Quantitative Economics 10 (2019), 457–485 1759-7331/20190457

On optimal inference in the linear IV model

Donald W. K. Andrews
Cowles Foundation, Yale University

Vadim Marmer
Vancouver School of Economics, University of British Columbia

Zhengfei Yu
Faculty of Humanities and Social Sciences, University of Tsukuba

This paper considers tests and confidence sets (CSs) concerning the coefficient on the endogenous variable in the linear IV regression model with homoskedastic normal errors and one right-hand side endogenous variable. The paper derives a finite-sample lower bound function for the probability that a CS constructed using a two-sided invariant similar test has infinite length and shows numerically that the conditional likelihood ratio (CLR) CS of Moreira (2003) is not always "very close," say 0.005 or less, to this lower bound function. This implies that the CLR test is not always very close to the two-sided asymptotically-efficient (AE) power envelope for invariant similar tests of Andrews, Moreira, and Stock (2006) (AMS).

On the other hand, the paper establishes the finite-sample optimality of the CLR test when the correlation between the structural and reduced-form errors, or between the two reduced-form errors, goes to 1 or -1 and other parameters are held constant, where optimality means achievement of the two-sided AE power envelope of AMS. These results cover the full range of (nonzero) IV strength.

The paper investigates in detail scenarios in which the CLR test is not on the two-sided AE power envelope of AMS. Also, theory and numerical results indicate that the CLR test is close to having the greatest average power, where the average is over a specified grid of concentration parameter values and over a pair of alternative hypothesis values of the parameter of interest, uniformly over all such pairs of alternative hypothesis values and uniformly over the correlation between the structural and reduced-form errors. Here, "close" means 0.015 or less for $k \leq 20$, where k denotes the number of IVs, and 0.025 or less for $0 < k \leq 40$.

The paper concludes that, although the CLR test is not always very close to the two-sided AE power envelope of AMS, CLR tests and CSs have very good overall properties.

Donald W. K. Andrews: donald.andrews@yale.edu
Vadim Marmer: vadim.marmer@ubc.ca
Zhengfei Yu: yu.zhengfei.gn@u.tsukuba.ac.jp
Andrews gratefully acknowledges the research support of the National Science Foundation via Grants SES-1355504 and SES-1656313. Marmer gratefully acknowledges the research support of the SSHRC via Grants 410-2010-1394 and 435-2013-0331. The authors thank Yanqin Fan, Patrik Guggenberger, Marcelo Moreira, James MacKinnon, Chen Zhang, and three referees for helpful comments.

 https://doi.org/10.3982/QE1082

Keywords. Conditional likelihood ratio test, confidence interval, infinite length, linear instrumental variables, optimal test, weighted average power, similar test.
JEL classification. C12, C36.

1. Introduction

The linear instrumental variables (IV) regression model is one of the most widely used models in economics. It has been widely studied and considerable effort has been made to develop good estimation and inference methods for it. In particular, following the recognition that standard two stage least squares t tests and confidence sets (CSs) can perform quite poorly under weak IVs (see Dufour (1997), Staiger and Stock (1997), and references therein), inference procedures that are robust to weak IVs have been developed, for example, see Kleibergen (2002) and Moreira (2003, 2009). The focus has been on models with one right-hand side endogenous variable, because this arises most frequently in applications, and on over-identified models, because Anderson and Rubin (1949) (AR) tests and CSs are robust to weak IVs and perform very well in exactly-identified models.

Andrews, Moreira, and Stock (2006) (AMS) develop a finite-sample two-sided AE power envelope for invariant similar tests concerning the coefficient on the right-hand side endogenous variable in the linear IV model under homoskedastic normal errors and known reduced-form variance matrix. They show via numerical simulations that the conditional likelihood ratio (CLR) test of Moreira (2003) has power that is essentially (i.e., up to simulation error) on the power envelope. Chernozhukov, Hansen, and Jansson (2009) show that this power envelope also applies to noninvariant tests provided the envelope is for power averaged over certain direction vectors in a unit sphere. Chernozhukov, Hansen, and Jansson (2009) also showed that the invariant similar tests that generate the two-sided AE power envelope are α-admissible and d-admissible. Mikusheva (2010) provided approximate optimality results for CLR-based CSs that utilize the testing results in AMS. Chamberlain (2007), Andrews, Moreira, and Stock (2008), and Hillier (2009) provided related results.

It is shown in Dufour (1997) that any CS with correct size $1-\alpha$ must have positive probability of having infinite length at every point in the parameter space. The AR and CLR CSs have this property. In fact, simulation results show that in some over-identified contexts the AR CS has a lower probability of having an infinite length than the CLR CS does. For example, consider a model with one right-hand side endogenous variable, k IVs, a concentration parameter λ_v (which is a measure of the strength of the IVs), homoskedastic normal errors, a correlation ρ_{uv} between the structural-equation error and the reduced-form error (for the first-stage equation) equal to zero, and no covariates. When (k, λ_v) equals $(2, 7)$, $(5, 10)$, $(10, 15)$, $(20, 15)$, and $(40, 20)$, the differences between the probabilities that the 95% CLR and AR CSs have infinite length are 0.013, 0.027, 0.037, 0.043, and 0.049, respectively.[1] In fact, one obtains positive differences for all combinations of (k, λ_v) for $k = 2, 5, 10, 20, 40$ and $\lambda_v = 1, 5, 10, 15, 20$. Hence, in these over-identified scenarios the AR CS outperforms the CLR CS in terms of its infinite-length

[1] See Table SM-I in the Online Supplementary Material 2 (Andrews, Marmer, and Yu (2019)) for other parameter combinations.

behavior, which is an important property for CSs. Similarly, one obtains positive (but smaller) differences also when $\rho_{uv} = 0.3$ for the same range of (k, λ_v) values. On the other hand, for $\rho_{uv} = 0.5, 0.7$, and 0.9, the differences are negative over the same range of (k, λ_v) values.

The AR and CLR CSs are based on inverting AR and CLR tests that fall into the class of invariant similar tests considered in AMS. Hence, the simulation results for $\rho_{uv} = 0.0$ and 0.3 raise the question: How can these results be reconciled with the near optimal CLR test and CS results described above? In this paper, we answer this question and related questions concerning the optimality of the CLR test and CS.

The contributions of the paper are as follows. First, the paper shows that the probability that an invariant similar CS has infinite length for a fixed true parameter value β_* equals one minus the power against β_* of the test used to construct the CS as the null value β_0 goes to ∞ or $-\infty$. This leads to explicit formulae for the probabilities that the AR and CLR CSs have infinite length. This result is established in the paper for homoskedastic errors. It is extended in Section 24 in the Online Supplementary Material 1 (Andrews, Marmer, and Yu (2019)) to the case of heteroskedastic and autocorrelated errors.

Second, the paper determines a finite-sample lower bound function on the probabilities that a CS has infinite length for CSs based on invariant similar tests. This lower bound is obtained by using the first result and finding the limit of the power bound in AMS as the null value β_0 goes to ∞ or $-\infty$. The lower bound function is found to be very simple. It is a function only of $|\rho_{uv}|$, λ_v, and k. These results allow one to compare the probabilities that the AR and CLR CSs have infinite length with the lower bound.

Third, simulation results show that the AR and CLR CSs are not always close to the lower bound. This is not surprising for the AR CS, but it is surprising for the CLR CS in light of the AMS results. The probabilities that the CLR CS has infinite length are found to be off the lower bound function by a magnitude that is decreasing in $|\rho_{uv}|$, increasing in k, and are maximized over λ_v at values that correspond to somewhat weak IVs, but not irrelevant IVs. For $\rho_{uv} = 0$, the paper shows (analytically) that the AR test achieves the lower bound function. Hence, for $\rho_{uv} = 0$, the probabilities that the CLR CS has infinite length exceed the lower bound by the same amounts as reported above for the difference between the infinite length probabilities of the CLR and AR CSs for several (k, λ_v) values. On the other hand, for values of $|\rho_{uv}| \geq 0.7$, the CLR CS has probabilities of having infinite length that are close to the lower bound function, 0.010 or less and typically much less, for all (k, λ_v) combinations considered. For values of $|\rho_{uv}| \geq 0.7$, the AR CS has probabilities of having infinite length that are often far from the lower bound. For $|\rho_{uv}| = 0.9$ and certain values of λ_v, they are as large as 0.089, 0.207, 0.288, 0.357, and 0.426 for $k = 2, 5, 10, 20$, and 40, respectively.[2]

The AMS numerical results did not detect scenarios where the CLR test's power is off the two-sided power envelope because AMS focused on power for a fixed null hypothesis and a wide range of alternative values, whereas the probability that a CS has infinite length depends on the underlying tests' power for a fixed true parameter and arbitrar-

[2]See Table SM-I in the Online Supplementary Material 2.

ily distant null hypothesis values. As discussed in Section 4 below, power in these two scenarios is different.

Fourth, the paper derives new optimality properties of the CLR and Lagrange multiplier (LM) tests when $\rho_{uv} \to \pm 1$ or $\rho_{\Omega} \to \pm 1$ with other parameters fixed at any values (with nonzero concentration parameter), where $\rho_{uv} \to \pm 1$ denotes $\rho_{uv} \to 1$ or $\rho_{uv} \to -1$ and likewise for $\rho_{\Omega} \to \pm 1$. In particular, optimality holds for fixed finite nonzero values of the concentration parameter. Optimality here is in the class of invariant similar tests or similar tests and employs the two-sided AE power envelope of AMS. These results are empirically relevant because they are consistent with the numerical results that show that the CLR test is close to the power envelope when $|\rho_{uv}|$ is large, namely, 0.7 and 0.9, but not extremely close to one.

These optimality results hold because taking $\rho_{uv} \to \pm 1$ or $\rho_{\Omega} \to \pm 1$ with other parameters fixed drives the length of the mean vector of the conditioning statistic T, as defined in AMS and below, to infinity. This is the same mechanism that yields asymptotic optimality of the CLR and LM tests when the concentration parameter goes to infinity as $n \to \infty$ (i.e., under strong or semistrong IVs). The results show that arbitrarily large values of the concentration parameter are not needed for limiting optimality of the CLR and LM tests.

Fifth, we simulate power differences between the two-sided AE power envelope of AMS and the power of the CLR test for a fixed alternative value β_* and a range of finite null values β_0 (rather than the power differences as $\beta_0 \to \pm\infty$ discussed above). These power differences are equivalent to the false coverage probability differences between the CLR CS and the corresponding infeasible optimal CS for a fixed true value β_* at incorrect values β_0. We consider a wide range of $(\beta_0, \lambda_v, \rho_{uv}, k)$ values. The maximum (over β_0 and λ_v values) power differences range between $[0.016, 0.061]$ over the (ρ_{uv}, k) values considered. On the other hand, the average (over β_0 and λ values) power differences only range between $[0.002, 0.016]$. This indicates that, although there are some (β_0, λ) values at which the CLR test is noticeably off the power envelope, on average the CLR test's power is not far from the power envelope. The maximum power differences over (β_0, λ) are found to increase in k and decrease in $|\rho_{uv}|$. The λ_v values at which the maxima are obtained are found to (weakly) increase with k and decrease in $|\rho_{uv}|$. The $|\beta_0|$ values at which the maxima are obtained are found to be independent of k and decrease in $|\rho_{uv}|$.

Sixth, the paper considers a weighted average power (WAP) envelope with a uniform weight function over a grid of concentration parameter values λ_v and the same two-point AE weight function over (β, λ) as in AMS. We refer to this as the WAP2 envelope. We determine numerically how close the power of the CLR test is to the WAP2 envelope. We find that the difference between the WAP2 envelope and the average power of the CLR test is in the range of $[0.001, 0.007]$ over all of the $(\beta_0, \beta_*, \rho_{uv}, k)$ values that we consider. Hence, the average power of the CLR test is quite close to the WAP2 envelope.

Other papers in the literature that consider WAP include Wald (1943), Andrews and Ploberger (1994), Andrews (1998), Moreira and Moreira (2013, 2015), Elliott, Müller, and Watson (2015), and papers referenced above. The WAP2 envelope considered here is closest to the WAP envelopes in Wald (1943), AMS, and Chernozhukov, Hansen, and

Jansson (2009) because the other papers listed put a weight function over all of the parameters in the alternative hypothesis, which yields a single weighted alternative density. In contrast, the WAP2 envelope, Wald (1943), AMS, and Chernozhukov, Hansen, and Jansson (2009) consider a family of weight functions over disjoint sets of parameters in the alternative hypothesis, which yields a WAP envelope.

In conclusion, based on our findings, we recommend use of the CLR test and CS in settings with homoskedastic uncorrelated errors. The CLR CS has higher probability of having infinite length than the AR CS in some scenarios, and the CLR test is not a UMP two-sided invariant similar test. But, no such UMP test exists and the CLR CS is close to the two-sided AE power envelope for invariant similar tests when $|\rho_{uv}|$ is not close to zero and is close to the WAP2 envelope for all values of $|\rho_{uv}|$. In settings where the errors may be heteroskedastic or autocorrelated, tests exist that reduce to the CLR test under homoskedastic and uncorrelated errors, for example, see Andrews, Moreira, and Stock (2004), Andrews and Guggenberger (2018), and I. Andrews and Mikusheva (2016). Other tests designed for the heteroskedastic and/or autocorrelated errors are given in Moreira and Moreira (2015) and I. Andrews (2016).

Finally, we point out that the results of this paper illustrate a point that applies more generally than in the linear IV model. In weak identification scenarios, where CSs may have infinite length (or may be bounded only due to bounds on the parameter space), good test performance at a priori implausible parameter values is important for good CS performance at plausible parameter values. More specifically, the probability under an a priori plausible parameter value β_* that a CS has infinite length depends on the power of the test used to construct the CS against β_* when the null value $|\beta_0|$ is arbitrarily large, which may be an a priori implausible null value.

For the computation of CLR CSs, see Mikusheva (2010). For a formula for the power of the CLR test, see Hillier (2009).

The paper is organized as follows. Section 2 specifies the model. Section 3 defines the class of invariant similar tests. Section 4 contrasts the power properties of tests in the scenario where β_0 is fixed and β_* takes on large (absolute) values, with the scenario where β_* is fixed and β_0 takes on large (absolute) values. Section 5 provides a formula for the probability that a CS has infinite length. Section 6 derives a lower bound on the probability that a CS constructed using two-sided invariant similar tests has infinite length. Section 7 reports differences between the probability that the CLR CS has infinite length and the lower bound derived in the previous section. Section 8 proves the optimality results for the CLR test described above. Section 9 reports differences between the power of CLR tests and the two-sided AE power bound of AMS for a wide range of parameter configurations. Section 10 provides comparisons of the power of the CLR test to the WAP2 power envelope described above.

Proofs and additional theoretical results are given in the Online Supplementary Material 1. Additional numerical results are given in the Online Supplementary Material 2.

2. Model

We consider the same model as in Andrews, Moreira, and Stock (2004, 2006) (AMS04, AMS) but, for simplicity and without loss of generality, without any exogenous variables.

The model has one right-hand side endogenous variable, k instrumental variables (IVs), and normal errors with known reduced-form error variance matrix. The model consists of a structural equation and a reduced-form equation:

$$y_1 = y_2\beta + u \quad \text{and} \quad y_2 = Z\pi + v_2, \tag{1}$$

where $y_1, y_2 \in R^n$ and $Z \in R^{n\times k}$ are observed variables; $u, v_2 \in R^n$ are unobserved errors; and $\beta \in R$ and $\pi \in R^k$ are unknown parameters. The IV matrix Z is fixed (i.e., nonstochastic) and has full column rank k. The $n \times 2$ matrix of errors $[u : v_2]$ is i.i.d. across rows with each row having a mean zero bivariate normal distribution with positive variances.

The two corresponding reduced-form equations are

$$\begin{aligned} &Y := [y_1 : y_2] := [Z\pi\beta + v_1 : Z\pi + v_2] = Z\pi a' + V, \quad \text{where} \\ &V := [v_1 : v_2] = [u + v_2\beta : v_2], \quad \text{and} \quad a := (\beta, 1)'. \end{aligned} \tag{2}$$

The distribution of $Y \in R^{n\times 2}$ is multivariate normal with mean matrix $Z\pi a'$, independence across rows, and reduced-form variance matrix $\Omega \in R^{2\times 2}$ for each row. For the purposes of obtaining exact finite-sample results, we suppose Ω is known. As in AMS, asymptotic results for unknown Ω and weak IVs are the same as the exact results with known Ω. The parameter space for $\theta = (\beta, \pi')'$ is R^{k+1}.

We are interested in tests of the null hypothesis $H_0 : \beta = \beta_0$ and CSs for β.

As shown in AMS, $Z'Y$ is a sufficient statistic for $(\beta, \pi')'$. As in Moreira (2003) and AMS, we consider a one-to-one transformation $[S : T]$ of $Z'Y$:

$$\begin{aligned} &S := (Z'Z)^{-1/2} Z'Yb_0 \cdot (b_0'\Omega b_0)^{-1/2} \sim N(c_\beta(\beta_0, \Omega)\cdot \mu_\pi, I_k) \quad \text{and} \\ &T := (Z'Z)^{-1/2} Z'Y\Omega^{-1}a_0 \cdot (a_0'\Omega^{-1}a_0)^{-1/2} \sim N(d_\beta(\beta_0, \Omega)\cdot \mu_\pi, I_k), \quad \text{where} \\ &b_0 := (1, -\beta_0)', \qquad a_0 := (\beta_0, 1)', \qquad \mu_\pi := (Z'Z)^{1/2}\pi \in R^k, \\ &c_\beta(\beta_0, \Omega) := (\beta - \beta_0)\cdot (b_0'\Omega b_0)^{-1/2} \in R, \\ &d_\beta(\beta_0, \Omega) := b'\Omega b_0 \cdot (b_0'\Omega b_0)^{-1/2} \det(\Omega)^{-1/2} \in R, \quad \text{and} \quad b = (1, -\beta)'. \end{aligned} \tag{3}$$

As defined, S and T are independent. Note that S and T depend on the null hypothesis value β_0.

3. Invariant similar tests

As in AMS, we consider tests that are invariant to orthonormal transformations of $[S : T]$, that is, $[S : T] \to [FS : FT]$ for a $k \times k$ orthogonal matrix F. The 2×2 matrix Q is a maximal invariant, where

$$Q = [S : T]'[S : T] = \begin{bmatrix} S'S & S'T \\ S'T & T'T \end{bmatrix} = \begin{bmatrix} Q_S & Q_{ST} \\ Q_{ST} & Q_T \end{bmatrix} \quad \text{and} \quad Q_1 = \begin{pmatrix} S'S \\ S'T \end{pmatrix} = \begin{pmatrix} Q_S \\ Q_{ST} \end{pmatrix}, \tag{4}$$

for example, see Theorem 1 of AMS. Note that Q_1 is the first column of Q and the matrix Q depends on the null value β_0.

The statistic Q has a noncentral Wishart distribution because $[S:T]$ is a multivariate normal matrix that has independent rows and common covariance matrix across rows. The distribution of Q depends on π only through the scalar

$$\lambda := \pi' Z' Z \pi \geq 0. \tag{5}$$

Leading examples of invariant identification-robust tests in the literature include the AR test, the LM test of Kleibergen (2002) and Moreira (2009), and the CLR test of Moreira (2003). The latter test depends on the standard LR test statistic coupled with a "conditional" critical value that depends on Q_T. The LR, LM, and AR test statistics are

$$\begin{aligned} LR &:= \frac{1}{2}\Big(Q_S - Q_T + \sqrt{(Q_S - Q_T)^2 + 4Q_{ST}^2}\Big), \\ LM &:= Q_{ST}^2/Q_T = \big(S'T\big)^2/T'T, \quad \text{and} \quad AR := Q_S/k = S'S/k. \end{aligned} \tag{6}$$

The critical values for the LM and AR tests are $\chi^2_{1,1-\alpha}$ and $\chi^2_{k,1-\alpha}/k$, respectively, where $\chi^2_{m,1-\alpha}$ denotes the $1-\alpha$ quantile of the χ^2 distribution with m degrees of freedom.

A test based on the maximal invariant Q is similar if its null rejection rate does not depend on the parameter π that determines the strength of the IVs Z. As in Moreira (2003), the class of invariant similar tests is specified as follows. Let the $[0,1]$-valued statistic $\phi(Q)$ denote a (possibly randomized) test that depends on the maximal invariant Q. An invariant test $\phi(Q)$ is similar with significance level α if and only if $E_{\beta_0}(\phi(Q)|Q_T = q_T) = \alpha$ for almost all $q_T > 0$ (with respect to Lebesgue measure), where $E_{\beta_0}(\cdot|Q_T = q_T)$ denotes conditional expectation given $Q_T = q_T$ when $\beta = \beta_0$ (which does not depend on π).

The CLR test rejects the null hypothesis when

$$LR > \kappa_{LR,\alpha}(Q_T), \tag{7}$$

where $\kappa_{LR,\alpha}(Q_T)$ is defined to satisfy $P_{\beta_0}(LR > \kappa_{LR,\alpha}(Q_T)|Q_T = q_T) = \alpha$ and the conditional distribution of $Q_1 = (Q_S, Q_{ST})'$ given Q_T is specified in AMS and in (26) in the Online Supplementary Material 1.

The invariance condition discussed above is a *rotational* invariance condition. In some cases, we also consider a *sign* invariance condition. A test that depends on $[S:T]$ is sign invariant if it is invariant to the transformation $[S:T] \to [-S:T]$. A rotation invariant test is also sign invariant if it depends on Q_{ST} only through $|Q_{ST}|$. Tests that are sign invariant are two-sided tests. In fact, AMS shows that the two-sided AE power envelope is identical to the power envelope generated by sign and rotation invariant tests; see (4.11) in AMS.

For simplicity, we will use the term invariant test to mean a rotation invariant test and the term sign and rotation invariant test to describe a test that satisfies both invariance conditions.

The paper also provides some results that apply to tests that satisfy no invariance properties. A test $\phi([S:T])$ (that is not necessarily invariant) is similar with significance level α if and only if $E_{\beta_0}(\phi([S:T])|T=t)=\alpha$ for almost all t (with respect to Lebesgue measure), where $E_{\beta_0}(\cdot|T=t)$ denotes conditional expectation given $T=t$ when $\beta=\beta_0$ (which does not depend on π); see Moreira (2009).

4. Power against distant alternatives compared to distant null hypotheses

In this section, we consider the power properties of tests when $|\beta_*-\beta_0|$ is large, where β_* denotes the true value of β. We compare scenario 1, where β_0 and Ω are fixed, and β_* takes on large (absolute) values, to scenario 2, where β_* and Ω are fixed, and β_0 takes on large (absolute) values. Scenario 1 yields the power function of a test against distant alternatives. Scenario 2 yields the false coverage probabilities of the CS constructed using the test for distant null hypotheses (from the true parameter value β_*). We show that, while power goes to one in scenario 1 as $\beta_*\to\pm\infty$ for fixed β_0 for standard tests, it is not true that power goes to one in scenario 2 as $\beta_0\to\pm\infty$ for fixed β_*. Hence, the power properties of tests are quite different in scenarios 1 and 2.

The numerical power function and power envelope calculations in AMS are all of the types in scenario 1. The difference in power properties of tests between scenarios 1 and 2 suggests that it is worth exploring the properties of tests in scenarios of the latter type as well. We do this in the paper and show that the finding of AMS that the CLR test is essentially on the two-sided AE power envelope and is always at least as powerful as the AR test does not hold when one considers a broader range of null and alternative hypothesis values (β_0,β_*) than considered in the numerical results in AMS.

It is convenient to consider the AR test, which is the simplest test. The AR test rejects $H_0:\beta=\beta_0$ when $S'S>\chi^2_{k,\alpha}$. When the true value is β, the distribution of the $S'S$ statistic is noncentral χ^2 with noncentrality parameter

$$c^2_\beta(\beta_0,\Omega)\cdot\lambda \tag{8}$$

and k degrees of freedom. For the fixed null hypothesis $H_0:\beta=\beta_0$, fixed Ω, and fixed λ, the power at the alternative hypothesis value β_* is determined by $c^2_{\beta_*}(\beta_0,\Omega)$. We have

$$\lim_{|\beta_*|\to\infty} c^2_{\beta_*}(\beta_0,\Omega)=\lim_{|\beta_*|\to\infty}(\beta_*-\beta_0)^2\cdot\left(b_0'\Omega b_0\right)^{-1}=\infty. \tag{9}$$

Hence, the power of the AR test goes to one as $|\beta_*|\to\infty$.

On the other hand, if one fixes the alternative hypothesis value β_* and one considers the limit as $|\beta_0|\to\infty$, then one obtains

$$\begin{aligned}\lim_{|\beta_0|\to\infty} c^2_{\beta_*}(\beta_0,\Omega)&=\lim_{|\beta_0|\to\infty}(\beta_*-\beta_0)^2\cdot\left(b_0'\Omega b_0\right)^{-1}\\&=\lim_{|\beta_0|\to\infty}(\beta_*-\beta_0)^2\cdot\left(\omega_1^2-2\omega_{12}\beta_0+\omega_2^2\beta_0^2\right)^{-1}\\&=1/\omega_2^2,\end{aligned} \tag{10}$$

where ω_1^2, ω_2^2, and ω_{12} denote the $(1,1)$, $(2,2)$, and $(1,2)$ elements of Ω, respectively. Hence, the power of the AR test does not go to one as $|\beta_0|\to\infty$ even though $|\beta_0-\beta_*|\to$

∞. This occurs because the structural equation error variance, $\text{Var}(u_i) = b_0'\Omega b_0$, diverges to infinity as $|\beta_0| \to \infty$.

The differing results in (9) and (10) is easy to show for the AR test, but it also holds for Kleibergen's and Moreira's LM test and Moreira's CLR test. For brevity, we do not provide such results here.

Note that Davidson and MacKinnon (2008, Section 4) provided different, but somewhat related, results to those in this section.[3] They consider power when β_0 is fixed and β_* takes on large (absolute) values (as in scenario 1) but when the correlation ρ_{uv} (between the structural-equation error u and the reduced-form error v_2) is held fixed and the structural equation error variance is estimated. In contrast, the results given here are for the case where the correlation ρ_Ω (between the reduced-form errors v_1 and v_2) is held fixed because ρ_Ω can be consistently estimated, and hence, in large samples can be treated as fixed and known. This is not true for ρ_{uv}. In the Davidson and MacKinnon (2008) scenario, power does not go to one as $\beta_* \to \pm\infty$ for fixed β_0.

5. Probability that a confidence set has infinite length

In this section, we show that the probability that a CS has infinite length is given by one minus the power of the test used to construct the CS as the null value β_0 of the test goes to ∞ or $-\infty$. This provides motivation for interest in the power of tests as $\beta_0 \to \pm\infty$. It shows why high power against distant null hypotheses is highly desirable.

We sometimes make the dependence of Q, S, and T on Y and β_0 explicit and write

$$Q = Q_{\beta_0}(Y) = \left[S_{\beta_0}(Y) : T_{\beta_0}(Y)\right]'\left[S_{\beta_0}(Y) : T_{\beta_0}(Y)\right]. \tag{11}$$

We denote the $(1,1)$, $(1,2)$, and $(2,2)$ elements of $Q_{\beta_0}(Y)$ by $Q_{S,\beta_0}(Y)$, $Q_{ST,\beta_0}(Y)$, and $Q_{T,\beta_0}(Y)$, respectively.

Let

$$\phi\left(Q_{\beta_0}(Y)\right) = 1\left(\mathcal{T}\left(Q_{\beta_0}(Y)\right) > \text{cv}\left(Q_{T\beta_0}(Y)\right)\right) \tag{12}$$

be a (nonrandomized) invariant similar level α test for testing $H_0 : \beta = \beta_0$ for fixed known Ω, where $\mathcal{T}(Q_{\beta_0}(Y))$ is a test statistic and $\text{cv}(Q_{T\beta_0}(Y))$ is a (possibly data-dependent) critical value. Examples include the AR, LM, and CLR tests in (6). Let CS_ϕ be the level $1-\alpha$ CS corresponding to ϕ. That is,

$$\text{CS}_\phi(Y) = \left\{\beta_0 : \phi\left(Q_{\beta_0}(Y)\right) = 0\right\}. \tag{13}$$

We say $\text{CS}_\phi(Y)$ has right (or left) infinite length, which we denote by $\text{RLength}(\text{CS}_\phi(Y)) = \infty$ (or $\text{LLength}(\text{CS}_\phi(Y)) = \infty$), if

$$\exists K(Y) < \infty \text{ such that } \beta \in \text{CS}_\phi(Y)\ \forall \beta \geq K(Y)\ \left(\text{or } \forall \beta \leq -K(Y)\right). \tag{14}$$

We say $\text{CS}_\phi(Y)$ has infinite length, which we denote by $\text{Length}(\text{CS}_\phi(Y)) = \infty$, if it has right and left infinite lengths. A CS with infinite length contains a set of the form $(-\infty, K_1(Y)) \cup (K_2(Y), \infty)$ for some $-\infty < K_1(Y) \leq K_2(Y) < \infty$.

[3] Davidson and MacKinnon (2008) do not consider the probabilities of unbounded CSs or provide optimality results for tests, which are the main focus of this paper.

Let $P_{\beta_*,\pi,\Omega}(\cdot)$ denote probability for events determined by Y when Y has a multivariate normal distribution with means matrix $[\pi\beta_* : \pi] \in R^{2k}$, independence across rows, and variance matrix Ω for each row. Let $P_{\beta_*,\beta_0,\lambda,\Omega}(\cdot)$ denote probability for events determined by Q when $Q := [S : T]'[S : T]$ and $[S : T]$ has the multivariate normal distribution in (3) with $\beta = \beta_*$ and $\lambda = \mu_\pi'\mu_\pi$. In this case, Q has a noncentral Wishart distribution whose density is given in (25) in the Online Supplementary Material 1.

For fixed true value β_* and reduced-form variance matrix Ω, let Σ_* denote the corresponding structural variance matrix of each row of $[u : v_2]$. Let ρ_{uv} denote the correlation between the structural and reduced-form errors, that is, the correlation corresponding to Σ_*. Some calculations show that

$$\begin{aligned} \rho_{uv} &= \frac{\omega_{12} - \omega_2^2\beta_*}{\left(\omega_1^2 - 2\omega_{12}\beta_* + \omega_2^2\beta_*^2\right)^{1/2}\omega_2} \quad \text{and} \\ \Sigma_* &= \begin{bmatrix} \sigma_u^2 & \sigma_u\sigma_v\rho_{uv} \\ \sigma_u\sigma_v\rho_{uv} & \sigma_v^2 \end{bmatrix} = \begin{bmatrix} \omega_1^2 - 2\omega_{12}\beta_* + \omega_2^2\beta_*^2 & \omega_{12} - \omega_2^2\beta_* \\ \omega_{12} - \omega_2^2\beta_* & \omega_2^2 \end{bmatrix}, \end{aligned} \tag{15}$$

where ω_1^2, ω_2^2, and ω_{12} are the elements of Ω; see (32) in the Online Supplementary Material 1. By the first equality in the second line of (15), $\sigma_u^2 = \mathrm{Var}(u_i)$, $\sigma_v^2 = \mathrm{Var}(v_{2i})$, and $\rho_{uv} = \mathrm{Corr}(u_i, v_{2i})$.

It is shown in Lemma 16.1 in the Online Supplementary Material 1 that the limit as $\beta_0 \to \pm\infty$ of $Q_{\beta_0}(Y)$ with Y fixed is

$$\begin{aligned} &Q_{\pm\infty}(Y) \\ &:= \begin{bmatrix} e_2'Y'P_ZYe_2 \cdot \dfrac{1}{\sigma_v^2} & -e_2'Y'P_ZY\Omega^{-1}e_1 \cdot \dfrac{\left(1-\rho_{uv}^2\right)^{1/2}\sigma_u}{\sigma_v} \\ -e_2'Y'P_ZY\Omega^{-1}e_1 \cdot \dfrac{\left(1-\rho_{uv}^2\right)^{1/2}\sigma_u}{\sigma_v} & e_1'\Omega^{-1}Y'P_ZY\Omega^{-1}e_1 \cdot \left(1-\rho_{uv}^2\right)\sigma_u^2 \end{bmatrix}, \end{aligned} \tag{16}$$

which is the same whether $\beta_0 \to +\infty$ or $-\infty$, where $P_Z := Z(Z'Z)^{-1}Z'$, $e_1 := (1, 0)'$, and $e_2 := (0, 1)'$. Let $Q_{T,\pm\infty}(Y)$ denote the $(2, 2)$ element of $Q_{\pm\infty}(Y)$. It is also shown in Lemma 16.1 in the Online Supplementary Material 1 that $Q_{\pm\infty}(Y)$ has a noncentral Wishart distribution with means matrix $\mu_\pi(1/\sigma_v, \rho_{uv}/(\sigma_v(1-\rho_{uv}^2)^{1/2})) \in R^{k\times 2}$ and identity variance matrix.[4]

Theorem 5.1. *Suppose* $\mathrm{CS}_\phi(Y)$ *is a* CS *based on invariant level* α *tests* $\phi(Q_{\beta_0}(Y))$ *whose test statistic and critical value functions,* $\mathcal{T}(q)$ *and* $\mathrm{cv}(q_T)$, *respectively, are continuous at all positive definite* 2×2 *matrices* q *and positive constants* q_T, $P_{\beta_*,\pi,\Omega}(\mathcal{T}(Q_c(Y)) = \mathrm{cv}(Q_{T,c}(Y))) = 0$ *for* $c = +\infty$ *in parts* (a) *and* (c) *below and* $c = -\infty$ *in parts* (b) *and* (c) *below. Then, for all* $\beta_* \in R$, $\lambda \geq 0$, *and* Ω *positive definite*;

(a) $P_{\beta_*,\pi,\Omega}(\mathrm{RLength}(\mathrm{CS}_\phi(Y)) = \infty) = 1 - \lim_{\beta_0\to\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1)$,

(b) $P_{\beta_*,\pi,\Omega}(\mathrm{LLength}(\mathrm{CS}_\phi(Y)) = \infty) = 1 - \lim_{\beta_0\to-\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1)$, *and*

[4]The density of this distribution is given in (27) in the Online Supplementary Material 1.

(c) $P_{\beta_*,\pi,\Omega}(\text{Length}(\text{CS}_\phi(Y)) = \infty) = 1 - \lim_{\beta_0 \to \pm\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1)$.

Comments. (i). For the AR, LM, and LR tests, the continuity conditions on $\mathcal{T}(q)$ and $\text{cv}(q_T)$ hold given their simple functional forms in (6) using the assumption that $q_T > 0$ for the LM statistic and using the continuity of $\kappa_{LR,\alpha}(q_T)$, which holds by the argument in the proof of Theorem 10.1 in Andrews and Guggenberger (2017). We have $P_{\beta_*,\pi,\Omega}(\mathcal{T}(Q_{\pm\infty}(Y)) = \text{cv}(Q_{T,\pm\infty}(Y))) = 0$ for the AR and LM tests because $\text{cv}(Q_{T,\pm\infty}(Y))$ is a constant and $\mathcal{T}(Q_{\pm\infty}(Y))$ is absolutely continuous with respect to Lebesgue measure. For the CLR test, $P_{\beta_*,\pi,\Omega}(\mathcal{T}(Q_{\pm\infty}(Y)) = \text{cv}(Q_{T,\pm\infty}(Y))) = 0$ by the argument given in the proof of Theorem 6.4 in the Online Supplementary Material 1. The AR, LM, and CLR test statistics are sign invariant. Hence, parts (a)–(c) of Theorem 5.1 apply to these tests. Theorem 6.4(a)–(c) below provides formulae for the quantities $\lim_{\beta_0 \to \pm\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1)$, which appear in Theorem 5.1, for the AR, LM, and CLR tests.

(ii). Comment (iii) to Theorem 6.2 below provides a lower bound on $1 - \lim_{\beta_0 \to \infty} P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1)$ over all sign and rotation invariant similar level α tests. Combining this with Theorem 5.1(c) yields a lower bound on the probability that a CS $\text{CS}_\phi(Y)$ based on such tests has Length $= \infty$.

Theorem 13.1 in the Online Supplementary Material 1 provides lower bounds on $1 - \lim_{\beta_0 \to \pm\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1)$ over all invariant similar level α tests. Combining these with Theorem 5.1(a) and (b) yields lower bounds on the probabilities that a CS $\text{CS}_\phi(Y)$ has RLength $= \infty$ based on $\beta_0 \to \infty$ and LLength $= \infty$ based on $\beta_0 \to -\infty$.

(iii). Note that Theorem 5.1 does not impose similarity, just invariance. The results of Theorem 5.1(a) and (b) also hold for a $\text{CS}_\phi(Y)$ that is based on level α tests that are not invariant. Denote such tests by $\phi(S_{\beta_0}(Y), T_{\beta_0}(Y))$ and suppose their test statistic and critical value functions, $\mathcal{T}(s, t)$ and $\text{cv}(t)$, respectively, are continuous at all $k \times 2$ matrices $[s : t]$ and k vectors t and satisfy $P_{\beta_*,\pi,\Omega}(\mathcal{T}(S_c(Y), T_c(Y)) = \text{cv}(T_c(Y))) = 0$ for $c = +\infty$, where $S_{\pm\infty}(Y) := \mp(Z'Z)^{-1/2}Z'Ye_2/\sigma_v$ and $T_{\pm\infty}(Y) := \pm(Z'Z)^{-1/2}Z'Y\Omega^{-1}e_1 \cdot (1 - \rho_{uv}^2)^{1/2}\sigma_u$. In this case, $P_{\beta_*,\pi,\Omega}(\text{RLength}(\text{CS}_\phi(Y)) = \infty) = 1 - \lim_{\beta_0 \to \infty} P_{\beta_*,\beta_0,\pi,\Omega}(\phi([S : T]) = 1)$ for all $\beta_* \in R$, $\lambda \geq 0$, and Ω positive definite, and likewise with LLength$(\cdot)$, $\beta_0 \to -\infty$, and $c = -\infty$ in place of RLength$(\cdot)$, $\beta_0 \to \infty$, and $c = +\infty$.

(iv). By Dufour (1997), all CSs for β with correct size must have positive probability of having infinite length (assuming π is not bounded away from 0). In consequence, expected CS length, which is a standard measure of the performance of a CS, is infinite for all identification-robust CSs. Due to this, Mikusheva (2010) compared CSs based on their expected truncated lengths for various truncation values. The result of Theorem 6.2 below implies that, for two CSs where the right-hand side of Theorem 6.2(c) is smaller for the first CS than the second, the first CS has smaller expected truncated length than the second for sufficiently large truncation values.

(v). Section 24 in the Online Supplementary Material 1 extends Theorem 5.1 to the linear IV model that allows for heteroskedasticity and/or autocorrelation in the errors.

(vi). As indicated in part (c), the right-hand side expressions in parts (a) and (b) are equal.

6. Power bound as $\beta_0 \to \pm\infty$

In this section, we provide two-sided AE power bounds for invariant similar tests as $\beta_0 \to \pm\infty$ for fixed β_*. We obtain these bounds by finding the limit of the power bounds in Theorem 3 of AMS as $\beta_0 \to \pm\infty$ using the dominated convergence theorem. The power bounds also apply to the larger class of similar tests for which invariance is not imposed, provided power is averaged over $\mu_\pi/\|\mu_\pi\|$ vectors using the uniform distribution on the unit sphere in R^k, as in Chernozhukov, Hansen, and Jansson (2009).

Using Theorem 5.1, these results are used to obtain bounds on the probabilities that CSs constructed using sign and rotation invariant similar tests have infinite length. They also are used to obtain bounds on certain average probabilities that similar invariant tests and similar tests have infinite right (or left) length.

This section also determines the power of the AR, LM, and CLR tests as $\beta_0 \to \pm\infty$ and the probabilities that AR, LM, and CLR CSs have infinite length.

6.1 *Density of Q as $\beta_0 \to \pm\infty$*

The density of $Q := [S:T]'[S:T]$ when $[S:T]$ has the multivariate normal distribution in (3) only depends on π through $\lambda := \mu_\pi'\mu_\pi$. Let $f_Q(q;\beta_*,\beta_0,\lambda,\Omega)$ denote this density when $\beta = \beta_*$. It is a noncentral Wishart density with means matrix of rank one and identity covariance matrix, which was first derived by Anderson (1946), equation (6). An explicit expression for $f_Q(q;\beta_*,\beta_0,\lambda,\Omega)$ is given in (25) in the Online Supplementary Material 1.

Now, we determine the limit of the density $f_Q(q;\beta_*,\beta_0,\lambda,\Omega)$ as $\beta_0 \to \pm\infty$. This, the power bound of AMS, and the dominated convergence theorem yield the power bound as $\beta_0 \to \pm\infty$ given below. Define

$$r_{uv} := \frac{\rho_{uv}}{\left(1-\rho_{uv}^2\right)^{1/2}} \quad \text{and} \quad \lambda_v := \lambda/\sigma_v^2 = \mu_\pi'\mu_\pi/\sigma_v^2. \tag{17}$$

Note that λ_v is the concentration parameter, which indexes the strength of the IVs. Let $f_Q(q;\rho_{uv},\lambda_v)$ denote the density of $Q := [S:T]'[S:T]$ when $[S:T]$ has a multivariate normal distribution with means matrix

$$\mu_\pi \cdot (1/\sigma_v, r_{uv}/\sigma_v) \in R^{k\times 2}, \tag{18}$$

all variances equal to one, and all covariances equal to zero. This density also is a noncentral Wishart density with means matrix of rank one and identity covariance matrix. The density depends on r_{uv}, σ_v, and π only through ρ_{uv} and λ_v. An explicit expression for $f_Q(q;\rho_{uv},\lambda_v)$ is given in (27) in Section 12.1 of the Online Supplementary Material 1.

Lemma 6.1. *For any fixed $\beta_* \in R$, $\lambda \geq 0$, Ω positive definite, and 2×2 variance matrix q, $\lim_{\beta_0\to\pm\infty} f_Q(q;\beta_*,\beta_0,\lambda,\Omega) = f_Q(q;\rho_{uv},\lambda_v)$, where ρ_{uv} and λ_v are defined in* (15) *and* (17), *respectively.*

Let $P_{\beta_*,\beta_0,\lambda,\Omega}(\cdot)$ and $P_{\rho_{uv},\lambda_v}(\cdot)$ denote probabilities under the alternative hypothesis densities $f_Q(q;\beta_*,\beta_0,\lambda,\Omega)$ and $f_Q(q;\rho_{uv},\lambda_v)$, respectively, defined above.

6.2 *Two-sided AE power bound as* $\beta_0 \to \pm\infty$

AMS provides a two-sided power envelope for invariant similar tests based on maximizing average power against two points in the alternative hypothesis: (β_*, λ) and (β_{2*}, λ_2). AMS refers to this as the two-sided AE power envelope because given one point (β_*, λ), the second point (β_{2*}, λ_2) is the unique point such that the test that maximizes average power against these two points is a two-sided AE test under strong IV asymptotics. This power envelope is a function of (β_*, λ).

Given (β_*, λ), the second point (β_{2*}, λ_2) satisfies

$$\beta_{2*} = \beta_0 - \frac{d_{\beta_0}(\beta_* - \beta_0)}{d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)} \quad \text{and} \quad \lambda_2 = \lambda \frac{\left(d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)\right)^2}{d_{\beta_0}^2}, \tag{19}$$

where $r_{\beta_0} := e_1'\Omega^{-1}a_0 \cdot (a_0'\Omega^{-1}a_0)^{-1/2}$; see (4.2) of AMS. We let $\text{POIS2}(Q; \beta_0, \beta_*, \lambda)$ denote the optimal average-power test statistic for testing $\beta = \beta_0$ against (β_*, λ) and (β_{2*}, λ_2). Its conditional critical value is denoted by $\kappa_{2,\beta_0}(Q_T)$. For brevity, the formulas for $\text{POIS2}(Q; \beta_0, \beta_*, \lambda)$ and $\kappa_{2,\beta_0}(Q_T)$ are given in Section 17 in the Online Supplementary Material 1.

The limit as $\beta_0 \to \pm\infty$ of the $\text{POIS2}(Q; \beta_0, \beta_*, \lambda)$ statistic is shown in (70) in the Online Supplementary Material 1 to be

$$\begin{aligned}
\text{POIS2}\big(Q; \infty, |\rho_{uv}|, \lambda_v\big) &:= \frac{\psi(Q; \rho_{uv}, \lambda_v) + \psi(Q; -\rho_{uv}, \lambda_v)}{2\psi_2\big(Q_T; |\rho_{uv}|, \lambda_v\big)}, \quad \text{where} \\
\psi(Q; \rho_{uv}, \lambda_v) &:= \exp\big(-\lambda_v\big(1 + r_{uv}^2\big)/2\big)\big(\lambda_v \xi(Q; \rho_{uv})\big)^{-(k-2)/4} \\
&\quad \times I_{(k-2)/2}\big(\sqrt{\lambda_v \xi(Q; \rho_{uv})}\big), \\
\psi_2\big(Q_T; |\rho_{uv}|, \lambda_v\big) &:= \exp\big(-\lambda_v r_{uv}^2/2\big)\big(\lambda_v r_{uv}^2 Q_T\big)^{-(k-2)/4} I_{(k-2)/2}\Big(\sqrt{\lambda_v r_{uv}^2 Q_T}\Big), \quad \text{and} \\
\xi(Q; \rho_{uv}) &:= Q_S + 2r_{uv}Q_{ST} + r_{uv}^2 Q_T,
\end{aligned} \tag{20}$$

where Q, Q_S, Q_{ST}, and Q_T are defined in (4), ρ_{uv} is defined in (15), r_{uv} and λ_v are defined in (17), and $I_\nu(\cdot)$ denotes the modified Bessel function of the first kind of order ν (e.g., see Comment (ii) to Lemma 3 of AMS for more details regarding $I_\nu(\cdot)$).

Let $\kappa_{2,\infty}(q_T)$ denote the conditional critical value of the $\text{POIS2}(Q; \infty, |\rho_{uv}|, \lambda_v)$ test statistic. That is, $\kappa_{2,\infty}(q_T)$ is defined to satisfy

$$P_{Q_1|Q_T}\big(\text{POIS2}\big(Q; \infty, |\rho_{uv}|, \lambda_v\big) > \kappa_{2,\infty}(q_T)|q_T\big) = \alpha \tag{21}$$

for all $q_T \geq 0$, where $P_{Q_1|Q_T}(\cdot|q_T)$ denotes probability under the null density $f_{Q_1|Q_T}(\cdot|q_T)$, which is specified explicitly in (26) in the Online Supplementary Material 1 and does not depend on β_0.

When $\rho_{uv} = 0$, the test based on $\text{POIS2}(Q; \infty, |\rho_{uv}|, \lambda_v)$ is the AR test. This follows because $\xi(Q; 0) = Q_S$, $\psi(Q; 0, \lambda_v)$ is monotone increasing in $\xi(Q; 0)$, and $\psi_2(Q_T; 0, \lambda_v)$ is a constant. Some intuition for this is that $EQ_{ST} = 0$ under the null and $\lim_{|\beta_0| \to \infty} EQ_{ST} =$

0 under any fixed alternative β_* when $\rho_{uv} = 0$.[5] In consequence, Q_{ST} is not useful for distinguishing between H_0 and H_1 when $|\beta_0| \to \infty$ and $\rho_{uv} = 0$. Furthermore, it is shown in (37) and Theorem 13.1 in the Online Supplementary Material 1 that the AR test is also the best one-sided invariant similar test as $\beta_0 \to +\infty$ and as $\beta_0 \to -\infty$ when $\rho_{uv} = 0$.

The following theorem shows that the POIS2$(Q; \infty, |\rho_{uv}|, \lambda_v)$ test provides a two-point average-power bound as $\beta_0 \to \pm\infty$ for any invariant similar test for any fixed (β_*, λ) and Ω.

THEOREM 6.2. *Let* $\{\phi_{\beta_0}(Q) : \beta_0 \to \pm\infty\}$ *be any sequence of invariant similar level* α *tests of* $H_0 : \beta = \beta_0$ *for fixed known* Ω. *For fixed* (β_*, λ), (β_{2*}, λ_2) *defined in* (19), *and* Ω, *the two-sided* AE *power envelope test* POIS2$(Q; \infty, |\rho_{uv}|, \lambda_v)$ *defined in* (20) *and* (21) *satisfies*

$$\begin{aligned}
&\limsup_{\beta_0\to\pm\infty}\big(P_{\beta_*,\beta_0,\lambda,\Omega}\big(\phi_{\beta_0}(Q)=1\big)+P_{\beta_{2*},\beta_0,\lambda_2,\Omega}\big(\phi_{\beta_0}(Q)=1\big)\big)/2\\
&\quad\le P_{\rho_{uv},\lambda_v}\big(\mathrm{POIS2}\big(Q;\infty,|\rho_{uv}|,\lambda_v\big)>\kappa_{2,\infty}(Q_T)\big)\\
&\quad= P_{-\rho_{uv},\lambda_v}\big(\mathrm{POIS2}\big(Q;\infty,|\rho_{uv}|,\lambda_v\big)>\kappa_{2,\infty}(Q_T)\big).
\end{aligned}$$

Comments. (i). The power bound in Theorem 6.2 only depends on (β_*, λ), (β_{2*}, λ_2), and Ω through $|\rho_{uv}|$, which is the absolute magnitude of endogeneity under β_*, and λ_v, which is the concentration parameter.

(ii). The power bound in Theorem 6.2 is strictly less than one. Hence, it is informative.

(iii). For sign and rotation invariant similar tests $\phi_{\beta_0}(Q)$, the lim sup on the left-hand side in Theorem 6.2 is the average of two equal quantities.

(iv). Theorem 6.2 can be extended to cover sequences of similar tests $\{\phi_{\beta_0}(S, T) : \beta_0 \to \pm\infty\}$ that satisfy no invariance properties, using the proof of Theorem 1 in Chernozhukov, Hansen, and Jansson (2009). In this case, the left-hand side probabilities in Theorem 6.2 depend on π or, equivalently $(\lambda, \mu_\pi/\|\mu_\pi\|)$, rather than just λ. In this case, Theorem 6.2 holds with $P_{\beta_*,\beta_0,\lambda,\Omega}(\phi_{\beta_0}(Q) = 1)$ replaced by $\int P_{\beta_*,\lambda,\mu_\pi/\|\mu_\pi\|,\Omega}(\phi_{\beta_0}(S, T) = 1)\, d\,\mathrm{Unif}(\mu_\pi/\|\mu_\pi\|)$ and analogously for the term that depends on (β_{2*}, λ_2), where $P_{\beta_*,\lambda,\mu_\pi/\|\mu_\pi\|,\Omega}(\cdot)$ denotes probability under $(\beta_*, \lambda, \mu_\pi/\|\mu_\pi\|, \Omega)$ and $\mathrm{Unif}(\cdot)$ denotes the uniform measure on the unit sphere in R^k.

6.3 *Lower bound on the probability that a CS has infinite length*

Next, we combine Theorems 5.1 and 6.2 to provide a lower bound on the probability that a sign and rotation invariant similar CS has infinite length. The same lower bound applies to the average probability over (β_*, λ) and (β_{2*}, λ) that a rotation invariant similar CS has right (left) infinite length. For a similar CS with no invariance properties, the same lower bound applies to a different average probability that the CS has right (left) infinite length.

[5]We have $EQ_{ST} = ES'ET$ by independence of S and T, $EQ_{ST} = 0$ under H_0 because $ES = 0$, and $\lim_{|\beta_0|\to\infty} EQ_{ST} = 0$ under β_* because $ET = \mu_\pi d_{\beta_*}(\beta_0, \Omega)$, $\lim_{\beta_0\to\pm\infty} d_{\beta_*}(\beta_0, \Omega) \to \mp r_{uv}/\sigma_v$ and $\|ES\|$ is bounded as $\beta_0 \to \pm\infty$ by Lemma 15.1 in the Online Supplementary Material 1, and $r_{uv} = 0$ when $\rho_{uv} = 0$.

Let $P_{\beta_*,\lambda,\Omega}(\cdot)$ denote probability for events determined by $(Z'Z)^{1/2}Z'Y$ that depend on π only through λ, such as events that are determined by a CS based on invariant tests.

Corollary 6.3. *Suppose* $\mathrm{CS}_\phi(Y)$ *is a* CS*based on invariant similar level* α *tests* $\phi(Q_{\beta_0}(Y))$ *that satisfy the continuity condition in Theorem* 5.1. (a) *For any fixed* $(\beta_*, \lambda, \Omega)$,

$$\begin{aligned}
&\big(P_{\beta_*,\lambda,\Omega}\big(\mathrm{RLength}\big(\mathrm{CS}_\phi(Y)\big)=\infty\big)+P_{\beta_{2*},\lambda_2,\Omega}\big(\mathrm{RLength}\big(\mathrm{CS}_\phi(Y)\big)=\infty\big)\big)/2\\
&\quad\geq 1-P_{\rho_{uv},\lambda_v}\big(\mathrm{POIS2}\big(Q;\infty,|\rho_{uv}|,\lambda_v\big)>\kappa_{2,\infty}(Q_T)\big)\quad \textit{and}\\
&\big(P_{\beta_*,\lambda,\Omega}\big(\mathrm{LLength}\big(\mathrm{CS}_\phi(Y)\big)=\infty\big)+P_{\beta_{2*},\lambda_2,\Omega}\big(\mathrm{LLength}\big(\mathrm{CS}_\phi(Y)\big)=\infty\big)\\
&\quad\geq 1-P_{\rho_{uv},\lambda_v}\big(\mathrm{POIS2}\big(Q;\infty,|\rho_{uv}|,\lambda_v\big)>\kappa_{2,\infty}(Q_T)\big).
\end{aligned}$$

(b) *If the tests* $\phi(Q_{\beta_0}(Y))$ *also are sign invariant, then for any fixed* $(\beta_*, \lambda, \Omega)$,

$$P_{\beta_*,\pi,\Omega}\big(\mathrm{Length}\big(\mathrm{CS}_\phi(Y)\big)=\infty\big)\geq 1-P_{\rho_{uv},\lambda_v}\big(\mathrm{POIS2}\big(Q;\infty,|\rho_{uv}|,\lambda_v\big)>\kappa_{2,\infty}(Q_T)\big).$$

Comments. (i). All three lower bounds in Corollary 6.3 are the same. The different parts of Corollary 6.3 specify different probabilities or average probabilities that have this lower bound.

(ii). A version of Corollary 6.3(a) also holds for a similar CS that does not satisfy any invariance properties. In this case, $P_{\beta_*,\lambda,\Omega}(\mathrm{RLength}(\mathrm{CS}_\phi(Y))=\infty)$ is replaced by $\int P_{\beta_*,\lambda,\mu_\pi/\|\mu_\pi\|,\Omega}(\mathrm{RLength}(\mathrm{CS}_\phi(Y))=\infty)\,d\,\mathrm{Unif}(\mu_\pi/\|\mu_\pi\|)$ and analogously for the other three left-hand side terms that depend on $\mathrm{LLength}(\mathrm{CS}_\phi(Y))$ or (β_{2*}, λ_2). This holds provided the similar level α tests $\phi(S_{\beta_0}(Y), T_{\beta_0}(Y))$ that define the CS satisfy the conditions in Comment (iii) to Theorem 5.1.

6.4 *Power of the AR, LM, and CLR tests as* $\beta_0 \to \pm\infty$

Here, we provide the power of the AR, LM, and CLR tests as $\beta_0 \to \pm\infty$ for fixed (β_*, Ω).

Theorem 6.4. *For fixed true* $(\beta_*, \lambda, \Omega)$, *the* AR, LM, *and* CLR *tests satisfy*

(a) $\lim_{\beta_0\to\pm\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(AR > \chi^2_{k,1-\alpha}/k) = P_{\rho_{uv},\lambda_v}(AR > \chi^2_{k,1-\alpha}/k) = P(\chi^2_k(\lambda_v) > \chi^2_{k,1-\alpha})$,

(b) $\lim_{\beta_0\to\pm\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(LM > \chi^2_{1,1-\alpha}) = P_{\rho_{uv},\lambda_v}(LM > \chi^2_{1,1-\alpha})$, *and*

(c) $\lim_{\beta_0\to\pm\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(LR > \kappa_{LR,\alpha}(Q_T)) = P_{\rho_{uv},\lambda_v}(LR > \kappa_{LR,\alpha}(Q_T))$,

where AR, LM, *and* LR *are defined as functions of* Q *in* (6), $\chi^2_{m,1-\alpha}$ *is the* $1-\alpha$ *quantile of the* χ^2_m *distribution, and* $\chi^2_m(\lambda_v)$ *is a noncentral* χ^2_m *random variable with noncentrality parameter* λ_v.

Comments. (i). By Theorem 5.1(c), Theorem 6.4 provides the probabilities that the AR, LM, and CLR CSs have infinite length when the true parameters are $(\beta_*, \lambda, \Omega)$. These probabilities depend only on $(|\rho_{uv}|, \lambda_v)$. For the AR CS, they only depend on λ_v.

(ii). As pointed out by a referee, the AR CS has infinite length when the first-stage F test strictly fails to reject $H_0 : \pi = 0_k$, meaning that $y_2' P_Z y_2/\omega_2 < \chi^2_{k,1-\alpha}$ (with a strict inequality). When the first-stage F test rejects $H_0 : \pi = 0_k$, that is, $y_2' P_Z y_2/\omega_2 > \chi^2_{k,1-\alpha}$, the AR CS has finite length. When $y_2' P_Z y_2/\omega_2 = \chi^2_{k,1-\alpha}$, the AR CS can have infinite length, right length, or left length, or have finite length.[6] Results in Mikusheva (2010, Proofs of Theorems 1 and 2) provide expressions for the cases where the LM and CLR CSs have infinite lengths, but they do not seem to have as simple intuitive interpretations as for the AR CS.

7. Comparisons of probabilities that confidence sets have infinite length

Next, we investigate how close are the probabilities that the CLR CS has infinite length to the lower bound in Corollary 6.3. The probabilities that the CLR CS has infinite length are given by Theorems 5.1(c) and 6.4(c). Without loss of generality, we take $\sigma_v^2 = \omega_{22} = 1$ and $\omega_{11} = 1$ in this section. Let POIS2 refer to the tests that generate the two-sided AE power envelope of AMS. These tests depend on the alternative (β_*, λ) considered and Ω. Let POIS2$_\infty$ refer to the tests in (20), which are the limits as $\beta_0 \to \pm\infty$ of the POIS2 tests. These tests depend on β_* (through $|\rho_{uv}|$) and λ. Let POIS2 and POIS2$_\infty$ CSs refer to the CSs constructed by inverting the POIS2 and POIS2$_\infty$ tests. These CSs are infeasible because they depend on knowing (β_*, λ).

Table 1 reports differences in simulated probabilities that the CLR and POIS2$_\infty$ CSs have infinite lengths. The latter provide a lower bound on infinite-length probabilities for CSs based on sign and rotation invariant tests, such as the CLR CS, by Corollary 6.3(b). Hence, these differences are necessarily nonnegative. The results cover $k = 2, 5, 10, 20, 40$, a range of λ values between 1 and 60 depending on the value of k, and $\rho_{uv} = 0, 0.3, 0.5, 0.7, 0.9$. Table 1 also reports the probabilities that the POIS2 CS has infinite length for the same k and λ values and a subset of the ρ_{uv} values, namely, $0, 0.7, 0.9$. The true value of β_* is taken to be 0 without loss of generality by Section 22 in the Online Supplementary Material 1. The results for negative and positive ρ_{uv} values are the same by Section 22 in the Online Supplementary Material 1, and hence, results for negative ρ_{uv} are not reported. The number of simulation repetitions employed is 50,000. The critical values are determined using 100,000 simulation repetitions.

The results show that the CLR CS is not close to the lower bound in some parameter scenarios. In particular, the differences in probabilities of infinite length (DPIL's) between the CLR and the POIS2$_\infty$ CSs are positive for numerous combinations of (k, λ, ρ_{uv}). The bold face numbers in Table 1 give the largest DPIL for each (k, λ, ρ_{uv}) combination. The DPILs are increasing in k, decreasing in $|\rho_{uv}|$, and maximized in the middle of the range of λ values considered. For example, for $(k, \rho_{uv}) = (2, 0)$, DPIL $\in [0.002, 0.013]$ over the λ values considered, whereas for $(k, \rho_{uv}) = (5, 0)$, DPIL $\in$

[6] These results hold because (i) the AR test strictly fails to reject $H_0 : \beta = \beta_0$ when $S'S < \chi^2_{k,1-\alpha}$ iff $b_0' Y' P_Z Y b_0 < b_0' \Omega b_0 \chi^2_{k,1-\alpha}$ iff $a\beta_0^2 + b\beta_0 + c < 0$, where $a := y_2' P_Z y_2 - \omega_2 \chi^2_{k,1-\alpha}$, $b := -2(y_1' P_Z y_2 - \omega_{12} \chi^2_{k,1-\alpha})$, and $c := y_1' P_Z y_1 - \omega_1 \chi^2_{k,1-\alpha}$, using (3) and some calculations, and (ii) the AR CS has infinite length when $a < 0$. When $a = 0$, the AR CS has infinite right length if $b < 0$, infinite left length if $b > 0$, infinite length if $b = 0$ and $c \leq 0$, and finite length if $b = 0$ and $c > 0$. For related results, see Dufour and Taamouti (2005).

TABLE 1. Differences in probabilities of infinite-length CI's for the CLR and POIS2$_\infty$ CI's, and Probabilities of Infinite-Length POIS2$_\infty$ CI's as Functions of k, λ and ρ_{uv}.

		CLR–POIS2$_\infty$					POIS2$_\infty$		
k	λ	$\rho_{uv}=0$	0.3	0.5	0.7	0.9	$\rho_{uv}=0$	0.7	0.9
2	1	0.002	0.003	0.003	0.001	0.002	0.867	0.862	0.851
2	3	0.007	0.008	0.003	0.004	**0.004**	0.680	0.654	0.614
2	5	0.011	**0.010**	**0.005**	**0.004**	0.002	0.497	0.452	0.407
2	7	**0.013**	0.009	0.004	0.004	0.003	0.345	0.291	0.256
2	10	0.012	0.007	0.004	0.003	0.002	0.182	0.138	0.117
2	15	0.007	0.004	0.002	0.001	0.001	0.056	0.034	0.029
2	20	0.003	0.002	0.001	0.000	0.000	0.015	0.008	0.006
5	1	0.003	0.002	0.001	0.001	0.003	0.902	0.900	0.884
5	3	0.010	0.007	0.003	0.001	0.005	0.779	0.752	0.670
5	5	0.020	0.010	0.003	0.004	**0.004**	0.639	0.571	0.459
5	7	0.026	0.013	0.005	**0.006**	0.002	0.502	0.404	0.295
5	10	**0.027**	**0.014**	**0.006**	0.005	0.001	0.323	0.214	0.139
5	12	0.027	0.013	0.006	0.004	0.001	0.230	0.133	0.082
5	15	0.023	0.011	0.005	0.003	0.000	0.132	0.061	0.035
5	20	0.012	0.005	0.003	0.001	0.000	0.047	0.014	0.008
5	25	0.006	0.003	0.001	0.000	0.000	0.015	0.003	0.002
10	1	0.002	0.002	0.001	0.001	0.003	0.918	0.917	0.904
10	5	0.018	0.011	0.005	0.003	**0.007**	0.733	0.673	0.526
10	10	0.035	**0.018**	0.008	**0.005**	0.002	0.461	0.317	0.173
10	15	**0.037**	0.017	**0.008**	0.005	0.001	0.242	0.110	0.046
10	17	0.034	0.016	0.007	0.004	0.000	0.177	0.069	0.026
10	20	0.026	0.015	0.006	0.002	0.000	0.109	0.033	0.011
10	25	0.016	0.008	0.003	0.001	0.000	0.043	0.008	0.002
10	30	0.008	0.004	0.002	0.000	0.000	0.016	0.002	0.000
20	1	0.003	0.002	0.001	0.000	0.002	0.929	0.930	0.921
20	5	0.017	0.012	0.004	0.003	**0.008**	0.806	0.768	0.617
20	10	0.035	0.021	0.008	0.008	0.003	0.597	0.462	0.240
20	15	**0.043**	**0.023**	**0.010**	**0.009**	0.002	0.393	0.211	0.070
20	20	0.042	0.021	0.009	0.005	0.001	0.226	0.079	0.018
20	25	0.033	0.016	0.007	0.003	0.000	0.116	0.024	0.004
20	30	0.023	0.011	0.004	0.002	0.000	0.053	0.007	0.001
20	40	0.007	0.003	0.001	0.000	0.000	0.010	0.001	0.000
40	1	0.001	0.000	0.000	−0.000	−0.001	0.936	0.936	0.932
40	5	0.011	0.008	0.005	0.003	**0.010**	0.861	0.837	0.717
40	10	0.030	0.016	0.006	0.010	0.004	0.721	0.615	0.354
40	15	0.046	0.024	0.011	**0.011**	0.002	0.553	0.371	0.128
40	20	**0.049**	**0.028**	**0.013**	0.010	0.001	0.394	0.186	0.038
40	30	0.043	0.022	0.010	0.004	0.000	0.155	0.029	0.002
40	40	0.022	0.010	0.004	0.001	0.000	0.046	0.003	0.000
40	60	0.003	0.001	0.000	0.000	0.000	0.002	0.000	0.000

$[0.003, 0.027]$ and for $(k, \rho_{uv}) = (40, 0)$, DPIL $\in [0.001, 0.049]$.[7] Hence, k has a noticeable effect on the magnitude of nonoptimality of the CLR CS with larger values of k leading to larger non-optimality. For $(k, \lambda) = (5, 10)$, we have DPIL $\in [0.001, 0.027]$ over the ρ_{uv} values considered, and for $(k, \lambda) = (20, 15)$, we have DPIL $\in [0.002, 0.043]$ over the ρ_{uv} values considered. Hence, $|\rho_{uv}|$ also has a noticeable effect on the magnitude of nonoptimality of the CLR CS in terms of DPILs with nonoptimality greatest at $\rho_{uv} = 0$.[8]

8. Optimality of CLR and LM tests as $\rho_{uv} \to \pm 1$ or $\rho_{\Omega} \to \pm 1$

The results of Table 1 show that the magnitude of nonoptimality of the CLR CS decreases as $|\rho_{uv}|$ increases to 1. This raises the question of whether CLR tests are optimal in some sense in the limit as $|\rho_{uv}| \to 1$. In this section, we show that this is indeed the case, not just for power as $\beta_0 \to \pm\infty$, but uniformly over all (β_0, β_*) parameter values in a two-sided AE power sense.

Let ρ_{Ω} denote the correlation parameter corresponding to the reduced-form variance matrix Ω, that is, $\rho_{\Omega} := \omega_{12}/(\omega_1 \omega_2)$.

In this section, we provide parameter configurations under which the CLR and LM tests have optimality properties. The results cover the case of strong and semistrong identification (where $\lambda \to \infty$). They cover the cases where $\rho_{uv} \to \pm 1$ or $\rho_{\Omega} \to \pm 1$ for (almost) any fixed values of the other parameters, which includes weak identification of any strength. And, they cover the cases where $(\rho_{uv}, \beta_0) \to (\pm 1, \pm\infty)$ or $(\rho_{\Omega}, \beta_0) \to (\pm 1, \pm\infty)$ and the other parameters are fixed at (almost) any values, which also includes weak identification.

In somewhat related results, Chernozhukov, Hansen, and Jansson (2009) showed that the CLR and LM tests can be written as the limits of certain WAP LR tests, which indicate that they are at least close to being admissible.

Let $d^2_{\beta_*} := d^2_{\beta_*}(\beta_0, \Omega)$ and $c^2_{\beta_*} := c^2_{\beta_*}(\beta_0, \Omega)$, where $d_{\beta}(\beta_0, \Omega)$ and $c_{\beta}(\beta_0, \Omega)$ are defined in (3). As in Section 6.2, let $\mathrm{POIS2}(Q; \beta_0, \beta_*, \lambda)$ and $\kappa_{2,\beta_0}(Q_T)$ denote the optimal average-power test statistic and its data-dependent critical value. Let $\chi^2_1(c^2_\infty)$ denote a noncentral χ^2_1 random variable with noncentrality parameter c^2_∞.

Theorem 8.1. *Consider any sequence of null parameters β_0 and true parameters $(\beta_*, \lambda, \Omega)$ such that $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty \in R\backslash\{0\}$. Then,*

(a) $P_{\beta_*,\beta_0,\lambda,\Omega}(\mathrm{POIS2}(Q; \beta_0, \beta_*, \lambda) > \kappa_{2,\beta_0}(Q_T)) \to P(\chi^2_1(c^2_\infty) > \chi^2_{1,1-\alpha})$,

(b) $P_{\beta_*,\beta_0,\lambda,\Omega}(LR > \kappa_{LR,\alpha}(Q_T)) \to P(\chi^2_1(c^2_\infty) > \chi^2_{1,1-\alpha})$, *and*

(c) $P_{\beta_*,\beta_0,\lambda,\Omega}(LM > \chi^2_{1,1-\alpha}) \to P(\chi^2_1(c^2_\infty) > \chi^2_{1,1-\alpha})$.

[7]The simulation standard deviations of the DPILs are in the range of $[0.0000, 0.0014]$ with most being in the range of $[0.0004, 0.0012]$; see Table SM-I in the Online Supplementary Material 2.

[8]Table SM-I in the Online Supplementary Material 2 shows that the differences in probabilities that the AR and POIS2 CSs have infinite length are very large for large ρ_{uv} values for some λ values. For example, for $\rho_{uv} = 0.9$, they are as large as 0.089, 0.207, 0.288, 0.357, 0.426 for $k = 2, 5, 10, 20, 40$, respectively, for some λ values. As shown above, $AR = \mathrm{POIS2}$ when $\rho_{uv} = 0$, so the differences are zero in this case and they increase in $|\rho_{uv}|$ for given (k, λ).

Comments. (i). Theorem 8.1 shows that the CLR and LM tests have the same limit power as the POIS2 test. Theorem 8.1 provides both finite-sample limiting optimality results, where n is fixed and the limits are determined by sequences of parameters $(\beta_0, \beta_*, \lambda, \Omega)$, and large-sample limiting optimality results, where the limits are determined by sequences of sample sizes n and parameters $(\beta_0, \beta_*, \lambda, \Omega)$.

(ii). By Corollary 1 of AMS, for any invariant similar test $\phi(Q)$, for any $(\beta_*, \beta_0, \lambda, \Omega)$,

$$\frac{1}{2}\big(P_{\beta_*,\beta_0,\lambda,\Omega}\big(\phi(Q)=1\big)+P_{\beta_{2*},\beta_0,\lambda_2,\Omega}\big(\phi(Q)=1\big)\big) \leq P_{\beta_*,\beta_0,\lambda,\Omega}\big(\mathrm{POIS2}(Q;\beta_0,\beta_*,\lambda)>\kappa_{2,\beta_0}(Q_T)\big). \tag{22}$$

That is, the POIS2 test determines the two-sided AE average power envelope of AMS for invariant similar tests, where the average is over (β_*, λ) and (β_{2*}, λ_2). A fortiori, by Theorem 1 of Chernozhukov, Hansen, and Jansson (2009), for any similar test $\phi([S:T])$ (that is not necessarily invariant), for any $(\beta_*, \beta_0, \lambda, \Omega)$, (22) holds with $P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q)=1)$ replaced by the power average $\int P_{\beta_*,\beta_0,\lambda,\mu_\pi/\|\mu_\pi\|,\Omega}(\phi([S:T])=1)\,d\,\mathrm{Unif}(\mu_\pi/\|\mu_\pi\|)$ and likewise for the second left-hand side summand in (22). Hence, the POIS2 test also determines this average power envelope for similar tests.

These results and Theorem 8.1 show that the CLR and LM tests achieve these average power envelopes for all $(\beta_*, \beta_0, \lambda, \Omega)$ asymptotically when $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty \neq 0$.

(iii). The power envelopes in Comment (ii) translate immediately into false-coverage-probability lower bounds for CSs based on invariant similar tests and similar tests. Specifically, one minus the left-hand side in (22), which equals the average false-coverage probability of the point β_0 by the CS based on $\phi(Q)$, where the average is over the truth being (β_*, λ) and (β_{2*}, λ_2), is greater than or equal to one minus the right-hand side in (22). In the case of noninvariant similar tests, the bound is on the average of the false-coverage probabilities of the CS with averaging over (β_*, λ) and (β_{2*}, λ_2) and $\mu_\pi/\|\mu_\pi\|$ in the unit sphere in R^k. Thus, Theorem 8.1 shows that the CLR and LM CSs have optimal average false-coverage-probability properties asymptotically when $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty \neq 0$.

(iv). Theorem 8.1 does not apply when the IVs are completely irrelevant, that is, $\lambda = 0$, because $\lambda = 0$ implies that $c_\infty = 0$. However, Theorem 8.1 does cover some cases where the IVs can be arbitrarily weak; see Theorem 8.2 below.

Next, we provide conditions under which $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty \in R\backslash\{0\}$, as is assumed in Theorem 8.1. First, if β_0 and Ω are fixed, Ω is nonsingular, and (β_*, λ) satisfy $\lambda \to \infty$ and

$$\lambda^{1/2}(\beta_* - \beta_0) \to L \in R\backslash\{0\} \quad \text{as } \lambda \to \infty, \tag{23}$$

then $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty \in R\backslash\{0\}$ with $c_\infty = L(b_0'\Omega b_0)^{-1/2}$. Here, L indexes the local alternatives against which the tests have nontrivial power. This result covers the usual strong IV case in which π is fixed, $Z'Z$ depends on n, and $\lambda = \pi' Z' Z\pi \to \infty$ as $n \to \infty$.

The scenario in (23) also covers cases where $\pi = \pi_n \to 0$ as $n \to \infty$, but sufficiently slowly that $\lambda = \pi_n' Z' Z\pi_n \to \infty$ as $n \to \infty$, which covers "semistrong" identification. As far as we are aware, this is the only optimality property in the literature for tests under semistrong identification. The scenario in (23) also covers a finite-sample, that is, fixed n, cases in which $Z'Z$ is fixed, π diverges, that is, $\|\pi\| \to \infty$, and $\lambda_{\min}(Z'Z) > 0$. In these cases, $\lambda = \pi' Z' Z \pi \to \infty$ as $\|\pi\| \to \infty$.

Second, the most novel cases in which Theorem 8.1 applies are when $\rho_{uv} \to \pm 1$ or $\rho_\Omega \to \pm 1$. The next result shows that $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty \in R\backslash\{0\}$ when $\rho_{uv} \to \pm 1$ or $\rho_\Omega \to \pm 1$ and the other parameters are fixed at (almost) any values. It also shows that this holds when $(\rho_{uv}, \beta_0) \to (1, \pm\infty)$ or $(-1, \pm\infty)$ or $(\rho_\Omega, \beta_0) \to (1, \pm\infty)$ or $(-1, \pm\infty)$ and the other parameters are fixed at (almost) any values.

Theorem 8.2. (a) *Suppose the parameters* β_0, β_*, $\sigma_u > 0$, $\sigma_v > 0$, *and* $\lambda > 0$ *are fixed,* $\rho_{uv} \in (-1, 1)$, *and* $\rho_{uv} \to \pm 1$. *Then* (i) $\lim_{\rho_{uv} \to \pm 1} \lambda^{1/2} c_{\beta_*} = \lambda^{1/2}(\beta_* - \beta_0)/|\sigma_u \pm (\beta_* - \beta_0)\sigma_v|$ *and* (ii) $\lim_{\rho_{uv} \to \pm 1} \lambda d^2_{\beta_*} = \infty$ *provided* $\beta_* - \beta_0 \neq \mp \sigma_u/\sigma_v$.

(b) *Suppose the parameters* β_0, β_*, $\omega_1 > 0$, $\omega_2 > 0$, *and* $\lambda > 0$ *are fixed,* $\rho_\Omega \in (-1, 1)$, *and* $\rho_\Omega \to \pm 1$. *Then* (i) $\lim_{\rho_\Omega \to \pm 1} \lambda^{1/2} c_{\beta_*} = \lambda^{1/2}(\beta_* - \beta_0)/|\omega_1 \mp \omega_2 \beta_0|$ *provided* $\beta_0 \neq \pm\omega_1/\omega_2$ *and* (ii) $\lim_{\rho_\Omega \to \pm 1} \lambda d^2_{\beta_*} = \infty$ *provided* $\beta_0 \neq \pm\omega_1/\omega_2$ *and* $\beta_* \neq \pm\omega_1/\omega_2$.

(c) *Suppose the parameters are as in part* (a) *except* $(\rho_{uv}, \beta_0) \to (1, \pm\infty)$ *or* $(-1, \pm\infty)$. *Then* (i) $\lim_{(\rho_{uv}, \beta_0) \to (1, \pm\infty)} \lambda^{1/2} c_{\beta_*} = \lim_{(\rho_{uv}, \beta_0) \to (-1, \pm\infty)} \lambda^{1/2} c_{\beta_*} = \pm\lambda^{1/2}/\sigma_v$ *and* (ii) $\lim_{(\rho_{uv}, \beta_0) \to (1, \pm\infty)} \lambda d^2_{\beta_*} = \lim_{(\rho_{uv}, \beta_0) \to (-1, \pm\infty)} \lambda d^2_{\beta_*} = \infty$.

(d) *Suppose the parameters are as in part* (b) *except* $(\rho_\Omega, \beta_0) \to (1, \pm\infty)$ *or* $(-1, \pm\infty)$. *Then* (i) $\lim_{(\rho_\Omega, \beta_0) \to (1, \pm\infty)} \lambda^{1/2} c_{\beta_*} = \lim_{(\rho_\Omega, \beta_0) \to (-1, \pm\infty)} \lambda^{1/2} c_{\beta_*} = \mp\lambda^{1/2}/\omega_2$ *and* (ii) $\lim_{(\rho_\Omega, \beta_0) \to (1, \pm\infty)} \lambda d^2_{\beta_*} = \infty$ *provided* $\beta_* \neq \omega_1/\omega_2$ *and* $\lim_{(\rho_\Omega, \beta_0) \to (-1, \pm\infty)} \lambda d^2_{\beta_*} = \infty$ *provided* $\beta_* \neq -\omega_1/\omega_2$.

Comments. (i). Combining Theorems 8.1 and 8.2 provides analytic finite-sample limiting optimality results for the CLR and LM tests and CSs as $\rho_{uv} \to \pm 1$ or $\rho_\Omega \to \pm 1$ with β_0 fixed or jointly with $\beta_0 \to \pm\infty$ for (almost) any fixed values of the other parameters. These results apply for any strength of the IVs except $\lambda = 0$. These results are much stronger than typical weighted average power (WAP) results because they hold for (almost) any fixed values of the parameters β_0, β_*, σ_1, σ_v, and $\lambda > 0$ when $\rho_{uv} \to \pm 1$ and (almost) any fixed values of the parameters β_0, β_*, ω_1, ω_2, and $\lambda > 0$ when $\rho_\Omega \to \pm 1$.

(ii). The cases $\rho_{uv} \to \pm 1$ and $\rho_\Omega \to \pm 1$ are closely related because $(1 - \rho^2_\Omega)^{1/2}\omega_1 = (1 - \rho^2_{uv})^{1/2}\sigma_u$ by (59) in the Online Supplementary Material 1. Thus, $\rho_{uv} \to \pm 1$ implies $|\rho_\Omega| \to 1$ or $\omega_1 \to 0$. And, $\rho_\Omega \to \pm 1$ implies $|\rho_{uv}| \to 1$ or $\sigma_u \to 0$.

(iii). The asymptotic results of Theorem 8.2 as $\rho_{uv} \to \pm 1$ or $\rho_\Omega \to \pm 1$ are empirically relevant because they reflect the behavior of the CLR test even when $|\rho_{uv}|$ or $|\rho_\Omega|$ is not very close to one. See the results in Table 1 when ρ_{uv} $(= \rho_\Omega)$ equals 0.7 and 0.9. The results of Theorem 8.2 indicate that it would be informative for empirical papers to report estimates of ρ_Ω (which is consistently estimable even under weak IVs).

9. General power/false-coverage-probability comparisons

Section 7 shows that the probability that the CLR CS has infinite length is higher than the lower bound for this probability in some parameter scenarios. By Theorem 5.1, this implies that the power differences between the CLR and POIS2 tests are not always close

TABLE 2. Maximum and Average Power Differences over λ and β_0 Values between POIS2 and CLR Tests for Fixed Alternative $\beta^* = 0$.

(a) Across k patterns for fixed ρ_{uv}

ρ_{uv}	k	$\lambda_{\max}$	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	POIS2–CLR Max	POIS2–CLR Average
0.0	2	7	−10,000.00	1.00	0.66	0.021	0.006
0.0	5	10	−50.00	1.00	0.68	0.030	0.009
0.0	10	15	−50.00	1.00	0.76	0.038	0.012
0.0	20	15	10.00	−1.00	0.60	0.042	0.014
0.0	40	22	−50.00	1.00	0.66	**0.059**	**0.016**
0.3	2	10	3.75	−0.96	0.86	0.019	0.005
0.3	5	10	3.50	−0.96	0.73	0.034	0.008
0.3	10	10	3.00	−0.94	0.59	0.032	0.009
0.3	20	15	3.50	−0.96	0.66	0.045	0.012
0.3	40	22	4.00	−0.97	0.72	**0.061**	**0.014**
0.5	2	5	2.00	−0.87	0.64	0.016	0.004
0.5	5	10	2.25	−0.90	0.82	0.029	0.005
0.5	10	10	2.00	−0.87	0.70	0.037	0.007
0.5	20	10	1.75	−0.82	0.53	0.046	0.009
0.5	40	15	1.75	−0.82	0.59	**0.050**	**0.012**
0.7	2	5	1.50	−0.75	0.81	0.016	0.002
0.7	5	5	1.50	−0.75	0.67	0.033	0.003
0.7	10	7	1.50	−0.75	0.71	0.036	0.005
0.7	20	7	1.25	−0.61	0.54	0.042	0.006
0.7	40	15	1.50	−0.75	0.84	**0.050**	**0.008**
0.9	2	0.9	1.25	−0.63	0.46	0.017	0.002
0.9	5	0.9	1.00	−0.22	0.33	0.017	0.002
0.9	10	3	1.25	−0.63	0.77	0.027	0.003
0.9	20	3	1.00	−0.22	0.61	0.032	0.003
0.9	40	5	1.25	−0.63	0.75	**0.040**	**0.004**

(*Continues*)

to zero as the null value $\beta_0 \to \pm\infty$ for a fixed true value $\beta_* = 0$. In this section, we investigate these power differences for a fixed true value $\beta_* = 0$ and a wide range of null hypothesis values β_0, not just $\beta_0 \to \pm\infty$. The results show that in some parameter scenarios these power differences are not close to zero for finite β_0 and can be larger than the power differences as $\beta_0 \to \pm\infty$.

Without loss of generality, we take $\sigma_v^2 = \omega_{22} = 1$ and $\omega_{11} = 1$ in this section. Table 2 reports maximum and average power differences over $\beta_0 \in R$ and $\lambda > 0$ for a fixed true value $\beta_* = 0$ for a range of values of (ρ_{uv}, k). As above, the choice of $\beta_* = 0$ is without loss of generality. These power differences are equivalent to false coverage probability differences between the CLR and POIS2 CSs for a fixed true value β_* at incorrect values β_0. They are necessarily nonnegative.

The λ values considered are 1, 3, 5, 7, 10, 15, 20, as well as 22, 25 when $k = 20$ and 40, and as well as 0.7, 0.8, 0.9 when $k = 2$ and 5 and $\rho_{uv} = 0.9$. The positive and negative

Table 2. *Continued.*

(b) Across ρ_{uv} patterns for fixed k

k	ρ_{uv}	$\lambda_{\max}$	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	POIS2–CLR	
						Max	Average
2	0.0	7	−10,000.00	1.00	0.66	**0.021**	**0.006**
2	0.3	10	3.75	−0.96	0.86	0.019	0.005
2	0.5	5	2.00	−0.87	0.64	0.016	0.004
2	0.7	5	1.50	−0.75	0.81	0.016	0.002
2	0.9	0.9	1.25	−0.63	0.46	0.017	0.002
5	0.0	10	−50.00	1.00	0.68	0.030	**0.009**
5	0.3	10	3.50	−0.96	0.73	**0.034**	0.008
5	0.5	10	2.25	−0.90	0.82	0.029	0.005
5	0.7	5	1.50	−0.75	0.67	0.033	0.003
5	0.9	0.9	1.00	−0.22	0.33	0.017	0.002
10	0.0	15	−50.00	1.00	0.76	**0.038**	**0.012**
10	0.3	10	3.00	−0.94	0.59	0.032	0.009
10	0.5	10	2.00	−0.87	0.70	0.037	0.007
10	0.7	7	1.50	−0.75	0.71	0.036	0.005
10	0.9	3	1.25	−0.63	0.77	0.027	0.003
20	0.0	15	10.00	−1.00	0.60	0.042	**0.014**
20	0.3	15	3.50	−0.96	0.66	0.045	0.012
20	0.5	10	1.75	−0.82	0.53	**0.046**	0.009
20	0.7	7	1.25	−0.61	0.54	0.042	0.006
20	0.9	3	1.00	−0.22	0.61	0.032	0.003
40	0.0	22	−50.00	1.00	0.66	0.059	**0.016**
40	0.3	22	4.00	−0.97	0.72	**0.061**	0.014
40	0.5	15	1.75	−0.82	0.59	0.050	0.012
40	0.7	15	1.50	−0.75	0.84	0.050	0.008
40	0.9	5	1.25	−0.63	0.75	0.040	0.004

β_0 values considered are those with $|\beta_0| \in \{0.25, 0.5, \ldots, 3.75, 4, 5, 7.5, 10, 50, 100, 1000, 10{,}000\}$. These (λ, β_0) values were chosen, based on preliminary simulations, to ensure that changes in the power differences in Table 2 (and Tables 3 and 4 below) across neighboring values (λ, β_0) are small.

The number of simulation repetitions employed is 5000. The critical values are determined using 100,000 simulation repetitions. For example, the simulation standard deviations for the power differences for $(\rho_{uv}, k) = (0, 20)$ and any fixed (β_0, λ) value range from $[0.0013, 0.0040]$ across different (β_0, λ) values, which compares to simulated average of the power differences over (β_0, λ) values that equals 0.014.

Tables 2(a) and 2(b) contain the same numbers, but are reported differently to make the patterns in the table more clear. Table 2(a) shows variation across k for fixed ρ_{uv}, whereas Table 2(b) shows variation across ρ_{uv} for fixed k. The third and fourth columns in each table report the values of λ and β_0 at which the maximum power difference is obtained. The fifth column in each table reports $\rho_{uv,0}$, which is the correlation between the structural-equation and reduced-form errors when β_0 is the true value (based on the assumption that the consistently-estimable reduced-form variance matrix is the same whether the truth is β_0 or β_*). In contrast, ρ_{uv} is the same correlation, but when β_* is the true value—which is the true β value in the power difference simulations. The sixth column in the tables reports the power of the CLR test at the (β_0, λ) values that maximize the power difference for given (ρ_{uv}, k), that is, at $(\beta_{0,\max}, \lambda_{\max})$.

Table 2 shows that the maximum (over (β_0, λ)) power differences between the POIS2 and CLR tests range between [0.016, 0.061] over the (ρ_{uv}, k) values. On the other hand, the average (over (β_0, λ)) power differences only range between [0.002, 0.016] over the (ρ_{uv}, k) values. This indicates that, although there are some (β_0, λ) values at which the CLR test is noticeably off the two-sided AE power envelope, on average the CLR test's power is not far from the power envelope. The numbers in boldface in Table 2(a) give the largest over k maximum or average power difference for a given value of ρ_{uv}. The numbers in boldface in Table 2(a) give the largest over ρ_{uv} maximum or average power difference for a given value of k.

In contrast, the analogous maximum and average power difference ranges for the AR test are [0.079, 0.513] and [0.012, 0.179]; see Table SM-III in the Online Supplementary Material 2. For the LM test, they are [0.242, 0.784] and [0.010, 0.203]; see Table SM-IV in the Online Supplementary Material 2. Hence, the power of AR and LM tests is very much farther from the POIS2 power envelope than is the power of the CLR test.

Table 2(a) shows that the maximum and average (over (β_0, λ)) power differences for the CLR test are clearly increasing in k. Table 2(a) shows that for $\rho_{uv} \geq 0.3$, the power differences are maximized at more or less the same β_0 regardless of the value of k. For $\rho_{uv} = 0$, this is also true to a certain extent, because the sign of β_0 is irrelevant (when $\rho_{uv} = 0$) and the values 50 and 10,000 are both large values. Table 2(a) also shows that for each ρ_{uv}, the power differences are maximized at λ values that (weakly) increase with k. The increase is particularly evident going from $k = 20$ to 40.

Table 2(b) shows that for $k \geq 5$, the maximum power differences are more or less the same for $\rho_{uv} \leq 0.7$, but noticeably lower for $\rho_{uv} = 0.9$. For $k = 2$, the maximum power differences are more or less the same for all ρ_{uv} considered. Table 2(b) shows that, for each k, the power differences are maximized at $|\beta_0|$ values that become closer to 0 as ρ_{uv} increases. Table 2(b) also shows that, for each k, the power differences are maximized at λ values that become closer to 0 as ρ_{uv} increases.[9]

In sum, the maximum power differences over (β_0, λ) are found to increase in k ceteris paribus and decrease in ρ_{uv} ceteris paribus. The λ values at which the maxima are obtained are found to (weakly) increase with k ceteris paribus and decrease in ρ_{uv} ceteris paribus. The $|\beta_0|$ values at which the maxima are obtained are found to be independent of k ceteris paribus and decrease in ρ_{uv} ceteris paribus.

Next, Figure 1 provides a picture of how the power of the CLR, AR, and POIS2 tests differ as a function of β_0 when other parameters are held fixed. Results given are for

[9] See Table SM-II in the Online Supplementary Material 2 for how the maximum PD's over β_0 vary with λ for the (ρ_{uv}, k) values in Table 2.

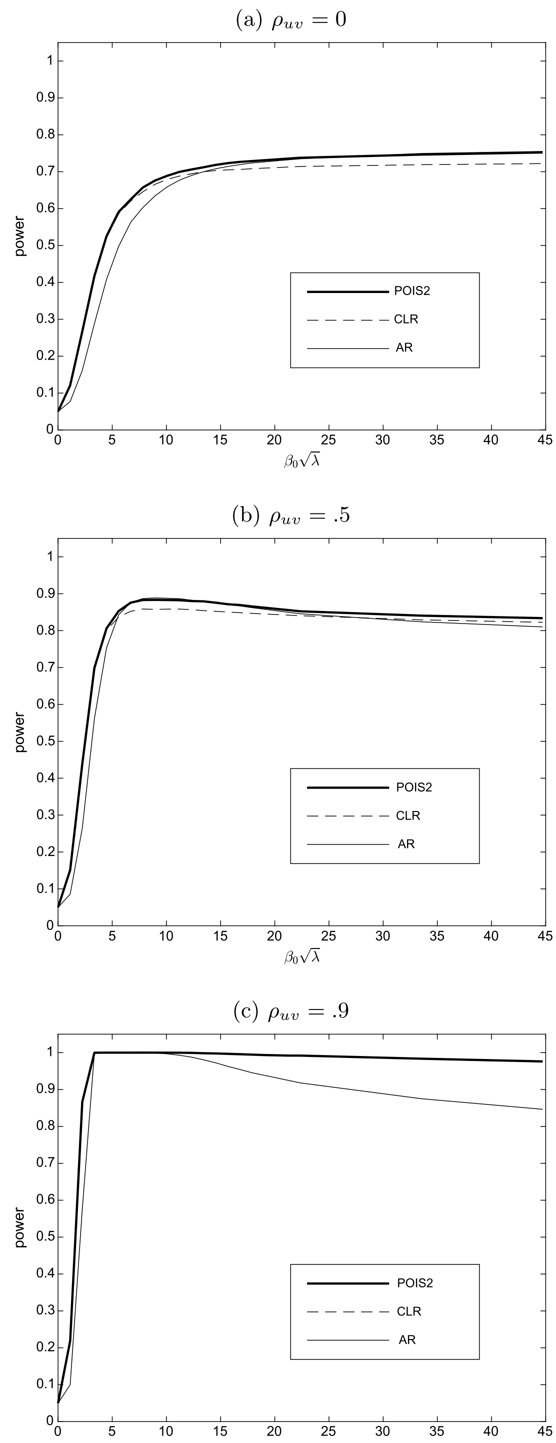

FIGURE 1. The power functions of the POIS2, CLR, and AR tests for $k = 10$, $\lambda = 15$, and $\rho_{uv} = 0, 0.5, 0.9$.

three parameter configurations $\rho_{uv} = 0, 0.5, 0.9$ with $\lambda = 15$, $k = 10$, and $\beta_* = 0$ in all three configurations. These parameter configurations are chosen because they are ones in which the power of the CLR test is noticeably off the power envelope for sufficiently large β_0 when $\rho_{uv} = 0$ and 0.5.

Figure 1(a) for $\rho_{uv} = 0$ shows: (i) the power of all three tests does not go to one as $\beta_0 \to \infty$ (the limit value depends on the magnitude of λ, which is 15 in Figure 1), (ii) the CLR test is off the power envelope and the AR test is on the power envelope (up to the numerical accuracy) for large $\beta_0 - \beta_*$, and (iii) the reverse is true for smaller $\beta_0 - \beta_*$.

Figure 1(b) for $\rho_{uv} = 0.5$ shows: (i) the power of all three tests does not go to one as $\beta_0 \to \infty$, (ii) the CLR test is off the power envelope for large $\beta_0 - \beta_*$ and on the power envelope (up to the numerical accuracy) for smaller $\beta_0 - \beta_*$, and (iii) the AR test is on the power envelope (up to the numerical accuracy) for intermediate values of $\beta_0 - \beta_*$ and off the power envelope for larger and smaller $\beta_0 - \beta_*$.

Figure 1(c) for $\rho_{uv} = 0.9$ shows: (i) the power of all three tests does not go to one as $\beta_0 \to \infty$, but the powers of the CLR and POIS2 tests are quite close to one for β_0 large, (ii) the CLR test is on the POIS2 power envelope (up to the numerical accuracy) for all β_0 values, and (iii) the AR test is off the POIS2 power envelope for most of the β_0 values considered, including small and large β_0 values.

In all of the simulations considered (across the parameters scenarios considered in Table 2), the CLR test was found to be on the POIS2 power envelope (up to the numerical accuracy) for small values of $\beta_0 - \beta_*$.

The numerical results in this section show that the finding of AMS that the CLR test is essentially on the two-sided AE power envelope does not hold when one considers a broader range of null and alternative hypothesis values (β_0, β_*) than those considered in the numerical results in AMS.

10. Differences between CLR power and an average over λ power envelope

Results above show that the CLR test is not close to the two-sided power envelope for all β values in some scenarios. In contrast, in this section, we show numerically the weaker property that the CLR test is close to the *average* two-sided power envelope for a wide range of β values, where the average is taken over a grid of λ values which index the strength of identification.

We introduce a "WAP2" power envelope for similar tests with weight functions over: (i) a finite grid of λ values, $\{\lambda_j > 0 : j \leq J\}$, (ii) the same two-points (β_*, λ_j) and $(\beta_{2*}, \lambda_{2j})$ as in AMS for each λ_j for $j \leq J$, and (iii) the same uniform weight function over $\mu_\pi / \|\mu_\pi\|$ as in Chernozhukov, Hansen, and Jansson (2009). In particular, we use the uniform weight function over the 36 values of λ in $\{2.5, 5.0, \ldots, 90.0\}$. This grid of λ values includes a comprehensive range of empirically relevant λ values that includes the λ values in Table 2 in which the power of the CLR test is noticeably off the power envelope.

The WAP2 envelope is a function of (β_0, β_*). The WAP2(Q, β_0, β_*) test statistic that generates this envelope is of the form $\sum_{j=1}^{J} (\psi(Q; \beta_0, \beta_{*j}, \lambda_j) + \psi(Q; \beta_0, \beta_{2*j}, \lambda_{2j}))/ \sum_{j=1}^{J} 2\psi_2(Q_T; \beta_0, \beta_*, \lambda_j)$, where the functions $\psi(Q; \beta_0, \beta, \lambda)$ and $\psi_2(Q_T; \beta_0, \beta, \lambda)$ are as in AMS (and as in (28) in Supplementary Material 1). The WAP2(Q, β_0, β_*) conditional critical value $\kappa_{2,\beta_0,J}(q_T)$ is defined to satisfy $P_{Q_1|Q_T}(\text{WAP2}(Q, \beta_0, \beta_*) > \kappa_{2,\beta_0,J}(q_T)|$

$q_T) = \alpha$ for all $q_T \geq 0$, where $P_{Q_1|Q_T}(\cdot|q_T)$ denotes probability under the density $f_{Q_1|Q_T}(\cdot|q_T)$, which is specified in (26) in the Online Supplementary Material 1.

To be consistent with Tables 1 and 2, we report power differences between the WAP2(Q, β_0, β_*) and CLR tests for $\beta_* = 0$ and a range of β_0 values. These power differences are equivalent to the false coverage probability differences between the CLR and WAP2 CSs for fixed true β_* and varying incorrect β_0 values. The differences are necessarily nonnegative.

We consider $\rho_{uv} \in \{0, 0.3, 0.5, 0.7, 0.9, 0.95, 0.99\}$, $k \in \{2, 5, 10, 20, 40\}$, the same β_0 values as in Table 2, and $\omega_1^2 = \omega_2^2 = 1$. (The large ρ_{uv} values of 0.95 and 0.99 are included to show that the results are not sensitive to ρ_{uv} being close to one.) Since $\beta_* = 0$, $\rho_\Omega = \rho_{uv}$. Section 22 in the Online Supplementary Material 1 shows that taking $\beta_* = 0$ and $\omega_1^2 = \omega_2^2 = 1$ is without loss of generality provided the support of the weight function for λ is scaled by ω_2^2 when $\omega_2 \neq 1$. The number of simulation repetitions employed is 1000 for each λ_j value. With power averaged over the 36 λ_j values and independence of the simulation draws across λ_j, this yields simulation SDs that are comparable to using 36,000 simulation repetitions. The critical values are determined using 100,000 simulation repetitions for $k = 5$ and 10,000 for other values of k.

For brevity, Table 3 reports results only for $k = 5$ for a subset of the β_0 values considered. Results for all values of k and β_0 considered are given in Table SM-V in the Online

TABLE 3. Average (over λ) Power Differences for $\lambda \in \{2.5, 5.0, \dots, 90.0\}$ between the WAP2 and CLR Tests for $k = 5$.

	$\rho_{uv,0}$		WAP2–CLR						
β_0	$\rho_{uv} = 0$	0.9	$\rho_{uv} = 0$	0.3	0.5	0.7	0.9	0.95	0.99
−10,000.00	1.00	1.00	0.005	0.002	0.001	0.001	0.000	−0.000	0.000
−100.00	1.00	1.00	**0.005**	0.002	0.001	0.001	0.000	−0.001	−0.000
−10.00	1.00	1.00	0.005	0.002	0.001	0.000	0.000	−0.000	−0.000
−4.00	0.97	1.00	0.003	0.001	0.000	−0.000	0.000	0.000	−0.000
−3.00	0.95	0.99	0.003	0.001	0.000	0.000	−0.000	**0.001**	**0.000**
−2.00	0.89	0.99	0.002	0.001	0.000	0.001	−0.000	−0.001	−0.000
−1.50	0.83	0.98	0.001	0.001	0.001	0.000	0.000	−0.001	−0.000
−1.00	0.71	0.97	0.001	0.000	−0.000	−0.000	−0.000	0.000	−0.000
−0.75	0.60	0.97	0.000	−0.000	0.001	−0.000	−0.000	0.000	0.000
−0.50	0.45	0.95	−0.000	−0.000	−0.001	−0.001	−0.000	−0.001	−0.000
−0.25	0.24	0.94	−0.001	−0.001	−0.001	−0.000	−0.000	0.001	−0.001
0.25	−0.24	0.83	−0.000	−0.001	−0.001	−0.000	−0.001	0.000	0.000
0.50	−0.45	0.68	0.001	0.000	0.000	0.000	0.000	−0.001	0.000
0.75	−0.60	0.33	0.000	0.001	0.001	0.001	0.000	0.000	0.000
1.00	−0.71	−0.22	0.002	0.001	0.001	0.001	0.000	0.000	0.000
1.50	−0.83	−0.81	0.001	0.002	0.003	**0.003**	**0.001**	−0.000	0.000
2.00	−0.89	−0.93	0.002	0.003	**0.004**	0.002	0.000	−0.001	−0.000
3.00	−0.95	−0.98	0.003	**0.005**	0.003	0.001	0.000	0.000	0.000
4.00	−0.97	−0.99	0.004	0.005	0.002	0.001	0.000	0.001	0.000
10.00	−1.00	−1.00	0.005	0.003	0.001	0.001	0.000	0.000	0.000
100.00	−1.00	−1.00	0.005	0.003	0.001	0.000	0.000	−0.001	0.000
10,000.00	−1.00	−1.00	0.005	0.002	0.001	0.001	0.000	−0.000	0.000

TABLE 4. Average (over λ) Power Differences between the WAP2 and CLR Tests.

	(a) Maxima over β_0							(b) Averages over β_0						
k	$\rho_{uv}=0$	0.3	0.5	0.7	0.9	0.95	0.99	$\rho_{uv}=0$	0.3	0.5	0.7	0.9	0.95	0.99
2	0.004	0.003	0.002	0.002	0.001	0.001	0.001	0.002	0.002	0.001	0.001	0.000	0.000	0.000
5	0.005	0.005	0.004	0.003	0.001	0.001	0.000	0.003	0.002	0.001	0.001	0.000	0.000	0.000
10	0.011	0.010	0.008	0.005	0.004	0.003	0.003	0.007	0.006	0.004	0.002	0.001	0.001	0.001
20	0.013	0.012	0.010	0.007	0.002	0.001	0.002	0.008	0.007	0.005	0.002	0.000	0.000	0.000
40	0.024	0.021	0.017	0.011	0.004	0.001	0.000	0.013	0.011	0.007	0.004	0.000	0.000	0.000

Supplementary Material 2. Table 4 reports summary results for all values of k. In particular, Table 4(a) provides the maxima over β_0 of the average over λ power differences for each (ρ_{uv}, k). Table 4(b) provides the average over β_0 of the average over λ power differences for each (ρ_{uv}, k).

The boldface numbers in Table 3 are the largest values in each column. Table 3 shows that the CLR test has power quite close to the WAP2 power envelope for $k=5$. The power differences for $\rho_{uv} \in \{0, 0.3, 0.5, 0.7\}$ are in $[0.000, 0.005]$ (with simulation standard deviations in $[0.0003, 0.0007]$) across all β_0 values. For $\rho_{uv} \in \{0.9, 0.95, 0.99\}$, the power differences are in $[0.000, 0.001]$ (with simulation standard deviations in $[0.0000, 0.0003]$) across all β_0 values.

Table 4 shows that power differences between the WAP2 power envelope and the CLR power are increasing in k and decreasing in $|\rho_{uv}|$. For $k=2$, the maximum power difference over β_0 and ρ_{uv} values is very small: 0.004. In the worst case for CLR, which is when $(k, \rho_{uv}) = (40, 0)$, the maximum power difference over β_0 values is substantially larger: 0.024. The average (over β_0 values) power difference in this case is 0.013, which is not very large. For $k=40$ and $\rho_{uv} \geq 0.9$, the maximum power difference (over β_0 and ρ_{uv} values) is very small: 0.004. This is consistent with the theoretical optimality properties of the CLR test as $\rho_{uv} \to \pm 1$ described in Section 8. For $k=40$ and $\rho_{uv} \geq 0.9$, the average power difference (over β_0 values and the five ρ_{uv} values) is very small: 0.000. The second worst case for CLR in Table 4 is when $(k, \rho_{uv}) = (20, 0)$. In this case, the maximum power difference over β_0 values is 0.013, which is noticeably lower than 0.024 for $(k, \rho_{uv}) = (40, 0)$.

In conclusion, the results in Tables 3 and 4 show that the CLR test is very close to the WAP2 power envelope for most (k, ρ_{uv}, β_0) values, but can deviate from it by as much as 0.024 for some β_0 values when $(k, \rho_{uv}) = (40, 0)$.

REFERENCES

Anderson, T. W. (1946), "The non-central Wishart distribution and certain problems of multivariate statistics." *Annals of Mathematical Statistics*, 17, 409–431. [468]

Anderson, T. W. and H. Rubin (1949), "Estimation of the parameters of a single equation in a complete system of stochastic equations." *Annals of Mathematical Statistics*, 20, 46–63. [458]

Andrews, D. W., V. Marmer, and Z. Yu (2019), "Supplement to 'On optimal inference in the linear IV model'." *Quantitative Economics Supplemental Material*, 10, https://doi.org/10.3982/QE1082. [458, 459]

Andrews, D. W. K. (1998), "Hypothesis testing with a restricted parameter space." *Journal of Econometrics*, 84, 155–199. [460]

Andrews, D. W. K. and P. Guggenberger (2018), "Identification- and singularity-robust inference for moment condition models." Cowles Foundation Discussion Paper No. 1978R, Yale University. [461]

Andrews, D. W. K. and P. Guggenberger (2017), "Asymptotic size of Kleibergen's LM and conditional LR tests for moment condition models." *Econometric Theory*, 33, 1046–1080. [467]

Andrews, D. W. K., M. J. Moreira, and J. H. Stock (2004), "Optimal invariant similar tests for instrumental variables regression with weak instruments." Cowles Foundation Discussion Paper No. 1476, Yale University. [461]

Andrews, D. W. K., M. J. Moreira, and J. H. Stock (2006), "Optimal two-sided invariant similar tests for instrumental variables regression." *Econometrica*, 74, 715–752. [457, 458, 461]

Andrews, D. W. K., M. J. Moreira, and J. H. Stock (2008), "Efficient two-sided nonsimilar invariant tests in IV regression with weak instruments." *Journal of Econometrics*, 146, 241–254. [458]

Andrews, D. W. K. and W. Ploberger (1994), "Optimal tests when a nuisance parameter is present only under the alternative." *Econometrica*, 62, 1383–1414. [460]

Andrews, I. (2016), "Conditional linear combination tests for weakly identified models." *Econometrica*, 84, 2155–2182. [461]

Andrews, I. and A. Mikusheva (2016), "Conditional inference with a functional nuisance parameter." *Econometrica*, 84, 1571–1612. [461]

Chamberlain, G. (2007), "Decision theory applied to an instrumental variables model." *Econometrica*, 75, 609–652. [458]

Chernozhukov, V., C. Hansen, and M. Jansson (2009), "Admissible invariant similar tests for instrumental variables regression." *Econometric Theory*, 25, 806–818. [458, 460, 461, 468, 470, 474, 475, 481]

Davidson, R. and J. G. MacKinnon (2008), "Bootstrap inference in a linear equation estimated by instrumental variables." *Econometrics Journal*, 11, 443–477. [465]

Dufour, J.-M. (1997), "Some impossibility theorems in econometrics with applications to structural and dynamic models." *Econometrica*, 65, 1365–1387. [458, 467]

Dufour, J.-M. and M. Taamouti (2005), "Projection-based statistical inference in linear structural models with possibly weak instruments." *Econometrica*, 73, 1351–1365. [472]

Elliott, G., U. K. Müller, and M. W. Watson (2015), "Nearly optimal tests when a nuisance parameter is present under the null hypothesis." *Econometrica*, 83, 771–811. [460]

Hillier, G. H. (2009), "Exact properties of the conditional likelihood ratio test in an IV regression model." *Econometric Theory*, 25, 915–957. [458, 461]

Kleibergen, F. (2002), "Pivotal statistics for testing structural parameters in instrumental variables regression." *Econometrica*, 70, 1781–1803. [458, 463]

Mikusheva, A. (2010), "Robust confidence sets in the presence of weak instruments." *Journal of Econometrics*, 157, 236–247. [458, 461, 467, 472]

Moreira, H. and M. J. Moreira (2013), "Contributions to the theory of optimal tests." Unpublished manuscript, available at Ensaios Economicos, http://hdl.handle.net/10438/11263. [460]

Moreira, H. and M. J. Moreira (2015), "Optimal two-sided tests for instrumental variables regression with heteroskedastic and autocorrelated errors." Unpublished manuscript, available at arXiv:1505.06644. [460, 461]

Moreira, M. J. (2003), "A conditional likelihood ratio test for structural models." *Econometrica*, 71, 1027–1048. [457, 458, 462, 463]

Moreira, M. J. (2009), "Tests with correct size when instruments can be arbitrarily weak." *Journal of Econometrics*, 152, 131–140. [458, 463, 464]

Staiger, D. and J. H. Stock (1997), "Instrumental variables regression with weak instruments." *Econometrica*, 65, 557–586. [458]

Wald, A. (1943), "Tests of statistical hypotheses concerning several parameters when the number of observations is large." *Transactions of the American Mathematical Society*, 54, 426–482. [460, 461]

Co-editor Andres Santos handled this manuscript.

Manuscript received 28 February, 2018; final version accepted 20 August, 2018; available online 2 October, 2018.