Massachusetts Institute of Technology
Department of Economics
Working Paper Series

ADMISSIBLE INVARIANT SIMILAR TESTS FOR INSTRUMENTAL VARIABLES REGRESSION

Victor Chernozhukov
Christian Hansen
Michael Jansson

Working Paper 07-09
March 12, 2007

Room E52-251
50 Memorial Drive
Cambridge, MA 02142

Admissible Invariant Similar Tests for Instrumental Variables Regression*

VICTOR CHERNOZHUKOV
DEPARTMENT OF ECONOMICS
M.I.T.

CHRISTIAN HANSEN
GRADUATE SCHOOL OF BUSINESS
UNIVERSITY OF CHICAGO

MICHAEL JANSSON
DEPARTMENT OF ECONOMICS
UC BERKELEY

March 12, 2007

ABSTRACT. This paper studies a model widely used in the weak instruments literature and establishes admissibility of the weighted average power likelihood ratio tests recently derived by Andrews, Moreira, and Stock (2004). The class of tests covered by this admissibility result contains the Anderson and Rubin (1949) test. Thus, there is no conventional statistical sense in which the Anderson and Rubin (1949) test "wastes degrees of freedom". In addition, it is shown that the test proposed by Moreira (2003) belongs to the closure of (i.e., can be interpreted as a limiting case of) the class of tests covered by our admissibility result.

1. INTRODUCTION

Conducting valid (and preferably "optimal") inference on structural coefficients in instrumental variables (IVs) regression models is known to be nontrivial when the IVs are weak.[1] Influential papers on this subject include Dufour (1997) and Staiger and Stock (1997), both of which highlight the inadequacy of conventional asymptotic approximations to the behavior of two-stage least squares and point out that valid inference can be based on the Anderson and Rubin (1949, henceforth AR) test. Several methods intended to enjoy improved power properties relative to the AR test have been proposed, prominent examples being the conditional likelihood ratio (CLR) and Lagrange multiplier (LM) tests of Moreira (2003) and Kleibergen (2002), respectively.

All of the abovementioned methods and results have been or can be deduced within an IV regression model with a single endogenous regressor, fixed (i.e., non-stochastic) IVs, and *i.i.d.* homoskedastic Gaussian errors. Studying that model, Andrews, Moreira, and Stock (2004, henceforth AMS) obtain a family of tests, the so-called weighted

*The authors thank Jim Powell, Paul Ruud, a co-editor, and (in particular) two referees for very valuable comments.

[1]Recent reviews include Stock, Wright, and Yogo (2002), Dufour (2003), Hahn and Hausman (2003), and Andrews and Stock (2006).

average power likelihood ratio (WAP-LR) tests, each member of which enjoys demonstrable optimality properties within the class of tests satisfying a certain invariance restriction. The invariance restriction in question, namely that inference is invariant to transformations of the minimal sufficient statistic (S, T) (defined in (3) below) corresponding to a rotation of the instruments, is satisfied by the AR, CLR, and LM tests. Furthermore, testing problems involving the structural coefficient are rotation invariant under the distributional assumptions employed by AMS. For these reasons (and others), the rotation invariance restriction seems "natural", in which case AMS's numerical finding that the CLR test is "nearly" efficient relative to the class of rotation invariant tests provides strong evidence in favor of the CLR test.

Nevertheless, because best invariant procedures can fail to be admissible even if they exist, it is not entirely obvious whether it is "natural" to confine attention to rotation invariant tests when developing optimality theory for hypothesis tests in the model of AMS. In particular, it would appear to be an open question whether the members of the WAP-LR family, upon which the construction of the two-sided power envelope of AMS is based, are even admissible (in the Gaussian model with fixed IVs). We show that all members of the WAP-LR family are indeed admissible, essentially because the defining optimality property of these tests can be reformulated in such a way that rotation invariance becomes a conclusion rather than an assumption. The AR test belongs to the WAP-LR family and is therefore admissible. In contrast, the CLR and LM tests do not seem to admit WAP-LR representations, though we demonstrate here that these test statistics "nearly" admit WAP-LR interpretations in the sense that they belong to the closure (appropriately defined) of the class of WAP-LR tests.

Section 2 introduces the model and defines some terminology needed for the development of the formal results of the paper, all of which are stated in Section 3 and proven in Section 4.

2. PRELIMINARIES

Consider the model

$$\begin{aligned} y_1 &= y_2\beta + u, \\ y_2 &= Z\pi + v_2, \end{aligned} \tag{1}$$

where y_1, $y_2 \in \mathbb{R}^n$ and $Z \in \mathbb{R}^{n\times k}$ are observed variables (for some $k \geq 2$); $\beta \in \mathbb{R}$ and $\pi \in \mathbb{R}^k$ are unknown parameters; and u, $v_2 \in \mathbb{R}^n$ are unobserved errors.[2]

[2]It is straightforward to accommodate exogenous regressors in (1). We tacitly assume that any such regressors have been "partialed out". This assumption is made for simplicity and entails no loss of generality (for details, see Section 2 of AMS).

Suppose β is the parameter of the interest. Specifically, suppose we are interested in a testing problem of the form

$$H_0 : \beta = \beta_0 \qquad \text{vs.} \qquad H_1 : \beta \neq \beta_0.$$

Writing the model in reduced form, we have:

$$\begin{aligned} y_1 &= Z\pi\beta + v_1, \\ y_2 &= Z\pi + v_2, \end{aligned} \tag{2}$$

where $v_1 = u + v_2\beta$. Following AMS (and many others), we treat Z as a fixed $n \times k$ matrix with full column rank and we assume that $(v_1', v_2')' \sim \mathcal{N}(0, \Omega \otimes I_n)$, where Ω is a known, positive definite 2×2 matrix. Without loss generality we normalize a variety of unimportant constants by assuming that $\beta_0 = 0$, $Z'Z = I_k$, and that Ω is of the form

$$\Omega = \begin{pmatrix} 1 & \delta \\ \delta & 1 + \delta^2 \end{pmatrix},$$

where $\delta \in \mathbb{R}$ is known.[3]

Under the stated assumptions, the model is fully parametric, the (multivariate normal) distribution of (y_1, y_2) being completely specified up to the parameters β and π. A minimal sufficient statistic for β and π is given by

$$\begin{pmatrix} S \\ T \end{pmatrix} = \begin{pmatrix} Z'y_1 \\ Z'(y_2 - \delta y_1) \end{pmatrix} \sim \mathcal{N}(\eta_\beta \otimes \pi, I_{2k}), \tag{3}$$

where

$$\eta_\beta = \begin{pmatrix} \beta \\ 1 - \delta\beta \end{pmatrix}.$$

Because (S, T) is sufficient, the totality of attainable power functions is spanned

[3]These assumptions correspond to the model in which (Z, y_1, y_2) has been replaced by

$$\left(\tilde{Z}, \tilde{y}_1, \tilde{y}_2\right) = (Z'Z)^{-1/2}\left[Z, \omega_{11}^{-1/2}(y_1 - y_2\beta_0), \omega_{22.1}^{-1/2} y_2\right],$$

where ω_{ij} is the (i,j) element of Ω and $\omega_{22.1} = \omega_{22} - \omega_{11}^{-1}\omega_{12}^2$. In the chosen parameterization, the parameter δ is related to the correlation coefficient ρ computed from Ω through the formula $\delta = \rho/\sqrt{1-\rho^2}$.

by the set of power functions associated with (possibly randomized) tests based on (S,T). Any such test can be represented by means of a $[0,1]$-valued function $\phi(\cdot)$ such that H_0 is rejected with probability $\phi(s,t)$ if $(S,T)=(s,t)$. The power function of this test is the function (with arguments β and π) $E_{\beta,\pi}\phi(S,T)$, where the subscript on E indicates the distribution with respect to which the expectation is taken.

For any $\alpha \in (0,1)$, a test with test function ϕ is of level α if

$$\sup_{\pi} E_{0,\pi}\phi(S,T) \leq \alpha. \tag{4}$$

A level α test with test function ϕ is said to be α-admissible if

$$E_{\beta,\pi}\phi(S,T) \leq E_{\beta,\pi}\varphi(S,T) \qquad \forall(\beta,\pi) \tag{5}$$

implies

$$E_{\beta,\pi}\phi(S,T) = E_{\beta,\pi}\varphi(S,T) \qquad \forall(\beta,\pi) \tag{6}$$

whenever the test associated with φ is of level α.[4]

The main purpose of the present paper is to investigate the α-admissibility properties of certain recently developed rotation invariant, α-similar tests. By definition, a rotation invariant test is one whose test function ϕ satisfies $\phi(OS,OT)=\phi(S,T)$ for every orthogonal $k\times k$ matrix O and an α-similar test is one for which

$$E_{0,\pi}\phi(S,T) = \alpha \qquad \forall\pi. \tag{7}$$

The rotation invariant, α-similar tests under consideration here are the WAP-LR tests of AMS. By construction, a size α WAP-LR test maximizes a WAP criterion of the form

$$\int_{\mathbb{R}^{k+1}} E_{\beta,\pi}\phi(S,T)\,dW(\beta,\pi) \tag{8}$$

among rotation invariant, α-similar tests, where the weight function W is some cumulative distribution function (cdf) on $\mathbb{R}^{k+1}$.

In spite of the fact that the defining property of a WAP-LR test is an optimality property, it is not obvious whether a WAP-LR is α-admissible, the reason being that

[4] In the model under study here, the present notion of α-admissibility agrees with that of Lehmann and Romano (2005, Section 6.7) (in which (5) and (6) are required to hold for all $\beta\neq 0$ and all π) because (i) all power functions are continuous and (ii) the set $\{\beta:\beta\neq 0\}$ is a dense subset of $\mathbb{R}$.

admissibility of optimum invariant tests cannot be taken for granted [e.g., Lehmann and Romano (2005, Section 6.7)]. We show in Section 3 that any WAP-LR test maximizes a WAP criterion among the class of α-similar tests. In the present context, that optimality property is sufficiently strong to imply that any WAP-LR test is α-admissible.

Remark. The fact that the alternative hypothesis is dense in the maintained hypothesis implies that any α-admissible test is d-admissible in the sense of Lehmann and Romano (2005). The (rotation invariant) posterior odds ratio tests of Chamberlain (2006) are d-admissible almost by construction [e.g., Lehmann and Romano (2005, Theorem 6.7.2 (i))], but because these tests are not necessarily similar it is unclear whether these tests are also α-admissible (for some α).

3. RESULTS

3.1. Admissible α-similar tests. Two basic facts about exponential families greatly simplify the construction of α-admissible, α-similar tests. First, because the power function of the test with test function ϕ can be represented as

$$E_{\beta,\pi}\phi\left(S,T\right)=\int_{\mathbb{R}^k}\int_{\mathbb{R}^k}\phi\left(s,t\right)f_S\left(s|\beta,\pi\right)f_T\left(t|\beta,\pi\right)dsdt, \tag{9}$$

where $f_S\left(\cdot|\beta,\pi\right)$ and $f_T\left(\cdot|\beta,\pi\right)$ denote the densities (indexed by β and π) of S and T, it follows from Lehmann and Romano (2005, Theorem 2.7.1) that $E_{\beta,\pi}\phi\left(S,T\right)$ is a continuous function of $\left(\beta,\pi\right)$. Therefore, an α-similar similar test cannot be dominated by a level α test which is not α-similar. By implication, an α-similar test with test function ϕ is α-admissible if (5) implies (6) whenever the test associated with φ is α-similar.

Second, because π is unrestricted, T is a complete, sufficient statistic for π under H_0 [e.g., Moreira (2001)]. As a consequence, a test with test function ϕ is α-similar if and only if it is conditionally α-similar in the sense that, almost surely,

$$E_{0,\pi}\left[\phi\left(S,T\right)|T\right]=\alpha \qquad \forall\pi. \tag{10}$$

It follows from the preceding considerations and the Neyman-Pearson lemma that if W is a cdf on $\mathbb{R}^{k+1}$, then the WAP criterion (8) is maximized among α-similar tests by the test with test function given by

$$\phi_{LR}^{W}\left(s,t;\alpha\right)=1\left[LR^{W}\left(s,t\right)>\kappa_{LR}^{W}\left(t;\alpha\right)\right], \tag{11}$$

where $1\left[\cdot\right]$ is the indicator function,

$$LR^W(s,t) = \frac{\int_{\mathbb{R}^{k+1}} f_S(s|\beta,\pi) f_T(t|\beta,\pi) \, dW(\beta,\pi)}{f_S(s|0,0) f_T(t|0,0)} \tag{12}$$

and $\kappa_{LR}^W(t;\alpha)$ is the $1-\alpha$ quantile of the distribution of $LR^W(\mathcal{Z}_k, t)$, where $\mathcal{Z}_k$ is distributed $\mathcal{N}(0, I_k)$.[5] Because the maximizer ϕ_{LR}^W is essentially unique (in the measure theoretic sense), the test with test function $\phi_{LR}^W(\cdot;\alpha)$ is α-admissible.[6] To demonstrate α-admissibility of a test, it therefore suffices to show its test function can be represented as $\phi_{LR}^W(\cdot;\alpha)$ for some W. Section 3.2 uses that approach to show that the WAP-LR tests of AMS are all α-admissible.

3.2. Admissibility of WAP-LR tests. The WAP-LR tests are indexed by cdfs on $\mathbb{R} \times \mathbb{R}_+$.[7] Accordingly, let w be a cdf on $\mathbb{R} \times \mathbb{R}_+$ and define

$$\mathcal{L}^w(s,t) = \int_{\mathbb{R}\times\mathbb{R}_+} \exp\left(-\frac{\lambda\eta_\beta'\eta_\beta}{2}\right) {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{\lambda}{4}\eta_\beta' Q(s,t)\eta_\beta\right] dw(\beta,\lambda), \tag{13}$$

where ${}_0\tilde{F}_1$ is the regularized confluent hypergeometric function,

$$Q(s,t) = \begin{pmatrix} s's & s't \\ t's & t't \end{pmatrix}, \tag{14}$$

and η_β is defined as in Section 2. By Corollary 1 of AMS, the test function of the size α WAP-LR test associated with w is

$$\phi_{AMS}^w(s,t;\alpha) = 1\left[\mathcal{L}^w(s,t) > \kappa_{AMS}^w(t;\alpha)\right], \tag{15}$$

where $\kappa_{AMS}^w(t;\alpha)$ is the $1-\alpha$ quantile of the distribution of $\mathcal{L}^w(\mathcal{Z}_k, t)$.[8,9]

[5] Details are provided in an Appendix, available from the authors upon request.

[6] As pointed out by a referee, this fact is a special case of the more general decision theoretic result that (essentially) unique Bayes rules are admissible [e.g., Ferguson (1967, Theorem 2.3.1)].

[7] Because the power of an invariant test depends on π only through the scalar $\pi'\pi$, the corresponding WAP criterion (8) depends on W only through the cdf on $\mathbb{R} \times \mathbb{R}_+$ given by

$$w(\beta,\lambda) = \int_{b\leq\beta,\pi'\pi\leq\lambda} dW(b,\pi).$$

[8] $\mathcal{L}^w(s,t)$ is proportional to $\psi_w(q_1,q_T)$ in Lemma 1 of AMS because ${}_0\tilde{F}_1(;k/2;z/4)$ is proportional to $z^{-(k-2)/4} I_{(k-2)/2}(\sqrt{z})$, where $I_\nu(\cdot)$ denotes the modified Bessel function of the first kind of order ν.

[9] The size α WAP-LR test is α-similar by construction and is rotation invariant because $\kappa_{AMS}^w(t;\alpha)$ depends on t only through $t't$.

For our purposes, it is convenient to characterize the defining optimality property of ϕ^w_{AMS} as follows. Let W^w_{AMS} denote the cdf (on $\mathbb{R}^{k+1}$) of $\left(\mathcal{B}, \sqrt{\Lambda}\mathcal{U}_k'\right)'$, where $(\mathcal{B}, \Lambda)'$ has cdf w and $\mathcal{U}_k$ is uniformly distributed on the unit sphere in $\mathbb{R}^k$ (independently of $\mathcal{B}$ and Λ). In terms of W^w_{AMS}, Theorem 3 of AMS asserts that the test with test function $\phi^w_{AMS}(\cdot;\alpha)$ maximizes

$$\int_{\mathbb{R}^{k+1}} E_{\beta,\pi}\phi(S,T)\, dW^w_{AMS}(\beta,\pi)$$

among rotation invariant, α-similar tests. In this optimality result, the assumption of rotation invariance is unnecessary because it turns out that

$$\phi^w_{AMS}(s,t;\alpha) = \phi_{LR}^{W^w_{AMS}}(s,t;\alpha). \tag{16}$$

The displayed equality follows from a calculation performed in the proof of the following strengthening of Theorem 3 of AMS.

Theorem 1. *Let w be a cdf on $\mathbb{R} \times \mathbb{R}_+$. If ϕ satisfies (7), then*

$$\int_{\mathbb{R}^{k+1}} E_{\beta,\pi}\phi(S,T)\, dW^w_{AMS}(\beta,\pi) \leq \int_{\mathbb{R}^{k+1}} E_{\beta,\pi}\phi^w_{AMS}(S,T;\alpha)\, dW^w_{AMS}(\beta,\pi),$$

where the inequality is strict unless $\Pr_{\beta,\pi}\left[\phi(S,T) = \phi^w_{AMS}(S,T;\alpha)\right] = 1$ for some (and hence for all) (β,π). In particular, the size α WAP-LR test associated with w is α-admissible.

The testing function of the size α AR test (for known Ω) is given by

$$\phi_{AR}(s,t;\alpha) = 1\left[s's > \chi^2_\alpha(k)\right], \tag{17}$$

where $\chi^2_\alpha(k)$ is the $1-\alpha$ quantile of the χ^2 distribution with k degrees of freedom. As remarked by AMS, $\mathcal{L}^w(s,t)$ is an increasing function of $s's$ whenever the weight function w assigns unit mass to the set $\left\{(\beta,\lambda) \in \mathbb{R} \times \mathbb{R}_+ : \beta = \delta^{-1}\right\}$. As a consequence, the following is an immediate consequence of Theorem 1.

Corollary 2. *The size α AR test is α-admissible.*

In spite of the fact that the AR test is a k degrees of freedom test applied to a testing problem with a single restriction, a fact which suggests that its power properties should be poor [e.g., Kleibergen (2002, p. 1781), Moreira (2003, p. 1031)], Corollary 2 implies that there is no conventional statistical sense in which the AR test "wastes degrees of freedom".

In addition to the AR test, Theorem 1 also covers the two-point optimal invariant similar tests of AMS, the power functions of which trace out a two-sided power envelope for rotation invariant, α-similar tests. On the other hand, the CLR and LM tests do not seem to belong to the class of WAP-LR tests. Indeed, it would appear to be an open question whether one or both of these tests even belong to the closure (appropriately defined) of the class WAP-LR tests. Section 3.3 provides an affirmative answer to that question.

Remark. As pointed out by a referee, two distinct generalizations of the results of this section seem feasible. First, the conclusion that a Bayes rule corresponding to a distribution w on $\mathbb{R} \times \mathbb{R}_+$ can be "lifted" to a Bayes rule corresponding to a distribution on $\mathbb{R}^{k+1}$ (by introducing a $\mathcal{U}_k$ which is uniformly distributed on the unit sphere in $\mathbb{R}^k$ and independent of (β, λ)) applies to more general decision problems than the one considered. Second, using Muirhead (1982, Theorems 2.1.14 and 7.4.1) it should be possible to generalize the results to a model with multiple endogenous regressors. To conserve space, we do not pursue these extensions here.

3.3. The CLR test. The test function of the size α CLR test is

$$\phi_{CLR}(s,t;\alpha) = 1\left[LR(s,t) > \kappa_{CLR}(t;\alpha)\right], \tag{18}$$

where

$$LR(s,t) = \frac{1}{2}\left(s's - t't + \sqrt{(s's + t't)^2 - 4\left[(s's)(t't) - (s't)^2\right]}\right) \tag{19}$$

and $\kappa_{CLR}(t;\alpha)$ is the $1-\alpha$ quantile of the distribution of $LR(\mathcal{Z}_k, t)$.

In numerical investigations, the CLR test has been found to perform remarkably well in terms of power. For instance, AMS find that the power of the size α CLR test is "essentially the same" as the two-sided power envelope for rotation invariant, α-similar tests. In light of this numerical finding, it would appear to be of interest to analytically characterize the relation (if any) between the CLR test and the class of WAP-LR tests.

Theorem 3 shows that $CLR(s,t)$ can be represented as the limit as $N \to \infty$ of a suitably normalized version of $\mathcal{L}^{w_{CLR,N}}(s,t)$, where $\{w_{CLR,N} : N \in \mathbb{N}\}$ is a carefully chosen collection of cdfs on $\mathbb{R} \times \mathbb{R}_+$.

To motivate this representation and the functional form of $w_{CLR,N}$, it is convenient to express the test function of the CLR test as

$$\phi_{CLR}(s,t;\alpha) = 1\left[CLR^*(s,t) > \kappa_{CLR^*}(t;\alpha)\right],$$

where

$$\begin{aligned} CLR^*(s,t) &= \sqrt{s's + t't + \sqrt{(s's+t't)^2 - 4\left[(s's)(t't) - (s't)^2\right]}} \\ &= \sqrt{2\left[LR(s,t) + t't\right]} \end{aligned} \tag{20}$$

and $\kappa_{CLR^*}(t;\alpha) = \sqrt{2\left[\kappa_{CLR}(t;\alpha) + t't\right]}$.

The statistic $CLR^*(s,t)$ is the square root of the largest eigenvalue of $Q(s,t)$ (defined in (14)) and therefore admits the following characterization:

$$CLR^*(s,t) = \sqrt{\max_{\eta \in \mathbb{R}^2 : \eta'\eta = 1} \eta' Q(s,t)\eta}. \tag{21}$$

Moreover, the integrand in (13) can be written as

$$\exp\left(-\frac{\lambda\eta_\beta'\eta_\beta}{2}\right) {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{\lambda\eta_\beta'\eta_\beta}{4}\tilde{\eta}_\beta' Q(s,t)\tilde{\eta}_\beta\right],$$

where $\tilde{\eta}_\beta = \eta_\beta / \sqrt{\eta_\beta'\eta_\beta}$ is a vector of unit length proportional to η_β. If $w_{CLR,N}$ is such that its support is the set of all pairs (β,λ) for which $\lambda\eta_\beta'\eta_\beta = N$, then the integrand is maximized (over the support of $w_{CLR,N}$) by setting $\tilde{\eta}_\beta$ equal to the eigenvector associated with the largest eigenvalue of $Q(s,t)$. This observation, and the fact that the tail behavior of ${}_0\tilde{F}_1[;k/2;\cdot/4]$ is similar to that of $\exp(\sqrt{\cdot})$, suggests that the large N behavior of $\mathcal{L}^{w_{CLR,N}}(s,t)$ "should" depend on $Q(s,t)$ only through $CLR^*(s,t)$.

For any $N > 0$, let $w_{CLR,N}$ denote the cdf of $\left(\mathcal{B}, N/\sqrt{\eta_\mathcal{B}'\eta_\mathcal{B}}\right)'$, where $\mathcal{B} \sim \mathcal{N}(0,1)$.[10] By construction, $w_{CLR,N}$ is such that its support is the set of all pairs (β,λ) for which $\lambda\eta_\beta'\eta_\beta = N$. Using that property, the relation (21), and basic facts about ${}_0\tilde{F}_1(;k/2;\cdot)$, we obtain the following result.

[10]The distributional assumption $\mathcal{B} \sim \mathcal{N}(0,1)$ is made for concreteness. An inspection of the proof of Theorem 3 shows that (22) is valid for any distribution (of $\mathcal{B}$) whose support is $\mathbb{R}$. Furthermore, as pointed out by a referee it is possible to obtain analogous results without making the distribution of $\lambda\eta_\beta'\eta_\beta$ degenerate.

Theorem 3. *For any* $(s', t')' \in \mathbb{R}^{2k}$,

$$CLR^* (s,t) = \lim_{N \to \infty} \frac{1}{\sqrt{N}} \left[\log \mathcal{L}^{w_{CLR,N}} (s,t) + \frac{N}{2} \right]. \tag{22}$$

In particular,

$$CLR(s,t) = \lim_{N \to \infty} \frac{1}{2N} \left[\log \mathcal{L}^{w_{CLR,N}} (s,t) + \frac{N}{2} \right]^2 - t't. \tag{23}$$

Define $\mathcal{L}^*_{CLR,N}(s,t) = [\log \mathcal{L}^{w_{CLR,N}}(s,t) + N/2] / \sqrt{N}$. The proof of Theorem 3 shows that as $N \to \infty$, $\mathcal{L}^*_{CLR,N}(\cdot)$ converges to $CLR^*(\cdot)$ in the topology of uniform convergence on compacta. Using this result, it follows that

$$\lim_{N \to \infty} E_{\beta,\pi} \left| \phi^{w_{CLR,N}}_{AMS}(S,T;\alpha) - \phi_{CLR}(S,T;\alpha) \right| = 0 \qquad \forall (\beta, \pi).$$

In particular, the power function of the WAP-LR test associated with $w_{CLR,N}$ converges (pointwise) to the power function of the CLR test (as $N \to \infty$).

In light of the previous paragraph, it seems plausible that the CLR test enjoys an "admissibility at ∞" property reminiscent of Andrews (1996, Theorem 1(c)). Verifying this conjecture is not entirely trivial, however, because the unboundedness (in t) of $\kappa_{CLR^*}(t;\alpha)$ makes it difficult (if not impossible) to adapt the proofs of Andrews and Ploberger (1995) and Andrews (1996) to the present situation.

Theorem 3 also suggests a method of constructing tests which share the nice numerical properties of the CLR test and furthermore enjoy demonstrable optimality properties, namely tests based on $\mathcal{L}^{w_{CLR,N}}(s,t)$ for some "large" (possibly sample-dependent) value of N. Because the power improvements (if any) attainable in this way are likely to be slight, the properties of tests constructed in this way are not investigated in this paper.

Remark. The test function of the size α LM test is

$$\phi_{LM}(s,t;\alpha) = 1 \left[\frac{(s't)^2}{t't} > \chi^2_\alpha(1) \right],$$

where $\chi^2_\alpha(1)$ is the $1-\alpha$ quantile of the χ^2 distribution with 1 degree of freedom. AMS show that a one-sided version of the LM test can be interpreted as a limit of WAP-LR tests (and is locally most powerful invariant). On the other hand, we are not aware of any such results which cover the (two-sided) LM test.

A slight variation on the argument used in the previous subsection can be used to obtain a WAP-LR interpretation of the LM test. Indeed, letting $w_{LM,N}$ denote the cdf of $\left[\mathcal{B}_N, N/\sqrt{\eta_{\mathcal{B}_N}'\eta_{\mathcal{B}_N}}\right]'$, where the distribution of $N^{1/3}\mathcal{B}_N$ is uniform on $[-1,1]$, it can be shown that[11]

$$|s't|/\sqrt{t't} = \lim_{N\to\infty} \frac{1}{N^{1/6}}\left[\log \mathcal{L}^{w_{LM,N}}(s,t) + \frac{N}{2} - \sqrt{N}\sqrt{t't}\right]. \quad (24)$$

implying in particular that

$$\frac{(s't)^2}{t't} = \lim_{N\to\infty} \frac{1}{N^{1/3}}\left[\log \mathcal{L}^{w_{LM,N}}(s,t) + \frac{N}{2} - \sqrt{N}\sqrt{t't}\right]^2. \quad (25)$$

4. PROOFS

Proof of Theorem 1. It suffices to establish (16). To do so, it suffices to show that $LR^{W_{AMS}^w}(s,t)$ is proportional to $\mathcal{L}^w(s,t)$. Now,

$$\begin{aligned}
\frac{f_S(s|\beta,\pi)\, f_T(t|\beta,\pi)}{f_S(s|0,0)\, f_T(t|0,0)} &= \exp\left(-\frac{1}{2}\left[\|s-\beta\pi\|^2 - \|s\|^2 + \|t-(1-\delta\beta)\pi\|^2 - \|t\|^2\right]\right) \\
&= \exp\left(-\frac{\lambda_\pi \eta_\beta'\eta_\beta}{2}\right)\exp\left(\sqrt{\lambda_\pi}\left[\beta s + (1-\delta\beta)\, t\right]'\tilde{\pi}\right),
\end{aligned}$$

where $\|\cdot\|$ signifies the Euclidean norm, $\lambda_\pi = \pi'\pi$, and $\tilde{\pi} = \lambda_\pi^{-1/2}\pi$.

Because $\tilde{\pi}$ has unit length, it follows from Muirhead (1982, Theorem 7.4.1) that

$$\begin{aligned}
LR^{W_{AMS}^w}(s,t) &= \int_{\mathbb{R}^{k+1}} \frac{f_S(s|\beta,\pi)\, f_T(t|\beta,\pi)}{f_S(s|0,0)\, f_T(t|0,0)}\, dW_{AMS}^w(\beta,\pi) \\
&\propto \int_{\mathbb{R}\times\mathbb{R}_+} \exp\left(-\frac{\lambda\eta_\beta'\eta_\beta}{2}\right)\ {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{\lambda}{4}\eta_\beta' Q(s,t)\,\eta_\beta\right] dw(\beta,\lambda),
\end{aligned}$$

as was to be shown. (We are grateful to a referee for suggesting the use of Muirhead (1982, Theorem 7.4.1).) ■

Proof of Theorem 3. The result is obvious if $Q(s,t) = 0$, so suppose $Q(s,t) \neq 0$.

[11] Details are provided in an Appendix, available from the authors upon request. An inspection of the proof given there shows that the distributional assumption on $N^{1/3}\mathcal{B}_N$ is made for concreteness insofar as the preceding representations are valid whenever $N^{1/3}\mathcal{B}_N$ has a (fixed) distribution whose support is $[-1,1]$.

The proof will make use of the fact [e.g., Andrews and Ploberger (1995, Lemma 2)] that $0 < \underline{C} \leq \overline{C} < \infty$, where

$$\underline{C} = \inf_{z \geq 0} \frac{{}_0\tilde{F}_1\left[; k/2; z/4\right]}{\exp\left(\sqrt{z}\right) \max\left(z, 1\right)^{-(k-1)/4}}, \qquad \overline{C} = \sup_{z \geq 0} \frac{{}_0\tilde{F}_1\left[; k/2; z/4\right]}{\exp\left(\sqrt{z}\right)}.$$

By construction,

$$\begin{aligned} \mathcal{L}^{w_{CLR,N}}\left(s,t\right) &= \int_{\mathbb{R}\times\mathbb{R}_+} \exp\left(-\frac{\lambda \eta_\beta' \eta_\beta}{2}\right) {}_0\tilde{F}_1\left[; \frac{k}{2}; \frac{\lambda}{4}\eta_\beta' Q\left(s,t\right)\eta_\beta\right] dw_{CLR,N}\left(\beta, \lambda\right) \\ &= \exp\left(-\frac{N}{2}\right) \int_{\mathbb{R}} {}_0\tilde{F}_1\left[; \frac{k}{2}; \frac{N}{4}\tilde{\eta}_\beta' Q\left(s,t\right)\tilde{\eta}_\beta\right] d\Phi\left(\beta\right), \end{aligned}$$

where $\Phi\left(\cdot\right)$ is the standard normal cdf. Using this representation, the relation (21), and monotonicity of ${}_0\tilde{F}_1\left[; k/2; \cdot\right]$,

$$\begin{aligned} \mathcal{L}^{w_{CLR,N}}\left(s,t\right) \exp\left(\frac{N}{2}\right) &= \int_{\mathbb{R}} {}_0\tilde{F}_1\left[; \frac{k}{2}; \frac{N}{4}\tilde{\eta}_\beta' Q\left(s,t\right)\tilde{\eta}_\beta\right] d\Phi\left(\beta\right) \\ &\leq {}_0\tilde{F}_1\left[; \frac{k}{2}; \frac{N}{4} CLR^*\left(s,t\right)^2\right] \\ &\leq \overline{C} \exp\left[\sqrt{N} CLR^*\left(s,t\right)\right], \end{aligned}$$

from which it follows immediately that

$$\overline{\lim}_{N\to\infty} \frac{1}{\sqrt{N}} \left[\log \mathcal{L}^{w_{CLR,N}}\left(s,t\right) + \frac{N}{2}\right] \leq CLR^*\left(s,t\right). \tag{26}$$

On the other hand, for any $0 < \varepsilon < CLR^*\left(s,t\right)$,

$$\begin{aligned} \mathcal{L}^{w_{CLR,N}}\left(s,t\right) \exp\left(\frac{N}{2}\right) &= \int_{\mathbb{R}} {}_0\tilde{F}_1\left[; \frac{k}{2}; \frac{N}{4}\tilde{\eta}_\beta' Q\left(s,t\right)\tilde{\eta}_\beta\right] d\Phi\left(\beta\right) \\ &\geq \int_{\mathbb{B}_\varepsilon(s,t)} {}_0\tilde{F}_1\left[; \frac{k}{2}; \frac{N}{4}\tilde{\eta}_\beta' Q\left(s,t\right)\tilde{\eta}_\beta\right] d\Phi\left(\beta\right) \\ &\geq \exp\left[\sqrt{N}\left(CLR^*\left(s,t\right) - \varepsilon\right)\right] \\ &\quad \times \underline{C} \max\left[N \cdot CLR^*\left(s,t\right)^2, 1\right]^{-(k-1)/4} \int_{\mathbb{B}_\varepsilon(s,t)} d\Phi\left(\beta\right), \end{aligned}$$

where

$$\mathbb{B}_{\varepsilon}(s,t) = \left\{\beta : \sqrt{\tilde{\eta}_{\beta}' Q(s,t)\,\tilde{\eta}_{\beta}} \geq CLR^{*}(s,t) - \varepsilon\right\},$$

the first inequality uses positivity of ${}_0\tilde{F}_1\left[;k/2;\cdot\right]$, and the last inequality uses (21) and monotonicity of ${}_0\tilde{F}_1\left[;k/2;\cdot\right]$. The integral $\int_{\mathbb{B}_{\varepsilon}(s,t)} d\Phi(\beta)$ is strictly positive because $\Phi(\cdot)$ has full support, so

$$\underline{\lim}_{N\to\infty}\frac{1}{\sqrt{N}}\left[\log\mathcal{L}^{wCLR,N}(s,t)+\frac{N}{2}\right] \geq CLR^{*}(s,t)-\varepsilon.$$

Letting ε tend to zero in the displayed inequality, we obtain an inequality which can be combined with (26) to yield (22).

Indeed, because

$$\sup_{(s',t')'\in K} CLR^{*}(s,t) < \infty \qquad \text{and} \qquad \inf_{(s',t')'\in K}\int_{\mathbb{B}_{\varepsilon}(s,t)} d\Phi(\beta) > 0$$

for any compact set $K \subset \mathbb{R}^{2k}$, the result (22) can be strengthened as follows:

$$\lim_{N\to\infty}\sup_{(s',t')'\in K}\left|\frac{1}{\sqrt{N}}\left[\log\mathcal{L}^{wCLR,N}(s,t)+\frac{N}{2}\right]-CLR^{*}(s,t)\right| = 0. \qquad \blacksquare$$

REFERENCES

ANDERSON, T. W., AND H. RUBIN (1949): "Estimation of the Parameters of a Single Equation in a Complete Set of Stochastic Equations," *Annals of Mathematical Statistics*, 20, 46–63.

ANDREWS, D. W. K. (1996): "Admissibility of the Likelihood Ratio Test When the Parameter Space is Restricted under the Alternative," *Econometrica*, 64, 705–718.

ANDREWS, D. W. K., M. J. MOREIRA, AND J. H. STOCK (2004): "Optimal Invariant Similar Tests for Instrumental Variables Regression," NBER Technical Working Paper No. 299.

ANDREWS, D. W. K., AND W. PLOBERGER (1995): "Admissibility of the Likelihood Ratio Test When a Nuisance Parameter is Present Only under the Altenative," *Annals of Statistics*, 23, 1609–1629.

ANDREWS, D. W. K., AND J. H. STOCK (2006): "Inference with Weak Instruments," Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society, forthcoming.

CHAMBERLAIN, G. (2006): "Decision Theory Applied to an Instrumental Variables Model," Econometrica, forthcoming.

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

——— (2003): "Identification, Weak Instruments, and Statistical Inference in Econometrics," *Canadian Journal of Economics*, 36, 767–808.

FERGUSON, T. S. (1967): *Mathematical Statistics: A Decision Theoretic Approach.* New York: Academic Press.

HAHN, J., AND J. HAUSMAN (2003): "Weak Instruments: Diagnosis and Cures in Empirical Econometrics," *American Economic Review*, 93, 118–125.

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

LEHMANN, E. L., AND J. P. ROMANO (2005): *Testing Statistical Hypotheses.* Third Edition. New York: Springer.

MOREIRA, M. J. (2001): "Tests with Correct Size When Instruments Can Be Arbitrarily Weak," CLE Working Paper No. 37, UC Berkeley.

——— (2003): "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica*, 71, 1027–1048.

MUIRHEAD, R. J. (1982): *Aspects of Multivariate Statistical Theory.* New York: Wiley.

STAIGER, D., AND J. H. STOCK (1997): "Instrumental Variables Estimation with Weak Instruments," *Econometrica*, 65, 557–586.

STOCK, J. H., J. H. WRIGHT, AND M. YOGO (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, 20, 518–529.

5. APPENDIX: OMITTED PROOFS

This Appendix provides a proof of (24) and a derivation of (11) – (12).

5.1. Proof of (24). To motivate the result, notice that if $w_{LM,N}$ is such that its support consists of all pairs (β, λ) for which $|\beta| \leq N^{-1/3}$ and $\lambda \eta_\beta' \eta_\beta = N$, then the large N behavior of the integrand in (13) is reminiscent of the behavior of

$$\exp\left(-\frac{N}{2}\right) {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\left(t't + 2\beta s't\right)\right].$$

Because

$$\begin{aligned}\max_{|\beta|\leq N^{-1/3}} \sqrt{t't + 2\beta s't} &= \sqrt{t't + 2N^{-1/3}|s't|} \\ &= \sqrt{t't} + N^{-1/3}|s't| / \sqrt{t't} + o\left(N^{-1/3}\right),\end{aligned}$$

reasoning similar to that leading to Theorem 3 therefore suggests that $|s't| / \sqrt{t't}$ can be represented as the limit as $N \to \infty$ of a suitably normalized version of $\mathcal{L}^{w_{LM,N}}(s,t)$.

As in Section 3.3, let $w_{LM,N}$ denote the cdf of $\left[\mathcal{B}_N, N/\sqrt{\eta_{\mathcal{B}_N}' \eta_{\mathcal{B}_N}}\right]'$, where the distribution of $N^{1/3}\mathcal{B}_N$ is uniform on $[-1, 1]$. Then

$$\begin{aligned}\mathcal{L}^{w_{LM,N}}(s,t) &= \int_{\mathbb{R}\times\mathbb{R}_+} \exp\left(-\frac{\lambda \eta_\beta' \eta_\beta}{2}\right) {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{1}{4}\lambda \eta_\beta' Q(s,t)\,\eta_\beta\right] dw_{LM,N}(\beta,\lambda) \\ &= \exp\left(-\frac{N}{2}\right)\frac{N^{1/3}}{2}\int_{[-N^{-1/3},N^{-1/3}]} {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\frac{\eta_\beta' Q(s,t)\,\eta_\beta}{\eta_\beta'\eta_\beta}\right] d\beta.\end{aligned}$$

The integrand can be written as

$${}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\left(t't + 2\beta s't + R_N(\beta;s,t)\right)\right],$$

where, for some finite constant $\overline{R}$ (depending only on s and t),

$$\sup_{|\beta|\leq N^{-1/3}} |R_N(\beta;s,t)| = \sup_{|\beta|\leq N^{-1/3}} \left|\frac{\eta_\beta' Q(s,t)\,\eta_\beta}{N\eta_\beta'\eta_\beta} - \left(t't + 2\beta s't\right)\right| \leq N^{-2/3}\overline{R}.$$

As a consequence, by monotonicity of ${}_0\tilde{F}_1\left[;k/2;\cdot\right]$ and the defining property of $\overline{C}$ (appearing in the proof of Theorem 3),

$$
\begin{aligned}
& \mathcal{L}^{wLM,N}(s,t)\exp\left(\frac{N}{2}\right) \\
\leq\ & \frac{\sqrt[3]{N}}{2}\int_{[-1/\sqrt[3]{N},1/\sqrt[3]{N}]} {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\left(t't+2\beta s't+N^{-2/3}\overline{R}\right)\right]d\beta \\
\leq\ & \sup_{|\beta|\leq N^{-1/3}} {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\left(t't+2\beta s't+N^{-2/3}\overline{R}\right)\right] \\
=\ & {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\left(t't+2N^{-1/3}\left|s't\right|+N^{-2/3}\overline{R}\right)\right] \\
\leq\ & \overline{C}\exp\left[\sqrt{N}\sqrt{t't+2N^{-1/3}\left|s't\right|+N^{-2/3}\overline{R}}\right],
\end{aligned}
$$

from which it follows that

$$
\begin{aligned}
& N^{-1/6}\left[\log\mathcal{L}^{wLM,N}(s,t)+\frac{N}{2}-\sqrt{N}\sqrt{t't}\right] \\
\leq\ & N^{-1/6}\log\overline{C}+N^{1/3}\sqrt{t't+2N^{-1/3}\left|s't\right|+N^{-2/3}\overline{R}}-N^{1/3}\sqrt{t't} \\
\leq\ & N^{-1/6}\log\overline{C}+\left|s't\right|/\sqrt{t't}+N^{-1/3}\frac{\overline{R}}{2\sqrt{t't}} \\
=\ & \left|s't\right|/\sqrt{t't}+O\left(N^{-1/6}\right),
\end{aligned}
$$

where the last inequality uses concavity of $\sqrt{\cdot}$.

To obtain an inequality in the opposite direction, we distinguish between the cases where $t=0$ and $t\neq 0$. If $t=0$, then

$$
\begin{aligned}
& N^{-1/6}\left[\log \mathcal{L}^{w_{LM,N}}(s,t)+\frac{N}{2}-\sqrt{N}\sqrt{t't}\right] \\
= \;& \frac{N^{1/6}}{2}\int_{[-N^{-1/3},N^{-1/3}]} {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\frac{\eta_\beta' Q(s,t)\eta_\beta}{\eta_\beta'\eta_\beta}\right] d\beta \\
\geq \;& N^{-1/6}\,{}_0\tilde{F}_1\left[;\frac{k}{2};0\right] = O\left(N^{-1/6}\right),
\end{aligned}
$$

which completes the proof of (24) in the case where $t=0$.

On the other hand, if $t \neq 0$, then it follows from monotonicity and positivity of ${}_0\tilde{F}_1\left[;k/2;\cdot\right]$ that

$$
\begin{aligned}
& \mathcal{L}^{w_{LM,N}}(s,t)\exp\left(\frac{N}{2}\right) \\
\geq \;& \frac{N^{1/3}}{2}\int_{[-N^{-1/3},N^{-1/3}]} {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\left(t't+2\beta s't-N^{-2/3}\overline{R}\right)\right] d\beta \\
\geq \;& \frac{N^{1/3}}{2}\int_{[0,N^{-1/3}]} {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\left(t't+2\beta\left|s't\right|-N^{-2/3}\overline{R}\right)\right] d\beta
\end{aligned}
$$

for every $N > \underline{N} = \inf\left\{N : t't - 2N^{-1/3}\left|s't\right| - N^{-2/3}\overline{R} > 0\right\}$. Therefore, for any $0<\varepsilon<1$ and any $N>\underline{N}$,

$$
\begin{aligned}
& \mathcal{L}^{w_{LM,N}}(s,t)\exp\left(\frac{N}{2}\right) \\
\geq \;& \frac{N^{1/3}}{2}\int_{[(1-\varepsilon)N^{-1/3},N^{-1/3}]} {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\left(t't+2\beta\left|s't\right|-N^{-2/3}\overline{R}\right)\right] d\beta \\
\geq \;& \frac{\varepsilon}{2}\inf_{\beta\in[(1-\varepsilon)N^{-1/3},N^{-1/3}]} {}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\left(t't+2\beta\left|s't\right|-N^{-2/3}\overline{R}\right)\right] \\
= \;& \frac{\varepsilon}{2}\,{}_0\tilde{F}_1\left[;\frac{k}{2};\frac{N}{4}\left(t't+2(1-\varepsilon)N^{-1/3}\left|s't\right|-N^{-2/3}\overline{R}\right)\right] \\
\geq \;& \frac{\varepsilon}{2}\underline{C}\left[Nt't-N^{1/3}\overline{R}\right]^{-(k-1)/4}\exp\left[\sqrt{N}\sqrt{t't+2(1-\varepsilon)N^{-1/3}\left|s't\right|-N^{-2/3}\overline{R}}\right],
\end{aligned}
$$

where the first inequality uses positivity of ${}_0\tilde{F}_1\left[;k/2;\cdot\right]$ and the last inequality uses the defining property of $\underline{C}$ (appearing in the proof of Theorem 3) and monotonicity of $[\cdot]^{-(k-1)/4}$.

By implication,

$$\begin{aligned}
& N^{-1/6}\left[\log \mathcal{L}^{w_{LM,N}}(s,t)+\frac{N}{2}-\sqrt{N}\sqrt{t't}\right] \\
\geq\ & N^{-1/6}\left[\log\left(\underline{C}\varepsilon/2\right)-\frac{k-1}{4}\log\left(Nt't-N^{1/3}\overline{R}\right)\right] \\
& +N^{1/3}\sqrt{t't+2\left(1-\varepsilon\right)N^{-1/3}\left|s't\right|-N^{-2/3}\overline{R}}-N^{1/3}\sqrt{t't} \\
\geq\ & \frac{\left(1-\varepsilon\right)\left|s't\right|-N^{-1/3}\overline{R}/2}{\sqrt{t't+\max\left[2\left(1-\varepsilon\right)N^{-1/3}\left|s't\right|-N^{-2/3}\overline{R},0\right]}}+O\left(N^{-1/6}\log N\right) \\
\geq\ & \left(1-\varepsilon\right)\left|s't\right|/\sqrt{t't}+O\left(N^{-1/6}\log N\right).
\end{aligned}$$

The proof of (24) can now be completed by letting ε tend to zero in the preceding display.

5.2. Derivation of (11) – (12). The test based on $\phi(S,T)$ is α-similar if and only if

$$\int_{\mathbb{R}^k}\int_{\mathbb{R}^k}\phi(s,t)\,f_S(s|0,\pi)\,f_T(t|0,\pi)\,dsdt=\alpha\qquad\forall\pi.$$

Because T is complete under $H_0:\beta=0$, the preceding condition holds if and only if

$$\int_{\mathbb{R}^k}\phi(s,t)\,f_S(s|0,\pi)\,ds=\alpha$$

for almost every $t\in\mathbb{R}^k$. Furthermore, $f_S(s|0,\pi)$ does not depend on π, so the test based on $\phi(S,T)$ is α-similar if and only if

$$\int_{\mathbb{R}^k}\phi(s,t)\,f_S(s|0,0)\,ds=\alpha$$

for almost every $t\in\mathbb{R}^k$.

Using (the Fubini theorem and) the representation (9) , we can write (8) as

$$\int_{\mathbb{R}^k}\int_{\mathbb{R}^k}\phi(s,t)\left[\int_{\mathbb{R}^{k+1}} f_S(s|\beta,\pi)\, f_T(t|\beta,\pi)\, dW(\beta,\pi)\right] dsdt.$$

To maximize this expression subject to the condition that the test based on $\phi(S,T)$ is α-similar, we can proceed on a "t by t" basis and consider the problem of maximizing

$$\int_{\mathbb{R}^k}\phi(s,t)\left[\int_{\mathbb{R}^{k+1}} f_S(s|\beta,\pi)\, f_T(t|\beta,\pi)\, dW(\beta,\pi)\right] ds$$

subject to the condition

$$\int_{\mathbb{R}^k}\phi(s,t)\, f_S(s|0,0)\, ds = \alpha.$$

By the Neyman-Pearson lemma, this maximization problem is solved by

$$\tilde{\phi}^W_{LR}(s,t;\alpha) = 1\left[\widetilde{LR}^W(s,t) > \tilde{\kappa}^W_{LR}(t;\alpha)\right],$$

where

$$\widetilde{LR}^W(s,t) = \frac{\int_{\mathbb{R}^{k+1}} f_S(s|\beta,\pi)\, f_T(t|\beta,\pi)\, dW(\beta,\pi)}{f_S(s|0,0)}$$

and $\tilde{\kappa}^W_{LR}(t;\alpha)$ is the $1-\alpha$ quantile of the distribution of $\widetilde{LR}^W(\mathcal{Z}_k,t)$. Finally, because $f_T(t|0,0)$ is positive we obtain the equality

$$\tilde{\phi}^W_{LR}(s,t;\alpha) = \phi^W_{LR}(s,t;\alpha).$$