SIMILAR-ON-THE-BOUNDARY TESTS FOR MOMENT INEQUALITIES EXIST, BUT HAVE POOR POWER

By

Donald W. K. Andrews

August 2011
Revised March 2012

COWLES FOUNDATION DISCUSSION PAPER NO. 1815R

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

http://cowles.econ.yale.edu/

Similar-on-the-Boundary Tests for Moment Inequalities Exist, But Have Poor Power

Donald W. K. Andrews*

Cowles Foundation for Research in Economics

Yale University

April 2010

Revised: March 2012

*The author gratefully acknowledges the research support of the National Science Foundation via grant numbers SES-0751517 and SES-1021101. The author thanks Patrik Guggenberger, Kei Hirano, a referee, and the co-editor for very helpful comments and references.

Abstract

This paper shows that moment inequality tests that are asymptotically similar on the boundary of the null hypothesis exist, but have very poor power. Hence, existing tests in the literature, which are asymptotically non-similar on the boundary, are not deficient. The results are obtained by first establishing results for the finite-sample multivariate normal one-sided testing problem. Then, these results are shown to have implications for more general moment inequality tests that are used in the literature on partial identification.

Keywords: Moment inequality, one-sided test, power, similar, test.

JEL Classification Numbers: C12, C15.

1. Introduction

This paper is concerned with tests of moment inequalities. This has been an active area of econometrics recently because of the usefulness of such tests in carrying out inference in models that may be only partially identified. They are also useful in other contexts, see below. All of the tests proposed in the econometrics literature on moment inequalities are asymptotically non-similar on the boundary (in a uniform sense) and hence are asymptotically biased. This raises the question of whether asymptotically similar-on-the-boundary tests exist and have desirable power properties.

We answer this question by first posing it in a finite-sample setting under the assumption of normality. Then, results for this case are converted into results for the asymptotic problem. The relevant finite-sample problem is that of testing a multivariate one-sided null hypothesis based on a normal random vector with known covariance matrix. Suppose $X \sim N(\mu, \Sigma)$, where $X, \mu \in R^p$ and $\Sigma \in R^{p \times p}$ is a known positive-definite matrix. The hypotheses of interest are

$$H_0 : \mu \geq 0 \text{ and } H_1 : \mu \ngeq 0. \tag{1.1}$$

In this paper, we show that there exist similar-on-the-boundary tests of H_0, but that their power properties are very poor in the sense that their power against some alternatives that are arbitrarily far from the null is equal to their size. These results are established using the properties of complete sufficient statistics. Then, we show that these results imply analogous asymptotic results for tests of moment inequalities. We conclude that existing moment inequality tests are not deficient due to their property of asymptotic non-similarity on the boundary of the null. Rather, any test with good overall power necessarily must be asymptotically non-similar on the boundary.

In some cases that arise in the moment inequality literature, the matrix Σ that arises is singular. For example, this occurs in the missing data example in Imbens and Manski (2004) because the lower and upper bounds are determined by the same random variable. Also see Stoye (2009), whose "super-efficient" case corresponds to a singular matrix Σ asymptotically. In such cases, the results of this paper do not apply.[1]

The results of this paper are relevant not just to the partial-identification moment inequality literature. They also apply to (asymptotic) tests of (i) stochastic dominance, (ii) model superiority based on predictive performance, see Hansen (2005), (iii) concavity

[1] This is not true of all, or even most, missing data problems.

and other restrictions in models of consumer behavior and producer technology, and (iv) multiperiod inequalities due to liquidity risk premiums in financial models. See Chen and Szroeter (2009) for references concerning these applications.

Now we discuss some related literature. An early paper by Fraser (1952) shows that for $X \sim N(\mu, I_p)$ no upper confidence bound for $\theta = \max\{\mu_1, ..., \mu_p\}$ of the form $(-\infty, g(X)]$ exists whose coverage probability is $1 - \alpha$ for all $\mu \in R^p$, under some monotonicity restrictions on $g(\cdot)$. Upper confidence bounds for $\max_{j \leq p} \mu_j$ of the form $(-\infty, g(X)]$ are obtained by inverting tests of $H_0^* : \mu \not< 0$ versus $H_1^* : \mu < 0$.[2] These hypotheses are not the same as the hypotheses in (1.1). For example, for $p = 2$, H_0 can written as $H_0 : \mu_1 \geq 0$ & $\mu_2 \geq 0$, whereas H_0^* is $H_0^* : \mu_1 \geq 0$ or $\mu_2 \geq 0$. Hence, Fraser's (1952) results do not apply to the hypotheses of interest in the moment inequality literature in econometrics that is concerned with partial identification.[3]

The paper by Blumenthal and Cohen (1968) is related to the present paper, but it focuses on estimators, not tests. Blumenthal and Cohen (1968) establish the non-existence of unbiased estimators of $h(\mu) = \min\{\mu_1, \mu_2\}$ when $X \sim N(\mu, I_2)$. Hirano and Porter (2012) provide results that encompass this result. Their results apply when (i) $X \sim N(\mu, \Sigma)$ for known, positive-definite, $p \times p$ variance matrix Σ, (ii) $h(\mu)$ is any non-differentiable function of μ, (iii) unbiased or quantile-unbiased estimators are considered, and/or (iv) the problem of interest has a limit experiment of the form just specified. These results can be used to establish the non-existence of mean- and quantile-unbiased estimators of the endpoint of an identified set based on moment inequalities. The results of the present paper also can be used to establish the non-existence of quantile-unbiased estimators of the endpoint of an identified set based on moment inequalities, see Sections 3 and 4. Asymptotically half-median-unbiased estimators of identified sets have been considered in Andrews and Shi (2007) and Chernozhukov, Lee, and Rosen (2008). The

[2]For a parameter $\theta \in R$ and data vector Y, an upper confidence bound of the form $(-\infty, L(Y)]$ can be constructed by inverting one-sided tests of the form $H_0 : \theta \geq \theta^*$ versus $H_1 : \theta < \theta^*$ for arbitrary $\theta^* \in R$. For example, if one observes $Y \sim N(\theta, 1)$, then a test of $H_0 : \theta \geq \theta^*$ versus $H_1 : \theta < \theta^*$ is a test that rejects H_0 when Y is small. One rejects H_0 if $Y - \theta^* < z_\alpha$ or one accepts H_0 if $\theta^* \leq Y - z_\alpha = Y + z_{1-\alpha}$. The latter yields the upper confidence bound $(-\infty, Y + z_{1-\alpha}]$. Analogously, the tests that correspond to the upper confidence bound $(-\infty, g(X)]$ of Fraser (1952) for $\theta = \max_{j \leq p} \mu_j$ are tests of $H_0^* : \max_{j \leq p} \mu_j \geq a$ versus $H_1^* : \max_{j \leq p} \mu_j < a$ for $a \in R$, which is equivalent to $H_0^* : \mu \not< 0$ versus $H_1^* : \mu < 0$ when $a = 0$.

[3]Fraser (1952) also considers confidence intervals $[h(X), g(X)]$ for which $P_\mu(h(X) \leq \min_{j \leq p} \mu_j \leq \max_{j \leq p} \mu_j \leq g(X)) \geq 1 - \alpha$. Note that $[h(X), g(X)]$ is not a two-sided confidence interval for $\max_{j \leq p} \mu_j$. The hypotheses that correspond to the lower bound $h(X)$ on $\min_{j \leq p} \mu_j$ are analogous to the hypotheses for the upper bound $g(X)$ on $\max_{j \leq p} \mu_j$. In consequence, Fraser's (1952) results for confidence intervals of this type also are not relevant to the testing problem in (1.1).

results just mentioned imply that asymptotically (fully) median-unbiased estimators do not exist.

The results of Hirano and Porter (2012) also have implications for inference. Their results show that there exist no locally asymptotically similar one-sided confidence intervals for parameters of the form $\min\{\mu_1, ..., \mu_p\}$ under their conditions. Their confidence intervals are restricted to be of the form $(-\infty, T]$ or $[T, \infty)$, where T is some possibly-randomized statistic. Although natural, this restriction rules out confidence intervals that are disconnected, bounded above and below, and/or randomized in complicated ways. In contrast, our testing results (and corresponding confidence set results) place no restrictions on the form of the test (or corresponding confidence interval). In consequence, instead of the non-existence result in Hirano and Porter (2012), we obtain an existence result for a test that is similar-on-the-boundary, but also show that it and all other similar-on-the-boundary tests have very poor power. Note that the methods of proof in Hirano and Porter (2012) and the present paper are quite different.

The results of this paper stand in contrast somewhat to results in the weak instruments (IV's) literature for the linear IV regression model. In the weak instruments literature, it has been shown that standard tests, such as the likelihood ratio test based on a fixed critical value, are not asymptotically similar. Nevertheless, asymptotically similar tests exist, such as the conditional likelihood ratio test which employs the likelihood ratio statistic and a data-dependent critical value, see Moreira (2003, 2009). In addition, it has been shown that the latter test has very good power properties, see Andrews, Moreira, and Stock (2006, 2008). In the testing problems considered here, however, tests that are similar on the boundary have poor power. See Moreira and Moreira (2011) for recent results concerning tests that maximize weighted average power among similar tests for a broad class of models.

The remainder of the paper is organized as follows. Section 2 motivates interest in similar-on-the-boundary tests. To most clearly elucidate the results of the paper and their proof, Section 3 considers the bivariate normal X case with known variance matrix I_2. Section 4 generalizes the results to the finite-sample case of real interest for the asymptotics of moment inequality tests, which is the p-variate case with known positive-definite covariance matrix Σ. Finally, Section 5 develops the implications of the preceding sections for general moment inequality tests, which may involve non-normal random variables and unknown covariance matrix.

2. Motivation

We now provide motivation for interest in similar-on-the-boundary tests from a power perspective. For simplicity, suppose $\Sigma = I_p$. For illustrative purposes, consider the LR test statistic for testing the hypotheses in (1.1). When $\Sigma = I_p$, it equals

$$LR = \sum_{J=1}^{p} X_j^2 1(X_j < 0). \tag{2.1}$$

If one uses a non-data-dependent critical value, say cv_α, where α is the significance level, then the least-favorable null parameter value is $\mu = 0$. This critical value yields a test that is markedly non-similar on the boundary of the null hypothesis. For example, for $\alpha = .05$ and all $\mu = (0, m_2, ..., m_p)' \in R^p$ with $m_j \in [1.5, \infty)$ for $2 \leq j \leq p$, the "fixed-critical-value" LR test has null rejection probabilities in $[.020, .021]$ for $p = 2$, in $[.0029, .0032]$ for $p = 5$, and in $[.00038, .00043]$ for $p = 10$.[4] Note that these null rejection probabilities are much less than the significance level .05 and decrease rapidly as p increases.

These results show that the bias of the LR test can be substantial. In consequence, the LR test has poor power against alternatives of the form $\mu = (-c, m_2, ..., m_p)' \in R^p$ for $c > 0$ and $m_j \in [1.5, \infty)$ for $2 \leq j \leq p$. More generally, the LR test has relatively low power for alternatives with some non-violated inequalities (SNVI), i.e., μ vectors with some negative elements and some elements that are positive and moderately large.

The reason for the non-similarity on the boundary is that the least-favorable critical value is too large when some elements of μ are positive and moderately large because the corresponding elements of X do not contribute to the test statistic LR (with high probability) or its distribution. Thus, the question arises: Can the critical value be altered in a data-dependent way to reduce, or even eliminate, the non-similarity on the boundary of the test and to improve its power against SNVI alternatives?

A moderately large positive value of X_1, say, indicates that $\mu_1 > 0$. So, in this case one would want to use a critical value that is smaller than otherwise. This motivates consideration of tests of the form: Reject H_0 if

$$LR > m(X), \tag{2.2}$$

[4] These results are determined via simulation using $100,000$ simulation repetitions.

where the data-dependent critical value $m(X)$ satisfies $\sup_{\mu \in R^p_+} P(LR > m(X)) \leq \alpha$. A function $m(x)$ that reduces the magnitude of non-similarity on the boundary and improves the power against SNVI alternatives has the property that it is decreasing in x_j given x_{-j} (which equals x with x_j deleted) for x_j large. For example, for $p = 2$, a good choice of function $m(x)$ to reduce non-similarity and bias is one that decreases in x_1 or x_2 for large enough values and asymptotes to a value slightly larger than $z^2_{1-\alpha}$ $(< \chi^2_{1,1-\alpha})$ as $x \to \infty$.

The finite-sample versions of the tests in Andrews and Soares (2010), Bugni (2010), Canay (2010), and Andrews and Barwick (2012) all are of the form in (2.2). These tests have noticeably reduced non-similarity on the boundary of the null compared to the fixed-critical-value LR test and higher power against SNVI alternatives. However, none is similar on the boundary. This raises the question. Do tests that are similar on the boundary exist? If so, do they have good power properties? These are the questions addressed in this paper.

3. Independent Bivariate Normal Mean Model

In this section, we provide finite-sample results for the simplest moment inequality model—a bivariate normal model with mean $\mu \in R^2$ and variance matrix I_2. Let $X \sim N(\mu, I_2)$. We consider tests of the null hypothesis $H_0 : \mu \geq 0$ versus the alternative hypothesis $H_1 : \mu \ngeq 0$.

The boundary of the null hypothesis is

$$B = \{\mu = (\mu_1, \mu_2)' : \mu \geq 0 \ \& \ \mu_1 = 0 \text{ or } \mu_2 = 0\}. \tag{3.1}$$

Theorem 1. *Let $X \sim N(\mu, I_2)$. Any (possibly randomized) test of the null hypothesis $H_0 : \mu \geq 0$ that is similar on the boundary B with rejection probability $\alpha \in (0, 1)$ on B has rejection probability α for all μ in $B^* = \{\mu = (\mu_1, \mu_2)' : \mu_1 = 0$ or $\mu_2 = 0\}$.*

Comments. 1. Theorem 1 says that a similar-on-the-boundary test (with rejection probability α on the boundary) has trivial power (i.e., power equal to α, which is less than or equal to the size of the test) for all alternatives that consist of the violation of one inequality with the other being binding, such as $\mu_1 = 0$ and $\mu_2 < 0$. Such alternatives include alternatives that are arbitrarily far from the null hypothesis. In consequence, Theorem 1 implies that the power properties of tests that are similar on the boundary are very poor.

2. Theorem 1 also holds for the mixed equality/inequality null hypothesis $H_0 : \mu_1 = 0$ & $\mu_2 \geq 0$ and the alternative $H_1 : \mu_1 \neq 0$ or $\mu_2 < 0$.[5] In this case, the boundary of the null is the null itself, i.e., $B = \{\mu = (\mu_1, \mu_2)' : \mu_1 = 0 \text{ and } \mu_2 \geq 0\}$, and $B^* = \{\mu = (\mu_1, \mu_2)' : \mu_1 = 0\}$. In the mixed equality/inequality case, a similar-on-the-boundary test has power equal to size α for all alternatives that do not involve a violation of the equality restriction $\mu_1 = 0$. For such alternatives, power equals size no matter how far μ is from the null hypothesis. Hence, in this case too, similar-on-the-boundary tests have very poor power properties.

3. Theorem 1 also holds if the null hypothesis H_0 is restricted such that its boundary includes just the set $\{\mu : \mu_1 = 0 \ \& \ \mu_2 \in [a_2, b_2]\} \cup \{\mu : \mu_2 = 0 \ \& \ \mu_1 \in [a_1, b_1]\}$ for some $0 \leq a_j < b_j < \infty$ for $j = 1, 2$.[6] Furthermore, if the null hypothesis H_0 is restricted such that its boundary includes just the set $\{\mu : \mu_1 = 0 \ \& \ \mu_2 \in [a_2, b_2]\}$ for some $0 \leq a_2 < b_2 < \infty$, then the result of Theorem 1 holds with B^* replaced by $B_1^* = \{\mu = (\mu_1, \mu_2)' : \mu_1 = 0\}$. These extensions have useful implications in the moment inequality model discussed in Section 5.

4. Theorem 1 can be used to show that median- and quantile-unbiased estimators of $\kappa = \min\{\mu_1, \mu_2\}$ do not exist. To see this, suppose a median-unbiased estimator $\widehat{\kappa}$ of κ exists. By definition, it has the property that $P_\mu(\widehat{\kappa} \leq \kappa) = 1/2 = P_\mu(\widehat{\kappa} \geq \kappa) \ \forall \mu \in R^2$. For any $a \in R$, consider the test of $H_0^\dagger : \kappa = a$ versus $H_1^\dagger : \kappa < a$ that rejects $H_0^\dagger$ if $\widehat{\kappa} \leq a$. This test has level α because $P_\mu(\widehat{\kappa} \leq a) = 1/2$ for all μ with $\kappa = \min\{\mu_1, \mu_2\} = a$. In other words, $P_\mu(\widehat{\kappa} \leq a) = 1/2 \ \forall \mu \in B + (a, a)'$. By Theorem 1, the latter implies that $P_\mu(\widehat{\kappa} \leq a) = 1/2 \ \forall \mu \in B^* + (a, a)'$.[7] In turn, for any $c > 0$, this gives: for $\mu_{ac} = (a, a - c)'$ (which is in $B^* + (a, a)'$), $P_{\mu_{ac}}(\widehat{\kappa} \leq a) = 1/2$. But, the value of κ corresponding to μ_{ac} is $\kappa_{\mu_{ac}} = \min\{a, a - c\} = a - c$. So, $P_{\mu_{ac}}(\widehat{\kappa} \leq \kappa_{\mu_{ac}}) = P_{\mu_{ac}}(\widehat{\kappa} \leq a - c) = 1/2$ by median unbiasedness. Combining these results gives $P_{\mu_{ac}}(\widehat{\kappa} \in (a - c, a]) = 0 \ \forall a \in R, \forall c > 0$. That is, $P_{\mu_{ac}}(\widehat{\kappa} \in R) = 0$, which is a contradiction. Hence, no median-unbiased estimator of $\min\{\mu_1, \mu_2\}$ exists.

[5] The proof is essentially the same as the proof of Theorem 1.

[6] In the proof of Theorem 1, the smaller boundary of the null considered here implies that X_2 is a complete sufficient statistic for μ_2 in the model $X_2 \sim N(\mu_2, 1)$ and $\mu_2 \in [a_2, b_2]$. This is enough for the rest of the proof to go through unchanged.

[7] More precisely, this holds by Comment 3 to Theorem 1 with the null hypothesis given by $H_0 : \mu \in B$ or its translation $B + (a, a)'$.

Proof of Theorem 1. Let $\xi(X)$ be a randomized test. (That is, the test ξ rejects H_0 with probability $\xi(X)$ ($\in [0,1]$)). Suppose $\xi(X)$ is similar on the boundary B with rejection probability $\alpha \in (0,1)$ on B. That is,

$$E_\mu \xi(X) = \int\int \xi(x_1, x_2)\phi(x_1 - \mu_1)\phi(x_2 - \mu_2)dx_1 dx_2 = \alpha \ \forall \mu \in B, \tag{3.2}$$

where ϕ denotes the standard normal density.

For $\mu_1 = 0$ and all $\mu_2 \geq 0$, this gives

$$\begin{aligned} E_{0,\mu_2}\xi(X) &= E_{\mu_2} g(X_2) = \int g(x_2)\phi(x_2 - \mu_2)dx_2 = \alpha, \text{ where} \\ g(x_2) &= \int \xi(x_1, x_2)\phi(x_1)dx_1. \end{aligned} \tag{3.3}$$

In the model $X_2 \sim N(\mu_2, 1)$ and $\mu_2 \geq 0$, the random variable X_2 is a complete sufficient statistic for μ_2.[8] This, (3.3), and the definition of completeness give

$$g(x_2) = \alpha \ \forall x_2 \in \mathcal{X}_2 \text{ for some set } \mathcal{X}_2 \text{ with } P_{\mu_2}(X_2 \in \mathcal{X}_2) = 1 \ \forall \mu_2 \geq 0. \tag{3.4}$$

By the absolute continuity of any (nondegenerate) normal distribution with respect to any other (nondegenerate) normal distribution, we have

$$P_{\mu_2}(X_2 \in \mathcal{X}_2) = 1 \ \forall \mu_2 \in R. \tag{3.5}$$

Hence, (3.3) holds for all $\mu_2 \in R$. It also holds with the roles of μ_1 and μ_2 reversed and the proof is complete. □

The result of Theorem 1 begs the question of whether any non-trivial similar-on-the-boundary test exists. By non-trivial, we mean a level α test whose power function is greater than α somewhere in the alternative.[9] The answer is yes. We provide a constructive proof.

Theorem 2. *Let $X \sim N(\mu, I_2)$. There exists a non-trivial test of level $\alpha \in (0,1)$ for the null hypothesis $H_0 : \mu \geq 0$ that is similar on the boundary B.*

[8] As is well known, this holds because the normal distribution with unknown mean and known variance is in the exponential family and the parameter space for μ_2 includes a (one-dimensional) rectangle, see Theorem 4.3.1 of Lehmann and Romano (2005, p. 117).

[9] A randomized test that rejects the null with probability α regardless of the data X obviously is similar on the boundary and level α. But, it is a *trivial* similar-on-the-boundary level α test.

Comment. The result of Theorem 2 holds for any nonsingular diagonal 2×2 matrix Σ. The same test as considered in the proof of Theorem 2 has the desired properties.[10]

Proof of Theorem 2. Consider the following (randomized) test

$$\xi(X_1, X_2) = \begin{cases} 2\alpha & \text{if } (X_1 \geq 0 \ \& \ X_2 \leq 0) \text{ or } (X_1 \leq 0 \ \& \ X_2 \geq 0) \\ 0 & \text{elsewhere.} \end{cases} \tag{3.6}$$

The power of this test against μ is

$$\begin{aligned} E_\mu \xi(X) &= 2\alpha \int_{-\infty}^{0} \int_{0}^{\infty} \phi(x_1 - \mu_1)\phi(x_2 - \mu_2) dx_1 dx_2 \\ &\quad + 2\alpha \int_{0}^{\infty} \int_{-\infty}^{0} \phi(x_1 - \mu_1)\phi(x_2 - \mu_2) dx_1 dx_2 \\ &= 2\alpha(1 - \Phi(\mu_2))\Phi(\mu_1) + 2\alpha\Phi(\mu_2)(1 - \Phi(\mu_1)) \\ &= 2\alpha(\Phi(\mu_1) + \Phi(\mu_2) - 2\Phi(\mu_1)\Phi(\mu_2)), \end{aligned} \tag{3.7}$$

where the second equality holds by change of variables with $z_j = -(x_j - \mu_j)$ for $j = 1, 2$.

If $\mu_1 = 0$ or $\mu_2 = 0$, then the right-hand side (rhs) of (3.7) equals α. Hence, the test $\xi(X)$ is similar on the boundary.

If μ_2 is arbitrarily large and $\mu_1 = -\mu_2$, then the rhs of (3.7) is arbitrarily close to 2α. Hence, the test has non-trivial power.

The derivative of the test's power with respect to μ_2 is negative when $\mu_1 > 0$ and vice versa:

$$\frac{\partial}{\partial \mu_2}[2\alpha(\Phi(\mu_1) + \Phi(\mu_2) - 2\Phi(\mu_1)\Phi(\mu_2))] = 2\alpha\phi(\mu_2)(1 - 2\Phi(\mu_1)) < 0, \tag{3.8}$$

where the inequality holds for all $\mu_1 > 0$. This implies that the power of the test is maximized under the null at the boundary and the test has level α. $\square$

Although the test in (3.6) is level α and similar on the boundary, its power properties are poor. The supremum of its power function is 2α and its power is α or less in the negative orthant.

To illustrate that nonsimilar tests exist with good overall power properties, Tables 1 and 2 report the power of the the similar-on-the-boundary test in (3.6) and the (rec-

[10] To prove this, the proof of Theorem 2 only needs to be altered by replacing μ_j by μ_j/σ_j, where σ_j^2 is the variance of X_j, for $j = 1, 2$.

ommended) Andrews and Barwick (2012) refined moment selection (RMS) test, respectively, in the $N(\mu, I_2)$ model for a grid of μ values. Both tests have size .05. The power of the test in (3.6) is computed via the formula in (3.7). The power of the RMS test is computed via simulation using $500,000$ simulation repetitions (for both the critical value calculations,

Table 1. Power of the similar-on-the-boundary test in (3.6) for the $N((\mu_1, \mu_2)', I_2)$ model

	μ_1		
μ_2	-4.0	-1.0	0.0
7.0	.100	.084	0.05
5.0	.100	.084	0.05
3.0	.100	.084	0.05
2.0	.098	.083	0.05
1.0	.084	.073	0.05
.75	.077	.069	0.05
.50	.069	.063	0.05
.25	.060	.057	0.05
0.0	.050	.050	0.05
-.50	.031	.037	0.05
-1.0	.016	.027	0.05
-2.0	.002	.017	0.05
-3.0	.000	.016	0.05
-4.0	.000	.016	0.05
-5.0	.000	.016	0.05
-7.0	.000	.016	0.05

which only have to be computed once, and the rejection probability calculations).

All (μ_1, μ_2) combinations in Tables 1 and 2 are in the alternative hypothesis except the first nine rows of the last column. The latter entries in Table 1 show that the test in (3.6) is similar on the boundary. In Tables 1 and 2, the distance from the null hypothesis increases as one moves from right to left and top to bottom. The bottom seven rows of the last column of Table 1 show that the test in (3.6) has power equal to its size for all μ_2 when $\mu_1 = 0$, which is in accord with Theorem 1. Table 1 also shows that the similar-on-the-boundary test has very poor power in general. Its power lies in $[.00, .10]$. As $\mu_1 \to -\infty$ and $\mu_2 \to \infty$, its power approaches $2\alpha = .10$. As $\mu_1 \to -\infty$ and $\mu_2 \to -\infty$, its power approaches .00.

In Table 2, the first nine rows of the last column show that the Andrews-Barwick RMS test is non-similar on the boundary of the null. Its rejection probability lies in $[.029, .050]$ on the boundary. This causes the test to be biased, but the bias disappears quickly as μ_1 decreases. It is essentially gone for $\mu_1 = -.25$.

In Table 2, the power of the RMS test increases to one as $\mu_1 \to -\infty$ and/or $\mu_2 \to -\infty$. Not surprisingly, even for $\mu_1 = -3.0$, the difficulty in determining whether one or two moment inequalities are binding (which is the root cause for the nonsimilarity of the test), causes its power to be less for $\mu_2 \in [-.5, 1.0]$ than for $\mu_2 \in [2, \infty) \cup (-\infty, -1.0]$. The RMS test has good overall power.

Table 2. Power of the Andrews-Barwick RMS test for the $N((\mu_1, \mu_2)', I_2)$ model

	μ_1								
μ_2	-4.0	-3.0	-2.0	-1.5	-1.0	-.50	-.25	-.125	0.0
7.0	.99	.91	.63	.43	.25	.119	.076	.059	.046
5.0	.99	.91	.63	.43	.25	.119	.076	.059	.046
3.0	.99	.90	.62	.42	.24	.114	.073	.057	.044
2.0	.98	.88	.58	.38	.22	.099	.063	.049	.037
1.0	.98	.85	.52	.33	.18	.079	.049	.038	.029
.75	.98	.85	.52	.33	.18	.077	.049	.038	.029
.50	.98	.85	.51	.33	.18	.079	.051	.041	.032
.25	.98	.85	.52	.33	.18	.086	.057	.047	.039
0.0	.98	.85	.53	.34	.20	.099	.070	.059	.050
-.50	.98	.87	.57	.40	.25	.151	.119	.108	.098
-1.0	.99	.90	.65	.49	.35	.250	.217	.205	.195
-2.0	1.00	.96	.83	.74	.65	.573	.547	.537	.528
-3.0	1.00	.99	.96	.93	.90	.869	.858	.854	.851
-4.0	1.00	1.00	1.00	.99	.99	981	.979	.979	.978
-5.0	1.00	1.00	1.00	1.00	1.00	.999	.999	.999	.999
-7.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

For the mixed equality/inequality null hypothesis $H_0 : \mu_1 = 0$ and $\mu_2 \geq 0$, it is trivial to construct a non-trivial similar-on-the-boundary level α test. Let

$$\xi(X) = 1(|X_1| > z_{1-\alpha/2}), \tag{3.9}$$

where $z_{1-\alpha/2}$ is the $1-\alpha$ quantile of the standard normal distribution. This test ignores information in X_2 and, hence, its power properties are not desirable. But, it is level α and is similar on the boundary $B = \{\mu = (\mu_1, \mu_2)' : \mu_1 = 0 \text{ and } \mu_2 \geq 0\}$.

4. Multivariate Normal Mean Model

In this section, we extend the result given in Theorem 1 to dimensions p greater than two and non-spherical variance matrices Σ.

Theorem 3. *Let $X \sim N(\mu, \Sigma)$, where Σ is a positive-definite $p \times p$ matrix for $p \geq 2$. Any (possibly randomized) test of the null hypothesis $H_0 : \mu \geq 0$ that is similar on the boundary $B = \{\mu = (\mu_1, ..., \mu_p)' : \mu \geq 0 \ \& \ \mu_j = 0$ for some $j \leq p\}$ with rejection probability $\alpha \in (0, 1)$ on B has rejection probability α for all μ in $B^* = \{\mu = (\mu_1, ..., \mu_p)' : \mu_j = 0$ for some $j \leq p\}$.*

Comments. 1. Theorem 3 says that a similar-on-the-boundary test (with rejection probability α on the boundary) has trivial power (i.e., power equal α) for all alternatives for which some inequality is satisfied and binding, (i.e., $\mu_j = 0$). Such alternatives include a host of alternatives that are arbitrarily far from the null hypothesis. Thus, Theorem 1 implies that the power properties of tests that are similar on the boundary are very poor.

2. Theorem 3 also holds for the mixed equality/inequality null hypothesis $H_0 : \mu_1 = 0 \ \& \ \mu_2 \geq 0$, where $\mu_1 \in R^p$, $\mu_2 \in R^q$, and $q \geq 1$, and the alternative hypothesis $H_1 : \mu_1 \neq 0$ or $\mu_2 \ngeq 0$. In this case, the boundary of the null is the null itself, i.e., $B = \{\mu = (\mu_1', \mu_2')' : \mu_1 = 0 \text{ and } \mu_2 \geq 0\}$, and $B^* = \{\mu = (\mu_1', \mu_2')' : \mu_1 = 0\}$.[11] Hence, for all alternatives where the equality restriction is not violated, i.e., $\mu = (0', \mu_2')'$ for $\mu_2 \ngeq 0$, similar-on-the-boundary tests have power equal to size α. In consequence, similar-on-the-boundary tests have poor power properties.

[11] The proof is the same as the proof of Theorem 3, but with μ_1 being a p-vector, rather than a scalar, with $p + q$ in place of p, with q in place of $p - 1$, and with the last paragraph of the proof deleted.

3. Comment 3 to Theorem 1 also applies in the context of Theorem 3. Specifically, suppose the null hypothesis H_0 is restricted such that its boundary includes just the set $B_1 = \{\mu : \mu_1 = 0 \ \& \ \mu_2 \in S\}$ for some (nondegenerate) rectangle S in the positive orthant of R^{p-1}, where $\mu = (\mu_1, \mu_2')'$, $\mu_1 \in R$, and $\mu_2 \in R^{p-1}$.Then, the result of Theorem 3 holds with B^* replaced by $B_1^* = \{\mu = (\mu_1, \mu_2')' : \mu_1 = 0\}$. In this situation, there are alternative parameters μ that are arbitrarily far from the null hypothesis for which power equals size α. By symmetry, the same result holds with any element of μ, say μ_j, in place of the first element, μ_1, with μ_2 re-defined accordingly, and with B_1 and B_1^* by B_j and B_j^*, which are defined accordingly.

4. Comment 4 to Theorem 1 also applies in the context of Theorem 3. Hence, median- and quantile-unbiased estimators of $\min\{\mu_1, ..., \mu_p\}$ do not exist when $X \sim N(\mu, \Sigma)$ and Σ is a known, positive-definite matrix.

5. If Σ is unknown and can take on more than one value, say $\Sigma \in \mathcal{S}$, then the null hypothesis is larger than in the known Σ case and the similarity-on-the-boundary condition is stronger. In consequence, in this case, the result of Theorem 1 holds for all $\mu \in B^* \backslash B$ and all positive definite $\Sigma \in \mathcal{S}$.

Proof of Theorem 3. For notational convenience, for any vector $v \in R^p$, we write $v = (v_1, v_2')'$ for $v_1 \in R$ and $v_2 \in R^{p-1}$.

By the Cholesky decomposition, there exists a unique nonsingular lower triangular matrix L with positive diagonal elements such that $\Sigma = LL'$. Let $M = L^{-1}$. Then, M is lower triangular (triangularity is preserved under inverses) and $M\Sigma M' = L^{-1}LL'L^{-1\prime} = I_p$. Let $Y = MX \sim N(\widetilde{\mu}, I_p)$, where $\widetilde{\mu} = M\mu$. By the lower triangular feature of M, $\widetilde{\mu}_1 = M_{11}\mu_1 \geq 0$, where M_{11} (> 0) denotes the $(1,1)$ element of M. Note that $\mu_1 = 0$ iff $\widetilde{\mu}_1 = 0$. Also, $\widetilde{\mu}_2 = M_2\mu$, where M_2 $(\in R^{(p-1)\times p})$ equals M with its first row deleted. M_2 is full row rank.

Suppose a test $\xi(X)$ is similar on the boundary B with rejection probability $\alpha \in (0,1)$ on B. That is,

$$E_\mu \xi(X) = \alpha \ \forall \mu \in B. \tag{4.1}$$

We can write the power of $\xi(X)$ against μ as

$$E_\mu \xi(X) = E_{\widetilde{\mu}} \xi(M^{-1}Y) = \int \int \xi(M^{-1}y)\phi(y_1 - \widetilde{\mu}_1)dy_1 \phi(y_2 - \widetilde{\mu}_2)dy_2, \tag{4.2}$$

where $\phi(y_2) = \Pi_{j=1}^{p-1}\phi(y_{2,j})$ and $y_2 = (y_{2,1}, ..., y_{2,p-1})'$.

For $\mu_1 = 0$ and $\mu_2 \geq 0$, (4.1), (4.2), and $\widetilde{\mu}_1 = M_{11}\mu_1 = 0$ give

$$\begin{aligned} E_{0,\mu_2}\xi(X) &= E_{\widetilde{\mu}_2} g(Y_2) \\ &= \int g(y_2)\phi(y_2 - \widetilde{\mu}_2)dy_2 \\ &= \alpha, \text{ where} \\ g(y_2) &= \int \xi(M^{-1}(y_1, y_2')')\phi(y_1)dy_1. \end{aligned} \tag{4.3}$$

Let

$$\Lambda = \{\lambda \in R^{p-1} : \lambda = M_2 b, \ b \in R^p, \ \& \ b \geq 0\}. \tag{4.4}$$

The second and third equalities in (4.3) hold for all $\widetilde{\mu}_2 \in \Lambda$ because $\widetilde{\mu}_2 = M_2\mu$ and μ is an arbitrary element of R^p with $\mu \geq 0$.

Consider the model $Y_2 \sim N(\lambda, I_{p-1})$ for $\lambda \in \Lambda$. The $N(\lambda, I_{p-1})$ distribution is in the exponential family. The columns of M_2 span R^{p-1} (because M_2 with its first column removed is a full rank triangular matrix since M is). In consequence, Λ contains a $(p-1)$-dimensional rectangle. Hence, in this model, the random vector Y_2 is a complete sufficient statistic for λ, e.g., see Theorem 4.3.1 in Lehmann and Romano (2005, p. 117).

Completeness of Y_2, (4.3), and the definition of completeness give

$$g(y_2) = \alpha \ \forall y_2 \in \mathcal{Y}_2 \text{ for some set } \mathcal{Y}_2 \text{ with } P_\lambda(Y_2 \in \mathcal{Y}_2) = 1 \ \forall \lambda \in \Lambda. \tag{4.5}$$

By the absolute continuity of any (nondegenerate) multivariate normal distribution with respect to any other (nondegenerate) multivariate normal distribution with the same dimension, we have

$$P_\lambda(Y_2 \in \mathcal{Y}_2) = 1 \ \forall \lambda \in R^{p-1}. \tag{4.6}$$

Consider $\mu = (0, \mu_2')' \in R^p$ for arbitrary $\mu_2 \in R^{p-1}$ (so μ is not necessarily in the null hypothesis). By the first two equalities in (4.3), $E_\mu\xi(X) = E_{\widetilde{\mu}_2} g(Y_2)$. This, (4.5), and (4.6) give

$$E_\mu\xi(X) = E_{\widetilde{\mu}_2} g(Y_2) = \alpha. \tag{4.7}$$

Consider any $\mu \in R^p$ with one element equal to zero, i.e., any $\mu \in B^*$. By the same argument as above that gives (4.7), but with $\mu_1 = 0$ replaced by $\mu_j = 0$ for some $j \leq p$, we obtain $E_\mu\xi(X) = \alpha$. This completes the proof. $\square$

5. Tests Based on Moment Inequalities

In this section, we consider tests concerning a parameter θ in a moment inequality model. The parameter θ need not be identified. By inverting the tests, one can construct confidence sets for the true value θ in the usual manner. We use Theorem 3 to show that any test that is asymptotically similar on the boundary of the null hypothesis has poor asymptotic power properties.

The moment inequality model is defined as follows. The true value θ_0 ($\in R^d$) is assumed to satisfy the moment inequalities:

$$E_{F_0} m_j(W_i, \theta_0) \geq 0 \text{ for } j = 1, ..., p, \tag{5.1}$$

where $\{m_j(\cdot, \theta) : j = 1, ..., p\}$ are known real-valued moment functions, and $\{W_i : i \geq 1\}$ are i.i.d. random vectors with true joint distribution F_0. The observed sample is $\{W_i : i \leq n\}$. The true value θ_0 is not necessarily identified. That is, knowledge of $E_{F_0} m_j(W_i, \theta)$ for $j = 1, ..., p$ for all $\theta \in \Theta$ does not necessarily imply knowledge of θ_0. Even knowledge of F_0 does not necessarily imply knowledge of the true value θ_0. To identify the true parameter θ_0, one may need to observe more variables than just $\{W_i : i \leq n\}$.

The null and alternative hypotheses are

$$H_0 : \theta_0 = \theta_{null} \text{ and } H_1 : \theta_0 \neq \theta_{null} \tag{5.2}$$

for a specified (known) value θ_{null}.

The parameter space for θ is a set $\Theta \subset R^d$. The parameter space for F given θ, $\mathcal{F}_\theta$, is the set of all F that satisfy:

$$\begin{aligned}
&\text{(i) } E_F m_j(W_i, \theta) \geq 0 \text{ for } j = 1, ..., p, \\
&\text{(ii) } \sigma^2_{F,j}(\theta) = Var_F(m_j(W_i, \theta)) \in (0, \infty) \text{ for } j = 1, ..., p, \\
&\text{(iii) } Corr_F(m(W_i, \theta)) \in \Psi, \text{ and} \\
&\text{(iv) } E_F |m_j(W_i, \theta)/\sigma_{F,j}(\theta)|^{2+\delta} \leq M \text{ for } j = 1, ..., p,
\end{aligned} \tag{5.3}$$

where Ψ is some set of $p \times p$ correlation matrices and $M < \infty$ and $\delta > 0$ are constants. The parameter space for F is $\mathcal{F} = \cup_{\theta \in \Theta} \mathcal{F}_\theta$. The set of distributions F in the null hypothesis is $\mathcal{F}_{\theta_{null}}$. Thus, the null hypothesis can be re-written as $H_0 : F_0 \in \mathcal{F}_{\theta_{null}}$.

The set of distributions F that are on the boundary of the null hypothesis is

$$\mathcal{F}_{Bdy} = \{F \in \mathcal{F}_{\theta_{null}} : E_F m_j(W_i, \theta_{null}) = 0 \text{ for some } j \leq p\}. \tag{5.4}$$

Given $\mu \in R^p$ and a symmetric positive-definite $p \times p$ matrix Σ, let $\{F_{n,\mu,\Sigma} \in \mathcal{F} : n \geq 1\}$ be a sequence of distributions for which

$$n^{1/2} E_{F_{n,\mu,\Sigma}} m(W_i, \theta_{null}) \to \mu \text{ and } Var_{F_{n,\mu,\Sigma}}(m(W_i, \theta_{null})) \to \Sigma. \tag{5.5}$$

We consider such sequences because typical tests statistics (and data-dependent critical values) have asymptotic distributions that depend on $\lim_{n\to\infty} n^{1/2} E_{F_{n,\mu,\Sigma}} m(W_i, \theta_{null})$, e.g., see Andrews and Soares (2010, p. 130). In brief, the reason is that the test statistics (and critical values) are functions of $n^{-1/2} \sum_{i=1}^n m(W_i, \theta_{null})$ and the asymptotic behavior of the latter is determined by its mean $n^{1/2} E_{F_{n,\mu,\Sigma}} m(W_i, \theta_{null})$ and its variance $Var_{F_{n,\mu,\Sigma}}(m(W_i, \theta_{null}))$ (in the i.i.d. case). The asymptotic distributions of typical test statistics (and critical values) are a function of a $N(\mu, \Sigma)$ distribution under $\{F_{n,\mu,\Sigma} \in \mathcal{F} : n \geq 1\}$.

Define B and B^* as in Section 4.

We impose the following assumptions:

Assumption 1. For some symmetric positive-definite $p \times p$ matrix Σ, all $\mu \in B$, and some $\mu^* \in B^* \backslash B$, there exist sequences $\{F_{n,\mu,\Sigma} \in \mathcal{F}_{Bdy} : n \geq 1\}$ and $\{F_{n,\mu^*,\Sigma} \in \mathcal{F} : n \geq 1\}$ such that (5.5) holds (with μ^* in place of μ in (5.5) for $\{F_{n,\mu^*,\Sigma} : n \geq 1\}$).

Assumption 2. The sequence of tests $\{\phi_n : n \geq 1\}$ is asymptotically similar on the boundary of the null hypothesis with asymptotic rejection probability $\alpha \in (0, 1)$. That is, $\lim_{n\to\infty} \sup_{F \in \mathcal{F}_{Bdy}} P_F(\phi_n \text{ rejects } H_0) = \lim_{n\to\infty} \inf_{F \in \mathcal{F}_{Bdy}} P_F(\phi_n \text{ rejects } H_0) = \alpha$.

Assumption 3. The tests $\{\phi_n : n \geq 1\}$ satisfy: For all sequences $\{F_{n,\mu,\Sigma} : n \geq 1\}$ and $\{F_{n,\mu^*,\Sigma} : n \geq 1\}$ as in Assumption 1,

$$\lim_{n\to\infty} P_{F_{n,\mu,\Sigma}}(\phi_n \text{ rejects } H_0) = P_{\mu,\Sigma}(\phi(Z, \Sigma) \text{ rejects } H_0^\infty)$$

and likewise with μ^* in place of μ, for some test ϕ that depends on (Z, Σ), where $Z \sim N(\mu, \Sigma)$ and where the null hypothesis for ϕ is $H_0^\infty : \mu \geq 0$.

Assumption 1 requires that null distributions F exist such that the vector of moment functions evaluated at θ_{null} and under F can take any p-vector value in some neighborhood of 0 ($\in R^p$) intersected with the non-negative orthant. This holds in many examples, but not all. For example, if one moment inequality is binding implies that some other moment inequality is slack by at least $c > 0$ for all $F \in \mathcal{F}$, then Assumption 1 fails to hold. This occurs in the interval-outcome regression model considered in Manski and Tamer (2002). On the other hand, suppose one moment inequality is binding implies that some other moment inequality cannot be binding, but the amount of slackness can be arbitrarily close to zero. In this case, Assumption 1 fails to hold at $\mu = 0 \in B$, but it can be weakened to cover this case. For convenience, we discuss this extension in Comment 2 to Theorem4 below.

Assumption 2 states that the tests under consideration satisfy the asymptotic similarity-on-the-boundary condition in a uniform sense. A test ϕ_n is similar on the boundary in finite samples if

$$\sup_{F \in \mathcal{F}_{Bdy}} P_F(\phi_n \text{ rejects } H_0) = \inf_{F \in \mathcal{F}_{Bdy}} P_F(\phi_n \text{ rejects } H_0). \tag{5.6}$$

The uniform asymptotic version of this condition just adds $\lim_{n\to\infty}$ on both sides of the equality. Without uniformity, e.g., if the condition is simply $\lim_{n\to\infty} P_F(\phi_n$ rejects $H_0) = \alpha$ for all $F \in \mathcal{F}_{Bdy}$, the condition is quite weak and does not imply that the tests are close to being similar on the boundary for finite n no matter how large n is.

Assumption 3 holds for a wide range of tests. For example, it holds for the class of moment selection tests in Andrews and Soares (2010) (using the asymptotic distribution or bootstrap distribution in the construction of critical values), the refined moment selection tests in Andrews and Barwick (2012), the subsampling tests that have been considered by Chernozhukov, Hong, and Tamer (2007), Romano and Shaikh (2008), and Andrews and Guggenberger (2009).[12] It also holds for generalized empirical likelihood (GEL) based tests with plug-in least favorable critical values, moment selection critical values based on the asymptotic distribution or bootstrap distribution, and subsampling

[12] For example, for the class of tests in Andrews and Soares (2010), Assumption 3 holds by Lemma 2 and its proof in the Supplement to Andrews and Soares (2010). This follows because the test statistic $T_n(\theta_{n,h})$ converges in distribution $S(Z, \Omega_{h_{22}})$, where $Z \sim N(h_1, \Omega_{h_{22}})$, under sequences $\{\gamma_{n,h} : n \geq 1\}$ that are analogous to those in Assumption 1 above by equation (S1.19). In addition, the moment selection critical value $\widehat{c}_n(\theta_{n,h}, 1-\alpha)$ converges in probability to a constant under these sequences. This holds because $\pi_1 = 0$ in condition (ii) of Lemma 2 (in the Supplement to Andrews and Soares (2010)) given that μ in Assumption 3 above, which corresponds to h_1 in Andrews and Soares (2010), has all elements finite. This implies that $\pi_1 = \pi_1^* = 0$, the inequality $\widehat{c}_n(\theta_{n,h}, 1-\alpha) \geq c_n^*$ in Lemma 2(a) holds as an equality, and the critical value $\widehat{c}_n(\theta_{n,h}, 1-\alpha)$ converges in probability to the constant $c_{\pi^*}(1-\alpha)$.

critical values, as in Andrews and Guggenberger (2009), Canay (2010), and Andrews and Soares (2010). It also holds for the tests in Rozen (2008) and Bugni (2010).

Theorem 4. *Under Assumptions* 1-3, *for all sequences* $\{F_{n,\mu^*,\Sigma} \in \mathcal{F} : n \geq 1\}$ *as in Assumption* 1 *with* $\mu^* \in B^* \backslash B$,

$$\lim_{n \to \infty} P_{F_{n,\mu^*,\Sigma}}(\phi_n \text{ rejects } H_0) = \alpha.$$

Comments. 1. Theorem 4 shows that an asymptotically similar-on-the-boundary test (with asymptotic rejection probability α on the boundary) has local power equal to α, which is less than or equal to its asymptotic size, for all alternatives with asymptotic mean vector μ^* in $B^* \backslash B$.[13] In consequence, such tests have very poor asymptotic power properties.

2. For the conclusion of Theorem 4 to hold, Assumption A can be weakened. Any given alternative vector $\mu^* \in B^* \backslash B$ lies in the set $B_j^* \backslash B$ for some $j \leq p$, where $B_j^* = \{\mu = (\mu_1, ..., \mu_p)' : \mu_j = 0\}$. Assumption A does not need to hold for all $\mu \in B$, it just needs to hold for all $\mu \in B_j$ for some set $B_j = \{\mu = (\mu_1, ..., \mu_p)' : \mu_j = 0, \mu_{-j} \in M\}$ for some (nondegenerate) rectangle M in R_+^{p-1}, where $\mu_{-j} = (\mu_1, ...\mu_{j-1}, \mu_{j+1}, ..., \mu_p)'$ and $R_+ = \{x \in R : x \geq 0\}$. The proof of this extension uses Comment 3 to Theorem 3.

3. The results of Theorem 4 can be extended to models with moment inequalities and equalities. Suppose the true value θ_0 $(\in \Theta \subset R^d)$ satisfies the moment conditions:

$$\begin{aligned} &E_{F_0} m_j(W_i, \theta_0) \geq 0 \text{ for } j = 1, ..., p \text{ and} \\ &E_{F_0} m_j(W_i, \theta_0) = 0 \text{ for } j = p+1, ..., p+q, \end{aligned} \tag{5.7}$$

where $\{m_j(\cdot, \theta) : j = 1, ..., p+q\}$ are known real-valued moment functions and $q \geq 1$. In this case, the parameter space $\mathcal{F}_\theta$ for F given θ contains the additional conditions $E_F m_j(W_i, \theta) = 0$ for $j = p+1, ..., p+q$ and p is replaced by $p+q$ in conditions (ii)-(iv) of (5.3). The sets B and B^* are defined to be $B = \{\mu = (\mu_1', \mu_2')' : \mu_1 = 0 \text{ and } \mu_2 \geq 0\}$, and $B^* = \{\mu = (\mu_1', \mu_2')' : \mu_1 = 0\}$, as in Comment 2 to Theorem 3. In this model, the boundary of the null is the null itself. That is, $\mathcal{F}_{Bdy} = \mathcal{F}_{\theta_{null}}$. Assumptions 1-3 apply in this model with p replaced by $p+q$. With the above changes, the result of Theorem 4 holds in the moment inequalities and equalities model.[14] In consequence, tests that

[13] By definition, the asymptotic size of a test is the limit of its finite-sample size, which is its maximum rejection probability under the null.

[14] To obtain this result, in the proof of Theorem 4, one just replaces the use of Theorem 3 with the

are asymptotically similar on the boundary have poor asymptotic power properties in models with moment inequalities and equalities.

4. A version of the result of Theorem 4 applies to tests of stochastic dominance provided the null and alternative hypotheses include distributions with finite support.

Proof of Theorem 4. By Assumption 1, for all $\mu \in B$, there exists a sequence $\{F_{n,\mu,\Sigma} \in \mathcal{F}_{Bdy} : n \geq 1\}$ such that (5.5) holds. This and Assumption 2 imply that

$$\lim_{n\to\infty} P_{F_{n,\mu,\Sigma}}(\phi_n \text{ rejects } H_0) = \alpha \ \forall \mu \in B. \tag{5.8}$$

By Assumption 3,

$$\lim_{n\to\infty} P_{F_{n,\mu,\Sigma}}(\phi_n \text{ rejects } H_0) = P_{\mu,\Sigma}(\phi(Z,\Sigma) \text{ rejects } H_0^\infty). \tag{5.9}$$

Combining (5.8) and (5.9) gives

$$P_{\mu,\Sigma}(\phi(Z,\Sigma) \text{ rejects } H_0^\infty) = \alpha \ \forall \mu \in B. \tag{5.10}$$

Thus, ϕ is a test based on $Z \sim N(\mu,\Sigma)$ (and possibly Σ) that is similar on the boundary of the null hypothesis $H_0^\infty : \mu \geq 0$. Hence, by Theorem 3, $P_{\mu^*,\Sigma}(\phi(Z,\Sigma) \text{ rejects } H_0^\infty) = \alpha$ for all $\mu^* \in B^*\backslash B$. Combining this with Assumption 3 and taking $\mu^* \in B^*\backslash B$ as in Assumption 1 gives

$$\lim_{n\to\infty} P_{F_{n,\mu^*,\Sigma}}(\phi_n \text{ rejects } H_0) = P_{\mu^*,\Sigma}(\phi(Z,\Sigma) \text{ rejects } H_0^\infty) = \alpha, \tag{5.11}$$

which is the result of Theorem 4. □

References

Andrews, D. W. K. & P. Guggenberger (2009) Validity of subsampling and 'plug-in asymptotic' inference for parameters defined by moment inequalities. *Econometric Theory* **25**, 669-709.

Andrews, D. W. K. & P. J. Barwick (2012) Inference for parameters defined by moment inequalities: a recommended moment selection procedure. *Econometrica* **80**, forth-

use of the result in Comment 2 to Theorem 3.

coming. Also available as Andrews and Jia (2011), Cowles Foundation Discussion Paper No. 1676R, Yale University.

Andrews, D. W. K., M. J. Moreira, & J. H. Stock (2006) Optimal two-sided invariant similar tests for instrumental variables regression. *Econometrica* **74**, 715-752.

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

Andrews, D. W. K. & X. Shi (2007) Inference based on conditional moment inequalities. Cowles Foundation Discussion Paper No. 1761R, Yale University.

Andrews, D. W. K. & G. Soares (2010) Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica* **78**, 2010, 119-157.

Blumenthal, S. & A. Cohen (1968) Estimation of the larger of two normal means. *Journal of the American Statistical Association* **63**, 861-876.

Bugni, F. A. (2010) Bootstrap inference in partially identified models defined by moment inequalities: coverage of the identified set. *Econometrica* **78**, 735-753.

Canay, I. A. (2010) EL inference for partially identified models: large deviations optimality and bootstrap validity. *Journal of Econometrics* **156**, 408-425.

Chen, L.-Y. & J. Szroeter (2009) Hypothesis testing of multiple inequalities: the method of constraint chaining. CEMMAP Working Paper No. 13/09, Institute for Fiscal Studies, Department of Economics, University College London.

Chernozhukov, V., H. Hong, & E. Tamer (2007) Estimation and confidence regions for parameter sets in econometric models. *Econometrica* **75**, 1243-1284.

Chernozhukov, V., S. Lee, & A. Rozen (2008) Inference with intersection bounds. Unpublished manuscript, Department of Economics, University College London.

Fraser, D. A. S. (1952) Confidence bounds for a set of means. *Annals of Mathematical Statistics* **23**, 525-575.

Hansen, P. R. (2005) A test for superior predictive ability. *Journal of Business & Economic Statistics* **23,** 365-380.

Hirano, K. & J. Porter (2012) Impossibility results for nondifferentiable functionals. *Econometrica* **80**, forthcoming.

Imbens, G. & C. Manski (2004) Confidence intervals for partially identified parameters. *Econometrica* **72**, 1845-1857.

Lehmann E. L. & J. P. Romano (2005) *Testing Statistical Hypotheses,* 3rd edition. New York: Springer.

Manski, C. F. & E. Tamer (2002) Inference on regression with interval data on a regressor or outcome. *Econometrica* **70**, 519-546.

Moreira, H. & M. J. Moreira (2011) Contributions to the Theory of Similar Tests. Unpublished manuscript, Department of Economics, Columbia University.

Moreira, M. J. (2003) A conditional likelihood ratio test for structural models. *Econometrica* **71**, 1027-1048.

Moreira, M. J. (2009) Tests with correct size when instruments can be arbitrarily weak. *Journal of Econometrics* **152**, 131-140.

Romano, J. P. & A. M. Shaikh (2008) Inference for identifiable parameters in partially identified econometric models. *Journal of Statistical Inference and Planning (Special Issue in Honor of T. W. Anderson)* **138**, 2786-2807.

Rosen, A. M. (2008) Confidence sets for partially identified parameters that satisfy a finite number of moment inequalities. *Journal of Econometrics* **146**, 107-117.

Stoye, J. (2009) More on confidence intervals for partially identified parameters. *Econometrica* **77**, 1299-1315.