# A Practical Method for Testing Many Moment Inequalities*

Yuehao Bai
Department of Economics
University of Chicago
ybai@uchicago.edu

Andres Santos
Department of Economics
University of California–Los Angeles
andres@econ.ucla.edu

Azeem M. Shaikh
Department of Economics
University of Chicago
amshaikh@uchicago.edu

September 13, 2019

**Abstract**

This paper considers the problem of testing a finite number of moment inequalities. For this problem, Romano et al. (2014) propose a two-step testing procedure. In the first step, the procedure incorporates information about the location of moments using a confidence region. In the second step, the procedure accounts for the use of the confidence region in the first step by adjusting the significance level of the test appropriately. An important feature of the proposed method is that it is "practical" in the sense that it remains computationally feasible even if the number of moments is large. Its justification, however, has so far been limited to settings in which the number of moments is fixed with the sample size. In this paper, we provide weak assumptions under which the same procedure remains valid even in settings in which there are "many" moments in the sense that the number of moments grows rapidly with the sample size. We confirm the practical relevance of our theoretical guarantees in a simulation study. We additionally provide both numerical and theoretical evidence that the procedure compares favorably with the method proposed by Chernozhukov et al. (2019), which has also been shown to be valid in such settings.

KEYWORDS: High-dimensional inference, partial identification, bootstrap, moment inequalities, multi-sided hypothesis

JEL classification codes: C12, C14

*We thank Roy Allen and Denis Chetverikov for helpful comments. The research of the third author is supported by NSF Grant SES-1530661.

# 1 Introduction

Let $X_i, i = 1, \ldots, n$ be an independent and identically distributed (i.i.d.) sequence of random variables with distribution $P \in \mathbf{P}_n$ on $\mathbf{R}^{p_n}$ and consider the problem of testing

$$H_0 : P \in \mathbf{P}_{0,n} \text{ versus } H_1 : P \in \mathbf{P}_{1,n} \ , \tag{1}$$

where

$$\mathbf{P}_{0,n} \equiv \{P \in \mathbf{P}_n : E_P[X_i] \leq 0\} \tag{2}$$

and $\mathbf{P}_{1,n} = \mathbf{P}_n \setminus \mathbf{P}_{0,n}$. Here, the inequality in (2) is intended to be interpreted component-wise and $\mathbf{P}_n$ is a "large" class of possible distributions for the observed data. By indexing both the number of moments, $p_n$, and the class of possible distributions, $\mathbf{P}_n$, by the sample size $n$, we anticipate asymptotic results that allow the the number of moments $p_n$ to grow rapidly with the sample size $n$. In this way, our asymptotic framework can accommodate settings in which it is desired to test possibly "many" moment inequalities. Our goal is to construct tests $\phi_n = \phi_n(X_1, \ldots, X_n)$ of (1) that are uniformly consistent in level, i.e.,

$$\limsup_{n\to\infty} \sup_{P \in \mathbf{P}_{0,n}} E_P[\phi_n] \leq \alpha \tag{3}$$

for some pre-specified value of $\alpha \in (0, 1)$.

In many instances where the testing problem described above arises in economics, the number of moments is large. Examples include entry models, as in Ciliberto and Tamer (2009), in which $p_n$ is on the order of $2^{m+1}$, where $m$ is the number of firms, and dynamic models of imperfect competition, as in Bajari et al. (2007), where $p_n$ may even be as large as 500. Yet, with the notable exception of Chernozhukov et al. (2019), tests of (1) that have been proposed have only been shown to satisfy (3) under restrictions on $\mathbf{P}_n$ that require number of moments $p_n$ to be small in the sense that it is independent of the sample size $n$. Canay and Shaikh (2017) provide a detailed review of these tests. In this paper, we focus on one particular such test of (1): the two-step testing procedure proposed by Romano et al. (2014). This test was shown to satisfy (3) under assumptions on $\mathbf{P}_n$ that restrict $p_n$ to not depend on $n$. Romano et al. (2014) emphasize, however, that the test is "practical" in the sense that it remains computationally feasible even if the number of moments is large, thereby permitting its implementation in examples such as those described above. In this paper, we show that the test of Romano et al. (2014) in fact continues to satisfy (3) for a large class of distributions that permits the number of moments $p_n$ to grow exponentially with the sample size $n$. In this way, our results establish the validity of the methodology for testing "many" moment inequalities, thereby supporting its application in examples such as those described above.

Our theoretical analysis relies crucially on the seminal work of Chernozhukov et al. (2013, 2017) on the high-dimensional central limit theorem. The high-dimensional central limit theorem had previously been applied to study tests of (1) by Chernozhukov et al. (2019), who, as mentioned previously, develop tests that satisfy (3) for a large class of distributions $\mathbf{P}_n$ that permits the number of moments $p_n$ to grow rapidly with the sample size $n$. Prior to the results in this paper, however, it was unclear whether it was sensible to compare the power of tests developed by Chernozhukov et al. (2019) with the one proposed by Romano et al. (2014) because it was not known whether the latter test continued to satisfy (3) when the number of moments $p_n$ was permitted to grow rapidly with the sample size $n$. In light of the results in this paper, such a comparison is now theoretically justified. For brevity, we restrict our comparison to the test of (1) most preferred by Chernozhukov et al. (2019) among those in their paper. We recall a result by Allen (2018), who argues that the test proposed by Romano et al. (2014) is more powerful in finite samples than the test proposed by Chernozhukov et al. (2019). See Remark 2.1 for further discussion of this point, where we clarify that this comparison is only valid for certain implementations of the two tests. We therefore supplement this theoretical comparison with simulation evidence for other implementations of the two tests. In our simulations, we find that the test proposed by Romano et al. (2014) continues to compare favorably, both in terms of size and power, with the test proposed by Chernozhukov et al. (2019).

The remainder of the paper is organized as follows. In Section 2, we provide a detailed description of the testing procedure in Romano et al. (2014) and the assumptions that will underlie our analysis. In our discussion of the assumptions, we emphasize that they permit the number of moments $p_n$ to grow rapidly with the sample size $n$. We then establish that the test satisfies (3) under these assumptions. The proof of this result is relegated to the Appendix. We also briefly discuss the sense in which the test proposed by Romano et al. (2014) is more powerful in finite samples than the test proposed by Chernozhukov et al. (2019) in Remark 2.1. In Section 3, we examine the practical relevance of our theoretical results via a simulation study, which includes further comparisons with the test proposed by Chernozhukov et al. (2019).

## 2 Main Result

We begin this section by describing the testing procedure in Romano et al. (2014). In order to do so, it is useful to introduce some further notation. For $1 \leq j \leq p_n$, let $X_{i,j}$ denote the $j$th component of $X_i$ and set

$$\bar{X}_{j,n} \equiv \frac{1}{n} \sum_{1 \leq i \leq n} X_{i,j} \tag{4}$$

$$S^2_{j,n} \equiv \frac{1}{n} \sum_{1 \leq i \leq n} (X_{i,j} - \bar{X}_{j,n})^2 \ . \tag{5}$$

We will also make use of the notation $\mu_j(P) \equiv E_P[X_{i,j}]$ and $\sigma^2_j(P) \equiv \mathrm{Var}_P[X_{i,j}]$, so (4) may be equivalently expressed as $\mu_j(\hat{P}_n)$ and (5) as $\sigma^2_j(\hat{P}_n)$, where $\hat{P}_n$ is the empirical distribution of $\{X_i\}_{i=1}^n$. While Romano et al. (2014) consider a variety of test statistics, we focus on the test that rejects for large values of

$$T_n \equiv \max\left\{ \max_{1 \leq j \leq p_n} \frac{\sqrt{n}\bar{X}_{j,n}}{S_{j,n}}, 0 \right\} \ .$$

In order to define the critical value with which we will compare $T_n$, it will be useful to introduce an i.i.d. sequence of random variables with distribution $\hat{P}_n$ conditional on $\{X_i\}_{i=1}^n$, which we will denote by $X_i^*, i = 1, \ldots, n$. We further define $\bar{X}^*_{j,n}$ and $(S^*_{j,n})^2$ by analogy with $\bar{X}_{j,n}$ in (4) and $S^2_{j,n}$ in (5) but substituting $X_i^*$

for $X_i$. Using this notation, the critical value with which we will compare $T_n$ is given by

$$\hat{c}_n^{(2)}(1-\alpha+\beta) \equiv \inf\left\{c \in \mathbf{R} : P\left\{\max\left\{\max_{1\le j\le p_n} \frac{\sqrt{n}(\bar{X}^*_{j,n} - \bar{X}_{j,n} + \hat{u}_{j,n})}{S^*_{j,n}}, 0\right\} \le c \Big| \{X_i\}_{i=1}^n\right\} \ge 1-\alpha+\beta\right\}, \quad (6)$$

where $\alpha \in (0, \frac{1}{2})$ is the nominal level of the test, $0 < \beta < \alpha$, and

$$\hat{u}_{j,n} \equiv \min\left\{\bar{X}_{j,n} + \frac{S_{j,n}}{\sqrt{n}}\hat{c}_n^{(1)}(1-\beta), 0\right\} \quad (7)$$

with

$$\hat{c}_n^{(1)}(1-\beta) \equiv \inf\left\{c \in \mathbf{R} : P\left\{\max_{1\le j\le p_n} \frac{\sqrt{n}(\bar{X}_{j,n} - \bar{X}^*_{j,n})}{S^*_{j,n}} \le c \Big| \{X_i\}_{i=1}^n\right\} \ge 1-\beta\right\}.$$

The test $\phi_n^{\mathrm{RSW}}$ of the null hypothesis in (1) we consider rejects whenever $T_n$ exceeds $\hat{c}_n^{(2)}(1-\alpha+\beta)$, i.e.,

$$\phi_n^{\mathrm{RSW}} \equiv I\left\{T_n > \hat{c}_n^{(2)}(1-\alpha+\beta)\right\}. \quad (8)$$

In order to motivate this choice of critical value, it is useful to note the test statistic $T_n$ satisfies

$$T_n = \max\left\{\max_{1\le j\le p_n}\left(\frac{\sqrt{n}(\bar{X}_{j,n} - \mu_j(P))}{S_{j,n}} + \frac{\sqrt{n}\mu_j(P)}{S_{j,n}}\right), 0\right\}. \quad (9)$$

The decomposition of $T_n$ in (9) highlights that the main impediment to approximating the distribution of $T_n$ is the presence of the nuisance parameters $\sqrt{n}\mu_j(P)$ for $1 \le j \le p_n$. Even though these nuisance parameters cannot be consistently estimated, Romano et al. (2014) observe that it may still be possible to construct a suitably valid confidence region for them. Lemma 4.1 in the Appendix employs their insight and the high-dimensional central limit theorem of Chernozhukov et al. (2017) to show, under conditions that permit $p_n$ to grow rapidly with the sample size $n$, that $\sqrt{n}\mu_j(P) \le \sqrt{n}\hat{u}_{j,n}$ for all $1 \le j \le p_n$ with probability approximately no less than $1-\beta$ whenever the null hypothesis in (1) is true. Since $T_n$ is monotonically increasing in the nuisance parameters $\sqrt{n}\mu_j(P)$ for all $1 \le j \le p_n$ it follows that, viewed as a function of these nuisance parameters, any quantile of $T_n$ is maximized over said confidence region by setting $\sqrt{n}\mu_j(P) = \sqrt{n}\hat{u}_{j,n}$ for all $1 \le j \le p_n$. Thus, the critical value $\hat{c}_n^{(2)}(1-\alpha+\beta)$ is a bootstrap estimate of the $1-\alpha+\beta$ quantile of $T_n$ under the "least favorable" nuisance parameter value $\sqrt{n}\mu_j(P) = \sqrt{n}\hat{u}_{j,n}$ for all $1 \le j \le p_n$. Here, the $1-\alpha+\beta$ quantiles is employed instead of $1-\alpha$, to account for the possibility that, with probability approximately no greater than $\beta$, we may find $\sqrt{n}\mu_j(P) > \sqrt{n}\hat{u}_{j,n}$ for some $1 \le j \le p_n$.

Our analysis of the test defined in (8) requires the following assumption:

**Assumption 2.1.** (i) $\{X_i\}_{i=1}^n$ is an i.i.d. sample with $X_i \in \mathbf{R}^{p_n}$ and $X_i \sim P \in \mathbf{P}_n$; (ii) $\sigma_j(P) > 0$ for all $1 \le j \le p_n$ and $P \in \mathbf{P}_n$; (iii) For $k = 1, 2$, there is a $M_{k,n} < \infty$ such that $E_P[|X_{i,j} - \mu_j(P)|^{2+k}] \le \sigma_j^{2+k}(P)M_{k,n}^k$ for all $1 \le j \le p_n$ and $P \in \mathbf{P}_n$; (iv) There exists a $B_n < \infty$ such that $E_P[\max_{1\le j\le p_n}|X_{i,j} - \mu_j(P)|^4/\sigma_j^4(P)] \le B_n^4$ for all $P \in \mathbf{P}_n$; (v) $(M_{1,n}^2 \vee M_{2,n}^2 \vee B_n^2)\log^{3.5}(p_n n) = o(n^{(1-\delta)/2})$ for some $\delta \in (0,1)$.

Assumption 2.1(i) simply formalizes the requirement that $\{X_i\}_{i=1}^n$ be an i.i.d. sample, while Assumption 2.1(ii) requires the variance of $X_{i,j}$ to be positive for all $P \in \mathbf{P}_n$ and $1 \leq j \leq p_n$. In Assumption 2.1(iii), we impose a uniform in $P \in \mathbf{P}_n$ and $1 \leq j \leq p_n$ bound on the (standardized) moments of $X_{i,j}$. This condition is a strengthening of the (standardized) uniform integrability condition imposed by Romano et al. (2014), which we require in order to study a setting in which $p_n$ diverges to infinity. Assumption 2.1(iv) bounds the fourth moments of the maximum of the (standardized) $X_{i,j}$. If, for example, the support of the standardized $X_{i,j}$ under $P$ is bounded uniformly in $P \in \mathbf{P}_n$, $1 \leq j \leq p_n$, and $n$, then $B_n$ can be taken to be a constant independent of $n$. In contrast, if the standardized $X_{i,j}$ have exponential tails uniformly in $P \in \mathbf{P}_n$, $1 \leq j \leq p_n$, and $n$, then $B_n$ can be set proportional to a power of $\log(p_n)$. Finally, Assumption 2.1(v) states the main condition governing the relationship between the dimension $p_n$ and the sample size $n$. Importantly, we note that under suitable moment restrictions on $X_{i,j}$, $p_n$ may grow exponentially with $n$.

Under Assumption 2.1, we are able to establish the main result of this paper.

**Theorem 2.1.** *If Assumption 2.1 holds, $\alpha \in (0, \frac{1}{2})$, and $0 < \beta < \alpha$, then $\phi_n^{\mathrm{RSW}}$ defined in (8) satisfies (3).*

Theorem 2.1 verifies that the test proposed in Romano et al. (2014) is indeed able to satisfy (3) even in settings in which $p_n$ grows rapidly with the sample size. In this manner, Theorem 2.1 provides theoretical support for applying the test $\phi_n^{\mathrm{RSW}}$ is empirical applications with "many" moment inequalities. The ability of the test in Romano et al. (2014) to control size in such high-dimensional settings had previously been conjectured, but not established, by Chernozhukov et al. (2019). We also note that while Theorem 2.1 applies for any fixed value of $\beta \in (0, \alpha)$, we note that Romano et al. (2014) recommend setting $\beta = \alpha/10$ in practice based on evidence from a simulation study.

**Remark 2.1.** Chernozhukov et al. (2019) propose several different tests of (1). In our comparisons, we restrict attention to their most preferred test, which is similar in spirit to the "generalized moment selection" tests developed in Andrews and Soares (2010). The proposed test rejects for large values of

$$\tilde{T}_n \equiv \max_{1 \leq j \leq p_n} \frac{\sqrt{n}\bar{X}_{j,n}}{S_{j,n}} \ .$$

In order to describe the critical value with which they compare $\tilde{T}_n$, for $I \subseteq \{1, \ldots, p_n\}$ and $\gamma \in (\frac{1}{2}, 1)$, define

$$\tilde{c}_n(I, \gamma) \equiv \inf \left\{ c \in \mathbf{R} : P\left\{ \max_{j \in I} \frac{\sqrt{n}(\bar{X}^*_{j,n} - \bar{X}_{j,n})}{S^*_{j,n}} \leq c \middle| \{X_i\}_{i=1}^n \right\} \geq \gamma \right\} \ . \tag{10}$$

Using this notation, the proposed test $\phi_n^{\mathrm{CCK}}$ rejects whenever $\tilde{T}_n$ exceeds $\tilde{c}_n(\hat{I}_n, 1 - \alpha + 2\beta)$, where

$$\hat{I}_n \equiv \left\{ 1 \leq j \leq p_n : \frac{\sqrt{n}\bar{X}_{j,n}}{S_{j,n}} > -2\tilde{c}_n(\{1, \ldots, p_n\}, 1 - \beta) \right\} \ ,$$

$\alpha \in (0, \frac{1}{2})$ and $0 < \beta < \frac{\alpha}{2}$, i.e.,

$$\phi_n^{\mathrm{CCK}} \equiv I\{\tilde{T}_n > \tilde{c}_n(\hat{I}_n, 1 - \alpha + 2\beta)\} \ . \tag{11}$$

We also consider the test $\phi_n^{\mathrm{CCK2}}$ defined as above, but in which $S^*_{j,n}$ in (10) is replaced with $S_{j,n}$. It is worth emphasizing that the formal analysis in Chernozhukov et al. (2019) concerns $\phi_n^{\mathrm{CCK2}}$, but we include both tests in our comparisons for completeness. Allen (2018) argues that the version of $\phi_n^{\mathrm{RSW}}$ that also replaces $S^*_{j,n}$ in (6) and (7) with $S_{j,n}$, which we denote by $\phi_n^{\mathrm{RSW2}}$, is more powerful than $\phi_n^{\mathrm{CCK2}}$ in the sense that $\phi_n^{\mathrm{RSW2}} \geq \phi_n^{\mathrm{CCK2}}$. An inspection of the proof of this claim reveals that it is true in finite samples only when a Gaussian multiplier bootstrap is used instead of the empirical bootstrap. By similar arguments to those in Allen (2018) it is also possible to show that $\phi_n^{\mathrm{RSW}} \geq \phi_n^{\mathrm{CCK}}$ provided that a Gaussian multiplier bootstrap is again used instead of the empirical bootstrap. ■

**Remark 2.2.** In some cases, it may be of interest to determine not just whether $\mu_j(P) \leq 0$ for all $1 \leq j \leq p_n$ or not, but the specific values of $1 \leq j \leq p_n$ for which $\mu_j(P) > 0$. For this purpose, it is natural to consider the problem of simultaneously testing $H_j : P \in \mathbf{P}_{j,n}$ versus $H'_j : P \in \mathbf{P}'_{j,n}$ for $j = 1, \ldots, p_n$, where $\mathbf{P}_{j,n} \equiv \{P \in \mathbf{P}_n : \mu_j(P) \leq 0\}$ and $\mathbf{P}'_{j,n} \equiv \mathbf{P}_n \setminus \mathbf{P}_j$. In order to account for the multiplicity of decisions being made, it is common to require control of the familywise error rate in the sense that

$$\limsup_{n\to\infty} \sup_{P \in \mathbf{P}_n} FWER_P \leq \alpha \; , \tag{12}$$

where

$$FWER_P = P\{\text{reject any } H_j \text{ with } P \in \mathbf{P}_{j,n}\} \; .$$

Using Theorem 2.1, it is possible to develop procedures that satisfy (12) under Assumption 2.1. For instance, it is straightforward to show that the procedure that rejects any $H_j$ with $\sqrt{n}\bar{X}_{j,n}/S_{j,n} > \hat{c}_n^{(2)}(1-\alpha+\beta)$ satisfies (12) under Assumption 2.1. By combining Theorem 2.1 with results in Romano and Wolf (2005), iterative improvements upon such a procedure are also possible. Indeed, one may simply apply this procedure and then repeat it with the set of null hypotheses that are *not* rejected after the first application, continuing in this fashion until no further null hypotheses are rejected. For some results in settings in which $p_n$ remains fixed with the sample size $n$, see Romano and Wolf (2018). ■

# 3 Simulations

In this section, we examine the finite-sample behavior of the test of (1) described in Section 2 via a small simulation study. We also compare its behavior with tests described in Remark 2.1.

We begin by describing the distribution of $X_i$. Following Chernozhukov et al. (2019), we specify that

$$X_{i,j} = \theta(I\{1 < j \leq 0.05p_n\} + \varepsilon_{i,j}) - bI\{0.1p_n < j \leq p_n\} + \varepsilon_{i,j}$$

for $1 \leq i \leq n$ and $1 \leq j \leq p_n$, where $\varepsilon_i, i = 1, \ldots, n$ are i.i.d. with distribution $N(0, \Sigma)$. We consider four different models, which differ according to the values of $b$ and $\Sigma$.

**Model 1**: $b = 0$, $\Sigma_{j,k} = 1$ for $1 \leq j,k \leq p_n$ with $j = k$ and $\rho$ otherwise.

**Model 2**: $b = 0.8$, $\Sigma_{j,k} = 1$ for $1 \leq j,k \leq p_n$ with $j = k$ and $\rho$ otherwise.

**Model 3**: $b = 0$, $\Sigma_{j,k} = \rho^{|j-k|}$ for $1 \leq j,k \leq p_n$.

**Model 4**: $b = 0.8$, $\Sigma_{j,k} = \rho^{|j-k|}$ for $1 \leq j, k \leq p_n$.

In Chernozhukov et al. (2019), Models 1 and 2 are referred to as "equicorrelated" and Models 3 and 4 are referred to as "autocorrelated." For each model, we further consider the following different values of $\rho$, $p_n$ and $\theta$: $\rho \in \{0, 0.5, 0.9\}$, $p_n \in \{40, 100, 200\}$, and $\theta \in \{0, 0.2\}$. In all designs, the sample size $n$ is set to equal one hundred, and all tests are implemented at a $\alpha = 0.05$ nominal level. Finally, we observe that the null hypothesis is true when $\theta = 0$ and the alternative hypothesis is true when $\theta = 0.2$. In this way, our designs permit us to study both the size and power of the tests under consideration.

In our simulations below, we consider three different tests:

**RSW**: The test $\phi_n^{\text{RSW}}$ defined in (8).

**RSW2**: The test $\phi_n^{\text{RSW2}}$ described in Remark 2.1.

**CCK**: The test $\phi_n^{\text{CCK}}$ defined in (11).

**CCK2**: The test $\phi_n^{\text{CCK2}}$ described in Remark 2.1.

Recall that the only distinction between $\phi_n^{\text{RSW}}$ and $\phi_n^{\text{RSW2}}$ is that the former employs $S^*_{j,n}$ in the bootstrap samples, while the latter employs $S_{j,n}$. The same distinction differentiates $\phi_n^{\text{CCK}}$ and $\phi_n^{\text{CCK2}}$. Following the recommendation in Romano et al. (2014), we choose $\beta = 0.005$ when implementing $\phi_n^{\text{RSW}}$. Following the recommendation in Chernozhukov et al. (2019), we choose $\beta = 0.001$ when implementing $\phi_n^{\text{CCK}}$ and $\phi_n^{\text{CCK2}}$.

The results of our simulations are presented in Tables 1–4, which correspond to Models 1–4, respectively. Columns labeled 'RSW', 'RSW2', 'CCK' and 'CCK2' display rejection probabilities (in percentage points) for the corresponding test. Columns labeled '≥CCK' and '≥CCK2' display, respectively, the percentage of replications where $\phi_n^{\text{RSW}} \geq \phi_n^{\text{CCK}}$ and $\phi_n^{\text{RSW2}} \geq \phi_n^{\text{CCK2}}$. Rows correspond to different values of $p_n \in \{40, 100, 200\}$ and $\rho \in \{0, 0.05, 0.9\}$. In all designs, we use 10,000 replications and 1,000 bootstrap samples. We emphasize that we employ the same bootstrap samples for all tests.

We summarize our findings from the simulations as follows:

- Both $\phi_n^{\text{RSW}}$ and $\phi_n^{\text{CCK}}$ exhibit good size control even in settings where $p_n$ exceeds the sample size $n = 100$, but $\phi_n^{\text{CCK}}$ tends to under-reject the null hypothesis more severely than $\phi_n^{\text{RSW}}$. See, e.g., Model 2, $p = 200$, $\rho = 0$, and $\theta = 0$, in which case $\phi_n^{\text{CCK}}$ has rejection probability 0.66%, whereas $\phi_n^{\text{RSW}}$ has rejection probability 4.63%. In contrast, the tests $\phi_n^{\text{RSW2}}$ and $\phi_n^{\text{CCK2}}$ have considerably worse size control, over-rejecting the null hypothesis in some cases quite severely. See, e.g., Model 3, $p_n = 200$, $\rho = 0$, and $\theta = 0$, in which case $\phi_n^{\text{RSW2}}$ has rejection probability 7.57% and $\phi_n^{\text{CCK2}}$ has rejection probability 7.86%.

- The tests $\phi_n^{\text{RSW2}}$ and $\phi_n^{\text{CCK2}}$ are generally the more powerful than $\phi_n^{\text{RSW}}$ and $\phi_n^{\text{CCK}}$, but this feature must be weighed against their considerably worse size control. The test $\phi_n^{\text{RSW}}$ is generally at least as powerful as $\phi_n^{\text{CCK}}$, and, at times, quite a bit more powerful. These instances tend to coincide with the values of $p_n$ and $\rho$ for which $\phi_n^{\text{CCK}}$ under-rejects the null hypothesis. See, e.g., Model 2, $p_n = 200$,

$\rho = 0$, and $\theta = 0.2$, in which case $\phi_n^{\text{CCK}}$ has rejection probability only 26.44%, whereas $\phi_n^{\text{RSW}}$ has rejection probability 66.70%. The power ranking between $\phi_n^{\text{CCK2}}$ and $\phi_n^{\text{RSW2}}$ is more ambiguous with each test outperforming the other in some design. The maximal power difference in favor of $\phi_n^{\text{CCK2}}$ occurs in Model 2 with $p_n = 40$ and $\rho = 0.9$, in which the difference in the rejection probabilities equals 5.43%. In contrast, the maximal power difference in favor of $\phi_n^{\text{RSW2}}$ occurs in Model 4 with $p_n = 200$ and $\rho = 0$, in which the difference in the rejection probabilities equals 11.18%.

- In nearly every replication, $\phi_n^{\text{RSW}}$ rejects the null hypothesis whenever $\phi_n^{\text{CCK}}$ does and $\phi_n^{\text{RSW2}}$ rejects the null hypothesis whenever $\phi_n^{\text{CCK2}}$ does. These results suggest that even though the analysis in Allen (2018) require the use of a Gaussian multiplier bootstrap, they may also hold approximately when employing the empirical bootstrap.

| | | $\theta = 0$ | | | | | | $\theta = 0.2$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **RSW** | **RSW2** | **CCK** | **CCK2** | **≥CCK** | **≥CCK2** | **RSW** | **RSW2** | **CCK** | **CCK2** | **≥CCK** | **≥CCK2** |
| $p_n = 40$ | $\rho = 0$ | 4.86 | 6.38 | 5.06 | 6.79 | 99.80 | 99.59 | 18.14 | 22.87 | 19.22 | 23.73 | 98.92 | 99.14 |
| | $\rho = 0.5$ | 4.53 | 5.87 | 4.78 | 6.31 | 99.75 | 99.56 | 12.51 | 15.11 | 13.02 | 15.90 | 99.49 | 99.21 |
| | $\rho = 0.9$ | 4.47 | 5.63 | 4.73 | 6.24 | 99.74 | 99.39 | 10.46 | 12.50 | 10.92 | 13.35 | 99.54 | 99.15 |
| $p_n = 100$ | $\rho = 0$ | 4.52 | 7.01 | 4.84 | 7.36 | 99.68 | 99.65 | 23.14 | 30.22 | 24.18 | 31.24 | 98.96 | 98.98 |
| | $\rho = 0.5$ | 4.28 | 5.99 | 4.48 | 6.36 | 99.80 | 99.63 | 14.33 | 18.18 | 14.97 | 18.97 | 99.36 | 99.21 |
| | $\rho = 0.9$ | 4.64 | 6.29 | 4.87 | 6.67 | 99.77 | 99.62 | 11.34 | 14.47 | 11.98 | 15.24 | 99.36 | 99.23 |
| $p_n = 200$ | $\rho = 0$ | 4.24 | 7.12 | 4.52 | 7.51 | 99.72 | 99.61 | 29.02 | 38.85 | 30.34 | 40.17 | 98.68 | 98.68 |
| | $\rho = 0.5$ | 4.55 | 6.68 | 4.79 | 7.13 | 99.76 | 99.55 | 15.38 | 20.48 | 16.14 | 21.07 | 99.24 | 99.41 |
| | $\rho = 0.9$ | 4.41 | 6.28 | 4.66 | 6.77 | 99.75 | 99.51 | 13.33 | 17.13 | 14.05 | 17.85 | 99.28 | 99.28 |

Table 1: Rej. prob. and perc. of $\phi_n^{\text{RSW}} \geq \phi_n^{\text{CCK}}$ and $\phi_n^{\text{RSW2}} \geq \phi_n^{\text{CCK2}}$ in Model 1.

| | | $\theta = 0$ | | | | | | $\theta = 0.2$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **RSW** | **RSW2** | **CCK** | **CCK2** | **≥CCK** | **≥CCK2** | **RSW** | **RSW2** | **CCK** | **CCK2** | **≥CCK** | **≥CCK2** |
| $p_n = 40$ | $\rho = 0$ | 4.65 | 5.22 | 0.86 | 5.35 | 100.00 | 99.77 | 46.29 | 49.36 | 16.65 | 47.66 | 100.00 | 98.66 |
| | $\rho = 0.5$ | 3.54 | 4.20 | 0.53 | 4.46 | 100.00 | 99.24 | 22.28 | 25.93 | 9.59 | 27.55 | 100.00 | 95.53 |
| | $\rho = 0.9$ | 2.37 | 3.20 | 0.78 | 4.43 | 100.00 | 98.42 | 13.93 | 17.26 | 8.73 | 22.69 | 100.00 | 93.34 |
| $p_n = 100$ | $\rho = 0$ | 4.54 | 5.61 | 0.76 | 5.64 | 100.00 | 99.67 | 56.98 | 62.82 | 21.87 | 56.47 | 100.00 | 99.28 |
| | $\rho = 0.5$ | 3.28 | 4.50 | 0.65 | 4.33 | 100.00 | 99.36 | 25.09 | 31.21 | 12.34 | 31.63 | 100.00 | 95.81 |
| | $\rho = 0.9$ | 1.88 | 2.85 | 0.88 | 3.87 | 100.00 | 98.58 | 15.14 | 19.94 | 11.10 | 25.10 | 100.00 | 93.37 |
| $p_n = 200$ | $\rho = 0$ | 4.63 | 5.93 | 0.66 | 5.58 | 100.00 | 99.80 | 66.70 | 74.50 | 26.44 | 63.46 | 100.00 | 99.70 |
| | $\rho = 0.5$ | 3.10 | 4.61 | 0.93 | 4.29 | 100.00 | 99.38 | 26.42 | 35.21 | 15.15 | 34.07 | 100.00 | 96.17 |
| | $\rho = 0.9$ | 1.91 | 3.18 | 1.05 | 3.93 | 100.00 | 98.78 | 15.48 | 21.50 | 12.29 | 25.53 | 99.97 | 94.23 |

Table 2: Rej. prob. and perc. of $\phi_n^{\text{RSW}} \geq \phi_n^{\text{CCK}}$ and $\phi_n^{\text{RSW2}} \geq \phi_n^{\text{CCK2}}$ in Model 2.

| | | $\theta = 0$ | | | | | | $\theta = 0.2$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **RSW** | **RSW2** | **CCK** | **CCK2** | **≥CCK** | **≥CCK2** | **RSW** | **RSW2** | **CCK** | **CCK2** | **≥CCK** | **≥CCK2** |
| $p_n = 40$ | $\rho = 0$ | 4.73 | 6.42 | 5.07 | 6.94 | 99.66 | 99.48 | 18.51 | 22.44 | 19.25 | 23.39 | 99.26 | 99.05 |
| | $\rho = 0.5$ | 4.71 | 6.33 | 5.07 | 6.63 | 99.64 | 99.70 | 17.06 | 20.56 | 17.78 | 21.40 | 99.28 | 99.16 |
| | $\rho = 0.9$ | 4.55 | 5.64 | 4.89 | 6.19 | 99.66 | 99.45 | 18.99 | 22.12 | 19.78 | 23.36 | 99.21 | 98.76 |
| $p_n = 100$ | $\rho = 0$ | 4.27 | 6.79 | 4.61 | 7.16 | 99.66 | 99.63 | 23.30 | 30.07 | 24.33 | 30.98 | 98.97 | 99.09 |
| | $\rho = 0.5$ | 4.27 | 6.56 | 4.64 | 6.99 | 99.63 | 99.57 | 20.66 | 26.82 | 21.48 | 27.73 | 99.18 | 99.09 |
| | $\rho = 0.9$ | 4.20 | 5.92 | 4.51 | 6.23 | 99.69 | 99.69 | 18.20 | 22.22 | 18.91 | 23.23 | 99.29 | 98.99 |
| $p_n = 200$ | $\rho = 0$ | 4.52 | 7.57 | 4.85 | 7.86 | 99.67 | 99.71 | 29.39 | 39.63 | 30.70 | 40.74 | 98.69 | 98.89 |
| | $\rho = 0.5$ | 4.43 | 7.44 | 4.75 | 7.79 | 99.68 | 99.65 | 25.19 | 33.34 | 26.09 | 34.25 | 99.10 | 99.09 |
| | $\rho = 0.9$ | 4.64 | 6.38 | 4.90 | 6.74 | 99.74 | 99.64 | 19.43 | 25.02 | 20.17 | 25.82 | 99.26 | 99.20 |

Table 3: Rej. prob. and perc. of $\phi_n^{\text{RSW}} \geq \phi_n^{\text{CCK}}$ and $\phi_n^{\text{RSW2}} \geq \phi_n^{\text{CCK2}}$ in Model 3.

| | | RSW | RSW2 | CCK | CCK2 | ≥CCK | ≥CCK2 | RSW | RSW2 | CCK | CCK2 | ≥CCK | ≥CCK2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\theta = 0$ | | | | | | $\theta = 0.2$ | | | | | |
| $p_n = 40$ | $\rho = 0$ | 4.17 | 4.78 | 0.81 | 4.95 | 100.00 | 99.64 | 45.86 | 48.81 | 16.20 | 46.76 | 100.00 | 98.81 |
| | $\rho = 0.5$ | 4.38 | 4.88 | 0.71 | 4.98 | 100.00 | 99.72 | 41.48 | 43.86 | 15.47 | 42.78 | 100.00 | 98.77 |
| | $\rho = 0.9$ | 4.50 | 4.90 | 1.13 | 5.47 | 100.00 | 99.37 | 42.13 | 43.94 | 17.39 | 44.92 | 100.00 | 97.95 |
| $p_n = 100$ | $\rho = 0$ | 4.82 | 5.66 | 0.56 | 5.61 | 100.00 | 99.72 | 57.38 | 62.71 | 22.13 | 56.36 | 100.00 | 99.22 |
| | $\rho = 0.5$ | 4.25 | 5.21 | 0.46 | 5.13 | 100.00 | 99.76 | 48.76 | 53.33 | 18.85 | 48.68 | 100.00 | 99.00 |
| | $\rho = 0.9$ | 4.47 | 5.02 | 0.73 | 5.15 | 100.00 | 99.70 | 40.75 | 43.84 | 16.11 | 42.64 | 100.00 | 98.30 |
| $p_n = 200$ | $\rho = 0$ | 4.18 | 5.48 | 0.45 | 5.08 | 100.00 | 99.86 | 67.02 | 74.82 | 26.67 | 63.64 | 100.00 | 99.72 |
| | $\rho = 0.5$ | 4.61 | 5.76 | 0.50 | 5.48 | 100.00 | 99.77 | 56.46 | 63.81 | 23.77 | 54.43 | 100.00 | 99.67 |
| | $\rho = 0.9$ | 4.31 | 5.11 | 0.60 | 5.07 | 100.00 | 99.74 | 41.24 | 45.95 | 16.52 | 42.07 | 100.00 | 98.89 |

Table 4: Rej. prob. and perc. of $\phi_n^{\text{RSW}} \geq \phi_n^{\text{CCK}}$ and $\phi_n^{\text{RSW2}} \geq \phi_n^{\text{CCK2}}$ in Model 4.

# 4 Appendix

*Proof of Theorem 2.1:* For any vector $(\lambda_1, \ldots, \lambda_{p_n})' \equiv \lambda \in \mathbf{R}^{p_n}$, measure $P$, and $x \in \mathbf{R}$ define

$$F_n(x, \lambda, P) \equiv P\left\{0 \vee \sqrt{n}(\bar{X}_{j,n} - \mu_j(P) + \lambda_j) \leq x S_{j,n} \text{ for all } 1 \leq j \leq p_n\right\}$$
$$J_n(x, \lambda, P) \equiv P\left\{\sqrt{n}(\bar{X}_{j,n} - \mu_j(P) + \lambda_j) \leq x S_{j,n} \text{ for all } 1 \leq j \leq p_n\right\},$$

and for any function $f : \mathbf{R} \to [0,1]$ let $f^{-1}(x) \equiv \inf\{c : f(c) \geq x\}$ with $f^{-1}(x) = +\infty$ whenever $\{c : f(c) \geq x\}$ is empty. Further define the event $\Omega_n(P)$ according to

$$\Omega_n(P) \equiv \{\mu_j(P) \leq \hat{u}_{j,n} \text{ for all } 1 \leq j \leq p_n\}, \tag{13}$$

and note that, for $(\hat{u}_{1,n}, \ldots, \hat{u}_{p_n,n})' \equiv \hat{u}_n \in \mathbf{R}^{p_n}$, the event $\Omega_n(P)$ implies $F_n(x, \mu(P), \hat{P}_n) \geq F_n(x, \hat{u}_n, \hat{P}_n)$ for all $x \in \mathbf{R}$, which yields $F_n^{-1}(x, \mu(P), \hat{P}_n) \leq F_n^{-1}(x, \hat{u}_n, \hat{P}_n)$ for all $x \in [0,1]$. In particular, by definition of $\hat{c}_n^{(2)}(1-\alpha+\beta)$ we obtain that $\Omega_n(P)$ implies $F_n^{-1}(1-\alpha+\beta, \mu(P), \hat{P}_n) \leq \hat{c}_n^{(2)}(1-\alpha+\beta)$, and hence Lemma 4.1 yields

$$\begin{aligned}\limsup_{n\to\infty} \sup_{P \in \mathbf{P}_{0,n}} P\left\{T_n > \hat{c}_n^{(2)}(1-\alpha+\beta)\right\} &\leq \limsup_{n\to\infty} \sup_{P \in \mathbf{P}_{0,n}} P\left\{T_n > \hat{c}_n^{(2)}(1-\alpha+\beta);\ \Omega_n(P)\right\} + \beta \\ &\leq \limsup_{n\to\infty} \sup_{P \in \mathbf{P}_{0,n}} P\left\{T_n > F_n^{-1}(1-\alpha+\beta, \mu(P), \hat{P}_n)\right\} + \beta. \end{aligned} \tag{14}$$

Next, note that $S_{j,n} \geq 0$ almost surely implies $F_n(x, \lambda, P) = J_n(x, \lambda, P)$ for any $\lambda, P$, and $x \geq 0$, while for any $\lambda, P$ and $x < 0$ we have $F_n(x, \lambda, P) \leq P\{S_{j,n} = 0 \text{ for all } 1 \leq j \leq p_n\}$. Hence, it follows that

$$\begin{aligned}&\sup_{x \in \mathbf{R}} \left|F_n(x, \mu(P), P) - F_n(x, \mu(P), \hat{P}_n)\right| \\ &\qquad \leq \sup_{x \geq 0} \left|J_n(x, \mu(P), P) - J_n(x, \mu(P), \hat{P}_n)\right| + P\left\{\max_{1 \leq j \leq p_n} S_{j,n} = 0\right\} + \hat{P}_n\left\{\max_{1 \leq j \leq p_n} S_{j,n} = 0\right\},\end{aligned}$$

which together with Lemmas 4.2 and 4.3 implies there are sequence $\xi_n \downarrow 0$ and $\delta_n \downarrow 0$ such that

$$\inf_{P \in \mathbf{P}_n} P\left\{\sup_{x \in \mathbf{R}} \left|F_n(x, \mu(P), P) - F_n(x, \mu(P), \hat{P}_n)\right| \leq \xi_n\right\} \geq 1 - \delta_n. \tag{15}$$

Moreover, since $F_n(F_n^{-1}(1-\alpha+\beta, \mu(P), \hat{P}_n), \mu(P), \hat{P}_n) \geq 1-\alpha+\beta$, it follows that

$$\left\{\sup_{x \in \mathbf{R}} \left|F_n(x, \mu(P), P) - F_n(x, \mu(P), \hat{P}_n)\right| \leq \xi_n\right\} \subseteq \left\{F_n(F_n^{-1}(1-\alpha+\beta, \mu(P), \hat{P}_n), \mu(P), P) \geq 1-\alpha+\beta-\xi_n\right\}$$

$$\subseteq \left\{ F_n^{-1}(1-\alpha+\beta, \mu(P), \hat{P}_n) \geq F_n^{-1}(1-\alpha+\beta-\xi_n, \mu(P), P) \right\} . \quad (16)$$

Thus, since $P\{T_n \leq x\} = F_n(x, \mu(P), P)$, results (15) and (16) together establish that

$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}_{0,n}} P\left\{ T_n > F_n^{-1}(1-\alpha+\beta, \mu(P), \hat{P}_n) \right\}$$
$$\leq \limsup_{n\to\infty} \sup_{P\in\mathbf{P}_{0,n}} P\left\{ T_n > F_n^{-1}(1-\alpha+\beta-\xi_n, \mu(P), P) \right\} + \delta_n \leq \limsup_{n\to\infty} \alpha - \beta - \xi_n + \delta_n . \quad (17)$$

The claim of the theorem therefore follows from (14), (17), $\xi_n \downarrow 0$, and $\delta_n \downarrow 0$. ■

**Lemma 4.1.** *Let Assumption 2.1 hold. If $\beta \in (0, 0.5)$, then it follows that*

$$\liminf_{n\to\infty} \inf_{P\in\mathbf{P}_{0,n}} P\left\{ \mu_j(P) \leq \hat{u}_{j,n} \text{ for all } 1 \leq j \leq p_n \right\} \geq 1-\beta .$$

*Proof:* The proof follows from Lemma 4.2 and arguments in the proof of Lemma A.1 in Romano and Shaikh (2012). First note that for any $P \in \mathbf{P}_{0,n}$ we have $\mu_j(P) \leq 0$ for all $1 \leq j \leq p_n$, and therefore by definition of $\hat{u}_{j,n}$

$$P\left\{ \mu_j(P) \leq \hat{u}_{j,n} \text{ for all } 1 \leq j \leq p_n \right\} = P\left\{ \sqrt{n}(\mu_j(P) - \bar{X}_{j,n}) \leq S_{j,n}\hat{c}_n^{(1)}(1-\beta) \text{ for all } 1 \leq j \leq p_n \right\} . \quad (18)$$

Next, for any measure $P$ we define the function $F_n(\cdot, P) : \mathbf{R} \to [0,1]$ to be given by

$$F_n(x, P) \equiv P\left\{ \sqrt{n}(\mu_j(P) - \bar{X}_{j,n}) \leq S_{j,n}x \text{ for all } 1 \leq j \leq p_n \right\} . \quad (19)$$

Then note that if $\{X_i\}_{i=1}^n$ satisfies Assumption 2.1, then so does $\{-X_i\}_{i=1}^n$. Hence, we may apply Lemma 4.2 to conclude there exist sequences $\xi_n \downarrow 0$ and $\delta_n \downarrow 0$ such that

$$\inf_{P\in\mathbf{P}_n} P\left\{ \sup_{x\geq 0} \left| F_n(x, P) - F_n(x, \hat{P}_n) \right| \leq \xi_n \right\} \geq 1-\delta_n . \quad (20)$$

Further let $\Phi$ denote the c.d.f. of a standard normal random variable and note that Theorem 1.1. Bentkus and Götze (1996) and Assumption 2.1(iii) imply

$$\sup_{P\in\mathbf{P}_n} F_n(0, P) \leq \sup_{P\in\mathbf{P}_n} P\left\{ \sqrt{n}(\mu_1(P) - \bar{X}_{1,n}) \leq S_{1,n} \times 0 \right\} \leq 0.5 + \frac{KM_{1,n}}{\sqrt{n}} \quad (21)$$

for some finite constant $K \in \mathbf{R}$. Next, for any $f : \mathbf{R} \to [0,1]$ let $f^{-1}(x) \equiv \inf\{c : f(c) \geq x\}$ with $f^{-1}(x) = +\infty$ if $\{c : f(c) \geq x\} = \emptyset$, and define the event $\Omega_n(P)$ to be given by

$$\Omega_n(P) \equiv \left\{ \sup_{x\geq 0} \left| F_n(x, P) - F_n(x, \hat{P}_n) \right| \leq \xi_n \right\} . \quad (22)$$

Then note that since $\beta < 0.5$ and $M_{1,n}/\sqrt{n} = o(1)$ by hypothesis, result (21) implies that

$$\sup_{P\in\mathbf{P}_n} F_n(0, P) + \xi_n < 1-\beta \quad (23)$$

for $n$ sufficiently large. Therefore, the definitions of $\hat{c}_n^{(1)}(1-\beta)$ and $\Omega_n(P)$ yield

$$\Omega_n(P) \subseteq \{F_n(0, \hat{P}_n) < 1-\beta\} \subseteq \{\hat{c}_n^{(1)}(1-\beta) \geq 0\} \quad (24)$$

for $n$ sufficiently large. Combining definition (22) and result (24) further implies

$$\Omega_n(P) \subseteq \left\{F_n(\hat{c}_n^{(1)}(1-\beta), P) \geq F_n(\hat{c}_n^{(1)}(1-\beta), \hat{P}_n) - \xi_n\right\}$$
$$\subseteq \left\{F_n(\hat{c}_n^{(1)}(1-\beta), P) \geq 1-\beta-\xi_n\right\} \subseteq \left\{\hat{c}_n^{(1)}(1-\beta) \geq F_n^{-1}(1-\beta-\xi_n, P)\right\}, \quad (25)$$

where the second and third set inclusions follow by definition of $\hat{c}_n^{(1)}(1-\beta)$ and $F_n^{-1}(\cdot, P)$. Hence, results (18), (20), and the definitions of $F_n^{-1}(\cdot, P)$ and $\Omega_n(P)$ yield

$$\liminf_{n\to\infty} \inf_{P\in\mathbf{P}_{0,n}} P\left\{\mu_j(P) \leq \hat{u}_{j,n}\ \forall 1 \leq j \leq p_n\right\}$$
$$\geq \liminf_{n\to\infty} \inf_{P\in\mathbf{P}_n} P\left\{\sqrt{n}(\mu_j(P) - \bar{X}_{j,n}) \leq S_{j,n} F_n^{-1}(1-\beta-\xi_n, P)\ \forall 1 \leq j \leq p_n\right\} - \delta_n$$
$$\geq \liminf_{n\to\infty} 1-\beta-\xi_n-\delta_n,$$

which establishes the claim of the lemma because $\xi_n \downarrow 0$ and $\delta_n \downarrow 0$. ■

**Lemma 4.2.** *Let Assumption 2.1 hold and for any $(\lambda_1, \ldots, \lambda_{p_n})' \equiv \lambda \in \mathbf{R}_-^{p_n}$, $P \in \mathbf{P}_n$, and $x \in \mathbf{R}$ define*

$$J_n(x, \lambda, P) \equiv P\left\{\sqrt{n}(\bar{X}_{j,n} - \mu_j(P) + \lambda_j) \leq x S_{j,n} \text{ for all } 1 \leq j \leq p_n\right\}.$$

*Then, there exists a sequence $\xi_n \downarrow 0$ such that*

$$\liminf_{n\to\infty} \inf_{P\in\mathbf{P}_n} P\left\{\sup_{x\geq 0} \sup_{\lambda\in\mathbf{R}_-^{p_n}} \left|J_n(x, \lambda, \hat{P}_n) - J_n(x, \lambda, P)\right| \leq \xi_n\right\} = 1.$$

*Proof:* We first note that $\sigma_j(P) > 0$ for all $1 \leq j \leq p_n$ by Assumption 2.1(ii) implies that

$$J_n(x, \lambda, P) = P\left\{\frac{\sqrt{n}(\bar{X}_{j,n} - \mu_j(P))}{\sigma_j(P)} \leq x\frac{S_{j,n}}{\sigma_j(P)} - \frac{\sqrt{n}\lambda_j}{\sigma_j(P)} \text{ for all } 1 \leq j \leq p_n\right\}$$
$$J_n(x, \lambda, \hat{P}_n) = \hat{P}_n\left\{\frac{\sqrt{n}(\bar{X}_{j,n} - \mu_j(\hat{P}_n))}{\sigma_j(P)} \leq x\frac{S_{j,n}}{\sigma_j(P)} - \frac{\sqrt{n}\lambda_j}{\sigma_j(P)} \text{ for all } 1 \leq j \leq p_n\right\}.$$

Next, let $(Z_1, \ldots, Z_{p_n})' \equiv Z \in \mathbf{R}^{p_n}$ be a Gaussian vector satisfying $E[Z_j] = 0$ and $E[Z_j Z_k] = E_P[(X_{i,j} - \mu_j(P))(X_{i,k} - \mu_k(P))]/\sigma_j(P)\sigma_k(P)$ for any $1 \leq j, k \leq p_n$, and for any measure $P$, $(\lambda_1, \ldots, \lambda_{p_n})' \equiv \lambda \in \mathbf{R}_-^{p_n}$ and $(\omega_1, \ldots, \omega_{p_n})' \equiv \omega \in \mathbf{R}^{p_n}$ satisfying $\omega_j > 0$ for all $1 \leq j \leq p_n$, define $F_n(x, \lambda, \omega, P)$ and $G_n(x, \lambda, \omega, P)$ to equal

$$F_n(x, \lambda, \omega, P) \equiv P\left\{\frac{\sqrt{n}(\bar{X}_{j,n} - \mu_j(P))}{\omega_j} \leq x - \frac{\sqrt{n}\lambda_j}{\omega_j} \text{ for all } 1 \leq j \leq p_n\right\} \quad (26)$$
$$G_n(x, \lambda, \omega, P) \equiv P\left\{Z_j \leq x - \frac{\sqrt{n}\lambda_j}{\omega_j} \text{ for all } 1 \leq j \leq p_n\right\}. \quad (27)$$

Since $B_n^2 \log^{3.5}(p_n)/n^{(1-\delta)/2} = o(1)$ for some $\delta > 0$ by Assumption 2.1(v), we may find an $\epsilon_n \downarrow 0$ satisfying

$$\frac{B_n^2 \log^2(p_n)}{n^{(1-\delta)/2}} = o(\epsilon_n) \qquad \log(p_n)\epsilon_n = o(1).$$

In particular, the condition $B_n^2 \log^2(p_n)/n^{(1-\delta)/2} = o(\epsilon_n)$ implies that the sequence $\eta_n$ defined by

$$\eta_n \equiv \sup_{P\in\mathbf{P}_n} P\left\{\max_{1\leq j\leq p_n} \left|\frac{S_{j,n}}{\sigma_j(P)} - 1\right| > \epsilon_n\right\} \quad (28)$$

satisfies $\eta_n = o(1)$ by Lemma 4.3(i). Moreover, by definitions (26) and (28) we can conclude that

$$F_n(x(1-\epsilon_n), \lambda, \sigma(P), P) - \eta_n \le J_n(x, \lambda, P) \le F_n(x(1+\epsilon_n), \lambda, \sigma(P), P) + \eta_n \tag{29}$$

for all $x \ge 0$, $P \in \mathbf{P}_n$, and $\lambda \in \mathbf{R}_-^{p_n}$. Next note $(M_{1,n}^2 \vee M_{2,n}^2 \vee B_n^2)\log^{3.5}(p_n n)/\sqrt{n} = o(1)$ by Assumption 2.1(v), Assumptions 2.1(i)(iii)(iv) and Proposition 2.1 in Chernozhukov et al. (2017) imply that

$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}_n} \sup_{x\in\mathbf{R}} \sup_{\lambda\in\mathbf{R}_-^{p_n}} |F_n(x, \lambda, \sigma(P), P) - G_n(x, \lambda, \sigma(P), P)| = 0. \tag{30}$$

On the other hand, we may further conclude by Lemma 4.4 and $\epsilon_n \log(p_n) = o(1)$ by construction that

$$\begin{aligned}
&\limsup_{n\to\infty} \sup_{P\in\mathbf{P}_n} \sup_{x\ge 0} \sup_{\lambda\in\mathbf{R}_-^{p_n}} G_n((1+\epsilon_n)x, \lambda, \sigma(P), P) - G_n((1-\epsilon_n)x, \lambda, \sigma(P), P) \\
&\qquad \le \limsup_{n\to\infty} \sup_{P\in\mathbf{P}_n} \sup_{x\ge 0} \sup_{\lambda\in\mathbf{R}_-^{p_n}} P\left\{\left|\max_{1\le j\le p_n} Z_j + \frac{\sqrt{n}\lambda_j}{\sigma_j(P)} - x\right| \le 2\epsilon_n x\right\} = 0.
\end{aligned} \tag{31}$$

Therefore, combining results (28), (29), (30), and (31) and employing that $\eta_n = o(1)$ we obtain

$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}_n} \sup_{x\ge 0} \sup_{\lambda\in\mathbf{R}_-^{p_n}} |J_n(x, \lambda, P) - G_n(x, \lambda, \sigma(P), P)| = 0. \tag{32}$$

To conclude the proof, we set $\bar{M}_n \equiv M_{1,n} \vee M_{2,n} \vee B_n$ and define the events $\Omega_{1,n}(P)$ and $\Omega_{2,n}(P)$ according to

$$\begin{aligned}
\Omega_{1,n}(P) &\equiv \left\{P\left\{\max_{1\le j\le p_n}\left|\frac{S_{j,n}^*}{\sigma_j(P)} - 1\right| > \epsilon_n \Big| \{X_i\}_{i=1}^n\right\} \le \frac{K}{n^\delta}\right\} \\
\Omega_{2,n}(P) &\equiv \left\{\sup_{x\in\mathbf{R}} \sup_{\lambda\in\mathbf{R}_-^{p_n}} \left|F_n(x, \lambda, \sigma(P), \hat{P}_n) - G_n(x, \lambda, \sigma(P), P)\right| \le K\left(\frac{\bar{M}_n^2 \log^{3.5}(p_n n)}{n^{(1-\delta)/2}}\right)^{1/6}\right\}
\end{aligned}$$

and note that for $\Omega_n(P) \equiv \Omega_{1,n}(P) \cap \Omega_{2,n}(P)$, for appropriately selected $K < \infty$, Lemma 4.3(ii) and Proposition 4.3 in Chernozhukov et al. (2017) (applied with $\alpha = n^{-\delta}$) allow us to conclude that

$$\liminf_{n\to\infty} \inf_{P\in\mathbf{P}_n} P\{\Omega_n(P)\} = 1. \tag{33}$$

Furthermore, observe that under $\Omega_n(P)$ we may argue as in result (29) to obtain that for all $x \ge 0$ and $\lambda \in \mathbf{R}_-^{p_n}$

$$\begin{aligned}
J_n(x, \lambda, \hat{P}_n) &\le F_n((1+\epsilon_n)x, \lambda, \sigma(P), \hat{P}_n) + \frac{K}{n^\delta} \\
J_n(x, \lambda, \hat{P}_n) &\ge F_n((1-\epsilon_n)x, \lambda, \sigma(P), \hat{P}_n) - \frac{K}{n^\delta}.
\end{aligned}$$

Therefore, employing results (31) and (33) imply that there exists a sequence $\xi_n \downarrow 0$ such that

$$\liminf_{n\to\infty} \inf_{P\in\mathbf{P}_n} P\left\{\sup_{x\ge 0} \sup_{\lambda\in\mathbf{R}_-^{p_n}} |J_n(x, \lambda, \hat{P}_n) - G_n(x, \lambda, \sigma(P), P)| \le \xi_n\right\} = 1. \tag{34}$$

The lemma thus follows from results (32) and (34). ■

**Lemma 4.3.** *Let Assumption 2.1(i)(ii)(iv) hold. Then: (i) For any sequence $\epsilon_n \downarrow 0$ satisfying $B_n^2 \log^2(p_n)/n^{(1-\delta)/2} =$*

*$o(\epsilon_n)$ for some $\delta \in (0,1)$ it follows that*

$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}_n} P\left\{\max_{1\le j\le p_n}\left|\frac{S_{j,n}}{\sigma_j(P)}-1\right| > \epsilon_n\right\} = 0. \tag{35}$$

*(ii) For any $\epsilon_n \downarrow 0$ satisfying the condition of part (i) there is a $K < \infty$ such that*

$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}_n} P\left\{P\left\{\max_{1\le j\le p_n}\left|\frac{S^*_{j,n}}{\sigma_j(P)}-1\right| > \epsilon_n \Big| \{X_i\}_{i=1}^n\right\} \le \frac{K}{n^\delta}\right\} = 1.$$

*Proof:* The first claim of the lemma corresponds to Lemma D.5 in Chernozhukov et al. (2019), which we may apply by Assumptions 2.1(i)(ii)(iv). In order to establish the second claim of the lemma we first define the event

$$\Omega_{1,n}(P) \equiv \left\{\max_{1\le j\le p_n}\left|\frac{S_{j,n}}{\sigma_j(P)}-1\right| \le \frac{\epsilon_n}{2}\right\},$$

where $\epsilon_n$ satisfies $B_n^2\log^2(p_n)/n^{(1-\delta)/2} = o(\epsilon_n)$ for some $\delta \in (0,1)$ by hypothesis. We further define $\hat{B}_n^4 \in \mathbf{R}$ to equal

$$\hat{B}_n^4 \equiv \frac{1}{n}\sum_{i=1}^n \max_{1\le j\le p_n}\left(\frac{X_{i,j}-\bar{X}_{j,n}}{S_{j,n}}\right)^4$$

and note that since $\epsilon_n \downarrow 0$ it follows that, for $n$ sufficiently large, $\Omega_{1,n}(P)$ implies $S_{j,n}$ is positive for all $1 \le j \le p_n$. Furthermore, Lemma D.5 in Chernozhukov et al. (2019) implies there are finite positive $K_1, K_2 \in \mathbf{R}$ satisfying

$$I\{\Omega_{1,n}(P)\} \times P\left\{\max_{1\le j\le p_n}\left|\frac{S^*_{j,n}}{S_{j,n}}-1\right| > K_1\frac{\hat{B}_n^2\log^2(p_n)}{n^{(1-\delta)/2}}\Big|\{X_i\}_{i=1}^n\right\} \le I\{\Omega_{1,n}(P)\} \times \frac{K_2}{n^\delta}. \tag{36}$$

Moreover, the definition of the event $\Omega_{1,n}(P)$ and the inequality $(a+b)^4 \le 8(a^4+b^4)$ also yield that

$$\begin{aligned} I\{\Omega_{1,n}(P)\} \times \hat{B}_n^4 &\le I\{\Omega_{1,n}(P)\} \times \max_{1\le j\le p_n}\frac{\sigma_j^4(P)}{S_{j,n}^4} \times \frac{1}{n}\sum_{i=1}^n \max_{1\le j\le p_n}\left(\frac{X_{i,j}-\bar{X}_{j,n}}{\sigma_j(P)}\right)^4 \\ &\le 8\left(1+\frac{\epsilon_n}{2}\right)^4 \times \frac{1}{n}\sum_{i=1}^n\left(\max_{1\le j\le p_n}\left(\frac{X_{i,j}-\mu_j(P)}{\sigma_j(P)}\right)^4 + \max_{1\le j\le p_n}\left(\frac{\bar{X}_{j,n}-\mu_j(P)}{\sigma_j(P)}\right)^4\right). \end{aligned} \tag{37}$$

Next note that for any sequence $\ell_n \downarrow 0$, Assumption 2.1(iv) and Markov's inequality imply that

$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}_n} P\left\{\frac{1}{n}\sum_{i=1}^n \max_{1\le j\le p_n}\left(\frac{X_{i,j}-\mu_j(P)}{\sigma_j(P)}\right)^4 > \frac{B_n^4}{\ell_n}\right\} = 0. \tag{38}$$

Furthermore, since $B_n \ge 1$ by Jensen's inequality, we note that $\epsilon_n \downarrow 0$ and the condition $B_n^2\log^2(p_n)/n^{(1-\delta)/2} = o(\epsilon_n)$ together imply that $\log^2(p_n)/n = o(1)$. Therefore, $\ell_n \downarrow 0$ and equation (73) in Chernozhukov et al. (2019) yield

$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}_n} P\left\{\max_{1\le j\le p_n}\left|\frac{1}{n}\sum_{i=1}^n \frac{X_{i,j}-\mu_j(P)}{\sigma_j(P)}\right|^4 > \frac{B_n^4}{\ell_n}\right\} = 0. \tag{39}$$

Combining results (37), (38), (39), and that $P\{\Omega_{1,n}(P)\} = 1 + o(1)$ uniformly in $P \in \mathbf{P}_n$ by part (i) of this lemma, it follows that there exists a constant $K_3 < \infty$ independent of the sequence $\ell_n$ with

$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}_n} P\left\{\hat{B}_n^4 > K_3\frac{B_n^4}{\ell_n}\right\} = 0.$$

Thus, by selecting $\ell_n \downarrow 0$ to satisfy $B_n^2\log^2(p_n)/(\sqrt{\ell_n}n^{(1-\delta)/2}) = o(\epsilon_n)$, which is possible due to $B_n^2\log^2(p_n)/n^{(1-\delta)/2} =$

$o(\epsilon_n)$ by hypothesis, we are able to conclude from result (36) that

$$\limsup_{n\to\infty}\sup_{P\in\mathbf{P}_n} P\left\{P\left\{\max_{1\le j\le p_n}\left|\frac{S^*_{j,n}}{S_{j,n}}-1\right|>\frac{\epsilon_n}{4}\Big|\{X_i\}_{i=1}^n\right\}\le\frac{K_2}{n^\delta}\right\}=1. \tag{40}$$

Finally, note that for any $(a_1,\ldots,a_{p_n})'\in\mathbf{R}^{p_n}$, we obtain by definition of the event $\Omega_{1,n}(P)$ that

$$\begin{aligned} I\{\Omega_{1,n}(P)\}\times\max_{1\le j\le p_n}\left|\frac{a_j}{\sigma_j(P)}-1\right| &\le I\{\Omega_{1,n}(P)\}\times\left(\max_{1\le j\le p_n}\left|\frac{a_j}{S_{j,n}}-1\right|\frac{S_{j,n}}{\sigma_j(P)}+\max_{1\le j\le p_n}\left|\frac{S_{j,n}}{\sigma_j(P)}-1\right|\right)\\ &\le I\{\Omega_{1,n}(P)\}\times\left(\max_{1\le j\le p_n}\left|\frac{a_j}{S_{j,n}}-1\right|(1+\frac{\epsilon_n}{2})+\frac{\epsilon_n}{2}\right). \end{aligned} \tag{41}$$

Thus, $P\{\Omega_{1,n}(P)\}=1+o(1)$ uniformly in $P\in\mathbf{P}_n$ by part (i) of this lemma, and results (40) and (41) imply

$$\begin{aligned} &\limsup_{n\to\infty}\sup_{P\in\mathbf{P}_n} P\left\{P\left\{\max_{1\le j\le p_n}\left|\frac{S^*_{j,n}}{\sigma_j(P)}-1\right|>\epsilon_n\Big|\{X_i\}_{i=1}^n\right\}\le\frac{K_2}{n^\delta}\right\}\\ &\qquad\ge\limsup_{n\to\infty}\sup_{P\in\mathbf{P}_n} P\left\{P\left\{\max_{1\le j\le p_n}\left|\frac{S^*_{j,n}}{S_{j,n}}-1\right|>\frac{\epsilon_n}{4}\Big|\{X_i\}_{i=1}^n\right\}\le\frac{K_2}{n^\delta}\right\}=1, \end{aligned}$$

which establishes the second claim of the lemma. ■

**Lemma 4.4.** *Let $(Z_1,\ldots,Z_p)'\equiv Z\in\mathbf{R}^p$ be Gaussian with $E[Z_j]=0$ and $E[Z_j^2]=1$ for all $1\le j\le p$, and $(s_1,\ldots,s_p)\equiv s\in\mathbf{R}^p_-$. Then, there is a constant $C<\infty$ such that for all $\delta\in(0,0.5]$ and $t>0$:*

$$\sup_{x\ge 0}P\left\{\left|\max_{1\le j\le p}(Z_j+s_j)-x\right|\le\delta x\right\}\le C\delta(1+\sqrt{\log(p)}+t)^2+\exp\left\{-\frac{t^2}{2}\right\}.$$

*Proof:* Let $\mathrm{m}_p$ denote the median of $\max_{1\le j\le p}Z_j$, and note that by Kwapień (1994) $\mathrm{m}_p\le E[\max_{1\le j\le p}Z_j]$. Since in addition $E[\max_{1\le j\le p}Z_j]\le\sqrt{2\log(p)}$ by Lemmas 2.2.1 and 2.2.2 in van der Vaart and Wellner (1996), we obtain

$$\mathrm{m}_p\le\sqrt{2\log(p)}. \tag{42}$$

Next, for any $t>0$ we set $a\equiv 2(\sqrt{2\log(p)}+t)$ and observe the union bound allows us to conclude that

$$\begin{aligned} \sup_{0\le x\le a}P\left\{\left|\max_{1\le j\le p}(Z_j+s_j)-x\right|\le\delta x\right\} &\le\sup_{0\le x\le a}P\left\{\left|\max_{1\le j\le p:s_j\le -a/2}(Z_j+s_j)-x\right|\le\delta x\right\}\\ &\quad+\sup_{0\le x\le a}P\left\{\left|\max_{1\le j\le p:s_j>-a/2}(Z_j+s_j)-x\right|\le\delta x\right\}. \end{aligned} \tag{43}$$

Moreover, we note that $\delta\in(0,0.5]$ and $x>0$ imply $x(1-\delta)>0$, and hence we obtain

$$\begin{aligned} \sup_{0\le x\le a}P\left\{\left|\max_{1\le j\le p:s_j\le -a/2}(Z_j+s_j)-x\right|\le\delta x\right\} &\le P\left\{\max_{1\le j\le p:s_j\le -a/2}(Z_j+s_j)\ge 0\right\}\\ &\le P\left\{\max_{1\le j\le p}Z_j\ge\sqrt{2\log(p)}+t\right\}\le\exp\left\{-\frac{t^2}{2}\right\}, \end{aligned} \tag{44}$$

where the second inequality holds by definition of $a$, while the final inequality follows from Borell's inequality (see, e.g., the Corollary in pg. 82 of Davydov et al. (1998)), result (42), and $1-\Phi(t)\le\exp\{-t^2/2\}$ for any $t>0$ and $\Phi$ the c.d.f. of a standard normal random variable. Next note that Lemma A.1 in Chernozhukov et al. (2017) yields

$$\sup_{0\le x\le a}P\left\{\left|\max_{1\le j\le p:s_j>-a/2}(Z_j+s_j)-x\right|\le\delta x\right\}\lesssim\delta a\sqrt{\log(p)}. \tag{45}$$

Moreover, since $s_j \leq 0$ for all $1 \leq j \leq p$ and $\delta \leq 0.5$ we can additionally conclude that

$$\sup_{x \geq a} P\left\{\left|\max_{1\leq j\leq p}(Z_j + s_j) - x\right| \leq \delta x\right\} \leq \sup_{x\geq a} P\left\{\max_{1\leq j \leq p} Z_j \geq x(1-\delta)\right\} \leq P\left\{\max_{1\leq j\leq p} Z_j \geq \frac{a}{2}\right\} \leq \exp\left\{-\frac{t^2}{2}\right\}, \qquad (46)$$

where the final inequality follows by another application of Borell's inequality and the arguments employed in (44). The lemma follows from (43), (44), (45), and (46). ■

# References


Allen, R. (2018). Testing moment inequalities: Selection versus recentering. *Economics Letters*, **162** 124–126.

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

Bajari, P., Benkard, C. L. and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, **75** 1331–1370.

Bentkus, V. and Götze, F. (1996). The Berry-Esseen bound for Student's statistic. *The Annals of Probability*, **24** 491–503.

Canay, I. A. and Shaikh, A. M. (2017). Practical and Theoretical Advances in Inference for Partially Identified Models. In *Advances in Economics and Econometrics: Eleventh World Congress* (B. Honoré, A. Pakes, M. Piazzesi and L. Samuelson, eds.), vol. 2 of *Econometric Society Monographs*. Cambridge University Press, 271–306.

Chernozhukov, V., Chetverikov, D. and Kato, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *The Annals of Statistics*, **41** 2786–2819.

Chernozhukov, V., Chetverikov, D. and Kato, K. (2017). Central limit theorems and bootstrap in high dimensions. *The Annals of Probability*, **45** 2309–2352.

Chernozhukov, V., Chetverikov, D. and Kato, K. (2019). Inference on causal and structural parameters using many moment inequalities. *Review of Economic Studies (forthcoming)*.

Ciliberto, F. and Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, **77** 1791–1828.

Davydov, Y. A., Lifshits, M. A. and Smorodina, N. V. (1998). *Local properties of distributions of stochastic functionals*. American Mathematical Society.

Kwapień, S. (1994). A remark on the median and the expectation of convex functions of Gaussian vectors. In *Probability in Banach spaces, 9*. Springer, 271–272.

Romano, J. P. and Shaikh, A. M. (2012). On the uniform asymptotic validity of subsampling and the bootstrap. *The Annals of Statistics*, **40** 2798–2822.

Romano, J. P., Shaikh, A. M. and Wolf, M. (2014). A practical two-step method for testing moment inequalities. *Econometrica*, **82** 1979–2002.

Romano, J. P. and Wolf, M. (2005). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, **100** 94–108.

Romano, J. P. and Wolf, M. (2018). Multiple testing of one-sided hypotheses: combining Bonferroni and the bootstrap. In *International Conference of the Thailand Econometrics Society*. Springer, 78–94.

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