases, and (iii) the approximation error of t_K^F is not always smaller than that of t_K^{LI} . In particular, if the number of endogenous regressors is one (i.e., $G_1 = 1$), the absolute value of the second term of t_K^F is always larger than that of t_K^{LI} unless $q_2 = 0$. The last point implies that improvement of the Fuller estimator over the LIML estimator does not necessarily imply improvement of the size property of the t -test under the large- K asymptotics.

It should be noted that the finite sample distributions of the t -ratios and corresponding estimators may be different due to estimation of the asymptotic variances in the denominators of t -ratios. In fact, we can derive the asymptotic expansions of the distributions of the LIML and Fuller estimators as follows. Let f_{LI} and f_F be the density functions of $\sqrt{n}((\hat{\beta}_{LI} - \beta)', (\hat{\gamma}_{LI} - \gamma)')'$ and $\sqrt{n}((\hat{\beta}_F - \beta)', (\hat{\gamma}_F - \gamma)')'$, respectively, and ϕ_Ψ be the density function of $N(0, \Psi)$.

Theorem 2. *Consider the setup of Section 2. Suppose that the assumptions in Theorem 1 hold true. The asymptotic expansions of f_{LI} and f_F under the large- K asymptotics (5) are given by*

$$\begin{aligned} f_{LI}(\xi) &= \phi_\Psi(\xi) \left\{ 1 + \frac{1}{\sqrt{n}} (q_2' \xi) (G_1 + K_1 + 1 - \xi' \Psi \xi) \right\} + O(n^{-1}), \\ f_F(\xi) &= \phi_\Psi(\xi) \left\{ 1 + \frac{1}{\sqrt{n}} \left[(q_2' \xi) (G_1 + K_1 + 1 - \xi' \Psi \xi) + \frac{a}{1-c} q_2' \sigma^2 Q^{-1} \Psi \xi \right] \right\} + O(n^{-1}), \end{aligned}$$

for each $\xi \in \mathbb{R}^{G_1+K_1}$, respectively.

We note that the expansion for the LIML estimator is derived based on Fujikoshi *et al.* (1982), but the expansion for the Fuller estimator is new in the literature. Based on Theorem 2, we can

see that under H_0 ,

$$P \left\{ \frac{1}{\sqrt{\Psi_j}} \iota' \sqrt{n} \begin{pmatrix} \hat{\beta}_{LI} \\ \hat{\gamma}_{LI} \end{pmatrix} \leq \tau \right\} = \Phi(\tau) + \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\Psi_j}} \tau^2 (\iota' \Psi q_2) \phi(\tau) + O(n^{-1}),$$

$$P \left\{ \frac{1}{\sqrt{\Psi_j}} \iota' \sqrt{n} \begin{pmatrix} \hat{\beta}_F \\ \hat{\gamma}_F \end{pmatrix} \leq \tau \right\} = \Phi(\tau) + \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\Psi_j}} \left[\tau^2 \iota' \Psi + \frac{a}{1-c} \sigma^2 \iota' Q^{-1} \right] q_2 \phi(\tau) + O(n^{-1}),$$

for each $\tau \in \mathbb{R}$. Comparing with Theorem 1, we can see that the distributions of the LIML and Fuller estimators and large- K t -ratio are distorted in opposite directions.³

3.2. Asymptotic expansions under local alternative hypothesis and adjusted t -ratios.

In order to investigate power properties of the large- K t -ratios, we now derive asymptotic expansions of their distributions under the local alternative hypothesis:

$$H_{1n} : \iota' \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \frac{1}{\sqrt{n}} \iota' \zeta,$$

for some $\zeta \in \mathbb{R}^{G_1+K_1}$. Let $\Phi_{\zeta}^{-}(\tau) = 1 - \{\Phi_{\zeta}(\tau) - \Phi_{\zeta}(-\tau)\}$, where Φ_{ζ} and ϕ_{ζ} are the cumulative distribution and density functions of $N(\Psi_j^{-1/2} \iota' \zeta, 1)$, respectively.

Theorem 3. *Consider the setup of Section 2. Suppose that the assumptions in Theorem 1 hold true. Then asymptotic expansions of the power functions of the one-sided and two-sided large- K t -tests under H_{1n} are given by*

$$P\{t_K \geq \tau\} = 1 - \Phi_{\zeta}(\tau) + \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\Psi_j}} \left[(\iota' \Psi q_2) \{\tau^2 - (\iota' \zeta / \sqrt{\Psi_j})^2\} + \frac{a\sigma^2}{1-c} \iota' Q^{-1} q_2 \right] \phi_{\zeta}(\tau) + O(n^{-1}),$$

$$P\{|t_K| \geq \tau\} = \Phi_{\zeta}^{-}(\tau) + \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\Psi_j}} \left[(\iota' \Psi q_2) \{\tau^2 - (\iota' \zeta / \sqrt{\Psi_j})^2\} + \frac{a\sigma^2}{1-c} \iota' Q^{-1} q_2 \right] \{\phi_{\zeta}(\tau) - \phi_{\zeta}(-\tau)\} + O(n^{-1}),$$

for each $\tau \in \mathbb{R}$, respectively. Here $a = 0$ for t_K^{LI} and $a > 0$ for t_K^F .

This theorem says that the large- K t -tests are locally biased up to the order $O(n^{-1/2})$ unless the degree of endogeneity q_2 is zero because $\frac{\partial}{\partial \zeta} P\{t_K \geq \tau\} = O(n^{-1/2})$ and $\frac{\partial}{\partial \zeta} P\{|t_K| \geq \tau\} = O(n^{-1/2})$ unless $q_2 = 0$ (Rao, 1973, pp.454).

Based on the asymptotic expansions in Theorems 1 and 3, we propose a simple adjustment to the t -ratio, which does not include the term of order $O(n^{-1/2})$ in the expansion of the null distribution:

$$t_K^{adj} = t_K - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\hat{\Psi}_j}} \left\{ (\iota' \hat{\Psi} \hat{q}_2) \min\{t_K^2, n^\varepsilon\} + \frac{an}{n-K} \hat{\sigma}^2 \iota' \hat{Q}^{-1} \hat{q}_2 \right\}, \quad (11)$$

³Bekker (1994) provided a skewed approximation for the distribution of the LIML estimator, not its t -ratio. Because the directions of the skewness of the distributions of the LIML estimator and the t -ratio under the null are opposite, Bekker's skewed approximation might make the size property of the test even worse.

for some $\varepsilon \in (0, 1/2)$. The trimming term, n^ε , has no impact on the second-order properties under the null and local alternative hypotheses. This term is introduced to guarantee consistency of the t -test based on t_K^{adj} under fixed alternatives. In our simulation study below, we set as $\varepsilon = 0.1$.

The asymptotic expansion of the null distribution of the adjusted statistic is obtained as follows.

Theorem 4. *Consider the setup of Section 2. Suppose that the assumptions in Theorem 1 hold true. Then the asymptotic expansion of the null distribution of the adjusted t -ratio (11) under the large- K asymptotics (5) is given by*

$$P\{t_K^{adj} \leq \tau\} = \Phi(\tau) + O(n^{-1}),$$

for each $\tau \in \mathbb{R}$.

This theorem says that (i) the approximation error in rejection probability of the one-sided test using the adjusted t -ratio t_K^{adj} is of order $O(n^{-1})$, which improves the (unadjusted) large- K t -test whose error is $O(n^{-1/2})$, and (ii) the two-sided test with the adjusted t -ratio have the approximation error $O(n^{-1})$, which is the same as that of the (unadjusted) large- K t -test.

Next we present the local power properties of the adjusted t -tests.

Theorem 5. *Under the same assumption in Theorem 1, the asymptotic expansions of the power functions of the one-sided and two-sided adjusted t -tests under H_{1n} are given by*

$$P\{t_K^{adj} \geq \tau\} = 1 - \Phi_\zeta(\tau) + \frac{1}{\sqrt{n}} \left[-\frac{1}{\sqrt{\Psi_j}} (\iota' \Psi q_2) \left(\frac{\iota' \zeta}{\sqrt{\Psi_j}} \right)^2 \right] \phi_\zeta(\tau) + O(n^{-1}),$$

$$P\{|t_K^{adj}| \geq \tau\} = \Phi_\zeta^-(\tau) + \frac{1}{\sqrt{n}} \left[-\frac{1}{\sqrt{\Psi_j}} (\iota' \Psi q_2) \left(\frac{\iota' \zeta}{\sqrt{\Psi_j}} \right)^2 \right] \{\phi_\zeta(\tau) - \phi_\zeta(-\tau)\} + O(n^{-1}),$$

for each $\tau \in \mathbb{R}$, respectively.

Note that the adjusted t -tests do not dominate the unadjusted t -tests in terms of the second-order power uniformly in ζ . However, the two-sided adjusted t -test has some desirable power property with regard to the second-order local maximinity, i.e., it is more powerful on the basis of the minimum power attainable for alternatives within a given distance from the null hypothesis (see, e.g., Mukerjee, 1994). To be precise, let $P_1(\tau, \zeta)$ and $P_1^{adj}(\tau, \zeta)$ be the terms of order $O(n^{-1/2})$ in the local power functions of the two-sided tests in Theorems 3 and 5, respectively. We obtain the following result.

Corollary 1. *Under the same assumption in Theorem 1 and H_{1n} , it holds*

$$\min_{\zeta \in \mathbb{R}^{G_1+K_1}; (\iota' \zeta)^2 = \delta} P_1^{adj}(\tau, \zeta) \geq \min_{\zeta \in \mathbb{R}^{G_1+K_1}; (\iota' \zeta)^2 = \delta} P_1(\tau, \zeta),$$

for each $\delta > 0$.

3.3. Discussion: Heteroskedastic and non-elliptically contoured distributions. This paper should be considered as a starting point toward more general higher-order theory for inference on instrumental variable regression models. In particular, it is interesting to extend

our analysis for the case of heteroskedastic and/or non-elliptically contoured error terms. It should be noted that the LIML and Fuller estimators can be inconsistent with many instruments and heteroskedasticity of unknown form (see, e.g., Hausman *et al.*, 2012). Hausman *et al.* (2012) proposed heteroskedasticity and many instrument robust versions of the LIML and Fuller estimators, and t -tests based on their asymptotic variance estimators.

Let

$$t_K^{HLI} = \frac{1}{\sqrt{\hat{\Psi}_{H,j}}} l' \sqrt{n} \begin{pmatrix} \hat{\beta}_{HLI} \\ \hat{\gamma}_{HLI} \end{pmatrix},$$

be the t -ratio for testing H_0 based on the heteroskedasticity robust LIML (HLIM) estimator $(\hat{\beta}'_{HLI}, \hat{\gamma}'_{HLI})'$ and its variance estimator $\hat{\Psi}_H$ defined in p. 217 and p. 215 of Hausman *et al.* (2012), respectively. The higher-order analyses are very different and more complicated not only for the HLIM estimator but also for the variance estimator. Thus, we leave such analysis for future research. However, motivated by the adjustment term in (11), we consider the following statistic in the simulation study in the next section

$$t_K^{Hadj} = t_K^{HLI} - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\hat{\Psi}_j}} (l' \hat{\Psi} \hat{q}_2) \min\{t_K^2, n^\varepsilon\}.$$

Since the second term is asymptotically negligible in the first-order, the t -ratio t_K^{Hadj} is still asymptotically valid. Although the second-order properties of t_K^{Hadj} is unknown, we expect that the above adjustment term might improve the finite sample performances of the t -test as far as the effects of heteroskedasticity and non-elliptically contoured errors are mild.

4. SIMULATION

In this section, we conduct a simulation study to examine quality of the preceding asymptotic approximations to the finite sample distributions of the t -ratios. We consider the data generating process (DGP):

$$\begin{aligned} y_{1i} &= y_{2i}\beta_0 + z_{1i}\gamma_0 + u_i, \\ y_{2i} &= z_i'\pi_2 + v_{2i}, \end{aligned} \tag{12}$$

for $i = 1, \dots, n$, where $\pi = (d, \dots, d)'$, $z_i = (z_{1i}, z_{2i})'$, $z_{1i} = 1$, and $z_{2i} \sim N(0, I_{K-1})$. The error terms are generated as $(u_i, v_{2i}) = (\epsilon_{1i}, \rho\epsilon_{1i} + \sqrt{1 - \rho^2}\epsilon_{2i})$, where ϵ_{1i} and ϵ_{2i} are independent and drawn from $N(0, 1)$.⁴ We set $\beta_0 = \gamma_0 = 0$, $\rho = 0.2, 0.4$, and $n = 100$ for the sample size in all cases. For each Monte Carlo replication, we set the value of d to fix the value of the concentration parameter (given the realized values of $\{z_i\}$)

$$\delta^2 = \frac{\pi_2' \left[\sum_{i=1}^n z_{2i} z_{2i}' - \sum_{i=1}^n z_{2i} z_{1i}' (\sum_{i=1}^n z_{1i} z_{1i}')^{-1} \sum_{i=1}^n z_{1i} z_{2i}' \right] \pi_2}{\text{Var}(v_{2i})}.$$

Null distributions of t -ratios for β_0 . First, we investigate the null distributions of the five types of t -ratios – the standard t -ratio with the LIML estimator ($t^{LI} = \sqrt{n} l' \hat{\beta}_{LI} / \sqrt{\hat{\sigma}^2 (\hat{Q}^{-1})_j}$), the

⁴In our preliminary simulation, we also consider the t_5 distribution as another example of the elliptically contoured distribution, but the results are similar to the normal case.

large- K t -ratio with the LIML estimator (t_K^{LI}), the large- K t -ratio with the Fuller estimator (t_K^F), the adjusted large- K t -ratio with the LIML estimator ($t_K^{adj,LI}$), and the adjusted large- K t -ratio with Fuller estimator ($t_K^{adj,F}$). For the adjusted statistics, we set as $\varepsilon = 0.1$. The number of Monte Carlo repetitions in each experiment is 10,000.

Tables 1 and 2 report the null rejection frequencies of the one-sided and two-sided tests at the nominal 5% significance level, respectively. Our findings are summarized as follows.

- i):** The size distortions of the standard t -ratio (t^{LI}) tend to be large when δ^2 is small and K is large. As K increases, the tails of t^{LI} become thicker and its rejection frequencies tend to be larger than the nominal level.
- ii):** Compared to t^{LI} , the rejection frequencies of the large- K t -ratios (t_K^{LI} and t_K^F) are smaller and avoid over-rejection. Although t_K^{LI} and t_K^F work well for two-sided testing, they show some asymmetric behaviors for one-sided testing. More precisely, t_K^{LI} and t_K^F (sometimes severely) under-reject for one-sided testing against $H_1 : \beta < 0$. Overall t_K^{LI} and t_K^F show similar performances, but t_K^{LI} is slightly better than t_K^F for testing against $H_1 : \beta < 0$.
- iii):** Compared to t^{LI} and the (unadjusted) large- K t -ratios (t_K^{LI} and t_K^F), the proposed adjusted t -ratios ($t_K^{adj,LI}$ and $t_K^{adj,F}$) work well for all cases. Their rejection frequencies are overall close to the nominal level, and do not show undesirable asymmetries for one-sided testing as in the unadjusted test statistics t_K^{LI} and t_K^F . The performances of $t_K^{adj,LI}$ and $t_K^{adj,F}$ are similar.

Overall, the adjusted ratios, $t_K^{adj,LI}$ and $t_K^{adj,F}$, perform well for all cases. In particular, the adjustments improve the size distortions for one-sided testing because the null distributions of the adjusted t -ratios are less skewed and closer to the standard normal distribution, which agrees with our results of the asymptotic expansions in Section 2.

Power comparison. Next, we conduct power comparisons of the two-sided unadjusted and adjusted large- K t -tests. We generate 2,000 datasets from the DGP in (12) for various values of β and report power curves at 5% significance level. Figures 1-2 display the power curves. Among various cases tried in preliminary simulations, we present the cases of $\delta^2 = 16$ as typical examples. From these figures, we can see that: (i) for negative values of β , the adjusted large- K t -tests ($t_K^{adj,LI}$ and $t_K^{adj,F}$) are more powerful than the (unadjusted) large- K t -tests, and (ii) for positive values of β , all tests show similar power properties.

Heteroskedastic errors. Finally, we study finite sample performances of the heteroskedastic versions of the large- K t -ratios discussed in Section 3.3. In particular, we consider heteroskedastic error terms with $u_i^H = (1 + 0.01z_{2i}^{*2})u_i$, where z_{2i}^* is the first element of z_{2i} , and report only for the case of $\rho = 0.4$ for brevity in Tables 3 and 4. We can see that the null distributions of the adjusted t -ratios reduce the skewness (asymmetries in the rejection frequencies for the positive and negative alternatives) and are closer to the standard normal distribution. Also our preliminary simulation suggests that the power curves are similar to the case of homoskedastic errors.

APPENDIX A. PROOFS

A.1. **Proof of Theorem 2.** Let

$$\hat{e} = \begin{cases} \sqrt{n}((\hat{\beta}_{LI} - \beta)', (\hat{\gamma}_{LI} - \gamma)') & \text{if } \lambda = \hat{\lambda}, \\ \sqrt{n}((\hat{\beta}_F - \beta)', (\hat{\gamma}_F - \gamma)') & \text{if } \lambda = \hat{\lambda} - a/(n - K). \end{cases}$$

By the definitions of the LIML and Fuller estimators in (3), we have

$$(Y, Z_1)'(P_Z - \lambda M_Z)(Y_2, Z_1)\hat{e} = \sqrt{n}(Y, Z_1)'(P_Z - \lambda M_Z)(Y, Z_1) \begin{pmatrix} 1 \\ -\beta \\ -\gamma \end{pmatrix}. \quad (13)$$

Using the definition,

$$D = (D_1, D_2) = \left(\begin{pmatrix} \pi_{11} \\ \pi_{21} \end{pmatrix}, \begin{pmatrix} \Pi_{12} & I_{K_1} \\ \Pi_{22} & 0 \end{pmatrix} \right),$$

it can be written as

$$\begin{aligned} & (Y, Z_1)'(P_Z - \lambda M_Z)(Y, Z_1) \\ &= \{ZD + (V, 0)\}'(P_Z - \lambda M_Z)\{ZD + (V, 0)\} \\ &= D'Z'ZD + D'Z'(V, 0) + (V, 0)'ZD + (V, 0)'(P_Z - \lambda M_Z)(V, 0). \end{aligned} \quad (14)$$

Also we note that $\sqrt{K}(V'P_ZV/K - \Omega) = O_p(1)$, and $\sqrt{n-K}(V'M_ZV/(n-K) - \Omega) = O_p(1)$.

By substituting (14) into (13) and putting

$$\begin{aligned} \hat{e} &= e^{(0)} + \frac{1}{\sqrt{n}}e^{(1)} + O_p(n^{-1}), \\ \hat{\lambda} &= \lambda^{(0)} + \frac{1}{\sqrt{n}}\lambda^{(1)} + O_p(n^{-1}), \end{aligned}$$

we can determine successively $(e^{(0)}, e^{(1)})$ and $(\lambda^{(0)}, \lambda^{(1)})$ as

$$\begin{aligned} \lambda^{(0)} &= c_*, \quad \lambda^{(1)} = \frac{c_*}{\sigma^2} \left[\frac{\sqrt{n}}{K} u'P_Z u - \frac{\sqrt{n}}{n-K} u'M_Z u \right], \\ e^{(0)} &= Q^{-1} \left[\frac{1}{\sqrt{n}} D_2' Z' u + \frac{\sqrt{c}}{\sqrt{K}} W_2' P_Z u - \frac{\sqrt{cc_*}}{\sqrt{n-K}} W_2' M_Z u \right], \\ e^{(1)} &= -Q^{-1} \left[\left\{ \frac{1}{\sqrt{n}} D_2' Z' (V_2, 0) + \frac{\sqrt{c}}{\sqrt{K}} W_2' P_Z (V_2, 0) - \frac{\sqrt{cc_*}}{\sqrt{n-K}} W_2' M_Z (V_2, 0) \right\} e^{(0)} \right. \\ &\quad \left. + \frac{1}{\sqrt{n}} W_2' Z D_2 e^{(0)} - (1-c)\lambda^{(1)} \left\{ \begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^2 \right\} e^{(0)} \right. \\ &\quad \left. + \sqrt{1-c} \lambda^{(1)} \frac{1}{\sqrt{n-K}} W_2' M_Z u \right] + a(1+c_*)\sigma^2 Q^{-1} q_2, \end{aligned} \quad (15)$$

where $W_2 = (V_2, 0) - uq_2'$, $q_2 = \frac{1}{\sigma^2}(\omega'_{21} - \beta'\Omega_{22}, 0)'$, and $c_* = c/(1-c)$. Each $\lambda^{(l)}$ is obtained by premultiplying $(1, -\beta', -\gamma')$ to (13). Each $e^{(l)}$ is obtained by using the last $G_1 + K_1$ rows of (13).

It should be noted that W_2 and u are uncorrelated when the rows of V are independently distributed. Moreover, we notice that $E[w_{2i}w'_{2i}w_{2i}u_i] = 0$ and $E[w_{2i}w'_{2i}w_{2i}u_i^3] = 0$ when W_2

and u follow some elliptically contoured distribution. Using these facts, the Cornish-Fisher expansions of $\frac{1}{\sqrt{n}}D'_2Z'u$, $\frac{1}{\sqrt{K}}W'_2P_Zu$, and $\frac{1}{\sqrt{n-K}}W'_2M_Zu$ may be written as

$$\begin{aligned}\frac{1}{\sqrt{n}}D'_2Z'u &= \mathcal{X} + O_p(n^{-1}), & \frac{1}{\sqrt{K}}W'_2P_Zu &= \mathcal{Y} + O_p(n^{-1}), \\ \frac{1}{\sqrt{n-K}}W'_2M_Zu &= \mathcal{Z} + O_p(n^{-1}),\end{aligned}$$

where \mathcal{X} , \mathcal{Y} , and \mathcal{Z} are some normally distributed random vectors. Thus, \hat{e} can be written as

$$\hat{e} = \mathcal{E} + \frac{1}{\sqrt{n}}e^{(1)} + O_p(n^{-1}), \quad (16)$$

where $\mathcal{E} = Q^{-1}(\mathcal{X} + \sqrt{c}\mathcal{Y} + \sqrt{cc_*}\mathcal{Z})$.

We derive an asymptotic expansion of the distribution of \hat{e} by inverting the characteristic function of \hat{e} up to order $n^{-1/2}$:

$$E[\exp(is'\mathcal{E})] + \frac{1}{\sqrt{n}}E[is'E[e^{(1)}|\mathcal{E}]\exp(is'\mathcal{E})] + O(n^{-1}), \quad (17)$$

where s is a $(G_1 + K_1) \times 1$ vector of real variables and $i = \sqrt{-1}$. The conditional expectation of $e^{(1)}$ given the first order term \mathcal{E} can be written as

$$E[e^{(1)}|\mathcal{E}] = -\{\mathcal{E}\mathcal{E}' - a(1 + c_*)Q^{-1}\sigma^2\}q_2 + O_p(n^{-1/2}).$$

Therefore, the probability $P(\hat{e} \leq \xi)$ is approximated to the order $n^{-1/2}$ by the Fourier inversion of the characteristic function (17). The inversion of the first term is $\Phi_\Psi(\xi)$. We also use the next Fourier inversion formula, which is a generalization of Fujikoshi *et al.* (1982): for $\mathcal{E} \sim N(\mu, \Sigma)$ and any polynomials $h(\cdot)$ and $g(\cdot)$,

$$\mathcal{F}^{-1}[h(-is)E[g(\mathcal{E})\exp(is'\mathcal{E})]]_{\mathcal{E}=\xi} = h\left(\frac{\partial}{\partial\xi}\right)g(\xi)\phi_{\mu,\Sigma}(\xi), \quad (18)$$

where $\phi_{\mu,\Sigma}(\xi)$ is the density function of $N(\mu, \Sigma)$ and $\partial/\partial\xi' = (\partial/\partial\xi_1, \dots, \partial/\partial\xi_{G_1+K_1})$. The conclusion follows by applying this formula.

A.2. Proof of Theorem 1. To derive the asymptotic expansion of the null distribution of the large- K t -ratio, we first expand each term of $\hat{\Psi}$ in (9). Let E_1 and E_2 such that

$$\begin{aligned}\frac{1}{K}(V, 0)'P_Z(V, 0) &= \begin{pmatrix} \Omega & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{\sqrt{K}}E_1, \\ \frac{1}{n-K}(V, 0)'M_Z(V, 0) &= \begin{pmatrix} \Omega & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{\sqrt{n-K}}E_2.\end{aligned}$$

Using (14) and a $(1 + G_1 + K_1) \times (G_1 + K_1)$ choice matrix $J_2 = (0, I_{G_1+K_1})'$, we have

$$\begin{aligned}\hat{Q} &= Q + \frac{1}{\sqrt{n}}\left[\frac{1}{\sqrt{n}}D'_2Z'(V_2, 0) + \frac{1}{\sqrt{n}}\begin{pmatrix} V'_2 \\ 0' \end{pmatrix}ZD_2\right. \\ &\quad \left. + \sqrt{c}J'_2E_1J_2 + \sqrt{cc_*}J'_2E_2J_2 - \lambda^{(1)}\begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix}\right] + O_p(n^{-1}),\end{aligned}$$

and thus \hat{Q}^{-1} is expanded as

$$\hat{Q}^{-1} = Q^{-1} - \frac{1}{\sqrt{n}}Q^{-1}BQ^{-1} + O_p(n^{-1}),$$

where

$$B = \frac{1}{\sqrt{n}}D'_2Z'(V_2, 0) + \frac{1}{\sqrt{n}}(V_2, 0)'ZD_2 + \sqrt{c}J'_2E_1J_2 + \sqrt{cc_*}J'_2E_2J_2 - \lambda^{(1)} \begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix}.$$

Also, note that

$$\begin{aligned} & \frac{1}{n-K} \hat{b}'Y'M_ZY\hat{b} \\ &= \left\{ b - \frac{1}{\sqrt{n}} \begin{pmatrix} 0 \\ e_\beta^{(0)} \end{pmatrix} + O_p(n^{-1}) \right\}' \\ & \quad \times \left\{ \Omega + \frac{1}{\sqrt{n}} \left[\sqrt{\frac{n-K}{1-c}} \left(\frac{1}{n-K} V'M_ZV - \Omega \right) \right] \right\} \left\{ b - \frac{1}{\sqrt{n}} \begin{pmatrix} 0 \\ e_\beta^{(0)} \end{pmatrix} + O_p(n^{-1}) \right\} \\ &= \sigma^2 + \frac{1}{\sqrt{n}} \left[-2(0, e_\beta^{(0)'})\Omega b + \sqrt{\frac{n-K}{1-c}} b' \left(\frac{1}{n-K} V'M_ZV - \Omega \right) b \right] + O_p(n^{-1}), \end{aligned}$$

where $b = (1, -\beta)'$ and $e_\beta^{(0)}$ is the first G_1 elements of $e^{(0)}$ in (15). Similarly,

$$\begin{aligned} & \frac{1}{(n-K)^2} Y'M_ZY\hat{b}\hat{b}'Y'M_ZY \\ &= \Omega b b' \Omega + \frac{1}{\sqrt{n}} \left[-\Omega b(0, e_\beta^{(0)'})\Omega + \sqrt{\frac{n-K}{1-c}} \Omega b b' \left(\frac{1}{n-K} V'M_ZV - \Omega \right) \right. \\ & \quad \left. - \Omega \begin{pmatrix} 0 \\ e_\beta^{(0)} \end{pmatrix} b' \Omega + \sqrt{\frac{n-K}{1-c}} \left(\frac{1}{n-K} V'M_ZV - \Omega \right) b b' \Omega \right] + O_p(n^{-1}). \end{aligned}$$

Combining these terms, we have

$$\hat{\Psi} = \Psi + \frac{1}{\sqrt{n}}\Psi^{(1)} + O_p(n^{-1}),$$

where

$$\begin{aligned}
\Psi^{(1)} &= Q^{-1} \left[-2e^{(0)'} q_2 \sigma^2 + \sqrt{\frac{n-K}{1-c}} \left(\frac{1}{n-K} u' M_Z u - \sigma^2 \right) \right] \\
&\quad - Q^{-1} B Q^{-1} \sigma^2 + (c_* + \kappa \eta) Q^{-1} A Q^{-1} - (c_* + \kappa \eta) Q^{-1} B Q^{-1} \left[\begin{pmatrix} \sigma^2 \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^4 \right] Q^{-1} \\
&\quad - (c_* + \kappa \eta) Q^{-1} \left[\begin{pmatrix} \sigma^2 \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^4 \right] Q^{-1} B Q^{-1} + \kappa^* \eta Q^{-1} \left[\begin{pmatrix} \sigma^2 \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^4 \right] Q^{-1}, \\
\kappa^* &= \frac{1}{3\sigma^2} \left[-\frac{4}{n} \sum_{i=1}^n u_i^3 w_{2i}' e^{(0)} - \frac{4}{n} \sum_{i=1}^n u_i^3 z_i' D_2 e^{(0)} \right. \\
&\quad \left. + \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n u_i^4 - E[u_i^4] \right) - \frac{2E[u_i^4]}{\sigma^2} \sqrt{\frac{n-K}{1-c}} \left(\frac{1}{n-K} u' M_Z u - \sigma^2 \right) \right], \\
A &= -2 \begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} e^{(0)'} q_2 \sigma^2 + \sqrt{\frac{n-K}{1-c}} \begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} \left(\frac{1}{n-K} u' M_Z u - \sigma^2 \right) \\
&\quad + \sqrt{\frac{1}{1-c}} J_2' E_2 J_2 \sigma^2 + q_2 \sigma^2 e^{(0)'} \begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} \\
&\quad - q_2 \sqrt{\frac{n-K}{1-c}} \left(\frac{1}{n-K} u' M_Z (V_2, 0) - q_2' \sigma^2 \right) \sigma^2 + \begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} e^{(0)} q_2' \sigma^2 \\
&\quad - \sqrt{\frac{n-K}{1-c}} \left(\frac{1}{n-K} (V_2, 0)' M_Z u - q_2 \sigma^2 \right) q_2' \sigma^2.
\end{aligned}$$

Under H_0 , the large- K t -ratio (10) is approximated as

$$\begin{aligned}
t_K &= \frac{\hat{e}}{\sqrt{\hat{\Psi}_j}} \sqrt{\frac{\Psi_j}{\hat{\Psi}_j}} = \frac{e' \left(e^{(0)} + \frac{1}{\sqrt{n}} e^{(1)} + O_p(n^{-1}) \right)}{\sqrt{\Psi_j}} \left(1 - \frac{1}{2\sqrt{n}} \frac{\Psi_j^{(1)}}{\Psi_j} + O_p(n^{-1}) \right) \\
&= \mathcal{T} + \frac{1}{\sqrt{n}} t^{(1)} + O_p(n^{-1}), \tag{19}
\end{aligned}$$

where

$$\mathcal{T} = \frac{e' e^{(0)}}{\sqrt{\Psi_j}}, \quad t^{(1)} = \frac{e' e^{(1)}}{\sqrt{\Psi_j}} - \frac{1}{2} \frac{\Psi_j^{(1)}}{\Psi_j} \mathcal{T}.$$

The first-order term \mathcal{T} is distributed as $N(0, 1)$. We derive an asymptotic expansion of the distribution function of the large- K t -ratio by inverting the characteristic function up to $O(n^{-1/2})$:

$$E[\exp(is\mathcal{T})] + \frac{1}{\sqrt{n}} E[isE[t^{(1)}|\mathcal{T}] \exp(is\mathcal{T})] + O(n^{-1}).$$

By using Kunitomo and Matsushita (2009, Lemma 4.3) and the fact that any odd moments of the elliptically contoured distribution is 0, the expectation of $t^{(1)}$ conditional on \mathcal{T} is calculated

as

$$\begin{aligned} E[t^{(1)}|\mathcal{T}] &= -\frac{1}{\sqrt{\Psi_j}} \left[(\iota' \Psi q_2) \mathcal{T}^2 - \frac{a}{1-c} \sigma^2 \iota' Q^{-1} q_2 \right] - \frac{1}{2} \left[-\frac{4}{\sqrt{\Psi_j}} (\iota' \Psi q_2) \mathcal{T}^2 \right] \\ &= \frac{1}{\sqrt{\Psi_j}} \left[\mathcal{T}^2 \iota' \Psi + \frac{a}{1-c} \sigma^2 \iota' Q^{-1} \right] q_2 + O_p(n^{-1/2}), \end{aligned}$$

where the first equality follows from Lemma 1. The probability $P\{t_K \leq \xi\}$ is approximated to the order $n^{-1/2}$ using the same formula as (18).

The validity of the expansion is given by similar arguments to those in Kunitomo and Matsushita (2009) and in Fujikoshi *et al.* (1982). The random variables that appear in our analyses are $x_1 = \frac{1}{\sqrt{n}} D_2' Z' u$, $x_2 = \frac{1}{\sqrt{n}} D_2' Z' W_2$, $x_3 = \sqrt{K} (u' P_Z u / K - \sigma^2)$, $x_4 = \sqrt{n-K} (u' M_Z u / (n-K) - \sigma^2)$, $x_5 = \frac{1}{\sqrt{K}} W_2' P_Z u$, $x_6 = \frac{1}{\sqrt{n-K}} W_2' M_Z u$, $x_7 = \sqrt{K} (W_2' P_Z W_2 / K - C_2)$, and $x_8 = \sqrt{n-K} (W_2' M_Z W_2 / (n-K) - C_2)$, where $C_2 = E[w_{2i} w_{2i}']$. We use the space J_n where each element of x_i (for $i = 1, \dots, 8$) is in the interval $(-2c\sqrt{\log n}, 2c\sqrt{\log n})$ and c is a standard deviation of each random variable. Then, if $E[\|v_i\|^8] < \infty$, we can take a positive constant d which satisfies

$$P\{\|x_j\| > \sqrt{\Lambda_n \log n}\} \leq \frac{d}{n(\log n)^2},$$

where Λ_n as the maximum of the characteristic roots of the covariance matrix of x_j ($j = 1, \dots, 8$) (Bhattacharya and Ghosh, 1978). Then, $P(J_n) = 1 - O(n^{-1})$, which can be proved in the same way as in Anderson (1974). We see that each element of $e^{(l)}$ and $t^{(l)}$ is a homogeneous polynomial of degree $l+1$ in the elements of x_j . The remainder terms of (16) and (19) are of the order $O(n^{-1})$ uniformly in J_n . Therefore, the analysis subsequent to (B.3) in Fujikoshi *et al.* (1982) is applicable.

Lemma 1. *Based on the setup and notation of the proof of Theorem 1, it holds $E[\Psi_j^{(1)}|\mathcal{T}] = -4\sqrt{\Psi_j}(q_2' \Psi \iota)$.*

Proof of Lemma 1. Decompose

$$\begin{aligned} \Psi_j^{(1)} &= \iota' \left\{ Q^{-1} \left[-2e^{(0)'} q_2 \sigma^2 + \sqrt{\frac{n-K}{1-c}} \left(\frac{1}{n-K} u' M_Z u - \sigma^2 \right) \right] + (c_* + \kappa\eta) Q^{-1} A Q^{-1} \right. \\ &\quad \left. - Q^{-1} B Q^{-1} \sigma^2 - (c_* + \kappa\eta) Q^{-1} B Q^{-1} \left[\begin{pmatrix} \sigma^2 \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^4 \right] Q^{-1} \right. \\ &\quad \left. - (c_* + \kappa\eta) Q^{-1} \left[\begin{pmatrix} \sigma^2 \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^4 \right] Q^{-1} B Q^{-1} + \kappa^* \eta Q^{-1} \left[\begin{pmatrix} \sigma^2 \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^4 \right] Q^{-1} \right\} \iota \\ &\equiv \Psi_{j1}^{(1)} + \Psi_{j2}^{(1)} + \Psi_{j3}^{(1)} + \Psi_{j4}^{(1)} + \Psi_{j5}^{(1)} + \Psi_{j6}^{(1)}. \end{aligned}$$

For $\Psi_{j6}^{(1)}$, we have $E[\Psi_{j6}^{(1)}|\mathcal{T}] = 0$. For $\Psi_{j1}^{(1)}$,

$$E[\Psi_{j1}^{(1)}|\mathcal{T}] = -\frac{2\sigma^2}{\sqrt{\Psi_j}} (\iota' Q^{-1} \iota) (\iota' E[e^{(0)} e^{(0)'}] q_2) \mathcal{T} = -\frac{2\sigma^2}{\sqrt{\Psi_j}} (\iota' Q^{-1} \iota) (\iota' \Psi q_2) \mathcal{T}. \quad (20)$$

For $\Psi_{j2}^{(1)}$, note that

$$\begin{aligned}
l'Q^{-1}AQ^{-1}l &= l'Q^{-1} \left[-2 \begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} e^{(0)'}q_2 + \sqrt{\frac{n-K}{1-c}} \begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} \left(\frac{1}{n-K}u'M_Zu - \sigma^2 \right) \right. \\
&\quad + \sqrt{\frac{1}{1-c}}J_2'E_2J_2\sigma^2 + q_2e^{(0)'} \begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} \\
&\quad - q_2\sqrt{\frac{n-K}{1-c}} \left(\frac{1}{n-K}u'M_Z(V_2, 0) - q_2'\sigma^2 \right) \sigma^2 + \begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} e^{(0)'}q_2' \\
&\quad \left. - \sqrt{\frac{n-K}{1-c}} \left(\frac{1}{n-K}(V_2, 0)'M_Zu - q_2\sigma^2 \right) q_2'\sigma^2 \right] Q^{-1}l \\
&\equiv A_{j1} + \dots + A_{j7}.
\end{aligned}$$

Observe that $E[A_{j2} + A_{j3} + A_{j5} + A_{j7}|\mathcal{T}] = 0$, and

$$\begin{aligned}
E[A_{j1}|\mathcal{T}] &= -\frac{2}{\sqrt{\Psi_j}} \left(l'Q^{-1} \begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} Q^{-1}l \right) (l'\Psi q_2)\mathcal{T}, \\
E[A_{j4} + A_{j6}|\mathcal{T}] &= \frac{2}{\sqrt{\Psi_j}} \left(l'Q^{-1} \begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} \Psi l \right) (l'Q^{-1}\sigma^2 q_2)\mathcal{T}.
\end{aligned}$$

Combining these results,

$$\begin{aligned}
E[\Psi_{j2}^{(1)}|\mathcal{T}] &= -(c^* + \kappa\eta) \frac{2}{\sqrt{\Psi_j}} \left(l'Q^{-1} \begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} Q^{-1}l \right) (l'\Psi q_2)\mathcal{T} \\
&\quad + (c^* + \kappa\eta) \frac{2}{\sqrt{\Psi_j}} \left(l'Q^{-1} \begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} \Psi l \right) (l'Q^{-1}\sigma^2 q_2)\mathcal{T}. \tag{21}
\end{aligned}$$

For $\Psi_{j3}^{(1)} + \Psi_{j4}^{(1)} + \Psi_{j5}^{(1)}$, we have

$$\begin{aligned}
&E[\Psi_{j3}^{(1)} + \Psi_{j4}^{(1)} + \Psi_{j5}^{(1)}|\mathcal{T}] \\
&= E[l'Q^{-1}BQ^{-1}l\sigma^2 - 2l'Q^{-1}B\Psi l|\mathcal{T}] \\
&= \frac{2}{\sqrt{\Psi_j}}(l'\Psi l)(q_2'Q^{-1}\sigma^2 l) - \frac{2}{\sqrt{\Psi_j}}(l'\Psi l)(q_2'\Psi l) \\
&\quad - \frac{2}{\sqrt{\Psi_j}} \left(l'Q^{-1} \left[Q + (c^* + \kappa\eta) \left\{ \begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - \sigma^4 q_2 q_2' \right\} \right] \Psi l \right) (q_2'Q^{-1}\sigma^2 l) \\
&= -(c^* + \kappa\eta) \frac{2}{\sqrt{\Psi_j}} \left(l'Q^{-1} \begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} \Psi l \right) (l'Q^{-1}\sigma^2 q_2)\mathcal{T} \\
&\quad + (c^* + \kappa\eta) \frac{2}{\sqrt{\Psi_j}} (q_2'Q^{-1}\sigma^2 l)^2 (q_2'\Psi l)\mathcal{T}, \tag{22}
\end{aligned}$$

where we have used the fact that $E[w_{2i}w_{2i}'] = \begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - \sigma^2 q_2 q_2'$.

Combining (20)-(22), the conclusion follows.

A.3. **Proof of Theorem 3.** Under H_{1n} , the large- K t -ratio can be written as

$$\begin{aligned} t_K &= \frac{1}{\sqrt{\hat{\Psi}_j}} l' \sqrt{n} \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = \frac{1}{\sqrt{\hat{\Psi}_j}} l' \sqrt{n} \begin{pmatrix} \hat{\beta} - \beta \\ \hat{\gamma} - \gamma \end{pmatrix} + \frac{l' \zeta}{\sqrt{\hat{\Psi}_j}} \\ &= \left(\mathcal{T} + \frac{l' \zeta}{\sqrt{\Psi_j}} \right) + \frac{1}{\sqrt{n}} \left(t^{(1)} - \frac{\Psi_j^{(1)} l' \zeta}{2 \Psi_j \sqrt{\Psi_j}} \right) + O_p(n^{-1}), \end{aligned} \quad (23)$$

where \mathcal{T} in the first-order term is distributed as $N(0, 1)$. Since the expectation of $\Psi_j^{(1)}$ conditional on \mathcal{T} is calculated as $E[\Psi_j^{(1)} | \mathcal{T}] = -4\sqrt{\Psi_j} (l' \Psi q_2) \mathcal{T} + O_p(n^{-1/2})$ by Lemma 1, we have

$$\begin{aligned} &E \left[t^{(1)} - \frac{\Psi_j^{(1)} l' \zeta}{2 \Psi_j \sqrt{\Psi_j}} \middle| \mathcal{T}_* = \mathcal{T} + \frac{l' \zeta}{\sqrt{\Psi_j}} \right] \\ &= \frac{1}{\sqrt{\Psi_j}} \left[\mathcal{T}_*^2 l' \Psi + \frac{a}{1-c} \sigma^2 l' Q^{-1} \right] q_2 - \frac{1}{\sqrt{\Psi_j}} (l' \Psi q_2) \left(\frac{l' \zeta}{\sqrt{\Psi_j}} \right)^2 + O_p(n^{-1/2}). \end{aligned} \quad (24)$$

Then the probability $P\{t_K \leq \tau\}$ is approximated to the order $O(n^{-1/2})$ using the inversion formula (18):

$$\begin{aligned} &P\{t_K < \tau\} \\ &= \Phi_\zeta(\tau) - \frac{1}{\sqrt{n}} \left[\frac{1}{\sqrt{\Psi_j}} \left\{ \tau^2 (l' \Psi) + \frac{a}{1-c} \sigma^2 Q^{-1} \right\} q_2 - \frac{1}{\sqrt{\Psi_j}} (l' \Psi q_2) \left(\frac{l' \zeta}{\sqrt{\Psi_j}} \right)^2 \right] \phi_\zeta(\tau) + O(n^{-1}). \end{aligned}$$

The result for the two-sided test is obtained by $P\{|t_K| \leq \tau\} = P\{t_K \leq \tau\} - P\{t_K \leq -\tau\}$.

A.4. **Proof of Theorems 4 and 5.** We only present the proof for Theorem 5 since the proof of Theorem 4 for the null distribution follows directly by setting $\zeta = 0$.

By using (23), the adjusted t -ratio under H_{1n} can be written as

$$\begin{aligned} t_K^{adj} &= t_K - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\hat{\Psi}_j}} \left\{ (l' \hat{\Psi} \hat{q}_2) t_K^2 + \frac{an}{n-K} \hat{\sigma}^2 l' \hat{Q}^{-1} \hat{q}_2 \right\} + O_p(n^{-1}) \\ &= \left(\mathcal{T} + \frac{l' \zeta}{\sqrt{\Psi_j}} \right) + \frac{1}{\sqrt{n}} \left(t^{(1)} - \frac{\Psi_j^{(1)} l' \zeta}{2 \Psi_j \sqrt{\Psi_j}} \right) \\ &\quad - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\Psi_j}} (l' \Psi q_2) \left(\mathcal{T} + \frac{l' \zeta}{\sqrt{\Psi_j}} \right)^2 - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\Psi_j}} \frac{a}{1-c} \sigma^2 l' Q^{-1} q_2 + O_p(n^{-1}), \end{aligned}$$

where the first equality follows from $\mathbb{I}\{t_K^2 \geq n^\varepsilon\} = O_p(n^{-1})$ (because of the fact that t_K^2 converges to a non-central χ^2 distribution under H_{1n}). Thus, using (24) with $a = 0$, the expectation of the $O(n^{-1/2})$ term conditional on $\mathcal{T}_* = \mathcal{T} + \frac{l' \zeta}{\sqrt{\Psi_j}}$ is calculated as

$$\begin{aligned} &E \left[\left(t^{(1)} - \frac{\Psi_j^{(1)} l' \zeta}{2 \Psi_j \sqrt{\Psi_j}} \right) - \frac{1}{\sqrt{\Psi_j}} (l' \Psi q_2) \left(\mathcal{T} + \frac{l' \zeta}{\sqrt{\Psi_j}} \right)^2 - \frac{1}{\sqrt{\Psi_j}} \frac{a}{1-c} \sigma^2 l' Q^{-1} q_2 \middle| \mathcal{T}_* \right] \\ &= -\frac{1}{\sqrt{\Psi_j}} (l' \Psi q_2) \left(\frac{l' \zeta}{\sqrt{\Psi_j}} \right)^2 + O_p(n^{-1/2}). \end{aligned}$$

The probability $P\{t_K^{adj} \leq \tau\}$ is approximated to the order $n^{-1/2}$ using the inversion formula (18):

$$P\{t_K^{adj} \leq \tau\} = \Phi_\zeta(\tau) - \frac{1}{\sqrt{n}} \left[-\frac{1}{\sqrt{\Psi_j}} (t' \Psi q_2) \left(\frac{t' \zeta}{\sqrt{\Psi_j}} \right)^2 \right] \phi_\zeta(\tau) + O(n^{-1}).$$

The result for the two-sided test is obtained by $P\{|t_K^{adj}| \leq \tau\} = P\{t_K^{adj} \leq \tau\} - P\{t_K^{adj} \leq -\tau\}$.

APPENDIX B. TABLES AND FIGURES

ρ	δ^2	K	t^{LI}		t_K^{LI}		t_K^F		$t_K^{adj,LI}$		$t_K^{adj,F}$		
			$\beta < 0$	$\beta > 0$	$\beta < 0$	$\beta > 0$	$\beta < 0$	$\beta > 0$	$\beta < 0$	$\beta > 0$	$\beta < 0$	$\beta > 0$	
0.2	64	5	0.052	0.056	0.045	0.050	0.040	0.050	0.057	0.050	0.058	0.050	
		10	0.056	0.062	0.043	0.048	0.038	0.048	0.054	0.049	0.055	0.049	
		20	0.072	0.083	0.038	0.052	0.033	0.051	0.054	0.054	0.056	0.054	
	32	5	0.045	0.061	0.033	0.048	0.027	0.046	0.055	0.053	0.057	0.053	
		10	0.057	0.078	0.032	0.052	0.025	0.050	0.056	0.058	0.057	0.058	
		20	0.077	0.107	0.027	0.048	0.021	0.047	0.058	0.060	0.060	0.060	
	16	5	0.032	0.066	0.017	0.046	0.012	0.040	0.057	0.060	0.057	0.059	
		10	0.042	0.088	0.015	0.042	0.010	0.038	0.057	0.068	0.055	0.066	
		20	0.069	0.133	0.013	0.044	0.010	0.039	0.058	0.078	0.057	0.077	
	0.4	64	5	0.044	0.061	0.039	0.055	0.031	0.057	0.056	0.051	0.057	0.051
			10	0.048	0.067	0.034	0.056	0.029	0.058	0.052	0.053	0.055	0.053
			20	0.060	0.084	0.034	0.056	0.026	0.058	0.053	0.053	0.056	0.052
32		5	0.036	0.069	0.026	0.058	0.020	0.060	0.058	0.053	0.060	0.054	
		10	0.045	0.078	0.024	0.056	0.017	0.057	0.060	0.053	0.062	0.054	
		20	0.054	0.111	0.018	0.062	0.012	0.061	0.053	0.063	0.055	0.062	
16		5	0.017	0.078	0.009	0.057	0.005	0.057	0.059	0.059	0.061	0.060	
		10	0.022	0.107	0.008	0.062	0.004	0.061	0.056	0.071	0.055	0.071	
		20	0.038	0.145	0.006	0.061	0.004	0.058	0.052	0.079	0.051	0.076	

TABLE 1. Null rejection frequencies of one-sided large- K t -tests at 5% significance level (homoskedastic case)

ρ	δ^2	K	t^{LI}	t_K^{LI}	t_K^F	$t_K^{adj,LI}$	$t_K^{adj,F}$	
0.2	64	5	0.053	0.045	0.045	0.052	0.053	
		10	0.063	0.045	0.044	0.053	0.053	
		20	0.082	0.039	0.038	0.050	0.051	
	32	5	0.053	0.038	0.037	0.054	0.055	
		10	0.070	0.035	0.033	0.057	0.059	
		20	0.105	0.033	0.032	0.058	0.061	
	16	5	0.044	0.026	0.024	0.055	0.059	
		10	0.061	0.024	0.022	0.058	0.059	
		20	0.108	0.022	0.020	0.068	0.068	
	0.4	64	5	0.051	0.044	0.045	0.050	0.050
			10	0.059	0.044	0.044	0.051	0.053
			20	0.079	0.041	0.042	0.050	0.053
32		5	0.050	0.036	0.036	0.051	0.054	
		10	0.065	0.037	0.037	0.054	0.058	
		20	0.096	0.038	0.038	0.058	0.062	
16		5	0.047	0.034	0.033	0.058	0.063	
		10	0.072	0.035	0.034	0.064	0.067	
		20	0.110	0.035	0.033	0.067	0.068	

TABLE 2. Null rejection frequencies of two-sided large- K t -tests at 5% significance level (homoskedastic case)

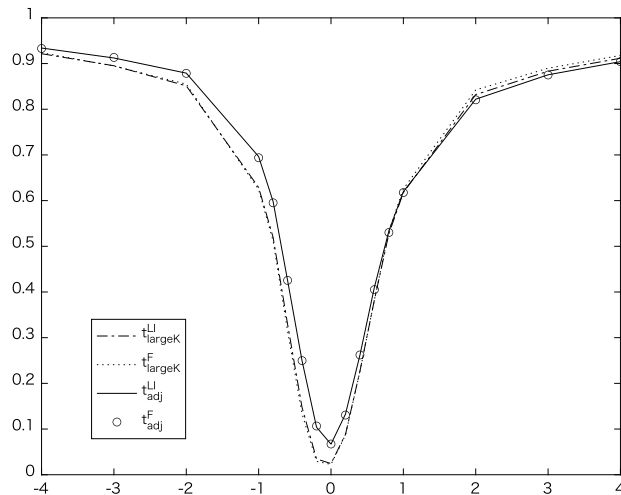


FIGURE 1. Power curves of two-sided tests: $n = 100, K = 20, \rho = 0.2, \delta^2 = 16$ (homoskedastic case)

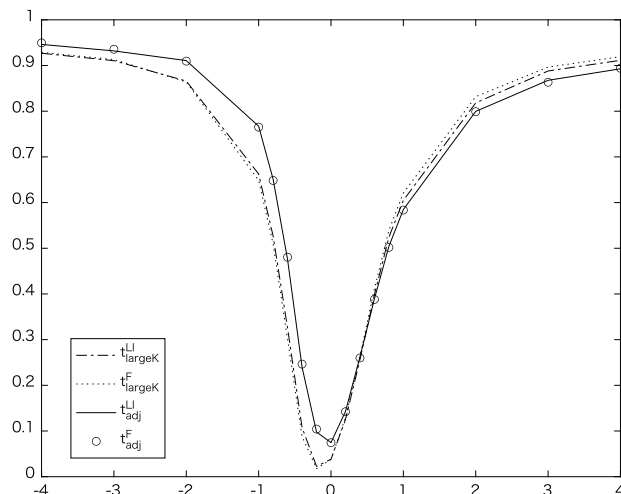


FIGURE 2. Power curves of two-sided tests: $n = 100, K = 20, \rho = 0.4, \delta^2 = 16$ (homoskedastic case)

ρ	δ^2	K	t_K^{HLI}		t_K^{HF}		$t_K^{adj,HLI}$		$t_K^{adj,HF}$	
			$\beta < 0$	$\beta > 0$	$\beta < 0$	$\beta > 0$	$\beta < 0$	$\beta > 0$	$\beta < 0$	$\beta > 0$
0.4	64	5	0.039	0.059	0.032	0.062	0.054	0.054	0.055	0.054
		10	0.030	0.062	0.023	0.065	0.050	0.058	0.052	0.058
		20	0.018	0.062	0.013	0.065	0.036	0.058	0.037	0.059
	32	5	0.019	0.062	0.013	0.065	0.050	0.059	0.051	0.059
		10	0.009	0.061	0.007	0.064	0.034	0.059	0.035	0.060
		20	0.003	0.069	0.002	0.071	0.018	0.069	0.018	0.069
	16	5	0.004	0.060	0.003	0.062	0.028	0.064	0.027	0.064
		10	0.003	0.063	0.001	0.062	0.017	0.070	0.017	0.070
		20	0.000	0.066	0.000	0.065	0.005	0.081	0.005	0.081

TABLE 3. Null rejection frequencies of one-sided large- K t -tests at 5% significance level (heteroskedastic case)

ρ	δ^2	K	t_K^{HLI}	t_K^{HF}	$t_K^{adj,HLI}$	$t_K^{adj,HF}$
0.4	64	5	0.048	0.051	0.054	0.055
		10	0.044	0.050	0.050	0.052
		20	0.040	0.048	0.044	0.045
	32	5	0.041	0.045	0.053	0.055
		10	0.038	0.043	0.045	0.047
		20	0.043	0.047	0.046	0.047
	16	5	0.034	0.035	0.045	0.047
		10	0.036	0.036	0.044	0.045
		20	0.040	0.039	0.050	0.051

TABLE 4. Null rejection frequencies of two-sided large- K t -tests at 5% significance level (heteroskedastic case)

REFERENCES

- [1] Anderson, T. W. (1974) An asymptotic expansion of the distribution of the limited information maximum likelihood estimate of a coefficient in a simultaneous equation system, *Journal of the American Statistical Association*, 69, 565-573.
- [2] Anderson, T. W. (2003) *An Introduction to Multivariate Statistical Analysis*, 3rd edition, John-Wiley.
- [3] Anderson, T. W., Kunitomo, N. and Y. Matsushita (2010) On the asymptotic optimality of the LIML estimator with possibly many instruments, *Journal of Econometrics*, 157, 191-204.
- [4] Anderson, T. W., Kunitomo, N. and Y. Matsushita (2011) On finite sample properties of alternative estimators of coefficients in a structural equation with many instruments, *Journal of Econometrics*, 165, 58-69.
- [5] Anderson, T. W., Kunitomo, N. and T. Sawa (1982) Evaluation of the distribution function of the limited information maximum likelihood estimator, *Econometrica*, 50, 1009-1027.
- [6] Angrist, J. D. and A. Krueger (1991) Does compulsory school attendance affect schooling and earnings, *Quarterly Journal of Economics*, 106, 979-1014.
- [7] Bekker, P. A. (1994) Alternative approximations to the distributions of instrumental variables estimators, *Econometrica*, 63, 657-681.
- [8] Chioda, L. and M. Jansson (2009) Optimal invariant inference when the number of instruments is large, *Econometric Theory*, 25, 793-805.
- [9] Fujikoshi, Y., Morimune, K., Kunitomo, N. and M. Taniguchi (1982) Asymptotic expansions of the distributions of the estimates of coefficients in a simultaneous equation system, *Journal of Econometrics*, 18, 191-205.
- [10] Fuller, W. A. (1977) Some properties of a modification of the limited information estimator, *Econometrica*, 45, 939-954.
- [11] Hahn, J., Hausman, J. A. and G. M. Kuersteiner (2004) Estimation with weak instruments: accuracy of higher-order bias and MSE approximations, *Econometrics Journal*, 7, 272-306.
- [12] Hansen, C., Hausman, J. A. and W. K. Newey (2008) Estimation with many instrumental variables, *Journal of Business & Economic Statistics*, 26, 398-422.
- [13] Hausman, J. A., Newey, W. K., Woutersen, T., Chao, J. and N. R. Swanson (2012) Instrumental variable estimation with heteroskedasticity and many instruments, *Quantitative Economics*, 3, 211-255.
- [14] Kunitomo, N. (1980) Asymptotic expansions of distributions of estimators in a linear functional relationship and simultaneous equations, *Journal of the American Statistical Association*, 75, 693-700.
- [15] Kunitomo, N. (1982) Asymptotic efficiency and higher order efficiency of the limited information maximum likelihood estimator in large econometric models, Technical Report No. 365, Institute for Mathematical Studies in the Social Sciences, Stanford University.
- [16] Kunitomo, N. and Y. Matsushita (2009) Asymptotic expansions of the distributions of semi-parametric estimators in a linear simultaneous equations system, *Journal of Multivariate Analysis*, 100, 1727-1751.
- [17] Morimune, K. (1983) Approximate distributions of k-class estimators when the degree of overidentification is large compared with sample size, *Econometrica*, 51, 821-841.
- [18] Morimune, K. (1989) t test in a structural equation, *Econometrica*, 57, 1341-1360.
- [19] Mukerjee, R. (1994) Comparison of tests in their original forms, *Sankhyā*, A 45, 118-127.
- [20] Phillips, P. C. B. (1977) A general theorem in the theory of asymptotic expansions as approximations to the finite sample distributions of econometric estimators, *Econometrica*, 45, 1517-1534.
- [21] Rao, C. R. (1973) *Linear Statistical Inference and Its Applications*, 2nd edition, John-Wiley.
- [22] Rothenberg, T. J. (1984) Approximating the distributions of econometric estimators and test statistics, in Griliches, Z. and M. D. Intriligator (eds.) *Handbook of Econometrics*, vol. II, ch. 15, Elsevier.
- [23] Rothenberg, T. J. (1988) Approximate power functions for some robust tests of regression coefficients, *Econometrica*, 56, 997-1019.

- [24] Sargan, J. D. (1975) Gram-Charlier approximations applied to t ratios of k-class estimators, *Econometrica*, 43, 327-346.
- [25] Ullah, A. (2004) *Finite Sample Econometrics*, Oxford University Press.
- [26] van Hasselt, M. (2010) Many instruments asymptotic approximations under non-normal error distributions, *Econometric Theory*, 26, 633-645.

GRADUATE SCHOOL OF ECONOMICS, HITOTSUBASHI UNIVERSITY, 2-1 NAKA, KUNITACHI, TOKYO 186-8601, JAPAN.

Email address: matsushita.y@r.hit-u.ac.jp

DEPARTMENT OF ECONOMICS, LONDON SCHOOL OF ECONOMICS, HOUGHTON STREET, LONDON, WC2A 2AE, UK.

Email address: t.otsu@lse.ac.uk