Efficient Inference in the Classical IV Regression Model with Weak Identification: Asymptotic Power Against Arbitrarily Large Deviations from the Null Hypothesis

Vadim Marmer∗
University of British Columbia

Zhengfei Yu†
University of British Columbia

September 1, 2015

Abstract

This paper considers efficient inference for the coefficient on the endogenous variable in linear regression models with weak instrumental variables (Weak-IV) and homoskedastic errors. We focus on the alternative hypothesis determined by an arbitrarily large deviation from the null hypothesis. The efficient rotation-invariant and asymptotically similar test turns out to be infeasible as it depends on the unknown correlation between structural and first-stage errors (the degree of endogeneity). We compare the asymptotic power properties of popular Weak-IV-robust tests, focusing on the Anderson-Rubin (AR) and the Conditional Likelihood Ratio (CLR) tests. We find that their relative power performance depends on the degree of endogeneity in the model and the number of IVs. Unexpectedly, the AR test outperforms the CLR when the degree of endogeneity is small and the number of IVs is large. We also describe a test that is optimal when IVs are strong and, when IVs are weak, has the same asymptotic power as the AR test against arbitrarily large deviations from the null.

Keywords: weak instruments; arbitrarily large deviations; power envelope; power comparisons

JEL Classification: C12; C26

∗We thank Mehmet Caner, Yanqin Fan, Mototsugu Shintani, and Kevin Song for helpful comments. Vadim Marmer gratefully acknowledges the financial support of the SSHRC under grants 410-2010-1394 and 435-2013-0331. Vancouver School of Economics, #997 - 1873 East Mall, Vancouver, BC Canada V6T 1Z1; E-mail: vadim.marmer@ubc.ca.

†Vancouver School of Economics, #997 - 1873 East Mall, Vancouver, BC Canada V6T 1Z1; E-mail: yuzhengfei@gmail.com.
1 Introduction

Weak instrumental variables (Weak-IV) have received a lot of attention in econometrics in particular following Staiger and Stock (1997), who developed an analytical framework for analysing the effect of weak instruments and constructing weak-identification-robust methods of inference, and Dufour (1997), who showed that usual bounded confidence intervals cannot be valid in the case of weak identification. Valid inference can be based on the Anderson-Rubin (AR) statistic (Anderson and Rubin, 1949) as suggested in Staiger and Stock (1997), and weak-identification robust confidence sets can be constructed by inversion of the AR test.

However, the AR test turned out to be inefficient when IVs are strong and the model is overidentified, and several modifications that address the AR’s deficiencies have been proposed in the literature, most notably Kleibergen’s Lagrange Multiplier (KLM) (Kleibergen, 2002, 2007) and the Conditional Likelihood Ratio (CLR) (Moreira, 2003) statistics and tests. Andrews, Moreira, and Stock (2006) showed that, in the normal linear instrumental variables regression models with homoskedastic errors, the KLM and the CLR tests are efficient if IVs are strong. However, when IVs are weak, there is no uniformly most powerful (UMP) test except for the just identified case, hence the efficiency depends on the optimizing criteria considered by the econometrician (even in the class of rotation-invariant and similar tests). Andrews, Moreira, and Stock (2006) considered weighted average power (WAP) using two carefully chosen values of parameters and calculated the power envelope. They also numerically demonstrated that in the Weak-IV scenario, the CLR test dominates the AR and KLM tests, and its power is numerically very close to the power envelope given by the optimal average power test. They recommended the CLR test for use in empirical work when the instruments may be weak and model is over-identified. Cattaneo, Crump, and Jansson (2012) extended the results of Andrews, Moreira, and Stock (2006) to non-normal errors by using an asymptotic framework of Gaussian experiments. Chernozhukov, Hansen, and Jansson (2009) showed that all members of the weighted average power likelihood ratio tests are admissible, including the AR test.

This paper is concerned with a different optimizing criterion and focuses on the alternative hypothesis determined by an arbitrarily large deviation from the null. Unlike in standard inference problems, in the Weak-IV scenario the power of robust tests is typically far below 100% even for
extremely large deviations from the null. Moreover, it converges as the deviation from the null hypothesis increases to infinity. Hence, by focusing on an arbitrarily large/infinite deviation from the null hypothesis, we can describe the nontrivial power of a test over vast regions of values for the true parameter.

Similarly to Cattaneo, Crump, and Jansson (2012), we use the asymptotic experiment approach. This framework substantially simplifies the analysis by reducing a complex inference problem to that based on a single normally distributed observation. Following Andrews, Moreira, and Stock (2006), we restrict our attention to rotationally invariant and (asymptotically) similar tests, and based on their results we first derive the optimal test against the alternative determined by an arbitrarily large deviation from the null hypothesis when IVs are weak. Focusing on such an alternative allows us to simplify further the asymptotic experiment and describe the optimal test. Unfortunately, the efficient test turns out to be infeasible as it depends on the unknown correlation between the structural and first-stage errors, which we refer to as the degree of endogeneity. We then compare the asymptotic power of popular Weak-IV robust tests (AR, KLM and CLR) against arbitrarily large deviations from the null. Those calculations have to be done numerically.

Surprisingly, we find that, when IVs are weak, the relative performance of the AR test versus the CLR test depends on the degree of endogeneity in the model. For a relatively low degree of endogeneity, the AR test outperforms the CLR test especially when the number of instruments (the degree of overidentification) is large. In that case, the power of the AR test also coincides with the power of the infeasible efficient test. On the other hand, for a relatively high degree of endogeneity, the CLR test outperforms the AR. However, the power of the CLR in that case is substantially below the power of the infeasible efficient test. Our findings show that the CLR test is in fact not nearly optimal even in the classical IV regression model with homoskedastic errors. For the KLM test, we find that when there is a high degree of endogeneity it is as powerful as the CLR test. However, for a low degree of endogeneity the performance of the KLM test is substantially worse than those of the CLR and AR tests.

While the AR test can dominate CLR when IVs are weak, it remains suboptimal if IVs are in fact strong and the model is overidentified. We therefore describe another Weak-IV-robust test, which attains efficiency when instruments are strong and, when IVs are weak, has the same asymptotic power against arbitrarily large deviations from the null as that of the AR test. The test is referred
to as the Conditional Lagrange Multiplier (CLM) test, since its critical values must be simulated on case by case basis using the conditioning device of Moreira (2003).

Our paper is not the first in the Weak-IV literature to consider the power of tests against arbitrarily large deviations from the null. Mills, Moreira, and Vilela (2014) studied such a problem in a finite-sample model with normal errors, and proposed a test that they claimed to be efficient against arbitrarily large deviations from the null. However, our asymptotic framework suggests that the power properties of their test resemble those of the KLM test, and is inferior to the AR and CLR tests when IVs are weak. Our Monte Carlo simulations confirm the prediction.

The paper proceeds as follows. Section 2 sets up the asymptotic experiment framework. Section 3 describes the infeasible optimal rotation-invariant and asymptotically similar test against arbitrarily large deviations from the null hypothesis when IVs are weak. Section 4 describes the asymptotic power properties of the AR, CLR and KLM tests under Weak-IV for arbitrarily large deviations from the null. Section 5 describes the CLM test and its power properties. Section 6 reports a Monte Carlo simulation study.

2 An Asymptotic Experiment for Linear IV Models

Consider the following linear IV model with a single endogenous regressor. The structural and first-stage equations are given by

\[
\begin{align*}
  y_1 &= y_2 \gamma + Z_2 \beta + u, \\
  y_2 &= Z_1 \pi_{1,n} + Z_2 \pi_2 + v,
\end{align*}
\]

where \( y_1, y_2 \in \mathbb{R}^n \) are endogenous variables, \( Z_1 \in \mathbb{R}^{n \times l_1} \) and \( Z_2 \in \mathbb{R}^{n \times l_2} \) are exogenous variables, and \( u, v \in \mathbb{R}^n \) are unobserved error terms. The coefficients \( \gamma \in \mathbb{R}, \beta \in \mathbb{R}^{l_2}, \pi_2 \in \mathbb{R}^{l_2}, \) and \( \pi_{1,n} \in \mathbb{R}^{l_1} \) are unknown parameters, among which the coefficient \( \gamma \) is the structural parameter of interest. Assumption 1 below describes the Weak-IV scenario.

**Assumption 1 (Weak-IV).** \( \pi_{1,n} = n^{-1/2}C, \) where \( C \in \mathbb{R}^{l_1} \) is fixed.

Throughout the paper, we assume that data are iid, the instrumental variables are uncorrelated with the unobserved error terms, the errors in the model are homoskedastic, and the instrumental
variables have finite second order moments:

**Assumption 2.** (a) \(\{(y_{1i}, y_{2i}, Z_{1i}, Z_{2i}), i = 1, ..., n\}\) are iid.

(b) \(E\begin{bmatrix} Z_{1i} \\ Z_{2i} \end{bmatrix} [ u_i \ v_i ] = 0.\)

(c) \(E\begin{bmatrix} u_i^2 & u_i v_i \\ u_i v_i & v_i^2 \end{bmatrix} | Z_{1i}, Z_{12} = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix}\) is a finite and positive definite matrix.

(d) \(E\begin{bmatrix} Z_{1i} Z_{1i}' & Z_{1i} Z_{12}' \\ Z_{2i} Z_{1i}' & Z_{2i} Z_{12}' \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ D_{12}' & D_{22} \end{bmatrix}\) is a finite and positive definite matrix.

By Assumption 2(c),

\[
\frac{1}{\sqrt{n}} \begin{bmatrix} Z_1' M_2 u \\ Z_1' M_2 v \end{bmatrix} \overset{d}{\rightarrow} N \left( 0, \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix} \otimes D_{12} \right), \tag{3}
\]

where \(M_2 = I_n - Z_2(Z_2'Z_2)^{-1}Z_2',\) and \(D_{12} = D_{11} - D_{12}D_{22}^{-1}D_{12}'.\) Note that the Kronecker product structure in (3) is characteristic for homoskedastic models.

Since econometricians are often interested in the coefficient on the endogenous regressor \(y_2,\) consider the null hypothesis \(H_0 : \Delta = \gamma - \gamma_0 = 0.\) The following two statistics, \(S_n\) and \(T_n,\) (and their studentized versions \(S^*_n\) and \(T^*_n\)) will be used repeatedly throughout the paper.\(^1\) We construct \(S_n \in \mathbb{R}^{l_1}\) as the sample covariance between the null restricted residuals and the instrumental variables:

\[
S_n = Z_1' M_2 (y_1 - y_2 \gamma_0) / n
= Z_1' M_2 Z_1 \Delta / \sqrt{n} + Z_1' M_2 (u + \Delta v) / n. \tag{4}
\]

Let \(\hat{\pi}_{1,n}\) be the OLS estimator of \(\pi_{1,n}\) in the first-stage regression. That is,

\[
\hat{\pi}_{1,n} = (Z_1' M_2 Z_1)^{-1} Z_1' M_2 y_2 = \left( \frac{Z_1' M_2 Z_1}{n} \right)^{-1} \frac{Z_1' M_2 v}{n} + \frac{C}{\sqrt{n}}. \tag{5}
\]

\(^1\)Our \(S^*_n\) and \(T^*_n\) correspond to \(S\) and \(T\) in Andrews, Moreira, and Stock (2006).
Define

$$\sigma^2(\Delta) = \sigma_u^2 + \Delta^2 \sigma_v^2 + 2\Delta \sigma_{uv}. $$

Note that $\sigma^2(\Delta)$ is the variance of the null-restricted error term in the expansion of $S_n$ in (4).

The asymptotic distribution implied by in (4) and (5) determines the asymptotic experiment under the Weak-IV scenario:

$$\begin{bmatrix} \sqrt{n}S_n \\ \sqrt{n}\hat{\pi}_{1,n} \end{bmatrix} \overset{a}{\sim} N \begin{bmatrix} \Delta D_{1:2}C \\ C \end{bmatrix}, \begin{bmatrix} \sigma^2(\Delta) D_{1:2} & (\sigma_{uv} + \Delta \sigma_v^2) I_{l_1} \\ (\sigma_{uv} + \Delta \sigma_v^2) I_{l_1} & \sigma_v^2 D_{1:2}^{-1} \end{bmatrix}. $$

As in Kleibergen (2002) and Moreira (2003), it is more convenient to work with an orthogonalized version of $\hat{\pi}_{1,n}$. We therefore construct the second statistic, $T_n \in \mathbb{R}^{l_1}$, to be asymptotically uncorrelated with $S_n$:

$$T_n = \hat{\pi}_{1,n} - \frac{\sigma_{uv} + \Delta \sigma_v^2}{\sigma^2(\Delta)} D_{1:2}^{-1} S_n. $$

The asymptotic distribution of $\sqrt{n} [S_n', T_n']$ is given by

$$N \begin{bmatrix} \Delta D_{1:2}C \\ \sigma_v^2 D_{1:2}^{-1} \end{bmatrix}, \begin{bmatrix} \sigma^2(\Delta) D_{1:2} & 0_{l_1} \\ 0_{l_1} & \sigma_v^2 D_{1:2}^{-1} - \frac{\sigma_{uv}^2 - \sigma_v^2}{\sigma^2(\Delta)} \end{bmatrix}. $$

(6)

Lastly, we studentize $S_n$ and $T_n$ and define

$$S_n^* = D_{1:2}^{-1/2} \sqrt{n} S_n / \sigma(\Delta),$$

$$T_n^* = \sigma(\Delta) D_{1:2}^{1/2} \sqrt{n} T_n / (\sigma_u^2 \sigma_v^2 - \sigma_{uv}^2)^{1/2}. $$

In the above construction, we assumed that $\sigma_{uv} + \Delta \sigma_v^2$, $\sigma^2(\Delta)$, and $\sigma_u^2 \sigma_v^2 - \sigma_{uv}^2$ were known. We can make this assumption because all three terms can be consistently estimated from data despite the fact that they involve unknown $u_i$'s and $\Delta$, which cannot be estimated consistently under the Weak-IV scenario. Let $\Omega$ denote the $2 \times 2$ covariance matrix of the null-restricted reduced-form errors.\(^2\) Then $\sigma^2(\Delta)$ is the upper left element of $\Omega$, $\sigma_{uv} + \Delta \sigma_v^2$ is the upper right element of $\Omega$, and $\sigma_u^2 \sigma_v^2 - \sigma_{uv}^2$ is the lower right element of $\Omega$.

\(^2\)The null-restricted reduced form equations are

$$\begin{bmatrix} y_1 - \gamma_0 y_2 \\ y_2 \end{bmatrix} = Z_1 \begin{bmatrix} \Delta \pi_{1,n} \\ \pi_{1,n} \end{bmatrix} + Z_2 \begin{bmatrix} \Delta \pi_{2} + \beta \\ \pi_{2} \end{bmatrix} + \begin{bmatrix} u + \Delta v \\ v \end{bmatrix}. $$

6
and $\sigma_u^2 \sigma_v^2 - \sigma_{uv}^2$ is the determinant of $\Omega$. Thus, an equivalent representation of our asymptotic experiment is
\[ \begin{bmatrix} S_n^* \\ T_n^* \end{bmatrix} \overset{d}{\sim} N\left( \begin{bmatrix} \frac{\Delta}{\sigma(\Delta)}D_{1/2}^{1/2}C \\ \frac{(\sigma_u^2 + \Delta \sigma_{uv})}{\sigma(\Delta)(\sigma_u^2 \sigma_v^2 - \sigma_{uv}^2)^{1/2}}D_{1/2}^{1/2}C \end{bmatrix}, I_{2\ell_1} \right). \tag{7} \]

Let $S^*$ and $T^*$ denote the limiting versions of $S_n^*$ and $T_n^*$ respectively. Focusing on arbitrarily large deviations from the null, that is, $\Delta \to \infty$, allows us to simplify the asymptotic experiment even further. We consider testing $H_0 : \Delta = 0$ against the alternative $H_1 : \Delta \to +\infty$. Under $H_0$, the asymptotic experiment takes the form:
\[ \begin{bmatrix} S^* \\ T^* \end{bmatrix} \overset{d}{\sim} N\left( \begin{bmatrix} 0 \\ \frac{1}{\sqrt{1-\rho^2}} \lambda \end{bmatrix}, I_{2\ell_1} \right), \tag{8} \]

where $\lambda = D_{1/2}^{1/2}C/\sigma_v$ and $\rho = \sigma_{uv}/(\sigma_u \sigma_v)$. Note that $\|\lambda\|$ is the so-called concentration parameter, which measures the strength of IVs. The correlation coefficient $\rho$ measures the degree of endogeneity in the model. Under $H_1 : \Delta \to +\infty$, our asymptotic experiment takes the following form:
\[ \begin{bmatrix} S^* \\ T^* \end{bmatrix} \overset{d}{\sim} N\left( \begin{bmatrix} \lambda \\ \frac{\rho}{\sqrt{1-\rho^2}} \lambda \end{bmatrix}, I_{2\ell_1} \right), \tag{9} \]

We can similarly analyze $H_1 : \Delta \to -\infty$, when the asymptotic experiment becomes
\[ \begin{bmatrix} S^* \\ T^* \end{bmatrix} \overset{d}{\sim} N\left( -\begin{bmatrix} \lambda \\ \frac{\rho}{\sqrt{1-\rho^2}} \lambda \end{bmatrix}, I_{2\ell_1} \right). \tag{10} \]

Our limiting Gaussian experiment is completely described by only three parameters: $\ell_1$, $\lambda$, and $\rho$. Testing $H_0 : \Delta = 0$ against $H_1 : \Delta \to \infty$ involves only simple hypothesis, and therefore the optimal test can be obtained by the means of the Neyman-Pearson Lemma. Unfortunately, the efficient test will depend on the unknown measure of endogeneity $\rho$. 

7
3 The Optimal Invariant and Asymptotically Similar Test for Arbitrarily Large Deviations

In this section, we derive the asymptotically optimal test against arbitrarily large deviations from the null among all rotation-invariant and asymptotically similar tests. Here rotational invariance means that a test is not affected by any orthonormal transformation of the instruments. Asymptotic similarity means that the asymptotic null rejection rate of a test is not affected by $\pi_{1,n}$, and therefore similar tests are robust to the Weak-IV problem.

Define the following statistics

$$Q_n = \begin{bmatrix} Q_{sn} & Q_{stn} \\ Q'_{stn} & Q_{tn} \end{bmatrix} = [S^*_n, T^*_n]'[S^*_n, T^*_n],$$

and

$$Q = \begin{bmatrix} Q_s & Q_{st} \\ Q'_{st} & Q_t \end{bmatrix} = [S^*, T^*]'[S^*, T^*].$$

Andrews, Moreira, and Stock (2006) showed that every rotation-invariant test can be written as a function of $Q_n$. In addition, they showed that an invariant test is asymptotically similar with significance level $\alpha$ if and only if the asymptotic null rejection rate of such a test equals $\alpha$, conditional on the value of $Q_{tn}$. Therefore, we can restrict our attention to tests that are functions of $Q_n$. The AR, KLM, and CLR tests all have such representations: the AR statistic is given by $Q_{sn}$, the KLM statistic depends on $Q_{stn}$ and $Q_{tn}$, while the CLR depends on $Q_{sn}, Q_{stn}$ and $Q_{tn}$.

From the asymptotic experiment (7), the $l_1 \times 2$ random matrix $[S^*, T^*]$ is multivariate normal with the mean matrix given by

$$M = \begin{bmatrix} \Delta & (\sigma_2^2 + \Delta \sigma_{uv})/\sigma^2 \sigma_v \sigma_w \\ \sigma^2 \sigma_v \sigma_w / (\Delta (\sigma^2_v \sigma^2_w \sigma^2 - \sigma^2_v \sigma^2_w)) + \Delta \sigma_{uv} \end{bmatrix} D_{1/2}^{1/2} C,$$

and the identity covariance matrix. The $2 \times 2$ random matrix $Q$ has a non-central Wishart distribution with the mean matrix of rank one and the identity covariance matrix. Therefore, we can calculate the density of $[Q_s, Q_{st}]$ conditional on $Q_t$, and then use the Neyman-Pearson Lemma to construct the optimal test against $H_1 : \Delta \to +\infty$ or $\Delta \to -\infty$. The method has been used by Andrews, Moreira, and Stock (2006) and Mills, Moreira, and Vilela (2014).
The following proposition describes the optimal rotation-invariant and asymptotically similar test in the Weak-IV scenario against arbitrarily large deviations from the null.

**Proposition 1.** Suppose that Assumptions 1 and 2 hold. Consider testing \( H_0 : \Delta = 0 \) against \( H_1 : \Delta \to \pm \infty \). The test that rejects \( H_0 \) when

\[
POIS_{\infty}(Q_{sn}, Q_{stn}) = Q_{sn} + \frac{2\rho}{\sqrt{1 - \rho^2}} Q_{stn} > \kappa_{\infty}(Q_{tn})
\]

maximizes the asymptotic power over all rotation-invariant and asymptotically similar tests that have asymptotic size \( \alpha \), where \( \kappa_{\infty}(Q_{tn}) \) is the \((1 - \alpha)\)-th quantile of the conditional asymptotic distribution of \( POIS_{\infty}(Q_{sn}, Q_{stn}) \) conditional on \( Q_{tn} \) under \( H_0 \).

**Remarks.**

1. The proposition shows that the efficient test statistic is a linear combination of the \( Q_{sn} \) (or the AR statistic) and \( Q_{stn} \) terms with the weights determined by \( \rho \). When the degree of endogeneity in the model is low, the \( Q_{stn} \) term receives less weight. Moreover, the efficient statistic reduces to the AR statistic when \( y_2 \) is in fact exogenous (when \( \rho = 0 \)). The intuition for this result can be easily understood from the asymptotic experiments in (8) and (9). When \( \rho = 0 \), \( Q_{st} \) has mean zero under both \( H_0 \) and \( H_1 \) because \( S^* \) has mean zero under \( H_0 \) while \( T^* \) has mean zero under \( H_1 \) for \( \rho = 0 \). Hence, \( Q_{st} \) cannot be used for distinguishing between the two hypotheses when \( \rho = 0 \). On the other hand, the mean of \( Q_s \) is different from zero under \( H_1 \) for any value of \( \rho \).

2. The efficient test does not depend neither on the concentration parameter nor the number of instruments.

3. A test statistic of similar form, which involves a linear combination of the AR statistic and another statistic based on \( Q_{stn} \), has been proposed recently in Andrews (2015). However, our alternative hypothesis is different from that in the aforementioned paper, and the resulting efficient test is different as well.

**Proof.** We first characterize the optimal invariant and similar test based on the asymptotic statistic \( Q \). If the optimal test statistic and the critical value are continuous in \( Q \), then by the Continuous
Mapping Theorem, replacing \( Q \) with \( Q_n \) yields a test that attains the asymptotic power envelope among all invariant and asymptotically similar tests.

By (14.17) of Andrews, Moreira, and Stock (2004), the density of \((Q_s, Q_{st}, Q_t)\) at \((q_s, q_{st}, q_t)\) is given by

\[
f_{Q}(q_s, q_{st}, q_t) = K_1 \exp \left(-\text{tr} \left(M'M \right)/2\right) \det(q)^{(l_1-3)/2} \exp \left(-\text{tr} \left(q / 2\right)\right) \\
\times \text{tr} \left(M'Mq \right)^{-(l_1-2)/4} I_{(l_1-2)/2} \sqrt{\text{tr} \left(M'Mq \right)},
\]

where

\[
q = \begin{bmatrix} q_s & q_{st} \\ q'_{st} & q_t \end{bmatrix},
\]

\[
I_v(x) = \left(\frac{x}{2}\right)^v \sum_{j=0}^{\infty} \frac{(x^2/4)^j}{j!\Gamma(v+j+1)},
\]

and \( K_1 \) is a constant only depending on \( l_1 \). Under \( H_0 : \Delta = 0 \), the value of the mean matrix is equal to \( M = [0, \lambda/\sqrt{1-\rho^2}] \). Under \( H_1 : \Delta \to +\infty, M = [\lambda, \rho\lambda/\sqrt{1-\rho^2}] \). Hence, under \( H_0 \) the density of \((Q_s, Q_{st}, Q_t)\) is

\[
f^0_{Q}(q_s, q_{st}, q_t) = K_1 \exp \left(-\frac{\|\lambda\|^2}{2(1-\rho^2)} \right) \det(q)^{(l_1-3)/2} \exp \left(-\frac{q_s + q_t}{2} \right) \\
\times \left(\frac{\|\lambda\|^2}{1-\rho^2} \right)^{-(l_1-2)/4} I_{(l_1-2)/2} \left(\sqrt{\|\lambda\|^2 \frac{q_t}{(1-\rho^2)}} \right).
\]

Under \( H_1 : \Delta \to +\infty \), the density of \((Q_s, Q_{st}, Q_t)\) becomes

\[
f^1_{Q}(q_s, q_{st}, q_t) = K_1 \exp \left(-\frac{\|\lambda\|^2}{2(1-\rho^2)} \right) \det(q)^{(l_1-3)/2} \exp \left(-\frac{q_s + q_t}{2} \right) \\
\times \left(\frac{\|\lambda\|^2 \zeta(q)}{1-\rho^2} \right)^{-(l_1-2)/4} I_{(l_1-2)/2} \left(\sqrt{\|\lambda\|^2 \zeta(q)} \right),
\]

where

\[
\zeta(q) = q_s + \frac{2\rho}{\sqrt{1-\rho^2}} q_{st} + \frac{\rho^2}{1-\rho^2} q_t.
\]
$H_0$ and $H_1$ can be written respectively as

$$f^0_{Q_t}(q_t) = \frac{1}{2} \exp \left(-\frac{\|\lambda\|^2}{2(1-\rho^2)}\right) q_t^{l_1/2} \exp \left(-\frac{q_t}{2}\right) \times \left(\frac{\|\lambda\|^2 q_t}{(1-\rho^2)}\right)^{\frac{l_1}{4}} I_{\frac{l_1}{2}} \left(\sqrt{\frac{\|\lambda\|^2 q_t}{(1-\rho^2)}}\right),$$

$$f^1_{Q_t}(q_t) = \frac{1}{2} \exp \left(-\frac{\rho^2 \|\lambda\|^2}{2(1-\rho^2)}\right) q_t^{l_1/2} \exp \left(-\frac{q_t}{2}\right) \times \left(\frac{\|\lambda\|^2 \rho^2 q_t}{(1-\rho^2)}\right)^{\frac{l_1}{4}} I_{\frac{l_1}{2}} \left(\sqrt{\frac{\|\lambda\|^2 \rho^2 q_t}{(1-\rho^2)}}\right).$$

We can now compute the densities of $(Q_s, Q_{st})$ conditionally on $Q_t$, under $H_0$ and $H_1$ respectively. The resulting likelihood ratio can be written as

$$LR(q) = \frac{f^1_Q(q_s, q_{st}, q_t) / f^1_{Q_t}(q_t)}{f^0_Q(q_s, q_{st}, q_t) / f^0_{Q_t}(q_t)} = \exp \left(-\frac{\|\lambda\|^2}{2}\right) \frac{\left(\frac{\|\lambda\|^2 q_t}{(1-\rho^2)}\right)^{-\frac{l_1}{4}} I_{\frac{l_1}{2}} \left(\sqrt{\frac{\|\lambda\|^2 q_t}{(1-\rho^2)}}\right)}{\left(\frac{\|\lambda\|^2 \rho^2 q_t}{(1-\rho^2)}\right)^{-\frac{l_1}{4}} I_{\frac{l_1}{2}} \left(\sqrt{\frac{\|\lambda\|^2 \rho^2 q_t}{(1-\rho^2)}}\right)} \sum_{j=0}^{\infty} \frac{(\|\lambda\|^2 \zeta(q)/4)^j}{j!(l_1/2+j)!}. \quad (12)$$

Note that the denominator in (12) is a function of $q_t$ and the numerator is an increasing function of $\zeta(q)$ in (11). Therefore, by the Neyman-Pearson Lemma, the optimal test for $H_1 : \Delta \to +\infty$ is to reject $H_0$ when

$$POIS_{\infty}(Q_s, Q_{st}) = Q_s + \frac{2\rho}{\sqrt{1-\rho^2}} Q_{st} > \kappa_{\infty}(Q_t),$$

where the critical value $\kappa_{\infty}(Q_t)$ is the $(1-\alpha)$-th quantile of the distribution of $POIS_{\infty}(Q_s, Q_{st})$ conditional on $Q_t$ and under $H_0$. Exactly the same result holds for $H_1 : \Delta \to -\infty$. □

In a recent paper, Mills, Moreira, and Vilela (2014) proposed a feasible test for arbitrarily large
deviations $\Delta$. The proposed test (see equation (5.9) of their paper) is to reject $H_0$ when

$$Q_{sn} + 2(\det(\Omega))^{-1/2}(\gamma_0 \omega_{22} - \omega_{12})Q_{stn} > \kappa_\infty(Q_{tn}),$$

(13)

where the $\kappa_\infty(Q_{tn})$ is the $(1 - \alpha)$-th quantile of null distribution of the left-hand side in (13) conditional on $Q_{tn}$, and

$$\Omega = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{bmatrix}$$

is the covariance matrix of the reduced-form errors $[u + \gamma v, v]$. However, one can show that with $\Delta \to \infty$, the weight assigned to $Q_{stn}$ also increases to infinity, and the behaviour of the statistic in (13) is effectively determined only by the second term. Hence, the power of their test\(^3\) for arbitrarily large deviations from the null should be equivalent to that of a test based solely on $Q_{stn}$. As discussed in Remark 1 after Proposition 1, such a test has a poor power for arbitrarily large deviations when the degree of endogeneity is low. This is confirmed by our the Monte Carlo results in Section 6 below.

4 Power Comparisons Among Weak-IV-Robust Tests

In this section, we numerically calculate and compare the asymptotic power of the three popular Weak-IV robust tests: AR, CLR, and KLM. Our asymptotic power calculations are based on the asymptotic experiments (8) and (9) constructed in Section 3 to analyze arbitrarily large deviations from the null hypothesis.

4.1 Popular Weak-IV-Robust Tests

The AR test rejects $H_0$ when $AR_n = Q_{sn} > \chi^2_{l_1,1-\alpha}$. By (9) and (10), under $H_1 : \Delta \to \infty$, $AR_n \sim \chi^2_{l_1}(\|\lambda\|^2)$, where $\chi^2_{l_1,1-\alpha}$ is the $(1 - \alpha)$-th quantile of the central $\chi^2$ distribution with $l_1$ degrees of freedom, and $\chi^2_{l_1}(\|\lambda\|^2)$ is the non-central $\chi^2$ distribution with the non-centrality parameter equal to $\|\lambda\|^2$ and $l_1$ degrees of freedom.

Though sometimes refereed to as an LM test, Kleibergen’s KLM is different from the usual

\(^3\)Mills, Moreira, and Vilela (2014) first computed the optimal test for a fixed $\gamma$, and then sent $\gamma$ to infinity. However, $\omega_{12}$ is a function of $\gamma$ and will also go to $\infty$. This is not taken into account in their test statistic in (3).
Lagrange Multiplier (LM) test as the former uses $T^*_n$ (which is asymptotically independent of $S^*_n$) instead of the usual $\hat{\pi}_{1,n}$ to estimate $\pi_{1,n}$. The KLM test rejects $H_0$ when $KLM_n = (Q_{stn})^2/Q_{tn} > \chi_{1,1-\alpha}^2$. Note that the denominator and numerator of KLM are asymptotically independent. Under $H_1: \Delta \rightarrow +\infty$, $KLM_n \approx \chi_1^2(C_{KLM})$. The non-centrality parameter $C_{KLM}$ can be described as

$$C_{KLM} = \lambda \left( \frac{N + \rho}{\sqrt{1-\rho^2}} \right) \left( \frac{N + \rho}{\sqrt{1-\rho^2}} \right)' \lambda,$$

where $N \sim N(0, I_{l_1})$. Note that if $\rho = 0$, the weighting matrix in $C_{KLM}$ becomes entirely random, which leads to poor power performance of the KLM test. We calculate numerically the asymptotic power of the KLM test against arbitrarily large deviations from the null in Section 4.2 using the asymptotic experiment in (9).

The CLR test was proposed by Moreira (2003) and was recommended by Andrews, Moreira, and Stock (2006) for practical use in homoskedastic linear IV models because its power was numerically very close to the power envelope obtained using the weighted average power approach. The test statistic can be written as $CLR_n = 0.5\left( Q_{sn} - Q_{tn} + \sqrt{(Q_{sn} - Q_{tn})^2 + 4Q_{stn}^2} \right)$. The critical value is the $(1-\alpha)$-th quantile of the conditional null distribution of $CLR_n$ given $T^*_n$, which can be easily computed by Monte Carlo methods. Denote such a critical value as $\kappa_{CLR,\alpha}(T^*_n)$. The CLR test rejects $H_0$ when $CLR_n > \kappa_{CLR,\alpha}(T^*_n)$. The asymptotic power of CLR test is difficult to compute analytically, and we numerically evaluate the CLR’s power against arbitrarily large deviations from the null in Section 4.2 for the asymptotic experiment in (9).

4.2 Power Calculations

In this section, we numerically compare the asymptotic power of the AR, KLM, CLR tests against arbitrarily large deviations from the null hypothesis under the Weak-IV scenario. According to our asymptotic experiments (9) and (10), the power depends on three parameters: the number of instruments $l_1$, the strength of identification $\|\lambda\|$, and the degree of endogeneity $\rho$.

Figures 1 to 3 below show the power of the AR, KLM, CLR tests as well as the power envelope derived in Proposition 1 (denoted as BD) for testing $H_0: \Delta = 0$ against $H_1: \Delta \rightarrow +\infty$. Figure 1 is for the model with 5 instruments while Figures 2 and 3 describe the power for 20 and 40
instruments respectively. In each graph, the horizontal axis displays the magnitude of $||\lambda||$ and the vertical axis displays asymptotic rejection probabilities. We allow the parameter $\rho$ (degree of endogeneity) to take values at 0.1, 0.3 and 0.7.

We find that in the Weak-IV scenario and when testing for arbitrarily large deviations from the null hypothesis, the CLR test no longer attains the power envelope (BD): for the values of the degree of endogeneity we considered, the power of the CLR test is strictly below the power envelope determined by the infeasible efficient test.

When the degree of endogeneity is sufficiently low (for example, $\rho = 0.1$), the power of the AR test is very close to the envelope. Moreover, in such cases the AR test outperforms the CLR test for any (even large) number of IVs. The superior performance of the AR for testing can be understood from the form of the asymptotic experiments in (8) and (9). Note that the mean of $Q_{st} = S^*T^*$ is close to zero for small values of $\rho$ under both $H_0$ and $H_1$. Consequently, $Q_{st}$ contains little useful information for distinguishing between $H_0$ and $H_1$ and reduces the testing power when used in the construction of a test statistic.

Surprisingly and somewhat counterintuitively, the difference in the performances of the AR and CLR tests increases with the number of instruments: with a large number of IVs the benefits of using the AR test instead of the CLR test can be very substantial if the degree of endogeneity is low.

In general, however, the relative power performance of the AR and CLR tests depends on the degree of endogeneity. When endogeneity is high ($\rho = 0.7$) the CLR test can substantially outperform the AR test for any number of instruments. The performance gains in this case also increase with the number of IVs.

When endogeneity is high, we find that the asymptotic power of the KLM test is very similar to that of the CLR test. However, for models with low endogeneity the KLM test is inferior to the AR and CLR tests. In those cases, the loss of power from using the KLM test can be very substantial and increasing with the strength of identification as the power the KLM test increases with $||\lambda||$ much slower than that of the AR and CLR tests.

Overall, our findings are very different from those of Andrews, Moreira, and Stock (2006) who, using the weighted average power approach, concluded that the CLR test is nearly efficient in the Weak-IV scenario. Contrary to their results, we find not only that the power of the CLR test lies
below the power envelope, but also that in certain practically relevant situations the AR test can outperform the CLR test.

5 Conditional Lagrange Multiplier (CLM) Test

The results in the previous section show that the AR test can outperform the CLR test in the Weak-IV scenario. Nevertheless, if IVs turns out to be strong, the AR test will be inefficient while the CLR and KLM tests both attain asymptotic efficiency. In this section, we describe a CLM test that attains asymptotic efficiency when IVs are strong and has the same asymptotic power against arbitrarily large deviations from the null as the AR test in the Weak-IV case.

We first describe the notion of efficiency (in presence of nuisance parameters) used in this section, which is adopted from Choi, Hall, and Schick (1996) and applied in the Strong-IV scenario, see also Cattaneo, Crump, and Jansson (2012). Suppose that the following assumption holds.

Assumption 3 (Strong-IV). \( \pi_{1,n} = \pi_1 \), where \( \pi_1 \in \mathbb{R}^{l_1} \setminus \{0_{l_1}\} \), and \( \gamma = \gamma_0 + \Delta/\sqrt{n} \).

In the Strong-IV scenario, the asymptotic experiment in (2) takes the following form:

\[
\begin{bmatrix}
\sqrt{n}S_n \\
\sqrt{n}(T_n - \pi_1)
\end{bmatrix} \overset{q}{\sim} N \left( \begin{bmatrix}
\Delta D_{1,2} \pi_1 \\
-\Delta \frac{\sigma_{uv}}{\sigma_u^2} \pi_1
\end{bmatrix}, \begin{bmatrix}
\sigma_u^2 D_{1,2} & 0_{l_1} \\
0_{l_1} & \frac{\sigma_u^2 \sigma_{uv}^2 - \sigma_u^2 \sigma_{uv}}{\sigma_u^2} D_{1,2}^{-1}
\end{bmatrix} \right). \tag{14}
\]

Note that in this case, the asymptotic variance is independent of \( \Delta \), can be consistently estimated, and therefore can be treated as known. Thus, the only unknown nuisance parameter is \( \pi_1 \). If \( \pi_1 \) were know, one could construct an asymptotically UMP unbiased test of \( H_0 : \Delta = 0 \) against \( H_1 : \Delta \neq 0 \). However, \( \pi_1 \) is unknown and can only be estimated at the \( 1/\sqrt{n} \)-rate. Thus, if the econometrician was to pretend that \( \pi_1 \) is known, he would be making an error of order \( 1/\sqrt{n} \), say \( \tau/\sqrt{n} \). We have effectively revised Assumption 4 as follows:

Assumption 4 (Revised Strong-IV). \( \pi_{1,n} = \pi_1 + \tau/\sqrt{n} \), where \( \pi_1 \in \mathbb{R}^{l_1} \setminus \{0_{l_1}\} \), and \( \gamma = \gamma_0 + \Delta/\sqrt{n} \).

Consider now testing \( H_0 : \Delta = 0, \tau = 0 \) against \( H_1 : \Delta \neq 0 \) or \( \tau \neq 0 \). The optimal test can be found by the Neyman-Pearson Lemma, and its asymptotic power will depend on \( \tau \). According to Choi, Hall, and Schick (1996) by minimizing this power with respect to \( \tau \), i.e. by choosing the
least favorable distortion to \( \pi_1 \), one can obtain the effective efficiency bound for testing \( H_0 : \Delta = 0 \) against \( H_1 : \Delta \neq 0 \). With Assumption 4, the asymptotic experiment becomes

\[
\begin{bmatrix}
\sqrt{n}S_n \\
\sqrt{n}(T_n - \pi_1)
\end{bmatrix} \overset{\sim}{\sim} N\left( \begin{bmatrix}
\Delta D_{1.2}\pi_1 \\
\tau - \Delta \frac{\sigma_{uv}\pi_1}{\sigma_u^2}
\end{bmatrix}, \begin{bmatrix}
\sigma_u^2 D_{1.2} & 0_{l_1} \\
0_{l_1} & \frac{\sigma_u^2\sigma_v^2 - \sigma_{uv}^2}{\sigma_u^2}D_{1.2}^{-1}
\end{bmatrix} \right).
\]

In this model, the likelihood ratio (LR) statistic is given by

\[
LR_n = \frac{\Delta^2 \pi_1' (\sqrt{n}S_n)}{\sigma_u^2} + \left( \tau - \Delta \frac{\sigma_{uv}\pi_1}{\sigma_u^2} \right)' \left( \frac{\sigma_u^2\sigma_v^2 - \sigma_{uv}^2}{\sigma_u^2}D_{1.2}^{-1} \right)^{-1} \sqrt{n}(T_n - \pi_1).
\]

By the Neyman-Pearson Lemma, an asymptotically efficient and unbiased size \( \alpha \) test rejects \( H_0 \) when \( LR_n / AsyVar(LR_n) > \chi^2_{1,1-\alpha} \), where \( \chi^2_{1,1-\alpha} \) is the \((1-\alpha)\)-th quantile of the \( \chi^2 \) distribution.

Under \( H_1 \), the test statistic has a non-central \( \chi^2 \) distribution with the non-centrality parameter given by \( AsyVar(LR_n) \).

Minimizing \( AsyVar(LR_n) \) with respect to \( \tau \) to find the least favourable distortion to \( \pi_1 \), it follows by the results of Choi, Hall, and Schick (1996) that, in the Strong-IV scenario, the effective efficiency bound for testing \( H_0 : \Delta = 0 \) against \( H_1 : \Delta \neq 0 \), for an asymptotically unbiased size \( \alpha \) test is given by

\[
P(\mathcal{X} > \chi^2_{1,1-\alpha}), \text{ where } \mathcal{X} \sim \chi^2_1(\Delta^2 \pi_1'D_{1.2}\pi_1/\sigma_u^2),
\]

and \( \chi^2_1(\delta^2) \) denotes the non-central \( \chi^2 \) distribution with the non-centrality parameter \( \delta^2 \). Note that this efficiency result does not require rotational invariance.

The CLR and KLM tests attain the efficiency bound in (15) in the Strong-IV case. However, the AR test is inefficient in this sense since its asymptotic power is captured by the non-central \( \chi^2 \) distribution that has the same non-centrality parameter \( AsyVar(LR_n) \) but a larger number of degrees of freedom \( l_1 > 1 \) (in the overidentified case). Below, we describe a test that attains the efficiency bound in (15) in the Strong-IV case, and has the same asymptotic power against arbitrarily large deviations from the null in the Weak-IV case as the AR test.
The classical LM test statistic for $H_0 : \gamma = \gamma_0$ in the linear IV model is given by

$$CLM_n = \frac{n (\hat{\pi}_{1,n}^t S_n)^2}{\hat{\omega}_{1,n} \hat{\pi}_{1,n}^t D_{1,2,n} \pi_1},$$

where $\hat{\omega}_{1,n}$ is the estimator of the upper left element of $\Omega$, which is the $2 \times 2$ matrix of the reduced-form errors,$^4$ and $\hat{D}_{1,2,n}$ denotes a consistent estimator of $D_{1,2}$. Let $\hat{S}_n^*$ denote a feasible version of $S_n^*$:

$$\hat{S}_n^* = \hat{D}_{1,2,n}^{-1/2} \sqrt{n} S_n / \hat{\omega}_{1,n}.$$

A feasible version of $T_n$ is

$$\hat{T}_n = \hat{\pi}_{1,n} - \frac{\hat{\omega}_{12,n}}{n^{1/2} \hat{\omega}_{11,n}} \hat{D}_{1,2,n}^{-1/2} \hat{S}_n^*,$$

where $\hat{\omega}_{12,n}$ is the estimator of the upper right element of $\Omega$. Critical values for the CLM test can be generated following the conditioning approach of Moreira (2003). Draw independently $\{\tilde{S}_r^* \sim N(0, I_l) : r = 1, \ldots, R\}$, and compute

$$\tilde{\pi}_{1,r,n} = \hat{T}_n + \frac{\hat{\omega}_{12,n}}{n^{1/2} \hat{\omega}_{11,n}} \hat{D}_{1,2,n}^{-1/2} \tilde{S}_r^*.$$

Next, compute

$$CLM_{r,n} = \frac{(\hat{\pi}_{1,r,n}^t \hat{D}_{1,2,n}^{1/2} \tilde{S}_r^*)^2}{\hat{\pi}_{1,r,n}^t \hat{D}_{1,2,n} \pi_1 \hat{\pi}_{1,r,n}}.$$

The critical value for the asymptotic size $\alpha$ CLM test is the $(1 - \alpha)$-th quantile of the empirical distribution of $\{CLM_{r,n} : r = 1, \ldots, R\}$. One rejects the null hypothesis when $CLM_n$ exceeds the critical value.

It is apparent from the equations in footnote 4 that in the Strong-IV scenario $\hat{\omega}_{11,n} \to_p \sigma_u^2$ and $\hat{\omega}_{12,n} \to_p \sigma_{uv}$. Furthermore, $\hat{\pi}_{1,n} \to_p \pi_1$, $CLM_{r,n} \to_d \chi^2_1$, $\hat{\pi}_{1,n} \to \pi_1$, and $CLM_n \to_d \chi^2_1(\Delta^2 \pi_{1} D_{1,2} \pi_{1} / \sigma_u^2)$. Consequently, the CLM test attains the effective efficiency bound in (15) when the instruments are strong.

---

$^4$Under Assumption 4, the null-restricted reduced-form equations are

$$\begin{bmatrix} y_1 - \gamma_0 y_2 \\ y_2 \end{bmatrix} = Z_1 \begin{bmatrix} n^{-1/2} \Delta \pi_1 \\ \pi_1 \end{bmatrix} + Z_2 \begin{bmatrix} n^{-1/2} \Delta \pi_2 + \beta \\ \pi_2 \end{bmatrix} + \begin{bmatrix} u + n^{-1/2} \Delta v \\ v \end{bmatrix}.$$
Lastly from the equations in footnote 2, it follows that in the Weak-IV scenario $\hat{\omega}_{11,n} \rightarrow p \sigma^2(\Delta)$ and $\hat{\omega}_{12,n} \rightarrow p \sigma_{uv}^2 + \Delta \sigma_v^2$. Hence, the relationship between $\hat{\pi}_{1,n}, S^*_n$, and $T^*_n$ is given by

$$\sqrt{n}\hat{\pi}_{1,n} = \frac{\sigma_{uv}^2 + \Delta \sigma_v^2}{\sigma(\Delta)} D_{1-2}^{-1/2} S^*_n + \frac{\sigma_{uv}^2 - \sigma_v^2}{\sigma(\Delta)} D_{1-2}^{-1/2} T^*_n + o_p(1).$$

Consequently, $\sqrt{n}\hat{\pi}_{1,n} = \sigma_v D_{1-2}^{-1/2} S^*_n + O_p(\Delta^{-1}) + o_p(1)$, and $CLM_n = S^*_n' S^*_n + O_p(\Delta^{-1}) + o_p(1)$. Similar relations hold when simulating the critical values for the CLM test, i.e. $CLM_{r,n} = \tilde{S}^r_{r,n}' \tilde{S}^r_{r,n} + O_p(\Delta^{-1}) + o_p(1)$. This establishes that the asymptotic power against arbitrarily large deviations from the null of the CLM test is equal to that of the AR test when instruments are weak.

6 Monte Carlo Results

In this section, we compare the performance of the AR, CLR, and KLM tests in Monte Carlo simulations. We also report the power of the CLM test proposed in Section 5, the power of the test proposed in Mills, Moreira, and Vilela (2014) for testing arbitrarily large deviations from the null (defined in (3) and denoted MMV in the tables below), and the power of the infeasible asymptotically optimal test for such alternatives (derived in Proposition 1 and denoted BD). In our simulations, we focus on the Weak-IV scenario and large deviations from the null hypothesis.

We generate data from the following specification of the classical IV regression model:

$$y_{1i} = y_{2i}\gamma + u_i,$$
$$y_{2i} = Z_i' \pi_1 + v_i,$$ 

for $i = 1, 2, ..., n$. We set $\pi_1 = (c, ..., c)'$, where $c = \sqrt{||\lambda||^2/n l_1}$, and $||\lambda||^2$ denotes the concentration parameter. The IVs $Z_{1i}$ are drawn independently from the errors from the $N(0, I_{l_1})$ distribution. The errors $(u_i, v_i)$ are generated as $(\epsilon_{1i}, \rho \epsilon_{1i} + \sqrt{1 - \rho^2} \epsilon_{2i})$, where $(\epsilon_{1i}, \epsilon_{2i})'$ are drawn from the $N(0, I_2)$ distribution. For all simulations in this section, we set the true value $\gamma = 0$.

The nominal significance level is set to 5%, and we use the following values of the parameters: the concentration parameter $||\lambda||^2 = 20$, the degree of endogeneity $\rho = 0.1, 0.7$ and the number of instruments $l_1 = 20, 40$. The sample size is $n = 1,000$, the number of Monte Carlo replications is 10,000, and the number of replications for simulating critical values is 5,000.
Table 1 reports the rejection frequencies for a low degree of endogeneity ($\rho = 0.1$) and a large number of IVs ($t_1 = 40$). In this case, when the alternative is far from the null (the deviation exceeds 5 in this example), the AR test outperforms the CLR test by about 5% in terms of its rejection frequencies. The advantage of the AR test over the CLR test remains quite stable for very large deviations (10000 and 50000). In addition, the rejection frequency of the AR test is very close to that of the infeasible optimal test for arbitrarily large deviations (BD). The power of the CLM test is very close to that of the AR test for large deviations from the null as predicted by our theory. As we expected, the performance of the MMV test resembles that of the KLM test when the deviation from the null is large. Both KLM and MMV are substantially less powerful than AR, CLM, and CLR in this scenario.

For Table 2, the number of IVs is reduced to 20. Qualitatively, the results are similar to those reported in Table 1. However, the advantage of the AR (and CLM) tests over CLR is diminished to approximately 4%.

In Table 3 we report the results for a large degree of endogeneity ($\rho = 0.7$). As expected, in this case the CLR test outperforms the AR and CLM test. The difference of their rejection frequencies for large deviations from the null is about 15%. In this case, AR and CLM are also outperformed by the KLM and MMV tests. Neither test is approximately efficient: the difference of the rejection frequencies between the CLR test and the infeasible efficient test is about 7%.

Overall, the Monte Carlo results confirm our theoretical findings.

References


Figure 1: Asymptotic power of the AR, CLR, and KLM tests, and the power of the infeasible efficient test (BD) for 5 IVs and different values of the measures of endogeneity $\rho$ and strength of IVs $\|\lambda\|$.
Figure 2: Asymptotic power of the AR, CLR, and KLM tests, and the power of the infeasible efficient test (BD) for 20 IVs and different values of the measures of endogeneity $\rho$ and strength of IVs $\|\lambda\|$.
Figure 3: Asymptotic power of the AR, CLR, and KLM tests, and the power of the infeasible efficient test (BD) for 40 IVs and different values of the measures of endogeneity $\rho$ and strength of IVs $\|\lambda\|$.
Table 1: Rejection frequencies of the AR, CLR, KLM, Mills, Moreira, and Vilela’s (MMV), CLM, and infeasible efficient (BD) tests at 5% level for $\rho = 0.1, l_1 = 40$

<table>
<thead>
<tr>
<th>$\gamma_0$</th>
<th>AR</th>
<th>CLR</th>
<th>KLM</th>
<th>MMV</th>
<th>CLM</th>
<th>BD</th>
<th>diff. between AR &amp; CLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0774</td>
<td>0.0710</td>
<td>0.0600</td>
<td>0.0778</td>
<td>0.0529</td>
<td>0.0733</td>
<td>0.0064</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1792</td>
<td>0.2586</td>
<td>0.2264</td>
<td>0.0497</td>
<td>0.3302</td>
<td>0.1357</td>
<td>-0.0794</td>
</tr>
<tr>
<td>1</td>
<td>0.3908</td>
<td>0.4734</td>
<td>0.3382</td>
<td>0.0384</td>
<td>0.5121</td>
<td>0.3236</td>
<td>-0.0826</td>
</tr>
<tr>
<td>5</td>
<td>0.6738</td>
<td>0.6235</td>
<td>0.1368</td>
<td>0.1037</td>
<td>0.6719</td>
<td>0.6587</td>
<td>0.0503</td>
</tr>
<tr>
<td>10</td>
<td>0.6803</td>
<td>0.6248</td>
<td>0.1292</td>
<td>0.1263</td>
<td>0.6804</td>
<td>0.6743</td>
<td>0.0555</td>
</tr>
<tr>
<td>50</td>
<td>0.6772</td>
<td>0.6224</td>
<td>0.1334</td>
<td>0.1440</td>
<td>0.6777</td>
<td>0.6805</td>
<td>0.0548</td>
</tr>
<tr>
<td>100</td>
<td>0.6766</td>
<td>0.6224</td>
<td>0.1351</td>
<td>0.1468</td>
<td>0.6774</td>
<td>0.6809</td>
<td>0.0542</td>
</tr>
<tr>
<td>500</td>
<td>0.6756</td>
<td>0.6227</td>
<td>0.1376</td>
<td>0.1499</td>
<td>0.6760</td>
<td>0.6812</td>
<td>0.0529</td>
</tr>
<tr>
<td>1000</td>
<td>0.6750</td>
<td>0.6227</td>
<td>0.1378</td>
<td>0.1507</td>
<td>0.6758</td>
<td>0.6811</td>
<td>0.0523</td>
</tr>
<tr>
<td>5000</td>
<td>0.6748</td>
<td>0.6227</td>
<td>0.1380</td>
<td>0.1510</td>
<td>0.6759</td>
<td>0.6809</td>
<td>0.0521</td>
</tr>
<tr>
<td>10000</td>
<td>0.6748</td>
<td>0.6227</td>
<td>0.1380</td>
<td>0.1510</td>
<td>0.6761</td>
<td>0.6809</td>
<td>0.0521</td>
</tr>
<tr>
<td>50000</td>
<td>0.6748</td>
<td>0.6227</td>
<td>0.1380</td>
<td>0.1510</td>
<td>0.6761</td>
<td>0.6809</td>
<td>0.0521</td>
</tr>
</tbody>
</table>

Table 2: Rejection frequencies of the AR, CLR, KLM, Mills, Moreira, and Vilela’s (MMV), CLM, and infeasible efficient (BD) tests for $\rho = 0.1, l_1 = 20$

<table>
<thead>
<tr>
<th>$\gamma_0$</th>
<th>AR</th>
<th>CLR</th>
<th>KLM</th>
<th>MMV</th>
<th>CLM</th>
<th>BD</th>
<th>diff. between AR &amp; CLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0633</td>
<td>0.0572</td>
<td>0.0508</td>
<td>0.0659</td>
<td>0.0479</td>
<td>0.0619</td>
<td>0.0061</td>
</tr>
<tr>
<td>0.5</td>
<td>0.2000</td>
<td>0.3228</td>
<td>0.3045</td>
<td>0.0355</td>
<td>0.4252</td>
<td>0.1372</td>
<td>-0.1228</td>
</tr>
<tr>
<td>1</td>
<td>0.4844</td>
<td>0.5942</td>
<td>0.4760</td>
<td>0.0282</td>
<td>0.6552</td>
<td>0.3906</td>
<td>-0.1098</td>
</tr>
<tr>
<td>5</td>
<td>0.7850</td>
<td>0.7475</td>
<td>0.1967</td>
<td>0.1341</td>
<td>0.7834</td>
<td>0.7693</td>
<td>0.0375</td>
</tr>
<tr>
<td>10</td>
<td>0.7887</td>
<td>0.7485</td>
<td>0.1760</td>
<td>0.1701</td>
<td>0.7882</td>
<td>0.7833</td>
<td>0.0402</td>
</tr>
<tr>
<td>50</td>
<td>0.7863</td>
<td>0.7452</td>
<td>0.1890</td>
<td>0.2022</td>
<td>0.7868</td>
<td>0.7876</td>
<td>0.0411</td>
</tr>
<tr>
<td>100</td>
<td>0.7855</td>
<td>0.7451</td>
<td>0.1917</td>
<td>0.2055</td>
<td>0.7857</td>
<td>0.7883</td>
<td>0.0404</td>
</tr>
<tr>
<td>500</td>
<td>0.7842</td>
<td>0.7449</td>
<td>0.1952</td>
<td>0.2100</td>
<td>0.7842</td>
<td>0.7884</td>
<td>0.0393</td>
</tr>
<tr>
<td>1000</td>
<td>0.7840</td>
<td>0.7451</td>
<td>0.1954</td>
<td>0.2105</td>
<td>0.7840</td>
<td>0.7885</td>
<td>0.0389</td>
</tr>
<tr>
<td>5000</td>
<td>0.7841</td>
<td>0.7450</td>
<td>0.1963</td>
<td>0.2109</td>
<td>0.7839</td>
<td>0.7885</td>
<td>0.0391</td>
</tr>
<tr>
<td>10000</td>
<td>0.7841</td>
<td>0.7450</td>
<td>0.1963</td>
<td>0.2109</td>
<td>0.7840</td>
<td>0.7885</td>
<td>0.0391</td>
</tr>
<tr>
<td>50000</td>
<td>0.7841</td>
<td>0.7450</td>
<td>0.1963</td>
<td>0.2111</td>
<td>0.7840</td>
<td>0.7885</td>
<td>0.0391</td>
</tr>
</tbody>
</table>
Table 3: Rejection frequencies of the AR, CLR, KLM, Mills, Moreira, and Vilela’s (MMV), CLM, and infeasible efficient (BD) tests at 5% level with $\rho = 0.7, l_1 = 40$

<table>
<thead>
<tr>
<th>$\gamma_0$</th>
<th>AR</th>
<th>CLR</th>
<th>KLM</th>
<th>MMV</th>
<th>CLM</th>
<th>BD</th>
<th>diff. between AR &amp; CLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0774</td>
<td>0.0665</td>
<td>0.0583</td>
<td>0.0677</td>
<td>0.0095</td>
<td>0.0647</td>
<td>0.0199</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3233</td>
<td>0.5750</td>
<td>0.5329</td>
<td>0.5574</td>
<td>0.0037</td>
<td>0.0041</td>
<td>-0.2517</td>
</tr>
<tr>
<td>1</td>
<td>0.9218</td>
<td>0.9270</td>
<td>0.5500</td>
<td>0.6136</td>
<td>0.7764</td>
<td>0.1837</td>
<td>-0.0052</td>
</tr>
<tr>
<td>5</td>
<td>0.8178</td>
<td>0.8880</td>
<td>0.7011</td>
<td>0.8358</td>
<td>0.8017</td>
<td>0.9337</td>
<td>-0.0702</td>
</tr>
<tr>
<td>10</td>
<td>0.7495</td>
<td>0.8600</td>
<td>0.7143</td>
<td>0.8192</td>
<td>0.7457</td>
<td>0.9192</td>
<td>-0.1105</td>
</tr>
<tr>
<td>50</td>
<td>0.6886</td>
<td>0.8344</td>
<td>0.7134</td>
<td>0.8027</td>
<td>0.6899</td>
<td>0.9026</td>
<td>-0.1458</td>
</tr>
<tr>
<td>100</td>
<td>0.6810</td>
<td>0.8306</td>
<td>0.7128</td>
<td>0.8004</td>
<td>0.6821</td>
<td>0.9004</td>
<td>-0.1496</td>
</tr>
<tr>
<td>500</td>
<td>0.6758</td>
<td>0.8280</td>
<td>0.7118</td>
<td>0.7982</td>
<td>0.6746</td>
<td>0.8983</td>
<td>-0.1522</td>
</tr>
<tr>
<td>1000</td>
<td>0.6747</td>
<td>0.8276</td>
<td>0.7117</td>
<td>0.7980</td>
<td>0.6736</td>
<td>0.8981</td>
<td>-0.1529</td>
</tr>
<tr>
<td>5000</td>
<td>0.6743</td>
<td>0.8274</td>
<td>0.7115</td>
<td>0.7979</td>
<td>0.6727</td>
<td>0.8978</td>
<td>-0.1531</td>
</tr>
<tr>
<td>10000</td>
<td>0.6743</td>
<td>0.8274</td>
<td>0.7115</td>
<td>0.7979</td>
<td>0.6726</td>
<td>0.8978</td>
<td>-0.1531</td>
</tr>
<tr>
<td>50000</td>
<td>0.6743</td>
<td>0.8274</td>
<td>0.7115</td>
<td>0.7979</td>
<td>0.6725</td>
<td>0.8978</td>
<td>-0.1531</td>
</tr>
</tbody>
</table>