The Limit of Finite Sample Size and a Problem with Subsampling

Donald W. K. Andrews[1]
Cowles Foundation for Research in Economics
Yale University

Patrik Guggenberger
Department of Economics
UCLA

June 2005
Revised: February 2006
Preliminary and Incomplete

[1] Andrews gratefully acknowledges the research support of the National Science Foundation via grant number SES-0417911. Guggenberger gratefully acknowledges research support from a faculty research grant from UCLA in 2005. For helpful comments, we thank Russell Davidson and the participants at the conference on Inference on Incomplete Models at Northwestern University, at the CIREQ conference at the University of Montreal, and at seminars at the Stern School of Business NYU, University of Washington, University of British Columbia, University of Bonn, and University of Konstanz.

Abstract

This paper considers tests and confidence intervals based on a test statistic that has a limit distribution that is discontinuous in a nuisance parameter or the parameter of interest. The paper shows that standard fixed critical value (FCV) tests and subsample tests often have asymptotic size—defined as the limit of the finite sample size—that is greater than the nominal level of the test. We determine precisely the asymptotic size of such tests under a general set of high-level conditions that are relatively easy to verify. Often the asymptotic size is determined by a sequence of parameter values that approach the point of discontinuity of the asymptotic distribution. The problem is not a small sample problem. For every sample size, there can be parameter values for which the test over-rejects the null hypothesis. Analogous results hold for confidence intervals.

We introduce a hybrid subsample/FCV test that alleviates the problem of over-rejection asymptotically and in some cases eliminates it. In addition, we introduce size-corrections to the FCV, subsample, and hybrid tests that eliminate over-rejection asymptotically. In some examples, these size corrections are computationally challenging or intractable. In other examples, they are feasible.

Keywords: Asymptotic size, finite sample size, hybrid test, over-rejection, size correction, subsample confidence interval, subsample test.

JEL Classification Numbers:

1 Introduction

THIS INTRODUCTION IS QUITE PRELIMINARY. IN PARTICULAR, REFERENCES TO THE LITERATURE ARE NOT INCLUDED. SUGGESTIONS FOR RELEVANT REFERENCES ARE WELCOME.

In this paper, we show that if a sequence of test statistics has an asymptotic distribution that is discontinuous in the true parameter, then a subsample test based on the test statistic does not necessarily yield the desired asymptotic level. That is, the limit of the finite sample size of the test can exceed its nominal level. The same is shown to be true for a standard fixed critical value (FCV) test. Analogous problems arise with subsample and FCV confidence intervals (CIs) and confidence regions.

Examples where the asymptotic distribution of a test statistic is discontinuous in the true parameter include (i) models where a parameter may be on a boundary, (ii) models where a parameter may be unidentified at one or more points, e.g., as occurs in an instrumental variables model with weak instruments and in models where some parameters are only set identified, rather than point identified, (iii) models where the "degree of identification" of a parameter differs across parameter values causing different rates of convergence of estimators, as occurs in an autoregressive model with possible unit root, and (iv) models where shrinkage or post-model selection procedures are employed.

The intuition for the result stated above for a subsample test is roughly as follows. Suppose for a parameter θ we are interested in testing $H_0 : \theta = \theta_0$, a nuisance parameter γ appears under the null hypothesis, and the asymptotic distribution of the test statistic of interest is discontinuous at $\gamma = 0$. Then, a subsample test statistic based on a subsample of size $b_n << n$ behaves like it is closer to the discontinuity point $\gamma = 0$ than does the full-sample test statistic. This occurs because the variability of the subsample statistic is greater than that of the full-sample statistic and, hence, its behavior at a fixed value $\gamma \neq 0$ is harder to distinguish from its behavior at $\gamma = 0$. In consequence, the subsample statistic can have a distribution that is close to the asymptotic distribution for $\gamma = 0$, whereas the full-sample statistic has a distribution that is close to the asymptotic distribution for $\gamma \neq 0$. If the asymptotic distribution of the test statistic for $\gamma \neq 0$ is more disperse than for $\gamma = 0$, then the subsample critical value is too small typically and the subsample test over-rejects the null hypothesis. On the other hand, if the asymptotic distribution of the test statistic for $\gamma \neq 0$ is less disperse than for $\gamma = 0$, then the subsample critical value is too large and the subsample test is not asymptotically similar. More precisely, the limit of the finite-sample size of a subsample test depends on the whole range of behavior of the test statistic and subsample statistic for parameter values close to $\gamma = 0$.

The intuition laid out in the previous paragraph is made rigorous by considering the behavior of subsampling tests under asymptotics in which the true parameter, γ_n, drifts to the point of discontinuity $\gamma = 0$ as $n \to \infty$. Since the finite-sample size of a test is based on the supremum of the null rejection rate over all parameter values γ for given n, the limit of the finite sample size of a test is always greater than or equal to its limit under a drifting sequence $\{\gamma_n : n \geq 1\}$. Hence, if the limit of the null rejection rate under a drifting sequence exceeds the nominal level, then the limit of the exact finite-sample null rejection rate exceeds the nominal level. Analogously, if the limit of the null rejection rate under a drifting sequence is less than the nominal level, then the limit of the exact finite-sample measure of nonsimilarity must be positive.

We show that there are two different rates of drift such that over-rejection and/or under-rejection (compared to the nominal level) can occur. The first rate is one under which the full-sample test statistic has an asymptotic distribution that depends on a localization parameter, h, and the subsample critical values behave like the critical value from the asymptotic distribution of the statistic under $\gamma = 0$. The second rate is one under which the full-sample test statistic has an asymptotic distribution that is the same as for fixed $\gamma \neq 0$ and the subsample critical value behaves like the critical value from the asymptotic distribution of the full-sample statistic under a drifting sequence with localization parameter h.

The results of the paper provide conditions under which sequences of these two types determine the limit of the finite-sample size of the test. In particular, under these conditions, we obtain necessary and sufficient conditions for the limit of the finite–sample size of a subsample test to exceed its nominal level. The paper gives corresponding results for standard tests that are based on a fixed critical value.

The paper introduces a hybrid test whose critical value is the maximum of a subsample critical value and a certain fixed critical value. The hybrid test has advantages over both the subsample test and a fixed critical value test in terms of its asymptotic size. In particular, in comparison to using the subsample critical value alone, we show that taking the maximum over the fixed critical value is either irrelevant asymptotically or it reduces over-rejection asymptotically somewhere in the null hypothesis. In some scenarios, the hybrid test has correct asymptotic size. In other scenarios, it does not.

We suggest three solutions to the asymptotic over-rejection problem discussed above. One solution is a size-corrected (SC) FCV test whose asymptotic size equals the desired nominal level. The second and third solutions are SC subsample and SC hybrid tests. These solutions can be computationally challenging and may be intractable in some contexts. But, they are feasible in some contexts. In some cases, these solutions do not work at all—i.e., the assumptions under which the solutions work are violated. Examples of this considered below include a CI based on a shrinkage estimator and tests and CIs in an IV regression model where the IVs are not completely exogenous.

We show that the SC-FCV test may be uniformly more powerful than the SC-subsample test or vice versa or the two tests cannot be unambiguously ranked, depending upon the shape of the quantile function of the asymptotic distributions of the test statistics (as a function of parameters of the model). The SC hybrid test has the nice power property that it is asymptotically equivalent to the SC-FCV test in the context where the SC-FCV test is uniformly more powerful than the SC-Sub test, and it is asymptotically equivalent to the SC-Sub test in the context in which the SC-Sub test is uniformly more powerful than the SC-FCV test.

The potential problem of subsampling tests, the hybrid procedure, and the SC methods outlined above carry over with some adjustments to confidence intervals (CIs) by the usual duality between tests and CIs. Some adjustments are needed because the limit of the finite-sample level of a CI depends on uniformity over $\theta \in \Theta$ and $\gamma \in \Gamma$, where Θ and Γ are the parameter spaces of θ and γ, respectively, whereas the limit of the finite-sample size of a test of $H_0 : \theta = \theta_0$ only depends on uniformity over $\gamma \in \Gamma$ for fixed θ_0.

The paper considers general test statistics, including t, LR, and LM statistics. The results cover one-sided, symmetric two-sided, and equal-tailed two-sided t tests and corresponding confidence intervals. The t statistics may be studentized (i.e., of the form $\tau_n(\widehat{\theta}_n - \theta_0)/\widehat{\sigma}_n$ for an estimator $\widehat{\theta}_n$, a scale estimator $\widehat{\sigma}_n$, and a normalization factor τ_n) or non-studentized (i.e., of the form $\tau_n(\widehat{\theta}_n - \theta_0)$). Non-studentized t statistics are often considered in the subsample literature, see Politis, Romano, and Wolf (1999) (PRW). But, studentized t statistics are needed in certain testing situations in which non-studentized statistics have rates of convergence that are parameter dependent. This occurs with unit root tests, see Romano and Wolf (2001), and with tests in the presence of weak instruments, see Guggenberger and Wolf (2004).

The results in the paper also apply to the case where the limit distribution of a test statistic is "continuous" in a nuisance parameter. In this case, sufficient conditions are given under which FCV, subsample, and hybrid tests and CIs have asymptotic levels (defined to be the limit of their finite sample levels) equal to their nominal levels. To the best of our knowledge, results of this sort are not available in the literature—results in the literature are pointwise asymptotic results.

NEED TO ADD DETAILS REGARDING THE EXTENSION OF THE RESULTS OF THE PAPER TO COVER THE $m < n$ BOOTSTRAP BY ADJUSTING ASSUMPTION B1 AND/OR B2 BELOW.

The main results given in the paper employ high-level assumptions. These assumptions are verified in several examples.

2 Basic Testing Set-up

We are interested in tests concerning a parameter $\theta \in R^d$ in the presence of a nuisance parameter $\gamma \in \Gamma$. Of special interest is the case where $d = 1$, but the results allow for $d > 1$. The null hypothesis of interest is $H_0 : \theta = \theta_0$. The alternative hypothesis of interest may be one-sided or two-sided.

2.1 Test Statistic

Let $T_n(\theta_0)$ denote a test statistic based on a sample of size n for testing $H_0 : \theta = \theta_0$ for some $\theta_0 \in R^d$. The leading case that we consider is when $T_n(\theta_0)$ is a t statistic, but the results also allow $T_n(\theta_0)$ to be an LR, LM, or some other statistic. Large values of

$T_n(\theta_0)$ indicate evidence against the null hypothesis, so a test based on $T_n(\theta_0)$ rejects the null hypothesis when $T_n(\theta_0)$ exceeds some critical value.

When $T_n(\theta_0)$ is a t statistic, it is defined as follows. Let $\widehat{\theta}_n$ be an estimator of a scalar parameter θ based on a sample of size n. Let $\widehat{\sigma}_n$ $(\in R)$ be an estimator of the scale of $\widehat{\theta}_n$. For alternatives of the sort (i) $H_1 : \theta > \theta_0$, (ii) $H_1 : \theta < \theta_0$, and (iii) $H_1 : \theta \neq \theta_0$, respectively, the t statistic is defined as follows:

Assumption t1. (i) $T_n(\theta_0) = \tau_n(\widehat{\theta}_n - \theta_0)/\widehat{\sigma}_n$, or (ii) $T_n(\theta_0) = -\tau_n(\widehat{\theta}_n - \theta_0)/\widehat{\sigma}_n$, or (iii) $T_n(\theta_0) = |\tau_n(\widehat{\theta}_n - \theta_0)/\widehat{\sigma}_n|$, where τ_n is some known normalization constant.

In many cases, $\tau_n = n^{1/2}$. For example, this is true in boundary examples and even in the unit root example. Note that τ_n is not uniquely defined because $\widehat{\sigma}_n$ could be scaled up or down to counteract changes in the scale of τ_n. In practice this is usually not an issue because typically there is a natural definition for $\widehat{\sigma}_n$, which determines its scale.

A common case considered in the subsample literature is when $T_n(\theta_0)$ is a *non-studentized* t statistic, see PRW. In this case, Assumption t1 and the following assumption hold.

Assumption t2. $\widehat{\sigma}_n = 1$.

There are cases, however, where a non-studentized test statistic has an asymptotic null distribution with a normalization factor τ_n that depends on a nuisance parameter γ. This causes problems for the standard theory concerning subsample methods, see PRW, Ch. 8. In such cases, a studentized test statistic often has the desirable property that the normalization factor τ_n does not depend on the nuisance parameter γ. This occurs with tests concerning unit roots in time series, see Romano and Wolf (2001), and with tests in the presence of weak instruments, see Guggenberger and Wolf (2004). The set-up that we consider allows for both non-studentized and studentized test statistics. Note that under Assumption t2 the order of magnitude of τ_n is uniquely determined.

The focus of this paper is on the behavior of tests when the asymptotic null distribution of $T_n(\theta_0)$ depends on the nuisance parameter γ and is discontinuous at some value(s) of γ. Without loss of generality, we take the point(s) of discontinuity to be γ values for which some subvector of γ is 0.

2.2 Fixed Critical Values

We consider two different types of critical value for use with the test statistic $T_n(\theta_0)$. The first is a *fixed critical value* (FCV) and is denoted $c_{Fix}(1-\alpha)$, where $\alpha \in (0,1)$ is the nominal size of the FCV test. The FCV test rejects H_0 when

$$T_n(\theta_0) > c_{Fix}(1-\alpha). \tag{2.1}$$

The results below allow $c_{Fix}(1-\alpha)$ to be any constant. However, if the discontinuity (or discontinuities) of the asymptotic null distribution of $T_n(\theta_0)$ is (are) not taken into account, one typically defines

$$c_{Fix}(1-\alpha) = c_\infty(1-\alpha), \tag{2.2}$$

where $c_\infty(1-\alpha)$ denotes the $1-\alpha$ quantile of J_∞ and J_∞ is the asymptotic null distribution of $T_n(\theta_0)$ when γ is not a point of discontinuity. For example, when Assumption t1(i), (ii), or (iii) holds, $c_\infty(1-\alpha)$ typically equals $z_{1-\alpha}$, $z_{1-\alpha}$, or $z_{1-\alpha/2}$, respectively, where $z_{1-\alpha}$ denotes the $1-\alpha$ quantile of the standard normal distribution. If $T_n(\theta_0)$ is an LR, LM, or Wald statistic, then $c_\infty(1-\alpha)$ typically equals the $1-\alpha$ quantile of a χ^2_d distribution, denoted $\chi^2_d(1-\alpha)$.

On the other hand, if a discontinuity at $\gamma = h^0$ is recognized, one might take the FCV to be

$$c_{Fix}(1-\alpha) = \max\{c_\infty(1-\alpha), c_{h^0}(1-\alpha)\}, \tag{2.3}$$

where $c_{h^0}(1-\alpha)$ denotes the $1-\alpha$ quantile of J_{h^0} and J_{h^0} is the asymptotic null distribution of $T_n(\theta_0)$ when $\gamma = h^0$. The FCV test based on this FCV is not likely to be asymptotically similar, but one might hope is that it has asymptotic level α. The results given below show that often even the latter is not true.

2.3 Subsample Critical Values

The second type of critical value that we consider is a subsample critical value. Let $\{b_n : n \geq 1\}$ be a sequence of subsample sizes. For brevity, we sometimes write b_n as b. Let $\{\widehat{T}_{n,b,i} : i = 1, ..., q_n\}$ be certain subsample statistics that are based primarily on subsamples of size b_n rather than the full sample. For example, with iid observations, there are $q_n = n!/((n-b_n)!b_n!)$ different subsamples of size b_n and $\widehat{T}_{n,b,i}$ is determined primarily by the observations in the ith such subsample. With time series observations, say $\{X_1, ..., X_n\}$, there are $q_n = n - b_n + 1$ subsamples of b_n consecutive observations, e.g., $Y_i = \{X_i, ..., X_{i+b_n-1}\}$, and $\widehat{T}_{n,b,i}$ is determined primarily by the observations in the ith subsample Y_i.

Let $L_{n,b}(x)$ and $c_{n,b}(1-\alpha)$ denote the empirical distribution function and $1-\alpha$ sample quantile, respectively, of the subsample statistics $\{\widehat{T}_{n,b,i} : i = 1, ..., q_n\}$. They are defined by

$$\begin{aligned} L_{n,b}(x) &= q_n^{-1} \sum_{i=1}^{q_n} 1(\widehat{T}_{n,b,i} \leq x) \text{ for } x \in R \text{ and} \\ c_{n,b}(1-\alpha) &= \inf\{x : L_{n,b}(x) \geq 1-\alpha\}. \end{aligned} \tag{2.4}$$

The subsample test rejects $H_0 : \theta = \theta_0$ if

$$T_n(\theta_0) > c_{n,b}(1-\alpha). \tag{2.5}$$

We now describe the subsample statistics $\{\widehat{T}_{n,b,i} : i = 1, ..., q_n\}$ in more detail. Let $\{T_{n,b,i}(\theta_0) : i = 1, ..., q_n\}$ be subsample statistics that are defined exactly as $T_n(\theta_0)$ is defined, but based on subsamples of size b_n rather than the full sample. For example, suppose Assumption t1 holds. Let $(\widehat{\theta}_{n,b,i}, \widehat{\sigma}_{n,b,i})$ denote the estimators $(\widehat{\theta}_b, \widehat{\sigma}_b)$ applied to the ith subsample. In this case, we have

$$\begin{aligned}
\text{(i) } T_{n,b,i}(\theta_0) &= \tau_b(\widehat{\theta}_{n,b,i} - \theta_0)/\widehat{\sigma}_{n,b,i}, \text{ or} \\
\text{(ii) } T_{n,b,i}(\theta_0) &= -\tau_b(\widehat{\theta}_{n,b,i} - \theta_0)/\widehat{\sigma}_{n,b,i}, \text{ or} \\
\text{(iii) } T_{n,b,i}(\theta_0) &= |\tau_b(\widehat{\theta}_{n,b,i} - \theta_0)/\widehat{\sigma}_{n,b,i}|. \qquad (2.6)
\end{aligned}$$

Below we make use of the empirical distribution of $\{T_{n,b,i}(\theta_0) : i = 1, ..., q_n\}$ defined by

$$U_{n,b}(x) = q_n^{-1} \sum_{i=1}^{q_n} 1(T_{n,b,i}(\theta_0) \leq x). \qquad (2.7)$$

In most cases, subsample critical values are based on a simple adjustment to the statistics $\{T_{n,b,i}(\theta_0) : i = 1, ..., q_n\}$, where the adjustment is designed to yield subsample statistics that behave similarly under the null and the alternative hypotheses. In particular, $\{\widehat{T}_{n,b,i} : i = 1, ..., q_n\}$ often are defined to satisfy the following condition.

Assumption Sub1. $\widehat{T}_{n,b_n,i} = T_{n,b_n,i}(\widehat{\theta}_n)$ for all $i \leq n$, where $\widehat{\theta}_n$ is an estimator of θ.

The estimator $\widehat{\theta}_n$ is usually chosen to be a consistent estimator of θ whether or not the null hypothesis holds. Assumption Sub1 can be applied to t statistics as well as to LR and LM statistics, among others.

If consistent estimation of θ is not possible when $\gamma = 0$, as occurs when θ is not identified when $\gamma = 0$, then taking $\{\widehat{T}_{n,b_n,i}\}$ to satisfy Assumption Sub1 is not desirable because $\widehat{\theta}_n$ is not necessarily close to θ_0 when γ is close to 0. For example, this occurs in the weak IV example, see Guggenberger and Wolf (2004). In such cases, it is preferable to take $\{\widehat{T}_{n,b_n,i}\}$ to satisfy the following assumption.[2]

Assumption Sub2. $\widehat{T}_{n,b_n,i} = T_{n,b_n,i}(\theta_0)$ for all $i \leq n$.

The results given below for subsample tests allow for subsample statistics $\{\widehat{T}_{n,b,i}\}$ that satisfy Assumption Sub1 or Sub2 or are defined in some other way.

2.4 Asymptotic Size

The exact size, $ExSz_n(\theta_0)$, of an FCV or subsample test is the supremum over $\gamma \in \Gamma$ of the null rejection probability under γ:

$$\begin{aligned}
ExSz_n(\theta_0) &= \sup_{\gamma \in \Gamma} RP_n(\theta_0, \gamma), \text{ where } RP_n(\theta_0, \gamma) = P_{\theta_0,\gamma}(T_n(\theta_0) > c_{1-\alpha}), \\
c_{1-\alpha} &= c_{Fix}(1-\alpha) \text{ or } c_{1-\alpha} = c_{n,b}(1-\alpha), \qquad (2.8)
\end{aligned}$$

and $P_{\theta,\gamma}(\cdot)$ denotes probability when the true parameters are (θ, γ).[3]

[2] When Assumption t1 holds, subsample statistics $\{\widehat{T}_{n,b_n,i}\}$ that satisfy Assumption Sub2 typically yield nontrivial power because the normalization constant τ_n satisfies $\tau_{b_n}/\tau_n \to 0$.

[3] We remind the reader that the *size* of a test is equal to the supremum of its rejection probability under the null hypothesis and a test is of *level* α if its size is less than or equal to α.

We are interested in the "asymptotic size" of the test defined by

$$AsySz(\theta_0) = \limsup_{n\to\infty} ExSz_n(\theta_0). \tag{2.9}$$

This definition should not be controversial. Our interest is in the exact finite sample size of the test. We use asymptotics to approximate this. Uniformity over $\gamma \in \Gamma$, which is built into the definition of $AsySz(\theta_0)$, is necessary for the asymptotic size to give a good approximation to the finite sample size.[4]

If $AsySz(\theta_0) > \alpha$, then the nominal level α test has asymptotic size greater than α and the test does not have correct asymptotic level.

To a lesser extent, we are also interested in the minimum rejection probability of the subsample test and its limit:

$$MinRP_n(\theta_0) = \inf_{\gamma\in\Gamma} RP_n(\theta_0,\gamma) \text{ and } AsyMinRP(\theta_0) = \liminf_{n\to\infty} MinRP_n(\theta_0). \tag{2.10}$$

The quantity $\alpha - MinRP_n(\theta_0)$ is the maximum amount of under-rejection of the test over points in the null hypothesis for fixed n. If $\alpha - AsyMinRP(\theta_0) > 0$, then the subsample test is not asymptotically similar and, hence, may sacrifice power.

3 Assumptions

This section introduces the assumptions that we employ. The assumptions are verified in several examples below.

3.1 Parameter Space

First, we introduce some notation. Let $\lfloor$ denote the left endpoint of an interval that may be open or closed at the left end. Define $\rfloor$ analogously for the right endpoint. Let $R_+ = \{x \in R : x \geq 0\}$, $R_- = \{x \in R : x \leq 0\}$, $R_{+,\infty} = R_+ \cup \{\infty\}$, $R_{-,\infty} = R_- \cup \{-\infty\}$, $R_\infty = R \cup \{\pm\infty\}$, $R_+^p = R_+ \times ... \times R_+$ (with p copies), and $R_\infty^p = R_\infty \times ... \times R_\infty$ (with p copies). Let $\mathrm{cl}(I)$ denote the closure of an interval $I = \lfloor I_1, I_2 \rfloor \subset R$ with respect to R_∞. (In particular, if $I = \lfloor I_1, \infty)$ for $I_1 > -\infty$, then $\mathrm{cl}(I) = [I_1,\infty) \cup \{\infty\}$; if $I = (-\infty, I_2\rfloor$ for $I_2 < \infty$, then $\mathrm{cl}(I) = (-\infty, I_2] \cup \{-\infty\}$; and if $I = R$, then $\mathrm{cl}(I) = R_\infty$.)

The model is indexed by a parameter γ that has up to three components: $\gamma = (\gamma_1, \gamma_2, \gamma_3)$. The points of discontinuity of the asymptotic distribution of the test statistic of interest are determined by the first component, $\gamma_1 \in R^p$. Through reparametrization we can assume without loss of generality that the discontinuity occurs at $\gamma_1 = 0$. Thus, γ_1 determines how close the parameter γ is to a point of discontinuity. The value of γ_1 affects the limit distribution of the test statistic of interest. The parameter space for γ_1 is $\Gamma_1 \subset R^p$.

[4] Note that the definition of the parameter space Γ is flexible. In some cases, one might want to define Γ so as to bound γ away from points in R^p that are troublesome. This is reasonable, of course, only if one has prior information that justifies the particular definition of Γ.

The second component, γ_2 ($\in R^q$), of γ also affects the limit distribution of the test statistic, but does not affect the distance of the parameter γ to the point of discontinuity. The parameter space for γ_2 is $\Gamma_2 \subset R^q$.

The third component, γ_3, of γ is assumed to be an element of an arbitrary space $\mathcal{T}_3$. Hence, it may be finite or infinite dimensional. By assumption, the third component γ_3 does not affect the limit distribution of the test statistic (which is why no structure on the space $\mathcal{T}_3$ is required). For example, in a linear model, a test statistic concerning one regression parameter may be invariant to the value of some other regression parameters. The latter parameters are then part of γ_3. Infinite dimensional γ_3 parameters also arise frequently. For example, error distributions are often part of γ_3. Due to the operation of the central limit theorem it is often the case that the asymptotic distribution of a test statistic does not depend on the particular error distribution—only on whether the error distribution has certain moments finite. Such error distributions are part of γ_3. The parameter space for γ_3 is $\Gamma_3(\gamma_1, \gamma_2)$ ($\subset \mathcal{T}_3$), which as indicated may depend on γ_1 and γ_2.

The parameter space for γ is

$$\Gamma = \{(\gamma_1, \gamma_2, \gamma_3) : \gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2, \gamma_3 \in \Gamma_3(\gamma_1, \gamma_2)\}. \tag{3.1}$$

In Section 4 below we provide two main theorems. The first theorem relies on weaker assumptions than the second, but gives weaker results. We label the assumptions to correspond to these two theorems. An Assumption that ends with 1 is used in Theorem 1. An Assumption that ends in 2 is used in Theorem 2 and is stronger than a corresponding assumption that ends in 1. (For example, Assumption A2 implies Assumption A1.) All other assumptions are used in both Theorems 1 and 2.

Assumption A1. **(i)** Γ satisfies (3.1), where $\Gamma_1 \subset R^p$, $\Gamma_2 \subset R^q$, and $\Gamma_3(\gamma_1, \gamma_2) \subset \mathcal{T}_3$ for some arbitrary space $\mathcal{T}_3$. **(ii)** The zero p-vector, 0, is in Γ_1.

Assumption A2. **(i)** Assumption A1(i) holds. **(ii)** $\Gamma_1 = \prod_{j=1}^{p} \Gamma_{1,j}$, where $\Gamma_{1,j} = \lfloor a_j, b_j \rfloor$ for some $-\infty \leq a_j < b_j \leq \infty$ that satisfy $a_j \leq 0 \leq b_j$ for $j = 1, ..., p$.

Under Assumption A2, the parameter space Γ_1 includes the point $\gamma_1 = 0$ as well as values γ_1 that are arbitrarily close to 0.[5]

Next, we define an index set for the different asymptotic null distributions of the test statistic $T_n(\theta_0)$ of interest. Define

$$H = H_1 \times H_2,\ H_1 = \prod_{j=1}^{p} \begin{cases} R_{+,\infty} & \text{if } a_j = 0 \\ R_{-,\infty} & \text{if } b_j = 0 \\ R_{\infty} & \text{if } a_j < 0 \text{ and } b_j > 0, \end{cases} \quad \text{and } H_2 = \text{cl}(\Gamma_2), \tag{3.2}$$

where $\text{cl}(\Gamma_2)$ is the closure of Γ_2 with respect to R^q_∞. For example, if $p = 1$, $a_1 = 0$, and $\Gamma_2 = R^q$, then $H_1 = R_{+,\infty}$, $H_2 = R^q_\infty$, and $H = R_{+,\infty} \times R^q_\infty$.

[5] The results below allow for the case where there is no subvector γ_1 of γ, i.e., $p = 0$. In this case, there is no discontinuity of the asymptotic distribution of the test statistic of interest, see below.

3.2 Convergence Assumptions

This subsection and the next introduce the high-level assumptions that we employ. The high-level assumptions are verified in several examples below.

Throughout this section, the true value of θ is the null value θ_0 and all limits are as $n \to \infty$. For an arbitrary distribution G, let $G(\cdot)$ denote the distribution function (df) of G and let $C(G)$ denote the continuity points of $G(\cdot)$. Define the $1-\alpha$ quantile, $q(1-\alpha)$, of a distribution G by $q(1-\alpha) = \inf\{x : G(x) \geq 1-\alpha\}$. For a df $G(\cdot)$, let $G(x-) = \lim_{\varepsilon \searrow 0} G(x-\varepsilon)$, where "$\lim_{\varepsilon \searrow 0}$" denotes the limit as $\varepsilon > 0$ declines to zero. Note that $G(x+) = \lim_{\varepsilon \searrow 0} G(x+\varepsilon)$ equals $G(x)$ because dfs are right continuous. The distributions J_h and J_{h^0} considered below are distributions of proper random variables that are finite with probability one.

For a sequence of constants $\{\kappa_n : n \geq 1\}$, let $\kappa_n \to [\kappa_{1,\infty}, \kappa_{2,\infty}]$ denote that $\kappa_{1,\infty} \leq \liminf_{n\to\infty} \kappa_n \leq \limsup_{n\to\infty} \kappa_n \leq \kappa_{2,\infty}$.

Let $r > 0$ denote a *rate of convergence index* such that when the true value of γ_1 satisfies $n^r \gamma_1 \to h_1$, then the test statistic $T_n(\theta_0)$ has an asymptotic distribution that depends on the localization parameter h_1. In most examples, $r = 1/2$.

Definition of $\{\gamma_{n,h} : n \geq 1\}$**:** Given $r > 0$ and $h = (h_1, h_2) \in H$, let $\{\gamma_{n,h} = (\gamma_{n,h,1}, \gamma_{n,h,2}, \gamma_{n,h,3}) : n \geq 1\}$ denote a sequence of parameters in Γ for which $n^r \gamma_{n,h,1} \to h_1$, $\gamma_{n,h,2} \to h_2$, and $\gamma_{n,h,3} \in \Gamma_3(\gamma_{n,h,1}, \gamma_{n,h,2})$ for all $n \geq 1$.

The sequence $\{\gamma_{n,h} : n \geq 1\}$ is defined such that under $\{\gamma_{n,h} : n \geq 1\}$, the asymptotic distribution of $T_n(\theta_0)$ depends on h and only h, see Assumptions B1 and B2 below. For a given model, there is a single fixed $r > 0$. Hence, for notational simplicity, we do not index $\{\gamma_{n,h} : n \geq 1\}$ by r. In addition, the limit distributions under $\{\gamma_{n,h} : n \geq 1\}$ of the test statistics of interest do not depend on $\gamma_{n,h,3}$, so we do not make the dependence of $\gamma_{n,h}$ on $\gamma_{n,h,3}$ explicit.

In models in which the asymptotic distribution of the test statistic of interest is continuous in the model parameters, we apply the results below with no parameter γ_1 (or $\gamma_{n,h,1}$), i.e., $p = 0$. We refer to this as the *continuous limit* case. On the other hand, in the *discontinuous limit* case—which is the case of main interest in this paper, we apply the results with $p \geq 1$.

Given any $h = (h_1, h_2) \in H$, define $h^0 = (0, h_2) \in H$, where $h_2 \in H_2$.

We use the following assumptions.

Assumption B1. (i) For some $r > 0$, some $h \in R^p$, some sequence $\{\gamma_{n,h} : n \geq 1\}$, and some distribution J_h, $T_n(\theta_0) \to_d J_h$ under $\{\gamma_{n,h} : n \geq 1\}$, and (ii) for all sequences $\{\gamma_{n,h^0} : n \geq 1\}$ and some distribution J_{h^0}, $T_n(\theta_0) \to_d J_{h^0}$ under $\{\gamma_{n,h^0} : n \geq 1\}$ (where r is the same as in part (i) and h^0 depends on the vector h given in part (i)).

Assumption B2. For some $r > 0$, all $h \in H$, all sequences $\{\gamma_{n,h} : n \geq 1\}$, and some distributions J_h, $T_n(\theta_0) \to_d J_h$ under $\{\gamma_{n,h} : n \geq 1\}$.

If $\gamma_{n,h}$ does not depend on n (which necessarily requires $h_1 = 0$), Assumption B1(i) is a standard assumption in the subsampling literature. For example, it is imposed in the basic theorem in PRW (1999, Thm. 2.2.1, p. 43) for subsampling with iid observations and in their theorem for stationary strong mixing observations, see PRW, Thm. 3.2.1, p. 70. If $\gamma_{n,h}$ does depend on n, Assumption B1(i) usually can be verified using the same sort of argument as when it does not. Similarly, Assumption B1(ii) usually can be verified using the same sort of argument and, hence, is not restrictive.

Assumption B2 holds in many examples, but it can be restrictive. It is for this reason that we introduce Assumption B1. Theorem 1 only requires Assumption B1, whereas Theorem 2 requires Assumption B2. In the "continuous limit" case (where Assumption B2 holds with $p = 0$ and $H = H_2$), the asymptotic distribution J_h may depend on h but is continuous in the sense that one obtains the same asymptotic distribution for any sequence $\{\gamma_{n,h} : n \geq 1\}$ for which $\gamma_{n,h,2}$ converges to $h_2 \in H_2$.

3.3 Subsample Assumptions

The assumptions above are all that are needed for FCV tests. For subsample tests, we require the following additional assumptions:

Assumption C. (i) $b_n \to \infty$, and (ii) $b_n/n \to 0$.

Assumption D. (i) $\{T_{n,b_n,i}(\theta_0) : i = 1, ..., q_n\}$ are identically distributed under any $\gamma \in \Gamma$ for all $n \geq 1$, and (ii) $T_{n,b_n,i}(\theta_0)$ and $T_{b_n}(\theta_0)$ have the same distribution under any $\gamma \in \Gamma$ for all $n \geq 1$.

Assumption E. For all sequences $\{\gamma_n \in \Gamma : n \geq 1\}$, $U_{n,b_n}(x) - E_{\theta_0,\gamma_n} U_{n,b_n}(x) \to_p 0$ under $\{\gamma_n : n \geq 1\}$ for all $x \in R$.

Assumption F1. For all $\varepsilon > 0$, $J_{h^0}(c_{h^0}(1-\alpha) - \varepsilon) < 1 - \alpha$ and $J_{h^0}(c_{h^0}(1-\alpha) + \varepsilon) > 1 - \alpha$, where $c_{h^0}(1-\alpha)$ is the $1-\alpha$ quantile of J_{h^0} and h^0 is as in Assumption B1(ii).

Assumption F2. For all $\varepsilon > 0$ and $h \in H$, $J_h(c_h(1-\alpha) - \varepsilon) < 1 - \alpha$ and $J_h(c_h(1-\alpha) + \varepsilon) > 1 - \alpha$, where $c_h(1-\alpha)$ is the $1-\alpha$ quantile of J_h.

Assumption G1. For the sequence $\{\gamma_{n,h} : n \geq 1\}$ in Assumption B1(i), $L_{n,b_n}(x) - U_{n,b_n}(x) \to_p 0$ for all $x \in C(J_{h^0})$ under $\{\gamma_{n,h} : n \geq 1\}$.

Assumption G2. For all $h \in H$ and all sequences $\{\gamma_{n,h} : n \geq 1\}$, if $U_{n,b_n}(x) \to_p J_g(x)$ under $\{\gamma_{n,h} : n \geq 1\}$ for all $x \in C(J_g)$ for some $g \in R_\infty^p$, then $L_{n,b_n}(x) - U_{n,b_n}(x) \to_p 0$ under $\{\gamma_{n,h} : n \geq 1\}$ for all $x \in C(J_g)$.

Assumptions C and D are standard assumptions in the subsample literature, e.g., see PRW, Thm. 2.2.1, p. 43, and are not restrictive. The sequence $\{b_n : n \geq 1\}$ can always be chosen to satisfy Assumption C. Assumption D necessarily holds when the observations are iid or stationary and subsamples are constructed in the usual way (described above).

Assumption E holds quite generally. For iid observations, the condition in Assumption E when $\gamma_{n,1} = 0$ and $(\gamma_{n,2}, \gamma_{n,3})$ does not depend on n (where $\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3})$) is verified by PRW (1999, p. 44) using a U-statistic inequality of Hoeffding. It holds for any triangular array of row-wise iid [0,1]-valued random variables by the same argument. Hence, Assumption E holds automatically when the observations are iid for each fixed $\gamma \in \Gamma$ when the subsample statistics are defined as above.

For stationary strong mixing observations, the condition in Assumption E when $\gamma_{n,1} = 0$ and $(\gamma_{n,2}, \gamma_{n,3})$ does not depend on n (where $\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3})$) is verified by PRW (1999, pp. 71-72) by establishing L^2 convergence using a strong mixing covariance bound. It holds for any sequence $\{\gamma_n \in \Gamma : n \geq 1\}$ and, hence, Assumption E holds, by the same argument as in PRW provided

$$\sup_{\gamma \in \Gamma} \alpha_\gamma(m) \to 0 \text{ as } m \to \infty, \tag{3.3}$$

where $\{\alpha_\gamma(m) : m \geq 1\}$ are the strong mixing numbers of the observations when the true parameters are (θ_0, γ).

Assumptions F1 and F2 are not restrictive. They hold in all of the examples that we have considered. In particular, Assumption F1 holds if either (i) $J_{h^0}(x)$ is strictly increasing at $x = c_{h^0}(1-\alpha)$ or (ii) $J_{h^0}(x)$ has a jump at $x = c_{h^0}(1-\alpha)$ with $J_{h^0}(c_{h^0}(1-\alpha)) > 1-\alpha$ and $J_{h^0}(c_{h^0}(1-\alpha)-) < 1-\alpha$. Sufficient conditions for Assumption F2 are analogous. Condition (i) holds in most examples. But, if J_{h^0} is a pointmass, as occurs with the example of a CI based on a super-efficient estimator with constant $a = 0$ (see Section 11.2 below), then condition (i) fails, but condition (ii) holds.

Assumptions G1 and G2 hold automatically when $\{\widehat{T}_{n,b_n,i}\}$ satisfy Assumption Sub2.

To verify that Assumption G1 or G2 holds when $\{\widehat{T}_{n,b_n,i}\}$ satisfy Assumption Sub1 and $T_n(\theta_0)$ is a non-studentized t statistic (i.e., Assumptions t1 and t2 hold), we use the following assumption.

Assumption H. $\tau_{b_n}/\tau_n \to 0$.

This is a standard assumption in the subsample literature, e.g., see PRW, Thm. 2.2.1, p. 43. In the leading case where $\tau_n = n^s$ for some $s > 0$, Assumption H follows from Assumption C(ii) because $\tau_{b_n}/\tau_n = (b_n/n)^s \to 0$.

Lemma 1 (a) *Assumptions* B1(i), t1, t2, Sub1, *and* H *imply Assumption* G1.

(b) *Assumptions* B2, t1, t2, Sub1, *and* H *imply Assumption* G2.

Comment. Lemma 1 is a special case of Lemma 5, which is stated in Section 10 for expositional convenience and is proved in the Appendix. Lemma 5 does not impose Assumption t2 and, hence, covers studentized t statistics.

The assumptions above have been designed to avoid the requirement that $J_h(x)$ is continuous in x because this assumption is violated in some applications, such as boundary problems, for some values of h and some values of x.

4 Asymptotic Results

The first result of this section concerns the asymptotic null behavior of FCV and subsample tests under a single sequence $\{\gamma_{n,h} : n \geq 1\}$.

Theorem 1 *Let* $\alpha \in (0,1)$ *be given.*

(a) *Suppose Assumption* B1(i) *holds. Then,*

$$P_{\theta_0,\gamma_{n,h}}(T_n(\theta_0) > c_{Fix}(1-\alpha)) \to [1 - J_h(c_{Fix}(1-\alpha)), 1 - J_h(c_{Fix}(1-\alpha)-)].$$

(b) *Suppose Assumptions* A1, B1, C-E, F1, *and* G1 *hold. Then,*

$$P_{\theta_0,\gamma_{n,h}}(T_n(\theta_0) > c_{n,b}(1-\alpha)) \to [1 - J_h(c_{h^0}(1-\alpha)), 1 - J_{h^0}(c_{h^0}(1-\alpha)-)].$$

Comments. 1. If $1 - J_h(c_{Fix}(1-\alpha)) > \alpha$, then the FCV test has $AsySz(\theta_0) > \alpha$, i.e., its asymptotic size exceeds its nominal level α.

2. Analogously, for the subsample test, if $1 - J_h(c_{h^0}(1-\alpha)) > \alpha$, then the test has $AsySz(\theta_0) > \alpha$.

3. If $1 - J_h(c_{Fix}(1-\alpha)-) < \alpha$, then the FCV test has $AsyMinRP(\theta_0) < \alpha$ and it is not asymptotically similar. Analogously, if $1 - J_h(c_{h^0}(1-\alpha)-) < \alpha$, then the subsample test has $AsyMinRP(\theta_0) < \alpha$ and it is not asymptotically similar.

4. If $J_h(x)$ is continuous at $x = c_{Fix}(1-\alpha)$ (which typically holds in applications for most values h but not necessarily all), then the result of Theorem 1(a) becomes $P_{\theta_0,\gamma_{n,h}}(T_n(\theta_0) > c_{Fix}(1-\alpha)) \to 1 - J_h(c_{Fix}(1-\alpha))$. Analogously, if $J_h(x)$ is continuous at $x = c_{h^0}(1-\alpha)$, then the result of Theorem 1(b) becomes $P_{\theta_0,\gamma_{n,h}}(T_n(\theta_0) > c_{n,b}(1-\alpha)) \to 1 - J_h(c_{h^0}(1-\alpha))$.

5. In the "continuous limit" case, $h^0 = h$ because no parameter γ_1 appears. Hence, the result of Theorem 1(b) for the subsample test is $P_{\theta_0,\gamma_{n,h}}(T_n(\theta_0) > c_{n,b}(1-\alpha)) \to \alpha$, provided $J_h(x)$ is continuous at $x = c_h(1-\alpha)$. That is, the pointwise asymptotic null rejection rate is the desired nominal rate α.

6. Typically Assumption B1(i) holds for an infinite number of values h, say $h \in H^*$ ($\subset R^p$). In this case, Comments 1-5 apply for all $h \in H^*$.

We now use the stronger Assumptions A2, B2, F2, and G2 to establish more precisely the asymptotic sizes and asymptotic minimum rejection probabilities of sequences of FCV and subsample tests. For FCV tests, we define

$$\begin{aligned} Max_{Fix}(\alpha) &= \sup_{h \in H}[1 - J_h(c_{Fix}(1-\alpha))] \text{ and} \\ Max^-_{Fix}(\alpha) &= \sup_{h \in H}[1 - J_h(c_{Fix}(1-\alpha)-)]. \end{aligned} \tag{4.1}$$

Define $Min_{Fix}(\alpha)$ and $Min^-_{Fix}(\alpha)$ analogously with "inf" in place of "sup."

For subsample tests, define

$$GH = \{(g,h) \in H \times H : g = (g_1, g_2),\ h = (h_1, h_2),\ g_2 = h_2 \text{ and for } j = 1, ..., p,$$
$$\text{(i) } g_{1,j} = 0 \text{ if } |h_{1,j}| < \infty,\ \text{(ii) } g_{1,j} \in R_{+,\infty} \text{ if } h_{1,j} = +\infty, \text{ and}$$
$$\text{(iii) } g_{1,j} \in R_{-,\infty} \text{ if } h_{1,j} = -\infty\}, \tag{4.2}$$

where $g_1 = (g_{1,1}, ..., g_{1,p})' \in H_1$ and $h_1 = (h_{1,1}, ..., h_{1,p})' \in H_1$. Note that for $(g,h) \in GH$, we have $|g_{1,j}| \leq |h_{1,j}|$ for all $j = 1, ..., p$. In the "continuous limit" case (where there is no γ_1 component of γ) GH simplifies considerably: $GH = \{(g_2, h_2) \in H_2 \times H_2 : g_2 = h_2\}$.

Define

$$Max_{Sub}(\alpha) = \sup_{(g,h) \in GH} [1 - J_h(c_g(1-\alpha))] \text{ and}$$
$$Max_{Sub}^{-}(\alpha) = \sup_{(g,h) \in GH} [1 - J_h(c_g(1-\alpha)-)]. \tag{4.3}$$

Define $Min_{Sub}(\alpha)$ and $Min_{Sub}^{-}(\alpha)$ analogously with "inf" in place of "sup." In the "continuous limit" case, $Max_{Sub}(\alpha)$ simplifies to $\sup_{h \in H}[1 - J_h(c_h(1-\alpha))]$, which is less than or equal to α by the definition of $c_h(1-\alpha)$.

Theorem 2 *Let* $\alpha \in (0,1)$ *be given.*

(a) *Suppose Assumptions* A2 *and* B2 *hold. Then, an FCV test satisfies*

$$AsySz(\theta_0) \in [Max_{Fix}(\alpha), Max_{Fix}^{-}(\alpha)] \textit{ and}$$
$$AsyMinRP(\theta_0) \in [Min_{Fix}(\alpha), Min_{Fix}^{-}(\alpha)].$$

(b) *Suppose Assumptions* A2, B2, C-E, F2, *and* G2 *hold. Then, a subsample test satisfies*

$$AsySz(\theta_0) \in [Max_{Sub}(\alpha), Max_{Sub}^{-}(\alpha)] \textit{ and}$$
$$AsyMinRP(\theta_0) \in [Min_{Sub}(\alpha), Min_{Sub}^{-}(\alpha)].$$

Comments. 1. If $J_h(x)$ is continuous at the appropriate value(s) of x, then $Max_{Fix}(\alpha) = Max_{Fix}^{-}(\alpha)$ and $Max_{Sub}(\alpha) = Max_{Sub}^{-}(\alpha)$ and Theorem 2 gives the precise value of $AsySz(\theta_0)$ and analogously for $AsyMinRP(\theta_0)$. Even for FCV tests, we are not aware of general results in the literature that establish the limit of the finite sample size of tests. REFERENCES WOULD BE WELCOME. WE SUSPECT THAT THERE MUST BE SOME RESULTS OF THIS SORT IN THE LITERATURE FOR FCV TESTS.

2. Given Theorem 2(b) and the definition of $Max_{Sub}^{-}(\alpha)$, sufficient conditions for a nominal level α subsample test to have asymptotic level α are the following: (a) $c_g(1-\alpha) \geq c_h(1-\alpha)$ for all $(g,h) \in GH$ and (b) $Max_{Sub}^{-}(\alpha) = Max_{Sub}(\alpha)$.[6] Condition

[6] Under these conditions, $Max_{Sub}^{-}(\alpha) = Max_{Sub}(\alpha) = \sup_{(g,h) \in GH}[1 - J_h(c_g(1-\alpha))] \leq \sup_{h \in H}[1 - J_h(c_h(1-\alpha))] \leq \alpha$.

(a) necessarily holds in "continuous limit" examples and it holds in some "discontinuous limit" examples. But, it often fails in "discontinuous limit" examples. Condition (b) holds in most examples.

3. The same argument as used to prove Theorem 2 can be used to prove slightly stronger results than those of Theorem 2. Namely, for FCV and subsample tests,

$$ExSz_n(\theta_0) \to [Max_{Type}(\alpha), Max^-_{Type}(\alpha)] \tag{4.4}$$

for $Type = Fix$ and Sub, respectively. (These results are stronger because they imply that $\liminf_{n\to\infty} ExSz_n(\theta_0) \geq Max_{Type}(\alpha)$, rather than just $\limsup_{n\to\infty} ExSz_n(\theta_0) \geq Max_{Type}(\alpha)$.) Hence, when $Max_{Type}(\alpha) = Max^-_{Type}(\alpha)$, we have

$$\lim_{n\to\infty} ExSz_n(\theta_0) = Max_{Type}(\alpha) \text{ for } Type = Fix \text{ and } Sub. \tag{4.5}$$

4. Theorem 2 does not provide asymptotic confidence levels of CIs that are obtained by inverting FCV or subsample tests. The reason is that the level (and asymptotic level) of a CI depends on uniformity of the coverage probability over nuisance parameters *and* over the parameter of interest θ, whereas the level (and asymptotic level) of a test concerning θ only depends on uniformity of the null rejection rate over nuisance parameters (because the null value of the parameter of interest, θ, is fixed). See Section 9 for analogous results for the asymptotic levels of CIs.

5 Hybrid Tests

In this section, we define a hybrid test that is useful when a test statistic has a limit distribution that is discontinuous in some parameter and a FCV or subsample test over-rejects asymptotically under the null hypothesis. The critical value of the hybrid test is the maximum of the subsample critical value and a certain fixed critical value. The hybrid test is quite simple to compute, in some situations has asymptotic size equal to its nominal level α, and in other situations over-rejects the null asymptotically less than either the standard subsample test or a certain fixed critical value test. In addition, at least in some scenarios, the power of the hybrid test is quite good relative to FCV and subsample tests, see Section 7 below.

We suppose the following assumption holds.

Assumption J. The asymptotic distribution J_h in Assumption B2 is the same distribution, call it J_∞, for all $h = (h_1, h_2)' \in H$ for which $h_{1,j} = \pm\infty$ for all $j = 1, ..., p$, where $h_1 = (h_{1,1}, ..., h_{1,p})'$.

In examples, Assumption J often holds when $T_n(\theta_0)$ is a studentized statistic (i.e., Assumption t1 holds, but t2 does not) or an LM or LR statistic. In such cases, J_∞ typically is a standard normal, absolute standard normal, or chi-square distribution. Let $c_\infty(1-\alpha)$ denote the $1-\alpha$ quantile of J_∞.

The hybrid test with nominal level α rejects the null hypothesis $H_0 : \theta = \theta_0$ when

$$T_n(\theta_0) > c^*_{n,b}(1-\alpha), \text{ where}$$
$$c^*_{n,b}(1-\alpha) = \max\{c_{n,b}(1-\alpha), c_\infty(1-\alpha)\}. \tag{5.1}$$

The hybrid test simply takes the critical value to be the maximum of the usual subsample critical value and the critical value from the J_∞ distribution, which is usually known.[7] Hence, it is straightforward to compute. Obviously, the rejection probability of the hybrid test is less than or equal to those of the standard subsample test and the FCV test with $c_{Fix}(1-\alpha) = c_\infty(1-\alpha)$. Hence, the hybrid test over-rejects less often than either of these two tests. Furthermore, it is shown in Lemma 3 below that the hybrid test of nominal level α has asymptotic level α (i.e., $AsySz(\theta_0) \leq \alpha$) provided the quantile function $c_g(1-\alpha)$ is maximized at a boundary point. For example, this occurs if $c_g(1-\alpha)$ is monotone increasing or decreasing in g_1 for fixed g_2, where $g = (g_1, g_2)$ (i.e., $c_{(g_1,g_2)}(1-\alpha) \leq c_{(g_1^*,g_2)}(1-\alpha)$ when $g_1 \leq g_1^*$ element by element or $c_{(g_1,g_2)}(1-\alpha) \geq c_{(g_1^*,g_2)}(1-\alpha)$ when $g_1 \leq g_2$ element by element).

Define

$$Max^-_{Hyb}(\alpha) = \sup_{(g,h)\in GH} [1 - J_h(\max\{c_g(1-\alpha), c_\infty(1-\alpha)\}-)]. \tag{5.2}$$

Define $Max_{Hyb}(\alpha)$ analogously, but without "$-$" at the end of the expression.

The following Corollary to Theorems 1(b) and 2(b) establishes the asymptotic size of the hybrid test.

Corollary 1 *Let $\alpha \in (0,1)$ be given.*

(a) *Suppose Assumptions* A1, B1, C-E, F1, G1, *and* J *hold. Then,*

$$P_{\theta_0,\gamma_{n,h}}(T_n(\theta_0) > c^*_{n,b}(1-\alpha))$$
$$\to [1 - J_h(\max\{c_{h^0}(1-\alpha), c_\infty(1-\alpha)\}), 1 - J_{h^0}(\max\{c_{h^0}(1-\alpha), c_\infty(1-\alpha)\}-)].$$

(b) *Suppose Assumptions* A2, B2, C-E, F2, G2, *and* J *hold. Then, the hybrid test based on $T_n(\theta_0)$ has $AsySz(\theta_0) \in [Max_{Hyb}(\alpha), Max^-_{Hyb}(\alpha)]$.*

Comments. 1. If $1 - J_h(\max\{c_{h^0}(1-\alpha), c_\infty(1-\alpha)\}) > \alpha$, then the hybrid test has $AsySz(\theta_0) > \alpha$.

2. Assumption J is not actually needed for the results of Corollary 1 to hold—in the definition of $c^*_{n,b}(1-\alpha)$), $c_\infty(1-\alpha)$ could be any constant. Assumption J is just used to motivate the particular choice of $c_\infty(1-\alpha)$ given above, as the $1-\alpha$ quantile

[7] Hybrid tests can be defined even when Assumption J does not hold. For example, we can define $c^*_{n,b}(1-\alpha) = \max\{c_{n,b}(1-\alpha), \sup_{h\in H} c_{h_\infty}(1-\alpha)\}$, where $c_{h_\infty}(1-\alpha)$ is the $1-\alpha$ quantile of J_{h_∞} and, given $h \in H$, $h_\infty = (h_{\infty,1,1}, ..., h_{\infty,1,p}, h_{\infty,2}) \in H$ is defined by $h_{\infty,1,j} = +\infty$ if $h_{1,j} > 0$, $h_{\infty,1,j} = -\infty$ if $h_{1,j} < 0$, $h_{\infty,1,j} = +\infty$ or $-\infty$ (chosen so that $h_\infty \in H$) if $h_{1,j} = 0$ for $j = 1, ..., p$, and $h_{\infty,2} = h_2$. When Assumption J holds, this reduces to the hybrid critical value in (5.1). We utilize Assumption J because it leads to a particularly simple form for the hybrid test.

of J_∞. Assumption J also is used in results given below concerning the properties of the hybrid test.

3. Corollary 1 holds by the proof of Theorems 1(b) and 2(b) with $c_{n,b}(1-\alpha)$ replaced by $\max\{c_{n,b}(1-\alpha), c_\infty(1-\alpha)\}$ throughout using a slight variation of Lemma 6(b) in the Appendix.

The hybrid test has better size properties than the subsample test for the following reason.

Lemma 2 *Suppose Assumptions* A2, B2, C-E, F2, G2, *and* J *hold. Then, either* (i) *the addition of* $c_\infty(1-\alpha)$ *to the subsample critical value is irrelevant asymptotically (i.e.,* $c_h(1-\alpha) \geq c_\infty(1-\alpha)$ *for all* $h \in H$, $Max^-_{Hyb}(\alpha) = Max^-_{Sub}(\alpha)$, *and* $Max_{Hyb}(\alpha) = Max_{Sub}(\alpha)$*), or* (ii) *the nominal level* α *subsample test over-rejects asymptotically (i.e.,* $AsySz(\theta_0) > \alpha$*) and the hybrid test reduces the asymptotic over-rejection for at least some parameter values.*

Theorem 3 in Section 7 below shows that the nominal level α hybrid test has asymptotic level α (i.e., $AsySz(\theta_0) \leq \alpha$) if $c_h(1-\alpha)$ is monotone increasing or decreasing in h (see Comment 1 to Theorem 3). Here we show that if $p = 1$, which occurs in many examples, then the hybrid test has correct size asymptotically under the more general condition that $c_h(1-\alpha)$ is maximized at either h^0 or h^∞, where $h^\infty = (\infty, h_2')'$ or $(-\infty, h_2')'$. For example, the latter condition is satisfied if $c_h(1-\alpha)$ is monotone increasing or decreasing in h_1, is bowl-shaped in h_1, or is wiggly in h_1 with global maximum at 0 or $\pm\infty$. The precise condition is the following. (Here, "Quant" abbreviates "quantile.")

Assumption Quant0. Either (i) (a) for all $h \in H$, $c_\infty(1-\alpha) \geq c_h(1-\alpha)$ and (b) $\sup_{h\in H}[1 - J_h(c_\infty(1-\alpha)-)] = \sup_{h\in H}[1 - J_h(c_\infty(1-\alpha))]$; or (ii) (a) $p = 1$, (b) for all $h \in H$, $c_{h^0}(1-\alpha) \geq c_h(1-\alpha)$, (c) $J_\infty(c_\infty(1-\alpha)-) = J_\infty(c_\infty(1-\alpha))$, and (d) $\sup_{h\in H}[1 - J_h(c_h(1-\alpha)-)] = \sup_{h\in H}[1 - J_h(c_h(1-\alpha))]$.

(In Assumption Quant0(ii), $h^0 = (0, h_2)$ given $h = (h_1, h_2)$.) The main force of Assumption Quant0 is parts (i)(a), (ii)(a), and (ii)(b). Parts (i)(b), (ii)(c), and (ii)(d) only require suitable continuity of J_h.

Lemma 3 *Let* $\alpha \in (0, 1)$ *be given. Suppose Assumptions* A2, B2, C-E, F2, G2, J, *and* Quant0 *hold. Then, the hybrid test based on* $T_n(\theta_0)$ *has* $AsySz(\theta_0) \leq \alpha$.

6 Size-Corrected Tests

In this section, we use Theorem 2 to define size-corrected (SC) FCV, subsample, and hybrid tests. These tests are useful when a test statistic has a distribution with "discontinuous limit." The SC tests only apply when Assumption B2 holds. Typically they do not apply if the asymptotic size of the FCV, subsample, or hybrid test is one. The method of this section applies to CIs as well, see Section 9 below.

The size-corrected fixed critical value (SC-FCV), subsample (SC-Sub), and hybrid (SC-Hyb) tests with nominal level α are defined to reject the null hypothesis $H_0 : \theta = \theta_0$ when

$$\begin{aligned} T_n(\theta_0) &> cv(1-\alpha), \\ T_n(\theta_0) &> c_{n,b}(1-\xi(\alpha)) \text{ and} \\ T_n(\theta_0) &> c^*_{n,b}(1-\xi^*(\alpha)), \end{aligned} \tag{6.1}$$

respectively, where $cv(1-\alpha)$, $\xi(\alpha)$ $(\in (0,\alpha])$, and $\xi^*(\alpha)$ $(\in (0,\alpha])$ are constants defined such that

$$\begin{aligned} \sup_{h \in H}[1 - J_h(cv(1-\alpha)-)] &\leq \alpha, \\ Max^-_{Sub}(\xi(\alpha)) &\leq \alpha \text{ and} \\ Max^-_{Hyb}(\xi^*(\alpha)) &\leq \alpha, \end{aligned} \tag{6.2}$$

respectively. If more than one such value $cv(1-\alpha)$ exists, we take $cv(1-\alpha)$ to be the smallest value.[8] If more than one such value $\xi(\alpha)$ (or $\xi^*(\alpha)$) exists, we take $\xi(\alpha)$ (or $\xi^*(\alpha)$, respectively) to be the largest value.[9] If $\sup_{h \in H}[1 - J_h(c_{Fix}(1-\alpha)-)] \leq \alpha$ (or $Max^-_{Sub}(\alpha) = \alpha$ or $Max^-_{Hyb}(\alpha) = \alpha$), then (i) no size correction is needed, (ii) $cv(1-\alpha) = c_{Fix}(1-\alpha)$ (or $\xi(\alpha) = \alpha$ or $\xi^*(\alpha) = \alpha$, respectively), and (iii) the SC-FCV test (or SC-Sub test or SC-Hyb test, respectively) is just the uncorrected test.[10]

We assume that $cv(1-\alpha)$, $\xi(\alpha)$, and $\xi^*(\alpha)$ values exist that satisfy (6.2):

Assumption K. Given $\alpha \in (0,1)$, there exists a constant $cv(1-\alpha) < \infty$ such that $\sup_{h \in H}[1 - J_h(cv(1-\alpha)-)] \leq \alpha$.

Assumption L. Given $\alpha \in (0,1)$, there exists $\xi(\alpha) \in (0,\alpha]$ such that $Max^-_{Sub}(\xi(\alpha)) \leq \alpha$.

Assumption M. Given $\alpha \in (0,1)$, there exists $\xi^*(\alpha) \in (0,\alpha]$ such that $Max^-_{Hyb}(\xi^*(\alpha)) \leq \alpha$.

Assumptions K, L, and M usually are not restrictive given Assumption B2. On the other hand, Assumption B2 is restrictive and may be violated in some examples of interest. In such cases, it may not be possible to construct size-corrected tests.

In most cases, Assumption K holds with

$$cv(1-\alpha) = \sup_{h \in H} c_h(1-\alpha). \tag{6.3}$$

[8] If no such smallest value exists, we take some value that is arbitrarily close to the infimum of the values that satisfy (6.2).

[9] If no such largest value exists, we take some value that is arbitrarily close to the supremum of the values that satisfy (6.2).

[10] Note that $Max^-_{Sub}(\alpha) < \alpha$ or $Max^-_{Hyb}(\alpha) < \alpha$ is not possible because $(h^0, h^0) \in GH$ and $J_{h^0}(c_{h^0}(1-\alpha)-) \leq 1-\alpha$ for any $h^0 = (0, h^0_2) \in H$ for some $h^0_2 \in H_2$.

A sufficient condition for this is the following.

Assumption KK. (i) $cv(1-\alpha) = \sup_{h \in H} c_h(1-\alpha) < \infty$ and (ii) for any $h \in H$ for which $c_h(1-\alpha) = cv(1-\alpha)$, $J_h(x)$ is continuous at $x = cv(1-\alpha)$.

Lemma 4 *Assumption* KK *implies Assumption* K*.*

If Assumptions L and M hold and $Max^-_{Sub}(\xi)$ and $Max^-_{Hyb}(\xi)$ are continuous for ξ in $(0, \alpha]$, then Assumptions L and M hold with $Max^-_{Sub}(\xi(\alpha)) = \alpha$ and $Max^-_{Hyb}(\xi^*(\alpha)) = \alpha$ for some $\xi(\alpha)$ and $\xi^*(\alpha)$ in $(0, \alpha]$ by the intermediate value theorem.

The following Corollary to Theorem 2 shows that the SC tests have asymptotic size less than or equal to their nominal level α.

Corollary 2 *Let* $\alpha \in (0,1)$ *be given.*

(a) *Suppose Assumptions* A2, B2, *and* K *hold. Then, the SC-FCV test has* $AsySz(\theta_0) \leq \alpha$.

(b) *Suppose Assumptions* A2, B2, *and* KK *hold and* $\sup_{h \in H} c_h(1-\alpha)$ *is attained at some* $h^* \in H$*. Then, the SC-FCV test has* $AsySz(\theta_0) = \alpha$.

(c) *Suppose Assumptions* A2, B2, C-E, G2, *and* L *hold and Assumption* F2 *holds with* α *replaced by* $\xi(\alpha)$*. Then, the SC-Sub test has* $AsySz(\theta_0) \leq \alpha$*. In addition,* $AsySz(\theta_0) = \alpha$ *if* $Max_{Sub}(\xi(\alpha)) = \alpha$.

(d) *Suppose Assumptions* A2, B2, C-E, G2, J, *and* M *hold and Assumption* F2 *holds with* α *replaced by* $\xi^*(\alpha)$*. Then, the SC-Hyb test has* $AsySz(\theta_0) \leq \alpha$*. In addition,* $AsySz(\theta_0) = \alpha$ *if* $Max_{Hyb}(\xi^*(\alpha)) = \alpha$.

Comments. 1. The conditions in parts (b)-(d) under which $AsySz(\theta_0) = \alpha$ hold in most applications.

2. Projection-based tests (e.g., see Dufour and Jasiak (2001)), typically have asymptotic size less than their nominal level—often substantially less. Because SC tests typically have asymptotic size equal to their nominal level, they have potentially substantial power advantages over projection-based tests.

3. A CI for θ constructed by inverting an SC test does not necessarily have asymptotic confidence level greater than or equal to $1-\alpha$. The reason is that the level of a CI depends on uniformity over both θ and γ whereas a test of $H_0 : \theta = \theta_0$, such as a SC test, only requires uniformity over γ. See Section 9 for SC CIs.

4. Corollary 2(a) follows from Theorem 2(a) applied with $c_{Fix}(1-\alpha) = cv(1-\alpha)$ because in this case $Max^-_{Fix}(\alpha) = \sup_{h \in H}[1 - J_h(cv(1-\alpha)-)] \leq \alpha$ by Assumption K. Corollary 2(b) holds by Assumption KK and (12.22) in the proof of Lemma 4 with h replaced by h^* using the assumption in part (b). Corollary 2(c) holds by Theorem 2(b) with α replaced by $\xi(\alpha)$ because in this case $AsySz(\theta_0) \leq Max^-_{Sub}(\xi(\alpha)) \leq \alpha$ using Assumption L. Corollary 2(d) holds by Theorem 2(b) with $(\alpha, c_{n,b}(1-\alpha))$ replaced by $(\xi^*(\alpha), c^*_{n,b}(1-\xi^*(\alpha)))$ throughout, because in this case $AsySz(\theta_0) \leq Max^-_{Hyb}(\xi^*(\alpha)) \leq \alpha$ using Assumption M.

To compute $cv(1-\alpha) = \sup_{h \in H} c_h(1-\alpha)$, one needs to be able to compute the $1-\alpha$ quantile of J_h for each $h \in H$ and to find the maximum of the quantiles over $h \in H$. Computation of quantiles can be done analytically in some cases, by numerical integration if the density of J_h is available, or by simulation if simulating a random variable with distribution J_h is possible. The maximization step may range in difficulty from being very easy to nearly impossible depending on how many elements of h affect the asymptotic distribution J_h, the shape and smoothness of $c_h(1-\alpha)$ as a function of h, and the time needed to compute $c_h(1-\alpha)$ for any given h.

In some examples $cv(1-\alpha)$ can be tabulated for selected values of α. Once this is done, the SC-FCV test is as easy to apply as the non-corrected FCV test.

To compute $\xi(\alpha)$ for the SC-Sub test, one needs to be able to compute $Max^-_{Sub}(\xi)$ for $\xi \in (0, \alpha]$. For this, one needs to be able to compute $c_h(1-\xi)$ for $h \in H$. Next, one needs to be able to maximize $1 - J_h(c_g(1-\xi)-)$ over $(g, h) \in GH$. Depending on how many elements of h affect the asymptotic distribution J_h and on the shape and smoothness of $J_h(c_g(1-\xi)-)$ as a function of (g, h), this is more or less difficult to do.

Given a method for computing the function $Max^-_{Sub}(\xi)$, the scalar $\xi(\alpha)$ can be computed using a halving algorithm because $Max^-_{Sub}(\xi)$ is nondecreasing in ξ and $\xi(\alpha) \in (0, \alpha]$.[11] The halving algorithm has an error bound of $\alpha/2^a$, where a is the number of evaluations of $Max^-_{Sub}(\xi)$. Hence, $a = 7$ yields a percentage error in the computed value of $\xi(\alpha)$ of less than 1%.

In some examples, $\xi(\alpha)$ can be tabulated for selected values of α. In other examples, tabulation of $\xi(\alpha)$ is difficult.

Computation of $\xi^*(\alpha)$ for the SC-Hyb test is analogous to computation of $\xi(\alpha)$ with $1 - J_h(\max\{c_g(1-\alpha), c_\infty(1-\alpha)\}-)$ in place of $1 - J_h(c_g(1-\alpha)-)$.

7 Power Comparisons of Size-Corrected Tests

In this section, we compare the asymptotic power of the SC-FCV, SC-Sub, and SC-Hyb tests. Since all three tests employ the same test statistic $T_n(\theta_0)$, the asymptotic power comparison is based on a comparison of the magnitudes of $cv(1-\alpha)$, $c_{n,b}(1-\xi(\alpha))$, and $c^*_{n,b}(1-\xi^*(\alpha))$ for n large. The first of these is fixed. The latter two are random and their large sample behavior depends on the particular sequence $\{\gamma_n \in \Gamma : n \geq 1\}$ of true parameters and may depend on whether the null hypothesis is true or not. We focus on the case in which they do not depend on whether the null hypothesis is true or not. This typically holds when the subsample statistics are defined to satisfy Assumption Sub1 (and fails when they satisfy Assumption Sub2). We also focus on the case where Assumption KK holds, so that $cv(1-\alpha) = \sup_{h \in H} c_h(1-\alpha)$.

From the definitions of the critical values $c_{n,b}(1-\xi(\alpha))$ and $c^*_{n,b}(1-\xi^*(\alpha))$ of the SC-Sub and SC-Hyb tests and Lemma 7(e) (in the Appendix), the possible limits of

[11] The halving algorithm starts with the potential solution set being $(0, \alpha]$. First, one computes $Max^-_{Sub}(\alpha/2)$. If it exceeds α, the solution set becomes $(0, \alpha/2]$ and one computes $Max^-_{Sub}(\alpha/4)$; if not, the solution set becomes $[\alpha/2, \alpha)$ and one computes $Max^-_{Sub}(3\alpha/4)$. One continues to convergence.

the critical values are

$$c_g(1-\xi(\alpha)) \text{ and } c^*_{g,\infty}(1-\xi^*(\alpha)) \text{ for } g \in H, \text{ where}$$
$$c^*_{g,\infty}(1-\alpha) = \max\{c_g(1-\alpha), c_\infty(1-\alpha)\}. \tag{7.1}$$

Hence, we are interested in the relative magnitudes of $cv(1-\alpha)$, $c_g(1-\xi(\alpha))$, and $c^*_{g,\infty}(1-\xi^*(\alpha))$ for $g \in H$. These relative magnitudes are determined by the shapes of the quantiles $c_h(1-\alpha)$ and $c_h(1-\xi(\alpha))$ as functions of $h \in H$. Theorem 3 below essentially shows that (i) if $c_g(1-\xi(\alpha))$ is monotone decreasing in $g \in H$, then $c_g(1-\xi(\alpha)) = c^*_{g,\infty}(1-\xi^*(\alpha)) \leq cv(1-\alpha)$ for all $g \in H$ with strict inequality for some g, (ii) if $c_g(1-\xi(\alpha))$ is monotone increasing in $g \in H$, then $cv(1-\alpha) = c^*_{g,\infty}(1-\xi^*(\alpha)) \leq c_g(1-\xi(\alpha))$ for all $g \in H$ with strict inequality for some g, and (iii) if $c_g(1-\xi(\alpha))$ is not monotone in $g \in H$, then one can have $cv(1-\alpha) \leq c_g(1-\xi(\alpha))$ for some $g \in H$ and $cv(1-\alpha) \geq c_g(1-\xi(\alpha))$ for some $g \in H$ with strict inequalities for some g.

In case (i), the SC-Sub and SC-Hyb tests are equivalent and are more powerful than the SC-FCV test; in case (ii), the SC-FCV and SC-Hyb tests are equivalent and are more powerful than the SC-Sub test; and in case (iii), the SC-FCV and SC-Sub tests cannot be unambiguously ranked.

These results show that the SC-Hyb test has some nice power properties. When the SC-Sub test dominates the SC-FCV test, the SC-Hyb test behaves like the SC-Sub test. And when the SC-FCV test dominates the SC-Sub test, SC-Hyb test behave like SC-FCV test.

We now state three alternative assumptions concerning the shapes of $c_h(1-\alpha)$ and $c_h(1-\xi(\alpha))$, which correspond to cases (i)-(iii) above. ("Quant" refers to "quantile.")

Assumption Quant1. (i) $c_g(1-\alpha) \geq c_h(1-\alpha)$ for all $(g,h) \in GH$, and (ii) $Max^-_{Sub}(\alpha) = Max_{Sub}(\alpha)$.

Assumption Quant2. (i) $c_g(1-\alpha) \leq c_h(1-\alpha)$ for all $(g,h) \in GH$, and (ii) $c_g(1-\xi(\alpha)) \leq c_h(1-\xi(\alpha))$ for all $(g,h) \in GH$.

Assumption Quant3. (i) $c_h(1-\alpha)$ is maximized over $h \in H$ at some $h^* = (h^*_1, h^*_2) = (h^*_{1,1}, ..., h^*_{1,p}, h^*_2) \in H$ with $h^*_{1,j} \neq 0$ and $h^*_{1,j} \neq \pm\infty$ for some $j = 1, ..., p$, (ii) $c_{h^{*0}}(1-\alpha) < c_{h^*}(1-\alpha)$, where $h^{*0} = (0, h^*_2)$, (iii) $J_{h^*}(x)$ is continuous at $x = c_{h^*}(1-\alpha)$, and (iv) $c_{h^\dagger}(1-\xi(\alpha)) < c_{h^*}(1-\alpha)$ for some $h^\dagger = (h^\dagger_1, h^\dagger_2) = ((h^\dagger_{1,1}, ..., h^\dagger_{1,p}, h^\dagger_2) \in H$ with $h^\dagger_{1,j} \neq 0$ and $h^\dagger_{1,j} \neq \pm\infty$ for some $j = 1, ..., p$.

Theorem 3 (a) *Suppose Assumptions* J, KK, L, M, *and* Quant1 *hold. Then,* (i) $cv(1-\alpha) = \sup_{h \in H} c_{h^0}(1-\alpha)$ (*where* $h^0 = (0, h_2)$ *given* $h = (h_1, h_2)$), (ii) $\xi(\alpha) = \alpha$, (iii) $\xi^*(\alpha) = \alpha$, (iv) $c^*_{g,\infty}(1-\xi^*(\alpha)) = c_g(1-\xi(\alpha))$, *and* (v) $c_g(1-\xi(\alpha)) \leq cv(1-\alpha)$ *for all* $g \in H$.

(b) *Suppose Assumptions* J, KK, L, M, *and* Quant2 *hold. Then,* (i) $cv(1-\alpha) = c_\infty(1-\alpha)$, (ii) $\xi^*(\alpha) = \alpha$, (iii) $c^*_{g,\infty}(1-\xi^*(\alpha)) = cv(1-\alpha)$, *and* (iv) $cv(1-\alpha) \leq c_g(1-\xi(\alpha))$ *for all* $g \in H$.

(c) *Suppose Assumptions* KK, L, *and* Quant3 *hold. Then,* (i) $cv(1-\alpha) = c_{h^*}(1-\alpha)$, (ii) $c_{h^*}(1-\xi(\alpha)) > cv(1-\alpha)$, *and* (iii) $c_{h^\dagger}(1-\xi(\alpha)) < cv(1-\alpha)$.

Comments. 1. Theorem 3(a)(ii) shows that the standard subsample test (without size correction) has correct asymptotic size when the quantile function $c_g(1-\alpha)$ is monotone increasing in $g \in H$. Theorem 3(a)(iii) and (b)(ii) show that the hybrid test has correct asymptotic size when the quantile function $c_g(1-\alpha)$ is monotone increasing or decreasing in $g \in H$.

2. If Assumption Quant1(i) holds with a strict inequality for $(g,h) = (h^0, h)$ for some $h = (h_1, h_2) \in H$, where $h^0 = (0, h_2) \in H$, then Theorem 3(a)(v) holds with a strict inequality with g equal to this value of h. If Assumption Quant2(ii) holds with a strict inequality for $(g,h) = (h^0, h)$ for some $h = (h_1, h_2) \in H$, where $h^0 = (0, h_2) \in H$, then the inequality in Theorem 3(b)(iv) holds with a strict inequality with g equal to this value of h.

3. Assumption J is not needed in Theorem 3(a)(i), (ii), and (v).

The results above are relevant when the subsample statistics satisfy Assumption Sub1 (because then their asymptotic behavior is the same under the null and the alternative). On the other hand, if Assumption Sub2 holds, then the subsample critical values typically diverge to infinity under fixed alternatives (at rate $b_n^{1/2} << n^{1/2}$ when Assumption t1 holds). Hence, in this case, the SC-FCV test is more powerful asymptotically than the SC-Sub test for distant alternatives. For brevity, we do not investigate the relative magnitudes of the critical values of the SC-FCV and SC-Sub tests for local alternatives when Assumption Sub2 holds.

8 Equal-tailed t Tests

This section considers *equal-tailed* two-sided t tests. There are two reasons for considering such tests. First, equal-tailed tests and CIs are preferred to symmetric procedures by some statisticians, e.g., see Efron and Tibshirani (1993). Second, given the potential problems of symmetric t tests documented in Section 4, it is of interest to see whether equal-tailed tests are subject to the same problems and, if so, whether the problems are more or less severe than for symmetric procedures.

We suppose Assumption t1(i) holds, so that $T_n(\theta_0) = \tau_n(\widehat{\theta}_n - \theta_0)/\widehat{\sigma}_n$. An equal-tailed FCV, subsample, or hybrid t test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ of nominal level α ($\in (0, 1/2)$) rejects H_0 when

$$T_n(\theta_0) > c_{1-\alpha/2} \text{ or } T_n(\theta_0) < c_{\alpha/2}, \tag{8.1}$$

where $c_{1-\alpha}$ is defined in (2.8) for FCV and subsample tests. For hybrid tests, the critical values in (8.1) are

$$\begin{aligned} c_{1-\alpha/2} &= c^*_{n,b}(1-\alpha/2) = \max\{c_{n,b}(1-\alpha/2), c_\infty(1-\alpha/2)\} \text{ and} \\ c_{\alpha/2} &= c^{**}_{n,b}(\alpha/2) = \min\{c_{n,b}(\alpha/2), c_\infty(\alpha/2)\}. \end{aligned} \tag{8.2}$$

The exact size, $ExSz_n(\theta_0)$, of the equal-tailed t test is

$$ExSz_n(\theta_0) = \sup_{\gamma \in \Gamma} \left(P_{\theta_0,\gamma}(T_n(\theta_0) > c_{1-\alpha/2}) + P_{\theta_0,\gamma}(T_n(\theta_0) < c_{\alpha/2}) \right). \tag{8.3}$$

The asymptotic size of the test is $AsySz(\theta_0) = \limsup_{n\to\infty} ExSz_n(\theta_0)$. The minimum rejection probability, $MinRP_n(\theta_0)$, of the test is the same as $ExSz_n(\theta_0)$ but with "sup" replaced by "inf" and $AsyMinRP(\theta_0) = \liminf_{n\to\infty} MinRP_n(\theta_0)$.

For equal-tailed subsample t tests, we replace Assumptions F1 and F2 by the following assumptions, which are not very restrictive.

Assumption N1. For all $\varepsilon > 0$, $J_{h^0}(c_{h^0}(\kappa) - \varepsilon) < \kappa$ and $J_{h^0}(c_{h^0}(\kappa) + \varepsilon) > \kappa$ for $\kappa = \alpha/2$ and $\kappa = 1 - \alpha/2$, where $c_{h^0}(1-\alpha)$ is the $1-\alpha$ quantile of J_{h^0} and h^0 is as in Assumption B1.

Assumption N2. For all $\varepsilon > 0$ and $h \in H$, $J_h(c_h(\kappa) - \varepsilon) < \kappa$ and $J_h(c_h(\kappa) + \varepsilon) > \kappa$ for $\kappa = \alpha/2$ and $\kappa = 1 - \alpha/2$, where $c_h(1-\alpha)$ is the $1-\alpha$ quantile of J_h.

Define

$$\begin{aligned}
Max^{r-}_{ET,Fix}(\alpha) &= \sup_{h\in H}[1 - J_h(c_{Fix}(1-\alpha/2)) + J_h(c_{Fix}(\alpha/2)-)], \\
Max^{\ell-}_{ET,Fix}(\alpha) &= \sup_{h\in H}[1 - J_h(c_{Fix}(1-\alpha/2)-) + J_h(c_{Fix}(\alpha/2))], \\
Max^{r-}_{ET,Sub}(\alpha) &= \sup_{(g,h)\in GH}[1 - J_h(c_g(1-\alpha/2)) + J_h(c_g(\alpha/2)-)], \text{ and} \\
Max^{\ell-}_{ET,Sub}(\alpha) &= \sup_{(g,h)\in GH}[1 - J_h(c_g(1-\alpha/2)-) + J_h(c_g(\alpha/2))].
\end{aligned} \tag{8.4}$$

Here "$r-$" denotes that the limit from the left "$-$" appears in the *right* summand in the expression for $Max^{r-}_{ET,Fix}(\alpha)$. Analogously, "$\ell-$" denotes that it appears in the *left* summand in the expression for $Max^{\ell-}_{ET,Fix}(\alpha)$. Define $Max^{r-}_{ET,Hyb}(\alpha)$ and $Max^{\ell-}_{ET,Hyb}(\alpha)$ as $Max^{r-}_{ET,Sub}(\alpha)$ and $Max^{\ell-}_{ET,Sub}(\alpha)$ are defined, but with $\max\{c_g(1-\alpha/2), c_\infty(1-\alpha/2)\}$ in place of $c_g(1-\alpha/2)$ and with $\min\{c_g(\alpha/2), c_\infty(\alpha/2)\}$ in place of $c_g(\alpha/2)$. Define $Min^{r-}_{ET,Fix}(\alpha), ..., Min^{\ell-}_{ET,Hyb}(\alpha)$ analogously with "inf" in place of "sup."

In the "continuous limit" case, $Max^{r-}_{ET,Sub}(\alpha)$ simplifies to $\sup_{h\in H}[1 - J_h(c_h(1 - \alpha/2)) + J_h(c_h(\alpha/2)-)]$ and likewise for $Max^{\ell-}_{ET,Sub}(\alpha)$.

The proofs of Theorems 1 and 2 can be adjusted straightforwardly to yield the following results for equal-tailed FCV, subsample, and hybrid t tests.

Corollary 3 *Let $\alpha \in (0, 1/2)$ be given. Let $T_n(\theta_0)$ be defined as in Assumption* t1(i).

(a) *Suppose Assumption* B1(i) *holds. Then, an equal-tailed FCV t test satisfies*

$$\begin{aligned}
&P_{\theta_0,\gamma_{n,h}}(T_n(\theta_0) > c_{Fix}(1-\alpha/2) \text{ or } T_n(\theta_0) < c_{Fix}(\alpha/2)) \\
&\quad \to [1 - J_h(c_{Fix}(1-\alpha/2)) + J_h(c_{Fix}(\alpha/2)-), \\
&\qquad 1 - J_h(c_{Fix}(1-\alpha/2)-) + J_h(c_{Fix}(\alpha/2))].
\end{aligned}$$

(b) *Suppose Assumptions* A1, B1, C-E, G1, *and* N1 *hold. Then, an equal-tailed subsample t test satisfies*

$$P_{\theta_0,\gamma_{n,h}}(T_n(\theta_0) > c_{n,b}(1-\alpha/2) \text{ or } T_n(\theta_0) < c_{n,b}(\alpha/2))$$
$$\to [1 - J_h(c_{h^0}(1-\alpha/2)) + J_h(c_{h^0}(\alpha/2)-), 1 - J_h(c_{h^0}(1-\alpha/2)-) + J_h(c_{h^0}(\alpha/2))].$$

(c) *Suppose Assumptions* A1, B1, C-E, G1, J, *and* N1 *hold. Then, the result of part* (b) *holds for an equal-tailed hybrid test with* $c_{n,b}(1-\alpha/2)$, $c_{n,b}(\alpha/2)$, $c_{h^0}(1-\alpha/2)$, *and* $c_{h^0}(\alpha/2)$ *replaced by* $c^*_{n,b}(1-\alpha/2)$, $c^{**}_{n,b}(\alpha/2)$, $\max\{c_{h^0}(1-\alpha/2), c_\infty(1-\alpha/2)\}$, *and* $\min\{c_{h^0}(\alpha/2), c_\infty(\alpha/2)\}$, *respectively.*

(d) *Suppose Assumptions* A2 *and* B2 *hold. Then, an equal-tailed FCV t test satisfies*

$$AsySz(\theta_0) \in [Max^{r-}_{ET,Fix}(\alpha), Max^{\ell-}_{ET,Fix}(\alpha)] \text{ and}$$
$$AsyMinRP(\theta_0) \in [Min^{r-}_{ET,Fix}(\alpha), Min^{\ell-}_{ET,Fix}(\alpha)].$$

(e) *Suppose Assumptions* A2, B2, C-E, G2, *and* N2 *hold. Then, an equal-tailed subsample t test satisfies the result of part* (d) *with* Sub *in place of* Fix.

(f) *Suppose Assumptions* A2, B2, C-E, G2, J, *and* N2 *hold. Then, an equal-tailed hybrid t test satisfies the result of part* (d) *with* Hyb *in place of* Fix.

Comments. 1. If $J_h(x)$ is continuous at the appropriate value(s) of x, then $Max^{r-}_{ET,Fix}(\alpha) = Max^{\ell-}_{ET,Fix}(\alpha)$ etc. and Corollary 3 gives the precise value of $AsySz(\theta_0)$.

2. By Corollary 3(e) and the definition of $Max^{\ell-}_{ET,Sub}(\alpha)$, sufficient conditions for a nominal level α equal-tailed subsample test to have asymptotic level α are the following: (a) $c_g(1-\alpha/2) \geq c_h(1-\alpha/2)$ for all $(g,h) \in GH$, (b) $c_g(\alpha/2) \leq c_h(\alpha/2)$ for all $(g,h) \in GH$, and (c) $\sup_{h \in H}[1-J_h(c_h(1-\alpha/2)-)+J_h(c_h(\alpha/2))] = \sup_{h \in H}[1-J_h(c_h(1-\alpha/2))+J_h(c_h(\alpha/2)-)]$. Conditions (a) and (b) automatically hold in "continuous limit" cases. They also hold in some "discontinuous limit" cases, but often fail in such cases. Condition (c) holds in most examples. (Note that conditions (a)-(c) are not *necessary* for a subsample test to have asymptotic level α.)

3. Theorems 1 and 2 give results concerning the null rejection rates for each tail separately of an equal-tailed t test. If one is interested in an equal-tailed t test, rather than a symmetric t test, such rates are of interest.

Size-corrected equal-tailed subsample t tests can be constructed by finding $\xi(\alpha)$ such that $Max^{\ell-}_{ET,Sub}(\xi(\alpha)) \leq \alpha$. In particular, The equal-tailed SC-Sub test is defined by (8.1) with $c_{1-\alpha/2} = c_{n,b}(1-\xi(\alpha)/2)$ and $c_{\alpha/2} = c_{n,b}(\xi(\alpha/2))$. Analogously, size-corrected equal-tailed FCV and hybrid t tests can be constructed by finding $\xi_{Fix}(\alpha)$ and $\xi^*(\alpha)$, respectively, such that $Max^{\ell-}_{ET,Fix}(\xi_{Fix}(\alpha)) \leq \alpha$ and $Max^{\ell-}_{ET,Hyb}(\xi^*(\alpha)) \leq \alpha$. In each case, this guarantees that the "overall" size of the test is less than or equal to α. It does not guarantee that the maximum rejection probability in each tail is less than or equal to $\alpha/2$. If the latter is desired, then one should size correct the lower and upper critical values of the equal-tailed test in the same way as one-sided t tests are size corrected in Section 6. (This can yield the overall size of the test to be strictly less than α if the (g, h) vector that maximizes the rejection probability is different for the lower and upper critical values.)

The equal-tailed SC-FCV, SC-Sub, and SC-Hyb t tests based on $\xi_{Fix}(\alpha)$, $\xi(\alpha)$, and $\xi^*(\alpha)$, respectively, have $AsySz(\theta_0) \leq \alpha$ under assumptions guaranteeing the existence of the size-correcting values $\xi_{Fix}(\alpha), \xi(\alpha)$, and $\xi^*(\alpha)$ and under the assumptions in Corollary 3(d), (e), and (f), respectively (with Assumption $N2$ holding for $(1 - \xi(\alpha)/2, \xi(\alpha)/2)$ and $(1 - \xi^*(\alpha)/2, \xi^*(\alpha)/2)$ in place of $(1 - \alpha/2, \alpha/2)$ in parts (e) and (f), respectively).

9 Confidence Intervals

In this section, we consider CIs for a parameter $\theta \in R^d$ when nuisance parameters $\eta \in R^s$ and $\gamma_3 \in \mathcal{T}_3$ may appear. To avoid considerable repetition, we recycle the definitions, assumptions, and results given in earlier sections for tests, but with θ and η defined to be part of the vector γ. In previous sections, θ and γ are separate parameters. Here, θ is a sub-vector of γ. The reason for making this change is that the confidence level of a CI for θ by definition depends on uniformity over both the nuisance parameters η and γ_3 *and* the parameter of interest θ. In contrast, the level of a test concerning θ only depends on uniformity over the nuisance parameters and not over θ (because θ is fixed under the null hypothesis). By making θ a sub-vector of γ, the results from previous sections, which are uniform over $\gamma \in \Gamma$, give the uniformity results that we need for CIs for θ. Of course, with this change, the index parameter h, the asymptotic distributions $\{J_h : h \in H\}$, and the assumptions are different in any given model in this CI section from the earlier test sections.

Specifically, we partition θ into $(\theta_1', \theta_2')'$, where $\theta_j \in R^{dj}$ for $j = 1, 2$, and we partition η into $(\eta_1', \eta_2')'$, where $\eta_j \in R^{s_j}$ for $j = 1, 2$. Then, we consider the same set-up as in Section 2 where $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ with $\gamma_1 = (\theta_1', \eta_1')'$ and $\gamma_2 = (\theta_2', \eta_2')'$, where $p = d_1 + s_1$ and $q = p_2 + s_2$. Thus, θ and η are partitioned such that θ_1 and η_1 determine whether γ is close to the point of discontinuity of the asymptotic distribution of the test statistic $T_n(\theta)$, whereas θ_2 and η_2 do not, but they still may affect the limit distribution of $T_n(\theta)$. In most examples, either no parameter θ_1 or θ_2 appears (i.e., $d_2 = 0$ or $d_1 = 0$) and either no parameter η_1 or η_2 appears (i.e., $s_2 = 0$ or $s_1 = 0$).

9.1 Basic Results for Confidence Intervals

We consider the same test statistic $T_n(\theta_0)$ for testing the null hypothesis $H_0 : \theta = \theta_0$ as above. Fixed, subsample, and hybrid critical values are defined as above. We obtain CIs for θ by inverting tests based on $T_n(\theta_0)$. When a fixed critical value is employed, this yields a FCV CI. When a subsample or hybrid critical value is employed, it yields a subsample or hybrid CI, respectively. Let Θ $(\subset R^d)$ denote the parameter space for

θ and let Γ denote the parameter space for γ. The CI for θ contains all points $\theta_0 \in \Theta$ for which the test of $H_0 : \theta = \theta_0$ fails to reject the null hypothesis:

$$CI_n = \{\theta_0 \in \Theta : T_n(\theta_0) \leq c_{1-\alpha}\}, \tag{9.1}$$

where $c_{1-\alpha}$ is a critical value equal to $c_{Fix}(1-\alpha)$, $c_{n,b}(1-\alpha)$, or $c^*_{n,b}(1-\alpha)$.

For example, suppose $T_n(\theta_0)$ is a (i) upper one-sided, (ii) lower one-sided, or (iii) symmetric two-sided t test of nominal level α (i.e., Assumption t1(i), (ii), or (iii) holds). Then, the corresponding CI of nominal level α is defined by

$$\begin{aligned} CI_n &= [\widehat{\theta}_n - \tau_n^{-1}\widehat{\sigma}_n c_{1-\alpha}, \infty), \\ CI_n &= (-\infty, \widehat{\theta}_n + \tau_n^{-1}\widehat{\sigma}_n c_{1-\alpha}], \text{ or} \\ CI_n &= [\widehat{\theta}_n - \tau_n^{-1}\widehat{\sigma}_n c_{1-\alpha}, \widehat{\theta}_n + \tau_n^{-1}\widehat{\sigma}_n c_{1-\alpha}], \end{aligned} \tag{9.2}$$

respectively.

The coverage probability of the CI defined in (9.1) when γ is the true parameter is

$$P_\gamma(\theta \in CI_n) = P_\gamma(T_n(\theta) \leq c_{1-\alpha}) := 1 - RP_n(\gamma), \tag{9.3}$$

where probabilities are indexed by $\gamma = ((\theta_1', \eta_1')', (\theta_2', \eta_2')', \gamma_3)$ here, whereas they are indexed by (θ, γ) in earlier sections. The exact and asymptotic confidence levels of CI_n are

$$ExCL_n = \inf_{\gamma \in \Gamma}(1 - RP_n(\gamma)) \text{ and } AsyCL = \liminf_{n\to\infty} ExCL_n, \tag{9.4}$$

respectively. Note that the confidence level depends on uniformity over both θ and η because $\gamma = (\gamma_1, \gamma_2, \gamma_3) = ((\theta_1', \eta_1')', (\theta_1', \eta_1')', \gamma_3)$.

We employ the same assumptions as in Section 3 but with the following changes.

Assumption Adjustments for CIs: (i) θ is a sub-vector of γ, rather than a separate parameter from γ. In particular, $\gamma = (\gamma_1, \gamma_2, \gamma_3) = ((\theta_1', \eta_1')', (\theta_2', \eta_2')', \gamma_3)$ for $\theta = (\theta_1', \theta_2')'$ and $\eta = (\eta_1', \eta_2')'$.

(ii) Instead of the true probabilities under a sequence $\{\gamma_{n,h} : n \geq 1\}$ being $\{P_{\theta_0, \gamma_{n,h}}(\cdot) : n \geq 1\}$, they are $\{P_{\gamma_{n,h}}(\cdot) : n \geq 1\}$.

(iii) The test statistic $T_n(\theta_0)$ is replaced in the assumptions under a true sequence $\{\gamma_{n,h} : n \geq 1\}$ by $T_n(\theta_{n,h})$, where $\gamma_{n,h} = (\gamma_{n,h,1}, \gamma_{n,h,2}, \gamma_{n,h,3})' = ((\theta_{n,h,1}', \eta_{n,h,1}')', (\theta_{n,h,2}', \eta_{n,h,2}')', \gamma_{n,h,3})$.

(iv) In Assumption D, θ_0 in $T_{n,b_n,i}(\theta_0)$ and $T_{b_n}(\theta_0)$ is replaced by θ, where $\theta = (\theta_1', \theta_2')'$ and $\gamma = (\gamma_1, \gamma_2, \gamma_3) = ((\theta_1', \eta_1')', (\theta_2', \eta_2')', \gamma_3)$.

(v) θ_0 is replaced in the definition of $U_{n,b}(x)$ in (2.7) by θ_n when the true parameter is $\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) = ((\theta_{n,1}', \eta_{n,1}')', (\theta_{n,2}', \eta_{n,2}')', \gamma_{n,3})$ and $\theta_n = ((\theta_{n,1}', \theta_{n,2}')'$.

With these changes in the assumptions and corresponding changes in the proofs, the proofs of Theorems 1 and 2 go through.[12] This yields the following results for FCV, subsample, and hybrid CIs.

[12] In the proofs of Corollary 4(d) and (e), $AsySz(\theta_0)$ is replaced by $1 - AsyCL$, $RP_n(\theta_0, \gamma)$ is replaced by $RP_n(\gamma)$, and one makes use of the fact that $\inf_{h\in H} J_h(c_{Fix}(1-\alpha)-) = 1 - Max^-_{Fix}(\alpha)$, $\inf_{(g,h)\in GH} J_h(c_g(1-\alpha)-) = 1 - Max^-_{Sub}(\alpha)$, etc.

Corollary 4 *Let $\alpha \in (0,1)$ be given. Let the assumptions be adjusted as stated above.*

(a) *Suppose Assumption* B1(i) *holds. Then,* $P_{\gamma_{n,h}}(T_n(\theta_{n,h}) \leq c_{Fix}(1-\alpha)) \to [J_h(c_{Fix}(1-\alpha)-), J_h(c_{Fix}(1-\alpha))]$.

(b) *Suppose Assumptions* A1, B1, C-E, F1, *and* G1 *hold. Then,* $P_{\gamma_{n,h}}(T_n(\theta_{n,h}) \leq c_{n,b}(1-\alpha)) \to [J_h(c_{h^0}(1-\alpha)-), J_h(c_{h^0}(1-\alpha))]$.

(c) *Suppose Assumptions* A1, B1, C-E, F1, G1, *and* J *hold. Then, the result of part* (b) *holds with* $c^*_{n,b}(1-\alpha)$ *and* $\max\{c_{h^0}(1-\alpha), c_\infty(1-\alpha)\}$ *in place of* $c_{n,b}(1-\alpha)$ *and* $c_{h^0}(1-\alpha)$*, respectively.*

(d) *Suppose Assumptions* A2 *and* B2 *hold. Then, the* FCV CI *satisfies* $AsyCL \in [\inf_{h\in H} J_h(c_{Fix}(1-\alpha)-), \inf_{h \in H} J_h(c_{Fix}(1-\alpha))]$.

(e) *Suppose Assumptions* A2, B2, C-E, F2, *and* G2 *hold. Then, the subsample* CI *satisfies* $AsyCL \in [\inf_{(g,h)\in GH} J_h(c_g(1-\alpha)-), \inf_{(g,h)\in GH} J_h(c_g(1-\alpha))]$.

(f) *Suppose Assumptions* A2, B2, C-E, F2, G2, *and* J *hold. Then, the hybrid* CI *satisfies the result of part* (e) *with* $\max\{c_g(1-\alpha), c_\infty(1-\alpha)\}$ *in place of* $c_g(1-\alpha)$.

Comment. The result of part (a) shows that if $J_h(c_{Fix}(1-\alpha)) < 1-\alpha$ for some $h \in R^{p+q}$, then the FCV CI has asymptotic level less than its nominal level $1-\alpha$: $AsyCL < 1-\alpha$. Similarly, if $J_h(c_{h^0}(1-\alpha)) < 1-\alpha$ or $J_h(\max\{c_{h^0}(1-\alpha), c_\infty(1-\alpha)\}) < 1-\alpha$, then parts (b) or (c) show that the subsample or hybrid CI, respectively, has asymptotic level less than its nominal level $1-\alpha$: $AsyCL < 1-\alpha$. Parts (d)-(f) establish $AsyCL$ precisely.

9.2 Equal-tailed t Confidence Intervals

An equal-tailed FCV or subsample t CI for θ of nominal level α is defined by

$$CI_n = [\widehat{\theta}_n - \tau_n^{-1}\widehat{\sigma}_n c_{1-\alpha/2}, \widehat{\theta}_n - \tau_n^{-1}\widehat{\sigma}_n c_{\alpha/2}]. \tag{9.5}$$

The following result for such CIs follows from Corollary 3(d)-(f) for equal-tailed t tests, but with the assumptions adjusted as above. For brevity, we do not give analogues of Corollary 3(a)-(c) for CIs.

Corollary 5 *Let $\alpha \in (0, 1/2)$ be given. Let the assumptions be adjusted as described above Corollary* 4.

(a) *Suppose Assumptions* A2 *and* B2 *hold. Then, an equal-tailed FCV t CI satisfies*

$$AsyCL \in [1 - Max^{\ell-}_{ET,Fix}(\alpha), 1 - Max^{r-}_{ET,Fix}(\alpha)].$$

(b) *Suppose Assumptions* A2, B2, C-E, G2, *and* N2 *hold. Then, equal-tailed subsample t CIs satisfy the result of part* (a) *with* Sub *in place of* Fix. *If Assumption* J *also holds, then the equal-tailed hybrid t CIs satisfy the result of part* (a) *with* Hyb *in place of* Fix.

9.3 Size-Corrected Confidence Intervals

Size-corrected CIs are defined as in (9.1), but with their critical values $c_{1-\alpha}$ defined as in Section 6 for SC tests. The assumptions are altered as in the paragraph preceding Corollary 4. Hence, the asymptotic distributions $J_h(\cdot)$ and quantiles $c_h(1-\alpha)$ that arise in the formulae for the critical values are those that apply when $\gamma = (\gamma_1, \gamma_2, \gamma_3) = ((\theta_1', \eta_1')', (\theta_2', \eta_2')', \gamma_3)$ (and typically are different from those that apply in the testing sections which rely on the unaltered assumptions).

The SC CIs satisfy the following properties, which follow from Corollary 2.

Corollary 6 *Let* $\alpha \in (0,1)$ *be given. Let the assumptions be adjusted as described above Corollary* 4.

(a) *Suppose Assumptions* A2, B2, *and* K *hold. Then, the SC-FCV CI satisfies* $AsyCL \geq 1-\alpha$, *and* $AsyCL = 1-\alpha$ *if Assumption* KK *holds and* $\sup_{h\in H} c_h(1-\alpha)$ *is attained at some* $h^* \in H$.

(b) *Suppose Assumptions* A2, B2, C-E, G2, *and* L *hold and Assumption* F2 *holds with* α *replaced by* $\xi(\alpha)$. *Then, the SC-Sub CI satisfies* $AsyCL \geq 1-\alpha$, *and* $AsyCL = 1-\alpha$ *if* $Max^-_{Sub}(\xi(\alpha)) = Max_{Sub}(\xi(\alpha)) = \alpha$.

(c) *Suppose Assumptions* A2, B2, C-E, G2, J, *and* M *hold and Assumption* F2 *holds with* α *replaced by* $\xi^*(\alpha)$. *Then, the SC-Hyb CI satisfies* $AsyCL \geq 1-\alpha$, *and* $AsyCL = 1-\alpha$ *if* $Max^-_{Hyb}(\xi^*(\alpha)) = Max_{Hyb}(\xi(\alpha)) = \alpha$.

10 Studentized t Statistics

In this section we provide sufficient conditions for Assumption G2 for the case when T_n is a studentized t statistic and the subsample statistics satisfy Assumption Sub1. This result generalizes Lemma 1 because Assumption t2 is not imposed. The results apply to models with iid, stationary and weakly dependent, or nonstationary observations.

Just as $T_{n,b_n,i}(\theta_0)$ is defined, let $(\widehat{\theta}_{n,b_n,i}, \widehat{\sigma}_{n,b_n,i})$ be the subsample statistics that are defined exactly as $(\widehat{\theta}_n, \widehat{\sigma}_n)$ are defined, but based on the ith subsample of size b_n. In analogy to $U_{n,b_n}(x)$ defined in (2.7), we define

$$U^{\sigma}_{n,b_n}(x) = q_n^{-1} \sum_{i=1}^{q_n} 1(d_{b_n} \widehat{\sigma}_{n,b_n,i} \leq x) \qquad (10.1)$$

for a sequence of normalization constants $\{d_n : n \geq 1\}$. Although $U^{\sigma}_{n,b_n}(x)$ depends on $\{d_n : n \geq 1\}$, we suppress the dependence for notational simplicity.

We now state modified versions of Assumptions B2, D, E, and H that are used with studentized statistics when Assumption Sub1 holds.

Assumption BB2. (i) For some $r > 0$, all $h \in H$, all sequences $\{\gamma_{n,h} : n \geq 1\}$, some normalization sequences of positive constants $\{a_n : n \geq 1\}$ and $\{d_n : n \geq 1\}$ such that $\tau_n = a_n/d_n$, and some distribution (V_h, W_h) on R^2, $(a_n(\widehat{\theta}_n - \theta_0), d_n\widehat{\sigma}_n) \to_d (V_h, W_h)$

under $\{\gamma_{n,h} : n \geq 1\}$, (ii) $P_{\theta_0,\gamma_{n,h}}(\{\widehat{\sigma}_{n,b_n,i} > 0$ for all $i = 1, ..., q_n\}) \to 1$ under all sequences $\{\gamma_{n,h} : n \geq 1\}$ and all $h \in H$, and (iii) $W_h(0) = 0$ for all $h \in H$.

Assumption DD. (i) $\{(\widehat{\theta}_{n,b_n,i}, \widehat{\sigma}_{n,b_n,i}) : i = 1, ..., q_n\}$ are identically distributed under any $\gamma \in \Gamma$ for all $n \geq 1$, and (ii) $(\widehat{\theta}_{n,b_n,1}, \widehat{\sigma}_{n,b_n,1})$ and $(\widehat{\theta}_{b_n}, \widehat{\sigma}_{b_n})$ have the same distribution under any $\gamma \in \Gamma$ for all $n \geq 1$.

Assumption EE. For all $h \in H$ and all sequences $\{\gamma_{n,h} : n \geq 1\}$ with corresponding normalization $\{d_n : n \geq 1\}$ as in Assumption BB2, $U^{\sigma}_{n,b}(x) - E_{\theta_0,\gamma_{n,h}} U^{\sigma}_{n,b}(x) \to_p 0$ under $\{\gamma_{n,h} : n \geq 1\}$ for all $x \in R$.

Assumption HH. $a_{b_n}/a_n \to 0$.

The normalization sequences $\{a_n : n \geq 1\}$ and $\{d_n : n \geq 1\}$ in Assumption BB2 may depend on $\{\gamma_{n,h} : n \geq 1\}$. (For notational simplicity, this dependence is suppressed.) For example, this occurs when the observations may be stationary or nonstationary depending on the value of γ. In particular, it occurs in an autoregressive model with a root that is less than or equal to one. In a model with iid or stationary strong mixing observations, one often takes $d_n = 1$ for all n, W_h is a pointmass distribution with pointmass at the probability limit of $\widehat{\sigma}_n$, and $a_n = n^{1/2}$. MAY NEED TO DELETE THE FOLLOWING SENTENCE IF IT TURNS OUT THAT ASSUMPTION HH CANNOT BE CHANGED. REASON: IF a_n IS DEFINED AS IN THE FOLLOWING SENTENCE, THEN ASSUMPTION HH WILL FAIL FOR SOME SEQUENCES OF σ_n VALUES. Alternatively, in a such a model one can take $d_n = 1/\sigma_n$ and $a_n = n^{1/2}/\sigma_n$, where σ_n is the population analogue of $\widehat{\sigma}_n$, and W_h is a pointmass distribution at one. This is useful to handle cases in which $\sigma_n \to 0$ as $n \to \infty$.

Assumption BB2 implies Assumption B2 (by the continuous mapping theorem using Assumption BB2(ii)). Note that there is a certain redundancy of normalization constants in Assumption BB2. Without any loss of generality, one could absorb d_n into the definition of $\widehat{\sigma}_n$ and take $d_n = 1$ for all n. We do not do this for two reasons. First, if there is a conventional definition $\widehat{\sigma}_n$, then this may preclude its use. Second, it is convenient to keep the assumptions as close as possible to those of PRW.

Assumption DD implies Assumption D. Assumption DD is not restrictive given the standard methods of defining subsamples. Assumption EE holds automatically for iid observations and for stationary strong mixing observations under the condition in (3.3) when the subsamples are constructed as described in Section 3 (for the same reason that Assumption E holds in these cases). Assumption HH holds in many examples when Assumption C holds, as is typically the case. However, it does not hold if θ is unidentified when $\gamma = 0$ (because consistent estimation of θ is not possible in this case and $a_n = 1$ in Assumption BB2(i)). For example, this occurs in a model with weak instruments.

The following Lemma generalizes Lemma 1. It does not impose Assumption t2.

Lemma 5 *Assumptions* t1, Sub1, A2, BB2, C, DD, EE, *and* HH *imply Assumption* G2.

Comments. 1. Given Lemma 5, the result of Theorem 2(b) holds for studentized t statistics under Assumptions t1, Sub1, A2, BB2, C, DD, E, EE, F2, and HH. These Assumptions imply Assumptions B2, D, and G2.

2. The proof of Lemma 5 is a variant of the proofs of Theorems 11.3.1(i) and 12.2.2(i) of PRW to allow for nuisance parameters $\{\gamma_{n,h} : n \geq 1\}$ that vary with n and t statistics that may be one- or two-sided.[13]

11 Examples

11.1 Testing with a Nuisance Parameter on the Boundary

Here we consider a testing problem where a nuisance parameter may be on the boundary of the parameter space under the null hypothesis.

Suppose $\{X_i \in R^2 : i \leq n\}$ are iid with distribution F,

$$X_i = \begin{pmatrix} X_{i1} \\ X_{i2} \end{pmatrix},\ E_F X_i = \begin{pmatrix} \theta \\ \mu \end{pmatrix},\ \text{and } Var_F(X_i) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}. \tag{11.2}$$

The null hypothesis is

$$H_0 : \theta = 0, \tag{11.3}$$

i.e., $\theta_0 = 0$. (The results below are invariant to the choice of θ_0.) The parameter space for the nuisance parameter μ is R_+.

We consider lower and upper one-sided tests and symmetric and equal-tailed two-sided tests of nominal level α. Each test is based on a studentized test statistic $T_n(\theta_0)$, where $T_n(\theta_0) = T_n^*(\theta_0), -T_n^*(\theta_0)$, or $|T_n^*(\theta_0)|$, $T_n^*(\theta_0) = \tau_n(\widehat{\theta}_n - \theta_0)/\widehat{\sigma}_{n1}$ and $\tau_n = n^{1/2}$. The estimators $(\widehat{\theta}_n, \widehat{\sigma}_{n1})$ of (θ, σ_1) are defined below. We consider subsample, FCV, and hybrid versions of all of these tests. The FCV tests employ the usual standard normal critical values that ignore the fact that μ may be on the boundary.

The size properties of the tests are given in Table I (as described in more detail below) and are summarized as follows. For the one-sided tests, we find large size distortions for the subsample tests, very small size distortions for the FCV tests, and no size distortions for the hybrid tests for all nominal sizes $\alpha \in [.01, .2]$ that we consider. (Only results for $\alpha = .05$ are reported.) The upper (lower) one-sided subsample test over-rejects the null most when the correlation ρ is close to -1 (respectively, 1). Monte Carlo simulations of its asymptotic null rejection probabilities indicate that its asymptotic size equals 1/2 for all nominal sizes $\alpha \in [.01, .2]$ that we consider. The symmetric two-sided subsample, FCV, and hybrid tests are all size-distorted. The Monte Carlo simulations for $\alpha \in [.01, .2]$ suggest that $AsySz(\theta_0) = 2\alpha$ for all version of these tests,

[13] Lemma 5 does not assume $\tau_{b_n}/\tau_n \to 0$ (only Assumption HH), although PRW's results do. A careful reading of their proof reveals that the assumption $a_{b_n}/a_n \to 0$ is enough to show that $U_{n,b}(x)$ and $L_{n,b}(x)$ have the same probability limits.

so size correction is possible. Finally, for the two-sided equal-tailed tests, we find extreme size distortion for the subsample test, no size distortion for the hybrid test (up to simulation error), and $AsySz(\theta_0)$ is approximately 2α for the FCV test.

We now define the estimators $(\widehat{\theta}_n, \widehat{\sigma}_{n1})$ of (θ, σ_1) used in the test statistic $T_n^*(\theta_0)$. Let $\widehat{\sigma}_{n1}, \widehat{\sigma}_{n2}$, and $\widehat{\rho}_n$ denote any consistent estimators of σ_1, σ_2, and ρ. Define $(\widehat{\theta}_n, \widehat{\mu}_n)$ to be the Gaussian quasi-ML estimator of (θ, μ) under the restriction that $\widehat{\mu}_n \geq 0$ and under the assumption that the standard deviations and correlation of X_i equal $\widehat{\sigma}_{n1}, \widehat{\sigma}_{n2}$, and $\widehat{\rho}_n$. This allows for the case where $(\widehat{\theta}_n, \widehat{\mu}_n, \widehat{\sigma}_{n1}, \widehat{\sigma}_{n2}, \widehat{\rho}_n)$ is the Gaussian quasi-ML estimator of $(\theta, \mu, \sigma_1, \sigma_2, \rho)$ under the restriction $\widehat{\mu}_n \geq 0$. Alternatively, $\widehat{\sigma}_{n1}, \widehat{\sigma}_{n2}$, and $\widehat{\rho}_n$ could be the sample standard deviations and correlation of X_{i1} and X_{i2}. A Kuhn-Tucker maximization shows that

$$\begin{aligned} \widehat{\theta}_n &= \overline{X}_{n1} - (\widehat{\rho}_n \widehat{\sigma}_{n1}) \min(0, \overline{X}_{n2}/\widehat{\sigma}_{n2}), \text{ where} \\ \overline{X}_{nj} &= n^{-1} \textstyle\sum_{i=1}^n X_{ij} \text{ for } j = 1, 2. \end{aligned} \tag{11.4}$$

Define the vector of nuisance parameters $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ by $\gamma_1 = \mu/\sigma_2$, $\gamma_2 = \rho$, and $\gamma_3 = (\sigma_1, \sigma_2, F)$. Let $r = 1/2$. In Assumption A2, set $\Gamma_1 = R_+$, $\Gamma_2 = (-1, 1)$, and $\Gamma_3(\gamma_1, \gamma_2) = (0, \infty) \times (0, \infty) \times \mathcal{F}(\mu, \rho, \sigma_1, \sigma_2)$, where

$$\begin{aligned} \mathcal{F}(\mu, \rho, \sigma_1, \sigma_2) = \{&F : E_F ||X_i||^{2+\delta} \leq M,\ E_F X_i = (0, \mu)',\ Var_F(X_{i1}) = \sigma_1^2, \\ &Var_F(X_{i2}) = \sigma_2^2,\ \&\ Corr_F(X_{i1}, X_{i2}) = \rho\} \end{aligned} \tag{11.5}$$

for some $M < \infty$ and $\delta > 0$.[14] Then, $H = R_{+,\infty} \times [-1, 1]$.

The following results are all under the null hypothesis, so the true parameter θ equals zero. For any $h = (h_1, h_2) \in H$ with $h_1 < \infty$ and any sequence $\{\gamma_{n,h} : n \geq 1\}$ of true parameters, consistency of $(\widehat{\sigma}_{n1}, \widehat{\sigma}_{n2}, \widehat{\rho}_n)$ and the CLT imply

$$n^{1/2}(\overline{X}_{n1}/\widehat{\sigma}_{n1}, \overline{X}_{n2}/\widehat{\sigma}_{n2}) \to_d (0, h_1)' + Z_{h_2}, \tag{11.6}$$

where $Z_{h_2} = (Z_{h_2,1}, Z_{h_2,2})' \sim N(0, V_{h_2})$ and V_{h_2} is a 2×2 matrix with diagonal elements 1 and off-diagonal elements h_2. (For this and the results below, we assume that $\widehat{\sigma}_{n1}$, $\widehat{\sigma}_{n2}$, and $\widehat{\rho}_n$ are consistent in the sense that $\widehat{\sigma}_{nj}/\sigma_{j,n,h} \to_p 1$ for $j = 1, 2$ and $\widehat{\rho}_n - \rho_{n,h} \to_p 0$ under $\{\gamma_{n,h} = (\theta_{n,h}, \rho_{n,h}, (\sigma_{1,n,h}, \sigma_{2,n,h}, F_{n,h})) : n \geq 1\}$.) By the continuous mapping theorem, we obtain

$$T_n^*(\theta_0) = n^{1/2} \widehat{\theta}_n / \widehat{\sigma}_{n1} \to_d J_h^* \text{ under } \{\gamma_{n,h}\}, \tag{11.7}$$

where J_h^* is defined by

$$Z_{h_2,1} - h_2 \min(0, Z_{h_2,2} + h_1) \sim J_h^*. \tag{11.8}$$

For $h \in H$ with $h_1 = \infty$, we have $\widehat{\theta}_n = \overline{X}_{n1}$ wp$\to$1 under $\{\gamma_{n,h}\}$ because $n^{1/2}\overline{X}_{n2}/\widehat{\sigma}_{n2}$ $\to_p \infty$ under $\{\gamma_{n,h}\}$. (The latter holds because $n^{1/2}\gamma_{n,h,1} = n^{1/2}\mu_n/\sigma_{n2} \to \infty$, $n^{1/2}(\overline{X}_{n2}$

[14] The condition $E_F||X_i||^{2+\delta} \leq M$ in $\mathcal{F}(\mu, \rho, \sigma_1, \sigma_2)$ ensures that the Liapunov CLT applies in (11.6)-(11.10) below. In $\mathcal{F}(\mu, \rho, \sigma_1, \sigma_2)$, $E_F X_{i1} = 0$ because the results given are all under the null hypothesis.

$-E\overline{X}_{n2})/\widehat{\sigma}_{n2} = O_p(1)$ by the CLT and $\widehat{\sigma}_{n2}/\sigma_{n2} \to_p 1$, and $n^{1/2}E\overline{X}_{n2}/\widehat{\sigma}_{n2} = n^{1/2}\mu_n/\widehat{\sigma}_{n2}$ $\to_p \infty$.) Therefore, under $\{\gamma_{n,h}\}$ with $h_1 = \infty$, we have

$$T_n^*(\theta_0) \to_d J_\infty^*, \text{ where } J_\infty^* \text{ is the } N(0,1) \text{ distribution.} \tag{11.9}$$

Note that the limit distributions J_h^* and J_∞^* do not depend on $\gamma_3 = (\sigma_1^2, \sigma_2^2, F)$.

For $T_n(\theta_0) = T_n^*(\theta_0), -T_n^*(\theta_0)$, and $|T_n^*(\theta_0)|$, we have

$$T_n(\theta_0) \to J_h \text{ under } \{\gamma_{n,h}\}, \text{ where } J_h = J_h^*, -J_h^*, \text{ and } |J_h^*|. \tag{11.10}$$

(If $Y \sim J_h^*$, then by definition, $-Y \sim -J_h^*$ and $|Y| \sim |J_h^*|$.) Note that J_h^* is stochastically increasing (decreasing) in h_1 for $h_2 < 0$ ($h_2 \geq 0$). Likewise, $-J_h^*$ is stochastically decreasing (increasing) in h_1 for $h_2 < 0$ ($h_2 \geq 0$).

The critical values for FCV tests are given by $z_{1-\alpha}$, $z_{1-\alpha}$, and $z_{1-\alpha/2}$, respectively, for the upper, lower, and symmetric versions. For subsample tests, critical values are given by $c_{n,b}(1-\alpha)$ obtained from the subsample statistics $\{T_{n,b,i}(\widehat{\theta}_n) : i \leq q_n\}$, defined in (2.6) (i)-(iii) for upper, lower, and two-sided symmetric tests, respectively.[15] For the hybrid tests, critical values are given by $\max\{c_{n,b}(1-\alpha), z_{1-\alpha}\}$ for the upper and lower one-sided tests and by $\max\{c_{n,b}(1-\alpha), z_{1-\alpha/2}\}$ for the symmetric two-sided test. For equal-tailed two-sided tests described in (8.1), critical values $(c_{\alpha/2}, c_{1-\alpha/2})$ for FCV, subsample, and hybrid tests are given by $(z_{\alpha/2}, z_{1-\alpha/2})$, $(c_{n,b}(\alpha/2), c_{n,b}(1-\alpha/2))$, and $(\min\{c_{n,b}(\alpha/2),\, z_{\alpha/2}\},\, \max\{c_{n,b}(1-\alpha/2), z_{1-\alpha/2}\})$, respectively.

Below we verify Assumptions A2-G2, J, and N2. By Theorem 2 (a)-(b), Corollary 1(b), and continuity of the distribution functions J_h, the asymptotic size of the lower and upper one-sided and symmetric two-sided tests is given by $AsySz(\theta_0) = Max_{Type}(\alpha)$ for $Type = Fix$, Sub, and Hyb. By Corollary 3(d)-(f), the asymptotic size of the two-sided equal-tailed versions of the test is given by $AsySz(\theta_0) = Max^{r-}_{ET,Type}(\alpha)$ in (8.4) for $Type = Fix$, Sub, and Hyb.

We first discuss upper one-sided tests. Given that $J_h = J_h^*$ is stochastically increasing (decreasing) in h_1 for fixed $h_2 < 0$ ($h_2 \geq 0$), we can show that

$$\begin{aligned}
Max_{Fix}(\alpha) &= \sup_{h \in H}[1 - J_h(c_{Fix}(1-\alpha))] = \sup_{h_2 \in [0,1]}(1 - J_{(0,h_2)}(z_{1-\alpha})),\\
Max_{Sub}(\alpha) &= \sup_{(g,h) \in GH}[1 - J_h(c_g(1-\alpha))] = \sup_{h_2 \in [-1,0]}(1 - J_\infty(c_{(0,h_2)}(1-\alpha))),\\
Max_{Hyb}(\alpha) &= \alpha.
\end{aligned} \tag{11.11}$$

The results for lower one-sided tests are analogous with $h_2 \in [0,1]$ and $h_2 \in [-1,0]$ replaced by $h_2 \in [-1,0]$ and $h_2 \in [0,1]$, respectively. To obtain the expression for $Max_{Sub}(\alpha)$ in (11.11), we use

$$\begin{aligned}
&\inf_{(g,h) \in GH: h_2 \in [0,1]} J_h(c_g(1-\alpha))\\
&= \min\{\inf_{h_1 \in [0,\infty), h_2 \in [0,1]} J_h(c_{(0,h_2)}(1-\alpha)), \inf_{h_1 \in [0,\infty], h_2 \in [0,1]} J_\infty(c_{(h_1,h_2)}(1-\alpha))\}\\
&= \min\{1-\alpha, 1-\alpha\} = 1-\alpha,
\end{aligned} \tag{11.12}$$

[15] The same results also hold under Assumption Sub2.

because $J_h = J_h^*$ is stochastically decreasing in h_1 for fixed $h_2 \geq 0$.

The values of $Max_{Fix}(\alpha)$ and $Max_{Sub}(\alpha)$ for the upper and lower one-sided tests are obtained by simulation. (All simulation results are based on 50,000 simulation repetitions and when maximization over h_1 is needed the upper bound is 12 and a grid of size 0.05 is used.) Table I reports $1 - J_{(0,h_2)}(z_{1-\alpha})$ and $1 - J_\infty(c_{(0,h_2)}(1-\alpha))$ for various values of h_2 and $\alpha = .05$ for upper one-sided tests. Because the results for lower one-sided tests are the same, but with h_2 replaced by $-h_2$, the results for lower one-sided tests are not reported in Table I. Simulation of $AsySz(\theta_0)$ is very fast because the two-dimensional maximization over (h_1, h_2) has been reduced to a one-dimensional maximization over h_2 in (11.11).

For the symmetric two-sided case, where $J_h = |J_h^*|$, we have

$$\begin{aligned} Max_{Fix}(\alpha) &= \sup_{h \in H}[1 - J_h(z_{1-\alpha/2})], \\ Max_{Sub}(\alpha) &= \max\{ \sup_{h_1 \in [0,\infty), h_2 \in [-1,1]} [1 - J_h(c_{(0,h_2)}(1-\alpha))], \\ &\qquad \sup_{h_1 \in [0,\infty], h_2 \in [-1,1]} [1 - J_\infty(c_{(h_1,h_2)}(1-\alpha))]\}, \end{aligned} \tag{11.13}$$

and similarly for $Max_{Hyb}(\alpha)$ with $c_{(0,h_2)}(1-\alpha)$ and $c_h(1-\alpha)$ replaced by $\max\{c_{(0,h_2)}(1-\alpha), z_{1-\alpha/2}\}$ and $\max\{c_h(1-\alpha), z_{1-\alpha/2}\}$, respectively. We use Monte Carlo simulation to calculate those quantities. Table I reports the values of $\sup_{h_1 \in [0,\infty]}[1 - J_{(h_1,h_2)}(z_{1-\alpha/2})]$ appearing in $Max_{Fix}(\alpha)$ for a range of ρ $(= h_2)$ values in $[-1, 1]$ and $\alpha = .05$. Note that these are the maximum asymptotic null rejection probabilities given ρ $(= h_2)$, where the maximum is over h_1 with h_2 fixed. Table I also reports the analogous expressions that depend on ρ $(= h_2)$ for the subsample and hybrid tests. The Table strongly suggests that $AsySz(\theta_0) = 2\alpha$ for all three types of symmetric two-sided tests. This implies that the tests are size-distorted but that size correction is possible by taking the nominal size equal to $\alpha/2$.

Finally, for the equal-tailed two-sided tests, we calculate $Max^{r-}_{ET,Type}(\alpha)$ in (8.4) for $Type = Fix$, Sub, and Hyb. Table I reports the maximum asymptotic null rejection probabilities for these tests given ρ $(= h_2)$ for a range of ρ values in $[-1, 1]$ and $\alpha = .05$. (The maximum is over $h \in H$ or $(g, h) \in GH$ with h_2 fixed.) We find extreme size distortion for the equal-tailed subsample test, correct size for the hybrid test, and $AsySz(\theta_0) = 2\alpha$ for the FCV test.

We now verify Assumptions A2-G2. Assumptions A2, C, and D clearly hold. Assumption B2 follows immediately from (11.7), (11.8), and (11.9) with J_h equal to $J_h^*, -J_h^*$, and $|J_h^*|$, respectively, for upper, lower, and symmetric tests. Assumption E holds by the general argument given in Section 3. For all $h \in H$, the distribution functions $J_h(x)$ are continuous for $x > 0$ and increasing at all of their quantiles $c_h(1-\alpha)$ for $\alpha < 1/2$. This establishes Assumptions F2 and N2 and shows that $Max^-_{Type}(\alpha) = Max_{Type}(\alpha)$ for any $\alpha < 1/2$ for $Type = Fix$, Sub, and Hyb. Assumption G2 follows by Lemma 5 noting that Assumptions BB2 and HH hold with $a_n = n^{1/2}/\sigma_{1,n,h}$, $d_n = 1/\sigma_{1,n,h}$, $\tau_n = n^{1/2}$, $V_h = J_h$, and W_h equal to pointmass at one. NEED TO FIX THE LAST STATEMMENT. ASSUMPTION HH DOES NOT NECESSARILY HOLD IF $a_n = n^{1/2}/\sigma_{1,n,h}$ BECAUSE $\sigma_{1,n,h}/\sigma_{1,b_n,h}$ COULD BE ILL-BEHAVED.

11.2 CI Based on a Super-Efficient Estimator

Here we show that the standard FCV, subsample, and hybrid CIs (of nominal level $1-\alpha$ for any $\alpha \in (0,1)$) based on a super-efficient estimator have $AsyCL = 0$ in a very simple regular model.

The model is

$$X_i = \theta + U_i, \text{ where } U_i \sim iid\ F \text{ for } i = 1, ..., n, \tag{11.1}$$

where F is some distribution with variance one. The super-efficient estimator, $\widehat{\theta}_n$, of θ and the test statistic, $T_n(\theta_0)$, are defined by

$$\begin{aligned} \widehat{\theta}_n &= \begin{cases} \overline{X}_n & \text{if } |\overline{X}_n| > n^{-1/4} \\ a\overline{X}_n & \text{if } |\overline{X}_n| \le n^{-1/4}, \end{cases} \text{ where } \overline{X}_n = n^{-1}\sum_{i=1}^n X_i, \\ T_n(\theta_0) &= |n^{1/2}(\widehat{\theta}_n - \theta_0)|, \end{aligned} \tag{11.2}$$

and $0 \le a < 1$. The constant a is a tuning parameter that determines the magnitude of shrinkage. The test statistic is a two-sided non-studentized t statistic, so that Assumptions t1(iii) and t2 hold with $\tau_n = n^{1/2}$.

The CI for θ is given by the third equation in (9.2) with $c_{1-\alpha}$ equal to the standard normal $1-\alpha/2$ quantile for the FCV CI: $c_{Fix}(1-\alpha) = z_{1-\alpha/2}$. For the subsample CI, $c_{1-\alpha}$ is equal to the subsample critical value $c_{n,b}(1-\alpha)$ based on subsample statistics $\{T_{n,b_n,i}(\widehat{\theta}_n) : i = 1, ..., q_n\}$ defined in equation (iii) of (2.6) with $\widehat{\sigma}_{n,b,i} = 1$. Note that Assumption Sub1 holds. (The results given below also hold if Assumption Sub2 holds.) For the hybrid CI, $c_{1-\alpha}$ is equal to the maximum of $c_{n,b}(1-\alpha)$ and $c_\infty(1-\alpha) = z_{1-\alpha/2}$. We take $\{b_n : n \ge 1\}$ so that Assumption C holds.

We apply Corollary 4(a)-(c) with $\gamma = \gamma_1 = \theta = \theta_1 \in R$, $p = d = 1$, and $\Gamma = \Gamma_1 = \Theta = R$. (No γ_2, γ_3, θ_2, or η parameters appear in this example.) The assumptions of Corollary 4(a)-(c) are verified below. We take $r = 1/2$ and $\gamma_{n,h}$ $(= \theta_{n,h}) = hn^{-1/2}$, where $h \in R$, in Assumption B1. When the true value is $\theta_{n,h}$, $\widehat{\theta}_n = a\overline{X}_n$ with probability that goes to one as $n \to \infty$ (wp$\to$ 1), see (11.7) below. Hence, wp$\to$ 1, we have

$$\begin{aligned} T_n(\theta_{n,h}) &= |n^{1/2}(a\overline{X}_n - \theta_{n,h})| \\ &= |an^{1/2}(\overline{X}_n - \theta_{n,h}) + (a-1)h| \\ &\to_d |aZ + (a-1)h| \sim J_h, \text{ where } Z \sim N(0,1) \text{ and} \\ J_h(x) &= \begin{cases} \Phi(a^{-1}(x + (1-a)h)) - \Phi(a^{-1}(-x + (1-a)h)) & \text{if } a \in (0,1) \\ 1(x \ge |h|) & \text{if } a = 0 \end{cases}, \end{aligned} \tag{11.3}$$

using the central limit theorem. Given that $p = d = 1$, we have $h^0 = 0$ and $J_{h^0} = J_0$. For $a \in (0,1)$, we have $J_0(x) = \Phi(a^{-1}x) - \Phi(-a^{-1}x)$ and $c_{h^0}(1-\alpha) = c_0(1-\alpha) = az_{1-\alpha/2}$. For $a = 0$, $J_0(x) = 1(x \geq 0)$ and $c_{h^0}(1-\alpha) = c_0(1-\alpha) = 0$.

For $a \in (0,1)$, Corollary 4(a)-(c) implies that the limits of the coverage probabilities of the FCV, subsample, and hybrid CIs under $\gamma_{n,h}$ $(= \theta_{n,h}) = hn^{-1/2}$ are

$$\begin{aligned} J_h(c_{Fix}(1-\alpha)) &= J_h(z_{1-\alpha/2}), \\ J_h(c_{h^0}(1-\alpha)) &= J_h(az_{1-\alpha/2}), \text{ and} \\ J_h(\max\{c_{h^0}(1-\alpha), c_\infty(1-\alpha)\}) &= J_h(\max\{az_{1-\alpha/2}, z_{1-\alpha/2}\}) = J_h(z_{1-\alpha/2}), \end{aligned} \tag{11.4}$$

respectively. Using (11.3), for $a \in (0,1)$ and $\xi = 1$ or $\xi = a$, we have

$$\lim_{h \to \infty} J_h(\xi z_{1-\alpha/2}) = 0. \tag{11.5}$$

Hence, for $a \in (0,1)$ and h sufficiently large, the asymptotic coverage probabilities of the symmetric two-sided FCV, subsample, and hybrid CIs is arbitrarily close to zero. Since $h \in R$ is arbitrary, this implies that $AsyCL = 0$ for these CIs.

For $a = 0$, the limits of the coverage probabilities of the FCV, subsample, and hybrid CIs under $\gamma_{n,h}$ $(= \theta_{n,h}) = hn^{-1/2}$ are

$$\begin{aligned} J_h(c_{Fix}(1-\alpha)) &= J_h(z_{1-\alpha/2}) = 1(z_{1-\alpha/2} \geq |h|) = 0 \text{ for } |h| > z_{1-\alpha/2}, \\ J_h(c_{h^0}(1-\alpha)) &= J_h(0) = 1(0 \geq |h|) = 0 \text{ for } |h| > 0, \text{ and} \\ J_h(\max\{c_{h^0}(1-\alpha), c_\infty(1-\alpha)\}) &= J_h(\max\{0, z_{1-\alpha/2}\}) = 1(z_{1-\alpha/2} \geq |h|) = 0 \\ &\quad \text{for } |h| > z_{1-\alpha/2}, \end{aligned} \tag{11.6}$$

respectively. Hence, for $a = 0$, $AsyCL = 0$ for the FCV, subsample, and hybrid CIs.

We obtain the same result that $AsyCL = 0$ if one-sided CIs or equal-tailed two-sided CIs are considered. Furthermore, the size-correction methods of Section 6 do not work in this example because Assumptions K, L, and M fail. (For example, Assumption K fails when $a = 0$ because (i) $H = R_\infty$ in this example and (ii) for any constant cv, we have $\sup_{h \in H}[1 - J_h(cv)] = \sup_{h \in R}[1 - 1(cv \geq |h|)] = 1$.)

PERHAPS MOVE THE NEXT PART TO THE APPENDIX.

It remains to verify Assumptions A1, B1, C-E, F1, and G1 for arbitrary choice of the parameter h. (We need not verify Assumption J because the result of Corollary 4(c) does not actually require Assumption J, see Comment 2 to Corollary 1.) Assumption A1 holds because $0 \in \Gamma = R$, Assumption C holds by assumption, Assumptions D and E hold because the observations are iid for each fixed $\theta \in R$, Assumption H holds because $\tau_{b_n}/\tau_n = b_n^{1/2}/n^{1/2} \to 0$ by Assummption C, and Assumption G1 holds by Lemma 1(a) using Assumption H. For $a \in (0,1)$, Assumption F1 holds because $J_{h^0}(x) = \Phi(a^{-1}x) - \Phi(-a^{-1}x)$ is strictly increasing at $c_{h^0}(1-\alpha) = az_{1-\alpha/2}$. For $a = 0$, Assumption F1 holds because $J_{h^0}(x) = 1(x \geq 0)$ has a jump at $x = c_{h^0}(1-\alpha) = 0$ with $J_{h^0}(c_{h^0}(1-\alpha)) = 1 > 1 - \alpha$ and $J_{h^0}(c_{h^0}(1-\alpha)-) = 0 < 1 - \alpha$.

Next, we verify Assumption B1. For any true sequence $\{\gamma_n : n \geq 1\}$ for which $n^{1/2}\gamma_n$ $(= n^{1/2}\theta_n) = O(1)$, we have

$$\begin{aligned} P_{\gamma_n}(|\overline{X}_n| \leq n^{-1/4}) &= P_{\gamma_n}(|n^{1/2}(\overline{X}_n - \theta_n) + n^{1/2}\theta_n| \leq n^{1/4}) \\ &= P_{\gamma_n}(|O_p(1) + O(1)| \leq n^{1/4}) \to 1 \text{ and} \\ P_{\gamma_n}(\widehat{\theta}_n = a\overline{X}_n) &\to 1, \end{aligned} \tag{11.7}$$

where the second equality uses the fact that $n^{1/2}(\overline{X}_n - \theta_n)$ has mean zero and variance one and the second convergence result uses the definition of $\widehat{\theta}_n$ in (11.2). For the particular sequence $\gamma_{n,h}$ $(= \theta_{n,h}) = hn^{-1/2}$ in Assumption B1(i), (11.3) and (11.7) imply that Assumption B1(i) holds with $J_h(x)$ defined as above. For any sequence $\{\gamma_{n,0} : n \geq 1\}$ as in Assumption B1(ii), we have $\gamma_{n,0}$ $(= \theta_{n,0}) = o(n^{-1/2})$ since $r = 1/2$. Hence, $P_{\gamma_{n,0}}(\widehat{\theta}_n = a\overline{X}_n) \to 1$ by (11.7), $|n^{1/2}\theta_{n,0}| = o(1)$, and the same argument as in (11.3) but with $h = 0$ implies that Assumption B1(ii) holds with $J_{h^0}(x) = J_0(x)$ defined as above.

11.3 CI for a Restricted Regression Parameter

Here we consider a multiple linear regression model where one regression parameter θ $(\in R)$ is restricted to be non-negative. We consider a studentized t statistic based on the least squares estimator of θ that is censored to be non-negative. We show that lower one-sided, symmetric two-sided, and equal-tailed two-sided subsample CIs for θ based on the studentized t statistic do not have correct asymptotic coverage probability. In particular, these three nominal $1 - \alpha$ CIs have asymptotic confidence levels of $1/2$, $1 - 2\alpha$, and $1/2 - \alpha/2$, respectively. Size-correction (of the type considered above) is possible for the symmetric subsample CI, but not for the lower one-sided or equal-tailed two-sided subsample CIs. We also show that upper and lower one-sided, symmetric two-sided, and equal-tailed two-sided FCV and hybrid CIs have correct $AsyCL$.

Consider the linear model with dependent variable $Y_i \in R$ and regressors $X_i \in R^k$ and $Z_i \in R$:

$$Y_i = X_i'\beta + Z_i\theta + U_i \tag{11.8}$$

for $i = 1, ..., n$. Assume $\{(U_i, X_i, Z_i) : i \geq 1\}$ are iid with distribution F and satisfy $E_F U_i^2 = \sigma_U^2 > 0$ and $E_F U_i(X_i', Z_i) = 0$. We also assume conditional homoskedasticity, that is, $E_F U_i^2 (X_i', Z_i)'(X_i', Z_i) = \sigma_U^2 Q_F$, where $Q_F = E_F(X_i', Z_i)'(X_i', Z_i) > 0$. Decompose Q_F into matrices Q_{XX}, Q_{XZ}, Q_{ZX}, and Q_{ZZ} in the obvious way. Denote by $Y, Z, U \in R^n$ and $X \in R^{n \times k}$ the matrices with rows Y_i, Z_i, U_i, and X_i', respectively, for $i = 1, ..., n$.

We consider inference concerning the parameter θ when the parameter space for θ is R_+ and that for β is R^k. Denote by $\widehat{\theta}_n$ the censored LS estimator of θ. That is,

$$\begin{aligned} \widehat{\theta}_n &= \max\{\widehat{\theta}_{LS}, 0\}, \text{ where} \\ \widehat{\theta}_{LS} &= (Z'M_X Z)^{-1} Z'M_X Y \text{ and } M_X = I - X(X'X)^{-1}X'. \end{aligned} \tag{11.9}$$

We have

$$n^{1/2}(\widehat{\theta}_n - \theta) = \max\{n^{1/2}(Z'M_XZ)^{-1}Z'M_XU, \ -n^{1/2}\theta\}. \quad (11.10)$$

By the law of large numbers and the CLT,

$$\begin{aligned} & n^{1/2}(Z'M_XZ)^{-1}Z'M_XU \to_d \zeta_\eta \sim N(0,\eta^2), \text{ where} \\ & \eta^2 = \sigma_U^2(Q_{ZZ} - Q_{ZX}Q_{XX}^{-1}Q_{XZ})^{-1}, \end{aligned} \quad (11.11)$$

under F. Denote by $\widehat{\eta}_n$ the consistent estimator of η that replaces population moments by sample averages and σ_U^2 by

$$\widehat{\sigma}_U^2 = \sum_{i=1}^n \widehat{U}_i^2/n, \text{ where } \widehat{U}_i = Y_i - X_i'\widehat{\beta}_n - Z_i\widehat{\theta}_n \quad (11.12)$$

and $(\widehat{\beta}_n, \widehat{\theta}_n)$ are the LS estimators of (β, θ) subject to the restriction $\theta \geq 0$.

Consider sequences $\{\gamma_{n,h} = (\gamma_{n,h,1}, \gamma_{n,h,2}, \gamma_{n,h,3}) = (\theta_{n,h}, \eta_{n,h}, (\beta_{n,h}, F_{n,h})) : n \geq 1\}$ of true parameters $(\theta, \eta, (\beta, F))'$ that satisfy $h_1 = \lim_{n\to\infty} n^{1/2}\theta_{n,h}$, $h_2 = \lim_{n\to\infty} \eta_{n,h}$, $\beta_{n,h} \in R^k$, and $F_{n,h} \in \mathcal{F}(\eta_{n,h})$ (defined below) for all $n \geq 1$. Let $h = (h_1, h_2) \in R_{+,\infty} \times [\eta_L, \eta_U]$ for some $0 < \eta_L < \eta_U < \infty$. Under a sequence $\{\gamma_{n,h}\}$, the Liapunov CLT, the continuous mapping theorem (CMT), and standard asymptotic calculations imply that the t statistic $T_n^*(\theta_0) = n^{1/2}(\widehat{\theta}_n - \theta_0)/\widehat{\eta}_n$ satisfies

$$T_n^*(\theta_{n,h}) \to_d \max\{\zeta, -h_1/h_2\}, \text{ where } \zeta = \zeta_\eta/\eta \sim N(0,1). \quad (11.13)$$

Note that there are no sequences $\{\theta_{n,h}\}$ of true parameters θ for which $h_1 < 0$. The distribution of $\max\{\zeta, -h_1/h_2\}$ depends on h only through h_1/h_2. Define the distribution J_h^* by

$$\max\{\zeta, -h_1/h_2\} \sim J_h^*. \quad (11.14)$$

As defined, J_h^* is standard normal when $h_1 = \infty$. When $h_1 = \infty$, we also write J_∞^* for J_h^*.

For $T_n(\theta_0) = T_n^*(\theta_0)$, $-T_n^*(\theta_0)$, and $|T_n^*(\theta_0)|$, we have

$$T_n(\theta_{n,h}) \to_d J_h, \text{ where } J_h = J_h^*, \ -J_h^*, \text{ and } |J_h^*|, \quad (11.15)$$

respectively, using the CMT. (Here $-J_h^*$ and $|J_h^*|$ denote the distributions of $-S$ and $|S|$ when $S \sim J_h^*$.) The dfs of J_h^*, $-J_h^*$, and $|J_h^*|$ are given by

$$\begin{aligned} & J_h^*(x) = \begin{cases} 0 & \text{for } x < -h_1/h_2 \\ \Phi(x) & \text{for } x \geq -h_1/h_2 \end{cases}, \quad (-J_h^*)(x) = \begin{cases} \Phi(x) & \text{for } x < h_1/h_2 \\ 1 & \text{for } x \geq h_1/h_2, \end{cases}, \text{ and} \\ & |J_h^*|(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ 2\Phi(x) - 1 & \text{for } 0 < x < h_1/h_2 \\ \Phi(x) & \text{for } x \geq h_1/h_2 \end{cases}, \end{aligned} \quad (11.16)$$

where $\Phi(x)$ is the standard normal df. A key property of J_h^* for the asymptotic properties of subsample CIs is that J_h^* is stochastically decreasing in h_1/h_2 and $-J_h^*$ and $|J_h^*|$ are stochastically increasing in h_1/h_2.

We consider upper and lower one-sided and symmetric and equal-tailed two-sided $1-\alpha$ CIs for θ for $\alpha < 1/2$. First, we discuss construction of these CIs. For upper and lower one-sided and symmetric two-sided CIs, FCV CIs for θ are given by the three equations of (9.2) with $c_{1-\alpha} = z_{1-\alpha}$, $c_{1-\alpha} = z_{1-\alpha}$, and $c_{1-\alpha} = z_{1-\alpha/2}$, respectively, $\widehat{\sigma}_n = \widehat{\eta}_n$, and $\tau_n = n^{1/2}$. For subsample CIs, $c_{1-\alpha}$ in (9.2) equals $c_{n,b}(1-\alpha)$ obtained from subsample statistics $T_{n,b,i}(\widehat{\theta}_n)$, defined in equations (2.6) (i)-(iii) for upper and lower one-sided and two-sided symmetric CIs, respectively.[16] For the hybrid CIs, we take $c_{1-\alpha} = \max\{c_{n,b}(1-\alpha), z_{1-\alpha}\}$ for the upper and lower one-sided CI and $c_{1-\alpha} = \max\{c_{n,b}(1-\alpha), z_{1-\alpha/2}\}$ for the symmetric two-sided CI.

The equal-tailed two-sided FCV, subsample, and hybrid CIs are defined in (9.5) with $(c_{\alpha/2}, c_{1-\alpha/2}) = (z_{\alpha/2}, z_{1-\alpha/2})$, $(c_{n,b}(\alpha/2), c_{n,b}(1-\alpha/2))$, and $(\min\{c_{n,b}(\alpha/2),\ z_{\alpha/2}\}, \max\{c_{n,b}(1-\alpha/2), z_{1-\alpha/2}\})$, respectively. Here the subsample quantiles $c_{n,b}(1-\alpha)$ are obtained from $T_{n,b,i}(\widehat{\theta}_n)$ (or $T_{n,b,i}(\theta_0)$) defined in equation (2.6) (i) with $\widehat{\sigma}_n = \widehat{\eta}_n$.

The parameter spaces for θ, η, and β are $\Theta = R_+$, $[\eta_L, \eta_U]$, and R^k, respectively. The parameter space for the distribution F of (U_i, X_i, Z_i) is

$$\begin{aligned}\mathcal{F}(\eta) = \{F : E_F|U_i|^{2+\delta} \leq M,\ E_F U_i^2 > 0,\ E_F U_i(X_i', Z_i) = 0,\ Q_F > 0, \\ E_F U_i^2 (X_i', Z_i)'(X_i', Z_i) = E_F U_i^2 Q_F,\ \sigma_U^2 (Q_{ZZ} - Q_{ZX} Q_{XX}^{-1} Q_{XZ})^{-1} = \eta^2\}\end{aligned} \tag{11.17}$$

for some $\delta > 0$ and $0 < M < \infty$. (The condition $E_F|U_i|^{2+\delta} \leq M$ in $\mathcal{F}(\eta)$ guarantees that the Liapunov CLT applies for sequences $\{\gamma_{n,h}\}$ in (11.13).) Hence, the parameter space for $\gamma = (\gamma_1, \gamma_2, \gamma_3) = (\theta, \eta, (\beta, F))$ is

$$\begin{aligned}\Gamma \ = \ & \{\gamma = (\gamma_1, \gamma_2, \gamma_3) = (\theta, \eta, (\beta, F)) : \gamma_1 = \theta \in R_+, \gamma_2 = \eta \in [\eta_L, \eta_U], \\ & \&\ \gamma_3 = (\beta, F) \in R^k \times \mathcal{F}(\eta)\}.\end{aligned} \tag{11.18}$$

We take $r = 1/2$ and $H = R_{+,\infty} \times [\eta_L, \eta_U]$.

We now apply Corollaries 4 and 5 to determine $AsyCL$ for each CI. Assumptions A2-G2 and N2 are verified below.

For upper one-sided CIs, $J_h(\cdot)$ $(= J_h^*(\cdot))$ is continuous at all $x > 0$ for all $h_1 \in R_{+,\infty}$ and $h_2 \in [\eta_L, \eta_U]$ using (11.16). Because the $1-\alpha$ quantile of J_h is positive for any $h \in H$ given $\alpha < 1/2$, the intervals for $AsyCL$ in Corollary 4(d)-(f) collapse to points. By Corollary 4(d)-(f), we find that the upper one-sided FCV, subsample, and hybrid CIs all have $AsyCL = 1-\alpha$ for $\alpha < 1/2$ because the $1-\alpha$ quantile of J_h for any $h \in H$ equals $z_{1-\alpha}$ using (11.16).

For the lower one-sided FCV CI, Corollary 4(d) implies that $AsyCL \in [\inf_{h\in H} J_h(z_{1-\alpha}-), \inf_{h\in H} J_h(z_{1-\alpha})]$. In this case, J_h $(= -J_h^*)$ is stochastically increasing in h_1/h_2. Hence, $\inf_{h\in H} J_h(z_{1-\alpha}) = J_\infty(z_{1-\alpha}) = \Phi(z_{1-\alpha}) = 1-\alpha$ using (11.16). Thus, $AsyCL = 1-\alpha$ for the lower one-sided FCV CI. For the lower one-sided hybrid CI, we have $\max\{c_g(1-\alpha), c_\infty(1-\alpha)\} = c_\infty(1-\alpha) = z_{1-\alpha}$ for all $g \in H$ because J_h $(= -J_h^*)$

[16] In the definition of $T_{n,b,i}(\widehat{\theta}_n)$ the role of $\widehat{\sigma}_{n,b,i}$ is played by the analogously defined expression $\widehat{\eta}_{n,b,i}$. The results also hold under Assumption Sub2 in which case the subsample statistics are $T_{n,b,i}(\theta_0)$.

is stochastically increasing in h_1/h_2. Hence, by Corollary 4(f) and the above result for the FCV CI, $AsyCL = 1 - \alpha$ for the lower one-sided hybrid CI.

For the lower one-sided subsample CI, Corollary 4(e) implies that $AsyCL \in [\inf_{(g,h)\in GH} J_h(c_g(1-\alpha)-), \inf_{(g,h)\in GH} J_h(c_g(1-\alpha))]$. We have

$$\inf_{(g,h)\in GH} J_h(c_g(1-\alpha)) = \inf_{h\in H} J_h(c_{h^0}(1-\alpha)) = \inf_{h\in H} J_h(0) = J_\infty(0) = 1/2, \quad (11.19)$$

where the first and third equalities hold because J_h $(= -J_h^*)$ is stochastically increasing in h_1/h_2, the second equality holds because $h^0 = (0, h_2)'$ and $J_{h^0}(x) = 1$ for all $x \geq 0$ using (11.16), and the last equality holds because $J_\infty(x) = \Phi(x)$ using (11.16). Therefore, $AsyCL = 1/2$ for the lower one-sided subsample CI.

We now discuss the results for symmetric two-sided CIs. By Corollary 4(d), we have $AsyCL \in [\inf_{h\in H} J_h(z_{1-\alpha/2}-), \inf_{h\in H} J_h(z_{1-\alpha/2})]$ for the FCV CI. Because J_h $(= |J_h^*|)$ is stochastically increasing in h_1/h_2 and using (11.16), we have $\inf_{h\in H} J_h(z_{1-\alpha/2}) = J_\infty(z_{1-\alpha/2}) = 2\Phi(z_{1-\alpha/2}) - 1 = 1 - \alpha$. Hence, $AsyCL = 1 - \alpha$ for the symmetric two-sided FCV CI. For the hybrid CI, we have $\max\{c_g(1-\alpha), c_\infty(1-\alpha)\} = c_\infty(1-\alpha) = z_{1-\alpha/2}$ for all $g \in H$ because J_h $(= |J_h^*|)$ is stochastically increasing in h_1/h_2 and using (11.16). Thus, using Corollary 4(f) and the above result for the FCV CI, we have $AsyCL = 1 - \alpha$ for the symmetric two-sided hybrid CI.

For the symmetric two-sided subsample CI, Corollary 4(e) implies that $AsyCL \in [\inf_{(g,h)\in GH} J_h(c_g(1-\alpha)-), \inf_{(g,h)\in GH} J_h(c_g(1-\alpha))]$. We have

$$\inf_{(g,h)\in GH} J_h(c_g(1-\alpha)) = \inf_{h\in H} J_h(c_{h^0}(1-\alpha)) = \inf_{h\in H} J_h(z_{1-\alpha}) = J_\infty(z_{1-\alpha}) = 1 - 2\alpha, \quad (11.20)$$

where the first and third equalities hold because J_h $(= |J_h^*|)$ is stochastically increasing in h_1/h_2, the second equality holds because $h^0 = (0, h_2)'$ and $J_{h^0}(x) = \Phi(x)$ for all $x \geq 0$ using (11.16), and the last equality holds because $J_\infty(x) = (|J_\infty^*|)(x) = 2\Phi(x) - 1$ using (11.16). Equation (11.20) holds with $c_g(1-\alpha)-$ in place of $c_g(1-\alpha)$. Hence, the (nominal $1-\alpha$) symmetric two-sided subsample CI has $AsyCL = 1 - 2\alpha$ and under-covers by α. An SC subsample CI can be constructed by taking $\xi(\alpha) = \alpha/2$.

Next, we discuss the results for the equal-tailed two-sided CIs. Here, $J_h = J_h^*$. By Corollary 5, $AsyCL \in [1 - Max_{ET,Type}^{\ell-}(\alpha), 1 - Max_{ET,Type}^{r-}(\alpha)]$ for $Type$ equal to Fix, Sub, and Hyb for the FCV, subsample, and hybrid CIs, respectively. For the FCV CI, $(c_{\alpha/2}, c_{1-\alpha/2}) = (z_{\alpha/2}, z_{1-\alpha/2})$ yields

$$\begin{aligned} Max_{ET,Fix}^{r-}(\alpha) &= \sup_{h\in H}[1 - J_h(z_{1-\alpha/2}) + J_h(z_{\alpha/2}-)] \\ &= \sup_{h\in H}[1 - \Phi(z_{1-\alpha/2}) + J_h(z_{\alpha/2}-)] = \alpha/2 + J_\infty(z_{\alpha/2}-) \\ &= \alpha/2 + \Phi(z_{\alpha/2}-) = \alpha, \end{aligned} \quad (11.21)$$

where the second equality holds by (11.16), the third equality holds because J_h $(= J_h^*)$ is stochastically decreasing in h_1/h_2, and the fourth equality holds by (11.16).

Analogously, $Max^{\ell-}_{ET,Type}(\alpha) = \alpha$. It follows that $AsyCL = 1 - \alpha$ for the equal-tailed FCV CI.

For the equal-tailed hybrid CI, the quantities $\max\{c_g(1-\alpha/2), c_\infty(1-\alpha/2)\}$ and $\min\{c_g(\alpha/2), c_\infty(\alpha/2)\}$ that appear in $Max^{r-}_{ET,Hyb}(\alpha)$ equal $z_{1-\alpha/2}$ and $z_{\alpha/2}$, respectively, because $c_g(1-\alpha/2) = z_{1-\alpha/2}$ for all $g \in H$ provided $\alpha \leq 1/2$ using (11.16) and $c_g(\alpha/2) \geq c_\infty(\alpha/2) = z_{\alpha/2}$ for all $g \in H$ using the fact that J_h $(= J^*_h)$ is stochastically decreasing in h_1/h_2 and (11.16). Hence, $Max^{r-}_{ET,Hyb}(\alpha) = Max^{r-}_{ET,Fix}(\alpha)$ and likewise with ℓ in place of r. In consequence, the result that $AsyCL = 1-\alpha$ for the equal-tailed FCV CI yields the same result for the equal-tailed hybrid CI.

Lastly, for the equal-tailed subsample CI, we have

$$\begin{aligned}
&Max^{r-}_{ET,Sub}(\alpha) = \sup_{(g,h)\in GH}[1 - J_h(z_{1-\alpha/2}) + J_h(c_g(\alpha/2)-)] \\
&= \sup_{(g,h)\in GH}[1 - \Phi(z_{1-\alpha/2}) + J_h(c_g(\alpha/2)-)] = \alpha/2 + \sup_{h\in H} J_h(0-) \\
&= \alpha/2 + J_\infty(0-) = \alpha/2 + \Phi(0) = \alpha/2 + 1/2, \qquad (11.22)
\end{aligned}$$

where the first equality holds because $c_g(1-\alpha/2) = z_{1-\alpha/2}$ for all $g \in H$ provided $\alpha \leq 1/2$ by (11.16), the second equality holds as in (11.21), the third equality holds because $c_g(\alpha/2) \leq 0$ with equality when $g = (0, h_2)'$ for $\alpha \leq 1/2$ using (11.16), the fourth equality holds because J_h $(= J^*_h)$ is stochastically decreasing in h_1/h_2, and the fifth equality holds because $J_\infty = J^*_\infty = \Phi$ using (11.16). Likewise, $Max^{\ell-}_{ET,Sub}(\alpha) = \alpha/2 + 1/2$. Therefore, $AsyCL = 1/2 - \alpha/2$ for the equal-tailed subsample CI. Clearly, size-correction (of the type discussed in the paper) is not possible here.

We now verify the assumptions needed to apply Corollaries 4 and 5. First, consider the case of an upper one-sided CI based on $T_n(\theta_0) = T^*_n(\theta_0)$. Assumption A2 holds by definition of Γ. Assumption B2 follows from (11.13). We choose $\{b_n : n \geq 1\}$ so that Assumption C holds. Assumption D holds by the iid assumption. Assumption E holds by the general argument given in Section 3. Assumption F2 holds because $J_h(x) = J^*_h(x)$ is strictly increasing for $x \geq 0$ and $c_h(1-\alpha) \geq 0$ for $\alpha \leq 1/2$ by (11.16). Assumption G2 follows by Lemma 5 under Assumption Sub1 and follows trivially under Assumption Sub2. The assumptions for Lemma 5 are verified as follows. Assumption BB2 holds with $(a_n, d_n) = (\tau_n, 1)$, where V_h is the distribution of $\max\{h_2\zeta, -h_1\}$, and W_h is a point mass distribution at $h_2 > 0$ under sequences $\gamma_{n,h} = (\theta_{n,h}, \eta_{n,h}, (\beta_{n,h}, F_{n,h}))'$ such that $h_1 \in R_{+,\infty}$ and $h_2 > 0$. Assumptions DD and EE hold by the same arguments as for Assumptions D and E. Assumption HH holds because $a_n = \tau_n = n^{1/2}$. The verification of the assumptions for the lower one-sided and two-sided cases is analogous with the exceptions of Assumptions F2 and N2. Using (11.16), one can verify that Assumption F2 holds for $J_h = -J^*_h$ and $J_h = |J^*_h|$ and Assumption N2 holds for $J_h = J^*_h$ because for all $h \in H$ either (i) $J_h(x)$ is strictly increasing at $x = c_h(1-\alpha)$ or (ii) $J_h(x)$ has a jump at $x = c_h(1-\alpha)$ with $J_h(c_h(1-\alpha)) > 1-\alpha$ and $J_h(c_h(1-\alpha)-) < 1-\alpha$ provided $\alpha \in (0,1)$.

11.4 Confidence Region Based on Moment Inequalities

Here we consider a confidence region (CR) for a true parameter θ_0 ($\in \Theta \subset R^d$) that is defined by moment inequalities and equalities. The true value need not be identified. The CR is obtained by inverting tests that are based on a generalized method of moments-type (GMM) criterion function. This method is introduced by Chernozhukov, Hong, and Tamer (2002) (CHT), who use subsampling to obtain a critical value.[17] Shaikh (2005) also considers this method and shows that the limit of finite sample size is the nominal level in one-sample and two-sample mean problems when a subsample critical value is employed. In this section, we show that this result holds quite generally for subsample CRs of this type—no specific assumptions concerning the form of the moment functions are necessary even though the asymptotic distribution of the test statistic is discontinuous (in the sense discussed above). Note that the results given here are for CRs for the true parameter, rather than for the "identified set."

We also consider CRs based on fixed and "plug-in" critical values—defined below. These critical values are bounded above as $n \to \infty$, whereas subsample critical values diverge to infinity at a rate that depends on the subsample size b_n. Hence, the use of fixed or plug-in critical values leads to more powerful tests and smaller CRs asymptotically than subsample critical values. Furthermore, the plug-in critical values (PCV) lead to more powerful tests and smaller CRs than the fixed critical values (FCV). Hence, the results here indicate that PCV is the best choice of critical value.

The test statistic considered below is similar to those considered by (i) Moon and Schorfheide (2004), who consider an empirical likelihood version of the GMM criterion function and assume identification of θ_0, (ii) Soares (2005), who allows for the plug-in of preliminary estimators in a GMM and/or empirical likelihood criterion function, and (iii) Rosen (2005), who considers a minimum distance version of the test statistic. By similar arguments to those given below, one can show that the limit of the finite sample size of a subsample CR based on any one of these test statistics equals its nominal level. The results for fixed and plug-in critical values also extend to these test statistics. Hence, for these versions of the test statistics as well, we find that the PCV has the best properties. For brevity, we only outline the arguments.

The model is as follows. The true value θ_0 ($\in \Theta \subset R^d$) is assumed to satisfy the following moment conditions:

$$\begin{aligned} E_F m_j(W_i, \theta_0) &\geq 0 \text{ for } j = 1, ..., p \text{ and} \\ E_F m_j(W_i, \theta_0) &= 0 \text{ for } j = p+1, ..., p+s, \end{aligned} \tag{11.23}$$

where $\{m_j(\cdot, \theta) : j = 1, ..., p+s\}$ are scalar-valued moment functions and $\{W_i : i \geq 1\}$ are stationary random vectors with joint distribution F.

[17] CHT focus on CRs for the identified set, rather than the true parameter. By definition, the identfied set, Θ_0, is the set of all $\theta \in \Theta$ that satisfy the moment inequalities and equalities when the true value is θ_0. Also, CHT consider a more general criterion function than that considered below. Their asymptotic results do not establish that the limit of the finite sample size of the CR for the true value is the nominal level, which is one of the results shown in this section.

The sample moment functions are

$$\overline{m}_{n,j}(\theta) = n^{-1} \sum_{i=1}^{n} m_j(W_i, \theta) \text{ for } j = 1, ..., p + s. \tag{11.24}$$

The test statistic that we consider for testing $H_0 : \theta = \theta_0$ is a Anderson–Rubin-type GMM statistic that gives positive weight to the moment inequalities only when they are violated:

$$T_n(\theta_0) = n \sum_{j=1}^{p} (\overline{m}_{n,j}(\theta_0)/\widehat{\sigma}_{n,j}(\theta_0))_-^2 + n \sum_{j=p+1}^{p+s} (\overline{m}_{n,j}(\theta_0)/\widehat{\sigma}_{n,j}(\theta_0))^2, \text{ where}$$
$$(x)_- = \begin{cases} x & \text{if } x < 0 \\ 0 & \text{if } x \geq 0 \end{cases} \tag{11.25}$$

and $\widehat{\sigma}_{n,j}^2(\theta_0)$ is a consistent estimator of $\sigma_{F,j}^2(\theta_0) = \lim_{n \to \infty} Var_F(n^{1/2}\overline{m}_{n,j}(\theta_0))$ for $j = 1, ..., p+s$. For example, with iid observations, one can define $\widehat{\sigma}_{n,j}^2(\theta_0) = n^{-1} \sum_{i=1}^{n} (m_j(W_i, \theta_0) - \overline{m}_{n,j}(\theta_0))^2$.

The subsample statistics are constructed such that Assumption Sub2 holds. (No consistent estimator of the true parameter exists when the latter is unidentified, so a subsample procedure that satisfies Assumption Sub1 is not available.)

We now specify $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ for this example. The moment conditions in (11.23) can be written as

$$\sigma_{F,j}^{-1}(\theta_0) E_F m_j(W_i, \theta_0) - \gamma_{1,j,0} = 0 \text{ for } j = 1, ..., p \text{ and}$$
$$\sigma_{F,j}^{-1}(\theta_0) E_F m_j(W_i, \theta_0) = 0 \text{ for } j = p + 1, ..., p + s \tag{11.26}$$

for some $\gamma_{1,0} = (\gamma_{1,1,0}, ..., \gamma_{1,p,0})' \in R_+^p$. Let $\Omega_0 = \lim_{n \to \infty} Corr_F(n^{1/2}\overline{m}_n(\theta_0))$, where $\overline{m}_n(\theta_0) = (\overline{m}_{n,1}(\theta_0), ..., \overline{m}_{n,p+s}(\theta_0))'$ and $Corr_F(n^{1/2}\overline{m}_n(\theta_0))$ denotes the $(p+s)\times(p+s)$ correlation matrix of $n^{1/2}\overline{m}_n(\theta_0)$. Let γ_1, Ω, and θ denote generic parameter values corresponding to the true null parameter values $\gamma_{1,0}$, Ω_0, and θ_0, respectively. We take $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ such that $\gamma_1 \in R_+^p$, $\gamma_2 = (\gamma_{2,1}', \gamma_{2,2}')' = (\theta', vech_*(\Omega)')' \in R^q$, where $vech_*(\Omega)$ denotes $vech(\Omega)$ with the diagonal elements of Ω deleted, $q = d + (p+s)(p+s-1)/2$, and $\gamma_3 = F$.

We take $r = 1/2$ and $h = (h_1, h_2)$, where $h_1 \in R_{+,\infty}^p$, $h_2 = (h_{2,1}', h_{2,2}')'$, $h_{2,1} \in \text{cl}(\Theta)$, $h_{2,2} \in \text{cl}(\Gamma_{2,2})$, and $\Gamma_{2,2}$ is some set of vectors $\gamma_{2,2}$ such that $\gamma_{2,2} = vech_*(C)$ for some $(p+s) \times (p+s)$ correlation matrix C. Hence, $H = R_{+,\infty}^p \times \text{cl}(\Theta) \times \text{cl}(\Gamma_{2,2})$. Note that h_1 corresponds to γ_1 and, hence, h_1 measures the extent to which the $j = 1, ..., p$ moment inequalities deviate from being equalities. Also, $h_{2,1}$ corresponds to θ and $h_{2,2}$ corresponds to $vech_*(\Omega)$.

The parameter spaces for γ_1 and γ_2 are $\Gamma_1 = R_+^p$ and $\Gamma_2 = \Theta \times \Gamma_{2,2}$, respectively. For given $(\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2$, the parameter space for F is

$$\mathcal{F}(\gamma_1, \gamma_2) = \{F : \sigma_{F,j}^{-1}(\theta) E_F m_j(W_i, \theta) - \gamma_{1,j} = 0 \text{ for } j = 1, ..., p, \tag{11.27}$$
$$E_F m_j(W_i, \theta) = 0 \text{ for } j = p + 1, ..., p + s\},$$

such that $\{\mathcal{F}(\gamma_1, \gamma_2) : (\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2\}$ satisfies the following "convergence condition." By definition, the convergence condition restricts $\{\mathcal{F}(\gamma_1, \gamma_2) : (\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2\}$ so that under any $\{\gamma_{n,h} = (\gamma_{n,h,1}, (\theta'_{n,h}, vech_*(\Omega_{n,h})')', F_{n,h}) : n \geq 1\}$ for any $h \in H$, we have

$$\begin{aligned}
&(A_{n,1}, ..., A_{n,p+s})' \to_d Z_{h_{2,2}} \sim N(0, V_{h_{2,2}}) \text{ as } n \to \infty, \text{ where} \\
&A_{n,j} = n^{1/2}(\overline{m}_{n,j}(\theta_{n,h}) - E_{F_{n,h}}\overline{m}_{n,j}(\theta_{n,h}))/\sigma_{F_{n,h},j}(\theta_{n,h}), \text{ and} \\
&\widehat{\sigma}_{nj}(\theta_{n,h})/\sigma_{F_{n,h},j}(\theta_{n,h}) \to_p 1 \text{ as } n \to \infty
\end{aligned} \tag{11.28}$$

for $j = 1, ..., p+s$, where $V_{h_{2,2}}$ is the $(p+s) \times (p+s)$ correlation matrix for which $vech_*(V_{h_{2,2}}) = h_{2,2}$. For example, if the observations $\{W_i : i \geq 1\}$ are iid under a fixed γ and $\widehat{\sigma}^2_{nj}(\theta)$ is defined as above, the convergence condition holds if $E_F|m_j(W_i, \theta)|^{2+\delta_1} \leq M$ and $\sigma_{F,j}(\theta) \geq \delta_2$ for $j = 1, ..., p+s$ for some constants $M < \infty$ and $\delta_1, \delta_2 > 0$ that do not depend on F. (This holds by straightforward calculations using the CLT and LLN for inid random vectors that satisfy a uniform $2 + \delta_1$ moment bound.) For dependent observations, one needs to specify a specific variance estimator $\widehat{\sigma}^2_{nj}(\theta)$, such as a HAC estimator, before a primitive convergence condition can be stated. For brevity, we do not do so here.

Given (11.25) and (11.28), Assumption B2 holds because under $\{\gamma_{n,h} : n \geq 1\}$ we have

$$T_n(\theta_{n,h}) \to_d \sum_{j=1}^{p}(Z_{h_{2,2},j} + h_{1,j})^2_- + \sum_{j=p+1}^{p+s} Z^2_{h_{2,2},j} \sim J_h \tag{11.29}$$

for all $h \in H$, where $Z_{h_{2,2}} = (Z_{h_{2,2},1}, ..., Z_{h_{2,2},p+s})'$. (Note that $Z_{h_{2,2},j'} + h_{1,j'} = 0$ for any j' in $\{1, ..., p\}$ for which $h_{1,j'} = \infty$.)

For $(g, h) \in GH$, we have

$$J_g \geq_{ST} J_h. \tag{11.30}$$

This holds because $\sum_{j=1}^{p}(Z_{h_{2,2},j} + g_{1,j})^2_- \geq_{ST} \sum_{j=1}^{p}(Z_{h_{2,2},j} + h_{1,j})^2_-$ a.s. for all $0 \leq g_1 \leq h_1$ due to the $(\cdot)_-$ function and, for all $(g, h) \in GH$, we have $0 \leq g_1 \leq h_1$ and $g_{2,2} = h_{2,2}$. Furthermore, J_h is continuous at its $1 - \alpha$ quantile for all $\alpha < 1/2$, see below. In consequence, by Corollary 4(e) and Comment 2 following Theorem 2, for the subsample CR, we have

$$AsyCL = \inf_{(g,h) \in GH} J_h(c_g(1-\alpha)) = 1 - \alpha. \tag{11.31}$$

Hence, in this example, discontinuity of the limit distribution does not cause size distortion for the subsample CR.

We now verify the remaining assumptions needed for Corollary 4. Assumption A2 holds with Γ defined as above. Assumption C holds by choice of b_n. Assumption D holds by stationarity and the standard definition of subsample statistics in the iid and dependent cases. Assumption E holds by the general argument given in Section 3 for iid observations and stationary strong mixing observations provided $\sup_{\gamma \in \Gamma} \alpha_\gamma(m) \to 0$ as

$m \to \infty$. Assumption F2 holds for all $\alpha < 1/2$ because if $s \geq 1$, then J_h is absolutely continuous, and if $s = 0$, then J_h has support R_+ and the df $J_h(x)$ has a jump of height greater than or equal to $1/2$ at $x = 0$ and no other jumps. Assumption G2 holds automatically because the subsample procedure satisfies Assumption Sub2.

Next, we discuss FCV CRs. Corollary 4(d), combined with the continuity results concerning J_h given in the discussion of Assumption F2 above and the result above that $J_g \geq_{ST} J_h$ for $(g, h) \in GH$, imply that for a FCV CR

$$AsyCL = \inf_{h \in H} J_h(c_{Fix}(1-\alpha)) = \inf_{h_2 \in H_2} J_{(0,h_2)}(c_{Fix}(1-\alpha)). \tag{11.32}$$

Hence, if $c_{Fix}(1-\alpha)$ is defined such that

$$\inf_{h_2 \in H_2} J_{(0,h_2)}(c_{Fix}(1-\alpha)) = 1-\alpha, \tag{11.33}$$

then the FCV CR has asymptotic level $1-\alpha$, as desired. By (11.29), the distribution $J_{(0,h_2)}$ only depends on

$$h_{2,2} = vech_*(\lim_{n \to \infty} Corr_{F_{n,h}}(\overline{m}_n(\theta_{n,h}))). \tag{11.34}$$

Hence, determination of the value $c_{Fix}(1-\alpha)$ that satisfies (11.33) only requires maximization over the possible asymptotic correlation matrices of $n^{1/2}\overline{m}_n(\theta_{n,h})$. For example if $p = 1$ and $s = 0$, then $J_{(0,h_2)}$ is the distribution of a random variable that is 0 with probability $1/2$ and is chi-squared with one degree of freedom with probability $1/2$. Hence, no unknown parameter appears. If $p+s = 2$, then $J_{(0,h_2)}$ depends on the scalar $h_{2,2}$, which is the asymptotic correlation between $n^{1/2}\overline{m}_{n,1}(\theta_{n,h})$ and $n^{1/2}\overline{m}_{n,2}(\theta_{n,h})$. For general $p+s$, one can determine $c_{Fix}(1-\alpha)$ such that the infimum in (11.32) equals $1-\alpha$ via simulation.

Given (11.32), one can design a data-dependent "plug-in" critical value (PCV) that yields a more powerful test than the FCV test and, hence, a smaller CR, because it is closer to being asymptotically similar. Let $c_{Plug}(h_{2,2}, 1-\alpha)$ denote the $1-\alpha$ quantile of $J_{(0,h_2)}(x)$ (which only depends on $h_{2,2}$). Let $\widehat{h}_{2,2,n}$ be a consistent estimator of $h_{2,2}$. The PCV is

$$c_{Plug}(\widehat{h}_{2,2,n}, 1-\alpha), \tag{11.35}$$

where $\widehat{h}_{2,2,n} - h_{2,2} \to_p 0$ and, hence, $c_{Plug}(\widehat{h}_{2,2,n}, 1-\alpha) - c_{Plug}(h_{2,2}, 1-\alpha) \to_p 0$ as $n \to \infty$ under $\{\gamma_{n,h} : n \geq 1\}$. For example, in the case of iid observations, one can take

$$\begin{aligned}
\widehat{h}_{2,2,n} &= vech_*\left(\widehat{D}_n^{-1/2}(\theta_0)\widehat{V}_n(\theta_0)\widehat{D}_n^{-1/2}(\theta_0)\right), \text{ where} \\
\widehat{V}_n(\theta_0) &= n^{-1}\sum_{i=1}^{n}(m(W_i,\theta_0) - \overline{m}_n(\theta_0))(m(W_i,\theta_0) - \overline{m}_n(\theta_0))', \\
m(W_i,\theta_0) &= (m_1(W_i,\theta_0), ..., m_{p+s}(W_i,\theta_0))' \text{ and} \\
\widehat{D}_n(\theta_0) &= Diag\{\widehat{\sigma}_{n,1}^2(\theta_0), ..., \widehat{\sigma}_{n,p+s}^2(\theta_0)\}.
\end{aligned} \tag{11.36}$$

The PCV, $c_{Plug}(\widehat{h}_{2,2,n}, 1-\alpha)$, can be computed by simulation.

The use of the PCV yields a CR for the true value θ_0 whose finite sample size has limit equal to $1-\alpha$ (using (11.32) with Fix replaced by $Plug$). The PCV CR is not asymptotically similar because the limit of its coverage probability exceeds $1-\alpha$ when $h_1 \neq 0$. However, it is closer to being asymptotically similar than the FCV CR is, because $c_{Plug}(h_{2,2}, 1-\alpha)) \leq \sup_{h_2 \in H_2} c_{Plug}(h_{2,2}, 1-\alpha) = c_{Fix}(1-\alpha)$ for all $h_{2,2}$ with strict inequality for some $h_{2,2}$.

NEED TO CHECK STATEMENTS IN THIS PARAGRAPH. For an empirical likelihood-based test statistic, as in Moon and Schorfheide (2004) or Soares (2005), the asymptotic distribution of the test statistic is the same as the GMM-based statistic above. Hence, the same argument as above leads to (11.31) and the subsample, FCV, and PCV CRs have the desired asymptotic level. The PCV CR yields the smallest CR based on an empirical likelihood test statistic.

Next, suppose the population moment functions are of the form $E_F m_j(W_i, \theta_0, \tau_0) \geq 0$ for $j = 1, ..., p$ and $E_F m_j(W_i, \theta_0, \tau_0) = 0$ for $j = p+1, ..., p+s$, where τ_0 is a parameter for which a preliminary asymptotically normal estimator $\widehat{\tau}_n(\theta_0)$ exists, as in Soares (2005). The sample moment functions are of the form $\overline{m}_{n,j}(\theta) = \overline{m}_{n,j}(\theta, \widehat{\tau}_n(\theta))$. In this case, the asymptotic variance of $n^{1/2}\overline{m}_{n,j}(\theta)$, as well as the quantities Ω and $h_{2,2}$, take different values than when τ_0 appears in place of $\widehat{\tau}_n(\theta_0)$. But, the form of the asymptotic distribution given in (11.29) is the same. (This relies on suitable smoothness of $E_F m_j(W_i, \theta_0, \tau_0)$ with respect to τ_0.) In consequence, by the same argument as above, we have $J_g \geq_{ST} J_h$ for $(g, h) \in GH$ and (11.31), (11.32), and the above PCV CR results hold. Hence, use of a preliminary estimator in the GMM criterion function does not cause size distortion for the subsample, FCV, or PCV CRs (provided J_h is properly defined and takes into account the estimation of τ_0 when computing $c_{Fix}(1-\alpha)$ or $c_{Plug}(\widehat{h}_{2,2,n}, 1-\alpha)$).

We now discuss CRs based on a minimum distance test statistic, as in Rosen (2005). (Rosen (2005) does not consider subsample critical values, but we do here.) The test statistic is

$$T_n(\theta) = \inf_{t=(t_1', 0_s)': t_1 \in R_+^p} n(\overline{m}_n(\theta) - t)' \widehat{V}_n^{-1} (\overline{m}_n(\theta) - t), \tag{11.37}$$

where $\overline{m}_n(\theta) = (\overline{m}_{n,1}(\theta), ..., \overline{m}_{n,p+s}(\theta))'$ and $\widehat{V}_n$ is a consistent estimator of $V = \lim_{n\to\infty} Var_F(n^{1/2}\overline{m}_n(\theta))$ when θ is the true parameter. For this test statistic, γ_1 and γ_3 are the same as above and $\gamma_2 = (\theta', vech_*(V)')'$. In this case, under $\{\gamma_{n,h} : n \geq 1\}$, we have

$$\begin{aligned} & T_n(\theta_{n,h}) \to_d L(h) \sim J_h, \text{ where} \\ & L(h) = \inf_{t=(t_1', 0_s)': t_1 \in R_+^p} \left(Z_{h_{2,2}}^* - \begin{pmatrix} t_1 - h_1 \\ 0_s \end{pmatrix} \right)' V^{-1} \left(Z_{h_{2,2}}^* - \begin{pmatrix} t_1 - h_1 \\ 0_s \end{pmatrix} \right) \text{ and} \\ & Z_{h_{2,2}}^* \sim N(0, V). \end{aligned} \tag{11.38}$$

If $0 \leq g_1 \leq h_1$, then algebra and $R^p_+ - g_1 \subset R^p_+ - h_1$ give

$$\begin{aligned} L(h) &= \inf_{t=(t_1^{*\prime},0_s)':t_1^* \in R^p_+ - h_1} \left(Z^*_{h_{2,2}} - \begin{pmatrix} t_1^* \\ 0_s \end{pmatrix} \right)' V^{-1} \left(Z^*_{h_{2,2}} - \begin{pmatrix} t_1^* \\ 0_s \end{pmatrix} \right) \\ &\leq \inf_{t=(t_1^{*\prime},0_s)':t_1^* \in R^p_+ - g_1} \left(Z^*_{h_{2,2}} - \begin{pmatrix} t_1^* \\ 0_s \end{pmatrix} \right)' V^{-1} \left(Z^*_{h_{2,2}} - \begin{pmatrix} t_1^* \\ 0_s \end{pmatrix} \right) \text{ a.s.(11.39)} \end{aligned}$$

Also, for $(g, h) \in GH$, we have $0 \leq g_1 \leq h_1$ and $g_{2,2} = h_{2,2}$. These results imply that $J_g \geq_{ST} J_h$ for $(g, h) \in GH$ and (11.31) holds. Hence, discontinuity of the limit distribution also does not cause size distortion for the CR based on the subsample minimum distance statistic. Analogous results to those above for the FCV and PCV CRs also hold.

All of the discussion above takes the parameter space for γ_1 to be $\Gamma_1 = R^p_+$. This is appropriate when any given moment inequality can hold as an equality regardless of whether other moment inequalities hold as equalitites or not. On the other hand, in some examples if one moment inequality holds as an equality, then some other moment inequality cannot hold as an equality. For example, consider a location model with interval outcomes. For simplicity, suppose the interval endpoints are integer values. The model is $y_i^* = \theta_0 + u_i$ and $y_i = [y_i^*]$ for $i = 1, ..., n$, where $[y_i^*]$ denotes the integer part of y_i^*, y_i^* is not observed, and y_i is observed. The interval outcome $[y_i, y_i + 1)$ necessarily includes unobserved outcome variable y_i^*. Two moment inequalities that place bounds on θ_0 are (i) $-E_{\theta_0} y_i + \theta_0 \geq 0$ and (ii) $E_{\theta_0} y_i + 1 - \theta_0 \geq 0$. If the first inequality holds as an equality, then the second cannot.

Analysis of the interval outcome model can be done using the general results of this paper as follows. We have $(-E_{\theta_0} y_i + \theta_0) \in [0, 1]$. We treat the two cases (a) $(-E_{\theta_0} y_i + \theta_0) \in [0, 1/2]$ and (b) $(-E_{\theta_0} y_i + \theta_0) \in (1/2, 1]$ separately because the asymptotic distribution of $T_n(\theta_{n,h})$ is discontinuous both at $-E_{\theta_0} y_i + \theta_0 = 0$ and at $-E_{\theta_0} y_i + \theta_0 = 1$. For case (a), we define γ_1 via $(-E_{\theta_0} y_i + \theta_0) + \gamma_1 = 0$ for $\gamma_1 \in [0, 1/2]$ and, in consequence, $(E_{\theta_0} y_i + 1 - \theta_0) - (1 - \gamma_1) = 0$. Using these equalities in place of the equalities in (11.26), we can analyze this model in the same way as above with $p = 2$ and $s = 0$. We obtain the same result as above that the limit of finite sample size is the nominal level for subsample CRs when $(-E_{\theta_0} y_i + \theta_0) \in [0, 1/2]$. For case (b), we define γ_1 via $(E_{\theta_0} y_i + 1 - \theta_0) - \gamma_1 = 0$ for $\gamma_1 \in [0, 1/2]$ and, in consequence, $(-E_{\theta_0} y_i + \theta_0) + (1 - \gamma_1) = 0$. Analogously, using these equalities in place of the equalities in (11.26), we can analyze this model in the same way as above. We obtain the same result as above that the limit of finite sample size is the nominal level for subsample CRs when $(-E_{\theta_0} y_i + \theta_0) \in (1/2, 1]$. Combining the results from cases (a) and (b) gives the same result for the model in which $(-E_{\theta_0} y_i + \theta_0) \in [0, 1]$.

11.5 Examples under Preparation

EXAMPLES THAT WE HAVEN'T FINSHED YET (OR IN SOME CASES HAVE NOT EVEN STARTED):

(1) CI FOR A UNIT ROOT AS IN Ch. 12 OF PRW. (ALTERNATIVE METHODS TO SUBSAMPLING ARE ANDREWS/STOCK TYPE CI METHODS. MIKUSHEVA's RESULTS ARE RELEVANT HERE.) HAVE RESULTS, INCLUDING SIMULATION RESULTS, BUT NOT WRITTEN UP.

(2) POOR IVS/WEAK EXOGENEITY: TESTS CONCERNING RHS PARAMETER ON ENDOGENOUS VARIABLE WHEN CORRELATION BETWEEN IV AND STRUCTURAL ERROR IS γ. SUPPOSE $\gamma = L/n^{1/2}$. CLOSELY RELATED TO MEHMET CANER's PAPER ON SUBSAMPLING IN THIS CONTEXT. HAVE RESULTS, BUT NOT FINISHED WRITE UP.

(3) WEAK IVs: RELATED TO GUGGENBERGER AND WOLF's PAPER ON SUBSAMPLING WITH WEAK IVs. HAVE SOME RESULTS, BUT NOT FINISHED WRITE UP. SUBSAMPLING, AFTER SIZE-CORRECTION, MAY BE GOOD FOR SUBSET INFERENCE ON ENDOGENOUS VARIABLES AND INFERENCE ON EXOGENOUS VARIABLES. NOT BEST FOR INFERENCE ON WHOLE VECTOR OF ENDOGENOUS VARIABLES—CONDITIONAL TESTS A LA MOREIRA (2003) ARE BETTER.

(4) TEST STATISTICS BASED ON POST-MODEL SELECTION ESTIMATORS.

(5) INFERENCE IN REGRESSIONS WITH NEARLY INTEGRATED REGRESSORS. PRW MENTION THIS AS A POTENTIAL APPLICATION OF SUBSAMPLING, p. 288. A BETTER METHOD THAN SUBSAMPLING MAY BE JANSSON AND MOREIRA'S METHOD.

12 Appendix of Proofs

The following Lemmas are used in the proofs of Theorems 1 and 2. (The expressions $\kappa_n \to [\kappa_{1,\infty}, \kappa_{2,\infty}]$ and $G(x-)$ used below are defined in Section 3.2.)

Lemma 6 *Suppose* (i) *for some df's* $L_n(\cdot)$ *and* $G_L(\cdot)$ *on* R, $L_n(x) \to_p G_L(x)$ *for all* $x \in C(G_L)$, (ii) $T_n \to_d G_T$, *where* T_n *is a scalar random variable and* G_T *is some distribution on* R, *and* (iii) *for all* $\varepsilon > 0$, $G_L(c_\infty - \varepsilon) < 1 - \alpha$ *and* $G_L(c_\infty + \varepsilon) > 1 - \alpha$, *where* c_∞ *is the* $1 - \alpha$ *quantile of* G_L *for some* $\alpha \in (0, 1)$. *Then for* $c_n := \inf\{x \in R : L_n(x) \geq 1 - \alpha\}$, (a) $c_n \to_p c_\infty$ *and* (b) $P(T_n \leq c_n) \to [G_T(c_\infty-), G_T(c_\infty)]$.

Comments. 1. Condition (iii) holds if $G_L(x)$ is strictly increasing at $x = c_\infty$ or if $G_L(x)$ has a jump at $x = c_\infty$ with $G_L(c_\infty) > 1 - \alpha$ and $G_L(c_\infty-) < 1 - \alpha$.

2. If $G_T(x)$ is continuous at c_∞, then the result of part (b) is $P(T_n \leq c_n) \to G_T(c_\infty)$.

Lemma 7 *Let* $\alpha \in (0, 1)$ *be given. Suppose Assumptions* A2, B2, C-E, F2, *and* G2 *hold. Let* $\{w_n : n \geq 1\}$ *be any subsequence of* $\{n\}$. *Let* $\{\gamma_{w_n} = (\gamma_{w_n,1}, \gamma_{w_n,2}, \gamma_{w_n,3}) : n \geq 1\}$ *be a sequence of points in* Γ *that satisfies* (i) $w_n^r \gamma_{w_n,1} \to h_1$ *for some* $h_1 \in R^p_\infty$, (ii) $b^r_{w_n} \gamma_{w_n,1} \to g_1$ *for some* $g_1 \in R^p_\infty$, *and* (iii) $\gamma_{w_n,2} \to h_2$ *for some* $h_2 \in R^q_\infty$. *Let* $h = (h_1, h_2)$, $g = (g_1, g_2)$, *and* $g_2 = h_2$. *Then, we have*

(a) $(g, h) \in GH$,

(b) $E_{\theta_0,\gamma_{w_n}} U_{w_n,b_{w_n}}(x) \to J_g(x)$ *for all* $x \in C(J_g)$,

(c) $U_{w_n,b_{w_n}}(x) \to_p J_g(x)$ *for all* $x \in C(J_g)$ *under* $\{\gamma_{w_n} : n \geq 1\}$,

(d) $L_{w_n,b_{w_n}}(x) \to_p J_g(x)$ *for all* $x \in C(J_g)$ *under* $\{\gamma_{w_n} : n \geq 1\}$,

(e) $c_{w_n,b_{w_n}}(1 - \alpha) \to_p c_g(1 - \alpha)$ *under* $\{\gamma_{w_n} : n \geq 1\}$,

(f) $P_{\theta_0,\gamma_{w_n}}(T_{w_n}(\theta_0) \leq c_{w_n,b_{w_n}}(1 - \alpha)) \to [J_h(c_g(1 - \alpha)-), J_h(c_g(1 - \alpha))]$, *and*

(g) *if* $|h_{1,j}| < \infty$ *for all* $j = 1, ..., p$ *and* $w_n = n$ *for all* $n \geq 1$, *then parts* (b)-(f) *hold with Assumptions* A2, B2, F2, *and* G2 *replaced by Assumptions* A1, B1, F1, *and* G1 *and* (g, h) *in parts* (b)-(f) *equal* (h^0, h) *of Assumption* B1.

Comment. If J_h is continuous at $c_g(1 - \alpha)$, $P_{\theta_0,\gamma_{w_n}}(T_{w_n}(\theta_0) \leq c_{w_n,b_{w_n}}(1 - \alpha)) \to J_h(c_g(1 - \alpha))$.

Lemma 8 *Let* $\alpha \in (0, 1)$ *be given. Suppose Assumptions* A2, B2, C-E, F2, *and* G2 *hold. Let* $(g, h) \in GH$ *be given. Then, there is a sequence* $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) : n \geq 1\}$ *of points in* Γ *that satisfy conditions* (i)-(iii) *of Lemma* 7 *and for this sequence parts* (b)-(f) *of Lemma* 7 *hold with* ω_n *replaced by* n.

Proof of Lemma 6. For $\varepsilon > 0$ such that $c_\infty \pm \varepsilon \in C(G_L) \cap C(G_T)$, we have

$$\begin{aligned} &L_n(c_\infty - \varepsilon) \to_p G_L(c_\infty - \varepsilon) < 1 - \alpha \text{ and} \\ &L_n(c_\infty + \varepsilon) \to_p G_L(c_\infty + \varepsilon) > 1 - \alpha \end{aligned} \tag{12.1}$$

by assumptions (i) and (iii). This and the definition of c_n yield

$$P(A_n(\varepsilon)) \to 1, \text{ where } A_n(\varepsilon) = \{c_\infty - \varepsilon \leq c_n \leq c_\infty + \varepsilon\}. \tag{12.2}$$

There exists a sequence $\{\varepsilon_k \in C(G_L) \cap C(G_T) : k \geq 1\}$ such that $\varepsilon_k \to 0$ as $k \to \infty$. Hence, part (a) holds.

Let $P(A, B)$ denote $P(A \cap B)$. For part (b), using the definition of $A_n(\varepsilon)$, we have

$$P(T_n \leq c_\infty - \varepsilon, A_n(\varepsilon)) \leq P(T_n \leq c_n, A_n(\varepsilon)) \leq P(T_n \leq c_\infty + \varepsilon). \tag{12.3}$$

Hence,

$$\begin{aligned} &\limsup_{n\to\infty} P(T_n \leq c_n) = \limsup_{n\to\infty} P(T_n \leq c_n, A_n(\varepsilon)) \\ &\leq \limsup_{n\to\infty} P(T_n \leq c_\infty + \varepsilon) = G_T(c_\infty + \varepsilon), \text{ and} \\ &\liminf_{n\to\infty} P(T_n \leq c_n) = \liminf_{n\to\infty} P(T_n \leq c_n, A_n(\varepsilon)) \\ &\geq \liminf_{n\to\infty} P(T_n \leq c_\infty - \varepsilon, A_n(\varepsilon)) = G_T(c_\infty - \varepsilon) \end{aligned} \tag{12.4}$$

using assumption (ii), $c_\infty \pm \varepsilon \in C(G_T)$, and (12.2). Given a sequence $\{\varepsilon_k : k \geq 1\}$ as above, (12.4) establishes part (b). □

Proof of Lemma 7. First, we prove part (a). We need to show that $g \in H$, $h \in H$, $g_2 = h_2$, and conditions (i)-(iii) in the definition of GH hold. For $j = 1, ..., p$, if $a_j = 0$, then $g_{1,j}, h_{1,j} \in R_{+,\infty}$ by conditions (i) and (ii) of the Lemma. Likewise, if $b_j = 0$, then $g_{1,j}, h_{1,j} \in R_{-,\infty}$. Otherwise, $g_{1,j}, h_{1,j} \in R_\infty$. Hence, by the definition of H_1, $g_1, h_1 \in H_1$. By condition (iii) of the Lemma, $h_2 \in \mathrm{cl}(\Gamma_2) = H_2$. Combining these results gives $g, h \in H$. By assumption of the Lemma, $g_2 = h_2$. By conditions (i) and (ii) of the Lemma and Assumption C(ii), conditions (i)-(iii) of GH hold. Hence, $(g, h) \in GH$.

Next, we prove part (b). For notational simplicity, we drop the subscript θ_0 from $P_{\theta_0,\gamma}$ and $E_{\theta_0,\gamma}$. We have

$$\begin{aligned} &E_{\gamma_{w_n}} U_{w_n,b_{w_n}}(x) = q_{w_n}^{-1} \sum_{i=1}^{q_{w_n}} P_{\gamma_{w_n}}(T_{w_n,b_{w_n},i}(\theta_0) \leq x) \\ &= P_{\gamma_{w_n}}(T_{w_n,b_{w_n},1}(\theta_0) \leq x) = P_{\gamma_{w_n}}(T_{b_{w_n}}(\theta_0) \leq x), \end{aligned} \tag{12.5}$$

where the first equality holds by definition of $U_{w_n,b_{w_n}}(x)$, the second equality holds by Assumption D(i), and the last equality holds by Assumption D(ii).

We now show that $P_{\gamma_{w_n}}(T_{b_{w_n}}(\theta_0) \leq x) \to J_g(x)$ for all $x \in C(J_g)$ by showing that any subsequence $\{t_n\}$ of $\{w_n\}$ has a sub-subsequence $\{s_n\}$ for which $P_{\gamma_{s_n}}(T_{b_{s_n}}(\theta_0) \leq x) \to J_g(x)$.

Given any subsequence $\{t_n\}$, select a sub-subsequence $\{s_n\}$ such that $\{b_{s_n}\}$ is strictly increasing. This can be done because $b_{w_n} \to \infty$ by Assumption C(i). Because $\{b_{s_n}\}$ is strictly increasing, it is a subsequence of $\{n\}$.

Below we show that Assumption B2 implies that for any subsequence $\{u_n\}$ of $\{n\}$ and any sequence $\{\gamma_{u_n}^* = (\gamma_{u_n,1}^*, \gamma_{u_n,2}^*, \gamma_{u_n,3}^*) \in \Gamma : n \geq 1\}$, that satisfies (i′) $u_n^r \gamma_{u_n,1}^* \to g_1$ and (ii′) $\gamma_{u_n,2}^* \to g_2 \in R^q$, we have

$$P_{\gamma_{u_n}^*}(T_{u_n}(\theta_0) \leq y) \to J_g(y), \tag{12.6}$$

for all $y \in C(J_g)$. We apply this result with $u_n = b_{s_n}$, $\gamma_{u_n}^* = \gamma_{s_n}$, and $y = x$ to obtain the desired result $P_{\gamma_{s_n}}(T_{b_{s_n}}(\theta_0) \leq x) \to J_g(x)$, where (i′) and (ii′) hold by assumptions (ii) and (iii) on $\{\gamma_{w_n} : n \geq 1\}$.

For the proof of part (b), it remains to show (12.6). Define a new sequence $\{\gamma_k^{**} = (\gamma_{k,1}^{**}, \gamma_{k,2}^{**}, \gamma_{k,3}^{**}) \in \Gamma : k \geq 1\}$ as follows. If $k = u_n$ set γ_k^{**} equal to $\gamma_{u_n}^*$. If $k \neq u_n$, define the jth component of $\gamma_{k,1}^{**}$ to be

$$\begin{array}{ll} \gamma_{k,1,j}^{**} = 0 & \text{if } g_{1,j} = 0 \\ \gamma_{k,1,j}^{**} = \max\{k^{-r} g_{1,j}, a_j/2\} & \text{if } -\infty < g_{1,j} < 0 \ \& \ a_j < 0 \\ \gamma_{k,1,j}^{**} = \min\{k^{-r} g_{1,j}, b_j/2\} & \text{if } 0 < g_{1,j} < \infty \ \& \ b_j > 0 \\ \gamma_{k,1,j}^{**} = a_j/2 & \text{if } g_{1,j} = -\infty \\ \gamma_{k,1,j}^{**} = b_j/2 & \text{if } g_{1,j} = +\infty \end{array} \tag{12.7}$$

for $j = 1, ..., p$, define $\gamma_{k,2}^{**} = \gamma_{u_{n_k},2}^*$ $(\in \Gamma_2)$, where $n_k = \max\{\ell \in N : u_\ell \leq k\}$, and define $\gamma_{k,3}^{**}$ to be any element of $\Gamma_3(\gamma_{k,1}^{**}, \gamma_{k,2}^{**})$. Note that the parameters $\{\gamma_k^{**} : k \geq 1\}$ are in Γ for all $k \geq 1$ and $k^r \gamma_{k,1}^{**} \to g_1$ and $\gamma_{k,2}^{**} \to g_2$ as $k \to \infty$. Hence, $\{\gamma_k^{**} : k \geq 1\}$ is of the form $\{\gamma_{n,g} : n \geq 1\}$ and Assumption B2 implies that $P_{\gamma_k^{**}}(T_k(\theta_0) \leq y) \to J_g(y)$ for all $y \in C(J_g)$. Because $\{u_n\}$ is a subsequence of $\{k\}$ and $\gamma_k^{**} = \gamma_{u_n}^*$ when $k = u_n$, the latter implies that $P_{\gamma_{u_n}^*}(T_{u_n}(\theta_0) \leq y) \to J_g(y)$, as desired.

Part (c) holds by part (b) and Assumption E.

To prove part (d), we show that Assumptions A2 and G2 imply that

$$L_{w_n, b_{w_n}}(x) - U_{w_n, b_{w_n}}(x) \to_p 0 \text{ under } \{\gamma_{w_n} : n \geq 1\} \text{ for all } x \in C(J_g). \tag{12.8}$$

This and part (c) of the Lemma establish part (d). To show (12.8), define a new sequence $\{\gamma_k^* = (\gamma_{k,1}^*, \gamma_{k,2}^*, \gamma_{k,3}^*) \in \Gamma : k \geq 1\}$ as follows. If $k = w_n$, set γ_k^* equal to γ_{w_n}. If $k \neq w_n$, define

$$\begin{array}{ll} \gamma_{k,1,j}^* = \max\{k^{-r} h_{1,j}, a_j/2\} & \text{if } g_{1,j} = 0 \ \& \ -\infty < h_{1,j} < 0 \\ \gamma_{k,1,j}^* = \min\{k^{-r} h_{1,j}, b_j/2\} & \text{if } g_{1,j} = 0 \ \& \ 0 \leq h_{1,j} < \infty \\ \gamma_{k,1,j}^* = \text{sgn}(h_{1,j})(b_k k)^{-r/2} & \text{if } g_{1,j} = 0 \ \& \ h_{1,j} = \pm\infty \\ \gamma_{k,1,j}^* = \max\{b_k^{-r} g_{1,j}, a_j/2\} & \text{if } -\infty < g_{1,j} < 0 \ \& \ h_{1,j} = -\infty \\ \gamma_{k,1,j}^* = \min\{b_k^{-r} g_{1,j}, b_j/2\} & \text{if } 0 < g_{1,j} < \infty \ \& \ h_{1,j} = \infty \\ \gamma_{k,1,j}^* = a_j/2 & \text{if } g_{1,j} = h_{1,j} = -\infty \\ \gamma_{k,1,j}^* = b_j/2 & \text{if } g_{1,j} = h_{1,j} = \infty, \end{array} \tag{12.9}$$

where $\gamma_{k,1}^* = (\gamma_{k,1,1}^*, ..., \gamma_{k,1,p}^*)'$, define $\gamma_{k,2}^* = \gamma_{w_{n_k},2}$, where $n_k = \max\{\ell \in N : w_\ell \leq k\}$, and define $\gamma_{k,3}^*$ to be any element of $\Gamma_3(\gamma_{k,1}^*, \gamma_{k,2}^*)$. As defined, $\gamma_k^* \in \Gamma$ for all $k \geq 1$

and straightforward calculations show that $\{\gamma_k^* : k \geq 1\}$ satisfies (i)-(iii) of Lemma 7 with $\{w_n\}$ replaced by $\{k\}$. Hence, by Lemma 7(b) with $\{w_n\}$ replaced by $\{k\}$, $U_{k,b_k}(x) \rightarrow_p J_g(x)$ as $k \rightarrow \infty$ under $\{\gamma_k^* : k \geq 1\}$ for all $x \in C(J_g)$. In consequence, because $\{\gamma_k^* : k \geq 1\}$ is of the form $\{\gamma_{n,h} : n \geq 1\}$, Assumption G2 implies that $L_{k,b_k}(x) - U_{k,b_k}(x) \rightarrow_p 0$ as $k \rightarrow \infty$ under $\{\gamma_k^* : k \geq 1\}$ for all $x \in C(J_g)$. Since $\gamma_k^* = \gamma_{w_n}$ for $k = w_n$, this implies that (12.8) holds.

Parts (e) and (f) are established by applying Lemma 6 with $L_n(x) = L_{w_n,b_{w_n}}(x)$ and $T_n = T_{w_n}(\theta_0)$ and verifying the conditions of Lemma 6 using (I) part (d), (II) $T_{w_n}(\theta_0) \rightarrow_d J_h$ under $\{\gamma_{w_n} : n \geq 1\}$ (which is verified below), and (III) Assumption F2. The result of (II) holds because $\{\gamma_k^* : k \geq 1\}$ in the proof of part (d) is of the form $\{\gamma_{n,h} : n \geq 1\}$ for h as defined in the statement of Lemma 7; this and Assumption B2 imply that $T_k(\theta_0) \rightarrow_d J_h$ as $k \rightarrow \infty$ under $\{\gamma_k^* : k \geq 1\}$; and the latter and $\gamma_k^* = \gamma_{w_n}$ for $k = w_n$ imply the result of (II).

Part (g) holds because (I) the proof of part (b) goes through with Assumptions A2 and B2 replaced by Assumptions A1 and B1 given that $|h_{1,j}| < \infty$ for all $j = 1, ..., p$, (II) the proof of part (c) holds without change, (III) part (d) holds immediately by part (c) and Assumption G1 (in place of Assumption G2) because $w_n = n$ for all $n \geq 1$, and (IV) the proof of parts (e) and (f) holds with Assumptions A2, B2, and F2 replaced by Assumptions A1, B1(i), and F1 given that $|h_{1,j}| < \infty$ for all $j = 1, ..., p$ (which implies that $g_{1,j} = 0$ for $j = 1, ..., p$ using conditions (i) and (ii) of the Lemma and Assumption C(ii)) and $w_n = n$. □

Proof of Lemma 8. Define $\gamma_{n,1,j}$ as in (12.9) with n in place of k for $j = 1, ..., p$ and let $\gamma_{n,1} = (\gamma_{n,1,1}, ..., \gamma_{n,1,p})'$. Define $\{\gamma_{n,2} : n \geq 1\}$ to be any sequence of points in Γ_2 such that $\gamma_{n,2} \rightarrow h_2$ as $n \rightarrow \infty$. Let $\gamma_{n,3}$ be any element of $\Gamma_3(\gamma_{n,1}, \gamma_{n,2})$ for $n \geq 1$. Then, $\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3})$ is in Γ for all $n \geq 1$ using Assumption A2. Also, using Assumption C, straightforward calculations show that $\{\gamma_n : n \geq 1\}$ satisfies conditions (i)-(iii) of Lemma 7 with $w_n = n$. Hence, parts (b)-(f) of Lemma 7 hold with $w_n = n$ for $\{\gamma_n : n \geq 1\}$ as defined above. □

Proof of Theorem 1. Part (a) holds by Assumption B1(i) and the definition of convergence in distribution by considering points of continuity of $J_h(\cdot)$ that are greater than $c_{Fix}(1-\alpha)$ and arbitrarily close to $c_{Fix}(1-\alpha)$ as well as continuity points that are less than $c_{Fix}(1-\alpha)$ and arbitrarily close to it. Part (b) follows from Lemma 7(g) because $|h_{1,j}| < \infty$ for $j = 1, ..., p$, $w_n = n$ for all $n \geq 1$, g in Lemma 7(g) equals $h^0 = (0, h_2)$, conditions (i) and (iii) of Lemma 7 hold by the definition of the sequence $\{\gamma_{n,h} : n \geq 1\}$, and condition (ii) of Lemma 7 holds because $n^r \gamma_{n,h,1} \rightarrow h_1$ with $||h_1|| < \infty$ implies that $b_n^r \gamma_{n,h,1} \rightarrow 0$ using Assumption C(ii). □

Proof of Theorem 2. The proof of part (a) is similar to that of part (b), but noticeably simpler because $c_{Fix}(1-\alpha)$ is a constant. Furthermore, the proof of the second result of part (b) is quite similar to that of the first result. Hence, for brevity, we only prove the first result of part (b).

We first show that $AsySz(\theta_0) \geq Max_{Sub}(\alpha)$. Equations (2.8) and (2.9) imply that for any sequence $\{\gamma_n \in \Gamma : n \geq 1\}$,

$$AsySz(\theta_0) \geq \limsup_{n\to\infty} [1 - P_{\theta_0,\gamma_n}(T_n(\theta_0) \leq c_{n,b}(1-\alpha))]. \tag{12.10}$$

In consequence, to show $AsySz(\theta_0) \geq Max_{Sub}(\alpha)$, it suffices to show that given any $(g,h) \in GH$ there exists a sequence $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) \in \Gamma : n \geq 1\}$ such that

$$\limsup_{n\to\infty} [1 - P_{\theta_0,\gamma_n}(T_n(\theta_0) \leq c_{n,b}(1-\alpha))] \geq 1 - J_h(c_g(1-\alpha)). \tag{12.11}$$

The latter inequality holds (as an equality) by Lemma 8.

It remains to show $AsySz(\theta_0) \leq Max^-_{Sub}(\alpha)$. Let $\{\gamma_n^* = (\gamma_{n,1}^*, \gamma_{n,2}^*, \gamma_{n,3}^*) \in \Gamma : n \geq 1\}$ be a sequence such that $\limsup_{n\to\infty} RP_n(\theta_0, \gamma_n^*) = \limsup_{n\to\infty} \sup_{\gamma\in\Gamma} RP_n(\theta_0,\gamma)$ $(= AsySz(\theta_0))$. Such a sequence always exists. Let $\{v_n : n \geq 1\}$ be a subsequence of $\{n\}$ such that $\lim_{n\to\infty} RP_{v_n}(\theta_0, \gamma_{v_n}^*)$ exists and equals $\limsup_{n\to\infty} RP_n(\theta_0, \gamma_n^*) = AsySz(\theta_0)$. Such a subsequence always exists.

Let $\gamma_{n,1,j}^*$ denote the jth component of $\gamma_{n,1}^*$. Either (1) $\limsup_{n\to\infty} |v_n^r \gamma_{v_n,1,j}^*| < \infty$ or (2) $\limsup_{n\to\infty} |v_n^r \gamma_{v_n,1,j}^*| = \infty$. If (1) holds, then for some subsequence $\{w_n\}$ of $\{v_n\}$

$$\begin{aligned} b_{w_n}^r \gamma_{w_n,1,j}^* &\to 0 \text{ and} \\ w_n^r \gamma_{w_n,1,j}^* &\to h_{1,j} \text{ for some } h_{1,j} \in R. \end{aligned} \tag{12.12}$$

If (2) holds, then either (2a) $\limsup_{n\to\infty} |b_{v_n}^r \gamma_{v_n,1,j}^*| < \infty$ or (2b) $\limsup_{n\to\infty} |b_{v_n}^r \gamma_{v_n,1,j}^*| = \infty$. If (2a) holds, then for some subsequence $\{w_n\}$ of $\{v_n\}$,

$$\begin{aligned} b_{w_n}^r \gamma_{w_n,1,j}^* &\to g_{1,j} \text{ for some } g_{1,j} \in R \text{ and} \\ w_n^r \gamma_{w_n,1,j}^* &\to h_{1,j}, \text{ where } h_{1,j} = \infty \text{ or } -\infty \text{ with } sgn(h_{1,j}) = sgn(g_{1,j}). \end{aligned} \tag{12.13}$$

If (2b) holds, then for some subsequence $\{w_n\}$ of $\{v_n\}$,

$$\begin{aligned} b_{w_n}^r \gamma_{w_n,1,j}^* &\to g_{1,j}, \text{ where } g_{1,j} = \infty \text{ or } -\infty, \text{ and} \\ w_n^r \gamma_{w_n,1,j}^* &\to h_{1,j}, \text{ where } h_{1,j} = \infty \text{ or } -\infty \text{ with } sgn(h_{1,j}) = sgn(g_{1,j}). \end{aligned} \tag{12.14}$$

In addition, for some subsequence $\{w_n\}$ of $\{v_n\}$,

$$\gamma_{w_n,2}^* \to h_2 \text{ for some } h_2 \in \text{cl}(\Gamma_2). \tag{12.15}$$

By taking successive subsequences over the p components of $\gamma_{v_n,1}^*$ and $\gamma_{v_n,2}^*$, we find that there exists a subsequence $\{w_n\}$ of $\{v_n\}$ such that for each $j = 1, ..., p$ exactly one of the cases (12.12)-(12.14) applies and (12.15) holds. In consequence, conditions (i)-(iii) of Lemma 7 hold. In addition, $\gamma_{w_n,3}^* \in \Gamma_3(\gamma_{w_n,1}^*, \gamma_{w_n,2}^*)$ for all $n \geq 1$ because $\gamma_{w_n}^* \in \Gamma$. Hence,

$$RP_{w_n}(\theta_0, \gamma_{w_n}^*) \to [1 - J_h(c_g(1-\alpha)), 1 - J_h(c_g(1-\alpha)-)] \tag{12.16}$$

by Lemma 7(f). Also, $(g,h) \in GH$ by Lemma 7(a). Since $\lim_{n\to\infty} RP_{v_n}(\theta_0, \gamma^*_{v_n}) = AsySz(\theta_0)$ and $\{w_n\}$ is a subsequence of $\{v_n\}$, we have $\lim_{n\to\infty} RP_{w_n}(\theta_0, \gamma^*_{w_n}) = AsySz(\theta_0)$. This, (12.16) and $(g,h) \in GH$ imply that $AsySz(\theta_0) \le Max^-_{Sub}(\alpha)$, which completes the proof of the first result of part (b). □

Proof of Lemma 2. If $c_h(1-\alpha) \ge c_\infty(1-\alpha)$ for all $h \in H$, then $Max^-_{Hyb}(\alpha) = Max^-_{Sub}(\alpha)$ and $Max_{Hyb}(\alpha) = Max_{Sub}(\alpha)$ follows immediately. On the other hand, suppose "$c_h(1-\alpha) \ge c_\infty(1-\alpha)$ for all $h \in H$" does not hold. Then, for some $g \in H$, $c_g(1-\alpha) < c_\infty(1-\alpha)$. Given g, define $h_1 = (h_{1,1}, ..., h_{1,p})' \in H_1$ by $h_{1,j} = +\infty$ if $g_{1,j} > 0$, $h_{1,j} = -\infty$ if $g_{1,j} < 0$, $h_{1,j} = +\infty$ or $-\infty$ (chosen so that $(g,h) \in GH$) if $g_{1,j} = 0$, and define $h_2 = g_2$. Let $h = (h_1, h_2)$. By construction, $(g,h) \in GH$. By Assumption J, $c_h(1-\alpha) = c_\infty(1-\alpha)$. Hence, we have

$$Max_{Sub}(\alpha) \ge 1 - J_h(c_g(1-\alpha)) > \alpha, \tag{12.17}$$

where the second inequality holds because $c_g(1-\alpha) < c_\infty(1-\alpha) = c_h(1-\alpha)$ and $c_h(1-\alpha)$ is the infimum of values x such that $J_h(x) \ge 1-\alpha$ or, equivalently, $1 - J_h(x) \le \alpha$. Equation (12.17) and Theorem 2(b) imply that $AsySz(\theta_0) > \alpha$ for the subsample test. The hybrid test reduces the over-rejection of the subsample test at (g,h) from being at least $1 - J_h(c_g(1-\alpha)) > \alpha$ to being at most $1 - J_h(c_\infty(1-\alpha)) = 1 - J_h(c_h(1-\alpha)) \le \alpha$ (with equality if $J_h(\cdot)$ is continuous at $c_h(1-\alpha)$). □

Proof of Lemma 3. Suppose Assumption Quant0(i) holds. Then, we have

$$\begin{aligned} Max^-_{Hyb}(\alpha) &= \sup_{(g,h)\in GH} [1 - J_h(\max\{c_g(1-\alpha), c_\infty(1-\alpha)\}-)] \\ &= \sup_{h\in H}[1 - J_h(c_\infty(1-\alpha)-)] = \sup_{h\in H}[1 - J_h(c_\infty(1-\alpha))] \\ &\le \sup_{h\in H}[1 - J_h(c_h(1-\alpha))] \le \alpha, \end{aligned} \tag{12.18}$$

where the second equality and first inequality hold by Assumption Quant0(i)(a), the third equality holds by Assumption Quant0(i)(b), and the last inequality holds by the definition of $c_h(1-\alpha)$.

Next, suppose Assumption Quant0(ii) holds. By Assumption Quant0(ii)(a), $p = 1$. Hence, given $(g,h) \in GH$ either (I) $|h_{1,1}| = \infty$ or (II) $|h_{1,1}| < \infty$. When (I) holds, $J_h = J_\infty$ by Assumption J and

$$\begin{aligned} &1 - J_h(\max\{c_g(1-\alpha), c_\infty(1-\alpha)\}-) \\ &\le 1 - J_\infty(c_\infty(1-\alpha)-) = 1 - J_\infty(c_\infty(1-\alpha)) \le \alpha, \end{aligned} \tag{12.19}$$

where the equality holds by Assumption Quant0(ii)(c). When (II) holds, g must equal h^0 by the definition of GH. Hence,

$$\begin{aligned} &1 - J_h(\max\{c_g(1-\alpha), c_\infty(1-\alpha)\}-) \le 1 - J_h(c_{h^0}(1-\alpha)-) \\ &\le \sup_{h\in H}[1 - J_h(c_h(1-\alpha)-)] = \sup_{h\in H}[1 - J_h(c_h(1-\alpha))] \le \alpha, \end{aligned} \tag{12.20}$$

where the second inequality holds because $c_{h^0}(1-\alpha) \geq c_h(1-\alpha)$ by Assumption Quant0(ii)(b) and the equality holds by Assumption Quant0(ii)(d). $\square$

Proof of Lemma 4. By Assumption KK(i), $cv(1-\alpha) < \infty$. Now, suppose $h \in H$ is such that $c_h(1-\alpha) < cv(1-\alpha)$. Then,

$$J_h(cv(1-\alpha)-) \geq J_h(c_h(1-\alpha)) \geq 1-\alpha, \tag{12.21}$$

where the first inequality holds because $J_h(x)$ is nondecreasing and $cv(1-\alpha)-\varepsilon > c_h(1-\alpha)$ for $\varepsilon > 0$ sufficiently small and the second inequality holds by the definition of the quantile $c_h(1-\alpha)$.

Next, suppose $h \in H$ is such that $c_h(1-\alpha) = cv(1-\alpha)$. Then,

$$J_h(cv(1-\alpha)-) = J_h(cv(1-\alpha)) = J_h(c_h(1-\alpha)) = 1-\alpha, \tag{12.22}$$

where the first equality holds by Assumption KK(ii), the second equality holds by the definition of h, and the third equality holds by the definition of the quantile $c_h(1-\alpha)$ and Assumption KK(ii). The case $h \in H$ and $c_h(1-\alpha) > cv(1-\alpha)$ is not possible because $cv(1-\alpha) = \sup_{g \in H} c_g(1-\alpha)$. Hence, (12.21) and (12.22) combine to show that for all $h \in H$, $1 - J_h(cv(1-\alpha)-) \leq \alpha$, as desired. $\square$

Proof of Theorem 3. The result of part (a)(i) follows from Assumption Quant1(i) with $g = h^0$ because $h \in H$ implies that $(h^0, h) \in GH$ and $cv(1-\alpha) = \sup_{h \in H} c_h(1-\alpha)$ by Assumption KK(i). Assumption Quant1(i) also implies that

$$1 - J_h(c_g(1-\alpha)) \leq 1 - J_h(c_h(1-\alpha)) \leq \alpha \tag{12.23}$$

for all $(g, h) \in GH$. Hence, using Assumption Quant1(ii), $Max^-_{Sub}(\alpha) = Max_{Sub}(\alpha) \leq \alpha$, $\xi(\alpha) = \alpha$, and part (a)(ii) holds.

The result of part (a)(iii) (i.e., $\xi^*(\alpha) = \alpha$) follows from

$$Max^-_{Hyb}(\alpha) \leq Max^-_{Sub}(\alpha) \leq \alpha, \tag{12.24}$$

where the second inequality holds by part (a)(ii).

Next, part (a)(iv) holds because

$$\begin{aligned} c^*_{g,\infty}(1-\xi^*(\alpha)) &= c^*_{g,\infty}(1-\alpha) = \max\{c_g(1-\alpha), c_\infty(1-\alpha)\} \\ &= c_g(1-\alpha) = c_g(1-\xi(\alpha)) \end{aligned} \tag{12.25}$$

for all $g \in H$, where the first through fourth equalities hold, respectively, by part (a)(iii), the definition of $c^*_{g,\infty}(1-\alpha)$, Assumption Quant1(i), and part (a)(ii).

Part (a)(v) of the Theorem holds because $c_g(1-\xi(\alpha)) = c_g(1-\alpha) \leq c_{g^0}(1-\alpha) \leq cv(1-\alpha)$ for all $g \in H$, where $g = (g_1, g_2)$ and $g^0 = (0, g_2)$, by part (a)(ii), Assumption Quant1(i), and part (a)(i), respectively.

Next, we prove part (b)(i). Given any $g = (g_1, g_2) = (g_{1,1}, ..., g_{1,p}, g_2) \in H$, let $g^\infty = (g_1^\infty, g_2) = (g_{1,1}^\infty, ..., g_{1,p}^\infty, g_2) \in H$ be such that $g_{1,j}^\infty = +\infty$ if $g_{1,j} > 0$, $g_{1,j}^\infty = -\infty$

if $g_{1,j} < 0$, $g_{1,j}^{\infty} = +\infty$ or $-\infty$ (chosen so that $g^{\infty} \in H$) if $g_{1,j} = 0$ for $j = 1, ..., p$. By Assumption Quant2(i), $c_g(1-\alpha) \leq c_{g^{\infty}}(1-\alpha)$ because $(g, g^{\infty}) \in GH$. By Assumption J, $c_{g^{\infty}}(1-\alpha) = c_{\infty}(1-\alpha)$ for all $g \in H$. Hence, $cv(1-\alpha) = \sup_{h \in H} c_h(1-\alpha) = c_{\infty}(1-\alpha)$, which proves part (b)(i).

We now prove part (b)(ii). Part (b)(i) and $c_{\infty}(1-\alpha) \leq c_{g,\infty}^{*}(1-\alpha)$ (by the definition of $c_{g,\infty}^{*}(1-\alpha)$), yields $cv(1-\alpha) \leq c_{g,\infty}^{*}(1-\alpha)$. The latter implies that $\xi^{*}(\alpha) = \alpha$, using the definition of $\xi^{*}(\alpha)$, which is the result of part (b)(ii). In turn, $\xi^{*}(\alpha) = \alpha$ gives

$$c_{g,\infty}^{*}(1 - \xi^{*}(\alpha)) = c_{g,\infty}^{*}(1 - \alpha) = \max\{c_g(1 - \alpha), c_{\infty}(1 - \alpha)\} \leq cv(1 - \alpha), \quad (12.26)$$

where the inequality holds by definition of $cv(1 - \alpha)$. Combining this with the reverse inequality that is given prior to the displayed equation yields $c_{g,\infty}^{*}(1-\xi^{*}(\alpha)) = cv(1-\alpha)$, which is the result of part (b)(iii).

We now prove part (b)(iv). By Assumption L, $Max_{Sub}^{-}(\xi(\alpha)) = \sup_{(g,h) \in GH}[1 - J_h(c_g(1 - \xi(\alpha))-)] \leq \alpha$. This implies that

$$c_g(1 - \xi(\alpha)) \geq c_h(1 - \alpha) \quad (12.27)$$

for all $(g, h) \in GH$. Given any $g \in H$, define g^0 and g^{∞} as in the previous three paragraphs. Equation (12.27) with $(g, h) = (g^0, g^{\infty})$, Assumption J (which implies that $c_h(1-\alpha) = c_{\infty}(1-\alpha)$ for all $h = (h_1, h_2) = (h_{1,1}, ..., h_{1,p}, h_2) \in H$ with $h_{1,j} = \pm\infty$ for all $j = 1, ..., p$), and part (b)(i) give

$$c_{g^0}(1 - \xi(\alpha)) \geq c_{g^{\infty}}(1 - \alpha) = c_{\infty}(1 - \alpha) = cv(1 - \alpha). \quad (12.28)$$

By Assumption Quant2(ii), $c_{g^0}(1-\xi(\alpha)) \leq c_g(1-\xi(\alpha))$ for all $g \in H$. This and (12.28) establish part (b)(iv).

Part (c)(i) and (c)(iii) follow immediately from Assumptions Quant3(i) and Quant3(iv), respectively. Part (c)(ii) is proved as follows. By (12.27) with $(g, h) = (h^{*0}, h^{*})$, and Assumption Quant3(ii), $c_{h^{*0}}(1 - \xi(\alpha)) \geq c_{h^{*}}(1 - \alpha) > c_{h^{*0}}(1 - \alpha)$. Hence, $\xi(\alpha) < \alpha$. This, Assumption Quant3(iii), and the result of part (c)(i) gives $c_{h^{*}}(1 - \xi(\alpha)) > c_{h^{*}}(1 - \alpha) = cv(1 - \alpha)$. □

Proof of Lemma 5. Assume $U_{n,b}(x) \to_p J_g(x)$ for all $x \in C(J_g)$ under $\{\gamma_{n,h} : n \geq 1\}$ for some $g \in H$ and $h \in H$. To show $L_{n,b}(x) - U_{n,b}(x) \to_p 0$ for all $x \in C(J_g)$ under $\{\gamma_{n,h}\}$, we use the argument in the proofs of Theorems 11.3.1(i) and 12.2.2(i) in PRW.

Define $R_n(t) := q_n^{-1} \sum_{i=1}^{q_n} 1(|\tau_{b_n}(\widehat{\theta}_n - \theta_0)/\widehat{\sigma}_{n,b_n,i}| \geq t)$. Using

$$U_{n,b}(x - t) - R_n(t) \leq L_{n,b}(x) \leq U_{n,b}(x + t) + R_n(t) \quad (12.29)$$

for any $t > 0$ (which holds for all versions (i)–(iii) of $T_n(\theta_0)$ in Assumption t1), the desired result follows once we establish that $R_n(t) \to_p 0$ under $\{\gamma_{n,h}\}$ for any fixed $t > 0$. By $\tau_n = a_n/d_n$, we have

$$|\tau_{b_n}(\widehat{\theta}_n - \theta_0)/\widehat{\sigma}_{n,b_n,i}| \geq t \text{ iff } (a_{b_n}/a_n)a_n|\widehat{\theta}_n - \theta_0| \geq d_{b_n}\widehat{\sigma}_{n,b_n,i}t \quad (12.30)$$

provided $\widehat{\sigma}_{n,b_n,i} > 0$, which by Assumption BB2(ii) holds uniformly in $i = 1, ..., n$ wp→1. By Assumption BB2(i) and HH, $(a_{b_n}/a_n)a_n|\widehat{\theta}_n - \theta_0| = o_p(1)$ under $\{\gamma_{n,h}\}$. Therefore, for any $\delta > 0$, $R_n(t) \leq q_n^{-1}\sum_{i=1}^{q_n} 1(\delta \geq d_{b_n}\widehat{\sigma}_{n,b_n,i}t) = U^{\sigma}_{n,b_n}(\delta/t)$ wp→1. Now, by an argument as in the proof of Lemma 7(a) and (b) (which uses Assumption EE, but does not use Assumption G2) applied to the statistic $d_n\widehat{\sigma}_n$ rather than $T_{w_n}(\theta_0)$, we have $U^{\sigma}_{n,b_n}(x) \rightarrow_p W_g(x)$ for all $x \in C(W_g)$ under $\{\gamma_{n,h}\}$, where $g \in H$ is defined as in Lemma 7 with $\{\gamma_{w_n}\}$ being equal to $\{\gamma_{n,h}\}$. Therefore, $U^{\sigma}_{n,b_n}(\delta/t) \rightarrow_p W_g(\delta/t)$ for $\delta/t \in C(W_g)$ under $\{\gamma_{n,h}\}$. By Assumption BB2(iii), W_g does not have positive mass at zero and, hence, $W_g(\delta/t) \rightarrow 0$ as $\delta \rightarrow 0$. We can therefore establish that $R_n(t) \rightarrow_p 0$ for any $t > 0$ by letting δ go to zero such that $\delta/t \in C(W_g)$. □

References

Chernozhukov, V., H. Hong, and E. Tamer (2002): "Parameter Set Inference in a Class of Econometric Models," unpublished working paper, Department of Economics, MIT.

Dufour, J.-M. and J. Jasiak (2001): "Finite Sample Limited Information Inference Methods for Structural Equations and Models with Generated Regressors," *International Economic Review*, 42, 815-843.

Efron, B. and R. J. Tibshirani (1993): *An Introduction to the Bootstrap.* New York: Chapman and Hall.

Guggenberger, P. and M. Wolf (2004): "Subsampling Tests of Parameter Hypotheses and Overidentifying Restrictions with Possible Failure of Identification", working paper, Dept. of Economics, UCLA.

Moon, H. R. and F. Schorfheide (2004): "Empirical Likelihood Estimation with Inequality Moment Conditions," unpublished working paper, Department of Economics, University of Southern California.

Politis, D. N. and J. P. Romano (1994): "Large Sample Confidence Regions Based on Subsamples Under Minimal Assumptions," *Annals of Statistics,* 22, 2031-2050.

Politis, D. N., J. P. Romano, and M. Wolf (1999): *Subsampling.* New York: Springer.

Romano, J. P. and M. Wolf (2001): "Subsampling Intervals in Autoregressive Models with Linear Time Trend", *Econometrica,* 69, 1283–1314.

Rosen, A. M. (2005): " Confidence Sets for Partially Identified Parameters That Satisfy a Finite Number of Moment Inequalities ," unpublished working paper, Department of Economics, Northwestern University.

Shaikh, A. M. (2005): "Inference for Partially Identified Econometric Models," unpublished working paper, Department of Economics, Stanford University.

Soares, G. (2005): "Inference with Inequality Moment Constraints," unpublished working paper, Department of Economics, Yale University.

Table I. Testing with a Nuisance Parameter on the Boundary: Maximum Asymptotic Null Rejection Probabilities (×100) as a Function of the True Correlation ρ for Tests with Nominal Size .05

	Upper 1-sided			Symmetric 2-sided			Equal-tailed 2-sided		
ρ	Sub	FCV	Hyb	Sub	FCV	Hyb	Sub	FCV	Hyb
-1.00	50.2	5.0	5.0	9.9	10.1	9.9	52.7	10.1	5.0
-.99	42.8	5.0	5.0	9.9	10.1	9.9	43.2	10.1	5.0
-.95	33.8	5.0	5.0	9.9	10.1	9.9	32.4	10.1	5.0
-.90	27.6	5.0	5.0	9.9	10.1	9.9	25.4	10.1	5.0
-.80	20.2	5.0	5.0	9.3	10.1	9.3	17.4	10.1	5.0
-.60	12.3	5.0	5.0	7.4	10.1	7.4	10.0	10.1	5.0
-.40	8.3	5.0	5.0	6.0	10.1	6.0	6.8	10.1	5.0
-.20	6.2	5.0	5.0	5.2	10.1	5.2	5.3	10.1	5.0
.00	5.0	5.0	5.0	5.0	10.1	5.0	5.0	10.1	5.0
.20	5.0	5.6	5.0	5.2	10.1	5.2	5.4	10.1	5.0
.40	5.0	5.8	5.0	5.9	10.1	5.9	6.7	10.1	5.0
.60	5.0	5.6	5.0	7.5	10.1	7.5	9.9	10.1	5.0
.80	5.0	5.1	5.0	9.6	10.1	9.6	17.3	10.1	5.0
.90	5.0	5.0	5.0	10.1	10.1	10.1	25.2	10.1	5.0
.95	5.0	5.0	5.0	10.1	10.1	10.1	32.4	10.1	5.0
.99	5.0	5.0	5.0	10.1	10.1	10.1	43.0	10.1	5.0
1.00	5.0	5.0	5.0	10.1	10.1	10.1	52.3	10.1	5.0