A NOTE ON OPTIMAL INFERENCE IN THE LINEAR IV MODEL

By

Donald W. K. Andrews, Vadim Marmer, and Zhengfei Yu

January 2017

COWLES FOUNDATION DISCUSSION PAPER NO. 2073

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

A Note 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

October 23, 2015
Revised: January 30, 2017*

*Andrews gratefully acknowledges the research support of the National Science Foundation via grant number SES-1355504. Marmer gratefully acknowledges the research support of the SSHRC via grant numbers 410-2010-1394 and 435-2013-0331.

Abstract

This paper considers tests and confidence sets (CS's) 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" 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 (non-zero) 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, the paper shows via theory and numerical work that the CLR test is close to having greatest average power, where the average is over a grid of concentration parameter values and over pairs alternative hypothesis values of the parameter of interest, uniformly over pairs of alternative hypothesis values and uniformly over the correlation between the structural and reduced-form errors.

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 CS's have very good overall properties.

Keywords: Conditional likelihood ratio test, confidence interval, infinite length, linear instrumental variables, optimal test, weighted average power, similar test.

JEL Classification Numbers: 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 (CS's) can perform quite poorly under weak IV's (see Dufour (1997), Staiger and Stock (1997), and references therein), inference procedures that are robust to weak IV's have been developed, e.g., see Kleibergen (2002) and Moreira (2003, 2009). The focus has been on models with one right-hand side (rhs) endogenous variable, because this arises most frequently in applications, and on over-identified models, because Anderson and Rubin (1949) (AR) tests and CS's are robust to weak IV's 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 rhs 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) (CHJ) show that this power envelope also applies to non-invariant tests provided the envelope is for power averaged over certain direction vectors in a unit sphere. CHJ also shows that the invariant similar tests that generate the two-sided AE power envelope are α-admissible and d-admissible. Mikusheva (2010) provides approximate optimality results for CLR-based CS's that utilize the testing results in AMS. Chamberlain (2007), Andrews, Moreira, and Stock (2008), and Hillier (2009) provide 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 CS's 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 rhs endogenous variable, k IV's, a concentration parameter λ_v (which is a measure of the strength of the IV's), 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 $(2, 7)$, $(5, 12)$, $(10, 15)$, $(20, 20)$, and $(40, 20)$, the differences between the probabilities that the 95% CLR and AR CS's have infinite length are .016, .029, .037, .044, and .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

[1] See Table SM-I in the Supplemental Material.

outperforms the CLR CS in terms of its infinite-length behavior, which is an important property for CS's. Similarly, one obtains positive (but smaller) differences also when $\rho_{uv} = .3$ for the same range of (k, λ_v) values. On the other hand, for $\rho_{uv} = .5, .7$, and $.9$, the differences are negative over the same range of (k, λ_v) values.

The AR and CLR CS's 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 $\rho_{uv} = .0$ and $.3$ 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 CS's have infinite length.

Second, the paper uses the first result to determine a finite-sample lower bound function on the probabilities that a CS has infinite length for CS's based on invariant similar tests. 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 CS's have infinite length with the lower bound.

Third, simulation results show that the AR and CLR CS's 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 IV's, but not irrelevant IV's. 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 CS's for several (k, λ_v) values. On the other hand, for values of $|\rho_{uv}| \geq .7$, the CLR CS has probabilities of having infinite length that are close to the lower bound function, .010 or less and typically much less, for all (k, λ_v) combinations considered. For values of $|\rho_{uv}| \geq .7$, the AR CS has probabilities of having infinite length that are often far from the lower bound. For $|\rho_{uv}| = .9$ and certain values of λ_v, they are as large as .084, .196, .280, .353, and .422 for $k = 2$, 5, 10, 20, and 40, respectively.[2]

Fourth, the paper derives new optimality properties of the CLR test when $\rho_{uv} \to \pm 1$ or $\rho_\Omega \to \pm 1$ with other parameters fixed at any values (with non-zero concentration parameter), where ρ_Ω

[2] See Table SM-I in the SM.

denotes the correlation between the reduced-form errors. 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 consistent with the numerical results that show that the CLR test is very close to the power envelope when $|\rho_{uv}|$ is large.

Fifth, we simulate power differences (PD's) 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 PD's as $\beta_0 \to \pm\infty$ discussed above). These PD's 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) PD's range between $[.016, .061]$ over the (ρ_{uv}, k) values considered. On the other hand, the average (over β_0 and λ values) PD's only range between $[.002, .016]$. 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 PD's 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 $[.001, .007]$ 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), 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 CHJ 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 CHJ 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. More specifically, we recommend using heteroskedasticity-robust versions of these procedures that have the same asymptotic properties as these procedures under homoskedasticity. For example, such tests are given in Andrews, Moreira, and Stock (2004) and Andrews and Guggenberger (2015). 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 $|\rho_{uv}|$.

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 CS's 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 CS's, 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 provides a formula for the probability that a CS has infinite length. Section 5 derives a lower bound on the probability that a CS constructed using two-sided invariant similar tests has infinite length. Section 6 reports differences between the probability that the CLR CS has infinite length and the lower bound derived in the previous section. Section 7 proves the optimality results for the CLR test described above. Section 8 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 9 provides comparisons of the power of the CLR test to the WAP2 power envelope described above. Proofs and additional theoretical and numerical results are given in the Supplemental Material (SM).

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 (wlog), without any exogenous variables. The model has one rhs endogenous variable, k instrumental variables (IV's), and normal errors with known reduced-form error variance matrix. The model consists of a structural equation and a reduced-form equation:

$$y_1 = y_2\beta + u \text{ and } y_2 = Z\pi + v_2, \tag{2.1}$$

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., non-stochastic) 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.

The two corresponding reduced-form equations are

$$\begin{aligned} Y &:= [y_1 : y_2] := [Z\pi\beta + v_1 : Z\pi + v_2] = Z\pi a' + V, \text{ where} \\ V &:= [v_1 : v_2] = [u + v_2\beta : v_2], \text{ and } a := (\beta, 1)'. \end{aligned} \tag{2.2}$$

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 IV's 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 CS's 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 $Z'Y$:

$$\begin{aligned} S &:= (Z'Z)^{-1/2} Z'Y b_0 \cdot (b_0'\Omega b_0)^{-1/2} \sim N(c_\beta(\beta_0, \Omega) \cdot \mu_\pi, I_k) \text{ and} \\ T &:= (Z'Z)^{-1/2} Z'Y \Omega^{-1} a_0 \cdot (a_0'\Omega^{-1} a_0)^{-1/2} \sim N(d_\beta(\beta_0, \Omega) \cdot \mu_\pi, I_k), \text{ where} \\ b_0 &:= (1, -\beta_0)', \ a_0 := (\beta_0, 1)', \ \mu_\pi := (Z'Z)^{1/2}\pi \in R^k, \\ c_\beta(\beta_0, \Omega) &:= (\beta - \beta_0) \cdot (b_0'\Omega b_0)^{-1/2} \in R, \\ d_\beta(\beta_0, \Omega) &:= b'\Omega b_0 \cdot (b_0'\Omega b_0)^{-1/2} \det(\Omega)^{-1/2} \in R, \text{ and } b = (1, -\beta)'. \end{aligned} \tag{2.3}$$

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 Hillier (1984) and AMS, we consider tests that are invariant to orthonormal transformations of $[S : T]$, i.e., $[S : T] \to [FS : FT]$ for a $k \times k$ orthogonal matrix F. The 2×2 matrix Q is a maximal invariant, where

$$Q = [S{:}T]'[S{:}T] = \begin{bmatrix} S'S & S'T \\ S'T & T'T \end{bmatrix} = \begin{bmatrix} Q_S & Q_{ST} \\ Q_{ST} & Q_T \end{bmatrix} \text{ and } Q_1 = \begin{pmatrix} S'S \\ S'T \end{pmatrix} = \begin{pmatrix} Q_S \\ Q_{ST} \end{pmatrix}, \tag{3.1}$$

e.g., 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 non-central 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

$$\lambda := \pi' Z' Z \pi \geq 0. \tag{3.2}$$

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

$$\begin{aligned} LR &:= \frac{1}{2}\left(Q_S - Q_T + \sqrt{(Q_S - Q_T)^2 + 4Q_{ST}^2}\right), \\ LM &:= Q_{ST}^2/Q_T \ = (S'T)^2/T'T, \text{ and } AR := Q_S/k \ = S'S/k. \end{aligned} \tag{3.3}$$

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 IV's 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

$$LR > \kappa_{LR,\alpha}(Q_T), \tag{3.4}$$

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 given Q_T is specified in AMS and in (11.3) in the SM.

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 Probability That a Confidence Set Has Infinite Length

In this section, we show that the probability that a confidence set (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

$$Q = Q_{\beta_0}(Y) = [S_{\beta_0}(Y) : T_{\beta_0}(Y)]'[S_{\beta_0}(Y) : T_{\beta_0}(Y)]. \tag{4.1}$$

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

$$\phi(Q_{\beta_0}(Y)) = 1(\mathcal{T}(Q_{\beta_0}(Y)) > cv(Q_{T\beta_0}(Y))) \tag{4.2}$$

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 $cv(Q_{T\beta_0}(Y))$ is a (possibly data-dependent) critical value. Examples include the AR, LM, and CLR tests in (3.3). Let CS_ϕ be the level $1-\alpha$ CS corresponding to ϕ. That is,

$$CS_\phi(Y) = \{\beta_0 : \phi(Q_{\beta_0}(Y)) = 0\}. \tag{4.3}$$

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

$$\exists K(Y) < \infty \text{ such that } \beta \in CS_\phi(Y)\ \forall \beta \geq K(Y) \text{ (or } \forall \beta \leq -K(Y)). \tag{4.4}$$

We say $CS_\phi(Y)$ has infinite length, which we denote by $Length(CS_\phi(Y)) = \infty$, 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 $-\infty < K_1(Y) \leq K_2(Y) < \infty$.

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 (2.3) with $\beta = \beta_*$ and $\lambda = \mu_\pi'\mu_\pi$. In this case, Q has a noncentral Wishart distribution whose density is given in (11.2) in the SM.

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, i.e., the correlation corresponding to Σ_*. Some calculations show that

$$
\begin{aligned}
\rho_{uv} &= \frac{\omega_{12} - \omega_2^2\beta_*}{(\omega_1^2 - 2\omega_{12}\beta_* + \omega_2^2\beta_*^2)^{1/2}\omega_2} \text{ and} \\
\Sigma_* &= \begin{bmatrix} \sigma_u^2 & \sigma_u\sigma_v\rho_{uv} \\ \sigma_u\sigma_v\rho_{uv} & \sigma_v^2 \end{bmatrix} = \begin{bmatrix} \omega_1^2 - 2\omega_{12}\beta_* + \omega_2^2\beta_*^2 & \omega_{12} - \omega_2^2\beta_* \\ \omega_{12} - \omega_2^2\beta_* & \omega_2^2 \end{bmatrix},
\end{aligned} \tag{4.5}
$$

where ω_1^2, ω_2^2, and ω_{12} denote the $(1,1)$, $(2,2)$, and $(1,2)$ elements of Ω, respectively, see (11.9) in the SM.

It is shown in Lemma 15.1 in the SM that the limit as $\beta_0 \to \pm\infty$ of $Q_{\beta_0}(Y)$ is

$$
Q_{\pm\infty}(Y) := \begin{bmatrix} e_2'Y'P_ZYe_2 \cdot \frac{1}{\sigma_v^2} & e_2'Y'P_ZY\Omega^{-1}e_1 \cdot \frac{\mp(1-\rho_{uv}^2)^{1/2}\sigma_u}{\sigma_v} \\ e_2'Y'P_ZY\Omega^{-1}e_1 \cdot \frac{\mp(1-\rho_{uv}^2)^{1/2}\sigma_u}{\sigma_v} & e_1'\Omega^{-1}Y'P_ZY\Omega^{-1}e_1 \cdot (1-\rho_{uv}^2)\sigma_u^2 \end{bmatrix}, \tag{4.6}
$$

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 15.1 in the SM that $Q_{\pm\infty}(Y)$ has a noncentral Wishart distribution with means matrix $\mp\mu_\pi(1/\sigma_v, \rho_{uv}/(\sigma_v(1-\rho_{uv}^2)^{1/2})) \in R^{k\times 2}$ and identity variance matrix.[3]

Theorem 4.1 *Suppose $CS_\phi(Y)$ 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 $cv(q_T)$, respectively, are continuous at all positive definite 2×2 matrices q and positive constants q_T, $P_{\beta_*,\pi,\Omega}(\mathcal{T}(Q_c(Y)) = cv(Q_{T,c}(Y))) = 0$ for $c = +\infty$ in parts* (a) *and* (c) *below and $c = -\infty$ in part* (b) *below. Then, for all $(\beta_*, \lambda, \Omega)$,*

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

[3] The density of this distribution is given in (11.4) in the SM.

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

(c) *if the tests are sign invariant, i.e.,* $\mathcal{T}(Q)$ *depends on* Q_{ST} *only through* $|Q_{ST}|$, *then* $P_{\beta_*,\pi,\Omega}(Length(CS_\phi(Y)) = \infty) = 1 - \lim_{\beta_0 \to \pm\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1)$.

Comments. (i). For the AR, LM, and LR tests, the continuity conditions on $\mathcal{T}(q)$ and $cv(q_T)$ hold given their simple functional forms in (3.3) 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 Thm. 10.1 in Andrews and Guggenberger (2016). We have $P_{\beta_*,\pi,\Omega}(\mathcal{T}(Q_{\pm\infty}(Y)) = cv(Q_{T,\pm\infty}(Y))) = 0$ for the AR and LM tests because $cv(Q_{T,\pm\infty}(Y))$ is a constant and $\mathcal{T}(Q_{\pm\infty}(Y))$ is absolutely continuous with respect to Lebesgue measure. For the CLR test, $P_{\beta_*,\pi,\Omega}(\mathcal{T}(Q_{\pm\infty}(Y)) = cv(Q_{T,\pm\infty}(Y))) = 0$ by the argument given in the proof of Theorem 5.4 in the SM. The AR, LM, and CLR test statistics are sign invariant. Hence, parts (a)-(c) of Theorem 4.1 apply to these tests. Theorem 5.4(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 4.1, for the AR, LM, and CLR tests.

(ii). Comment (iii) to Theorem 5.2 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 4.1(c) yields a lower bound on the probability that a CS $CS_\phi(Y)$ based on such tests has $Length = \infty$. The lower bound on the probability that $Length = \infty$ is greater than the lower bound on the probability that $RLength = \infty$ (or that $LLength = \infty$) unless $\rho_{uv} = 0$ (in which case it turns out that they are equal).

Theorem 12.1 in the SM 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 4.1(a) and (b) yields lower bounds on the probabilities that a CS $CS_\phi(Y)$ has $RLength = \infty$ based on $\beta_0 \to \infty$ and $LLength = \infty$ based on $\beta_0 \to -\infty$.

(iii). The results of Theorem 4.1(a) and (b) also hold for a $CS_\phi(Y)$ 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 $cv(t)$, respectively, are continuous at all $k \times 2$ matrices $[s : t]$ and k vectors t and satisfy $P_{\beta_*,\pi,\Omega}(\mathcal{T}(S_c(Y), T_c(Y)) = cv(T_c(Y))) = 0$ 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, $P_{\beta_*,\pi,\Omega}(RLength(CS_\phi(Y)) = \infty) = 1 - \lim_{\beta_0 \to \infty} P_{\beta_*,\beta_0,\pi,\Omega}(\phi([S : T]) = 1)$ 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 CS's 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 CS's. Due to this, Mikusheva (2010) compares CS's based on their expected truncated lengths for various truncation values. The result of Theorem 5.2 below implies that, for two CS's where the rhs of Theorem 5.2(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.

5 Power Bound as $\boldsymbol{\beta_0} \rightarrow \pm\infty$

In this section, we provide two-sided AE power bounds for invariant similar tests as $\beta_0 \rightarrow \pm\infty$ for fixed β_*. 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 CHJ.

Using Theorem 4.1, these results are used to obtain bounds on the probabilities that CS's 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 $\beta_0 \rightarrow \pm\infty$ and the probabilities that AR, LM, and CLR CS's have infinite length.

5.1 Density of Q as $\boldsymbol{\beta_0} \rightarrow \pm\infty$

The density of $Q := [S : T]'[S : T]$ when $[S : T]$ has the multivariate normal distribution in (2.3) 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, eqn. (6)). An explicit expression for $f_Q(q; \beta_*, \beta_0, \lambda, \Omega)$ is given in (11.2) in the SM.

Now, we determine the limit of the density $f_Q(q; \beta_*, \beta_0, \lambda, \Omega)$ as $\beta_0 \rightarrow \pm\infty$. Define

$$r_{uv} := \frac{\rho_{uv}}{(1 - \rho_{uv}^2)^{1/2}} \text{ and } \lambda_v := \lambda/\sigma_v^2 = \mu_\pi' \mu_\pi/\sigma_v^2. \tag{5.1}$$

Note that λ_v is the concentration parameter, which indexes the strength of the IV's. 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

$$\mu_\pi \cdot (1/\sigma_v, r_{uv}/\sigma_v) \in R^{k \times 2}, \tag{5.2}$$

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 (11.4) in Section 11.1 the SM.

Lemma 5.1 *For any fixed* (β_*,λ,Ω), $\lim_{\beta_0\to\pm\infty} f_Q(q;\beta_*,\beta_0,\lambda,\Omega) = f_Q(q;\rho_{uv},\lambda_v)$ *for all* 2×2 *variance matrices* q, *where* ρ_{uv} *and* λ_v *are defined in* (4.5) *and* (5.1), *respectively.*

Comment. Lemma 5.1 is proved by showing: $\lim_{\beta_0\to\pm\infty} c_{\beta_*}(\beta_0,\Omega) = \mp 1/\sigma_v$ and $\lim_{\beta_0\to\pm\infty} d_{\beta_*}(\beta_0,\Omega) = \pm\frac{\omega_2^2\beta_*-\omega_{12}}{\omega_2(\omega_1^2\omega_2^2-\omega_{12}^2)^{1/2}} = \mp\frac{\rho_{uv}}{\sigma_v(1-\rho_{uv}^2)^{1/2}}$, see Lemma 14.1 in the SM. When expressed in terms of Σ_*, the latter limit only depends on ρ_{uv}, σ_u, and σ_v and its functional form is of a relatively simple multiplicative form.

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.

5.2 Two-Sided AE Power Bound as $\boldsymbol{\beta_0}\to\pm\infty$

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 efficient 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

$$\beta_{2*} = \beta_0 - \frac{d_{\beta_0}(\beta_*-\beta_0)}{d_{\beta_0}+2r_{\beta_0}(\beta_*-\beta_0)} \text{ and } \lambda_2 = \lambda\frac{(d_{\beta_0}+2r_{\beta_0}(\beta_*-\beta_0))^2}{d_{\beta_0}^2}, \tag{5.3}$$

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 $POIS2(Q;\beta_0,\beta_*,\lambda)$ 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 $POIS2(Q;\beta_0,\beta_*,\lambda)$ and $\kappa_{2,\beta_0}(Q_T)$ are given in Section 16 in the SM.

The limit as $\beta_0\to\pm\infty$ of the $POIS2(Q;\beta_0,\beta_*,\lambda)$ statistic is shown in (16.6) in the SM to be

$$\begin{aligned} POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) &:= \frac{\psi(Q;\rho_{uv},\lambda_v)+\psi(Q;-\rho_{uv},\lambda_v)}{2\psi_2(Q_T;|\rho_{uv}|,\lambda_v)}, \text{ where} \\ \psi(Q;\rho_{uv},\lambda_v) &:= \exp(-\lambda_v(1+r_{uv}^2)/2)(\lambda_v\xi(Q;\rho_{uv}))^{-(k-2)/4}I_{(k-2)/2}(\sqrt{\lambda_v\xi(Q;\rho_{uv})}), \\ \psi_2(Q_T;|\rho_{uv}|,\lambda_v) &:= \exp(-\lambda_v r_{uv}^2/2)(\lambda_v r_{uv}^2 Q_T)^{-(k-2)/4}I_{(k-2)/2}(\sqrt{\lambda_v r_{uv}^2 Q_T}), \text{ and} \\ \xi(Q;\rho_{uv}) &:= Q_S+2r_{uv}Q_{ST}+r_{uv}^2Q_T, \end{aligned} \tag{5.4}$$

where Q, Q_S, Q_{ST}, and Q_T are defined in (3.1), ρ_{uv} is defined in (4.5), r_{uv} and λ_v are defined in (5.1), 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 $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v)$ test statistic. That is, $\kappa_{2,\infty}(q_T)$ is defined to satisfy

$$P_{Q_1|Q_T}(POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(q_T)|q_T) = \alpha \tag{5.5}$$

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 (11.3) in the SM and does not depend on β_0.

When $\rho_{uv} = 0$, the test based on $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v)$ 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 $\rho_{uv} = 0$.[4] 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 (12.5) and Theorem 12.1 in the SM that the AR test is also the best one-sided test as $\beta_0 \to +\infty$ and as $\beta_0 \to -\infty$.

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 5.2 *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* (5.3), *and* Ω, *the two-sided* AE *power envelope test* $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v)$ *defined in* (5.4) *and* (5.5) *satisfies*

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

Comments. **(i).** The power bound in Theorem 5.2 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 5.2 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 5.2 is the average of two equal quantities.

[4] 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 $ET = \mu_\pi d_{\beta_*}(\beta_0,\Omega)$, $\lim_{\beta_0\to\pm\infty} d_{\beta_*}(\beta_0,\Omega) \to \mp r_{uv}/\sigma_v$ by Lemma 14.1(e) in the SM, and $r_{uv} = 0$ when $\rho_{uv} = 0$.

(iv). Theorem 5.2 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 CHJ. In this case, the left-hand side (lhs) probabilities in Theorem 5.2 depend on π or, equivalently $(\lambda, \mu_\pi/||\mu_\pi||)$, rather than just λ. In this case, Theorem 5.2 holds with $P_{\beta_*,\beta_0,\lambda,\Omega}(\phi_{\beta_0}(Q) = 1)$ replaced by $\int P_{\beta_*,\lambda,\mu_\pi/||\mu_\pi||,\Omega}(\phi_{\beta_0}(S,T) = 1)dUnif(\mu_\pi/||\mu_\pi||)$ 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 $Unif(\cdot)$ denotes the uniform measure on the unit sphere in R^k.

5.3 Lower Bound on the Probability That a CS Has Infinite Length

Next, we combine Theorems 4.1 and 5.2 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.

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 5.3 *Suppose $CS_\phi(Y)$ is a* CS *based on invariant similar level α tests $\phi(Q_{\beta_0}(Y))$ that satisfy the continuity condition in Theorem* 4.1*.* (a) *For any fixed $(\beta_*, \lambda, \Omega)$,*

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

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

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

Comments. **(i).** All three lower bounds in Corollary 5.3 are the same. The different parts of Corollary 5.3 specify different probabilities or average probabilities that have this lower bound.

(ii). Corollary 5.3(a) also holds for a similar CS that does not satisfy any invariance properties. In this case, $P_{\beta_*,\lambda,\Omega}(RLength(CS_\phi(Y)) = \infty)$ is replaced by $\int P_{\beta_*,\lambda,\mu_\pi/||\mu_\pi||,\Omega}(RLength(CS_\phi(Y)) = \infty)dUnif(\mu_\pi/||\mu_\pi||)$ and analogously for the other three lhs terms that depend on $LLength(CS_\phi(Y))$

and/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 4.1.

5.4 Power of the AR, LM, and CLR Tests as $\boldsymbol{\beta_0} \rightarrow \pm\infty$

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

Theorem 5.4 *For fixed true* $(\beta_*, \lambda, \Omega)$, *the* AR, LM, *and* CLR *tests satisfy*

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

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

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

where AR, LM, *and* LR *are defined as functions of* Q *in* (3.3), $\chi^2_{m,1-\alpha}$ *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.

Comment. By Theorem 4.1(c), Theorem 5.4 provides the probabilities that the AR, LM, and CLR CS's 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.

6 Comparisons of Probabilities That Confidence Sets Have Infinite Length

Next, we investigate how close are the probabilities the CLR CS has infinite length to the lower bound in Corollary 5.3. Let POIS2 refer to the tests that generates 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 (5.4), which are the limits as $\beta_0 \rightarrow \pm\infty$ of the POIS2 tests. These tests depend on β_* (through $|\rho_{uv}|$) and λ_v. Let POIS2 and POIS2$_\infty$ CS's refer to the CS's constructed by inverting the POIS2 and POIS2$_\infty$ tests. These CS's are infeasible because they depend on knowing (β_*, λ).

Table I reports differences in simulated probabilities that the CLR and POIS2$_\infty$ CS's have infinite lengths. The latter provide a lower bound on infinite-length probabilities for CS's based on sign and invariant tests, such as the CLR CS, by Corollary 5.3(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 $\rho_{uv} = 0, .3, .5, .7, .9$. Table I also reports the probabilities that the CLR CS has infinite length for the same k and λ values and a subset of the ρ_{uv} values, viz., $0, .7, .9$. The true value of β_* is taken to be 0 wlog by Section 21 in the SM. The results for negative and positive ρ_{uv} values are the same by Section 21 in the SM, 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 optimal in some parameter scenarios. In particular, the differences in probabilities of infinite length (DPIL's) between the CLR and the POIS2$_\infty$ CS's are positive for numerous combinations of (k, λ, ρ_{uv}). The DPIL's 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$\in [.002, .016]$ over the λ values considered, whereas for $(k, \rho_{uv}) = (5, 0)$, DPIL$\in [.003, .031]$ and for $(k, \rho_{uv}) = (40, 0)$, DPIL$\in [.002, .049]$.[5] Hence, k has a noticeable effect on the magnitude of non-optimality of the CLR CS with larger values of k leading to larger non-optimality. For $(k, \lambda) = (5, 10)$, we have DPIL$\in [.002, .031]$ over the ρ_{uv} values considered, and for $(k, \lambda) = (20, 15)$, we have DPIL$\in [.001, .046]$ over the ρ_{uv} values considered. Hence, $|\rho_{uv}|$ also has a noticeable effect on the magnitude of non-optimality of the CLR CS in terms of DPIL's with non-optimality greatest at $\rho_{uv} = 0$.[6]

7 Optimality of CLR and LM Tests as $\boldsymbol{\rho}_{\mathbf{uv}} \to \pm 1$ or $\boldsymbol{\rho}_{\boldsymbol{\Omega}} \to \pm 1$

The results of Table I show that the magnitude of non-optimality 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, i.e., $\rho_\Omega := \omega_{12}/(\omega_1\omega_2)$.

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 semi-strong 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, CHJ show that the CLR and LM tests can be written as the limits of certain WAP LR tests, which indicates that they are at least close to being admissible.

[5] The simulation standard deviations of the DPIL's are in the range of $[.0000, 0014]$ with most being in the range of $[.0004, .0012]$, see Table SM-I in the SM.

[6] Table SM-I in the SM shows that the differences in probabilities that the AR and POIS2 CS's have infinite length are very large for large ρ_{uv} values for some λ values. For example, for $\rho_{uv} = .9$, they are as large as $.084, .196, .280, .353, .422$ for $k = 2, 5, 10, 20, 40$, respectively, for some λ values. As shown above, $AR = POIS2$ when $\rho_{uv} = 0$, so the differences are zero in this case and they increase in $|\rho_{uv}|$ for given (k, λ).

Let $d^2_{\beta_*} := d^2_{\beta_*}(\beta_0, \Omega)$ and $c^2_{\beta_*} := c^2_{\beta_*}(\beta_0, \Omega)$. As in Section 5.2, let $POIS2(Q; \beta_0, \beta_*, \lambda)$ 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 7.1 *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, as $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty$,*

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

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

Comments. (i). Theorem 7.1 shows that the CLR and LM tests have the same limit power as the POIS2 test. Theorem 7.1 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 $(\beta_*, \beta_0, \lambda, \Omega)$,

$$\frac{1}{2}(P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1) + P_{\beta_{2*},\beta_0,\lambda_2,\Omega}(\phi(Q) = 1) \leq P_{\beta_*,\beta_0,\lambda,\Omega}(POIS2(Q; \beta_0, \beta_*, \lambda). \tag{7.1}$$

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 CHJ, for any similar test $\phi([S : T])$ (that is not necessarily invariant), for any $(\beta_*, \beta_0, \lambda, \Omega)$, (7.1) holds with $P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1)$ replaced by the power average $\int P_{\beta_*,\beta_0,\lambda,\mu_\pi/||\mu_\pi||,\Omega}(\phi([S : T]) = 1) dUnif(\mu_\pi/||\mu_\pi||)$ and likewise for the second lhs summand in (7.1). Hence, the POIS2 test also determines this average power envelope for similar tests.

These results and Theorem 7.1 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 (FCP) lower bounds for CS's based on invariant similar tests and similar tests. Specifically, one minus the lhs in (7.1), which equals the average FCP 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 rhs in (7.1). In the case of non-invariant similar tests, the bound is on the average of the FCP's 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 7.1 shows that the CLR and LM CS's have optimal average FCP properties asymptotically when $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty \neq 0$.

(iv). Theorem 7.1 does not apply when the IV's are completely irrelevant, i.e., $\lambda = 0$, because $\lambda = 0$ implies that $c_\infty = 0$. However, Theorem 7.1 does cover some cases where the IV's can be arbitrarily weak, see Theorem 7.2 below.

Next, we provide conditions under which $\lambda d_{\beta_*}^2 \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty \in R \backslash \{0\}$, as is assumed in Theorem 7.1. First, if β_0 and Ω are fixed, Ω is nonsingular, and (β_*, λ) satisfy $\lambda \to \infty$ and

$$\lambda^{1/2}(\beta_* - \beta_0) \to L \in R \text{ as } \lambda \to \infty, \tag{7.2}$$

then $\lambda d_{\beta_*}^2 \to \infty$ 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 (7.2) 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 "semi-strong" identification. As far as we are aware, this is the only optimality property in the literature for tests under semi-strong identification. The scenario in (7.2) also covers finite-sample, i.e., fixed n, cases in which $Z'Z$ is fixed, π diverges, i.e., $||\pi|| \to \infty$, and $\lambda_{\min}(Z'Z) > 0$. In these cases, $\lambda = \pi' Z' Z \pi \to \infty$ as $||\pi|| \to \infty$.

The most novel cases in which Theorem 7.1 applies are when $\rho_{uv} \to \pm 1$ or $\rho_\Omega \to \pm 1$. The next result shows that $\lambda d_{\beta_*}^2 \to \infty$ 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 7.2 (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) $\lim_{\rho_{uv} \to \pm 1} \lambda d_{\beta_*}^2 = \infty$ *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) $\lim_{\rho_\Omega \to \pm 1} \lambda d_{\beta_*}^2 = \infty$ *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) $\lim_{(\rho_{uv}, \beta_0) \to (1, \pm\infty)} \lambda d_{\beta_*}^2 = \lim_{(\rho_{uv}, \beta_0) \to (-1, \pm\infty)} \lambda d_{\beta_*}^2 = \infty$.

(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) $\lim_{(\rho_\Omega, \beta_0) \to (1, \pm\infty)} \lambda d_{\beta_*}^2 = \infty$ *provided* $\beta_* \neq \omega_1/\omega_2$ *and* $\lim_{(\rho_\Omega, \beta_0) \to (-1, \pm\infty)} \lambda d_{\beta_*}^2 = \infty$ *provided* $\beta_* \neq -\omega_1/\omega_2$.

Comments. (i). Combining Theorems 7.1 and 7.2 provides analytic finite-sample limiting optimality results for the CLR and LM tests and CS's 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 IV's 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 $(1-\rho_\Omega^2)^{1/2}\omega_1 = (1-\rho_{uv}^2)^{1/2}\sigma_u$ by (15.10) in the SM. Thus, $\rho_{uv} \to \pm 1$ implies $|\rho_\Omega| \to 1$ and/or $\omega_1 \to 0$. And, $\rho_\Omega \to \pm 1$ implies $|\rho_{uv}| \to 1$ and/or $\sigma_u \to 0$.

8 General Power/False-Coverage-Probability Comparisons

By Theorem 4.1, the results in Table I equal power differences (PD's) between the POIS2 and CLR tests as the null value $\beta_0 \to \pm\infty$ for fixed true value $\beta_* = 0$. Here, we consider PD's between the POIS2 and CLR tests for finite β_0 values, rather than PD's as $\beta_0 \to \pm\infty$. Specifically, Table II reports maximum and average PD's 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 $\beta_* = 0$ (and $\omega_1^2 = \omega_2^2 = 1$) is wlog. These PD's are equivalent to false coverage probability differences (FCPD's) between the CLR and POIS2 CS's 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 $.7, .8, .9$ when $k = 2$ and 5 and $\rho_{uv} = .9$. The positive and negative β_0 values considered are those with $|\beta_0| \in \{.25, .5, ..., 3.75, 4, 5, 7.5, 10, 50, 100, 1000, 10000\}$.

The number of simulation repetitions employed is $5,000$. The critical values are determined using $100,000$ simulation repetitions. For example, the simulation standard deviations for the PD's for $(\rho_{uv}, k) = (0, 20)$ and any fixed (β_0, λ) value range from $[.0013, .0040]$ across different (β_0, λ) values, which compares to simulated averages of the PD's over (β_0, λ) values that are of the .014 order of magnitude.

Tables II(a) and II(b) contain the same numbers, but are reported differently to make the patterns in the table more clear. Table II(a) shows variation across k for fixed ρ_{uv}, whereas Table II(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 PD 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 PD simulations. The sixth column in the tables reports the power of the CLR test at the (β_0, λ) values that maximize the PD for given (ρ_{uv}, k), i.e., at $(\beta_{0,\max}, \lambda_{\max})$.

Table II shows that the maximum (over (β_0, λ)) PD's between the POIS2 and CLR tests range between $[.016, .061]$ over the (ρ_{uv}, k) values. On the other hand, the average (over (β_0, λ)) PD's only range between $[.002, .016]$ 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.

In contrast, the analogous maximum and average PD ranges for the AR test are $[.079, .513]$ and $[.012, .179]$, see Table SM-III in the SM. For the LM test, they are $[.242, .784]$ and $[.010, .203]$, see Table SM-IV in the SM. 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 II(a) shows that the maximum and average (over (β_0, λ)) PD's for the CLR test are clearly increasing in k. Table II(a) shows that for $\rho_{uv} \geq .3$, the PD's 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 II(a) also shows that for each ρ_{uv}, the PD's are maximized at λ values that (weakly) increase with k. The increase is particularly evident going from $k = 20$ to 40.

Table II(b) shows that for $k \geq 5$, the maximum PD's are more or less the same for $\rho_{uv} \leq .7$, but noticeably lower for $\rho_{uv} = .9$. For $k = 2$, the maximum PD's are more or less the same for all ρ_{uv} considered. Table II(b) shows that, for each k, the PD's are maximized at $|\beta_0|$ values that are closer to 0 as ρ_{uv} increases. Table II(b) also shows that, for each k, the PD's are maximized at λ values that are closer to 0 as ρ_{uv} increases.[7]

In sum, the maximum PD's 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.

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.

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

9 Differences between CLR Power and an Average Over λ Power Envelope

In this section, 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 CHJ. In particular, we use the uniform weight function over the 36 values of λ in $\{2.5, 5.0, ..., 90.0\}$.

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 (11.5) in the SM). The WAP2(Q, β_0, β_*) conditional critical value $\kappa_{2,\beta_0,J}(q_T)$ is defined to satisfy $P_{Q_1|Q_T}(WAP2(Q, \beta_0, \beta_*) > \kappa_{2,\beta_0,J}(q_T)|q_T) = \alpha$ 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 (11.3) in the SM.

To be consistent with Tables I and II, we report PD's between the WAP2(Q, β_0, β_*) and CLR tests for $\beta_* = 0$ and a range of β_0 values. These PD's are equivalent to the FCPD's between the CLR and WAP2 CS's for fixed true β_* and varying incorrect β_0 values. The differences are necessarily nonnegative.

We consider $\rho_{uv} \in \{0, .3, .5, .7, .9, .95, .99\}$, $k = 2, 5, 10, 20, 40$, the same β_0 values as in Table II, and $\omega_1^2 = \omega_2^2 = 1$. Since $\beta_* = 0$, $\rho_\Omega = \rho_{uv}$. Section 21 shows that taking $\beta_* = 0$ and $\omega_1^2 = \omega_2^2 = 1$ is wlog 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 $1,000$ 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 SD's 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 III 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 SM. Table IV reports summary results for all values of k. In particular, Table IV(a) provides the maxima over β_0 of the average over λ PD's for each (ρ_{uv}, k). Table IV(b) provides the average over β_0 of the average over λ PD's for each (ρ_{uv}, k).

Table III shows that the CLR test has power quite close to the WAP2 power envelope for $k = 5$. The PD's for $\rho_{uv} \in \{0, .3, .5, .7\}$, we have PD$\in$ $[.000, .005]$ and SD$\in$ $[.0003, .0007]$ across all β_0 values. For $\rho_{uv} \in \{.9, .95, .99\}$, we have PD$\in$ $[.000, .001]$ and SD$\in$ $[.0000, .0003]$ across all β_0 values.

Table IV shows that PD's between the WAP2 power envelope and the CLR power are increasing in k and decreasing in $|\rho_{uv}|$. For $k = 2$, the maximum PD over β_0 and ρ_{uv} values is very small:

.004. In the worst case for CLR, which is when $(k, \rho_{uv}) = (40, 0)$, the maximum PD over β_0 values is substantially larger: .024. The average (over β_0 values) PD in this case is .013, which is not very large. For $k = 40$ and $\rho_{uv} \geq .9$, the maximum PD (over β_0 and ρ_{uv} values) is very small: .004. This is consistent with the theoretical optimality properties of the CLR test as $\rho_{uv} \to \pm 1$ described in Section 7. For $k = 40$ and $\rho_{uv} \geq .9$, the average PD (over β_0 values and the five ρ_{uv} values) is very small: .000. The second worst case for CLR in Table V is when $(k, \rho_{uv}) = (20, 0)$. In this case, the maximum PD over β_0 values is .013, which is noticeably lower than .024 for $(k, \rho_{uv}) = (40, 0)$.

In conclusion, the results in Tables III and IV 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 .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.

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.

Andrews, D. W. K. (1998): "Hypothesis Testing with a Restricted Parameter Space," *Journal of Econometrics*, 84, 155–199.

Andrews, D. W. K., and P. Guggenberger (2015): "Identification- and Singularity-Robust Inference for Moment Condition Models," Cowles Foundation Discussion Paper No. 1978, Yale University.

——— (2016): "Asymptotic Size of Kleibergen's LM and Conditional LR Tests for Moment Condition Models," *Econometric Theory,* 32, forthcoming. Also available as Cowles Foundation Discussion Paper No. 1977, Yale University.

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.

——— (2006): "Optimal Two-sided Invariant Similar Tests for Instrumental Variables Regression," *Econometrica,* 74, 715–752.

Andrews, D. W. K. and W. Ploberger (1994): "Optimal Tests When a Nuisance Parameter Is Present Only Under the Alternative," *Econometrica,* 63, 1383–1414.

Chamberlain, G. (2007): "Decision Theory Applied to an Instrumental Variables Model," *Econometrica,* 75, 609–652.

Chernozhukov, V., C. Hansen, and M. Jansson (2009): "Admissible Invariant Similar Tests for Instrumental Variables Regression," *Econometric Theory,* 25, 806–818.

Dufour, J.-M. (1997): "Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models," *Econometrica,* 65, 1365–1387.

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.

Hillier, G. H. (1984): "Hypothesis Testing in a Structural Equation: Part I, Reduced Form Equivalence and Invariant Test Procedures," unpublished manuscript, Dept. of Econometrics and Operations Research, Monash University.

—— (2009): "Exact Properties of the Conditional Likelihood Ratio Test in an IV Regression Model," *Econometric Theory,* 25, 915–957.

Kleibergen, F. (2002): "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression," *Econometrica,* 70, 1781-1803.

Mikusheva, A. (2010): "Robust Confidence Sets in the Presence of Weak Instruments," *Journal of Econometrics,* 157, 236–247.

Mills, B., M. J. Moreira, and L. P. Vilela (2014): "Tests Based on t-Statistics for IV Regression with Weak Instruments," *Journal of Econometrics,* 182, 351–363.

Moreira, M. J. (2003): "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica,* 71, 1027-1048.

——— (2009): "Tests with Correct Size When Instruments Can Be Arbitrarily Weak, *Journal of Econometrics,* 152, 131–140.

Moreira, H. and M. J. Moreira (2013): "Contributions to the Theory of Optimal Tests," unpublished manuscript, FGV/EPGE, Brasil.

van der Vaart, A. W. (1998): *Asymptotic Statistics.* Cambridge, UK: Cambridge University Press.

van der Vaart, A. W. and J. A. Wellner (1996): *Weak Convergence and Empirical Processes.* New York: Springer.

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.

TABLE I. 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}

k	λ	CLR–POIS2$_\infty$					POIS2$_\infty$		
		$\rho_{uv}=0$	.3	.5	.7	.9	$\rho_{uv}=0$	.7	.9
2	1	.002	.004	**.003**	**.002**	**.002**	.870	.863	.850
2	3	.007	.003	.002	.001	.001	.680	.656	.613
2	5	.012	**.008**	**.003**	**.002**	.001	.497	.455	.410
2	7	**.016**	.005	.001	.001	-.000	.344	.298	.262
2	10	.013	.006	.002	.001	.000	.184	.142	.122
2	15	.008	.004	.002	.001	.001	.056	.035	.029
2	20	.003	.001	.001	.000	-.000	.014	.008	.007
5	1	.003	.003	.002	.002	.002	.900	.898	.882
5	3	.010	.010	**.006**	.002	**.004**	.778	.751	.671
5	5	.017	.012	.004	.003	-.001	.639	.576	.466
5	7	.026	.012	.003	.001	-.002	.502	.407	.304
5	10	**.031**	**.016**	.005	**.004**	.002	.323	.221	.144
5	12	.029	.011	.003	.000	-.001	.231	.140	.086
5	15	.023	.013	.005	.003	.001	.134	.062	.036
5	20	.013	.007	.003	.001	.000	.048	.016	.008
5	25	.006	.002	.001	.000	.000	.016	.004	.001
10	1	.001	.002	.002	.001	.000	.919	.916	.904
10	5	.018	.015	.008	**.005**	**.004**	.731	.667	.526
10	10	.032	.014	.003	**.005**	.001	.459	.320	.176
10	15	**.037**	**.020**	**.009**	.004	.000	.239	.111	.046
10	17	.035	.017	.008	.003	.000	.176	.069	.025
10	20	.027	.014	.006	.002	.000	.110	.032	.010
10	25	.017	.008	.003	.001	.000	.045	.008	.002
10	30	.009	.004	.002	.000	-.000	.017	.002	.000
20	1	.001	.003	.002	.002	.000	.929	.926	.919
20	5	.014	.011	.006	.003	**.006**	.809	.766	.620
20	10	.035	.022	**.011**	**.010**	.001	.603	.463	.246
20	15	**.046**	**.025**	**.011**	.008	.001	.390	.215	.073
20	20	.044	.021	**.011**	.005	.000	.226	.080	.018
20	30	.033	.012	.005	.001	.000	.056	.007	.001
20	40	.007	.003	.001	.000	.000	.010	.001	.000
40	1	.002	.001	.001	.001	.000	.937	.937	.934
40	5	.011	.008	.005	.001	**.009**	.859	.838	.717
40	10	.028	.020	.009	**.009**	.002	.721	.614	.354
40	15	.043	.022	.011	**.009**	.001	.555	.372	.129
40	20	**.049**	**.029**	**.013**	**.009**	.001	.394	.186	.038
40	30	.043	.021	.011	.004	.000	.155	.028	.002
40	40	.022	.012	.005	.001	.000	.046	.003	.000
40	60	.002	.001	.000	.000	.000	.003	.000	.000

TABLE II. 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}

						POIS2–CLR	
ρ_{uv}	k	$\lambda_{\max}$	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	max	average
.0	2	7	−10000.00	1.00	.66	.021	.006
.0	5	10	−50.00	1.00	.68	.030	.009
.0	10	15	−50.00	1.00	.76	.038	.012
.0	20	15	10.00	−1.00	.60	.042	.014
.0	40	22	−50.00	1.00	.66	**.059**	**.016**
.3	2	10	3.75	−0.96	.86	.019	.005
.3	5	10	3.50	−0.96	.73	.034	.008
.3	10	10	3.00	−0.94	.59	.032	.009
.3	20	15	3.50	−0.96	.66	.045	.012
.3	40	22	4.00	−0.97	.72	**.061**	**.014**
.5	2	5	2.00	−0.87	.64	.016	.004
.5	5	10	2.25	−0.90	.82	.029	.005
.5	10	10	2.00	−0.87	.70	.037	.007
.5	20	10	1.75	−0.82	.53	.046	.009
.5	40	15	1.75	−0.82	.59	**.050**	**.012**
.7	2	5	1.50	−0.75	.81	.016	.002
.7	5	5	1.50	−0.75	.67	.033	.003
.7	10	7	1.50	−0.75	.71	.036	.005
.7	20	7	1.25	−0.61	.54	.042	.006
.7	40	15	1.50	−0.75	.84	**.050**	**.008**
.9	2	0.9	1.25	−0.63	.46	.017	.002
.9	5	0.9	1.00	−0.22	.33	.017	.002
.9	10	3	1.25	−0.63	.77	.027	.003
.9	20	3	1.00	−0.22	.61	.032	.003
.9	40	5	1.25	−0.63	.75	**.040**	**.004**

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

						POIS2–CLR	
k	ρ_{uv}	$\lambda_{\max}$	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	max	average
2	.0	7	−10000.00	1.00	.66	**.021**	**.006**
2	.3	10	3.75	−0.96	.86	.019	.005
2	.5	5	2.00	−0.87	.64	.016	.004
2	.7	5	1.50	−0.75	.81	.016	.002
2	.9	0.9	1.25	−0.63	.46	.017	.002
5	.0	10	−50.00	1.00	.68	.030	**.009**
5	.3	10	3.50	−0.96	.73	**.034**	.008
5	.5	10	2.25	−0.90	.82	.029	.005
5	.7	5	1.50	−0.75	.67	.033	.003
5	.9	0.9	1.00	−0.22	.33	.017	.002
10	.0	15	−50.00	1.00	.76	**.038**	**.012**
10	.3	10	3.00	−0.94	.59	.032	.009
10	.5	10	2.00	−0.87	.70	.037	.007
10	.7	7	1.50	−0.75	.71	.036	.005
10	.9	3	1.25	−0.63	.77	.027	.003
20	.0	15	10.00	−1.00	.60	.042	**.014**
20	.3	15	3.50	−0.96	.66	.045	.012
20	.5	10	1.75	−0.82	.53	**.046**	.009
20	.7	7	1.25	−0.61	.54	.042	.006
20	.9	3	1.00	−0.22	.61	.032	.003
40	.0	22	−50.00	1.00	.66	.059	**.016**
40	.3	22	4.00	−0.97	.72	**.061**	.014
40	.5	15	1.75	−0.82	.59	.050	.012
40	.7	15	1.50	−0.75	.84	.050	.008
40	.9	5	1.25	−0.63	.75	.040	.004

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

β_0	$\rho_{uv,0}$		WAP2–CLR						
	$\rho_{uv} = 0$	.9	$\rho_{uv} = 0$	.3	.5	.7	.9	.95	.99
-10000.00	1.00	1.00	.005	.002	.001	.001	.000	-.000	.000
-100.00	1.00	1.00	**.005**	.002	.001	.001	.000	-.001	-.000
-10.00	1.00	1.00	.005	.002	.001	.000	.000	-.000	-.000
-4.00	.97	1.00	.003	.001	.000	-.000	.000	.000	-.000
-3.00	.95	.99	.003	.001	.000	.000	-.000	**.001**	**.000**
-2.00	.89	.99	.002	.001	.000	.001	-.000	-.001	-.000
-1.50	.83	.98	.001	.001	.001	.000	.000	-.001	-.000
-1.00	.71	.97	.001	.000	-.000	-.000	-.000	.000	-.000
-0.75	.60	.97	.000	-.000	.001	-.000	-.000	.000	.000
-0.50	.45	.95	-.000	-.000	-.001	-.001	-.000	-.001	-.000
-0.25	.24	.94	-.001	-.001	-.001	-.000	-.000	.001	-.001
0.25	$-.24$	.83	-.000	-.001	-.001	-.000	-.001	.000	.000
0.50	$-.45$	.68	.001	.000	.000	.000	.000	-.001	.000
0.75	$-.60$	.33	.000	.001	.001	.001	.000	.000	.000
1.00	$-.71$	$-.22$	.002	.001	.001	.001	.000	.000	.000
1.50	$-.83$	$-.81$	.001	.002	.003	**.003**	**.001**	-.000	.000
2.00	$-.89$	$-.93$	.002	.003	**.004**	.002	.000	-.001	-.000
3.00	$-.95$	$-.98$	.003	**.005**	.003	.001	.000	.000	.000
4.00	$-.97$	$-.99$	.004	.005	.002	.001	.000	.001	.000
10.00	-1.00	-1.00	.005	.003	.001	.001	.000	.000	.000
100.00	-1.00	-1.00	.005	.003	.001	.000	.000	-.001	.000
10000.00	-1.00	-1.00	.005	.002	.001	.001	.000	-.000	.000

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

k	(a) Maxima over β_0							(b) Averages over β_0						
	$\rho_{uv} = 0$	.3	.5	.7	.9	.95	.99	$\rho_{uv} = 0$	.3	.5	.7	.9	.95	.99
2	.004	.003	.002	.002	.001	.001	.001	.002	.002	.001	.001	.000	.000	.000
5	.005	.005	.004	.003	.001	.001	.000	.003	.002	.001	.001	.000	.000	.000
10	.011	.010	.008	.005	.004	.003	.003	.007	.006	.004	.002	.001	.001	.001
20	.013	.012	.010	.007	.002	.001	.002	.008	.007	.005	.002	.000	.000	.000
40	.024	.021	.017	.011	.004	.001	.000	.013	.011	.007	.004	.000	.000	.000

Supplemental Material to
A NOTE ON OPTIMAL INFERENCE IN THE LINEAR IV MODEL

By

Donald W. K. Andrews, Vadim Marmer, and Zhengfei Yu

January 2017

COWLES FOUNDATION DISCUSSION PAPER NO. 2073

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Supplemental Material

for

A Note on Optimal Inference in the Linear IV Model

Donald W. K. Andrews
Cowles Foundation for Research in Economics
Yale University

Vadim Marmer
Vancouver School of Economics
University of British Columbia

Zhengfei Yu
Faculty of Humanities and Social Sciences
University of Tsukuba

First Version: October 23, 2015
Revised: January 30, 2017

Contents

10 Outline

References to sections, theorems, and lemmas with section numbers less than 10 refer to sections and results in the main paper.

Section 11 of this Supplemental Material (SM) provides expressions for the densities $f_Q(q;\beta_*,\beta_0,\lambda,\Omega)$, $f_{Q_1|Q_T}(q_1|q_T)$, and $f_Q(q;\rho_{uv},\lambda_v)$, expressions for the POIS2 test statistic and critical value of AMS, and expressions for the one-to-one transformations between the reduced-form and structural variance matrices. Section 12 provides one-sided power bounds for invariant similar tests as $\beta_0 \to \pm\infty$, where β_0 denotes the null hypothesis value. Section 13 corrects (4.1) of AMS, which concerns the two-point weight function that defines AMS's two-sided AE power envelope.

Section 14 proves Lemma 5.1. Section 15 proves Theorem 4.1 and its Comment (v). Section 16 proves Theorem 5.2 and its Comment (iv), Corollary 5.3 and its Comment (ii), and Theorem 5.4. Section 17 proves Theorem 7.1. Section 18 proves Theorem 12.1 and Lemmas 13.1 and 13.2.

Section 19 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 20 shows how the model is transformed to go from a testing problem of $H_0:\beta=\beta_0$ versus $H_1:\beta=\beta_*$ for $\pi\in R^k$ and fixed Ω to a testing problem of $H_0:\overline{\beta}=0$ versus $H_1:\overline{\beta}=\overline{\beta}_*$ for some $\overline{\pi}\in R^k$ and some fixed $\overline{\Omega}$ with diagonal elements equal to one. This links the model considered here to the model used in the Andrews, Moreira, and Stock (2006) (AMS) numerical work.

Section 21 shows how the model is transformed to go from a testing problem of $H_0:\beta=\beta_0$ versus $H_1:\beta=\beta_*$ for $\pi\in R^k$ and fixed Ω to a testing problem of $H_0:\overline{\beta}=\overline{\beta}_0$ versus $H_1:\overline{\beta}=0$ for some $\overline{\pi}\in R^k$ and some fixed $\overline{\Omega}$ with diagonal elements equal to one. These transformation results imply that there is no loss in generality in the numerical results of the paper to taking $\omega_1^2=\omega_2^2=1$, $\beta_*=0$, and $\rho_{uv}\in[0,1]$ (rather than $\rho_{uv}\in[-1,1]$).

Section 22 provides numerical results that supplement the results given in Tables I-IV in the main paper.

11 Definitions

11.1 Densities of Q when $\boldsymbol{\beta} = \boldsymbol{\beta}_*$ and when $\boldsymbol{\beta_0} \rightarrow \pm\infty$

In this subsection, we provide expressions for (i) the density $f_Q(q;\beta_*,\beta_0,\lambda,\Omega)$ of Q when the true value of β is β_*, and the null value β_0 is finite, (ii) the conditional density $f_{Q_1|Q_T}(q_1|q_T)$ of Q_1 given $Q_T = q_T$, and (iii) the limit of $f_Q(q;\beta_*,\beta_0,\lambda,\Omega)$ as $\beta_0 \rightarrow \pm\infty$.

Let

$$\xi_{\beta_*}(q) = \xi_{\beta_*}(q;\beta_0,\Omega) := c_{\beta_*}^2 q_S + 2c_{\beta_*} d_{\beta_*} q_{ST} + d_{\beta_*}^2 q_T, \tag{11.1}$$

where $c_{\beta_*} = c_{\beta_*}(\beta_0,\Omega)$ and $d_{\beta_*} = d_{\beta_*}(\beta_0,\Omega)$. As in Section 5, $f_Q(q;\beta_*,\beta_0,\lambda,\Omega)$ denotes the density of $Q := [S:T]'[S:T]$ when $[S:T]$ has the multivariate normal distribution in (2.3) with $\beta = \beta_*$ and $\lambda = \mu_\pi'\mu_\pi$. This noncentral Wishart density is

$$\begin{aligned} f_Q(q;\beta_*,\beta_0,\lambda,\Omega) &= K_1 \exp(-\lambda(c_{\beta_*}^2 + d_{\beta_*}^2)/2)\det(q)^{(k-3)/2}\exp(-(q_S+q_T)/2) \\ &\quad \times(\lambda\xi_{\beta_*}(q))^{-(k-2)/4} I_{(k-2)/2}(\sqrt{\lambda\xi_{\beta_*}(q)}),\ \text{where} \\ q = \begin{bmatrix} q_S & q_{ST} \\ q_{ST} & q_T \end{bmatrix},\ q_1 &= \begin{pmatrix} q_S \\ q_{ST} \end{pmatrix} \in R^+ \times R,\ q_T \in R^+, \end{aligned} \tag{11.2}$$

$K_1^{-1} = 2^{(k+2)/2} pi^{1/2}\Gamma((k-1)/2)$, $I_\nu(\cdot)$ denotes the modified Bessel function of the first kind of order ν, $pi = 3.1415...$, and $\Gamma(\cdot)$ is the gamma function. This holds by Lemma 3(a) of AMS with $\beta = \beta_*$.

By Lemma 3(c) of AMS, the conditional density of Q_1 given $Q_T = q_T$ when $[S:T]$ is distributed as in (2.3) with $\beta = \beta_0$ is

$$f_{Q_1|Q_T}(q_1|q_T) := K_1 K_2^{-1}\exp(-q_S/2)\det(q)^{(k-3)/2} q_T^{-(k-2)/2}, \tag{11.3}$$

which does not depend on β_0, λ, or Ω.

By Lemma 5.1, the limit of $f_Q(q;\beta_*,\beta_0,\lambda,\Omega)$ as $\beta_0 \rightarrow \pm\infty$ is the density $f_Q(q;\rho_{uv},\lambda_v)$. As in Section 5, $f_Q(q;\rho_{uv},\lambda_v)$ denotes the density of $Q := [S:T]'[S:T]$ when $[S:T]$ has a multivariate normal distribution with means matrix in (5.2), all variances equal to one, and all covariances equal to zero. This is a noncentral Wishart density that has following form:

$$\begin{aligned} f_Q(q;\rho_{uv},\lambda_v) &= K_1\exp(-\lambda_v(1+r_{uv}^2)/2)\det(q)^{(k-3)/2}\exp(-(q_S+q_T)/2) \\ &\quad \times(\lambda_v\xi(q;\rho_{uv}))^{-(k-2)/4} I_{(k-2)/2}(\sqrt{\lambda_v\xi(q;\rho_{uv})}),\ \text{where} \\ \xi(q;\rho_{uv}) &:= q_S + 2r_{uv}q_{ST} + r_{uv}^2 q_T. \end{aligned} \tag{11.4}$$

This expression for the density holds by the proof of Lemma 3(a) of AMS with means matrix $\mu_\pi \cdot (1/\sigma_v, r_{uv}/\sigma_v)$ in place of the means matrix $\mu_\pi \cdot (c_\beta, d_\beta)$.

11.2 POIS2 Test

Here we define the $POIS2(q_1, q_T; \beta_0, \beta_*, \lambda)$ test statistic of AMS, which is analyzed in Section 5, and its conditional critical value $\kappa_{2,\beta_0}(q_T)$.

Given (β_*, λ), the parameters (β_{2*}, λ_2) are defined in (5.3), which is the same as (4.2) of AMS. By Cor. 1 of AMS, the optimal average-power test statistic against (β_*, λ) and (β_{2*}, λ_2) is

$$\begin{aligned} POIS2(Q; \beta_0, \beta_*, \lambda) &:= \frac{\psi(Q; \beta_0, \beta_*, \lambda) + \psi(Q; \beta_0, \beta_{2*}, \lambda_2)}{2\psi_2(Q_T; \beta_0, \beta_*, \lambda)}, \text{ where} \\ \psi(Q; \beta_0, \beta, \lambda) &:= \exp(-\lambda(c_\beta^2 + d_\beta^2)/2)(\lambda\xi_\beta(Q))^{-(k-2)/4} I_{(k-2)/2}(\sqrt{\lambda\xi_\beta(Q)}), \\ \psi_2(Q_T; \beta_0, \beta, \lambda) &:= \exp(-\lambda d_\beta^2/2)(\lambda d_\beta^2 Q_T)^{-(k-2)/4} I_{(k-2)/2}(\sqrt{\lambda d_\beta^2 Q_T}), \end{aligned} \tag{11.5}$$

Q and Q_T are defined in (3.1), $c_\beta = c_\beta(\beta, \Omega)$ and $d_\beta = d_\beta(\beta, \Omega)$ are defined in (2.3), $I_\nu(\cdot)$ is defined in (11.2), $\xi_\beta(Q)$ is defined in (11.1) with Q and β in place of q and β_*, and $\lambda := \mu_\pi' \mu_\pi$. Note that $\psi_2(Q_T; \beta_*, \lambda) = \psi_2(Q_T; \beta_{2*}, \lambda_2)$ by (5.3).

Let $\kappa_{2,\beta_0}(q_T)$ denote the conditional critical value of the $POIS2(Q; \beta_0, \beta_*, \lambda)$ test statistic. That is, $\kappa_{2,\beta_0}(q_T)$ is defined to satisfy

$$P_{Q_1|Q_T}(POIS2(Q; \beta_0, \beta_*, \lambda) > \kappa_{2,\beta_0}(q_T)|q_T) = \alpha \tag{11.6}$$

for all $q_T \geq 0$, where $P_{Q_1|Q_T}(\cdot|q_T)$ denotes probability under the density $f_{Q_1|Q_T}(\cdot|q_T)$ defined in (11.3). The critical value function $\kappa_{2,\beta_0}(\cdot)$ depends on $(\beta_0, \beta_*, \lambda, \Omega)$ and k (and (β_{2*}, λ_2) through (β_*, λ)).

11.3 Structural and Reduced-Form Variance Matrices

Let u_i, v_{1i}, and v_{2i} denote the ith elements of u, v_1, and v_2, respectively. We have

$$v_{1i} := u_i + v_{2i}\beta \text{ and } \Omega = \begin{bmatrix} \omega_1^2 & \omega_{12} \\ \omega_{12} & \omega_2^2 \end{bmatrix}, \tag{11.7}$$

where β denotes the true value.

Given the true value β and some structural error variance matrix Σ, the corresponding reduced-form error variance matrix $\Omega(\beta, \Sigma)$ is

$$\begin{aligned}
\Omega(\beta, \Sigma) &:= Var\left(\left(\begin{array}{c} v_{1i} \\ v_{2i} \end{array}\right)\right) = Var\left(\left(\begin{array}{c} u_i + v_{2i}\beta \\ v_{2i} \end{array}\right)\right) \\
&= \left[\begin{array}{cc} 1 & \beta \\ 0 & 1 \end{array}\right] \Sigma \left[\begin{array}{cc} 1 & 0 \\ \beta & 1 \end{array}\right] = \left[\begin{array}{cc} \sigma_u^2 + 2\sigma_{uv}\beta + \sigma_v^2\beta^2 & \sigma_{uv} + \sigma_v^2\beta \\ \sigma_{uv} + \sigma_v^2\beta & \sigma_v^2 \end{array}\right], \text{ where} \\
\Sigma &= \left[\begin{array}{cc} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{array}\right]. \qquad (11.8)
\end{aligned}$$

Given the true value β and the reduced-form error variance matrix Ω, the structural variance matrix $\Sigma(\beta, \Omega)$ is

$$\begin{aligned}
\Sigma(\beta, \Omega) &:= Var\left(\left(\begin{array}{c} u_i \\ v_{2i} \end{array}\right)\right) = Var\left(\left(\begin{array}{c} v_{1i} - v_{2i}\beta \\ v_{2i} \end{array}\right)\right) \qquad (11.9) \\
&= \left[\begin{array}{cc} 1 & -\beta \\ 0 & 1 \end{array}\right] \Omega \left[\begin{array}{cc} 1 & 0 \\ -\beta & 1 \end{array}\right] = \left[\begin{array}{cc} \omega_1^2 - 2\omega_{12}\beta + \omega_2^2\beta^2 & \omega_{12} - \omega_2^2\beta \\ \omega_{12} - \omega_2^2\beta & \omega_2^2 \end{array}\right].
\end{aligned}$$

Let $\sigma_u^2(\beta, \Omega)$, $\sigma_v^2(\beta, \Omega)$, and $\sigma_{uv}(\beta, \Omega)$ denote the $(1,1)$, $(2,2)$, and $(1,2)$ elements of $\Sigma(\beta, \Omega)$. Let $\rho_{uv}(\beta, \Omega)$ denote the correlation implied by $\Sigma(\beta, \Omega)$.

In the asymptotics as $\beta_0 \to \pm\infty$, we fix β_* and Ω and consider the testing problem as $\beta_0 \to \pm\infty$. Rather than fixing Ω, one can equivalently fix the structural variance matrix when $\beta = \beta_*$, say at Σ_*. Given β_* and Σ_*, there is a unique reduced-form error variance matrix $\Omega = \Omega(\beta_*, \Sigma_*)$ defined using (11.8). Significant simplifications in certain formulae occur when they are expressed in terms of Σ_*, rather than Ω, e.g., see Lemma 14.1(e) below.

For notational simplicity, we denote the $(1,1)$, $(2,2)$, and $(1,2)$ elements of Σ_* by σ_u^2, σ_v^2, and σ_{uv}, respectively, without any $*$ subscripts. As defined in (4.5), $\rho_{uv} := \sigma_{uv}/(\sigma_u\sigma_v)$. Thus, ρ_{uv} is the correlation between the structural and reduced-form errors u_i and v_{2i} when the true value of β is β_*. Note that ρ_{uv} does not change when (β_*, Σ_*) is fixed (or, equivalently, $(\beta_*, \Omega) = (\beta_*, \Omega(\beta_*, \Sigma_*))$ is fixed) and β_0 is changed. Also, note that $\sigma_v^2 = \omega_2^2$ because both denote the variance of v_{2i} under $\beta = \beta_*$ and $\beta = \beta_0$.

12 One-Sided Power Bound as $\beta_0 \to \pm\infty$

In this section, we provide one-sided power bounds for invariant similar tests as $\beta_0 \to \pm\infty$ for fixed β_*. The approach is the same as in Andrews, Moreira, and Stock (2004) (AMS04) except that we consider $\beta_0 \to \pm\infty$. Also see Mills, Moreira, and Vilela (2014).

12.1 Point Optimal Invariant Similar Tests for Fixed $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_*}$

First, we consider the point null and alternative hypotheses:

$$H_0 : \beta = \beta_0 \text{ and } H_1 : \beta = \beta_*, \tag{12.1}$$

where $\pi \in R^k$ (or, equivalently, $\lambda \geq 0$) under H_0 and H_1.

Point optimal invariant similar (POIS) tests for any given null and alternative parameter values β_0 and β_*, respectively, and any given Ω are constructed in AMS04, Sec. 5. Surprisingly, the same test is found to be optimal for all values of π under H_1, i.e., for all strengths of identification. The optimal test is constructed by determining the level α test that maximizes conditional power given $Q_T = q_T$ among tests that are invariant and have null rejection probability α conditional on $Q_T = q_T$, for each $q_T \in R$.

By AMS04 (Comment 2 to Cor. 2), the POIS test of $H_0 : \beta = \beta_0$ versus $H_1 : \beta = \beta_*$, for any $\pi \in R^k$ (or $\lambda \geq 0$) under H_1, rejects H_0 for large values of

$$POIS(Q; \beta_0, \beta_*) := Q_S + 2\frac{d_{\beta_*}(\beta_0, \Omega)}{c_{\beta_*}(\beta_0, \Omega)} Q_{ST}. \tag{12.2}$$

The critical value for the $POIS(Q; \beta_0, \beta_*)$ test is a conditional critical value given $Q_T = q_T$, which we denote by $\kappa_{\beta_0}(q_T)$. The critical value $\kappa_{\beta_0}(q_T)$ is defined to satisfy

$$P_{Q_1|Q_T}(POIS(Q; \beta_0, \beta_*) > \kappa_{\beta_0}(q_T)|q_T) = \alpha \tag{12.3}$$

for all $q_T \geq 0$, where $P_{Q_1|Q_T}(\cdot|q_T)$ denotes probability under the conditional density $f_{Q_1|Q_T}(q_1|q_T)$ defined in (11.3). Although the density $f_{Q_1|Q_T}(q_1|q_T)$ does not depend on β_0, $\kappa_{\beta_0}(q_T)$ depends on β_0, as well as (β_*, Ω, k), because $POIS(Q; \beta_0, \beta_*)$ does.

Note that, although the same $POIS(Q; \beta_0, \beta_*)$ test is best for all strengths of identification, i.e., for all $\lambda = \mu_\pi' \mu_\pi > 0$, the power of this test depends on λ.

12.2 One-Sided Power Bound When $\boldsymbol{\beta_0} \to \pm\infty$

Now we consider the best one-sided invariant similar test as $\beta_0 \to \pm\infty$ keeping (β_*, Ω) fixed. Lemma 14.1 below implies that

$$\lim_{\beta_0 \to \pm\infty} \frac{d_{\beta_*}(\beta_0, \Omega)}{c_{\beta_*}(\beta_0, \Omega)} = \left(\mp \frac{\rho_{uv}}{\sigma_v (1 - \rho_{uv}^2)^{1/2}} \right) / \left(\mp 1/\sigma_v \right) = \frac{\rho_{uv}}{(1 - \rho_{uv}^2)^{1/2}}, \tag{12.4}$$

where ρ_{uv}, defined in (4.5), is the correlation between the structural and reduced-form errors u_i and v_{2i} under β_*. Hence, the limit as $\beta_0 \to \pm\infty$ of the $POIS(Q; \beta_0, \beta_*)$ test statistic in (12.2) is

$$POIS(Q; \infty, \rho_{uv}) := \lim_{\beta_0 \to \pm\infty} \left(Q_S + 2\frac{d_{\beta_*}(\beta_0, \Omega)}{c_{\beta_*}(\beta_0, \Omega)} Q_{ST} \right) = Q_S + 2\frac{\rho_{uv}}{(1-\rho_{uv}^2)^{1/2}} Q_{ST}. \qquad (12.5)$$

Notice that (i) this limit is the same for $\beta_0 \to +\infty$ and $\beta_0 \to -\infty$, (ii) the $POIS(Q; \infty, \rho_{uv})$ statistic depends on $(\beta_*, \Omega) = (\beta_*, \Omega(\beta_*, \Sigma_*))$ only through $\rho_{uv} := Corr(\Sigma_*)$, and (iii) when $\rho_{uv} = 0$, the $POIS(Q; \infty, \rho_{uv})$ statistic is the AR statistic (times k). Some intuition for result (iii) is that $EQ_{ST} = 0$ under the null and $\lim_{|\beta_0| \to \infty} EQ_{ST} = 0$ under any fixed alternative β_* when $\rho_{uv} = 0$ (see the discussion in Section 5.2). In consequence, Q_{ST} is not useful for distinguishing between H_0 and H_1 when $|\beta_0| \to \infty$ and $\rho_{uv} = 0$.

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

$$P_{Q_1|Q_T}(POIS(Q; \infty, \rho_{uv}) > \kappa_\infty(q_T)|q_T) = \alpha \qquad (12.6)$$

for all $q_T \geq 0$. The density $f_{Q_1|Q_T}(\cdot|q_T)$ of $P_{Q_1|Q_T}(\cdot|q_T)$ only depends on the number of IV's k, see (11.3). The critical value function $\kappa_\infty(\cdot)$ depends on ρ_{uv} and k.

Let $\phi_{\beta_0}(Q)$ denote a test of $H_0 : \beta = \beta_0$ versus $H_1 : \beta = \beta_*$ based on Q that rejects H_0 when $\phi_{\beta_0}(Q) = 1$. In most cases, a test depends on β_0 because the distribution of Q depends on β_0, see (2.3) and (3.1), and not because $\phi_{\beta_0}(\cdot)$ depends on β_0. For example, this is true of the AR, LM, and CLR tests in (3.3) and (3.4). However, we allow for dependence of $\phi_{\beta_0}(\cdot)$ on β_0 in the following result in order to cover all possible sequences of (non-randomized) tests of $H_0 : \beta = \beta_0$.

Theorem 12.1 *Let $\{\phi_{\beta_0}(Q) : \beta_0 \to \pm\infty\}$ be any sequence of invariant similar level α tests of $H_0 : \beta = \beta_0$ versus $H_1 : \beta = \beta_*$ when Q has density $f_Q(q; \beta, \beta_0, \lambda, \Omega)$ for some $\lambda \geq 0$ and Ω is fixed and known. For fixed true $(\beta_*, \lambda, \Omega)$, the $POIS(Q; \infty, \rho_{uv})$ test satisfies*

$$\limsup_{\beta_0 \to \pm\infty} P_{\beta_*, \beta_0, \lambda, \Omega}(\phi_{\beta_0}(Q) = 1) \leq P_{\rho_{uv}, \lambda_v}(POIS(Q; \infty, \rho_{uv}) > \kappa_\infty(Q_T)).$$

Comments. (i). Theorem 12.1 shows that the $POIS(Q; \infty, \rho_{uv})$ test provides an asymptotic power bound as $\beta_0 \to \pm\infty$ for any invariant similar test for any fixed $(\beta_*, \lambda, \Omega)$. This power bound is strictly less than one. The reason is that $\lim_{\beta_0 \to \pm\infty} |c_{\beta_*}(\beta_0, \Omega)| \nrightarrow \infty$. This is the same reason that the AR test does not have power that converges to one in this scenario, see Section 19. Hence, the bound in Theorem 12.1 is informative.

(ii). The power bound in Theorem 12.1 only depends on $(\beta_*, \lambda, \Omega)$ through ρ_{uv}, the magnitude of endogeneity under β_*, and λ_v, the concentration parameter.

(iii). As an alternative to the power bound given in Theorem 12.1, one might consider developing a formal limit of experiments result, e.g., along the lines of van der Vaart (1998, Ch. 9). This approach does not appear to work for the sequence of experiments consisting of the two unconditional distributions of $[S : T]$ (or Q) for $\beta = \beta_0, \beta_*$ and indexed by β_0 as $\beta_0 \to \pm\infty$. The reason is that the likelihood ratio of these two distributions is asymptotically degenerate as $\beta_0 \to \pm\infty$ (either 0 or ∞ depending on which density is in the numerator) when the truth is taken to be $\beta = \beta_0$. This occurs because the length of the mean vector of T diverges to infinity as $\beta_0 \to \pm\infty$ (provided $\lambda = \mu_\pi' \mu_\pi > 0$) by (2.3) and Lemma 14.1(c) below. For the sequence of conditional distributions of Q given $Q_T = q_T$, it should be possible to obtain a formal limit of experiments result, but this would not very helpful because we are interested in the unconditional power of tests and a conditional limit of experiments result would not deliver this.

(iv). The proof of Theorem 12.1 is given in Section 18 below.

13 Equations (4.1) and (4.2) of AMS

This section corrects (4.1) of AMS, which concerns the two-point weight function that defines AMS's two-sided AE power envelope.

Equation (4.1) of AMS is:[8] given (β_*, λ), the second point (β_{2*}, λ_2) solves

$$\lambda_2^{1/2} c_{\beta_{2*}} = -\lambda^{1/2} c_{\beta_*} \ (\neq 0) \text{ and } \lambda_2^{1/2} d_{\beta_{2*}} = \lambda^{1/2} d_{\beta_*}. \tag{13.1}$$

AMS states that provided $\beta_* \neq \beta_{AR}$, the solutions to the two equations in (4.1) satisfy the two equations in (4.2) of AMS, which is the same as (5.3) and which we repeat here for convenience:[9]

$$\begin{aligned} &\beta_{2*} = \beta_0 - \frac{d_{\beta_0}(\beta_* - \beta_0)}{d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)} \text{ and } \lambda_2 = \lambda \frac{(d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0))^2}{d_{\beta_0}^2}, \text{ where} \\ &r_{\beta_0} := e_1' \Omega^{-1} a_0 \cdot (a_0' \Omega^{-1} a_0)^{-1/2} \text{ and } e_1 := (1, 0)'. \end{aligned} \tag{13.2}$$

Equation (4.2) is correct as stated, but (4.1) of AMS is not correct. More specifically, it is not complete. It should be: given (β_*, λ), the second point (β_{2*}, λ_2) solves either (13.1) or

$$\lambda_2^{1/2} c_{\beta_{2*}} = \lambda^{1/2} c_{\beta_*} \ (\neq 0) \text{ and } \lambda_2^{1/2} d_{\beta_{2*}} = -\lambda^{1/2} d_{\beta_*}. \tag{13.3}$$

[8] Note that (β_*, λ) and (β_{2*}, λ_2) in this paper correspond to (β^*, λ^*) and (β_2^*, λ_2^*) in AMS.

[9] The formulae in (5.3) and (13.2) only hold for $\beta_* \neq \beta_{AR}$, where $\beta_{AR} := (\omega_1^2 - \omega_{12}\beta_0)/(\omega_{12} - \omega_2^2\beta_0)$ provided $\omega_{12} - \omega_2^2\beta_0 \neq 0$ (which necessarily holds for $|\beta_0|$ sufficiently large because $\omega_2^2 > 0$).

For brevity, we write the "either or" conditions in (13.1) and (13.3) as

$$\lambda_2^{1/2} c_{\beta_{2*}} = \mp \lambda^{1/2} c_{\beta_*} \ (\neq 0) \text{ and } \lambda_2^{1/2} d_{\beta_{2*}} = \pm \lambda^{1/2} d_{\beta_*}. \tag{13.4}$$

The reason (4.1) of AMS needs to be augmented by (13.3) is that for some (β_*, λ), β_0, and Ω, (4.1) has no real solutions (β_{2*}, λ_2) and the expressions for (β_{2*}, λ_2) in (4.2) of AMS do not satisfy (4.1). Once (4.1) of AMS is augmented by (13.3), there exist real solutions (β_{2*}, λ_2) to the augmented conditions and they are given by the expressions in (4.2) of AMS, i.e., by (13.2). This is established in the following lemma.

Lemma 13.1 *The conditions in* (13.4) *hold iff the conditions in* (4.2) *of AMS hold, i.e., iff the conditions in* (13.2) *holds.*

With (4.1) of AMS replaced by (13.4), the results in Theorem 8(b) and (c) of AMS hold as stated. That is, the two-point weight function that satisfies (13.4) leads to a two-sided weighted average power (WAP) test that is asymptotically efficient under strong IV's. And, all other two-point weight functions lead to two-sided WAP tests that are not asymptotically efficient under strong IV's.

Lemma 13.2 *Under the assumptions of* Theorem 8 *of* AMS*, i.e.,* Assumptions SIV-LA *and* 1-4 *of* AMS, (a) *if* (β_{2*}, λ_2) *satisfies* (13.4)*, then* $LR^*(\widehat{Q}_{1,n}, \widehat{Q}_{T,n}; \beta_*, \lambda) = e^{-\frac{1}{2}(\tau^*)^2} \cosh(\tau^* LM_n^{1/2}) + o_p(1)$*, where* $\tau^* = \lambda^{1/2} c_{\beta_*}$*, which is a strictly-increasing continuous function of* LM_n*, and* (b) *if* (β_{2*}, λ_2) *does not satisfy* (13.4)*, then* $LR^*(\widehat{Q}_{1,n}, \widehat{Q}_{T,n}; \beta_*, \lambda) = \eta_2(Q_{ST,n}/Q_{T,n}^{1/2}) + o_p(1)$ *for a continuous function* $\eta_2(\cdot)$ *that is not even.*

Comments. (i). Lemma 13.2(a) is an extension of Theorem 8(b) of AMS; while Lemma 13.2(b) is a correction to Theorem 8(c) of AMS.

(ii). The proofs of Lemma 13.1 and 13.2 are given in Section 18 below.

Having augmented (4.1) by (13.3), the two-point weight function of AMS does not have the property that β_{2*} is necessarily on the opposite side of β_0 from β_*. However, it does have the properties that (i) for any (β_*, λ), (β_{2*}, λ_2) is the only point that yields a two-point WAP test that is asymptotic efficient in a two-sided sense under strong IV's, (ii) the marginal distributions of Q_S, Q_T, and Q_{ST} are the same under (β_*, λ) and (β_{2*}, λ_2), and (iii) the joint distribution of (Q_S, Q_{ST}, Q_T) under (β_*, λ) is the same as that of $(Q_S, -Q_{ST}, Q_T)$ under (β_{2*}, λ_2).

14 Proof of Lemma 5.1

The proof of Lemma 5.1 and other proofs below use the following lemma.

The distributions of $[S:T]$ under (β_0,Ω) and (β_*,Ω) depend on $c_\beta(\beta_0,\Omega)$ and $d_\beta(\beta_0,\Omega)$ for $\beta=\beta_0$ and β_*. The limits of these quantities as $\beta_0\to\pm\infty$ are given in the following lemma.[10]

Lemma 14.1 *For fixed β_* and positive definite matrix Ω, we have*

(a) $\lim_{\beta_0\to\pm\infty} c_{\beta_0}(\beta_0,\Omega)=0$.

(b) $\lim_{\beta_0\to\pm\infty} c_{\beta_*}(\beta_0,\Omega)=\mp 1/\sigma_v$.

(c) $\lim_{\beta_0\to\pm\infty} d_{\beta_0}(\beta_0,\Omega)=\infty$.

(d) $d_{\beta_0}(\beta_0,\Omega)/|\beta_0|=\frac{\omega_2}{(\omega_1^2\omega_2^2-\omega_{12}^2)^{1/2}}+o(1)=\frac{1}{\sigma_u(1-\rho_{uv}^2)^{1/2}}+o(1)$ *as* $|\beta_0|\to\infty$.

(e) $\lim_{\beta_0\to\pm\infty} d_{\beta_*}(\beta_0,\Omega)=\pm\frac{\omega_2^2\beta_*-\omega_{12}}{\omega_2(\omega_1^2\omega_2^2-\omega_{12}^2)^{1/2}}=\mp\frac{\rho_{uv}}{\sigma_v(1-\rho_{uv}^2)^{1/2}}$.

Comment. The limits in parts (d) and (e), expressed in terms of Σ_*, only depend on ρ_{uv}, σ_u, and σ_v and their functional forms are of a relatively simple multiplicative form. The latter provides additional simplifications of certain quantities that appear below.

Proof of Lemma 14.1. Part (a) holds because $c_{\beta_0}(\beta_0,\Omega)=0$ for all β_0. Part (b) holds by the following calculations:

$$\begin{aligned}\lim_{\beta_0\to\pm\infty} c_{\beta_*}(\beta_0,\Omega) &= \lim_{\beta_0\to\pm\infty}(\beta_*-\beta_0)\cdot(b_0'\Omega b_0)^{-1/2}\\ &= \lim_{\beta_0\to\pm\infty}(\beta_*-\beta_0)\cdot(\omega_1^2-2\omega_{12}\beta_0+\omega_2^2\beta_0^2)^{-1/2}\\ &= \mp 1/\omega_2\\ &= \mp 1/\sigma_v. \end{aligned}\tag{14.1}$$

Now, we establish part (e). Let $b_*:=(1,-\beta_*)'$. We have

$$\begin{aligned}\lim_{\beta_0\to\pm\infty} d_{\beta_*}(\beta_0,\Omega) &= \lim_{\beta_0\to\pm\infty} b_*'\Omega b_0\cdot(b_0'\Omega b_0)^{-1/2}\det(\Omega)^{-1/2}\\ &= \lim_{\beta_0\to\pm\infty}\frac{\omega_1^2-\omega_{12}\beta_*-\omega_{12}\beta_0+\omega_2^2\beta_*\beta_0}{(\omega_1^2-2\omega_{12}\beta_0+\omega_2^2\beta_0^2)^{1/2}(\omega_1^2\omega_2^2-\omega_{12}^2)^{1/2}}\\ &= \pm\frac{\omega_2^2\beta_*-\omega_{12}}{\omega_2(\omega_1^2\omega_2^2-\omega_{12}^2)^{1/2}}.\end{aligned}\tag{14.2}$$

Next, we write the limit in (14.2) in terms of the elements of the structural error variance matrix

[10] Throughout, $\beta_0\to\pm\infty$ means $\beta_0\to\infty$ or $\beta_0\to-\infty$.

Σ_*. The term in the square root in the denominator of (14.2) satisfies

$$\omega_1^2\omega_2^2 - \omega_{12}^2 = (\sigma_u^2 + 2\sigma_{uv}\beta_* + \sigma_v^2\beta_*^2)\sigma_v^2 - (\sigma_{uv} + \sigma_v^2\beta_*)^2 = \sigma_u^2\sigma_v^2 - \sigma_{uv}^2, \tag{14.3}$$

where the first equality uses $\omega_2^2 = \sigma_v^2$ (since both denote the variance of v_{2i}), $\omega_1^2 = \sigma_u^2 + 2\sigma_{uv}\beta_* + \sigma_v^2\beta_*^2)$, and $\omega_{12} = \sigma_{uv} + \sigma_v^2\beta_*$ (which both hold by (11.8) with $\beta = \beta_*$ and $\Sigma = \Sigma_*$), and the second equality holds by simple calculations. The limit in (14.2) in terms of the elements of Σ_* is

$$\pm\frac{\omega_2^2\beta_* - \omega_{12}}{\omega_2(\omega_1^2\omega_2^2 - \omega_{12}^2)^{1/2}} = \pm\frac{\sigma_v^2\beta_* - (\sigma_{uv} + \sigma_v^2\beta_*)}{\sigma_v(\sigma_u^2\sigma_v^2 - \sigma_{uv}^2)^{1/2}} = \mp\frac{\rho_{uv}}{\sigma_v(1-\rho_{uv}^2)^{1/2}}, \tag{14.4}$$

where the first equality uses (14.3), $\omega_2^2 = \sigma_v^2$, and $\omega_{12} = \sigma_{uv} + \sigma_v^2\beta_*$, and the second inequality holds by dividing the numerator and denominator by $\sigma_u\sigma_v$. This establishes part (e).

For part (c), we have

$$\begin{aligned}
\lim_{\beta_0\to\pm\infty} d_{\beta_0}(\beta_0,\Omega) &= \lim_{\beta_0\to\pm\infty} (b_0'\Omega b_0)^{1/2}\det(\Omega)^{-1/2} \\
&= \lim_{\beta_0\to\pm\infty} \frac{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}}{(\omega_1^2\omega_2^2 - \omega_{12}^2)^{1/2}} \\
&= \infty.
\end{aligned}\tag{14.5}$$

Part (d) holds because, as $|\beta_0| \to \infty$, we have

$$\begin{aligned}
d_{\beta_0}(\beta_0,\Omega)/|\beta_0| &= \frac{(\omega_1^2/\beta_0^2 - 2\omega_{12}/\beta_0 + \omega_2^2)^{1/2}}{(\omega_1^2\omega_2^2 - \omega_{12}^2)^{1/2}} \\
&= \frac{\omega_2}{(\omega_1^2\omega_2^2 - \omega_{12}^2)^{1/2}} + o(1) \\
&= \frac{1}{\sigma_u(1-\rho_{uv}^2)^{1/2}} + o(1),
\end{aligned}\tag{14.6}$$

where the last equality uses (14.3) and $\omega_2 = \sigma_v$. $\square$

Next, we prove Lemma 5.1, which states that for any fixed $(\beta_*, \lambda, \Omega)$, $\lim_{\beta_0\to\pm\infty} f_Q(q; \beta_*, \beta_0, \lambda, \Omega)$ $= f_Q(q; \rho_{uv}, \lambda_v)$.

Proof of Lemma 5.1. By Lemma 14.1(b) and (e) and (5.1), we have $\lim_{\beta_0\to\pm\infty} c_{\beta_*} = \mp 1/\sigma_v$ and

$\lim_{\beta_0\to\pm\infty} d_{\beta_*} = \mp r_{uv}/\sigma_v$. In consequence,

$$\begin{aligned}
\lim_{\beta_0\to\pm\infty} \lambda(c_{\beta_*}^2 + d_{\beta_*}^2) &= \lambda(1/\sigma_v^2)(1+r_{uv}^2) = \lambda_v(1+r_{uv}^2) \text{ and} \\
\lim_{\beta_0\to\pm\infty} \lambda\xi_{\beta_*}(q) &= \lim_{\beta_0\to\pm\infty} \lambda(c_{\beta_*}^2 q_S + 2c_{\beta_*}d_{\beta_*}q_{ST} + d_{\beta_*}^2 q_T) \\
&= \lambda(1/\sigma_v^2)(q_S + 2r_{uv}q_{ST} + r_{uv}^2 q_T) = \lambda_v \xi(q;\rho_{uv}),
\end{aligned} \tag{14.7}$$

using the definitions of λ_v and $\xi(q;\rho_{uv})$ in (5.1) and (11.4), respectively, where the first equality in the third line uses $(\mp 1)(\mp r_{uv}) = r_{uv}$. Combining this with (11.2) and (11.4) proves the result of the lemma. □

15 Proof of Theorem 4.1

The proof of Theorem 4.1 uses the following lemma.[11] Let

$$\begin{aligned}
S_{\pm\infty}(Y) &:= (Z'Z)^{-1/2}Z'Ye_2 \cdot \frac{\mp 1}{\sigma_v}, \\
T_{\pm\infty}(Y) &:= (Z'Z)^{-1/2}Z'Y\Omega^{-1}e_1 \cdot (\pm(1-\rho_{uv}^2)^{1/2}\sigma_u), \text{ and} \\
Q_{\pm\infty}(Y) &:= \begin{bmatrix} e_2'Y'P_ZYe_2 \cdot \frac{1}{\sigma_v^2} & e_2'Y'P_ZY\Omega^{-1}e_1 \cdot \frac{\mp(1-\rho_{uv}^2)^{1/2}\sigma_u}{\sigma_v} \\ e_2'Y'P_ZY\Omega^{-1}e_1 \cdot \frac{\mp(1-\rho_{uv}^2)^{1/2}\sigma_u}{\sigma_v} & e_1'\Omega^{-1}Y'P_ZY\Omega^{-1}e_1 \cdot (1-\rho_{uv}^2)\sigma_u^2 \end{bmatrix},
\end{aligned} \tag{15.1}$$

where $\rho_{uv} := Corr(u_i, v_{2i})$, $P_Z := Z(Z'Z)^{-1}Z'$, $e_1 := (1,0)'$, and $e_2 := (0,1)'$. Let $Q_{T,\pm\infty}(Y)$ denote the $(2,2)$ element of $Q_{\pm\infty}(Y)$. As defined in (5.1), $r_{uv} = \rho_{uv}/(1-\rho_{uv}^2)^{1/2}$.

Lemma 15.1 *For fixed β_* and positive definite matrix Ω, we have*

(a) $\lim_{\beta_0\to\pm\infty} S_{\beta_0}(Y) = S_{\pm\infty}(Y)$,

(b) $S_{\pm\infty}(Y) \sim N(\mp\frac{1}{\sigma_v}\mu_\pi, I_k)$,

(c) $\lim_{\beta_0\to\pm\infty} T_{\beta_0}(Y) = T_{\pm\infty}(Y) = (Z'Z)^{-1/2}Z'Y\Omega^{-1}e_1\cdot(\pm(1-\rho_\Omega^2)^{1/2}\omega_1)$, *where* $\rho_\Omega := Corr(v_{1i}, v_{2i})$,

(d) $T_{\pm\infty}(Y) \sim N\left(\mp\frac{r_{uv}}{\sigma_v}\mu_\pi, I_k\right)$,

(e) $S_{\pm\infty}(Y)$ *and* $T_{\pm\infty}(Y)$ *are independent,*

(f) $\lim_{\beta_0\to\pm\infty} Q_{\beta_0}(Y) = Q_{\pm\infty}(Y)$, *and*

(g) $Q_{\pm\infty}(Y)$ *has a noncentral Wishart distribution with means matrix* $\mp\mu_\pi(\frac{1}{\sigma_v}, \frac{r_{uv}}{\sigma_v}) \in R^{k\times 2}$, *identity variance matrix, and density given in* (11.4).

Comment. The convergence results in Lemma 15.1 hold for all realizations of Y.

[11] The proof of Comment (v) to Theorem 4.1 is the same as that of Theorem 4.1(a) and (b) with $[S_{\beta_0}(Y), T_{\beta_0}(Y)]$ and $T_{\beta_0}(Y)$ in place of $Q_{\beta_0}(Y)$ and $Q_{T,\beta_0}(Y)$, respectively.

Proof of Theorem 4.1. First, we prove part (a). We have

$$
\begin{aligned}
& 1(RLength(CS_\phi(Y)) = \infty) \\
&= 1(\mathcal{T}(Q_{\beta_0}(Y)) \leq cv(Q_{T,\beta_0}(Y)) \ \forall \beta_0 \geq K(Y) \text{ for some } K(Y) < \infty) \\
&= \lim_{\beta_0 \to \infty} 1(\mathcal{T}(Q_{\beta_0}(Y)) \leq cv(Q_{T,\beta_0}(Y))), \qquad (15.2)
\end{aligned}
$$

where the second equality holds provided the limit as $\beta_0 \to \infty$ on the rhs exists, the first equality holds by the definition of $CS_\phi(Y)$ in (4.1)-(4.3) and the definition of $RLength(CS_\phi(Q)) = \infty$ in (4.4), and the second equality holds because its rhs equals one (when the rhs limit exists) iff $\mathcal{T}(Q_{\beta_0}(Y)) \leq cv(Q_{T,\beta_0}(Y))$ for $\forall \beta_0 \geq K(Y)$ for some $K(Y) < \infty$, which is the same as its lhs.

Now, we use the dominated convergence theorem to show

$$
\begin{aligned}
& \lim_{\beta_0 \to \infty} E_{\beta_*,\pi,\Omega} 1(\mathcal{T}(Q_{\beta_0}(Y)) \leq cv(Q_{T,\beta_0}(Y))) \\
&= E_{\beta_*,\pi,\Omega} \lim_{\beta_0 \to \infty} 1(\mathcal{T}(Q_{\beta_0}(Y)) \leq cv(Q_{T,\beta_0}(Y))). \qquad (15.3)
\end{aligned}
$$

The dominated convergence theorem applies because (i) the indicator functions in (15.3) are dominated by the constant function equal to one, which is integrable, and (ii) $\lim_{\beta_0 \to \infty} 1(\mathcal{T}(Q_{\beta_0}(Y)) \leq cv(Q_{T,\beta_0}(Y))$ exists a.s.$[P_{\beta_*,\pi,\Omega}]$ and equals $1(\mathcal{T}(Q_\infty(Y)) \leq cv(Q_{T,\infty}(Y))$ a.s.$[P_{\beta_*,\pi,\Omega}]$. The latter holds because the assumption that $\mathcal{T}(q)$ and $cv(q_T)$ are continuous at positive definite (pd) q and positive q_T, respectively, coupled with the result of Lemma 15.1(f) (that $Q_{\beta_0}(Y) \to Q_\infty(Y)$ as $\beta_0 \to \infty$ for all sample realizations of Y, where $Q_\infty(Y)$ is defined in (15.1)), imply that (a) $\lim_{\beta_0 \to \infty} \mathcal{T}(Q_{\beta_0}(Y)) = \mathcal{T}(Q_\infty(Y))$ for all realizations of Y for which $Q_\infty(Y)$ is pd, (b) $\lim_{\beta_0 \to \infty} cv(Q_{T,\beta_0}(Y)) = cv(Q_{T,\infty}(Y))$ for all realizations of Y with $Q_{T,\infty}(Y) > 0$, and hence (c) $\lim_{\beta_0 \to \infty} 1(\mathcal{T}(Q_{\beta_0}(Y)) \leq cv(Q_{T\beta_0}(Y)) = 1(\mathcal{T}(Q_\infty(Y)) \leq cv(Q_{T,\infty}(Y))$ for all realizations of Y for which $\mathcal{T}(Q_\infty(Y)) \neq cv(Q_{T,\infty}(Y))$. We have $P_{\beta_*,\pi,\Omega}(\mathcal{T}(Q_\infty(Y)) = cv(Q_{T,\infty}(Y))) = 0$ by assumption, and $P_{\beta_*,\pi,\Omega}(Q_\infty(Y)$ is pd & $Q_{T,\infty}(Y) > 0) = 1$ (because $Q_\infty(Y)$ has a noncentral Wishart distribution by Lemma 15.1(g)). Thus, condition (ii) above holds and the DCT applies.

Next, we have

$$
\begin{aligned}
& 1 - \lim_{\beta_0 \to \infty} P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1) \\
&= \lim_{\beta_0 \to \infty} E_{\beta_*,\pi,\Omega} 1(\mathcal{T}(Q_{\beta_0}(Y)) \leq cv(Q_{T,\beta_0}(Y))) \\
&= E_{\beta_*,\pi,\Omega} \lim_{\beta_0 \to \infty} 1(\mathcal{T}(Q_{\beta_0}(Y)) \leq cv(Q_{T,\beta_0}(Y))) \\
&= P_{\beta_*,\pi,\Omega}(RLength(CS_\phi(Y)) = \infty), \qquad (15.4)
\end{aligned}
$$

where the first equality holds because the distribution of Q under $P_{\beta_*,\beta_0,\lambda,\Omega}(\cdot)$ equals the distribution of $Q_{\beta_0}(Y)$ under $P_{\beta_*,\pi,\Omega}(\cdot)$ and $\phi(Q)=0$ iff $\mathcal{T}(Q_{\beta_0})\leq cv(Q_T)$ by (4.2), the second equality holds by (15.3), and the last equality holds by (15.2). Equation (15.4) establishes part (a).

The proof of part (b) is the same as that of part (a), but with $LLength$, $\forall\beta_0\leq -K(Y)$, $\beta_0\to-\infty$, $Q_{-\infty}(Y)$, and $Q_{T,-\infty}(Y)$ in place of $RLength$, $\forall\beta_0\geq K(Y)$, $\beta_0\to\infty$, $Q_{\infty}(Y)$, and $Q_{T,\infty}(Y)$, respectively.

The proof of part (c) uses the following: (i) $Q_\infty(Y)$ and $Q_{-\infty}(Y)$ only differ in the sign of their off-diagonal elements by (15.1), (ii) $\mathcal{T}(Q_\infty(Y))$ does not depend on the sign of the off-diagonal element of $Q_\infty(Y)$ by assumption, and hence, (iii) $1(\mathcal{T}(Q_\infty(Y))\leq cv(Q_{T,\infty}(Y))=1(\mathcal{T}(Q_{-\infty}(Y))$ $\leq cv(Q_{T,-\infty}(Y))$ for all sample realizations of Y. We have

$$
\begin{aligned}
&1(RLength(CS_\phi(Y))=\infty\ \&\ LLength(CS_\phi(Y))=\infty)\\
=\ &1(\mathcal{T}(Q_{\beta_0}(Y))\leq cv(Q_{T\beta_0}(Y))\ \forall\beta_0\geq K(Y)\ \&\ \forall\beta_0\leq -K(Y)\text{ for some }K(Y)<\infty)\\
=\ &\lim_{\beta_0\to\infty}1(\mathcal{T}(Q_{\beta_0}(Y))\leq cv(Q_{T,\beta_0}(Y))\ \&\ \mathcal{T}(Q_{-\beta_0}(Y))\leq cv(Q_{T,-\beta_0}(Y)))\\
=\ &1(\mathcal{T}(Q_\infty(Y))\leq cv(Q_{T,\infty}(Y))\ \&\ \mathcal{T}(Q_{-\infty}(Y))\leq cv(Q_{T,-\infty}(Y)))\\
=\ &1(\mathcal{T}(Q_\infty(Y))\leq cv(Q_{T,\infty}(Y))\\
=\ &\lim_{\beta_0\to\infty}1(\mathcal{T}(Q_{\beta_0}(Y))\leq cv(Q_{T,\beta_0}(Y))
\end{aligned}
\tag{15.5}
$$

where the first two equalities hold for the same reasons as the equalities in (15.2), the third equality holds a.s.$[P_{\beta_*,\pi,\Omega}]$ by result (ii) that follows (15.3) and the same result with $-\beta_0$ and $-\infty$ in place of β_0 and ∞, respectively, the second last equality holds by condition (iii) immediately above (15.5), and the last equality holds by result (ii) that follows (15.3).

Now, we have

$$
\begin{aligned}
&P_{\beta_*,\pi,\Omega}(RLength(CS_\phi(Y))=\infty\ \&\ LLength(CS_\phi(Y))=\infty)\\
=\ &E_{\beta_*,\pi,\Omega}\lim_{\beta_0\to\infty}1(\mathcal{T}(Q_{\beta_0}(Y))\leq cv(Q_{T,\beta_0}(Y)))\\
=\ &1-\lim_{\beta_0\to\infty}P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q)=1),
\end{aligned}
\tag{15.6}
$$

where the first equality holds by (15.5) and the second equality holds by the first four lines of (15.4). This establishes the equality in part (c) when $\beta_0\to\infty$. The equality in part (c) when $\beta_0\to-\infty$ holds because (15.5) and (15.6) hold with $\beta_0\to\infty$ replaced by $\beta_0\to-\infty$ since the indicator function on the rhs of the second equality in (15.5) depends on β_0 only through $|\beta_0|$. $\square$

Proof of Lemma 15.1. Part (a) holds because

$$\begin{aligned}
\lim_{\beta_0 \to \pm\infty} S_{\beta_0}(Y) &= \lim_{\beta_0 \to \pm\infty} (Z'Z)^{-1/2} Z'Y b_0 \cdot (b_0' \Omega b_0)^{-1/2} \\
&= (Z'Z)^{-1/2} Z'Y \lim_{\beta_0 \to \pm\infty} \begin{pmatrix} 1 \\ -\beta_0 \end{pmatrix} / (\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2 \beta_0^2)^{1/2} \\
&= (Z'Z)^{-1/2} Z'Y e_2 (\mp 1/\sigma_v), \qquad (15.7)
\end{aligned}$$

where $e_2 := (0,1)'$, the first equality holds by (2.3), the second equality holds because $b_0 := (1, -\beta_0)'$, and the third equality holds using $\omega_2 = \sigma_v$.

Next, we prove part (b). The statistic $S_{\pm\infty}(Y)$ has a multivariate normal distribution because it is a linear combination of multivariate normal random variables. The mean of $S_{\pm\infty}(Y)$ is

$$ES_{\pm\infty}(Y) = (Z'Z)^{-1/2} Z'Z[\pi\beta_* : \pi] e_2 \cdot \frac{\mp 1}{\sigma_v} = (Z'Z)^{1/2} \pi \cdot \frac{\mp 1}{\sigma_v} = \mu_\pi \cdot \frac{\mp 1}{\sigma_v}, \qquad (15.8)$$

where the first equality holds using (2.2) with $a = (\beta_*, 1)'$ and (15.1). The variance matrix of $S_{\pm\infty}(Y)$ is

$$\begin{aligned}
& Var(S_{\pm\infty}(Y)) = Var((Z'Z)^{-1/2} Z'Y e_2)/\sigma_v^2 = Var\left(\sum_{i=1}^n (Z'Z)^{-1/2} Z_i Y_i' e_2 \right) / \sigma_v^2 \\
&= \sum_{i=1}^n Var((Z'Z)^{-1/2} Z_i Y_i' e_2)/\sigma_v^2 = \sum_{i=1}^n (Z'Z)^{-1/2} Z_i Z_i (Z'Z)^{-1/2} e_2' \Omega e_2 / \sigma_v^2 = I_k, \qquad (15.9)
\end{aligned}$$

where the third equality holds by independence across i and the last equality uses $\omega_2^2 = \sigma_v^2$. This establishes part (b).

To prove part (c), we have

$$\begin{aligned}
\lim_{\beta_0 \to \pm\infty} T_{\beta_0}(Y) &= \lim_{\beta_0 \to \pm\infty} (Z'Z)^{-1/2} Z'Y \Omega^{-1} a_0 \cdot (a_0' \Omega^{-1} a_0)^{-1/2} \\
&= (Z'Z)^{-1/2} Z'Y \Omega^{-1} \lim_{\beta_0 \to \pm\infty} \begin{pmatrix} \beta_0 \\ 1 \end{pmatrix} / (\omega^{11}\beta_0^2 + 2\omega^{12}\beta_0 + \omega^{22})^{1/2} \\
&= (Z'Z)^{-1/2} Z'Y \Omega^{-1} e_1 \cdot (\pm 1/\omega^{11})^{1/2} \\
&= (Z'Z)^{-1/2} Z'Y \Omega^{-1} e_1 \cdot (\pm(\omega_1^2\omega_2^2 - \omega_{12}^2)^{1/2}/\omega_2), \qquad (15.10)
\end{aligned}$$

where ω^{11}, ω^{12}, and ω^{22} denote the $(1,1)$, $(1,2)$, and $(2,2)$ elements of Ω^{-1}, respectively, $e_1 := (1,0)'$, the first equality holds by (2.3), the second equality holds because $a_0 := (\beta_0, 1)'$, and the fourth equality holds by the formula for ω^{11}. In addition, we have

$$(\omega_1^2\omega_2^2 - \omega_{12}^2)^{1/2}/\omega_2 = (1-\rho_\Omega^2)^{1/2}\omega_1 = (1-\rho_{uv}^2)^{1/2}\sigma_u, \tag{15.11}$$

where the first equality uses $\rho_\Omega := \omega_{12}/(\omega_1\omega_2)$ and the second equality holds because $\omega_1^2\omega_2^2 - \omega_{12}^2 = \sigma_u^2\sigma_v^2 - \sigma_{uv}^2$ by (14.3) and $\omega_2 = \sigma_v$. Equations (15.10) and (15.11), combined with (15.1), establish part (c).

Now, we prove part (d). Like $S_{\pm\infty}(Y)$, $T_{\pm\infty}(Y)$ has a multivariate normal distribution. The mean of $T_{\pm\infty}(Y)$ is

$$\begin{aligned} ET_{\pm\infty}(Y) &= (Z'Z)^{-1/2}Z'Z[\pi\beta_* : \pi]\Omega^{-1}e_1 \cdot (\pm(1-\rho_{uv}^2)^{1/2}\sigma_u) \\ &= (Z'Z)^{1/2}\pi(\beta_*\omega^{11} + \omega^{12}) \cdot (\pm(1-\rho_{uv}^2)^{1/2}\sigma_u), \end{aligned} \tag{15.12}$$

where the equality holds using (2.2) with $a = (\beta_*, 1)'$ and (15.1). In addition, we have

$$\beta_*\omega^{11} + \omega^{12} = \frac{\beta_*\omega_2^2 - \omega_{12}}{\omega_1^2\omega_2^2 - \omega_{12}^2} = \frac{-\sigma_{uv}}{\sigma_u^2\sigma_v^2 - \sigma_{uv}^2} = \frac{-\rho_{uv}}{(1-\rho_{uv}^2)\sigma_u\sigma_v}, \tag{15.13}$$

where the second equality uses $\omega_1^2\omega_2^2 - \omega_{12}^2 = \sigma_u^2\sigma_v^2 - \sigma_{uv}^2$ by (14.3) and $\beta_*\omega_2^2 - \omega_{12} = -\sigma_{uv}$ by (11.9) with $\beta = \beta_*$. Combining (15.12) and (15.13) gives

$$ET_{\pm\infty}(Y) = \mu_\pi \cdot \frac{\mp\rho_{uv}}{\sigma_v(1-\rho_{uv}^2)^{1/2}} = \mu_\pi \cdot \frac{\mp r_{uv}}{\sigma_v}. \tag{15.14}$$

The variance matrix of $T_{\pm\infty}(Y)$ is

$$\begin{aligned} &Var(T_{\pm\infty}(Y)) = Var((Z'Z)^{-1/2}Z'Y\Omega^{-1}e_1) \cdot (1-\rho_{uv}^2)\sigma_u^2 \\ &= Var\left(\sum_{i=1}^n (Z'Z)^{-1/2}Z_iY_i'\Omega^{-1}e_1\right) \cdot (1-\rho_{uv}^2)\sigma_u^2 = \sum_{i=1}^n Var((Z'Z)^{-1/2}Z_iY_i'\Omega^{-1}e_1) \cdot (1-\rho_{uv}^2)\sigma_u^2 \\ &= \sum_{i=1}^n (Z'Z)^{-1/2}Z_iZ_i(Z'Z)^{-1/2}e_1'\Omega^{-1}e_1 \cdot (1-\rho_{uv}^2)\sigma_u^2 = I_k\frac{\omega_2^2}{\omega_1^2\omega_2^2 - \omega_{12}^2} \cdot (1-\rho_{uv}^2)\sigma_u^2 \\ &= I_k\frac{\sigma_v^2}{\sigma_u^2\sigma_v^2 - \sigma_{uv}^2} \cdot (1-\rho_{uv}^2)\sigma_u^2 = I_k, \end{aligned}$$

where the first equality holds by (15.1), the third equality holds by independence across i, and the second last equality uses $\omega_1^2\omega_2^2 - \omega_{12}^2 = \sigma_u^2\sigma_v^2 - \sigma_{uv}^2$ by (14.3) and $\omega_2^2 = \sigma_v^2$.

Part (e) holds because

$$\begin{aligned} Cov(S_{\pm\infty}(Y), T_{\pm\infty}(Y)) &= -\sum_{i=1}^{n} Cov((Z'Z)^{-1/2} Z_i Y_i' e_2, (Z'Z)^{-1/2} Z_i Y_i' \Omega^{-1} e_1) \cdot (1-\rho_{uv}^2)^{1/2} \sigma_u/\sigma_v \\ &= \sum_{i=1}^{n} (Z'Z)^{-1/2} Z_i Z_i (Z'Z)^{-1/2} e_2' \Omega \Omega^{-1} e_1 \cdot (1-\rho_{uv}^2)^{1/2} \sigma_u/\sigma_v = 0^k. \quad (15.15) \end{aligned}$$

Part (f) follows from parts (a) and (c) of the lemma and (4.1).

Part (g) holds by the definition of the noncentral Wishart distribution and parts (b), (d), and (e) of the lemma. The density of $Q_{\pm\infty}(Y)$ equals the density in (11.4) because the noncentral Wishart density is invariant to a sign change in the means matrix. $\square$

16 Proofs of Theorems 5.2, Corollary 5.3, and Theorem 5.4

The following lemma is used in the proof of Theorem 5.2. As above, 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 $f_Q(q;\rho_{uv},\lambda_v)$, which are defined in Section 11.1. See (11.2) and (11.4) for explicit expressions for these noncentral Wishart densities.

Lemma 16.1 (a) $\lim_{\beta_0\to\pm\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(POIS2(Q;\beta_*,\beta_0,\lambda) > \kappa_{2,\beta_0}(Q_T)) = P_{\rho_{uv},\lambda_v}(POIS2(Q;\infty, |\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(Q_T))$,

(b) $\lim_{\beta_0\to\pm\infty} P_{\beta_{2*},\beta_0,\lambda_2,\Omega}(POIS2(Q;\beta_*,\beta_0,\lambda) > \kappa_{2,\beta_0}(Q_T)) = P_{-\rho_{uv},\lambda_v}(POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(Q_T))$,

(c) $P_{\rho_{uv},\lambda_v}(POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(Q_T)) = P_{-\rho_{uv},\lambda_v}(POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(Q_T))$,

(d) $\lim_{\beta_0\to\pm\infty} \beta_{2*} = -\beta_* + 2\frac{\omega_{12}}{\omega_2^2} = \beta_* + 2\frac{\sigma_u\rho_{uv}}{\sigma_v}$, *and*

(e) $\lim_{\beta_0\to\pm\infty} \lambda_2 = \lambda$.

The reason that Q has the density $f_Q(q;-\rho_{uv},\lambda_v)$ (defined in (11.4)) in the limit expression in Lemma 16.1(b) can be seen clearly from the following lemma.

Lemma 16.2 *For any fixed* (β_*,λ,Ω), $\lim_{\beta_0\to\pm\infty} f_Q(q;\beta_{2*},\beta_0,\lambda_2,\Omega) = f_Q(q;-\rho_{uv},\lambda_v)$ *for all* 2×2 *variance matrices* q, *where* β_{2*} *and* λ_2 *satisfy* (5.3) *and* ρ_{uv} *and* λ_v *are defined in* (4.5) *and* (5.1), *respectively.*

Proof of Lemma 16.2. Given (β^*, λ^*), suppose the second point (β_2^*, λ_2^*) solves (13.1). In this case, by Lemma 14.1(b) and (e), we have

$$\begin{aligned}\lim_{\beta_0\to\pm\infty} \lambda_2^{1/2} c_{\beta_{2*}}(\beta_0,\Omega) &= \lim_{\beta_0\to\pm\infty} -\lambda^{1/2} c_{\beta_*}(\beta_0,\Omega) = \pm\lambda^{1/2}/\sigma_v = \pm\lambda_v^{1/2} \text{ and} \\ \lim_{\beta_0\to\pm\infty} \lambda_2^{1/2} d_{\beta_{2*}}(\beta_0,\Omega) &= \lim_{\beta_0\to\pm\infty} \lambda^{1/2} d_{\beta_*}(\beta_0,\Omega) = \mp\lambda^{1/2}\frac{\rho_{uv}}{\sigma_v(1-\rho_{uv}^2)^{1/2}} = \mp\lambda_v^{1/2} r_{uv}.\end{aligned} \tag{16.1}$$

Using (11.1), (11.4), and (16.1), we obtain

$$\begin{aligned}\lim_{\beta_0\to\pm\infty} \lambda_2(c_{\beta_{2*}}^2 + d_{\beta_{2*}}^2) &= \lambda_v(1+r_{uv}^2) \text{ and} \\ \lim_{\beta_0\to\pm\infty} \lambda_2 \xi_{\beta_{2*}}(q) &:= \lim_{\beta_0\to\pm\infty} \lambda_2(c_{\beta_{2*}}^2 q_S + 2c_{\beta_{2*}} d_{\beta_{2*}} q_{ST} + d_{\beta_{2*}}^2 q_T) \\ &= \lambda_v(q_S - 2r_{uv}q_{ST} + r_{uv}^2 q_T) \\ &=: \lambda_v \xi(q; -\rho_{uv}),\end{aligned} \tag{16.2}$$

On the other hand, given (β^*, λ^*), suppose the second point (β_2^*, λ_2^*) solves (13.3). In this case, the minus sign on the rhs side of the first equality on the first line of (16.1) disappears, the quantity on the rhs side of the last equality on the first line of (16.1) becomes $\mp\lambda_v^{1/2}$, a minus sign is added to the rhs side of the first equality on the second line of (16.1), and the quantity on the rhs side of the last equality on the second line of (16.1) becomes $\pm\lambda_v^{1/2} r_{uv}$. These changes leave $\lambda_2 c_{\beta_{2*}}^2$, $\lambda_2 d_{\beta_{2*}}^2$, and $\lambda_2 c_{\beta_{2*}} d_{\beta_{2*}}$ unchanged from the case where (β_2^*, λ_2^*) solves (13.1). Hence, (16.2) also holds when (β_2^*, λ_2^*) solves (13.3).

Combining (16.2) with (11.2) (with (β_{2*}, λ_2) in place of (β_*, λ)) and (11.4) proves the result of the lemma. $\square$

Proof of Theorem 5.2. By Theorem 3 of AMS, for all $(\beta_*, \beta_0, \lambda, \Omega)$,

$$\begin{aligned}&P_{\beta_*,\beta_0,\lambda,\Omega}(\phi_{\beta_0}(Q)=1) + P_{\beta_{2*},\beta_0,\lambda_2,\Omega}(\phi_{\beta_0}(Q)=1) \\ \leq\ &P_{\beta_*,\beta_0,\lambda,\Omega}(POIS2(Q;\beta_0,\beta_*,\lambda) > \kappa_{2,\beta_0}(Q_T)) + P_{\beta_{2*},\beta_0,\lambda_2,\Omega}(POIS2(Q;\beta_0,\beta_*,\lambda) > \kappa_{2,\beta_0}(Q_T)).\end{aligned} \tag{16.3}$$

That is, the test on the rhs maximizes the two-point average power for testing $\beta = \beta_0$ against (β_*, λ) and (β_{2*}, λ_2) for fixed known Ω.

Equation (16.3) and Lemma 16.1(a)-(c) establish the result of Theorem 5.2 by taking the $\limsup_{\beta_0\to\pm\infty}$ of the lhs and the $\liminf_{\beta_0\to\pm\infty}$ of the rhs. $\square$

The proof of Comment (iv) to Theorem 5.2 is the same as that of Theorem 5.2, but in place of (16.3) it uses the inequality in Theorem 1 of CHJ i.e., $\int P_{\beta_*,\beta_0,\lambda,\mu_\pi/||\mu_\pi||,\Omega}(\phi_{\beta_0}(Q) =$

$1)dUnif(\mu_\pi/||\mu_\pi||) \leq \int P_{\beta_*,\beta_0,\lambda,\mu_\pi/||\mu_\pi||,\Omega}(POIS2(Q;\beta_0,\beta_*,\lambda) > \kappa_{2,\beta_0}(Q_T))dUnif(\mu_\pi/||\mu_\pi||)$, plus the fact that the rhs expression equals $P_{\beta_*,\beta_0,\lambda,\Omega}(POIS2(Q;\beta_0,\beta_*,\lambda) > \kappa_{2,\beta_0}(Q_T))$ because the distribution of Q only depends on μ_π through $\lambda = \mu_\pi'\mu_\pi$.

Proof of Lemma 16.1. To prove part (a), we write

$$\begin{aligned}
&P_{\beta_*,\beta_0,\lambda,\Omega}(POIS2(Q;\beta_0,\beta_*,\lambda) > \kappa_{2,\beta_0}(Q_T)) \\
&= \int\int 1(POIS2(q;\beta_0,\beta_*,\lambda) > \kappa_{2,\beta_0}(q_T))\phi_k(s - c_{\beta_*}\mu_\pi)\phi_k(t - d_{\beta_*}\mu_\pi)dsdt, \text{ and} \\
&P_{\rho_{uv},\lambda_v}(POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(Q_T)) \qquad (16.4) \\
&= \int\int 1(POIS2(q;\infty,|\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(q_T))\phi_k(s - (\mp 1/\sigma_v)\mu_\pi)\phi_k(t - (\mp r_{uv}/\sigma_v)\mu_\pi)dsdt,
\end{aligned}$$

where $\phi_k(x)$ for $x \in R^k$ denotes the density of k i.i.d. standard normal random variables, $\lambda = \mu_\pi'\mu_\pi$, $s,t \in R^k$, $q = [s:t]'[s:t]$, $q_T = t't$, $c_{\beta_*} = c_{\beta_*}(\beta_0,\Omega)$, $d_{\beta_*} = d_{\beta_*}(\beta_0,\Omega)$, the $\mp$ signs in the last line are both $+$ or both $-$, and the integral in the last line is the same whether both $\mp$ signs are $+$ or $-$ (by a change of variables calculation).

We have

$$\lim_{\beta_0\to\pm\infty} \phi_k(s - c_{\beta_*}(\beta_0,\Omega)\mu_\pi)\phi_k(t - d_{\beta_*}(\beta_0,\Omega)\mu_\pi) = \phi_k(s - (\mp 1/\sigma_v)\mu_\pi)\phi_k(t - (\mp r_{uv}/\sigma_v)\mu_\pi) \quad (16.5)$$

for all $s,t \in R^k$, by Lemma 14.1(b) and (e) and the smoothness of the standard normal density function. By (5.4) and (11.5) and Lemma 14.1(b) and (e), we have

$$\lim_{\beta_0\to\pm\infty} POIS2(q;\beta_0,\beta_*,\lambda) = POIS2(q;\infty,|\rho_{uv}|,\lambda_v) \quad (16.6)$$

for all for 2×2 variance matrices q, for given (β_*,λ,Ω). In addition, we show below that $\lim_{\beta_0\to\pm\infty}$ $\kappa_{2,\beta_0}(q_T) = \kappa_{2,\infty}(q_T)$ for all $q_T \geq 0$. Combining these results gives the following convergence result:

$$\begin{aligned}
&\lim_{\beta_0\to\pm\infty} 1(POIS2(q;\beta_0,\beta_*,\lambda) > \kappa_{2,\beta_0}(q_T)) \cdot \phi_k(s - c_{\beta_*}(\beta_0,\Omega)\mu_\pi)\phi_k(t - d_{\beta_*}(\beta_0,\Omega)\mu_\pi) \\
&= 1(POIS2(q;\infty,|\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(q_T)) \cdot \phi_k(s - (\mp 1/\sigma_v)\mu_\pi)\phi_k(t - (\mp r_{uv}/\sigma_v)\mu_\pi) \qquad (16.7)
\end{aligned}$$

for all $[s:t]$ for which $POIS2(q;\infty,|\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(q_T)$ or $POIS2(q;\infty,|\rho_{uv}|,\lambda_v) < \kappa_{2,\infty}(q_T)$, where $[s:t]$, q and (q_S,q_{ST},q_T) are functionally related by $q = [s:t]'[s:t]$ and the definitions in (11.2).

Given Lebesgue measure on the set of points $(s',t')' \in R^{2k}$, the induced measure on (q_S,q_{ST},q_T) $= (s's,s't,t't) \in R^3$ is absolutely continuous with respect to (wrt) Lebesgue measure on R^3 with positive density only for positive definite q. (This follows from change of variables calculations. These calculations are analogous to those used to show that if $[S:T]$ has the multivariate normal density $\phi_k(s-(\mp 1/\sigma_v)\mu_\pi)\phi_k(t-(\mp r_{uv}/\sigma_v)\mu_\pi)$, then Q has the density $f_Q(q;\rho_{uv},\lambda_v)$, which, viewed as a function of (q_S, q_{ST}, q_T), is a density wrt Lebesgue measure on R^3 that is positive only for positive definite q.) The Lebesgue measure of the set of (q_S, q_{ST}, q_T) for which $POIS2(q;\infty,|\rho_{uv}|,\lambda_v) = \kappa_{2,\infty}(q_T)$ is zero. (This holds because (i) the definition of $POIS2(q;\infty,|\rho_{uv}|,\lambda_v)$ in (5.4) implies that the Lebesgue measure of the set of (q_S, q_{ST}) for which $POIS2(q;\infty,|\rho_{uv}|,\lambda_v) = \kappa_{2,\infty}(q_T)$ is zero for all $q_T \geq 0$ and (ii) the Lebesgue measure of the set of (q_S, q_{ST}, q_T) for which $POIS2(q;\infty,|\rho_{uv}|,\lambda_v)$ $= \kappa_{2,\infty}(q_T)$ is obtained by integrating the set in (i) over $q_T \in R$ subject to the constraint that q is positive definite.) In turn, this implies that the Lebesgue measure of the set of $(s', t')'$ for which $POIS2(q;\infty,|\rho_{uv}|,\lambda_v) = \kappa_{2,\infty}(q_T)$ is zero. Hence, (16.7) verifies the a.s. (wrt Lebesgue measure on R^{2k}) convergence condition required for the application of the DCT to obtain part (a) using (16.4).

Next, to verify the dominating function requirement of the DCT, we need to show that

$$\sup_{\beta_0 \in R} |\phi_k(s - c_{\beta_*}(\beta_0,\Omega)\mu_\pi)\phi_k(t - d_{\beta_*}(\beta_0,\Omega)\mu_\pi)| \tag{16.8}$$

is integrable wrt Lebesgue measure on R^{2k} (since the indicator functions in (16.7) are bounded by one). For any $0 < c < \infty$ and $m \in R$, we have

$$\begin{aligned} \int \sup_{|m|\leq c} \exp\left(-(x-m)^2/2\right) dx &= 2\int_0^\infty \sup_{|m|\leq c} \exp\left(-x^2/2 + mx - m^2/2\right) dx \\ \leq 2\int_0^\infty \exp\left(-x^2/2 + cx\right) dx &= 2\int_0^\infty \exp\left(-(x-c)^2/2 + c^2/2\right) dx < \infty, \end{aligned} \tag{16.9}$$

where the first equality holds by symmetry. This result yields the integrability of the dominating function in (16.8) because $\phi_k(\cdot)$ is a product of univariate standard normal densities and $\sup_{\beta_0 \in R} |c_{\beta_*}(\beta_0,\Omega)| < \infty$ and $\sup_{\beta_0 \in R} |d_{\beta_*}(\beta_0,\Omega)| < \infty$ are finite by Lemma 14.1(b) and (e) and continuity of $c_{\beta_*}(\beta_0,\Omega)$ and $d_{\beta_*}(\beta_0,\Omega)$ in β_0.

Hence, the DCT applies and it yields part (a).

It remains to show $\lim_{\beta_0 \to \pm\infty} \kappa_{2,\beta_0}(q_T) = \kappa_{2,\infty}(q_T)$ for all $q_T \geq 0$. As noted above, $\lim_{\beta_0 \to \pm\infty}$ $POIS2(q;\beta_0,\beta_*,\lambda) = POIS(q;\infty,|\rho_{uv}|,\lambda_v)$ for all 2×2 variance matrices q. Hence, $1(POIS2(Q;\beta_0,\beta_*,\lambda) \leq x) \to 1(POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) \leq x)$ as $\beta_0 \to \pm\infty$ for all $x \in R$ for which $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) \neq x$. We have $P_{Q_1|Q_T}(POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) = x|q_T) = 0$ for all $q_T \geq 0$ by the absolute continuity of $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v)$ under $P_{Q_1|Q_T}(\cdot|q_T)$ (by the functional form of $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v)$ and the absolute continuity of Q_1 under $P_{Q_1|Q_T}(\cdot|q_T)$, whose density is given in (11.3)). Thus, by the DCT, for all $x \in R$,

$$\lim_{\beta_0 \to \pm\infty} P_{Q_1|Q_T}(POIS2(Q;\beta_0,\beta_*,\lambda) \leq x|q_T) = P_{Q_1|Q_T}(POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) \leq x|q_T) \text{ and}$$
$$POIS2(Q;\beta_0,\beta_*,\lambda) \to_d POIS2(Q;\infty,|\rho_{uv}|.\lambda_v) \text{ as } \beta_0 \to \pm\infty \text{ under } P_{Q_1|Q_T}(\cdot|q_T). \quad (16.10)$$

The second line of (16.10), coupled with the fact that $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v)$ has a strictly increasing distribution function at its $1-\alpha$ quantile under $P_{Q_1|Q_T}(\cdot|q_T)$ for all $q_T \geq 0$ (which is shown below), implies that the $1-\alpha$ quantile of $POIS2(Q;\beta_0,\beta_*,\lambda)$ under $P_{Q_1|Q_T}(\cdot|q_T)$ (i.e., $\kappa_{2,\beta_0}(q_T)$) converges as $\beta_0 \to \pm\infty$ to the $1-\alpha$ quantile of $POIS2(Q;\beta_0,\beta_*,\lambda)$ under $P_{Q_1|Q_T}(\cdot|q_T)$ (i.e., $\kappa_{2,\infty}(q_T)$). This can be proved by contradiction. First, suppose $\delta := \limsup_{j\to\infty}\kappa_{2,j}(q_T) - \kappa_{2,\infty}(q_T) > 0$ (where each $j \in R$ represents some value of β_0 here). Then, there exists a subsequence $\{m_j : j \geq 1\}$ of $\{j : j \geq 1\}$ such that $\delta = \lim_{j\to\infty}\kappa_{2,m_j}(q_T) - \kappa_{2,\infty}(q_T)$. We have

$$\begin{aligned}
\alpha &= \lim_{j\to\infty} P_{Q_1|Q_T}(POIS2(Q;m_j,\beta_*,\lambda) > \kappa_{2,m_j}(q_T)|q_T) \\
&\leq \lim_{j\to\infty} P_{Q_1|Q_T}(POIS2(Q;m_j,\beta_*,\lambda) > \kappa_{2,\infty}(q_T) + \delta/2|q_T) \\
&= P_{Q_1|Q_T}(POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(q_T) + \delta/2|q_T) \\
&< P_{Q_1|Q_T}(POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) > \kappa_{2,\infty}(q_T)|q_T) \\
&= \alpha, \qquad (16.11)
\end{aligned}$$

where the first equality holds by the definition of $\kappa_{2,\beta_0}(q_T)$, the first inequality holds by the expression above for δ, the second equality holds by the first line of (16.10) with $x = \kappa_{2,\infty}(q_T) + \delta/2$, the second inequality holds because $\delta > 0$ and the distribution function of $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v)$ is strictly increasing at its $1-\alpha$ quantile $\kappa_{2,\infty}(q_T)$ under $P_{Q_1|Q_T}(\cdot|q_T)$ for all $q_T \geq 0$, and the last equality holds by the definition of $\kappa_{2,\infty}(q_T)$. Equation (16.11) is a contradiction, so $\delta \leq 0$. An analogous argument shows that $\liminf_{\beta_0\to\infty}\kappa_{2,\beta_0}(q_T) - \kappa_{2,\infty}(q_T) < 0$ does not hold. Hence, $\lim_{\beta_0\to\infty}\kappa_{2,\beta_0}(q_T) = \kappa_{2,\infty}(q_T)$. An analogous argument shows that $\liminf_{\beta_0\to-\infty}\kappa_{2,\beta_0}(q_T) = \kappa_{2,\infty}(q_T)$.

It remains to show that the distribution function of $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v)$ is strictly increasing at its $1-\alpha$ quantile $\kappa_{2,\infty}(q_T)$ under $P_{Q_1|Q_T}(\cdot|q_T)$ for all $q_T \geq 0$. This holds because (i) $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v)$ is a nonrandom strictly increasing function of $(\xi(Q;\rho_{uv}),\xi(Q;-\rho_{uv}))$ conditional on $T = t$ (specifically, $POIS2(Q;\infty,|\rho_{uv}|,\lambda_v) = C_{q_T}\sum_{j=0}^{\infty}[(\lambda_v\xi(Q;\rho_{uv}))^j + (\lambda_v\xi(Q;-\rho_{uv}))^j]$ $/(4^j j!\Gamma(\nu+j+1))$, where C_{q_T} is a constant that may depend on q_T, $\nu := (k-2)/2$, and $\Gamma(\cdot)$ is the gamma function, by (5.4) and (4.8) of AMS, which provides an expression for the modified Bessel function of the first kind $I_\nu(x)$), (ii) $\xi(Q;\rho_{uv}) = (S + r_{uv}T)'(S + r_{uv}T)$ and

$\xi(Q; -\rho_{uv}) = (S - r_{uv}T)'(S - r_{uv}T)$ have the same noncentral χ^2_k distribution conditional on $T = t$ (because $[S : T]$ has a multivariate normal distribution with means matrix given by (5.2) and identity variance matrix), (iii) $(\xi(Q; \rho_{uv}), \xi(Q; -\rho_{uv}))$ has a positive density on R^2_+ conditional on $T = t$ and also conditional on $Q_T = q_T$ (because the latter conditional density is the integral of the former conditional density over t such that $t't = q_T$), and hence, (iv) $POIS2(Q; \infty, |\rho_{uv}|, \lambda_v)$ has a positive density on R_+ conditional on q_T for all $q_T \geq 0$. This completes the proof of part (a).

The proof of part (b) is the same as that of part (a), but with (i) $-c_{\beta_*}$ and $\pm 1/\sigma_v$ in place of c_{β_*} and $\mp 1/\sigma_v$, respectively, in (16.4), (16.5), and (16.7), and (ii) π_2 in place of π, where

$$\begin{aligned} \pi_2 &:= Me_{1,k},\ e_{1,k} := (1, 0, ...,)' \in R^k,\ M := \frac{\lambda^{1/2} g(\beta_0, \beta_*, \Omega)}{(e_{1,k}' Z' Z e_{1,k})^{1/2}}, \\ g(\beta_0, \beta_*, \Omega) &:= \frac{d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)}{d_{\beta_0}},\ \text{and } \lambda_2 := \mu_{\pi_2}' \mu_{\pi_2}. \end{aligned} \tag{16.12}$$

As defined, λ_2 satisfies (5.3) because

$$\lambda_2 := \mu_{\pi_2}' \mu_{\pi_2} = \pi_2' Z' Z \pi_2 = M^2 e_{1,k}' Z' Z e_{1,k} = \lambda g^2(\beta_0, \beta_*, \Omega). \tag{16.13}$$

In addition, $\lambda_2 \to \lambda$ as $\beta_0 \to \pm\infty$ by (16.17) below. With the above changes, the proof of part (a) establishes part (b).

Part (c) holds because the test statistic $POIS2(Q; \infty, |\rho_{uv}|, \lambda_v)$ and critical value $\kappa_{2,\infty}(Q_T)$ only depend on ρ_{uv} and q_{ST} through $|\rho_{uv}|$ and $|q_{ST}|$, respectively, and the density $f_Q(q; \rho_{uv}, \lambda_v)$ of Q only depends on the sign of ρ_{uv} through $r_{uv} q_{ST}$. In consequence, a change of variables from (q_S, q_{ST}, q_T) to $(q_S, -q_{ST}, q_T)$ establishes the result of part (c).

To prove part (d), we have

$$\begin{aligned} d_{\beta_0} &= (a_0' \Omega^{-1} a_0)^{1/2} = \frac{\omega_2^2 \beta_0^2 - 2\omega_{12}\beta_0 + \omega_1^2}{\omega_1^2 \omega_2^2 - \omega_{12}^2} (a_0' \Omega^{-1} a_0)^{-1/2} \text{ and} \\ r_{\beta_0} &= e_1' \Omega^{-1} a_0 (a_0' \Omega^{-1} a_0)^{-1/2} = \frac{\omega_2^2 \beta_0 - \omega_{12}}{\omega_1^2 \omega_2^2 - \omega_{12}^2} (a_0' \Omega^{-1} a_0)^{-1/2}, \end{aligned} \tag{16.14}$$

where the first equalities on lines one and two hold by (2.7) of AMS and (5.3), respectively. Next, we have

$$\begin{aligned}
\beta_{2*} &= \beta_0 - \frac{d_{\beta_0}(\beta_* - \beta_0)}{d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)} \\
&= \frac{d_{\beta_0}(2\beta_0 - \beta_*) + 2r_{\beta_0}(\beta_* - \beta_0)\beta_0}{d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)} \\
&= \frac{(\omega_2^2\beta_0^2 - 2\omega_{12}\beta_0 + \omega_1^2)(2\beta_0 - \beta_*) + 2(\omega_2^2\beta_0 - \omega_{12})(\beta_*\beta_0 - \beta_0^2)}{(\omega_2^2\beta_0^2 - 2\omega_{12}\beta_0 + \omega_1^2) + 2(\omega_2^2\beta_0 - \omega_{12})(\beta_* - \beta_0)} \\
&= \frac{\beta_0^2(-\omega_2^2\beta_* - 4\omega_{12} + 2\omega_2^2\beta_* + 2\omega_{12}) + O(\beta_0)}{\beta_0^2(\omega_2^2 - 2\omega_2^2) + O(\beta_0)} \\
&= \frac{(\omega_2^2\beta_* - 2\omega_{12}) + o(1)}{-\omega_2^2 + o(1)} \\
&= -\beta_* + \frac{2\omega_{12}}{\omega_2^2} + o(1),
\end{aligned} \tag{16.15}$$

where the third equality uses (16.14) and the two terms involving β_0^3 in the numerator of the rhs of the third equality cancel. Next, we have

$$-\beta_* + \frac{2\omega_{12}}{\omega_2^2} = \frac{2(\omega_{12} - \omega_2^2\beta_*) + \omega_2^2\beta_*}{\omega_2^2} = \frac{2\sigma_{uv} + \sigma_v^2\beta_*}{\sigma_v^2} = \beta_* + 2\frac{\sigma_{uv}}{\sigma_v^2} = \beta_* + 2\frac{\sigma_u\rho_{uv}}{\sigma_v}, \tag{16.16}$$

where the second equality uses (11.9) with $\beta = \beta_*$ and $\omega_2^2 = \sigma_v^2$.

Next, we prove part (e). We have

$$\begin{aligned}
\left(\frac{\lambda_2}{\lambda}\right)^{1/2} &= \left|\frac{d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)}{d_{\beta_0}}\right| \\
&= \left|\frac{\omega_2^2\beta_0^2 - 2\omega_{12}\beta_0 + \omega_1^2 + 2(\omega_2^2\beta_0 - \omega_{12})(\beta_* - \beta_0)}{\omega_2^2\beta_0^2 - 2\omega_{12}\beta_0 + \omega_1^2}\right| \\
&= \left|\frac{\beta_0^2(\omega_2^2 - 2\omega_2^2) + \beta_0(-2\omega_{12} + 2\omega_2^2\beta_* + 2\omega_{12}) + \omega_1^2 - 2\omega_{12}\beta_*}{\omega_2^2\beta_0^2 - 2\omega_{12}\beta_0 + \omega_1^2}\right| \\
&= 1 + o(1),
\end{aligned} \tag{16.17}$$

where the first equality holds by (5.3) and the second equality uses (16.14). $\square$

Proof of Corollary 5.3. We have

$$\begin{aligned}
&(P_{\beta_*,\lambda,\Omega}(RLength(CS_\phi(Y)) = \infty) + P_{\beta_{2*},\lambda_2,\Omega}(RLength(CS_\phi(Y)) = \infty))/2 \\
&= 1 - \lim_{\beta_0\to\infty}[P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1) + \lim_{\beta_0\to\infty} P_{\beta_{2*},\beta_0,\lambda_2,\Omega}(\phi(Q) = 1)]/2 \\
&\geq P_{\rho_{uv},\lambda_v}(POIS2(Q;\infty, |\rho_{uv}|, \lambda_v) > \kappa_{2,\infty}(Q_T)),
\end{aligned} \tag{16.18}$$

where the equality holds by Theorem 4.1(a) with (β_*, λ) and $(\beta_{2*}, \lambda_{2*})$, $P_{\beta_*,\lambda,\Omega}(\cdot)$ is equivalent to $P_{\beta_*,\pi,\Omega}(\cdot)$, which appears in Theorem 4.1(a) (because events determined by $CS_\phi(Y)$ only depend on π through λ, since $CS_\phi(Y)$ is based on rotation-invariant tests), and the inequality holds by Theorem 5.2(a). This establishes the first result of part (a).

The second result of part (a) holds by the same calculations as in (16.18), but with *LLength* and $\beta_0 \to -\infty$ in place of *RLength* and $\beta_0 \to \infty$, respectively, using Theorem 4.1(b) in place of Theorem 4.1(a).

Part (b) holds by combining Theorem 4.1(c) and Theorem 5.2 because, as noted in Comment (iii) to Theorem 5.2, the lim sup on the left-hand side in Theorem 5.2 is the average of two equal quantities. $\square$

Next, we prove Comment (ii) to Corollary 5.3. The proof is the same as that of Corollary 5.3, but using

$$\int P_{\beta_*,\lambda,\mu_\pi/||\mu_\pi||,\Omega}(RLength(CS_\phi(Y)) = \infty) dUnif(\mu_\pi/||\mu_\pi||) = 1 - \lim_{\beta_0\to\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(\phi(Q) = 1) \tag{16.19}$$

and likewise with (β_{2*}, λ) in place of (β_*, θ) in place of the first equality in (16.18). The proof of (16.19) is the same as the proof of Theorem 4.1(a) but with $Q_{\beta_0}(Y)$ and $Q_{T,\beta_0}(Y)$ replaced by $[S_{\beta_0}(Y), T_{\beta_0}(Y)]$, and $T_{\beta_0}(Y)$, respectively, throughout the proof, with $E_{\beta_*,\pi,\Omega}(\cdot)$ replaced by $\int E_{\beta_*,\lambda,\mu_\pi/||\mu_\pi||,\Omega}(\cdot) dUniform(\mu_\pi/||\mu_\pi||)$ in (15.3), and using Lemma 15.1(a) and (c) in place of Lemma 15.1(f) when verifying the limit property (ii) needed for the dominated convergence theorem following (15.3).

Proof of Theorem 5.4. The proof is quite similar to, but much simpler than, the proof of part (a) of Lemma 16.1 with $POIS2(q; \beta_0, \beta_*, \lambda) > \kappa_{2,\beta_0}(q_T)$ in (16.4) replaced by $q_S > \chi^2_{k,1-\alpha}/k$ for the AR test, $q^2_{ST}/q_T > \chi^2_{1,1-\alpha}$ for the LM test, and $q_S - q_T + ((q_S - q_T)^2 + 4q^2_{ST})^{1/2} > 2\kappa_{LR,\alpha}(q_T)$ for the CLR test. The proof is much simpler because for the latter three tests neither the test statistics nor the critical values depend on β_0. The parameter β_0, for which the limit as $\beta_0 \to \pm\infty$ is being considered, only enters through the multivariate normal densities in (16.4). The limits of these densities and an integrable dominating function for them have already been provided in the proof of Lemma 16.1(a). The indicator function that appears in (16.7) is bounded by one regardless of which test appears in the indicator function. In addition, $P_{\beta_*,\rho_{uv},\lambda_v}(AR = \chi^2_{k,1-\alpha}) = 0$ and $P_{\beta_*,\rho_{uv},\lambda_v}(LM = \chi^2_{1,1-\alpha}) = 0$ because the AR statistic has a noncentral χ^2_k distribution with noncentrality parameter λ_v under $P_{\beta_*,\rho_{uv},\lambda_v}$ (since $S \sim N(\mu_\pi/\sigma_v, I_k)$ by Lemma 5.1 and (5.2)) and the conditional distribution of the LM statistic given T under $P_{\beta_*,\rho_{uv},\lambda_v}$ is a noncentral χ^2

distribution.

Next, we show $P_{\beta_*,\rho_{uv},\lambda_v}(LR = \kappa_{LR,\alpha}(Q_T)) = 0$. Let $J = AR - LM$. Then, $2LR = J + LM - Q_T + ((J + LM - Q_T)^2 + 4LM \cdot Q_T)^{1/2}$. We can write $Q = [S : T]'[S : T]$, where $[S : T]$ has a multivariate normal distribution with means matrix given by (5.2) and identity variance matrix. As shown below, conditional on $T = t$, LM and J have independent noncentral χ^2 distributions with 1 and $k-1$ degrees of freedom, respectively. This implies that (i) the distribution of LR conditional on $T = t$ is absolutely continuous, (ii) $P_{\beta_*,\rho_{uv},\lambda_v}(LR = \kappa_{LR,\alpha}(Q_T)|T = t) = 0$ for all $t \in R^k$, and (iii) $P_{\beta_*,\rho_{uv},\lambda_v}(LR = \kappa_{LR,\alpha}(Q_T)) = 0$. It remains to show that conditional on $Q_T = q_T$, LM and J have independent noncentral χ^2 distributions. We can write $LM = S'P_TS$ and $J = S'(I_k - P_T)S$, where $P_T := T(T'T)^{-1}T'$ and S has a multivariate normal with identity variance matrix. This implies that P_TS and $(I_k - P_T)S$ are independent conditional on $T = t$ and LM and J have independent noncentral χ^2 distributions conditional on $T = t$ for all $t \in R^k$. This completes the proof. $\square$

17 Proof of Theorem 7.1

The proof of Theorem 7.1(a) uses the following lemma.

Lemma 17.1 *Suppose* $b_{1x} = 1 + \delta_x/x$ *and* $b_{2x} = 1 - \delta_x/x$, *where* $\delta_x \to \delta_\infty \neq 0$ *as* $x \to \infty$, $K_{j1x} = (b_{jx}x)^\eta$ *for some* $\eta \in R$ *for* $j = 1, 2$, *and* $K_{j2x} \to K_\infty \in (0, \infty)$ *as* $x \to \infty$ *for* $j = 1, 2$. *Then,* (a) *as* $x \to \infty$,

$$\begin{aligned}
&\log\left(K_{11x}K_{12x}e^{b_{1x}x} + K_{21x}K_{22x}e^{b_{2x}x}\right) - x - \eta\log x - \log K_\infty \\
&\to \delta_\infty + \log\left(1 + e^{-2\delta_\infty}\right) \text{ and}
\end{aligned}$$

(b) *the function* $s(y) := y + \log\left(1 + e^{-2y}\right)$ *for* $y \in R$ *is infinitely differentiable, symmetric about zero, strictly increasing for* $y > 0$, *and hence, strictly increasing in* $|y|$ *for* $|y| > 0$.

Proof of Lemma 17.1. Part (a) holds by the following:

$$\begin{aligned}
&\log\left(K_{11x}K_{12x}e^{b_{1x}x} + K_{21x}K_{22x}e^{b_{2x}x}\right) - x - \eta\log x - \log K_\infty \\
&= \log\left(K_{11x}K_{12x}e^{b_{1x}x}\left(1 + \frac{K_{21x}K_{22x}}{K_{11x}K_{12x}}e^{(b_{2x}-b_{1x})x}\right)\right) - x - \eta\log x - \log K_\infty \\
&= b_{1x}x + \log K_{11x} + \log(K_{12x}/K_\infty) + \log\left(1 + \frac{K_{21x}K_{22x}}{K_{11x}K_{12x}}e^{(b_{2x}-b_{1x})x}\right) - x - \eta\log x \\
&= \delta_x + \eta\log(b_{1x}) + \log(K_{12x}/K_\infty) + \log\left(1 + \frac{K_{21x}K_{22x}}{K_{11x}K_{12x}}e^{-2\delta_x}\right) \\
&\to \delta_\infty + \log\left(1 + e^{-2\delta_\infty}\right), \qquad (17.1)
\end{aligned}$$

where the third equality uses $b_{1x}x - x = \delta_x$, $\log K_{11x} = \eta \log(b_{1x}x) = \eta \log(b_{1x}) + \eta \log(x)$, and $b_{2x} - b_{1x} = -2\delta_x/x$, and the convergence uses $\log(b_{1x}) = \log(1 + o(1)) \to 0$, $K_{12x}/K_\infty \to 1$, $K_{21x}/K_{11x} = (b_{2x}/b_{1x})^\eta = 1 + o(1)$, and $K_{22x}/K_{12x} \to 1$.

The function $s(y)$ is infinitely differentiable because $\log(x)$ and e^{-2y} are. The function $s(y)$ is symmetric about zero because

$$\begin{aligned} & y + \log\left(1 + e^{-2y}\right) = -y + \log\left(1 + e^{2y}\right) \\ \Leftrightarrow\ & 2y = \log\left(1 + e^{2y}\right) - \log\left(1 + e^{-2y}\right) = \log\left(\frac{1 + e^{2y}}{1 + e^{-2y}}\right) = \log(e^{2y}) = 2y. \end{aligned} \tag{17.2}$$

The function $s(y)$ is strictly increasing for $y > 0$ because

$$\frac{d}{dy}s(y) = 1 - \frac{2e^{-2y}}{1 + e^{-2y}} = \frac{1 - e^{-2y}}{1 + e^{-2y}} = \frac{e^{2y} - 1}{e^{2y} + 1}, \tag{17.3}$$

which is positive for $y > 0$. We have $s(y) = s(|y|)$ because $s(y)$ is symmetric about zero, and $(d/d|y|)s(|y|) > 0$ for $|y| > 0$ by (17.3). Hence, $s(y)$ is strictly increasing in $|y|$ for $|y| > 0$. $\square$

Proof of Theorem 7.1. Without loss in generality, we prove the results for the case where $sgn(d_{\beta_*})$ is the same for all terms in the sequence as $\lambda d_{\beta_*}^2 \to \infty$. Given (2.3), without loss of generality, we can suppose that

$$S = c_{\beta_*}\mu_\pi + Z_S \text{ and } T = d_{\beta_*}\mu_\pi + Z_T, \tag{17.4}$$

where Z_S and Z_T are independent $N(0^k, I_k)$ random vectors.

We prove part (c) first. The distribution of Q depends on μ_π only through λ. In consequence, without loss of generality, we can assume that $\Upsilon := \mu_\pi/\lambda^{1/2} \in R^k$ does not vary as $\lambda d_{\beta_*}^2$ and $\lambda^{1/2}c_{\beta_*}$ vary. The following establishes the a.s. convergence of the one-sided LM test statistic: as

$\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty$,

$$\begin{aligned}
\frac{Q_{ST}}{Q_T^{1/2}} &= \frac{(c_{\beta_*}\mu_\pi + Z_S)'(d_{\beta_*}\mu_\pi + Z_T)}{((d_{\beta_*}\mu_\pi + Z_T)'(d_{\beta_*}\mu_\pi + Z_T))^{1/2}} \\
&= \frac{(c_{\beta_*}\mu_\pi + Z_S)'(d_{\beta_*}\mu_\pi + Z_T)}{(d^2_{\beta_*}\lambda)^{1/2}(1 + o_{a.s.}(1))} \\
&= \frac{(c_{\beta_*}\mu_\pi/\lambda^{1/2} + Z_S/\lambda^{1/2})'(sgn(d_{\beta_*})\mu_\pi + O_{a.s.}(1/|d_{\beta_*}|)}{(1 + o_{a.s.}(1))} \\
&= \left(sgn(d_{\beta_*})\Upsilon' Z_S + sgn(d_{\beta_*})\lambda^{1/2} c_{\beta_*} + O_{a.s.}\left(\frac{(\lambda c^2_{\beta_*})^{1/2}}{(\lambda d^2_{\beta_*})^{1/2}}\right) + O_{a.s.}\left(\frac{1}{(\lambda d^2_{\beta_*})^{1/2}}\right)\right) \\
&\quad \times (1 + o_{a.s.}(1)) \\
&\to_{a.s.} sgn(d_{\beta_*})\Upsilon' Z_S + sgn(d_{\beta_*}) c_\infty \\
&= : LM_{1\infty} \sim N(sgn(d_{\beta_*}) c_\infty, 1),
\end{aligned} \tag{17.5}$$

where the first equality holds by (3.1) and (17.4), the second equality holds using $d_{\beta_*}\mu_\pi + Z_T = (\lambda d^2_{\beta_*})^{1/2}(d_{\beta_*}\mu_\pi/(\lambda d^2_{\beta_*})^{1/2} + o_{a.s.}(1))$ since $\lambda d^2_{\beta_*} \to \infty$, the convergence holds because $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty$, and the limit random variable $LM_{1\infty}$ has a $N(sgn(d_{\beta_*})c_\infty, 1)$ distribution because $sgn(d_{\beta_*})\Upsilon' Z_S \sim N(0,1)$ (since $Z_S \sim N(0^k, I_k)$ and $||\Upsilon|| = 1$).

The a.s. convergence in (17.5) implies convergence in distribution by the dominated convergence theorem applied to $1(Q_{ST}/Q_T^{1/2} \leq y)$ for any fixed $y \in R$. In consequence, we have

$$P(LM > \chi^2_{1,1-\alpha}) = P((Q_{ST}/Q_T^{1/2})^2 > \chi^2_{1,1-\alpha}) \to P(LM^2_{1\infty} > \chi^2_{1,1-\alpha}) = P(\chi^2_1(c^2_\infty) > \chi^2_{1,1-\alpha}) \tag{17.6}$$

as $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty$, which establishes part (c).

To prove Theorem 7.1(a), we apply Lemma 17.1 to a realization of the random vectors Z_S and Z_T with

$$\begin{aligned}
x &:= (\lambda d^2_{\beta_*} Q_T)^{1/2}, \\
b_{1x} x &:= (\lambda \xi_{\beta_*}(Q; \beta_0, \Omega))^{1/2} := \lambda^{1/2}(c^2_{\beta_*} Q_S + 2 c_{\beta_*} d_{\beta_*} Q_{ST} + d^2_{\beta_*} Q_T)^{1/2}, \\
b_{2x} x &:= \lambda^{1/2}(c^2_{\beta_*} Q_S - 2 c_{\beta_*} d_{\beta_*} Q_{ST} + d^2_{\beta_*} Q_T)^{1/2}, \\
K_{11x} &:= (b_{1x} x)^{-(k-1)/2}, \\
K_{12x} &:= \frac{(b_{1x} x)^{1/2} I_{(k-2)/2}(b_{1x} x)}{e^{b_{1x} x}} \\
K_{21x} &:= (b_{2x} x)^{-(k-1)/2}, \text{ and} \\
K_{22x} &:= \frac{(b_{2x} x)^{1/2} I_{(k-2)/2}(b_{2x} x)}{e^{b_{2x} x}},
\end{aligned} \tag{17.7}$$

Thus, we take $\eta := -(k-1)/2$.

We have

$$Q_T = (d_{\beta_*}\mu_\pi + Z_T)'(d_{\beta_*}\mu_\pi + Z_T) = \lambda d_{\beta_*}^2(1 + o_{a.s.}(1)). \tag{17.8}$$

This implies that $x = (\lambda d_{\beta_*}^2)(1 + o_{a.s.}(1))$. Thus, $x \to \infty$ a.s. since $\lambda d_{\beta_*}^2 \to \infty$ by assumption.

The conditions $\lambda d_{\beta_*}^2 \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty \in R$ imply that $b_{1x}x \to \infty$ and $b_{2x}x \to \infty$ as $x \to \infty$. In consequence, by the properties of the modified Bessel function of the first kind, $I_{(k-2)/2}(x)$ for x large, e.g., see Lebedev (1965, p. 136),

$$\lim_{b_{1x}x \to \infty} K_{12x} = 1/(2\pi)^{1/2} \text{ and } \lim_{b_{2x}x \to \infty} K_{22x} = 1/(2\pi)^{1/2}. \tag{17.9}$$

Hence, the assumptions of Lemma 17.1 on K_{j2x} for $j = 1, 2$ hold with $K_\infty = 1/(2\pi)^{1/2}$.

Next, we have

$$\begin{aligned}
b_{1x} &= (\lambda c_{\beta_*}^2 Q_S + 2\lambda c_{\beta_*} d_{\beta_*} Q_{ST} + \lambda d_{\beta_*}^2 Q_T)^{1/2}/x \\
&= \left(1 + \frac{2\lambda c_{\beta_*} d_{\beta_*} Q_{ST}}{(\lambda d_{\beta_*}^2 Q_T)^{1/2} x} + \frac{\lambda c_{\beta_*}^2 Q_S}{x^2}\right)^{1/2} \\
&= \left(1 + \frac{2\lambda^{1/2} c_{\beta_*} sgn(d_{\beta_*})}{x}\frac{Q_{ST}}{Q_T^{1/2}} + \frac{\lambda c_{\beta_*}^2 Q_S}{x^2}\right)^{1/2} \\
&= 1 + (1 + o_{a.s.}(1))^{-1/2}\left(\frac{2\lambda^{1/2} c_{\beta_*} sgn(d_{\beta_*})}{x}\frac{Q_{ST}}{Q_T^{1/2}} + \frac{\lambda c_{\beta_*}^2 Q_S}{x^2}\right),
\end{aligned} \tag{17.10}$$

where the fourth equality holds by the mean value theorem because $\lambda^{1/2} c_{\beta_*} = O(1)$, $x \to \infty$ a.s., and $Q_{ST}/Q_T^{1/2} = O(1)$ a.s. (by (17.5)) imply that the term in parentheses on the last line of (17.10) is $o_{a.s.}(1)$.

From (17.10), we have

$$\begin{aligned}
\delta_x &= (1 + o_{a.s.}(1))^{-1/2}\left(2\lambda^{1/2} c_{\beta_*} sgn(d_{\beta_*})\frac{Q_{ST}}{Q_T^{1/2}} + \frac{\lambda c_{\beta_*}^2 Q_S}{x}\right) \\
&\to 2c_\infty sgn(d_{\beta_*}) LM_{1\infty} =: \delta_\infty \text{ a.s.}
\end{aligned} \tag{17.11}$$

using (17.5). This verifies the convergence condition of Lemma 17.1 on δ_x with $\delta_\infty \neq 0$ a.s. (by the absolute continuity of Z_S). Hence, Lemma 17.1 applies with x, $b_{1x}, \ldots$ as in (17.7).

Let ξ_{β_*} abbreviate $\xi_{\beta_*}(Q;\beta_0,\Omega) = c^2_{\beta_*}Q_S + 2c_{\beta_*}d_{\beta_*}Q_{ST} + d^2_{\beta_*}Q_T$. Let $\xi_{\beta_{2*}} = c^2_{\beta_*}Q_S - 2c_{\beta_*}d_{\beta_*}Q_{ST} + d^2_{\beta_*}Q_T$. So, $b_{1x}x = (\lambda\xi_{\beta_*})^{1/2}$ and $b_{2x}x = (\lambda\xi_{\beta_{2*}})^{1/2}$. Let

$$\begin{aligned}\tau(\beta_*,\lambda,Q_T) &:= -(\lambda d^2_{\beta_*}Q_T)^{1/2} + \frac{k-1}{2}\log((\lambda d^2_{\beta_*}Q_T)^{1/2}) - \log K_\infty \\ &= -x - \eta\log x - \log K_\infty, \end{aligned} \tag{17.12}$$

where the equality holds using the definitions in (17.7) and $K_\infty = 1/(2\pi)^{1/2}$ by (17.9).

Given the definitions of $POIS2(Q;\beta_0,\beta_*,\lambda)$ and $x, b_{1x},, ...$ in (11.5) and (17.7), respectively, Lemma 17.1(a) gives

$$\begin{aligned}&\log(POIS2(Q;\beta_0,\beta_*,\lambda)) + \log(2\psi_2(Q_T;\beta_0,\beta_*,\lambda)) + \tau(\beta_*,\lambda,Q_T) \\ =\ &\log\left((\lambda\xi_{\beta_*})^{-(k-2)/4}I_{(k-2)/2}((\lambda\xi_{\beta_*})^{1/2}) + (\lambda\xi_{\beta_{2*}})^{-(k-2)/4}I_{(k-2)/2}((\lambda\xi_{\beta_{2*}})^{1/2})\right) + \tau(\beta_*,\lambda,Q_T) \\ =\ &\log\Bigg((\lambda\xi_{\beta_*})^{-(k-1)/4}\frac{(\lambda\xi_{\beta_*})^{1/4}I_{(k-2)/2}((\lambda\xi_{\beta_*})^{1/2})}{e^{(\lambda\xi_{\beta_*})^{1/2}}}e^{(\lambda\xi_{\beta_*})^{1/2}} \\ &\quad + (\lambda\xi_{\beta_{2*}})^{-(k-1)/4}\frac{(\lambda\xi_{\beta_{2*}})^{1/4}I_{(k-2)/2}((\lambda\xi_{\beta_{2*}})^{1/2})}{e^{(\lambda\xi_{\beta_{2*}})^{1/2}}}e^{(\lambda\xi_{\beta_{2*}})^{1/2}}\Bigg) + \tau(\beta_*,\lambda,Q_T) \\ =\ &\log\left(K_{11x}K_{12x}e^{b_{1x}x} + K_{21x}K_{22x}e^{b_{2x}x}\right) - x - \eta\log x - \log K_\infty \\ \to_{a.s.}\ &\delta_\infty + \log\left(1 + e^{-2\delta_\infty}\right) \\ =\ &s(\delta_\infty) \\ =\ &s(2c_\infty|LM_{1\infty}|), \end{aligned} \tag{17.13}$$

where $\psi_2(Q_T;\beta_0,\beta_*,\lambda)$ is defined in (11.5), $LM^2_{1\infty} \sim \chi^2_1(c^2_\infty)$ is defined in (17.5), the first equality holds by the definition of $POIS2(Q;\beta_0,\beta_*,\lambda)$ in (11.5), the third equality uses the definitions in (17.7) and (17.12), the convergence holds by Lemma 17.1(a), the second last equality holds by the definition of $s(y)$ in Lemma 17.1(b), and the last equality holds because $\delta_\infty := 2c_\infty sgn(d_{\beta_*})LM_{1\infty}$, see (17.11), and $s(y)$ is symmetric around zero by Lemma 17.1(b).

Equation (17.13) and the dominated convergence theorem (applied to $1(\log(POIS2(Q;\beta_0,\beta_*,\lambda)) + \log(2\psi_2(Q_T;\beta_0,\beta_*,\lambda)) + \tau(\beta_*,\lambda,Q_T) \leq w)$ for any $w \in R$) give

$$\log(POIS2(Q;\beta_0,\beta_*,\lambda)) + \log(2\psi_2(Q_T;\beta_0,\beta_*,\lambda)) + \tau(\beta_*,\lambda,Q_T) \to_d s(\delta_\infty) = s(2c_\infty|LM_{1\infty}|). \tag{17.14}$$

Now we consider the behavior of the critical value function for the POIS2 test, $\kappa_{2,\beta_0}(q_T)$, where q_T denotes a realization of Q_T. We are interested in the power of the POIS2 test. So, we are interested in the behavior of $\kappa_{2,\beta_0}(q_T)$ for q_T sequences as $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2}c_{\beta_*} \to c_\infty$

that are generated when the true parameters are (β_*, λ). This behavior is given in (17.8) to be $q_T = \lambda d^2_{\beta_*}(1 + o(1))$ a.s. under (β_*, λ).

Up to this point in the proof, the parameters (β_*, λ) have played a duel role. First, they denote the parameter values against which the POIS2 test is designed to have optimal two-sided power and, hence, determine the form of the POIS2 test statistic. Second, they denote the true values of β and λ (because we are interested in the power of the POIS2 test when the (β_*, λ) values for which it is designed are the true values). Here, where we discuss the behavior of the critical value function $\kappa_{2,\beta_0}(\cdot)$, (β_*, λ) only play the former role. The true value of β is β_0 and the true value of λ we denote by λ_0. The function $\kappa_{2,\beta_0}(\cdot)$ depends on (β_*, λ) because the POIS2 test statistic does, but the null distribution that determines $\kappa_{2,\beta_0}(\cdot)$ does not depend on (β_*, λ). In spite of this, the values q_T which are of interest to us, do depend on (β_*, λ) as noted in the previous paragraph.

The function $\kappa_{2,\beta_0}(\cdot)$ is defined in (11.6). Its definition depends on the conditional null distribution of Q_1 given $Q_T = q_T$ whose density $f_{Q_1|Q_T}(\cdot|q_T)$ is given in (11.3). This density depends on k, but not on any other parameters, such as β_0, $\lambda_0 = \mu'_{\pi 0}\mu_{\pi_0}$, or Ω. In consequence, for the purposes of determining the properties of $\kappa_{2,\beta_0}(\cdot)$ we can suppose that $\beta_0 = 0$, $\mu_{\pi_0} = 1^k/||1^k||$, $\lambda_0 = 1$, and $\Omega = I_2$. In this case,

$$S = Z_S \sim N(0^k, I_k),\ T = \mu_{\pi_0} + Z_T \sim N(\mu_{\pi_0}, I_k), \tag{17.15}$$

and S and T are independent (using $d_{\beta_0}(\beta_0, \Omega) = b_0'\Omega b_0(b_0'\Omega b_0)^{-1/2}\det(\Omega)^{-1/2} = 1$ since $b_0 = (1, \beta_0)' = (1, 0)'$).

We now show that $\kappa_{2,\beta_0}(q_T)$ satisfies

$$\log(\kappa_{2,\beta_0}(q_T)) + \log(2\psi_2(q_T; \beta_0, \beta_*, \lambda)) + \tau(\beta_*, \lambda, q_T) \to s(2|c_\infty|(\chi^2_{1,1-\alpha})^{1/2}) \text{ as } q_T \to \infty \tag{17.16}$$

for any sequence of constants $q_T = \lambda d^2_{\beta_*}(1 + o(1))$ as $\lambda d^2_{\beta_*} \to \infty$.

Suppose random variables $\{W_m : m \geq 1\}$ and W satisfy: (i) $W_m \to_d W$ as $m \to \infty$, (ii) W has a continuous and strictly increasing distribution function at its $1 - \alpha$ quantile κ_∞, and (iii) $P(W_m > \kappa_m) = \alpha$ for all $m \geq 1$ for some constants $\{\kappa_m : m \geq 1\}$. Then, $\kappa_m \to \kappa_\infty$. This holds because if $\limsup_{m\to\infty} \kappa_m > \kappa_\infty$, then there is a subsequence $\{v_m\}$ of $\{m\}$ such that $\lim_{m\to\infty} \kappa_{v_m} = \kappa_{\infty+} > \kappa_\infty$ and $\alpha = P(W_{v_m} > \kappa_{v_m}) \to P(W > \kappa_{\infty+}) < P(W > \kappa_\infty) = \alpha$, which is a contradiction, and likewise $\liminf_{m\to\infty} \kappa_m < \kappa_\infty$ leads to a contradiction.

We apply the result in the previous paragraph with (a) $\{W_m : m \geq 1\}$ given by $\log(POIS2(Q; \beta_0, \beta_*, \lambda)) + \log(2\psi_2(q_T; \beta_0, \beta_*, \lambda)) + \tau(\beta_*, \lambda, q_T)$ under the null hypothesis and conditional on $T = t$ with $t = 1^k q_T^{1/2}/k^{1/2}$ for some sequence of constants $q_T = \lambda d^2_{\beta_*}(1 + o(1)) \to \infty$ as $\lambda d^2_{\beta_*} \to$

∞, (b) $W = s(2c_\infty|S'1^k/k^{1/2}|)$, where $S'1^k/k^{1/2} \sim N(0,1)$, (c) κ_m equal to $\log(\kappa_{2,\beta_0}(q_T)) + \log(2\psi_2(q_T;\beta_0,\beta_*,\lambda)) + \tau(\beta_*,\lambda,q_T)$, and (d) $\kappa_\infty = s(2|c_\infty|(\chi^2_{1,1-\alpha})^{1/2})$.

We need to show conditions (i)-(iii) above hold. Condition (ii) holds straightforwardly for W as in (b) given the normal distribution of S, the functional form of $s(y)$, and $c_\infty \neq 0$.

By definition of $\kappa_{2,\beta_0}(q_T)$, under the null hypothesis, $P_{Q_1|Q_T}(POIS2(Q;\beta_0,\beta_*,\lambda) > \kappa_{2,\beta_0}(q_T) |q_T) = \alpha$ for all $q_T \geq 0$, see (11.6). This implies that the invariant POIS2 test is similar. In turn, this implies that under the null hypothesis $P(POIS2(Q;\beta_0,\beta_*,\lambda) > \kappa_{2,\beta_0}(q_T)|T=t) = \alpha$ for all $t \in R^k$ because Theorem 1 of Moreira (2009) shows that any invariant similar test has null rejection probability α conditional on T. This verifies condition (iii) because the log function is monotone and the last two summands of W_m and κ_m defined in (a) and (c) above cancel.

Next, we show that condition (i) holds. Given (17.15) and $t = 1^k q_T^{1/2}/k^{1/2}$, under the null and conditional on $T = t$, we have

$$\frac{Q_{ST}}{Q_T^{1/2}} = \frac{S't}{(t't)^{1/2}} = S'1^k/k^{1/2} \sim \chi^2_1, \tag{17.17}$$

which does not depend on $\lambda d^2_{\beta_*}$ or $\lambda^{1/2}c_{\beta_*}$. Hence, in place of the a.s. convergence result for $Q_{ST}/Q_T^{1/2}$ as $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2}c_{\beta_*\to}c_\infty$ in (17.5), which applies under the alternative hypothesis with true parameters (β_*,λ), we have $Q_{ST}/Q_T^{1/2} = S'1^k/k^{1/2}$ under the null hypothesis for all $\lambda d^2_{\beta_*}$ and $\lambda^{1/2}c_{\beta_*}$. Using this in place of (17.5), the unconditional a.s. convergence result in (17.13), established in (17.7)-(17.13), goes through as a conditional on $T = t$ a.s. result without any further changes. In consequence, the convergence in distribution result in (17.14) also holds conditional on $T = t$ a.s., but with $s(2c_\infty|S'1^k/k^{1/2}|)$ in place of $s(2c_\infty|LM_{1\infty}|)$. This verifies condition (i).

Given that conditions (i)-(iii) hold, we obtain $\kappa_m \to \kappa_\infty$ as $\lambda d^2_{\beta_*} \to \infty$ for κ_m and κ_∞ defined in (c) and (d), respectively, above. This establishes (17.16).

Given (17.16), we have

$$\begin{aligned}
&P_{\beta_*,\beta_0,\lambda,\Omega}(POIS2(Q;\beta_0,\beta_*,\lambda) > \kappa_{2,\beta_0}(Q_T)) \\
=\ &P_{\beta_*,\beta_0,\lambda,\Omega}(\log(POIS2(Q;\beta_0,\beta_*,\lambda)) + \log(2\psi_2(Q_T;\beta_0,\beta_*,\lambda)) + \tau(\beta_*,\lambda,Q_T) \\
&\quad > \log(\kappa_{2,\beta_0}(Q_T)) + \log(2\psi_2(Q_T;\beta_0,\beta_*,\lambda)) + \tau(\beta_*,\lambda,Q_T)) \\
\to_d\ &P(s(2c_\infty|LM_{1\infty}|) > s(2c_\infty|\chi^2_{1,1-\alpha}|)) \\
=\ &P(LM^2_{1\infty} > \chi^2_{1,1-\alpha}) \\
=\ &P(\chi^2_1(c^2_\infty) > \chi^2_{1,1-\alpha}),
\end{aligned} \tag{17.18}$$

where the second last equality uses the fact that $s(y)$ is symmetric and strictly increasing for $y > 0$ by Lemma 17.1(b). Equation (17.18) establishes part (a) of the theorem.

Now we establish part (b) of the theorem. Let

$$J := S' M_T S, \tag{17.19}$$

where $M_T := I_k - P_T$ and $P_T := T(T'T)^{-1}T'$. It follows from (3.3) that

$$LM = S' P_T S \text{ and } Q_S = LM + J. \tag{17.20}$$

By (17.8), $Q_T = \lambda d^2_{\beta_*}(1 + o_{a.s.}(1)) \to \infty$ a.s. as $\lambda d^2_{\beta_*} \to \infty$ when the true parameters are (β_*, λ). By (17.20) and some algebra, we have $(Q_S - Q_T)^2 + 4LM \cdot Q_T = (LM - J + Q_T)^2 + 4LM \cdot J$. This and the definition of LR in (3.3) give

$$LR = \frac{1}{2}\left(LM + J - Q_T + \sqrt{(LM - J + Q_T)^2 + 4LM \cdot J}\right). \tag{17.21}$$

Using a mean-value expansion of the square-root expression in (17.21) about $(LM - J + Q_T)^2$, we have

$$\sqrt{(LM - J + Q_T)^2 + 4LM \cdot J} = LM - J + Q_T + (2\sqrt{\zeta})^{-1} 4LM \cdot J \tag{17.22}$$

for an intermediate value ζ between $(LM - J + Q_T)^2$ and $(LM - J + Q_T)^2 + 4LM \cdot J$. It follows that

$$LR = LM + o(1) \text{ a.s.} \tag{17.23}$$

because $Q_T \to \infty$ a.s., $LM = O(1)$ a.s., and $J = O(1)$ a.s. as $\lambda d^2_{\beta_*} \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty \in R$, which imply that $(\sqrt{\zeta})^{-1} = o(1)$ a.s. These properties of LM and J hold because $LM = S' P_T S \le S'S$, $J = S' M_T S \le S'S$, and, using (17.4), we have $S'S = (c_{\beta_*}\mu_\pi + Z_S)'(c_{\beta_*}\mu_\pi + Z_S) = O(1)$ a.s. because $||c_{\beta_*}\mu_\pi||^2 = \lambda c^2_{\beta_*} = O(1)$ by assumption.

The critical value function for the CLR test, $\kappa_{LR,\alpha}(\cdot)$, depends only on k and α, see Lemma 3(c) and (3.5) in AMS. It is well known in the literature that $\kappa_{LR,\alpha}(\cdot)$ satisfies $\kappa_{LR,\alpha}(q_T) \to \chi^2_{1,1-\alpha}$ as $q_T \to \infty$, e.g., see Moreira (2003, Proposition 1). Hence, we have

$$\begin{aligned} &P_{\beta_*,\beta_0,\lambda,\Omega}(LR > \kappa_{LR,\alpha}(Q_T)) = P_{\beta_*,\beta_0,\lambda,\Omega}(LM + o_{a.s.}(1) > \chi^2_{1,1-\alpha} + o_{a.s.}(1)) \\ &= P_{\beta_*,\beta_0,\lambda,\Omega}(LM + o_p(1) > \chi^2_{1,1-\alpha}) \to P(\chi^2_1(c^2_\infty) > \chi^2_{1,1-\alpha}) \end{aligned} \tag{17.24}$$

as $\lambda d_{\beta_*}^2 \to \infty$ and $\lambda^{1/2} c_{\beta_*} \to c_\infty$, where the first equality holds by (17.23), $Q_T \to \infty$ a.s. by (17.8), and $\lim_{q_T \to \infty} \kappa_{LR,\alpha}(q_T) = \chi^2_{1,1-\alpha}$ and the convergence holds by part (c) of the theorem. This establishes part (b) of the theorem. $\square$

Proof of Theorem 7.2. First, we establish part (a)(i) of the theorem. By (11.8) with $\beta = \beta_*$ and $\Sigma = \Sigma_*$, we have

$$\Omega(\beta_*, \Sigma_*) = \begin{bmatrix} \omega_1^2 & \omega_{12} \\ \omega_{12} & \omega_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_u^2 + 2\sigma_{uv}\beta_* + \sigma_v^2\beta_*^2 & \sigma_{uv} + \sigma_v^2\beta_* \\ \sigma_{uv} + \sigma_v^2\beta_* & \sigma_v^2 \end{bmatrix}. \tag{17.25}$$

Using this, we obtain, as $\rho_{uv} \to \pm 1$,

$$\begin{aligned} c_{\beta_*} = c_{\beta_*}(\beta_0, \Omega(\beta_*, \Sigma_*)) &= (\beta_* - \beta_0)(\omega_1^2 - 2\beta_0\omega_{12} + \omega_2^2\beta_0^2)^{-1/2} \\ &= (\beta_* - \beta_0)(\sigma_u^2 + 2\sigma_{uv}\beta_* + \sigma_v^2\beta_*^2 - 2\beta_0(\sigma_{uv} + \sigma_v^2\beta_*) + \sigma_v^2\beta_0^2)^{-1/2} \\ &= (\beta_* - \beta_0)(\sigma_u^2 + 2(\beta_* - \beta_0)\sigma_u\sigma_v\rho_{uv} + (\beta_* - \beta_0)^2\sigma_v^2)^{-1/2} \\ &\to (\beta_* - \beta_0)(\sigma_u^2 \pm 2(\beta_* - \beta_0)\sigma_u\sigma_v + (\beta_* - \beta_0)^2\sigma_v^2)^{-1/2} \\ &= (\beta_* - \beta_0)/|\sigma_u \pm (\beta_* - \beta_0)\sigma_v|, \end{aligned} \tag{17.26}$$

where the second equality uses (2.3), the convergence only holds if $\sigma_u \pm (\beta_* - \beta_0)\sigma_v \neq 0$, and the fourth equality uses $\sigma_{uv} = \sigma_u\sigma_v\rho_{uv}$. This proves part (a)(i).

To prove part (a)(ii), we have

$$\begin{aligned} d_{\beta_*} &= d_{\beta_*}(\beta_0, \Omega(\beta_*, \Sigma_*)) = b_*'\Omega b_0(b_0'\Omega b_0)^{-1/2}\det(\Omega)^{-1/2} \\ &= (\omega_1^2 - \omega_{12}(\beta_0 + \beta_*) + \omega_2^2\beta_0\beta_*) \cdot (\omega_1^2 - 2\beta_0\omega_{12} + \omega_2^2\beta_0^2)^{-1/2} \cdot (\omega_1^2\omega_2^2 - \omega_{12}^2)^{-1/2}, \end{aligned} \tag{17.27}$$

where the second equality holds by (2.3). The second multiplicand on the rhs of (17.27) converges to $|\sigma_u \pm (\beta_* - \beta_0)\sigma_v|^{-1}$ provided $\sigma_u \pm (\beta_* - \beta_0)\sigma_v \neq 0$ by the calculations in (17.26).

The first multiplicand on the rhs of (17.27) satisfies, as $\rho_{uv} \to \pm 1$,

$$\begin{aligned} &\omega_1^2 - \omega_{12}(\beta_0 + \beta_*) + \omega_2^2\beta_0\beta_* \\ = \ &\sigma_u^2 + 2\sigma_{uv}\beta_* + \sigma_v^2\beta_*^2 - (\sigma_{uv} + \sigma_v^2\beta_*)(\beta_0 + \beta_*) + \sigma_v^2\beta_0\beta_* \\ = \ &\sigma_u^2 + \sigma_u\sigma_v\rho_{uv}(\beta_* - \beta_0) \\ \to \ &\sigma_u(\sigma_u \pm \sigma_v(\beta_* - \beta_0)), \end{aligned} \tag{17.28}$$

where the first equality uses (17.25) and the second equality holds by simple algebra and $\sigma_{uv} = \sigma_u \sigma_v \rho_{uv}$.

The reciprocal of the square of the third multiplicand on the rhs of (17.27) satisfies, as $\rho_{uv} \to \pm 1$,

$$
\begin{aligned}
\omega_1^2 \omega_2^2 - \omega_{12}^2 &= (\sigma_u^2 + 2\sigma_u \sigma_v \rho_{uv} \beta_* + \sigma_v^2 \beta_*^2)\sigma_v^2 - (\sigma_u \sigma_v \rho_{uv} + \sigma_v^2 \beta_*)^2 \\
&\to (\sigma_u^2 \pm 2\sigma_u \sigma_v \beta_* + \sigma_v^2 \beta_*^2)\sigma_v^2 - (\pm \sigma_u \sigma_v + \sigma_v^2 \beta_*)^2 \\
&= (\sigma_u \pm \sigma_v \beta_*)^2 \sigma_v^2 - (\pm \sigma_u + \sigma_v \beta_*)^2 \sigma_v^2 \\
&= 0,
\end{aligned} \tag{17.29}
$$

where the first equality holds by (17.25) and $\sigma_{uv} = \sigma_u \sigma_v \rho_{uv}$.

Combining (17.27)-(17.29) and $\lambda > 0$ proves part (a)(ii).

Next, we establish part (b) of the theorem. Using the definition of $c_\beta(\beta_0, \Omega)$ in (2.3), we have

$$
\begin{aligned}
\lim_{\rho_\Omega \to \pm 1} c_{\beta_*}(\beta_0, \Omega) &= \lim_{\rho_\Omega \to \pm 1} (\beta_* - \beta_0)(b_0' \Omega b_0)^{-1/2} \\
&= \lim_{\rho_\Omega \to \pm 1} (\beta_* - \beta_0)(\omega_1^2 - 2\beta_0 \omega_1 \omega_2 \rho_\Omega + \omega_2^2 \beta_0^2)^{-1/2} \\
&= (\beta_* - \beta_0)/|\omega_1 \mp \omega_2 \beta_0|,
\end{aligned} \tag{17.30}
$$

where the third equality holds provided $\omega_1 \mp \omega_2 \beta_0 \neq 0$. This establishes part (b)(i) of the theorem.

Using the definition of $d_\beta(\beta_0, \Omega)$ in (2.3) and $b_* := (1, \beta_*)'$, we have

$$
\begin{aligned}
\lim_{\rho_\Omega \to \pm 1} d_{\beta_*}(\beta_0, \Omega) &= \lim_{\rho_\Omega \to \pm 1} b_*' \Omega b_0 (b_0' \Omega b_0)^{-1/2} \det(\Omega)^{-1/2} \\
&= \lim_{\rho_\Omega \to \pm 1} (\omega_1^2 - \omega_1 \omega_2 \rho_\Omega (\beta_0 + \beta_*) + \omega_2^2 \beta_0 \beta_*) \cdot (\omega_1^2 - 2\beta_0 \omega_1 \omega_2 \rho_\Omega + \omega_2^2 \beta_0^2)^{-1/2} \\
&\quad \cdot (\omega_1^2 \omega_2^2 - \omega_1^2 \omega_2^2 \rho_\Omega^2)^{-1/2} \\
&= (\omega_1 \mp \omega_2 \beta_0)(\omega_1 \mp \omega_2 \beta_*) \cdot \frac{1}{|\omega_1 \mp \omega_2 \beta_0|} \cdot \frac{1}{\omega_1 \omega_2} \cdot \lim_{\rho_\Omega \to \pm 1} \frac{1}{(1 - \rho_\Omega^2)^{1/2}} \\
&= sgn((\omega_1 \mp \omega_2 \beta_0)(\omega_1 \mp \omega_2 \beta_*)) \cdot \infty,
\end{aligned} \tag{17.31}
$$

where the third and fourth equalities hold provided $\omega_1 \mp \omega_2 \beta_0 \neq 0$ and $\omega_1 \mp \omega_2 \beta_* \neq 0$. This and $\lambda > 0$ establish part (b)(ii) of the theorem.

Part (c)(i) is proved as follows:

$$
c_{\beta_*} = \frac{\beta_* - \beta_0}{(\sigma_u^2 + 2(\beta_* - \beta_0)\sigma_u \sigma_v \rho_{uv} + (\beta_* - \beta_0)^2 \sigma_v^2)^{1/2}} \to \mp \frac{1}{\sigma_v} \text{ as } (\rho_{uv}, \beta_0) \to (1, \pm \infty), \tag{17.32}
$$

where the first equality holds by (17.26) and the convergence holds by considering only the dominant β_0 terms. The same result holds as $(\rho_{uv}, \beta_0) \to (1, \pm\infty)$ because ρ_{uv} enters the middle expression in (17.32) only through a term that does not affect the limit.

Part (c)(ii) is proved using the expression for d_{β_*} in (17.27). By (17.29), the third multiplicand in (17.27), which does not depend on β_0, diverges to infinity when $\rho_{uv} \to 1$ or -1. The product of the first two multiplicands on the rhs of (17.27) equals

$$\begin{aligned}\frac{\omega_1^2 - \omega_{12}(\beta_0 + \beta_*) + \omega_2^2\beta_0\beta_*}{(\omega_1^2 - 2\beta_0\omega_{12} + \omega_2^2\beta_0^2)^{1/2}} &= \frac{\sigma_u^2 + \sigma_u\sigma_v\rho_{uv}(\beta_* - \beta_0)}{(\sigma_u^2 + 2(\beta_* - \beta_0)\sigma_u\sigma_v\rho_{uv} + (\beta_* - \beta_0)^2\sigma_v^2)^{1/2}} \\ &\to \mp\frac{\sigma_u\sigma_v}{\sigma_v} = \mp\sigma_u \text{ as } (\rho_{uv}, \beta_0) \to (1, \pm\infty),\end{aligned} \tag{17.33}$$

where the equality uses the calculations in the first three lines of (17.26) and (17.28) and the convergence holds by considering only the dominant β_0 terms. When $(\rho_{uv}, \beta_0) \to (-1, \pm\infty)$, the limit in (17.33) is $\pm\sigma_u$ because ρ_{uv} enters multiplicatively in the dominant β_0 term in the numerator. In both cases, the product of the first two multiplicands on the rhs of (17.27) converges to a non-zero constant and the third multiplicand diverges to infinity. Hence, d_{β_*} diverges to $+\infty$ or $-\infty$ and $\lambda d_{\beta_*}^2 \to \infty$ since $\lambda > 0$, which completes the proof.

Part (d)(i) holds because

$$c_{\beta_*} = \frac{\beta_* - \beta_0}{(\omega_1^2 - 2\beta_0\omega_1\omega_2\rho_\Omega + \omega_2^2\beta_0^2)^{1/2}} \to \mp\frac{1}{\omega_2} \text{ as } (\rho_\Omega, \beta_0) \to (1, \pm\infty), \tag{17.34}$$

where the equality uses (17.30). The same convergence holds as $(\rho_\Omega, \beta_0) \to (1, \pm\infty)$ because ρ_{uv} enters the middle expression in (17.34) only through a term that does not affect the limit.

Part (d)(ii) is proved using the expression for d_{β_*} in (17.31):

$$\begin{aligned}&d_{\beta_*} = \frac{(\omega_1^2 - \omega_1\omega_2\rho_\Omega(\beta_0 + \beta_*) + \omega_2^2\beta_0\beta_*)}{(\omega_1^2 - 2\beta_0\omega_1\omega_2\rho_\Omega + \omega_2^2\beta_0^2)^{1/2}} \cdot (\omega_1^2\omega_2^2 - \omega_1^2\omega_2^2\rho_\Omega^2)^{-1/2}, \\ &\frac{(\omega_1^2 - \omega_1\omega_2\rho_\Omega(\beta_0 + \beta_*) + \omega_2^2\beta_0\beta_*)}{(\omega_1^2 - 2\beta_0\omega_1\omega_2\rho_\Omega + \omega_2^2\beta_0^2)^{1/2}} \to \frac{\pm(\omega_2^2\beta_* - \omega_1\omega_2)}{\omega_2} = \mp(\omega_1 - \omega_2\beta_*), \text{ and} \\ &(\omega_1^2\omega_2^2 - \omega_1^2\omega_2^2\rho_\Omega^2)^{-1/2} \to \infty \text{ as } (\rho_\Omega, \beta_0) \to (1, \pm\infty).\end{aligned} \tag{17.35}$$

Hence, $\lambda d_{\beta_*}^2 \to \infty$ as $(\rho_\Omega, \beta_0) \to (1, \pm\infty)$ provided $\omega_1 - \omega_2\beta_* \neq 0$. When $(\rho_\Omega, \beta_0) \to (-1, \pm\infty)$, the limit in the second line of (17.35) is $\pm(\omega_2^2\beta_* + \omega_1\omega_2)/\omega_2 = \pm(\omega_1 + \omega_2\beta_*)$ and, hence, $\lambda d_{\beta_*}^2 \to \infty$ provided $\omega_1 + \omega_2\beta_* \neq 0$, which completes the proof. □

18 Proofs of Theorem 12.1 and Lemmas 13.1 and 13.2

Proof of Theorem 12.1. By Cor. 2 and Comment 2 to Cor. 2 of Andrews, Moreira, and Stock (2004), for all $(\beta_*, \beta_0, \lambda, \Omega)$,

$$P_{\beta_*,\beta_0,\lambda,\Omega}(\phi_{\beta_0}(Q) = 1) \leq P_{\beta_*,\beta_0,\lambda,\Omega}(POIS(Q; \beta_0, \beta_*) > \kappa_{\beta_0}(Q_T)). \tag{18.1}$$

That is, the test on the rhs is the (one-sided) POIS test for testing $H_0 : \beta = \beta_0$ versus $H_1 : \beta = \beta_*$ for fixed known Ω and any $\lambda \geq 0$ under H_1.

We use the dominated convergence theorem (DCT) to show

$$\lim_{\beta_0 \to \pm\infty} P_{\beta_*,\beta_0,\lambda,\Omega}(POIS(Q; \beta_0, \beta_*) > \kappa_{\beta_0}(Q_T)) = P_{\rho_{uv},\lambda_v}(POIS(Q; \infty, \rho_{uv}) > \kappa_\infty(Q_T)). \tag{18.2}$$

Equations (18.1) and (18.2) imply that the result of Theorem 12.1 holds.

By (12.2), (12.5), and Lemma 14.1(b) and (e),

$$\lim_{\beta_0 \to \pm\infty} POIS(q; \beta_0, \beta_*) = POIS(q; \infty, \rho_{uv}) \tag{18.3}$$

for all 2×2 variance matrices q, for given (β_*, π, Ω).

The proof of (18.2) is the same as the proof of Lemma 16.1(a), but with $POIS(Q; \beta_0, \beta_*)$, $\kappa_{\beta_0}(Q_T)$, $POIS(Q; \infty, \rho_{uv})$, and $\kappa_\infty(Q_T)$ in place of $POIS2(Q; \beta_0, \beta_*, \lambda)$, $\kappa_{2,\beta_0}(Q_T)$, $POIS2(Q; \infty, |\rho_{uv}|, \lambda_v)$, and $\kappa_{2,\infty}(Q_T)$, respectively, using (18.3) in place of (16.6), and using the results (established below) that (i) the Lebesgue measure of the set of (q_S, q_{ST}, q_T) for which $POIS(q; \infty, \rho_{uv}) = \kappa_\infty(q_T)$ is zero, (ii) $P_{Q_1|Q_T}(POIS(Q; \infty, \rho_{uv}) = x | q_T) = 0$ for all $q_T \geq 0$, and (iii) the distribution function of $POIS(Q; \infty, \rho_{uv})$ is strictly increasing at its $1 - \alpha$ quantile $\kappa_\infty(q_T)$ under $P_{Q_1|Q_T}(\cdot | q_T)$ for all $q_T \geq 0$.

Condition (i) holds because (a) $POIS(q; \infty, \rho_{uv}) = q_S + 2r_{uv}q_{ST}$ (see (12.5)) implies that the Lebesgue measure of the set of (q_S, q_{ST}) for which $q_S + 2r_{uv}q_{ST} = \kappa_\infty(q_T)$ is zero for all q_T and (b) the Lebesgue measure of the set of (q_S, q_{ST}, q_T) for which $q_S + 2r_{uv}q_{ST} = \kappa_\infty(q_T)$ is obtained by integrating the set in (a) over $q_T \in R$ subject to the constraint that q is positive definite.

Condition (ii) holds by the absolute continuity of $POIS(Q; \infty, \rho_{uv})$ under $P_{Q_1|Q_T}(\cdot | q_T)$ (by the functional form of $POIS(Q; \infty, \rho_{uv})$ and the absolute continuity of Q_1 under $P_{Q_1|Q_T}(\cdot | q_T)$, whose density is given in (11.3)).

Condition (iii) holds because we can write $POIS(Q; \infty, \rho_{uv}) = S'S + 2r_{uv}S'T = (S + r_{uv}T)'(S + r_{uv}T) - r_{uv}^2 T'T$, where $[S : T]$ has a multivariate normal distribution with means matrix given by (5.2) and identity variance matrix and, hence, $POIS(Q; \infty, \rho_{uv})$ has a shifted noncentral χ^2 distribution conditional on $T = t$. In consequence, it has a positive density on $(r_{uv}^2 t't, \infty) = (r_{uv}^2 q_T, \infty)$ conditional on $T = t$ and also conditional on $Q_T = q_T$ (because the latter conditional density is the integral of the former conditional density over t such that $t't = q_T$). This completes the proof. □

Proof of Lemma 13.1. First, we show that (13.4) implies the equation for λ_2 in (13.2). By the expression $d_\beta = a'\Omega^{-1}a_0(a_0'\Omega^{-1}a_0)^{-1/2}$ given in (2.7) in AMS, where $a := (\beta, 1)'$ and $a_0 := (\beta_0, 1)'$, for any $\beta \in R$,

$$\begin{aligned} d_\beta - d_{\beta_0} &= (a - a_0)'\Omega^{-1}a_0(a_0'\Omega^{-1}a_0)^{-1/2} \\ &= (\beta - \beta_0)e_1'\Omega^{-1}a_0(a_0'\Omega^{-1}a_0)^{-1/2} := (\beta - \beta_0)r_{\beta_0}, \end{aligned} \tag{18.4}$$

where $e_1 := (1, 0)'$ and the last equality holds by the definition of r_{β_0}.

Substituting (18.4) into the second equation in (13.4) gives

$$\begin{aligned} &\lambda_2^{1/2} d_{\beta_{2*}} = \pm\lambda^{1/2} d_{\beta_*} \\ \text{iff } &\lambda_2^{1/2}(d_{\beta_0} + r_{\beta_0}(\beta_{2*} - \beta_0)) = \pm\lambda^{1/2}(d_{\beta_0} + r_{\beta_0}(\beta_* - \beta_0)) \\ \text{iff } &\lambda_2^{1/2} d_{\beta_0} = \pm\lambda^{1/2}(d_{\beta_0} + r_{\beta_0}(\beta_* - \beta_0)) - r_{\beta_0}\lambda_2^{1/2}(\beta_{2*} - \beta_0). \end{aligned} \tag{18.5}$$

Given the definition of c_β in (2.3), the first equation in (13.4) can be written as

$$\lambda_2^{1/2}(\beta_{2*} - \beta_0) = \mp\lambda^{1/2}(\beta_* - \beta_0). \tag{18.6}$$

Substituting this into (18.5) yields

$$\begin{aligned} &\lambda_2^{1/2} d_{\beta_{2*}} = \pm\lambda^{1/2} d_{\beta_*} \\ \text{iff } &\lambda_2^{1/2} d_{\beta_0} = \pm\lambda^{1/2}(d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)) \\ \text{iff } &\lambda_2^{1/2} = \pm\lambda^{1/2}\frac{d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)}{d_{\beta_0}}. \end{aligned} \tag{18.7}$$

The square of the equation in the last line in (18.7) is the equation for λ_2 in (13.2).

Next, we show that (13.4) implies the equation for β_{2*} in (13.2). Using (18.6), the first equation in (13.4) can be written as

$$\beta_{2*} = \beta_0 \mp \frac{\lambda^{1/2}}{\lambda_2^{1/2}}(\beta_* - \beta_0). \tag{18.8}$$

This combined with the equation for $\lambda^{1/2}/\lambda_2^{1/2}$ obtained from the last line of (18.7) gives

$$\beta_{2*} = \beta_0 - \frac{d_{\beta_0}}{d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)}(\beta_* - \beta_0), \tag{18.9}$$

where a minus sign appears because the $\mp$ sign in (18.8) gets multiplied by the $\pm$ sign in the last line of (18.7), which yields a minus sign in both cases. Equation (18.9) is the same as the first condition in (13.2). This completes the proof that (13.4) implies (13.2).

Now, we prove the converse. We suppose (13.2) holds. Taking the square root of the second equation in (13.2) gives

$$\lambda_2^{1/2} = \pm\lambda^{1/2}\frac{d_{\beta_0} + 2r_{\beta_0}(\beta_* - \beta_0)}{d_{\beta_0}}, \tag{18.10}$$

where the $\pm$ sign means that this equation holds either with $+$ or with $-$. Substituting this into the first equation in (13.2) gives (18.8), which is the same as (18.6), and (18.6) is the first equation in (13.4).

The second equation in (13.4) is given by (18.5). Given that the first equation in (13.4) holds, the second equation in (13.4) is given in (18.7). The last line of (18.7) holds by (18.10). This completes the proof that (13.2) implies (13.4). $\square$

Proof of Lemma 13.2. The proof of part (a) of the lemma is essentially the same as that of Theorem 8(b) in AMS. The only change is to note that when (β_{2*}, λ_2) satisfies (13.3), we have $\tau^* = \tau_2^*$, $\delta^* = -\delta_2^*$, and $\delta_{\max} = |\delta^*| = |\delta_2^*|$ (using the notation in AMS). Because $\delta_{\max} = |\delta^*| = |\delta_2^*|$, we obtain $\sqrt{\delta^2} - \sqrt{\delta_{\max}^2} = 0$ and the remainder of the proof of Theorem 8(b) goes through as is.

The proof of part (b) of the lemma is quite similar to the proof of Theorem 8(c) of AMS. The latter proof first considers the case where "(β_{2*}, λ_2) does not satisfy the second condition of (13.1)." This needs to be changed to "(β_{2*}, λ_2) does not satisfy the second condition of (13.1) or (13.3)." With this change, the rest of that part of the proof of Theorem 8(c) goes through unchanged.

The remaining cases (where both (13.1) and (13.3) fail) to consider are (i) when the second condition in (13.1) holds and the first condition in (13.1) fails and (ii) when the second condition in (13.3) holds and the first condition in (13.3) fails. These are mutually exclusive scenarios because the second conditions in (13.1) and (13.3) are incompatible. The proof of Theorem 8(c) of AMS considers case (i) and proves the result of Theorem 8(c) for that case. The proof of Theorem 8(c) for case (ii) is quite similar to that for case (i) using (A.21) in AMS because $\delta^* = -\delta_2^*$, $\delta_{\max} = |\delta^*| = |\delta_2^*| > 0$, and $\tau^* \neq \tau_2^*$ imply that $sgn(\delta^*) = -sgn(\delta_2^*)$ and $\tau^* sgn(\delta^*) \neq -\tau_2^* sgn(\delta_2^*)$. This last inequality shows that the expression in (A.21) in AMS is a continuous function of $Q_{ST}Q_T^{-1/2}$ that is not even. (Note that (A.21) in AMS has a typo: the quantity $\tau_2^* sgn(\delta^*)$ in its second summand should be $\tau_2^* sgn(\delta_2^*)$.) $\square$

19 Power Against Distant Alternatives Compared to Distant Null Hypotheses

19.1 Scenario 1 Compared to Scenario 2

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 is fixed and β_* takes on large (absolute) values, to scenario 2, where β_* is 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, it is not true that power goes to one in scenario 2 as $\beta_0 \to \pm\infty$ for fixed β_*.

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

$$c_\beta^2(\beta_0, \Omega) \cdot \lambda \tag{19.1}$$

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

$$\lim_{|\beta_*| \to \infty} c^2_{\beta_*}(\beta_0, \Omega) = \lim_{|\beta_*| \to \infty} (\beta_* - \beta_0)^2 \cdot (b_0' \Omega b_0)^{-1} = \infty. \tag{19.2}$$

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

$$\begin{aligned}
\lim_{|\beta_0| \to \infty} c^2_{\beta_*}(\beta_0, \Omega) &= \lim_{|\beta_0| \to \infty} (\beta_* - \beta_0)^2 \cdot (b_0' \Omega b_0)^{-1} \\
&= \lim_{|\beta_0| \to \infty} (\beta_* - \beta_0)^2 \cdot (\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2 \beta_0^2)^{-1} \\
&= 1/\omega_2^2.
\end{aligned} \tag{19.3}$$

Hence, the power of the AR test does not go to one as $|\beta_0| \to \infty$ even though $|\beta_0 - \beta_*| \to \infty$.

The differing results in (19.2) and (19.3) 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. The numerical power function and power envelope calculations in AMS are all of the type in scenario 1 in (19.2). The difference in the power of the AR (and other) tests between (19.2) and (19.3) 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.

19.2 Structural Error Variance Matrices under Distant Alternatives and Distant Null Hypotheses

Here, we compute the structural error variance matrices in scenarios 1 and 2 considered in (19.2) and (19.3). By design, the reduced-form variance matrix Ω is the same for β_0 and β_* and, hence, does not vary between these two scenarios.

In scenario 1 in (19.2), the structural error variance matrix under H_0 is $\Sigma(\beta_0, \Omega)$, defined in (11.9). Under $H_1 : \beta = \beta_*$, as $|\beta_*| \to \infty$, we have

$$\begin{aligned}\lim_{\beta_* \to \pm\infty} \rho_{uv}(\beta_*, \Omega) &= \lim_{\beta_* \to \pm\infty} \frac{\omega_{12} - \omega_2^2 \beta_*}{(\omega_1^2 - 2\omega_{12}\beta_* + \omega_2^2 \beta_*^2)^{1/2} \omega_2} = \mp 1 \text{ and} \\ \lim_{|\beta_*| \to \infty} \sigma_u^2(\beta_*, \Omega)/\sigma_v^2(\beta_*, \Omega) &= \frac{\omega_1^2 - 2\omega_{12}\beta_* + \omega_2^2 \beta_*^2}{\omega_2^2} = \infty, \end{aligned} \tag{19.4}$$

where $\rho_{uv}(\beta_*, \Omega)$, $\sigma_u^2(\beta_*, \Omega)$, and $\sigma_v^2(\beta_*, \Omega)$ are defined just below (11.9). Equation (19.4) shows that, for standard power envelope calculations, when the alternative hypothesis value β_* is large in absolute value the structural variance matrix under H_1 exhibits correlation close to one in absolute value and a large ratio of structural to reduced-form variances.

In scenario 2 in (19.3), the structural error variance error matrix under H_* is $\Sigma(\beta_*, \Omega)$. Under $H_0 : \beta = \beta_0$, by exactly the same argument as in (19.4) with β_0 in place of β_*, we obtain

$$\lim_{\beta_0 \to \pm\infty} \rho_{uv}(\beta_0, \Omega) = \mp 1 \text{ and } \lim_{|\beta_0| \to \infty} \sigma_u^2(\beta_0, \Omega)/\sigma_v^2(\beta_0, \Omega) = \infty. \tag{19.5}$$

So, in scenario 2, when the null hypothesis value β_0 is large in absolute value the structural variance matrix under H_0 exhibits correlation close to one in absolute value and a large ratio of structural to reduced-form variances.

From a testing perspective, it is natural and time honored to fix the null hypothesis value β_0 and consider power as the alternative hypothesis value β_* varies. On the other hand, a confidence set is the set of null hypothesis values β_0 for which one does not reject $H_0 : \beta = \beta_0$. Hence, for a given true value β_*, the false coverage probabilities of the confidence set equal one minus its power as one varies $H_0 : \beta = \beta_0$. Thus, from the confidence set perspective, it is natural to fix β_* and consider power as β_0 varies.

20 Transformation of the $\boldsymbol{\beta_0}$ Versus $\boldsymbol{\beta_*}$ Testing Problem to a 0 Versus $\overline{\boldsymbol{\beta}}_*$ Testing Problem

In this section, we transform the general testing problem of $H_0 : \beta = \beta_0$ versus $H_1 : \beta = \beta_*$ for $\pi \in R^k$ and fixed Ω to a testing problem of $H_0 : \overline{\beta} = 0$ versus $H_1 : \overline{\beta} = \overline{\beta}_*$ for some $\overline{\pi} \in R^k$ and some fixed $\overline{\Omega}$ whose diagonal elements equal one. This is done using the transformations given footnotes 7 and 8 of AMS, which argue that there is no loss in generality in the AMS numerical results to take $\omega_1^2 = \omega_2^2 = 1$ and $\beta_0 = 0$. These results help link the numerical work done in this paper with that done in AMS.

Starting with the model in (2.1), we transform the model based on (y_1, y_2) with parameters (β, π) and fixed reduced-form variance matrix Ω to a model based on $(\widetilde{y}_1, y_2)$ with parameters $(\widetilde{\beta}, \pi)$ and fixed reduced-form variance matrix $\widetilde{\Omega}$, where

$$\begin{aligned}
\widetilde{y}_1 &:= y_1 - y_2\beta_0, \\
\widetilde{\beta} &:= \beta - \beta_0, \text{ and} \\
\widetilde{\Omega} &:= Var\left(\begin{pmatrix} \widetilde{y}_1 \\ y_2 \end{pmatrix}\right) = Var\left(\begin{bmatrix} 1 & -\beta_0 \\ 0 & 1 \end{bmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) \\
&= \begin{bmatrix} \omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2 & \omega_{12} - \omega_2^2\beta_0 \\ \omega_{12} - \omega_2^2\beta_0 & \omega_2^2 \end{bmatrix}. \qquad (20.1)
\end{aligned}$$

The transformed testing problem is $H_0 : \widetilde{\beta} = 0$ versus $H_1 : \widetilde{\beta} = \widetilde{\beta}_*$, where $\widetilde{\beta}_* = \beta_* - \beta_0$, with parameter π and reduced-form variance matrix $\widetilde{\Omega}$.

The matrix $\widetilde{\Omega}$ does not have diagonal elements equal to one, so we transform the model based on $(\widetilde{y}_1, y_2)$ with parameters $(\widetilde{\beta}, \pi)$ and fixed reduced-form variance matrix $\widetilde{\Omega}$ to a model based on

$(\overline{y}_1, \overline{y}_2)$ with parameters $(\overline{\beta}, \overline{\pi})$ and fixed reduced-form variance matrix $\overline{\Omega}$, where[12]

$$\begin{aligned}
\overline{y}_1 &:= \frac{\widetilde{y}_1}{\widetilde{\omega}_1} = \frac{y_1 - y_2\beta_0}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}} \\
\overline{y}_2 &:= \frac{1}{\widetilde{\omega}_2} y_2 = \frac{1}{\omega_2} y_2, \\
\overline{\beta} &:= \frac{\widetilde{\omega}_2}{\widetilde{\omega}_1}\widetilde{\beta} = \frac{\omega_2}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}}(\beta - \beta_0), \text{ and} \\
\overline{\pi} &:= \frac{1}{\widetilde{\omega}_2}\pi = \frac{1}{\omega_2}\pi.
\end{aligned} \tag{20.2}$$

In addition, we have

$$\begin{aligned}
\overline{\Omega} &:= Var\begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix} = Var\left(\begin{bmatrix} 1/\widetilde{\omega}_1 & 0 \\ 0 & 1/\widetilde{\omega}_2 \end{bmatrix}\begin{pmatrix} \widetilde{y}_1 \\ y_2 \end{pmatrix}\right) \\
&= \begin{bmatrix} 1/\widetilde{\omega}_1 & 0 \\ 0 & 1/\widetilde{\omega}_2 \end{bmatrix} \widetilde{\Omega} \begin{bmatrix} 1/\widetilde{\omega}_1 & 0 \\ 0 & 1/\widetilde{\omega}_2 \end{bmatrix} \\
&= \begin{bmatrix} 1/\widetilde{\omega}_1 & 0 \\ 0 & 1/\omega_2 \end{bmatrix} \begin{bmatrix} \omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2 & \omega_{12} - \omega_2^2\beta_0 \\ \omega_{12} - \omega_2^2\beta_0 & \omega_2^2 \end{bmatrix} \begin{bmatrix} 1/\widetilde{\omega}_1 & 0 \\ 0 & 1/\omega_2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & \frac{\omega_{12} - \omega_2^2\beta_0}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}\omega_2} \\ \frac{\omega_{12} - \omega_2^2\beta_0}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}\omega_2} & 1 \end{bmatrix}.
\end{aligned} \tag{20.3}$$

The transformed testing problem is $H_0 : \overline{\beta} = 0$ versus $H_1 : \overline{\beta} = \overline{\beta}_*$, where

$$\overline{\beta}_* = \frac{\omega_2}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}}(\beta_* - \beta_0), \tag{20.4}$$

with parameter $\overline{\pi}$ and reduced-form variance matrix $\overline{\Omega}$.

Now, we consider the limit as $\beta_0 \to \pm\infty$ of the original model and see what it yields in terms of the transformed model. We have

$$\lim_{\beta_0 \to \pm\infty} \overline{\beta}_* = \mp 1 \text{ and } \lim_{\beta_0 \to \pm\infty} \overline{\Omega} = \begin{bmatrix} 1 & \mp 1 \\ \mp 1 & 1 \end{bmatrix}. \tag{20.5}$$

So, the asymptotic testing problem as $\beta_0 \to \pm\infty$ in terms of a model with a null hypothesis $\overline{\beta}$ value of 0 and a reduced-form variance matrix $\overline{\Omega}$ with ones on the diagonal is a test of $H_0 : \overline{\beta} = 0$ versus $H_1 : \overline{\beta} = \mp 1$.

[12] The formula $\overline{\beta} := (\widetilde{\omega}_2/\widetilde{\omega}_1)\widetilde{\beta}$ in (20.2) comes from $\overline{y}_1 := \widetilde{y}_1/\widetilde{\omega}_1 = (y_2\widetilde{\beta}+u)/\widetilde{\omega}_1 = y_2\widetilde{\beta}/\widetilde{\omega}_1 + u/\widetilde{\omega}_1 = (y_2/\widetilde{\omega}_2)\widetilde{\beta}(\widetilde{\omega}_2/\widetilde{\omega}_1) + u/\widetilde{\omega}_1 = \overline{y}_2\overline{\beta} + \overline{u}$, where the last equality holds when $\overline{\beta} := (\widetilde{\omega}_2/\widetilde{\omega}_1)\widetilde{\beta}$ and $\overline{u} := u/\widetilde{\omega}_1$.

We get the same expression for the limits as $\beta_0 \to \pm\infty$ of $c_{\beta_*}(\beta_0, \Omega)$ and $d_{\beta_*}(\beta_0, \Omega)$ written in terms of the transformed parameters $(\overline{\beta}_0, \overline{\beta}_*, \overline{\pi}, \overline{\Omega})$ as in Lemma 14.1 except they are multiplied by σ_v. This occurs because $\mu_{\overline{\pi}} = \mu_\pi/\sigma_v$. In consequence, the limits as $\beta_0 \to \pm\infty$ of $c_{\beta_*}(\beta_0, \Omega)\mu_\pi$ and $d_{\beta_*}(\beta_0, \Omega)\mu_\pi$ written in terms of the transformed parameters $(\overline{\beta}_0, \overline{\beta}_*, \overline{\pi}, \overline{\Omega})$ are the same as their limits without any transformation.

Lemma 20.1 *Let* $\overline{\beta}_* = \overline{\beta}_*(\beta_0)$ *and* $\overline{\Omega} = \overline{\Omega}(\beta_0)$ *be defined in* (20.4) *and* (20.3), *respectively. Let* $\overline{\beta}_0(\beta_0) = 0$.

(a) $\lim_{\beta_0 \to \pm\infty} c_{\overline{\beta}_*(\beta_0)}(\overline{\beta}_0(\beta_0), \overline{\Omega}(\beta_0)) = \mp 1$.

(b) $\lim_{\beta_0 \to \pm\infty} d_{\overline{\beta}_*(\beta_0)}(\overline{\beta}_0(\beta_0), \overline{\Omega}(\beta_0)) = \mp \frac{\rho_{uv}}{(1-\rho_{uv}^2)^{1/2}}$.

Comment. (i). By Lemmas 14.1 and 20.1, the distributions of all of the tests considered in this paper are the same in the model in Section 2 when β_* and Ω are fixed and the null hypothesis value β_0 satisfies $\beta_0 \to \pm\infty$, and in the transformed model of this section when the null hypothesis $\overline{\beta}_0$ is fixed at 0 and the alternative hypothesis value $\overline{\beta}_* = \overline{\beta}_*(\beta_0)$ and the reduced-form variance $\overline{\Omega} = \overline{\Omega}(\beta_0)$ converge as in (20.5) as $\beta_0 \to \pm\infty$. (This uses the fact that $\sigma_v = 1$ in Lemma 20.1.)

(ii). AMS footnote 5 notes that there is a special parameter value $\beta = \beta_{AR}$ at which the one-sided point optimal invariant similar test of $H_0 : \beta = \beta_0$ versus $H_1 : \beta = \beta_{AR}$ is the (two-sided) AR test. In footnote 5 β_{AR} is defined to be $\beta_{AR} = \frac{\omega_1^2 - \omega_{12}\beta_0}{\omega_{12} - \omega_2^2\beta_0}$. If we compute β_{AR} for the transformed model $(\overline{y}_1, \overline{y}_2)$ with parameters $(\overline{\beta}, \overline{\pi}, \overline{\Omega})$, where $\overline{\beta}_0 = 0$, we obtain

$$\overline{\beta}_{AR} = \frac{\overline{\omega}_1^2 - \overline{\omega}_{12}\overline{\beta}_0}{\overline{\omega}_{12} - \overline{\omega}_2^2\overline{\beta}_0} = \frac{1}{\overline{\omega}_{12}} = \mp 1, \tag{20.6}$$

which is the same as the limit of $\overline{\beta}_* = \overline{\beta}_*(\beta_0)$ as $\beta_0 \to \pm\infty$ in (20.2).

Proof of Lemma 20.1. First, we prove part (a). We have

$$\begin{aligned} c_{\overline{\beta}_*}(\overline{\beta}_0, \overline{\Omega}) &= (\overline{\beta}_* - \overline{\beta}_0)(\overline{b}_0'\overline{\Omega}\overline{b}_0)^{-1/2} \\ &= \frac{\omega_2}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}}(\beta_* - \beta_0)(1 - 2\overline{\omega}_{12}\overline{\beta}_0 + \overline{\beta}_0^2)^{-1/2} \\ &= \frac{\omega_2(\beta_* - \beta_0)}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}} \\ &\to \mp 1 \text{ as } \beta_0 \to \pm\infty, \end{aligned} \tag{20.7}$$

where the second equality uses (20.4) and the third equality uses $\overline{\beta}_0 = 0$. Hence, $c_{\overline{\beta}_*}(\overline{\beta}_0, \overline{\Omega})\mu_{\overline{\pi}} \to \mp(1/\sigma_v)\mu_\pi$ as $\beta_0 \to \pm\infty$ using the expression for $\overline{\pi}$ in (20.2) and $\omega_2 = \sigma_v$.

Next, we prove part (b). Let $\overline{b}_* = (1, -\overline{\beta}_*)'$ and $\overline{b}_0 = (1, -\overline{\beta}_0)'$. We have

$$
\begin{aligned}
\det(\overline{\Omega}) &= 1 - \overline{\omega}_{12}^2, \\
\overline{\omega}_{12} &= \frac{\omega_{12} - \omega_2^2\beta_0}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}\omega_2}, \text{ and} \\
\overline{b}_*'\overline{\Omega}\overline{b}_0(\overline{b}_0'\overline{\Omega}\overline{b}_0)^{-1/2} &= \frac{1 - \overline{\omega}_{12}\overline{\beta}_0 - \overline{\omega}_{12}\overline{\beta}_* + \overline{\beta}_0\overline{\beta}_*}{(1 - 2\overline{\omega}_{12}\overline{\beta}_0 + \overline{\beta}_0^2)^{1/2}} = 1 - \overline{\omega}_{12}\overline{\beta}_*,
\end{aligned} \tag{20.8}
$$

where the second equality on the third line uses $\overline{\beta}_0 = 0$. Next, we have

$$
\begin{aligned}
1 - \overline{\omega}_{12}\overline{\beta}_* &= 1 - \frac{\omega_{12} - \omega_2^2\beta_0}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}\omega_2} \frac{\omega_2}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}}(\beta_* - \beta_0) \\
&= 1 - \frac{(\omega_{12} - \omega_2^2\beta_0)(\beta_* - \beta_0)}{\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2} \\
&= \frac{\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2 - \omega_{12}\beta_* + \omega_{12}\beta_0 + \omega_2^2\beta_0\beta_* - \omega_2^2\beta_0^2}{\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2} \\
&= \frac{\omega_1^2 - \omega_{12}\beta_0 - \omega_{12}\beta_* + \omega_2^2\beta_0\beta_*}{\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2},
\end{aligned} \tag{20.9}
$$

where the first equality uses (20.3) and (20.4).

In addition, we have

$$
\begin{aligned}
1 - \overline{\omega}_{12}^2 &= 1 - \frac{(\omega_{12} - \omega_2^2\beta_0)^2}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)\omega_2^2} \\
&= \frac{\omega_1^2\omega_2^2 - 2\omega_{12}\omega_2^2\beta_0 + \omega_2^4\beta_0^2 - \omega_{12}^2 + 2\omega_{12}\omega_2^2\beta_0 - \omega^4\beta_0^2}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)\omega_2^2} \\
&= \frac{\omega_1^2\omega_2^2 - \omega_{12}^2}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)\omega_2^2},
\end{aligned} \tag{20.10}
$$

where the first equality uses (20.8).

Using (20.8)-(20.10), we have

$$
\begin{aligned}
d_{\overline{\beta}_*}(\overline{\beta}_0, \overline{\Omega}) &= \overline{b}_*'\overline{\Omega}\overline{b}_0(\overline{b}_0'\overline{\Omega}\overline{b}_0)^{-1/2} \det(\overline{\Omega})^{-1/2} \\
&= \frac{\omega_1^2 - \omega_{12}\beta_0 - \omega_{12}\beta_* + \omega_2^2\beta_0\beta_*}{\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2} \left(\frac{\omega_1^2\omega_2^2 - \omega_{12}^2}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)\omega_2^2} \right)^{-1/2} \\
&= \frac{(\omega_1^2 - \omega_{12}\beta_0 - \omega_{12}\beta_* + \omega_2^2\beta_0\beta_*)}{(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{1/2}(\omega_1^2\omega_2^2 - \omega_{12}^2)^{1/2}}\omega_2.
\end{aligned} \tag{20.11}
$$

The rhs of (20.11) is the same as the expression on the second line of (14.2) multiplied by $\omega_2 = \sigma_v$. In consequence, the calculations in (14.2)-(14.4) give the result of part (a) of Lemma 20.1. $\square$

21 Transformation of the $\boldsymbol{\beta_0}$ Versus $\boldsymbol{\beta_*}$ Testing Problem to a $\overline{\boldsymbol{\beta}}_{\mathbf{0}}$ Versus 0 Testing Problem

In this section, we transform the general testing problem of $H_0 : \beta = \beta_0$ versus $H_1 : \beta = \beta_*$ for $\pi \in R^k$ and fixed reduced-form variance matrix Ω to a testing problem of $H_0 : \overline{\beta} = \overline{\beta}_0$ versus $H_1 : \overline{\beta} = 0$ for some $\overline{\pi} \in R^k$ and some fixed $\overline{\Omega}$ with diagonal elements equal to one. These transformation results imply that there is no loss in generality in the numerical results of the paper to taking $\omega_1^2 = \omega_2^2 = 1$ and $\beta_* = 0$. We also show that there is no loss in generality in the numerical results of the paper to taking $\rho_{uv} \in [0, 1]$, rather than $\rho_{uv} \in [-1, 1]$, where ρ_{uv} is the structural variance matrix correlation defined in (4.5).

We consider the same transformations as in Section 20, but with β_* in place of β_0 in (20.1)-(20.3) and with the roles of β_* and β_0 reversed in (20.4) and (20.5). The transformed testing problem given the transformations in (20.1) (with β_* in place of β_0) is $H_0 : \widetilde{\beta} = \widetilde{\beta}_0$ versus $H_1 : \widetilde{\beta} = 0$, where $\widetilde{\beta}_0 = \beta_0 - \beta_*$, with parameter π and reduced-form variance matrix $\widetilde{\Omega}$. The transformed testing problem given the transformations in (20.1)-(20.3) (with β_* in place of β_0) is $H_0 : \overline{\beta} = \overline{\beta}_0$ versus $H_1 : \overline{\beta} = 0$, where $\widetilde{\beta}_0 = \beta_0 - \beta_*$, with parameters $\overline{\beta}$, $\overline{\pi}$, and $\overline{\Omega}$ defined in (20.2) and (20.3) (with the roles of β_* and β_0 reversed).

For example, a scenario in which a typical test has high power in the original scenario of testing $H_0 : \beta = \beta_0$ versus $H_1 : \beta = \beta_*$, such as $\beta_0 = 0$ and $|\beta_*|$ large, gets transformed into the testing problem of $H_0 : \overline{\beta} = \overline{\beta}_0$ versus $H_1 : \overline{\beta} = 0$ with correlation $\overline{\omega}_{12}$ (the $(1, 2)$ element of $\overline{\Omega}$) close to ± 1, because by (20.5) (with the roles of β_* and β_0 reversed) we have

$$\lim_{\beta_* \to \pm\infty} \overline{\Omega} = \begin{bmatrix} 1 & \mp 1 \\ \mp 1 & 1 \end{bmatrix}. \tag{21.1}$$

In this case, we also have $\lim_{\beta_* \to \pm\infty} \overline{\beta}_0 = \mp 1$ by (20.5). Also, note that the reduced-form and structural variances matrices are equal when the alternative hypothesis holds in the testing problem $H_0 : \overline{\beta} = \overline{\beta}_0$ versus $H_1 : \overline{\beta} = 0$, so the result in (21.1) also applies to the structural variance matrix $\Sigma(\overline{\beta}, \overline{\Omega})$ when $\overline{\beta} = 0$ whose correlation we denote by $\overline{\rho}_{uv}$, i.e., $\lim_{\beta_* \to \pm\infty} \overline{\rho}_{uv} = \mp 1$. Here the parameter $\overline{\rho}_{uv}$ is the parameter ρ_{uv} that appears in the tables in the paper. These results are useful in showing how the numerical results of the paper apply to general hypotheses of the form $H_0 : \beta = \beta_0$ versus $H_1 : \beta = \beta_*$.

Next, we show that there is no loss in generality in the numerical results of the paper to taking $\rho_{uv} \in [0, 1]$. We consider the hypotheses $H_0 : \beta = \beta_0$ versus $H_1 : \beta = 0$, as in the numerical results in the paper. When the true β equals 0 and Ω has ones on its diagonal, the reduced-form and structural variance matrices are equal, see (11.9). Hence, the correlation ω_{12} given by Ω equals the structural variance correlation ρ_{uv} in power calculations in the paper, and it suffices to show that there is no loss in generality in the numerical results of the paper to taking $\omega_{12} \in [0, 1]$.

By (2.3), the distributions of S and T only depend on $c_\beta(\beta_0, \Omega)$, $d_\beta(\beta_0, \Omega)$, and $\mu_\pi := (Z'Z)^{1/2}\pi$. The vector μ_π does not depend on β, β_0, or Ω. First, note that ω_{12} enters $c_\beta(\beta_0, \Omega) := (\beta - \beta_0)(b_0'\Omega b_0)^{-1/2} = (\beta - \beta_0)(\omega_1^2 - 2\omega_{12}\beta_0 + \omega_2^2\beta_0^2)^{-1/2}$ only through $\omega_{12}\beta_0$. In consequence, the distribution of S is the same under (β_0, ω_{12}) as under $(-\beta_0, -\omega_{12})$. Second, by (2.8) of AMS, $d_\beta(\beta_0, \Omega)$ can be written as $b'\Omega b_0(b_0'\Omega b_0)^{-1/2}\det(\Omega)^{-1/2}$, where $b := (1, -\beta)'$. The distribution of T when $\beta = 0$ depends on $d_0(\beta_0, \Omega) = (1 - \omega_{12}\beta_0)(b_0'\Omega b_0)^{-1/2}\det(\Omega)^{-1/2}$. The first two multiplicands depend on ω_{12} only through $\omega_{12}\beta_0$ and the third multiplicand only depends on ω_{12} through ω_{12}^2 (because $\det(\Omega) = 1 - \omega_{12}^2$). In addition, S and T are independent. Hence, the distribution of $[S : T]$ for given (β_0, ω_{12}) when $\beta = 0$ equals its distribution under $(-\beta_0, -\omega_{12})$ when $\beta = 0$. Thus, the power of a test of $H_0 : \beta = \beta_0$ versus $H_1 : \beta = 0$ when $\omega_{12} < 0$ equals its power for testing $H_0 : \beta = -\beta_0$ versus $H_1 : \beta = 0$ for $-\omega_{12} > 0$.

22 Additional Numerical Results

This section contains Tables SM-I,..., SM-V and Figure SM-I, which provide additional numerical results to those in Tables I,..., V in the main paper.

Table SM-I is analogous to Table I, but considers AR and POIS$_\infty$ CS's, in addition to the CLR and POIS2$_\infty$ CS's. Here, POIS$_\infty$ denotes the CS obtained from the optimal one-sided invariant similar test as $\beta_0 \to \pm\infty$ defined in (12.5) in Section 12.2. Table SM-I reports probabilities of infinite length, as well as differences in probabilities of infinite length (DPIL's) for CLR and AR, AR and POIS2$_\infty$, and CLR and POIS2$_\infty$ CS's. In addition, it reports simulation standard deviations for the first and third DPIL's.

The results for the DPIL's vary greatly with ρ_{uv}. When $\rho_{uv} = 0$, the AR CS is the same as the optimal POIS$_\infty$ CS and the CLR-AR DPIL's range over $[.001, .049]$ as (k, λ) vary. On the other hand, when $\rho_{uv} = .9$, the AR CS is far from optimal and the CLR-AR DPIL's range over $[-.002, -.421]$ as (k, λ) vary. In sum, when $\rho_{uv} \geq .5$, the AR CS can, and typically does, perform noticeably worse than the CLR CS in terms of DPIL's.

Table SM-II reports more detailed results than those given in Table II. Table SM-II reports the maximum power differences (PD's) over β_0 values between the POIS2 power envelope and the CLR test for a grid of (k, ρ_{uv}, λ) values. (In contrast, Table II reports maximum and average PD's over (β_0, λ) values for a grid of (k, ρ_{uv}) values.) Table SM-II shows that the maximum (over β_0) PD's vary substantially over λ values for $\rho_{uv} \leq .7$ values and less so for $\rho_{uv} = .9$. For example, for $k = 5$ and $\rho_{uv} = .0, .3, .5, .7, .9$, the PD's ranges (over λ values) are $[.004, .030]$, $[.008, .034]$, $[.007, .029]$, $[.005, .033]$, $[.001, .017]$, respectively.

Tables SM-III and SM-IV are the same as Table II except they consider the AR and LM tests, respectively, rather than the CLR test. As noted in the main paper, Tables SM-III and SM-IV show that the power of the AR and LM tests is much farther from the POIS2 power envelope than is the power of the CLR test. Table SM-III(a) shows that the maximum and average (over (β_0, λ)) PD's for the AR test are increasing in k up to $k = 20$, but drop going from $k = 20$ to 40. (This drop may be due to the choice of λ values considered. The choice yields the $\lambda_{\max}$ value for the AR test to be on the upper bound of the values considered.) Table SM-III(b) shows that the maximum and average (over (β_0, λ)) PD's for the AR test are increasing in ρ_{uv} for all values of k, which is the opposite of the pattern for the CLR test. Table SM-III(b) also shows that the $\lambda_{\max}$ values are at the boundary of the grid of λ values considered for all k.

Table SM-IV(a) shows that the maximum and average PD's (over (β_0, λ)) for the LM test are clearly increasing in k, except that for $\rho_{uv} = 0, .3$ there is a drop from $k = 20$ to 40 (which may be due to the choice of λ values considered, as for the AR test). Table SM-IV(b) shows that the maximum and average (over (β_0, λ)) PD's for the LM test are decreasing in ρ_{uv} for all values of k, as for the CLR test. Table SM-IV(b) also shows that the $\lambda_{\max}$ values decrease in ρ_{uv} for each k, as with the CLR test.

Table SM-V is the same as Table III except that it reports results for $k = 2, 5, 10, 20$, and 40, rather than just $k = 5$. It also reports results for a finer grid of β_0 values than in Table IV and it reports the power of the WAP2 test, in addition to the difference in power between the WAP2 and CLR tests.

Figure SM-I provides graphs that are the same as in AMS, but with $\rho_\Omega = 0$, rather than $\rho_\Omega = .5$ or .95. Specifically, these graphs provide the power of the significance level .05 CLR, LM, and AR tests and the POIS2 power envelope for fixed null value $\beta_0 = 0$, varying true value β^*, $k = 2, 5, 10$ and $\lambda = 5, 20$. The number of simulation repetitions used to construct the power functions is 5,000 and 100,000 repetitions are used to compute the null distribution of the POIS2 statistic to obtain its p-values.

Figure SM-I shows that the power of the CLR test is very close to the POIS2 power envelope in the scenarios considered. In fact, the maximum differences are .0074, .0040, .0110, .0062, .0102, and .0090 in the six graphs in Figure SM-I. Note that $\rho_\Omega = 0$ is the ρ_Ω value that yields many of the largest differences between the power of the CLR test and the POIS2 power envelope when the true $\beta^* = 0$ is fixed and the null value β_0 varies, as shown in Tables II and SM-II.[13] The results in Figure SM-I show that standard power graphs with $\beta_0 = 0$ fixed and true β^* varying, as in AMS, do not pick up the relatively large differences between the power of the CLR test and the POIS2 power envelope that appear in some $\rho_\Omega = 0$ parameter configurations considered in Tables II and SM-II.

[13] In Tables II and SM-II, $\rho_{uv} = \rho_\Omega$ for all β_0 values because the true value $\beta^* = 0$.

References

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.

——— (2006): "Optimal Two-sided Invariant Similar Tests for Instrumental Variables Regression," *Econometrica,* 74, 715–752.

Lebedev, N. N. (1965): *Special Functions and Their Application.* Englewood Cliffs, NJ: Prentice-Hall.

Mills, B., M. J. Moreira, and L. P. Vilela (2014): "Tests Based on t-Statistics for IV Regression with Weak Instruments," *Journal of Econometrics,* 182, 351–363.

Moreira, M. J. (2009): "Tests with Correct Size When Instruments Can Be Arbitrarily Weak," *Journal of Econometrics,* 152, 131-140.

TABLE SM-I(a). Probabilities of Infinite-Length Confidence Intervals for $\rho_{uv} = 0$

k	λ	AR	CLR	POIS_∞	POIS2_∞	CLR–AR	SD_1	AR–POIS2_∞	CLR–POIS2_∞	SD_2
2	1	.870	.872	.870	.870	.002	.0007	.000	.002	.0007
2	3	.680	.687	.680	.680	.007	.0009	.000	.007	.0009
2	5	.497	.509	.497	.497	.012	.0010	.000	.012	.0010
2	7	.344	.360	.344	.344	.016	.0010	.000	.016	.0010
2	10	.184	.196	.184	.184	.013	.0008	.000	.013	.0008
2	15	.056	.063	.056	.056	.008	.0006	.000	.008	.0006
2	20	.014	.018	.014	.014	.003	.0003	.000	.003	.0003
5	1	.900	.904	.900	.900	.003	.0008	.000	.003	.0008
5	3	.778	.788	.778	.778	.010	.0011	.000	.010	.0011
5	5	.639	.656	.639	.639	.017	.0012	.000	.017	.0012
5	7	.502	.528	.502	.502	.026	.0012	.000	.026	.0012
5	10	.323	.354	.323	.323	.031	.0012	.000	.031	.0012
5	12	.231	.260	.231	.231	.029	.0011	.000	.029	.0011
5	15	.134	.157	.134	.134	.023	.0009	.000	.023	.0009
5	20	.048	.061	.048	.048	.013	.0007	.000	.013	.0007
5	25	.016	.022	.016	.016	.006	.0004	.000	.006	.0004
10	1	.919	.920	.919	.919	.001	.0008	.000	.001	.0008
10	5	.731	.749	.731	.731	.018	.0012	.000	.018	.0012
10	10	.459	.491	.459	.459	.032	.0013	.000	.032	.0013
10	15	.239	.276	.239	.239	.037	.0012	.000	.037	.0012
10	17	.176	.212	.176	.176	.035	.0011	.000	.035	.0011
10	20	.110	.137	.110	.110	.027	.0010	.000	.027	.0010
10	25	.045	.062	.045	.045	.017	.0007	.000	.017	.0007
10	30	.017	.026	.017	.017	.009	.0005	.000	.009	.0005
20	1	.929	.931	.929	.929	.001	.0009	.000	.001	.0009
20	5	.809	.823	.809	.809	.014	.0012	.000	.014	.0012
20	10	.603	.637	.603	.603	.035	.0014	.000	.035	.0014
20	15	.390	.436	.390	.390	.046	.0014	.000	.046	.0014
20	20	.226	.270	.226	.226	.044	.0012	.000	.044	.0012
20	25	.120	.153	.120	.120	.033	.0010	.000	.033	.0010
20	30	.056	.080	.056	.056	.023	.0008	.000	.023	.0008
20	40	.010	.017	.010	.010	.007	.0004	.000	.007	.0004
40	1	.937	.939	.937	.937	.002	.0008	.000	.002	.0008
40	5	.859	.870	.859	.859	.011	.0011	.000	.011	.0011
40	10	.721	.748	.721	.721	.028	.0013	.000	.028	.0013
40	15	.555	.598	.555	.555	.043	.0014	.000	.043	.0014
40	20	.394	.443	.394	.394	.049	.0014	.000	.049	.0014
40	30	.155	.198	.155	.155	.043	.0012	.000	.043	.0012
40	40	.046	.069	.046	.046	.022	.0008	.000	.022	.0008
40	60	.003	.005	.003	.003	.002	.0002	.000	.002	.0002

TABLE SM-I(b). Probabilities of Infinite-Length Confidence Intervals for $\rho_{uv} = .3$

k	λ	AR	CLR	POIS_∞	POIS2_∞	CLR–AR	SD_1	AR–POIS2_∞	CLR–POIS2_∞	SD_2
2	1	.869	.871	.863	.867	.002	.0007	.003	.004	.0007
2	3	.680	.682	.654	.679	.002	.0010	.001	.003	.0008
2	5	.497	.502	.463	.494	.005	.0010	.002	.008	.0009
2	7	.346	.349	.306	.344	.004	.0011	.002	.005	.0008
2	10	.185	.188	.152	.181	.003	.0009	.003	.006	.0007
2	15	.056	.057	.040	.053	.001	.0006	.002	.004	.0005
2	20	.015	.015	.009	.014	.000	.0004	.001	.001	.0003
5	1	.901	.903	.896	.900	.002	.0008	.000	.003	.0008
5	3	.778	.783	.759	.774	.005	.0011	.005	.010	.0010
5	5	.639	.649	.608	.636	.010	.0013	.003	.012	.0010
5	7	.502	.513	.462	.501	.011	.0013	.001	.012	.0010
5	10	.322	.332	.280	.316	.009	.0013	.006	.016	.0010
5	12	.231	.239	.192	.228	.008	.0012	.003	.011	.0009
5	15	.135	.139	.103	.126	.005	.0010	.008	.013	.0008
5	20	.048	.050	.033	.044	.002	.0007	.005	.007	.0005
5	25	.016	.016	.009	.014	.001	.0005	.002	.002	.0003
10	1	.918	.920	.914	.917	.002	.0009	.000	.002	.0008
10	5	.730	.740	.705	.726	.010	.0013	.005	.015	.0011
10	10	.459	.469	.412	.455	.010	.0014	.003	.014	.0011
10	15	.238	.248	.196	.228	.009	.0013	.010	.020	.0010
10	17	.177	.184	.138	.168	.007	.0012	.009	.017	.0009
10	20	.108	.114	.080	.100	.006	.0010	.008	.014	.0007
10	25	.045	.047	.028	.039	.002	.0007	.005	.008	.0005
10	30	.017	.018	.010	.014	.001	.0005	.003	.004	.0004
20	1	.929	.931	.928	.927	.002	.0008	.001	.003	.0008
20	5	.808	.817	.791	.806	.009	.0012	.002	.011	.0011
20	10	.602	.619	.562	.597	.016	.0015	.005	.022	.0012
20	15	.393	.406	.340	.381	.012	.0015	.013	.025	.0012
20	20	.227	.237	.180	.216	.009	.0013	.012	.021	.0010
20	25	.119	.126	.086	.110	.007	.0011	.009	.016	.0008
20	30	.056	.059	.036	.047	.003	.0009	.009	.012	.0006
20	40	.010	.011	.005	.007	.001	.0004	.003	.003	.0003
40	1	.938	.940	.935	.939	.001	.0008	-.000	.001	.0008
40	5	.860	.867	.848	.858	.007	.0011	.001	.008	.0010
40	10	.718	.735	.691	.715	.017	.0014	.002	.020	.0012
40	15	.554	.570	.509	.547	.016	.0015	.007	.022	.0012
40	20	.392	.409	.337	.380	.017	.0016	.012	.029	.0012
40	30	.154	.163	.114	.143	.009	.0013	.012	.021	.0009
40	40	.045	.049	.028	.037	.004	.0008	.008	.012	.0006
40	60	.002	.003	.001	.001	.001	.0002	.001	.001	.0002

TABLE SM-I(c). Probabilities of Infinite-Length Confidence Intervals for $\rho_{uv} = .5$

k	λ	AR	CLR	POIS_∞	POIS2_∞	CLR–AR	SD_1	AR–POIS2_∞	CLR–POIS2_∞	SD_2
2	1	.868	.869	.847	.865	.000	.0008	.003	.003	.0006
2	3	.678	.674	.613	.672	-.004	.0010	.006	.002	.0006
2	5	.496	.486	.408	.483	-.010	.0012	.013	.003	.0006
2	7	.345	.330	.255	.329	-.015	.0012	.016	.001	.0006
2	10	.186	.168	.113	.166	-.018	.0010	.020	.002	.0005
2	15	.056	.047	.025	.046	-.008	.0007	.010	.002	.0003
2	20	.015	.012	.005	.011	-.003	.0004	.004	.001	.0002
5	1	.901	.902	.888	.900	.001	.0008	.001	.002	.0007
5	3	.779	.775	.726	.769	-.004	.0012	.010	.006	.0007
5	5	.641	.626	.552	.622	-.015	.0014	.019	.004	.0007
5	7	.501	.478	.394	.474	-.023	.0015	.026	.003	.0007
5	10	.322	.291	.215	.286	-.031	.0015	.036	.005	.0007
5	12	.233	.199	.136	.196	-.034	.0014	.037	.003	.0007
5	15	.135	.106	.065	.101	-.029	.0012	.034	.005	.0006
5	20	.048	.034	.017	.030	-.014	.0008	.018	.003	.0004
5	25	.016	.010	.004	.008	-.006	.0005	.007	.001	.0002
10	1	.919	.919	.907	.917	.000	.0009	.002	.002	.0008
10	5	.730	.720	.659	.712	-.010	.0014	.018	.008	.0008
10	10	.457	.420	.333	.417	-.037	.0016	.040	.003	.0008
10	15	.240	.194	.131	.185	-.045	.0015	.055	.009	.0007
10	17	.176	.136	.087	.129	-.040	.0014	.048	.008	.0007
10	20	.109	.077	.044	.071	-.032	.0012	.039	.006	.0006
10	25	.045	.027	.012	.023	-.018	.0008	.021	.003	.0004
10	30	.016	.008	.003	.006	-.008	.0005	.010	.002	.0002
20	1	.928	.929	.923	.927	.001	.0008	.002	.002	.0008
20	5	.808	.804	.758	.798	-.004	.0013	.009	.006	.0008
20	10	.601	.576	.484	.565	-.025	.0016	.036	.011	.0009
20	15	.393	.340	.255	.329	-.053	.0017	.063	.011	.0009
20	20	.228	.176	.113	.165	-.053	.0016	.064	.011	.0008
20	25	.118	.078	.044	.072	-.040	.0013	.046	.006	.0006
20	30	.057	.031	.014	.026	-.026	.0009	.031	.005	.0004
20	40	.010	.004	.002	.003	-.006	.0004	.007	.001	.0002
40	1	.938	.939	.932	.938	.001	.0008	.000	.001	.0008
40	5	.861	.859	.824	.855	-.001	.0012	.006	.005	.0008
40	10	.717	.705	.627	.696	-.012	.0016	.021	.009	.0009
40	15	.552	.513	.419	.502	-.038	.0018	.050	.011	.0009
40	20	.390	.334	.244	.321	-.056	.0018	.069	.013	.0009
40	30	.155	.103	.059	.092	-.052	.0014	.062	.011	.0007
40	40	.045	.022	.010	.017	-.023	.0009	.028	.005	.0004
40	60	.002	.001	.000	.000	-.001	.0002	.001	.000	.0001

TABLE SM-I(d). Probabilities of Infinite-Length Confidence Intervals for $\rho_{uv} = .7$

k	λ	AR	CLR	POIS_∞	POIS2_∞	CLR–AR	SD_1	AR–POIS2_∞	CLR–POIS2_∞	SD_2
2	1	.867	.865	.822	.863	-.002	.0008	.004	.002	.0003
2	3	.678	.657	.557	.656	-.021	.0012	.022	.001	.0005
2	5	.495	.457	.345	.455	-.038	.0013	.040	.002	.0005
2	7	.346	.299	.202	.298	-.047	.0013	.048	.001	.0005
2	10	.187	.143	.084	.142	-.043	.0012	.044	.001	.0004
2	15	.057	.037	.017	.035	-.020	.0008	.021	.001	.0003
2	20	.015	.008	.003	.008	-.007	.0004	.007	.000	.0002
5	1	.900	.899	.871	.898	-.001	.0009	.003	.002	.0004
5	3	.778	.752	.667	.751	-.026	.0014	.028	.002	.0006
5	5	.640	.578	.463	.576	-.062	.0016	.064	.003	.0008
5	7	.499	.409	.301	.407	-.091	.0018	.092	.001	.0007
5	10	.323	.224	.143	.221	-.099	.0017	.102	.004	.0006
5	12	.232	.140	.082	.140	-.092	.0016	.093	.000	.0005
5	15	.136	.065	.034	.062	-.070	.0013	.073	.003	.0004
5	20	.049	.017	.007	.016	-.031	.0009	.032	.001	.0003
5	25	.015	.004	.001	.004	-.011	.0005	.012	.000	.0001
10	1	.918	.917	.895	.916	-.001	.0009	.002	.001	.0005
10	5	.730	.673	.572	.667	-.057	.0017	.063	.005	.0008
10	10	.456	.325	.223	.320	-.131	.0019	.136	.005	.0008
10	15	.240	.115	.066	.111	-.124	.0017	.129	.004	.0006
10	17	.177	.072	.038	.069	-.104	.0016	.107	.003	.0005
10	20	.110	.035	.016	.032	-.076	.0013	.078	.002	.0004
10	25	.045	.009	.003	.008	-.036	.0009	.037	.001	.0002
10	30	.016	.002	.001	.002	-.014	.0005	.014	.000	.0001
20	1	.928	.927	.915	.926	-.001	.0009	.003	.002	.0006
20	5	.808	.769	.684	.766	-.039	.0015	.042	.003	.0008
20	10	.600	.472	.354	.463	-.127	.0020	.137	.010	.0009
20	15	.391	.223	.139	.215	-.168	.0020	.176	.008	.0008
20	20	.227	.085	.044	.080	-.142	.0018	.147	.005	.0006
20	25	.117	.027	.012	.025	-.090	.0014	.092	.002	.0004
20	30	.056	.008	.003	.007	-.048	.0010	.050	.001	.0002
20	40	.011	.001	.000	.001	-.010	.0005	.010	.000	.0001
40	1	.937	.938	.926	.937	.001	.0009	-.000	.001	.0007
40	5	.860	.839	.772	.838	-.021	.0014	.022	.001	.0007
40	10	.719	.623	.504	.614	-.095	.0019	.104	.009	.0010
40	15	.550	.382	.268	.372	-.168	.0021	.178	.009	.0009
40	20	.390	.195	.119	.186	-.195	.0021	.204	.009	.0008
40	30	.155	.032	.014	.028	-.123	.0016	.127	.004	.0004
40	40	.045	.004	.001	.003	-.041	.0009	.043	.001	.0002
40	60	.002	.000	.000	.000	-.002	.0002	.002	.000	.0000

TABLE SM-I(e). Probabilities of Infinite-Length Confidence Intervals for $\rho_{uv} = .9$

k	λ	AR	CLR	POIS$_\infty$	POIS2$_\infty$	CLR–AR	SD$_1$	AR–POIS2$_\infty$	CLR–POIS2$_\infty$	SD$_2$
2	1	.868	.852	.774	.850	-.016	.0010	.018	.002	.0006
2	3	.679	.613	.491	.613	-.065	.0015	.066	.001	.0004
2	5	.494	.411	.295	.410	-.083	.0016	.084	.001	.0003
2	7	.344	.262	.169	.262	-.083	.0015	.082	-.000	.0003
2	10	.183	.122	.069	.122	-.061	.0013	.061	.000	.0002
2	15	.055	.030	.014	.029	-.025	.0008	.025	.001	.0001
2	20	.015	.006	.002	.007	-.008	.0004	.008	-.000	.0001
5	1	.901	.884	.822	.882	-.017	.0012	.019	.002	.0007
5	3	.775	.675	.555	.671	-.101	.0018	.104	.004	.0006
5	5	.636	.465	.345	.466	-.172	.0021	.171	-.001	.0004
5	7	.500	.302	.202	.304	-.198	.0021	.196	-.002	.0004
5	10	.325	.146	.085	.144	-.180	.0019	.181	.002	.0003
5	12	.232	.085	.046	.086	-.147	.0017	.146	-.001	.0003
5	15	.136	.037	.018	.036	-.099	.0014	.099	.001	.0002
5	20	.049	.008	.003	.008	-.041	.0009	.041	.000	.0001
5	25	.015	.002	.000	.001	-.014	.0005	.014	.000	.0001
10	1	.918	.904	.853	.904	-.013	.0011	.013	.000	.0007
10	5	.730	.530	.406	.526	-.201	.0022	.205	.004	.0006
10	10	.456	.177	.106	.176	-.279	.0022	.280	.001	.0004
10	15	.241	.046	.022	.046	-.195	.0018	.195	.000	.0003
10	17	.180	.025	.012	.025	-.154	.0017	.155	.000	.0002
10	20	.110	.010	.004	.010	-.099	.0014	.099	.000	.0002
10	25	.044	.002	.001	.002	-.042	.0009	.042	.000	.0001
10	30	.016	.000	.000	.000	-.015	.0006	.015	-.000	.0000
20	1	.927	.919	.884	.919	-.008	.0010	.009	.000	.0007
20	5	.806	.626	.498	.620	-.180	.0021	.186	.006	.0007
20	10	.599	.247	.159	.246	-.352	.0023	.353	.001	.0004
20	15	.390	.073	.039	.073	-.317	.0022	.318	.001	.0004
20	20	.224	.019	.008	.018	-.205	.0018	.205	.000	.0002
20	25	.116	.004	.002	.004	-.112	.0014	.112	.000	.0001
20	30	.056	.001	.000	.001	-.055	.0010	.055	.000	.0001
20	40	.010	.000	.000	.000	-.010	.0005	.010	.000	.0000
40	1	.936	.934	.905	.934	-.002	.0010	.002	.000	.0006
40	5	.859	.727	.605	.717	-.132	.0020	.142	.009	.0009
40	10	.720	.356	.246	.354	-.364	.0024	.366	.002	.0006
40	15	.551	.130	.074	.129	-.421	.0023	.422	.001	.0005
40	20	.388	.039	.018	.038	-.349	.0022	.350	.001	.0003
40	30	.154	.002	.001	.002	-.152	.0016	.152	.000	.0001
40	40	.046	.000	.000	.000	-.046	.0009	.046	.000	.0000
40	60	.002	.000	.000	.000	-.002	.0002	.002	.000	.0000

TABLE SM-II(a). Maximum Power Differences over λ and β_0 Values between POIS2 and CLR Tests for Fixed Alternative $\beta^* = 0$ for $\rho_{uv} = 0.00$

k	λ	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	POIS2–CLR
2	1	−5.00	.98	.133	.009
2	3	100.00	−1.00	.314	.012
2	5	−10.00	1.00	.499	.019
2	**7**	**−10000.00**	**1.00**	**.663**	**.021**
2	10	7.50	−.99	.814	.018
2	15	10.00	−1.00	.950	.009
2	20	−1.25	.78	.923	.004
5	1	−10.00	1.00	.095	.004
5	3	−5.00	.98	.209	.012
5	5	−1000.00	1.00	.363	.017
5	7	−10.00	1.00	.501	.025
5	**10**	**−50.00**	**1.00**	**.680**	**.030**
5	15	−10000.00	1.00	.870	.025
5	20	−50.00	1.00	.953	.015
10	1	2.00	−.89	.078	.003
10	3	3.75	−.97	.161	.012
10	5	10.00	−1.00	.268	.021
10	7	100.00	−1.00	.379	.028
10	10	−10000.00	1.00	.540	.030
10	**15**	**−50.00**	**1.00**	**.760**	**.038**
10	20	50.00	−1.00	.888	.025
20	1	−2.75	.94	.064	.006
20	3	−100.00	1.00	.116	.007
20	5	100.00	−1.00	.180	.016
20	7	5.00	−.98	.252	.028
20	10	−100.00	1.00	.389	.040
20	**15**	**10.00**	**−1.00**	**.596**	**.042**
20	20	100.00	−1.00	.770	.040
20	22	100.00	−1.00	.820	.038
20	25	−10000.00	1.00	.878	.035
40	1	−0.25	.24	.054	.011
40	3	2.50	−.93	.095	.011
40	5	−7.50	.99	.149	.020
40	7	−50.00	1.00	.201	.024
40	10	−100.00	1.00	.287	.035
40	15	−7.50	.99	.441	.041
40	20	50.00	−1.00	.608	.058
40	**22**	**−50.00**	**1.00**	**.664**	**.059**
40	25	1000.00	−1.00	.742	.056

TABLE SM-II(b). Maximum Power Differences over λ and β_0 Values between POIS2 and CLR Tests for Fixed Alternative $\beta^* = 0$ for $\rho_{uv} = 0.30$

k	λ	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	POIS2–CLR
2	1	10.00	−1.00	.140	.009
2	3	7.50	−.99	.335	.012
2	5	3.50	−.96	.543	.016
2	7	2.25	−.90	.692	.018
2	**10**	**3.75**	**−.96**	**.860**	**.019**
2	15	2.25	−.90	.966	.009
2	20	−1.75	.91	.920	.004
5	1	3.75	−.96	.102	.008
5	3	2.75	−.93	.239	.017
5	5	7.50	−.99	.387	.018
5	7	3.00	−.94	.546	.023
5	**10**	**3.50**	**−.96**	**.732**	**.034**
5	15	2.75	−.93	.901	.023
5	20	4.00	−.97	.971	.014
10	1	3.00	−.94	.090	.003
10	3	3.00	−.94	.181	.014
10	5	2.75	−.93	.296	.023
10	7	3.75	−.96	.417	.026
10	**10**	**3.00**	**−.94**	**.590**	**.032**
10	15	3.50	−.96	.806	.032
10	20	3.50	−.96	.921	.025
20	1	−50.00	1.00	.067	.006
20	3	3.50	−.96	.126	.009
20	5	4.00	−.97	.195	.019
20	7	3.00	−.94	.285	.030
20	10	3.25	−.95	.432	.038
20	**15**	**3.50**	**−.96**	**.655**	**.045**
20	20	2.75	−.93	.817	.038
20	22	3.00	−.94	.863	.034
20	25	3.75	−.96	.915	.031
40	1	−0.25	.50	.051	.008
40	3	1.50	−.78	.097	.008
40	5	2.75	−.93	.153	.014
40	7	5.00	−.98	.215	.020
40	10	4.00	−.97	.312	.036
40	15	3.00	−.94	.485	.042
40	20	3.75	−.96	.663	.059
40	**22**	**4.00**	**−.97**	**.724**	**.061**
40	25	3.75	−.96	.798	.052

TABLE SM-II(c). Maximum Power Differences over λ and β_0 Values between POIS2 and CLR Tests for Fixed Alternative $\beta^* = 0$ for $\rho_{uv} = 0.50$

k	λ	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	POIS2–CLR
2	1	2.00	−.87	.162	.006
2	3	2.50	−.92	.399	.014
2	**5**	**2.00**	**−.87**	**.635**	**.016**
2	7	2.50	−.92	.782	.013
2	10	2.25	−.90	.922	.012
2	15	−10.00	1.00	.943	.004
2	20	−1.25	.90	.804	.003
5	1	1.75	−.82	.112	.007
5	3	2.00	−.87	.291	.019
5	5	2.50	−.92	.475	.022
5	7	2.50	−.92	.638	.028
5	**10**	**2.25**	**−.90**	**.821**	**.029**
5	15	1.75	−.82	.951	.014
5	20	1.00	−.50	.969	.007
10	1	2.00	−.87	.097	.003
10	3	1.75	−.82	.215	.014
10	5	2.00	−.87	.362	.028
10	7	1.75	−.82	.500	.028
10	**10**	**2.00**	**−.87**	**.697**	**.037**
10	15	2.00	−.87	.887	.023
10	20	2.25	−.90	.968	.018
20	1	5.00	−.98	.071	.008
20	3	2.00	−.87	.148	.011
20	5	3.00	−.94	.233	.024
20	7	2.00	−.87	.355	.034
20	**10**	**1.75**	**−.82**	**.533**	**.046**
20	15	2.00	−.87	.769	.040
20	20	2.00	−.87	.905	.031
20	22	2.25	−.90	.934	.025
20	25	2.25	−.90	.963	.014
40	1	−0.25	.65	.051	.007
40	3	1.50	−.76	.117	.008
40	5	2.00	−.87	.184	.014
40	7	3.25	−.95	.256	.026
40	10	2.00	−.87	.381	.035
40	**15**	**1.75**	**−.82**	**.594**	**.050**
40	20	1.75	−.82	.776	.049
40	22	2.00	−.87	.835	.049
40	25	2.00	−.87	.897	.040

TABLE SM-II(d). Maximum Power Differences over λ and β_0 Values between POIS2 and CLR Tests for Fixed Alternative $\beta^* = 0$ for $\rho_{uv} = 0.70$

k	λ	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	POIS2–CLR
2	1	1.75	−.83	.214	.006
2	3	2.25	−.91	.528	.013
2	**5**	**1.50**	**−.75**	**.811**	**.016**
2	7	1.25	−.61	.927	.009
2	10	−2.75	.98	.672	.005
2	15	−4.00	.99	.896	.003
2	20	−1.00	.92	.671	.002
5	1	2.00	−.88	.142	.005
5	3	1.75	−.83	.413	.020
5	**5**	**1.50**	**−.75**	**.669**	**.033**
5	7	1.50	−.75	.834	.029
5	10	1.50	−.75	.949	.016
5	15	7.50	−.99	.971	.003
5	20	−2.50	.98	.897	.003
10	1	1.25	−.61	.117	.004
10	3	1.50	−.75	.320	.028
10	5	1.25	−.61	.521	.030
10	**7**	**1.50**	**−.75**	**.711**	**.036**
10	10	1.50	−.75	.878	.021
10	15	1.75	−.83	.979	.010
10	20	0.50	.27	.757	.007
20	1	2.00	−.88	.087	.007
20	3	1.50	−.75	.212	.025
20	5	1.50	−.75	.377	.038
20	**7**	**1.25**	**−.61**	**.544**	**.042**
20	10	1.50	−.75	.754	.036
20	15	1.50	−.75	.935	.024
20	20	−1.25	.94	.582	.007
20	22	−1.00	.92	.546	.007
20	25	−7.50	1.00	.952	.005
40	1	−100.00	1.00	.071	.006
40	3	2.25	−.91	.158	.015
40	5	1.75	−.83	.275	.030
40	7	1.25	−.61	.393	.038
40	10	1.50	−.75	.588	.049
40	**15**	**1.50**	**−.75**	**.837**	**.050**
40	20	1.50	−.75	.948	.026
40	22	1.50	−.75	.967	.017
40	25	−3.25	.98	.817	.009

TABLE SM-II(e). Maximum Power Differences over λ and β_0 Values between POIS2 and CLR Tests for Fixed Alternative $\beta^* = 0$ for $\rho_{uv} = 0.90$

k	λ	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	POIS2–CLR
2	0.7	1.25	−.63	.359	.010
2	0.8	1.00	−.22	.412	.015
2	**0.9**	**1.25**	**−.63**	**.461**	**.017**
2	1	1.25	−.63	.505	.013
2	3	1.25	−.63	.947	.005
2	5	−7.50	1.00	.497	.002
2	7	−5.00	1.00	.600	.001
2	10	−3.50	1.00	.702	.001
2	15	−50.00	1.00	.967	.001
2	20	−3.75	1.00	.950	.001
5	0.7	1.00	−.22	.256	.013
5	0.8	1.00	−.22	.292	.017
5	**0.9**	**1.00**	**−.22**	**.331**	**.017**
5	1	1.00	−.22	.365	.015
5	3	1.00	−.22	.870	.015
5	5	1.00	−.22	.985	.004
5	7	0.75	.33	.975	.004
5	10	10.00	−1.00	.916	.002
5	15	−10.00	1.00	.934	.001
5	20	−1.75	.99	.815	.001
10	1	1.00	−.22	.273	.017
10	**3**	**1.25**	**−.63**	**.766**	**.027**
10	5	1.25	−.63	.956	.014
10	7	2.00	−.93	.964	.005
10	10	−50.00	1.00	.812	.004
10	15	−1.50	.98	.618	.003
10	20	−2.75	.99	.895	.003
20	1	1.00	−.22	.183	.015
20	**3**	**1.00**	**−.22**	**.607**	**.032**
20	5	1.25	−.63	.882	.022
20	7	5.00	−.99	.705	.008
20	10	3.00	−.98	.951	.005
20	15	−3.25	.99	.769	.003
20	20	−1.75	.99	.766	.003
20	22	−3.75	1.00	.931	.003
20	25	−3.25	.99	.946	.002
40	1	1.00	−.22	.146	.016
40	3	1.00	−.22	.440	.027
40	**5**	**1.25**	**−.63**	**.750**	**.040**
40	7	1.25	−.63	.919	.022
40	10	3.75	−.99	.844	.009
40	15	−3.00	.99	.674	.005
40	20	−2.25	.99	.770	.004
40	22	−2.75	.99	.857	.002
40	25	−0.25	.94	.151	.001

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

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

						POIS2–AR	
ρ_{uv}	k	$\lambda_{\max}$	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	max	average
.0	2	20	0.50	−.45	.49	.079	.012
.0	5	20	0.75	−.60	.66	.151	.014
.0	10	20	−0.75	.60	.57	.183	.015
.0	20	25	0.75	−.60	.57	**.217**	**.016**
.0	40	25	−0.75	.60	.42	.161	.011
.3	2	20	0.50	−.21	.60	.084	.014
.3	5	20	−0.75	.74	.56	.163	.020
.3	10	20	−1.00	.81	.61	.201	.022
.3	20	25	0.50	−.21	.47	**.231**	**.024**
.3	40	25	−1.00	.81	.47	.194	.019
.5	2	20	−0.75	.82	.58	.090	.020
.5	5	20	0.50	.00	.66	.182	.031
.5	10	20	−1.00	.87	.59	.232	.038
.5	20	25	−1.00	.87	.60	**.285**	**.044**
.5	40	25	−1.25	.90	.54	.248	.040
.7	2	20	−0.75	.90	.54	.094	.030
.7	5	20	−1.25	.94	.72	.208	.054
.7	10	20	−1.25	.94	.67	.281	.069
.7	20	25	−1.25	.94	.71	**.361**	.084
.7	40	25	−2.00	.97	.73	.351	**.085**
.9	2	20	−1.00	.97	.64	.105	.033
.9	5	20	−1.25	.98	.70	.237	.068
.9	10	20	−1.50	.98	.75	.340	.133
.9	20	25	−1.25	.98	.76	.455	.160
.9	40	25	−1.75	.99	.82	**.513**	**.179**

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

						POIS2–AR	
k	ρ_{uv}	$\lambda_{\max}$	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	max	average
2	.0	20	0.50	−.45	.49	.079	.012
2	.3	20	0.50	−.21	.60	.084	.014
2	.5	20	−0.75	.82	.58	.090	.020
2	.7	20	−0.75	.90	.54	.094	.030
2	.9	20	−1.00	.97	.64	**.105**	**.033**
5	.0	20	0.75	−.60	.66	.151	.014
5	.3	20	−0.75	.74	.56	.163	.020
5	.5	20	0.50	.00	.66	.182	.031
5	.7	20	−1.25	.94	.72	.208	.054
5	.9	20	−1.25	.98	.70	**.237**	**.068**
10	.0	20	−0.75	.60	.57	.183	.015
10	.3	20	−1.00	.81	.61	.201	.022
10	.5	20	−1.00	.87	.59	.232	.038
10	.7	20	−1.25	.94	.67	.281	.069
10	.9	20	−1.50	.98	.75	**.340**	**.133**
20	.0	25	0.75	−.60	.57	.217	.016
20	.3	25	0.50	−.21	.47	.231	.024
20	.5	25	−1.00	.87	.60	.285	.044
20	.7	25	−1.25	.94	.71	.361	.084
20	.9	25	−1.25	.98	.76	**.455**	**.160**
40	.0	25	−0.75	.60	.42	.161	.011
40	.3	25	−1.00	.81	.47	.194	.019
40	.5	25	−1.25	.90	.54	.248	.040
40	.7	25	−2.00	.97	.73	.351	.085
40	.9	25	−1.75	.99	.82	**.513**	**.179**

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

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

						POIS2–LM	
ρ_{uv}	k	$\lambda_{\max}$	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	max	average
.0	2	15	50.00	−1.00	.95	.312	.117
.0	5	20	1000.00	−1.00	.95	.538	.173
.0	10	20	10000.00	−1.00	.89	.611	.174
.0	20	25	10000.00	−1.00	.88	**.687**	**.203**
.0	40	25	10000.00	−1.00	.74	.621	.173
.3	2	15	3.75	−.96	.97	.309	.103
.3	5	20	3.25	−.95	.97	.536	.155
.3	10	20	3.50	−.96	.92	.628	.161
.3	20	25	3.25	−.95	.91	**.710**	**.187**
.3	40	25	3.25	−.95	.80	.671	.168
.5	2	10	2.00	−.87	.92	.311	.077
.5	5	15	2.00	−.87	.95	.538	.121
.5	10	20	2.00	−.87	.97	.639	.133
.5	20	25	2.00	−.87	.96	.734	**.153**
.5	40	25	2.00	−.87	.90	**.750**	.152
.7	2	7	1.50	−.75	.93	.310	.041
.7	5	10	1.50	−.75	.95	.531	.070
.7	10	15	1.50	−.75	.98	.621	.084
.7	20	20	1.50	−.75	.99	.721	.089
.7	40	22	1.50	−.75	.97	**.784**	**.107**
.9	2	3	1.25	−.63	.95	.242	.010
.9	5	3	1.00	−.22	.87	.422	.016
.9	10	5	1.25	−.63	.96	.474	.023
.9	20	5	1.25	−.63	.88	.562	.021
.9	40	7	1.25	−.63	.92	**.620**	**.031**

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

						POIS2–LM	
k	ρ_{uv}	$\lambda_{\max}$	$\beta_{0,\max}$	$\rho_{uv,0}$	POIS2	max	average
2	.0	15	50.00	−1.00	.95	**.312**	**.117**
2	.3	15	3.75	−.96	.97	.309	.103
2	.5	10	2.00	−.87	.92	.311	.077
2	.7	7	1.50	−.75	.93	.310	.041
2	.9	3	1.25	−.63	.95	.242	.010
5	.0	20	1000.00	−1.00	.95	.538	**.173**
5	.3	20	3.25	−.95	.97	.536	.155
5	.5	15	2.00	−.87	.95	**.538**	.121
5	.7	10	1.50	−.75	.95	.531	.070
5	.9	3	1.00	−.22	.87	.422	.016
10	.0	20	10000.00	−1.00	.89	.611	**.174**
10	.3	20	3.50	−.96	.92	.628	.161
10	.5	20	2.00	−.87	.97	**.639**	.133
10	.7	15	1.50	−.75	.98	.621	.084
10	.9	5	1.25	−.63	.96	.474	.023
20	.0	25	10000.00	−1.00	.88	.687	**.203**
20	.3	25	3.25	−.95	.91	.710	.187
20	.5	25	2.00	−.87	.96	**.734**	.153
20	.7	20	1.50	−.75	.99	.721	.089
20	.9	5	1.25	−.63	.88	.562	.021
40	.0	25	10000.00	−1.00	.74	.621	**.173**
40	.3	25	3.25	−.95	.80	.671	.168
40	.5	25	2.00	−.87	.90	.750	.152
40	.7	22	1.50	−.75	.97	**.784**	.107
40	.9	7	1.25	−.63	.92	.620	.031

TABLE SM-V(a). Average (over λ) Power Differences for $\lambda \in \{2.5, 5.0, ..., 90.0\}$ between the WAP2 and CLR Tests for $k = 2$

β_0	$\rho_{uv,0}$		WAP2					WAP2–CLR				
	$\rho_{uv} = 0$	.9	$\rho_{uv} = 0$	.3	.5	.7	.9	$\rho_{uv} = 0$	.3	.5	.7	.9
−10000.00	1.00	1.00	.946	.950	.951	.955	.958	.003	.002	.001	.001	**.001**
−1000.00	1.00	1.00	.946	.950	.951	.955	.958	.003	.002	.001	.001	**.001**
−100.00	1.00	1.00	.946	.949	.951	.954	.957	.003	.002	.001	.000	**.001**
−50.00	1.00	1.00	.946	.949	.950	.953	.956	.003	.002	.001	.001	**.001**
−10.00	1.00	1.00	.946	.945	.945	.947	.948	.003	.002	.001	.000	**.001**
−7.50	.99	1.00	.945	.944	.943	.944	.945	.003	.002	.001	.000	.000
−5.00	.98	1.00	.944	.941	.938	.938	.937	.003	.001	.001	.001	.000
−4.00	.97	1.00	.943	.938	.935	.933	.931	.003	.001	.001	.001	.000
−3.75	.97	1.00	.942	.937	.933	.931	.929	.003	.001	.001	.001	**.001**
−3.50	.96	1.00	.941	.936	.932	.929	.927	.002	.001	.001	.001	**.001**
−3.25	.96	.99	.940	.934	.930	.926	.923	.002	.001	.001	.001	**.001**
−3.00	.95	.99	.940	.933	.927	.923	.920	.003	.001	.001	.001	.000
−2.75	.94	.99	.938	.930	.925	.920	.916	.003	.001	.001	.000	.000
−2.50	.93	.99	.937	.928	.922	.916	.911	.002	.001	.001	.001	.000
−2.25	.91	.99	.935	.925	.918	.910	.904	.002	.001	.001	.001	.000
−2.00	.89	.99	.932	.920	.912	.904	.896	.002	.002	.001	.001	**.001**
−1.75	.87	.99	.928	.914	.904	.895	.885	.002	.001	.001	.001	.000
−1.50	.83	.98	.921	.905	.893	.881	.870	.001	.001	.001	.001	.000
−1.25	.78	.98	.910	.890	.876	.861	.847	.001	.002	.001	.001	**.001**
−1.00	.71	.97	.895	.866	.845	.827	.809	.002	.002	.001	.001	**.001**
−0.75	.60	.97	.851	.814	.788	.762	.738	.001	.002	**.002**	.001	**.001**
−0.50	.45	.95	.736	.681	.646	.613	.584	.002	.002	**.002**	**.002**	**.001**
−0.25	.24	.94	.366	.329	.310	.293	.279	.002	.002	**.002**	.001	**.001**
0.25	−.24	.83	.366	.410	.447	.488	.544	.001	.001	**.002**	**.002**	**.001**
0.50	−.45	.68	.737	.804	.848	.892	.937	.002	.002	.001	.001	.000
0.75	−.60	.33	.851	.899	.928	.959	.989	.002	.001	.001	.001	.000
1.00	−.71	−.22	.892	.928	.951	.975	.997	.001	.001	.001	.001	.000
1.25	−.78	−.63	.911	.942	.959	.979	.997	.001	.002	.001	.001	.000
1.50	−.83	−.81	.921	.948	.963	.980	.995	.001	.002	**.002**	.001	.000
1.75	−.87	−.89	.928	.950	.965	.979	.993	.002	.002	**.002**	**.002**	.000
2.00	−.89	−.93	.932	.953	.965	.978	.991	.001	.002	**.002**	**.002**	.000
2.25	−.91	−.95	.935	.953	.965	.976	.989	.002	.002	**.002**	.001	.000
2.50	−.93	−.96	.937	.954	.964	.975	.986	.002	**.003**	**.002**	.001	.000
2.75	−.94	−.97	.938	.955	.964	.973	.985	.003	**.003**	**.002**	.001	.000
3.00	−.95	−.98	.939	.955	.963	.973	.983	.002	**.003**	**.002**	.001	.000
3.25	−.96	−.98	.940	.954	.963	.972	.982	.003	**.003**	**.002**	.001	.000
3.50	−.96	−.99	.941	.954	.962	.970	.980	.003	**.003**	**.002**	.000	.000
3.75	−.97	−.99	.942	.955	.962	.970	.979	.003	**.003**	**.002**	.001	.000
4.00	−.97	−.99	.942	.954	.962	.969	.978	.003	.002	**.002**	.001	.000
5.00	−.98	−.99	.944	.954	.960	.967	.974	.003	.002	**.002**	.001	.000
7.50	−.99	−1.00	.945	.953	.957	.964	.970	.003	.002	.001	.001	.000
10.00	−1.00	−1.00	.946	.952	.956	.961	.967	**.004**	.002	**.002**	.001	.000
50.00	−1.00	−1.00	.946	.950	.952	.956	.960	.003	.002	.001	.000	**.001**
100.00	−1.00	−1.00	.946	.950	.952	.955	.959	.003	.002	**.002**	.001	**.001**
1000.00	−1.00	−1.00	.946	.950	.951	.955	.958	.003	.002	.001	.001	**.001**
10000.00	−1.00	−1.00	.946	.950	.951	.955	.958	.003	.002	.001	.001	**.001**

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

β_0	$\rho_{uv,0}$		WAP2					WAP2–CLR				
	$\rho_{uv} = 0$	.9	$\rho_{uv} = 0$	.3	.5	.7	.9	$\rho_{uv} = 0$	.3	.5	.7	.9
-10000.00	1.00	1.00	.923	.924	.929	.939	.953	.005	.002	.001	.001	.000
-1000.00	1.00	1.00	.923	.924	.929	.939	.953	.005	.002	.001	.001	.000
-100.00	1.00	1.00	.923	.923	.929	.939	.952	.005	.002	.001	.001	.000
-50.00	1.00	1.00	.923	.923	.929	.938	.951	**.005**	.003	.001	.001	.000
-10.00	1.00	1.00	.922	.920	.924	.931	.942	.005	.002	.001	.000	.000
-7.50	.99	1.00	.921	.918	.922	.929	.938	.004	.002	.000	.000	.000
-5.00	.98	1.00	.919	.915	.917	.923	.931	.004	.001	.000	.000	.000
-4.00	.97	1.00	.918	.912	.913	.917	.924	.003	.001	.000	-.000	.000
-3.75	.97	1.00	.917	.911	.911	.915	.922	.003	.002	.000	-.000	-.000
-3.50	.96	1.00	.917	.910	.910	.913	.920	.003	.001	-.000	-.000	.000
-3.25	.96	.99	.916	.909	.908	.910	.917	.003	.001	.000	.000	-.000
-3.00	.95	.99	.916	.907	.906	.907	.914	.003	.001	.000	.000	-.000
-2.75	.94	.99	.914	.905	.903	.904	.910	.002	.001	.000	.000	-.000
-2.50	.93	.99	.913	.902	.899	.900	.904	.002	.001	.001	-.000	.000
-2.25	.91	.99	.910	.898	.895	.894	.897	.002	.000	.001	-.000	-.000
-2.00	.89	.99	.907	.893	.888	.887	.888	.002	.001	.000	.001	-.000
-1.75	.87	.99	.903	.886	.880	.877	.877	.001	.001	.001	.000	.000
-1.50	.83	.98	.896	.877	.868	.863	.863	.001	.001	.001	.000	.000
-1.25	.78	.98	.885	.861	.850	.842	.839	.002	.001	.001	.000	.000
-1.00	.71	.97	.865	.836	.820	.808	.800	.001	.000	-.000	-.000	-.000
-0.75	.60	.97	.823	.783	.760	.741	.727	.000	-.000	.001	-.000	-.000
-0.50	.45	.95	.705	.649	.618	.592	.572	-.000	-.000	-.001	-.001	-.000
-0.25	.24	.94	.340	.311	.294	.282	.272	-.001	-.001	-.001	-.000	-.000
0.25	$-.24$	.83	.347	.390	.428	.476	.536	-.000	-.001	-.001	-.000	-.001
0.50	$-.45$	.68	.711	.777	.825	.877	.931	.001	.000	.000	.000	.000
0.75	$-.60$	.33	.827	.873	.908	.944	.985	.000	.001	.001	.001	.000
1.00	$-.71$	$-.22$	.868	.904	.930	.960	.994	.002	.001	.001	.001	.000
1.25	$-.78$	$-.63$	.887	.915	.940	.966	.994	.001	.001	.004	.003	**.001**
1.50	$-.83$	$-.81$	.897	.922	.943	.966	.992	.001	.002	.003	**.003**	.001
1.75	$-.87$	$-.89$	.904	.926	.945	.965	.989	.002	.003	.004	.003	-.000
2.00	$-.89$	$-.93$	.907	.928	.945	.963	.986	.002	.003	**.004**	.002	.000
2.25	$-.91$	$-.95$	.910	.930	.945	.961	.984	.001	.004	.004	.001	.000
2.50	$-.93$	$-.96$	.912	.931	.945	.960	.981	.001	**.005**	.004	.001	-.000
2.75	$-.94$	$-.97$	.914	.931	.944	.958	.979	.003	.005	.004	.001	.000
3.00	$-.95$	$-.98$	.915	.931	.943	.957	.977	.003	.005	.003	.001	.000
3.25	$-.96$	$-.98$	.916	.931	.942	.956	.976	.003	.004	.003	.001	.000
3.50	$-.96$	$-.99$	.917	.931	.942	.955	.974	.003	.005	.003	.001	-.000
3.75	$-.97$	$-.99$	.918	.931	.941	.954	.973	.003	.004	.002	.001	.000
4.00	$-.97$	$-.99$	.919	.931	.940	.954	.972	.004	.005	.002	.001	.000
5.00	$-.98$	$-.99$	.920	.930	.939	.951	.968	.004	.005	.002	.000	.000
7.50	$-.99$	-1.00	.922	.929	.936	.948	.963	.005	.004	.001	.001	.000
10.00	-1.00	-1.00	.922	.928	.935	.946	.960	.005	.003	.001	.001	.000
50.00	-1.00	-1.00	.923	.925	.930	.941	.955	.005	.003	.001	.000	-.000
100.00	-1.00	-1.00	.923	.924	.930	.940	.954	.005	.003	.001	.000	.000
1000.00	-1.00	-1.00	.923	.924	.929	.939	.953	.005	.002	.001	.001	.000
10000.00	-1.00	-1.00	.923	.924	.929	.939	.953	.005	.002	.001	.001	.000

TABLE SM-V(c). Average (over λ) Power Differences for $\lambda \in \{2.5, 5.0, ..., 90.0\}$ between the WAP2 and CLR Tests for $k = 10$

β_0	$\rho_{uv,0}$		WAP2					WAP2–CLR				
	$\rho_{uv} = 0$	.9	$\rho_{uv} = 0$	.3	.5	.7	.9	$\rho_{uv} = 0$	.3	.5	.7	.9
-10000.00	1.00	1.00	.901	.903	.910	.924	.946	**.011**	.006	.003	.001	.000
-1000.00	1.00	1.00	.901	.903	.910	.924	.946	**.011**	.006	.003	.001	.000
-100.00	1.00	1.00	.901	.902	.910	.924	.945	**.011**	.006	.003	.002	.001
-50.00	1.00	1.00	.901	.902	.909	.923	.944	**.011**	.006	.003	.002	.001
-10.00	1.00	1.00	.900	.898	.904	.916	.935	**.011**	.005	.003	.001	.001
-7.50	.99	1.00	.900	.896	.902	.913	.932	.010	.005	.003	.002	.001
-5.00	.98	1.00	.897	.892	.897	.907	.924	.008	.004	.003	.002	.001
-4.00	.97	1.00	.895	.889	.893	.902	.918	.007	.004	.002	.002	.001
-3.75	.97	1.00	.894	.888	.891	.900	.916	.007	.004	.002	.001	.001
-3.50	.96	1.00	.893	.886	.889	.898	.914	.007	.004	.002	.002	.001
-3.25	.96	.99	.892	.884	.888	.896	.911	.006	.003	.003	.002	.001
-3.00	.95	.99	.891	.882	.885	.893	.908	.006	.004	.002	.001	.001
-2.75	.94	.99	.890	.881	.882	.889	.904	.007	.005	.002	.001	.001
-2.50	.93	.99	.888	.878	.879	.886	.899	.006	.004	.002	.002	.001
-2.25	.91	.99	.885	.874	.874	.881	.892	.006	.003	.002	.002	.001
-2.00	.89	.99	.881	.869	.868	.873	.884	.004	.003	.002	.002	.001
-1.75	.87	.99	.876	.862	.860	.863	.872	.005	.003	.003	.002	.001
-1.50	.83	.98	.869	.853	.847	.848	.857	.005	.003	.002	.001	.002
-1.25	.78	.98	.858	.837	.829	.827	.834	.003	.003	.003	.002	.002
-1.00	.71	.97	.838	.809	.799	.794	.796	.003	.003	.002	.002	.002
-0.75	.60	.97	.793	.756	.738	.727	.724	.003	.003	.003	.002	.002
-0.50	.45	.95	.676	.623	.596	.580	.572	.004	.004	.003	.003	**.004**
-0.25	.24	.94	.323	.294	.282	.274	.272	.003	.004	.003	.002	.003
0.25	$-.24$	.83	.326	.370	.408	.460	.530	.004	.004	.004	.003	.002
0.50	$-.45$	.68	.674	.743	.796	.857	.923	.003	.003	.002	.002	.001
0.75	$-.60$	.33	.794	.845	.885	.929	.979	.003	.003	.003	.002	.001
1.00	$-.71$	$-.22$	.838	.878	.910	.948	.989	.003	.004	.003	.003	.002
1.25	$-.78$	$-.63$	.857	.891	.920	.953	.989	.003	.005	.005	**.005**	.002
1.50	$-.83$	$-.81$	.870	.899	.927	.955	.986	.004	.005	.007	**.005**	.001
1.75	$-.87$	$-.89$	.877	.904	.929	.953	.983	.005	.007	**.008**	.004	.001
2.00	$-.89$	$-.93$	.881	.906	.929	.951	.980	.005	.007	**.008**	.004	.001
2.25	$-.91$	$-.95$	.884	.909	.929	.950	.978	.005	.009	**.008**	.003	.001
2.50	$-.93$	$-.96$	.887	.910	.928	.948	.975	.006	.009	**.008**	.003	.001
2.75	$-.94$	$-.97$	.888	.910	.928	.946	.973	.006	.009	**.008**	.003	.001
3.00	$-.95$	$-.98$	.890	.911	.927	.945	.971	.006	**.010**	**.008**	.002	.001
3.25	$-.96$	$-.98$	.892	.911	.925	.944	.969	.007	**.010**	.007	.002	.000
3.50	$-.96$	$-.99$	.893	.912	.925	.943	.968	.007	**.010**	.007	.002	.001
3.75	$-.97$	$-.99$	.893	.912	.924	.942	.967	.007	**.010**	.006	.002	.000
4.00	$-.97$	$-.99$	.894	.912	.923	.942	.966	.008	**.010**	.006	.002	.000
5.00	$-.98$	$-.99$	.896	.910	.921	.939	.962	.009	.009	.005	.002	.001
7.50	$-.99$	-1.00	.899	.908	.918	.934	.957	.009	.008	.004	.002	.000
10.00	-1.00	-1.00	.900	.907	.916	.932	.954	**.011**	.008	.004	.002	.000
50.00	-1.00	-1.00	.901	.903	.912	.926	.948	.010	.006	.003	.002	.001
100.00	-1.00	-1.00	.901	.903	.911	.925	.947	**.011**	.006	.003	.002	.001
1000.00	-1.00	-1.00	.901	.903	.910	.925	.946	**.011**	.006	.003	.001	.000
10000.00	-1.00	-1.00	.901	.903	.910	.924	.946	**.011**	.006	.003	.001	.000

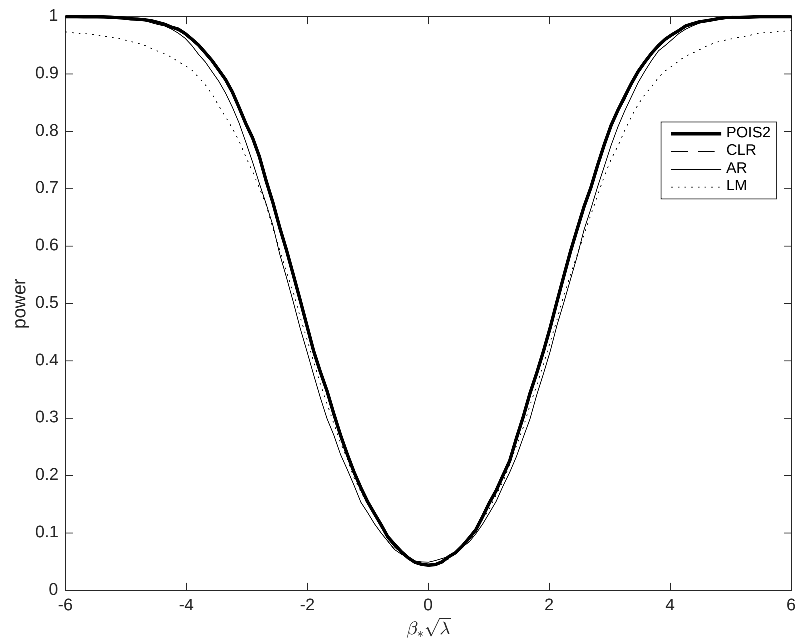
TABLE SM-V(d). Average (over λ) Power Differences for $\lambda \in \{2.5, 5.0, ..., 90.0\}$ between the WAP2 and CLR Tests for $k = 20$

β_0	$\rho_{uv,0}$		WAP2					WAP2–CLR				
	$\rho_{uv} = 0$	.9	$\rho_{uv} = 0$	.3	.5	.7	.9	$\rho_{uv} = 0$	.3	.5	.7	.9
−10000.00	1.00	1.00	.862	.865	.878	.901	.934	**.013**	.008	.004	.001	.000
−1000.00	1.00	1.00	.862	.865	.878	.900	.934	**.013**	.008	.004	.001	.000
−100.00	1.00	1.00	.862	.864	.877	.900	.933	**.013**	.007	.003	.002	.000
−50.00	1.00	1.00	.862	.864	.876	.899	.932	**.013**	.007	.003	.002	.000
−10.00	1.00	1.00	.860	.860	.870	.892	.923	.012	.007	.002	.002	.001
−7.50	.99	1.00	.859	.858	.867	.889	.919	.012	.007	.002	.002	.000
−5.00	.98	1.00	.856	.853	.862	.881	.911	.010	.006	.003	.001	.000
−4.00	.97	1.00	.854	.849	.858	.876	.905	.009	.005	.003	.002	.000
−3.75	.97	1.00	.853	.847	.856	.875	.902	.009	.005	.003	.003	-.000
−3.50	.96	1.00	.852	.845	.854	.872	.900	.009	.004	.003	.003	.000
−3.25	.96	.99	.851	.843	.852	.869	.897	.009	.004	.003	.002	-.000
−3.00	.95	.99	.849	.841	.849	.866	.893	.008	.004	.003	.002	.000
−2.75	.94	.99	.847	.839	.846	.862	.889	.008	.005	.003	.001	-.000
−2.50	.93	.99	.845	.836	.842	.857	.884	.007	.004	.003	.002	.000
−2.25	.91	.99	.842	.832	.837	.851	.877	.006	.004	.002	.002	.000
−2.00	.89	.99	.838	.826	.830	.843	.869	.006	.004	.003	.002	.000
−1.75	.87	.99	.833	.818	.820	.832	.857	.005	.004	.003	.002	.000
−1.50	.83	.98	.825	.807	.806	.817	.841	.005	.005	.003	.002	.000
−1.25	.78	.98	.811	.789	.787	.794	.816	.005	.004	.004	.002	.001
−1.00	.71	.97	.787	.760	.752	.757	.776	.004	.004	.003	.001	.001
−0.75	.60	.97	.739	.701	.688	.688	.704	.004	.002	.002	.001	.000
−0.50	.45	.95	.615	.567	.549	.543	.556	.002	.002	.001	.001	.001
−0.25	.24	.94	.286	.263	.258	.257	.263	.001	.001	.002	.001	.000
0.25	−.24	.83	.286	.328	.367	.428	.512	.002	.003	.001	.001	.001
0.50	−.45	.68	.617	.692	.757	.829	.912	.004	.003	.003	.002	.000
0.75	−.60	.33	.743	.800	.849	.906	.971	.004	.004	.003	.003	.001
1.00	−.71	−.22	.790	.837	.877	.927	.983	.004	.005	.004	.003	**.002**
1.25	−.78	−.63	.812	.853	.889	.936	.983	.004	.005	.005	**.007**	**.002**
1.50	−.83	−.81	.824	.861	.897	.936	.980	.004	.006	.008	**.007**	.001
1.75	−.87	−.89	.832	.867	.900	.934	.976	.006	.007	**.010**	.006	.001
2.00	−.89	−.93	.838	.870	.900	.931	.972	.006	.009	**.010**	.004	.001
2.25	−.91	−.95	.842	.872	.900	.930	.969	.007	.010	**.010**	.004	.001
2.50	−.93	−.96	.845	.873	.899	.927	.967	.008	.011	**.010**	.003	.001
2.75	−.94	−.97	.848	.874	.897	.925	.965	.008	.011	.009	.002	.001
3.00	−.95	−.98	.849	.875	.896	.924	.963	.008	.011	.008	.002	.001
3.25	−.96	−.98	.851	.875	.895	.922	.961	.009	**.012**	.008	.002	.000
3.50	−.96	−.99	.852	.875	.894	.921	.960	.009	**.012**	.007	.002	.001
3.75	−.97	−.99	.853	.875	.893	.920	.958	.009	**.012**	.007	.002	.001
4.00	−.97	−.99	.854	.875	.893	.919	.957	.009	**.012**	.007	.002	.001
5.00	−.98	−.99	.857	.874	.890	.916	.954	.010	**.012**	.006	.002	.001
7.50	−.99	−1.00	.859	.872	.887	.911	.948	.012	.010	.005	.002	.001
10.00	−1.00	−1.00	.860	.870	.885	.909	.945	.012	.010	.005	.002	.001
50.00	−1.00	−1.00	.862	.866	.879	.902	.936	**.013**	.008	.004	.002	.001
100.00	−1.00	−1.00	.862	.865	.879	.901	.935	**.013**	.008	.004	.001	.000
1000.00	−1.00	−1.00	.862	.865	.878	.901	.934	**.013**	.008	.004	.001	.000
10000.00	−1.00	−1.00	.862	.865	.878	.901	.934	**.013**	.008	.004	.001	.000

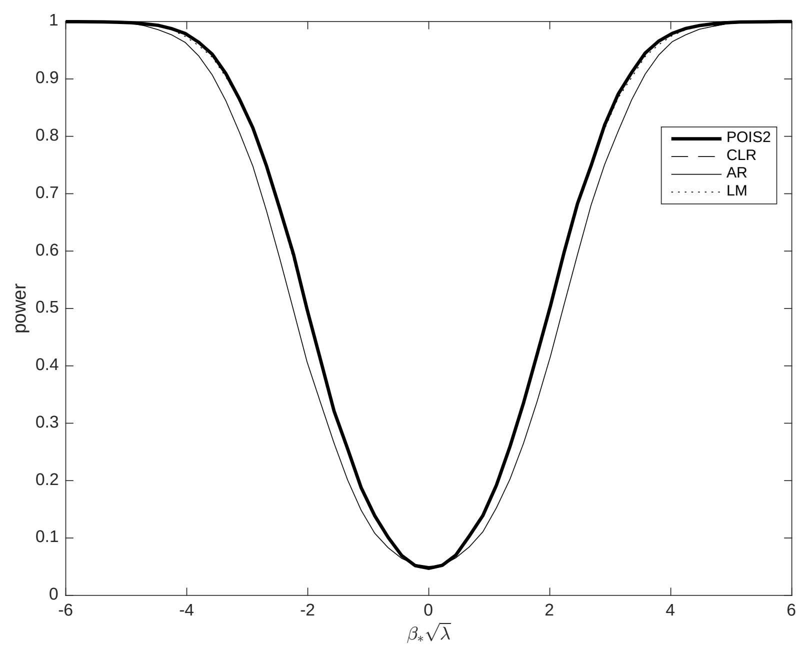
TABLE SM-V(e). Average (over λ) Power Differences for $\lambda \in \{2.5, 5.0, ..., 90.0\}$ between the WAP2 and CLR Tests for $k = 40$

β_0	$\rho_{uv,0}$		WAP2					WAP2–CLR				
	$\rho_{uv}=0$	.9	$\rho_{uv}=0$	.3	.5	.7	.9	$\rho_{uv}=0$	.3	.5	.7	.9
−10000.00	1.00	1.00	.817	.819	.835	.869	.919	**.024**	.013	.006	.004	.001
−1000.00	1.00	1.00	.817	.819	.835	.869	.919	.023	.013	.006	.004	.001
−100.00	1.00	1.00	.817	.818	.834	.867	.917	.023	.013	.006	.004	.001
−50.00	1.00	1.00	.817	.817	.833	.867	.916	.023	.012	.005	.004	.001
−10.00	1.00	1.00	.814	.810	.826	.858	.905	.022	.010	.005	.003	.000
−7.50	.99	1.00	.812	.807	.823	.853	.901	.020	.010	.005	.003	.001
−5.00	.98	1.00	.808	.802	.816	.845	.892	.018	.009	.005	.003	.000
−4.00	.97	1.00	.804	.797	.810	.838	.884	.016	.008	.004	.003	.000
−3.75	.97	1.00	.802	.796	.808	.836	.882	.015	.008	.004	.004	.000
−3.50	.96	1.00	.801	.794	.806	.834	.879	.014	.008	.004	.003	-.000
−3.25	.96	.99	.799	.792	.803	.831	.876	.013	.008	.004	.003	.000
−3.00	.95	.99	.797	.790	.800	.828	.872	.012	.008	.004	.004	.000
−2.75	.94	.99	.795	.786	.796	.823	.867	.012	.007	.004	.003	.001
−2.50	.93	.99	.792	.783	.792	.817	.861	.011	.007	.005	.002	.001
−2.25	.91	.99	.788	.777	.786	.811	.854	.011	.006	.004	.003	.000
−2.00	.89	.99	.783	.772	.778	.802	.846	.009	.007	.005	.004	.001
−1.75	.87	.99	.777	.761	.767	.790	.833	.009	.006	.004	.003	-.000
−1.50	.83	.98	.768	.749	.754	.773	.816	.008	.006	.004	.003	.001
−1.25	.78	.98	.752	.730	.730	.750	.791	.007	.006	.004	.003	.000
−1.00	.71	.97	.727	.698	.694	.709	.751	.007	.007	.004	.003	-.000
−0.75	.60	.97	.672	.633	.624	.632	.675	.006	.005	.004	.001	-.000
−0.50	.45	.95	.536	.490	.480	.489	.524	.004	.004	.002	.001	-.002
−0.25	.24	.94	.233	.217	.215	.221	.242	.001	.001	-.000	-.001	-.002
0.25	−.24	.83	.237	.275	.318	.383	.491	.001	.000	.000	-.002	-.002
0.50	−.45	.68	.539	.621	.697	.788	.892	.004	.004	.004	.003	.001
0.75	−.60	.33	.672	.741	.804	.876	.958	.004	.005	.005	.004	.001
1.00	−.71	−.22	.727	.784	.837	.901	.974	.005	.006	.006	.006	.003
1.25	−.78	−.63	.754	.803	.853	.911	.974	.006	.008	.010	.009	**.004**
1.50	−.83	−.81	.769	.815	.861	.913	.969	.007	.012	.014	**.011**	.002
1.75	−.87	−.89	.779	.824	.865	.910	.965	.008	.016	.016	.008	.001
2.00	−.89	−.93	.785	.828	.866	.905	.961	.009	.017	.016	.006	.001
2.25	−.91	−.95	.790	.831	.866	.903	.958	.010	.019	**.017**	.006	.001
2.50	−.93	−.96	.794	.833	.864	.901	.955	.011	.020	.016	.005	.001
2.75	−.94	−.97	.797	.834	.862	.898	.953	.012	.020	.015	.004	.001
3.00	−.95	−.98	.799	.834	.861	.896	.950	.013	**.021**	.014	.004	.001
3.25	−.96	−.98	.801	.834	.859	.894	.949	.014	**.021**	.013	.004	.001
3.50	−.96	−.99	.803	.835	.858	.893	.947	.015	**.021**	.013	.004	.001
3.75	−.97	−.99	.805	.834	.857	.892	.946	.016	**.021**	.013	.004	.001
4.00	−.97	−.99	.806	.834	.855	.890	.944	.016	**.021**	.012	.004	.001
5.00	−.98	−.99	.810	.831	.852	.887	.940	.018	.019	.011	.004	.001
7.50	−.99	−1.00	.814	.828	.847	.881	.933	.021	.017	.009	.004	.001
10.00	−1.00	−1.00	.815	.826	.844	.878	.929	.022	.016	.009	.004	.002
50.00	−1.00	−1.00	.817	.820	.837	.871	.921	.023	.014	.007	.004	.001
100.00	−1.00	−1.00	.817	.820	.836	.870	.920	.023	.014	.007	.005	.001
1000.00	−1.00	−1.00	.817	.819	.836	.869	.919	**.024**	.013	.006	.004	.001
10000.00	−1.00	−1.00	.817	.819	.835	.869	.919	**.024**	.013	.006	.004	.001

FIGURE SM-I(a). Power of the significance level .05 CLR, LM, and AR tests and the POIS2 power envelope for fixed null value $\beta_0 = 0$, varying true value β_*, $k = 2$, $\rho_\Omega = 0$, and $\lambda = 5, 20$

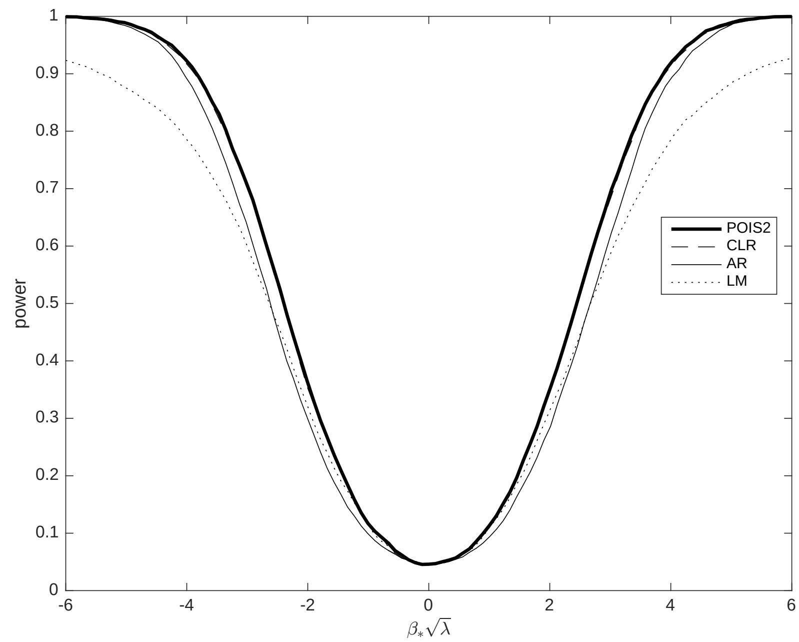


(i) $\lambda = 5$

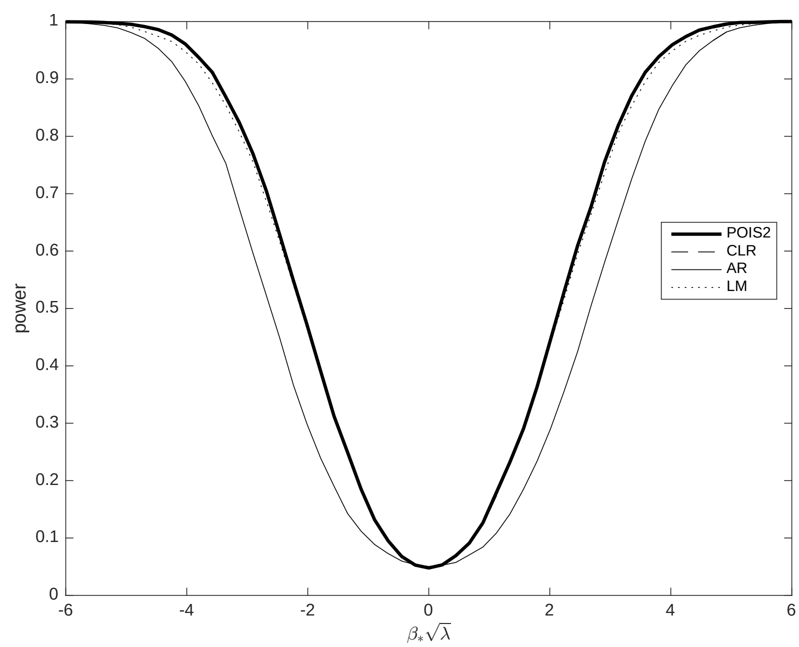


(ii) $\lambda = 20$

FIGURE SM-I(b). Power of the significance level .05 CLR, LM, and AR tests and the POIS2 power envelope for fixed null value $\beta_0 = 0$, varying true value β_*, $k = 5$, $\rho_\Omega = 0$, and $\lambda = 5, 20$

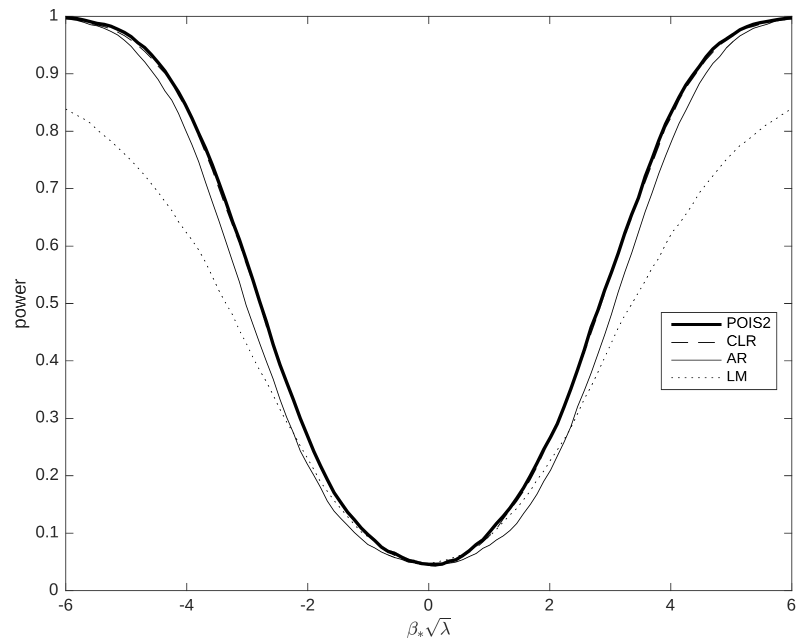


(i) $\lambda = 5$

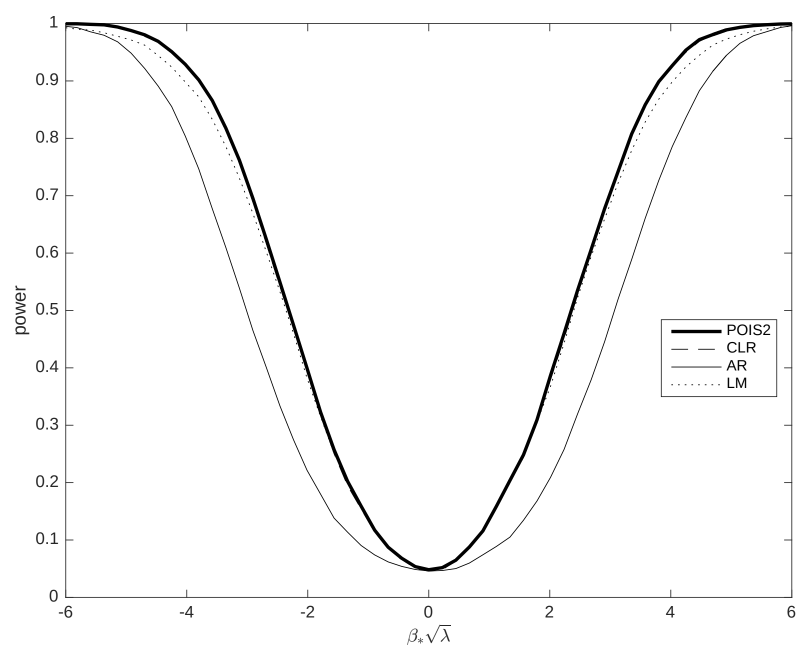


(ii) $\lambda = 20$

FIGURE SM-I(c). Power of the significance level .05 CLR, LM, and AR tests and the POIS2 power envelope for fixed null value $\beta_0 = 0$, varying true value β_*, $k = 10$, $\rho_\Omega = 0$, and $\lambda = 5, 20$



(i) $\lambda = 5$



(ii) $\lambda = 20$