HYBRID AND SIZE-CORRECTED SUBSAMPLE METHODS

By

Donald W. K. Andrews Patrik Guggenberger

March 2007

COWLES FOUNDATION DISCUSSION PAPER NO. 1606

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

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

Hybrid and Size-Corrected Subsample Methods

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

Patrik Guggenberger
Department of Economics
UCLA

June 2005
Revised: March 2007

* 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. The authors thank Hannes Leeb, Azeem Shaikh, Jim Stock, and the participants at a number of seminars and conferences at which the paper was presented.

Abstract

This paper considers the problem of constructing tests and confidence intervals (CIs) that have correct asymptotic size in a broad class of non-regular models. The models considered are non-regular in the sense that standard test statistics have asymptotic distributions that are discontinuous in some parameters. It is shown in Andrews and Guggenberger (2005a) that standard fixed critical value, subsample, and $b < n$ bootstrap methods often have incorrect size in such models. This paper introduces general methods of constructing tests and CIs that have correct size. First, procedures are introduced that are a hybrid of subsample and fixed critical value methods. The resulting hybrid procedures are easy to compute and have correct size asymptotically in many, but not all, cases of interest. Second, the paper introduces size-correction and "plug-in" size-correction methods for fixed critical value, subsample, and hybrid tests. The paper also introduces finite-sample adjustments to the asymptotic results of Andrews and Guggenberger (2005a) for subsample and hybrid methods and employs these adjustments in size-correction.

The paper discusses several examples in detail. The examples are: (i) tests when a nuisance parameter may be near a boundary, (ii) CIs in an autoregressive model with a root that may be close to unity, and (iii) tests and CIs based on a post-conservative model selection estimator.

Keywords: Asymptotic size, autoregressive model, $b < n$ bootstrap, finite-sample size, hybrid test, model selection, over-rejection, parameter near boundary, size correction, subsample confidence interval, subsample test.

JEL Classification Numbers: C12, C15.

1 Introduction

Non-regular models are becoming increasingly important in econometrics and statistics as developments in computation make it feasible to employ models with more nonlinearities. In a variety of non-regular models, however, methods based on a standard asymptotic fixed critical value (FCV) or the bootstrap do not yield tests or confidence intervals with the correct size even asymptotically. In such cases, the usual prescription in the literature is to use subsample or $b < n$ bootstrap methods (where b denotes the bootstrap sample size). For references, see Andrews and Guggenberger (2005a), hereafter denoted AG1. However, AG1 shows that in a fairly broad array of non-regular models these methods do not deliver correct asymptotic size (defined to be the limit of finite-sample size). The purpose of this paper is to provide some general methods of constructing tests and confidence intervals (CIs) that do have correct asymptotic size in such models. The results cover cases in which a test statistic has an asymptotic distribution that is discontinuous in some parameter. Examples are listed in the Abstract. Additional examples are given in AG1 and Andrews and Guggenberger (2005b).

To avoid considerable repetition, we use the same notation and definitions in this paper as in AG1. Assumptions A1, A2, B1, B2, etc. are stated in AG1.

Some of the methods we propose are as easy to use as subsample or bootstrap methods and do not require a detailed analysis of the model at hand (except to determine whether the method is asymptotically correct). Other methods require a detailed analysis of the model and computation of size-correction values. Once these size-correction values have been determined, the methods are easy to use. In a few models, size-correction methods (at least of the type we consider here) do not work.

The methods considered here apply to a wide variety of models. In fact, they may be the most general statistical methods available given that they apply more broadly than the bootstrap, the $b < n$ bootstrap, and subsampling. However, their usefulness is greatest in models in which other methods are not applicable. In models in which other methods work properly (in the sense that the limit of their finite-sample size equals their nominal level), such methods may be preferable to the methods considered here in terms of the accuracy of the asymptotic approximations and/or the power of the test or length of the CI they generate. For example, if standard asymptotic approximations or the bootstrap work properly in a given model, then they are preferable to the methods considered here. (However, note that the fixed critical value methods considered here reduce to standard asymptotic approximations when the latter work properly.)

The first method considered in the paper is a hybrid method that takes the critical value for a given test statistic to be the maximum of a subsample critical value and the FCV that applies when the true parameter is not near a point of discontinuity of the asymptotic distribution. The latter is usually a normal or chi-square critical value. By simply taking the maximum of these two critical values, one obtains a test or CI that has correct asymptotic size in quite a few cases where the FCV, subsample, or both methods have incorrect asymptotic size. Examples are given below. Furthermore, the paper shows that the hybrid method has the feature that relative to a subsample method either (i) the subsample method has correct size asymptotically and the subsample and hybrid critical values are the same asymptotically or (ii) the subsample method has incorrect size asymptotically and the hybrid method reduces the magnitude of over-rejection for at least some parameter values, sometimes eliminating size distortion.

The hybrid test also can be applied with a $b < n$ bootstrap critical value in place of a subsample critical value. The reason is that the $b < n$ bootstrap can be viewed as subsampling with replacement and the difference between sampling with and without replacement is asymptotically negligible if $b^2/n \to 0$, see Politis, Romano, and Wolf (1999, p. 48).

The second method considered in the paper is a size-correction (SC) method. This method can be applied to FCV, subsample, and hybrid procedures. The basic idea is to use the formulae given in AG1 for the asymptotic sizes of these procedures and to increase the magnitudes of the critical values (by adding a constant or reducing the nominal level) to achieve a test whose asymptotic size equals the desired asymptotic level. Closed form solutions are obtained for the SC values (based on adding a constant). Numerical work in a number of different examples shows that computation of the SC values is tractable. However, the more complicated is an example, the more difficult is computation. The computation of the SC values in a very complicated model could be difficult or intractable. The same SC values are applicable when one uses a $b < n$ bootstrap critical value in place of a subsample critical value (provided $b^2/n \to 0$).

The paper provides some analytical comparisons of the asymptotic power of different SC tests and finds that the SC hybrid test has advantages over FCV and subsample methods in most cases, but it does not dominate the SC subsample method. In Appendix B, we introduce a SC combined method that has power at least as good as that of the SC subsample and hybrid tests. But, it reduces to the SC hybrid test in most examples and, hence, may be of more interest theoretically than practically.

The SC methods that we consider are not asymptotically conservative, but typically are asymptotically non-similar. That is, for tests, the limit of the supremum of the finite-sample rejection probability over points in the null hypothesis equals the nominal level, but the limit of the infimum over points in the null hypothesis is less than the nominal level. Usually power can be improved in such cases by reducing the magnitude of asymptotic non-similarity. To do so, we introduce "plug-in" size-correction (PSC) methods for FCV, subsample, and hybrid tests. These methods are applicable if there is a parameter sub-vector that affects the asymptotic distribution of the test statistic under consideration and is consistently estimable. The PSC method makes the size-correction value depend on a consistent estimator of the parameter sub-vector. Closed-form solutions for the PSC values are given. In some examples, the PSC method is found to be very effective.

The asymptotic results for subsample methods derived in AG1, and utilized here for size correction, do not depend on the choice of subsample size b provided $b \to \infty$

and $b/n \to 0$ as $n \to \infty$. One would expect that this may lead to poor approximations in some cases. To improve the approximations, the paper introduces finite-sample adjustments to the asymptotic rejection probabilities of subsample and hybrid tests. The adjustments depend on the magnitude of $\delta_n = b/n$. The adjusted formulae for the asymptotic rejection probabilities are used to define adjusted SC (ASC) values and adjusted PSC (APSC) values. In some examples these adjustments are found to work very well, but in some others they do not perform well due to under-correction.

The bulk of the paper considers tests that reject the null hypothesis when a test statistic is large. This covers upper and lower one-sided t tests, symmetric two-sided t tests and tests based on likelihood ratio and Lagrange multiplier statistics, among others. In the non-regular cases considered here, asymptotic distributions often are not symmetric and equal-tailed t tests are of interest. We show how the methods outlined above can be generalized to equal-tailed t tests.

For expositional purposes, the focus of much of the paper is on tests, rather than CIs. CIs are very important, so we show how the results for tests extend to CIs that are obtained by inverting tests. We note that results for CIs do not follow immediately from the results for tests because CIs require uniformity in the asymptotics over all parameters in the parameter space for correct asymptotic size, while tests only require uniformity over all parameters in the null hypothesis. Nevertheless, using the same notational/definitional adjustments as described in AG1, the paper extends the test results to CIs in a succinct fashion.

We now discuss the literature that is related to the methods considered in this paper. The work of Politis and Romano (1994) and Politis, Romano, and Wolf (1999) on subsampling is quite relevant, as is the literature on the $b < n$ bootstrap, see AG1 for references. We are not aware of any methods in the literature that are analogous to the hybrid test or that consider size-correction of subsample or $b < n$ bootstrap methods. Nor are we aware of any general methods of size-correction for FCV tests for the type of non-regular cases considered in this paper. For specific models in the class considered here, however, some methods are available. For example, for CIs based on post-conservative model selection estimators in regression models, Kabaila (1998) suggests a method of size-correction. For models with weak instruments, Anderson and Rubin (1949), Dufour (1997), Staiger and Stock (1997), Kleibergen (2002), Moreira (2001, 2003), Guggenberger and Smith (2005), and Otsu (2006) suggest methods. A variant of Moreira's method also is applicable in predictive regressions with nearly integrated regressors, see Jansson and Moreira (2006). In autoregressive models, CI methods of Stock (1991), Andrews (1993), Horowitz (1997), and Hansen (1999) can be used in place of the least squares estimator combined with normal critical values or subsample critical values. Mikusheva (2005) shows that the former methods yield correct asymptotic size under normality and non-normality of the innovations. (She does not consider Horowitz's method.)

The paper considers three examples in detail. The objectives are to illustrate how the general results of the paper are applied in several models, to determine the asymptotic behavior of subsample, hybrid, and FCV methods in these models, and to see how well these methods and the size-corrected methods work in small samples. In each example, (i) the assumptions required for the general results to hold are verified, (ii) asymptotic and adjusted asymptotic sizes of tests or CIs are computed, (iii) SC, PSC, ASC, and/or APSC values are computed where applicable, and (iv) finite-sample rejection/coverage probabilities are computed. We summarize the main results here.

The first example concerns tests when a nuisance parameter may be near a boundary of the parameter space. Asymptotic results for FCV and subsample tests for this example are considered in AG1. Here we consider hybrid tests, PSC and APSC tests, and finite-sample results. Subsample tests are found to have substantial asymptotic and finite-sample size distortions. For nominal 5% tests the asymptotic (respectively, finite-sample for $n = 120$ and $b = 12$) sizes of upper, symmetric, and equal-tailed tests are 50.2 (49.8), 10.1 (8.4), and 52.3% (52.7%), respectively.

In this example, hybrid tests have asymptotically correct size for all types of tests and have finite-sample sizes that range from 3.4 to 5.2%. FCV tests have asymptotic size distortions only for upper tests and the distortions are very small: 5.8% compared to the nominal size of 5%. PSC and APSC subsample and PSC FCV tests are available based on the usual sample autocorrelation estimator. The upper PSC FCV test has finite-sample size of 5.2%. The PSC subsample tests also perform very well: the finite-sample sizes of the upper, symmetric, and equal-tailed tests are 5.3, 5.1, and 5.5%, respectively. The symmetric APSC subsample test performs well, but the upper and equal-tailed tests over-reject (with finite-sample sizes of 13.5 and 13.5%). To conclude, in this example, subsample tests display poor asymptotic and finite-sample size performance. On the other hand, hybrid, FCV, and PSC subsample tests perform quite well.

The second example concerns CIs in a first-order autoregressive (AR(1)) model with a root that may be near unity. Romano and Wolf (2001) also consider subsampling in this model. We consider two models: model 1 includes an intercept, and model 2 includes an intercept and time trend. The CIs considered are based on the least squares t statistic. In this example, we show that the upper and equal-tailed subsample and lower and two-sided FCV CIs have asymptotic size distortions with the distortions being larger in model 2 than in model 1. (In independent work, Mikusheva (2005) shows that equal-tailed subsample CIs have size-distortions in a no-intercept AR(1) model with normal innovations. Her results do not provide an expression for the asymptotic size.) On the other hand, symmetric subsample tests are shown to have correct asymptotic size. (An explanation is given below.) All types of hybrid tests have correct asymptotic size. We find that the asymptotic approximations for subsample CIs are very poor for a sample size of $n = 130$ and $b = 12$, but the adjusted asymptotic approximations are quite good. For example, the asymptotic, adjusted asymptotic, and finite-sample sizes of nominal 95% equal-tailed subsample CIs in model 1 are 60.1, 86.1, and 86.7%, respectively. In model 1, the finite-sample sizes of the hybrid tests are 94.8, 92.7, 92.7, and 95.6% for upper, lower, symmetric, and equal-tailed CIs, respectively. All of the CIs that have incorrect asymptotic size can be size-corrected. For example, the

ASC subsample upper and equal-tailed CIs have finite-sample sizes of 95.3 and 94.9%, respectively, in model 1.

None of the CIs discussed in the previous paragraph are similar asymptotically or in finite samples. Hence, we expect that these CIs are longer on average than other CIs that are asymptotically similar, see the references above.

The third example is a post-conservative model selection (CMS) example. We consider a LS t test concerning a regression parameter after model selection is used to determine whether another regressor should be included in the model. The model selection procedure uses a LS t test with nominal level 5%. This procedure, which is closely related to AIC, is conservative (i.e., it chooses a correct model, but not necessarily the most parsimonious model, with probability that goes to one). In this example, results for CIs are exactly the same as for tests because of location invariance of the tests. The asymptotic results for FCV tests in the CMS example are variations of those of Leeb (2006) and Leeb and Pötscher (2005) (and other papers referenced in these two papers).

In the CMS example, nominal 5% subsample, FCV, and hybrid tests have asymptotic and adjusted-asymptotic sizes between 90 and 96% for upper, symmetric, and equal-tailed tests.[1] The finite-sample sizes of these tests for $n = 120$ and $b = 12$ are close to the asymptotic values. This is especially true of the adjusted-asymptotic sizes for which the largest deviations are 1.8%. Most plug-in size-corrected tests perform very well in this example. For example, the 5% PSC hybrid and APSC hybrid tests have finite-sample size of 4.8% for upper, lower, and symmetric tests. Also, the PSC FCV tests have finite-sample sizes of 5.1, 5.3, and 5.2% for upper, lower, and symmetric tests, respectively. Hence, the PSC methods of this paper are quite successful in the CMS example.

Andrews and Guggenberger (2005b) applies the results of AG1 and this paper to two additional examples. The first is a weak instrumental variables (IV) example. This example considers tests based on the two-stage least squares (2SLS) estimator of a single included right-hand side (rhs) endogenous variable in a single equation IV regression model when the IVs may be weakly correlated with the rhs endogenous variable. Subsample tests and CIs are shown to have incorrect size asymptotically, but they can be size-corrected. The latter result is of particular interest given Dufour's (1997) result that the 2SLS CI based on a fixed critical value (as well as any CI that has finite length with probability one) has a finite-sample size of zero for all sample sizes. Subsample and size-corrected subsample tests are shown to have infinite length with positive probability in the weak IV example. Hence, the stated results are consistent with those of Dufour.

In the IV example, the asymptotic sizes of the subsample tests are found to provide poor approximations to the finite-sample sizes, but the adjusted asymptotic sizes are accurate. In consequence, the ASC subsample tests perform well. The hybrid test has

[1] This is for a parameter space of $[-.995, .995]$ for the (asymptotic) correlation between the LS estimators of the two regressors.

correct asymptotic size for all types of test in the IV example.

The second example considered in Andrews and Guggenberger (2005b) concerns CIs when the parameter of interest may be near a boundary in a regression model. In contrast to the first example discussed above, the parameter of interest, rather than a nuisance parameter, may be near a boundary. In this example, we find that lower one-sided and equal-tailed two-sided subsample CIs have asymptotic size far below their nominal level. Size-correction of these subsample CIs is possible. However, these CIs exhibit a relatively high degree of non-similarity, which is not desirable from a power perspective. Hybrid and FCV CIs are shown to have correct asymptotic size in this example, but these CIs are not asymptotically similar. The asymptotic results for this example can be generalized to a wide variety of nonlinear models using results in the literature, such as Andrews (1999, 2001).

In conclusion, the examples considered in the paper and in Andrews and Guggenberger (2005b) show that the general results of the paper apply to a wide variety of different models and test statistics. In many models, the hybrid test has correct asymptotic size and finite-sample size that is close to its nominal size. In most models either the asymptotic or adjusted asymptotic results provide good approximations to the finite-sample quantities of interest. In consequence, either the unadjusted or the adjusted size-correction methods introduced in the paper perform quite well. Thus, the examples demonstrate the usefulness of the methods analyzed in the paper.

Throughout the paper $\alpha \in (0,1)$ is a given constant.

The remainder of the paper is outlined as follows. Section 2 introduces the hybrid tests. Section 3 introduces the size-corrected tests. Section 4 compares the asymptotic power of size-corrected FCV, subsample, and hybrid tests. Section 5 introduces the plug-in size-corrected tests. Section 6 introduces the finite-sample adjustments to the asymptotic sizes of subsample and hybrid tests and the adjusted size-corrected tests. Section 7 extends the hybrid and size-correction results to equal-tailed tests. Section 8 extends all of the testing results to confidence intervals. Sections 9 and 10 provide the results for the autoregressive and post-conservative model selection examples. Appendix A provides (i) some assumptions not stated in the text, (ii) size-correction results for equal-tailed tests, and (iii) proofs of the general results. Appendix B gives (i) results for the combined size-corrected subsample and hybrid test and (ii) proofs for the examples, including the verification of assumptions.

2 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 an 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 the fixed critical value test at some null parameter values. In addition, at least in some scenarios, the power of the hybrid test is quite good relative to FCV and subsample tests, see Section 4 below.

We suppose the following assumption holds.

Assumption K. The asymptotic distribution J_h in Assumption B2 of AG1 is the same (proper) distribution, call it J_∞, for all $h = (h_1, h_2) \in H$ for which $h_{1,m} = +\infty$ or $-\infty$ for $m = 1, ..., p$, where $h_1 = (h_{1,1}, ..., h_{1,p})'$.

For notational simplicity, in Assumption K and below we write (h_1, h_2), rather than $(h_1', h_2')'$, even though $h = (h_1, h_2)$ is a $p + q$ column vector. In examples, Assumption K often holds when $T_n(\theta_0)$ is a studentized statistic (i.e., Assumption t1 holds, but t2 does not) or a likelihood ratio (LR), Lagrange multiplier (LM), or Wald 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

$$\begin{aligned} &T_n(\theta_0) > c^*_{n,b}(1-\alpha), \text{ where} \\ &c^*_{n,b}(1-\alpha) = \max\{c_{n,b}(1-\alpha), c_\infty(1-\alpha)\}. \end{aligned} \tag{2.1}$$

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.[2] 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 2 below that the hybrid test of nominal level α has asymptotic level α (i.e., $AsySz(\theta_0) \leq \alpha$) provided the quantile function $c_{(h_1,h_2)}(1-\alpha)$ is maximized at a boundary point of h_1 for each h_2. For example, this occurs if $c_h(1-\alpha)$ is monotone increasing or decreasing in h_1 for each fixed $h_2 \in H_2$, where $h = (h_1, h_2)$ (i.e., $c_{(h_1,h_2)}(1-\alpha) \leq c_{(h_1^*,h_2)}(1-\alpha)$ when $h_1 \leq h_1^*$ element by element or $c_{(h_1,h_2)}(1-\alpha) \geq c_{(h_1^*,h_2)}(1-\alpha)$ when $h_1 \leq h_1^*$).

Define

$$Max^-_{Hyb}(\alpha) = \sup_{(g,h)\in GH} [1 - J_h(\max\{c_g(1-\alpha), c_\infty(1-\alpha)\}-)]. \tag{2.2}$$

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

The following Corollary to Theorems 1(b) and 2(b) of AG1 establishes the asymptotic size of the hybrid test. By definition, $h^0 = (0, h_2)$.

[2] Hybrid tests can be defined even when Assumption K 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 $h = (h_1, h_2) \in H$, $h^\infty = (h^\infty_{1,1}, ..., h^\infty_{1,p}, h^\infty_2) \in H$ 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 K holds, this reduces to the hybrid critical value in (2.1). We utilize Assumption K because it leads to a particularly simple form for the hybrid test.

Corollary 1 (a) *Suppose Assumptions* A1, B1, C-E, F1, G1, *and* K *hold. Then,*

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

(b) *Suppose Assumptions* A2, B2, C-E, F2, G2, *and* K *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 K 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 K is just used to motivate the particular choice of $c_\infty(1-\alpha)$ given above, as the $1-\alpha$ quantile of J_∞. Assumption K also is used in results given below concerning the properties of the hybrid test.

3. Corollary 1 holds by the proofs of Theorems 1(b) and 2(b) of AG1 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 3(b) in Appendix A of AG1.

The following result shows that the hybrid test has better size properties than the subsample test.

Lemma 1 *Suppose Assumptions* A2, B2, C-E, F2, G2, *and* K *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.*

Next, we show that the hybrid test has correct size asymptotically if $c_h(1-\alpha)$ is maximized at h^∞ or is maximized at h^0 and $p = 1$, where p is the dimension of h_1 and $h^\infty = (\infty, h_2)$ or $(-\infty, h_2)$ for $h = (h_1, h_2)$. For example, the maximization 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. (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) $\sup_{h\in H}[1 - J_h(c_h(1-\alpha)-)] = \sup_{h\in H}[1 - J_h(c_h(1-\alpha))]$.

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 2 *Suppose Assumptions* A2, B2, C-E, F2, G2, K, *and* Quant0 *hold. Then, the hybrid test based on* $T_n(\theta_0)$ *has* $AsySz(\theta_0) \leq \alpha$.

(Alternative sufficient conditions for the Hybrid test to have $AsySz(\theta_0) \leq \alpha$ are given in Theorem 2 in Section 4 below, see Comment 1 to Theorem 2).

Figure 1 illustrates the asymptotic critical value (cv) functions of the hybrid, FCV, and subsample tests for the case where $\gamma = \gamma_1 \in R_+$, (i.e., no subvectors γ_2 or γ_3 appear, $p = 1$, and $H = R_{+,\infty}$). The argument of the cv functions is $g \in H$. For example, the asymptotic subsample cv function is $c_g(1-\alpha)$ for $g \in H$. In Figure 1, the curved line is the subsample cv function, the horizontal line is the FCV cv function, i.e., the constant $c_\infty(1-\alpha)$, and the hybrid cv function is the maximum of the two.

In Figure 1(a), the subsample and hybrid cv functions are the same and the corresponding tests have the desired asymptotic size α. (The latter holds because $c_\infty(1-\alpha)$ is $\leq$ the cv function at g for all $g \in R_+$ and $c_0(1-\alpha)$ is $\geq$ the cv function at g for all $g \in R_+$ and these two conditions are necessary and sufficient for a test to have asymptotic size α assuming continuity of $J_h(\cdot)$ by Theorem 2 of AG1). On the other hand, in Figure 1(a), the FCV test has asymptotic size $> \alpha$. In Figures 1(b) and 1(d), the hybrid cv function equals the FCV cv function, both of these tests have asymptotic size α, whereas the subsample test has asymptotic size $> \alpha$. Figures 1(a) and 1(b) illustrate the results of Lemma 1(i) and 1(ii), respectively.

Figure 1(c) illustrates a case where the hybrid test has asymptotic size α, but both the FCV and subsample tests have asymptotic size $> \alpha$. In Figures 1(a)-(d), Assumption Quant0 holds, so the hybrid test has correct asymptotic size, as established in Lemma 2.

Figures 1(e) and 1(f) illustrate cases in which the function $c_g(1-\alpha)$ is maximized at an interior point $g \in (0, \infty)$. In these cases, the hybrid, FCV, and subsample tests all have asymptotic size $> \alpha$. Figures 1(e) and 1(f) illustrate the results of Lemma 1(ii) and 1(i), respectively. In particular, in Figure 1(e), the over-rejection of the subsample test for g close to zero is reduced for the hybrid test because its cv function is larger.

Example 1. This example is a continuation of Example 1 of AG1. It is a testing problem where a nuisance parameter may be on the boundary of the parameter space under the null hypothesis. The observations are $\{X_i \in R^2 : i \leq n\}$, which are i.i.d. with distribution F, $X_i = (X_{i1}, X_{i2})'$, $E_F X_i = (\theta, \mu)'$, and (X_{i1}, X_{i2}) have correlation ρ. The null hypothesis is $H_0 : \theta = 0$, i.e., $\theta_0 = 0$. The parameter space for the nuisance parameter μ is $[0, \infty)$. The test statistic $T_n(\theta_0)$ equals $T_n^*(\theta_0)$, $-T_n^*(\theta_0)$, or $|T_n^*(\theta_0)|$, where $T_n^*(\theta_0)$ is a t statistic based on the Gaussian quasi-ML estimator of θ that imposes the restriction that $\mu \in [0, \infty)$, see AG1 for details. In AG1, Assumptions A2, B2, C-E, F2, and G2 are verified.

Table I reports maximum (over $h_1 = \lim_{n\to\infty} n^{1/2}\mu_{n,h}/\sigma_{n,h,2}$) null rejection probabilities ($\times 100$) for several fixed values of h_2 ($= \lim_{n\to\infty} \rho_{n,h}$) for hybrid and several other nominal 5% tests.[3] Depending on the column, the probabilities are asymptotic

[3] The results in Table I are based on 20,000 simulation repetitions. For the asymptotic results, the search over h_1 is done with stepsize 0.05 on $[0, 10]$ and also includes the two values $h_1 = \pm 9{,}999{,}999{,}999$. For the finite-sample results, the search over h_1 is done with stepsize .001 on $[0, 0.5]$, with stepsize 0.05 on $[0.5, 1.0]$, and with stepsize 1.0 on $[1.0, 10]$. Calculations indicate that

or finite-sample. The finite-sample results are for the case of $n = 120$ and $b = 12$ with $\widehat{\sigma}_{n1}, \widehat{\sigma}_{n2}$, and $\widehat{\rho}_n$ being the sample standard deviations and correlation of X_{i1} and X_{i2}. To dramatically increase computational speed, here and in all of the tables below, finite-sample subsample and hybrid results are based on $q_n = 119$ subsamples of consecutive observations.[4] Hence, only a small fraction of the "120 choose 12" available subsamples are used. In cases where subsample and hybrid tests have correct asymptotic size, their finite-sample performance is expected to be better when all available subsamples are used than when only $q_n = 119$ are used. This should be taken into account when assessing the results of the tables. Panels (a), (b), and (c) of Table I give results for upper one-sided, symmetric two-sided, and equal-tailed two-sided tests, respectively. The results for lower one-sided tests are the same as for the upper tests with the sign of h_2 changed (by symmetry) and, hence, are not given. For convenience, results for equal-tailed hybrid tests are given in Table I and are discussed here even though such tests are only defined in Section 7 below. (The definitions are analogous to those of one-sided and symmetric two-sided hybrid tests.) The rows labelled Max give the size (asymptotic or $n = 120$) of the test considered. For brevity, we refer below to the numbers given in the tables as though they are precise, but of course they are subject to simulation error.

Column 2 of Table I shows that subsample tests have very large asymptotic size distortions for upper one-sided and equal-tailed two-sided tests (nominal 5% tests have asymptotic levels 50.2 and 52.5, respectively), and moderate size distortions for symmetric two-sided tests (the nominal 5% test has asymptotic level 10.1). (These are a subset of results given in Table I of AG1.) Also, column 7 of Table I shows that the FCV tests have very small asymptotic size distortions for upper one-sided tests (the nominal 5% test has asymptotic level 5.8), and no size distortions for symmetric and equal-tailed two-sided tests.

Column 10 of Table I shows that the nominal 5% hybrid test has asymptotic size of 5% for upper, symmetric, and equal-tailed tests. So, the hybrid test has correct asymptotic size for all three types of tests in this example.

Finite-sample results for the Sub, FCV, and Hyb tests are given in columns 4, 8, and 12 of Table I, respectively. For Hyb tests, the asymptotic approximations are fairly accurate, but tend to over-estimate the finite-sample rejection rates somewhat for some values of h_2 with finite-sample values varying between 3.4 and 5.2 compared to the asymptotic values of 5.0. For FCV tests, the asymptotic approximations are found to be very accurate for upper tests and quite accurate for symmetric and equal-tailed

these stepsizes are sufficiently small to yield accuracy to within $\pm.1$.

[4] This includes 10 "wrap-around" subsamples that contain observations at the end and beginning of the sample, for example, observations indexed by $(110, ..., 120, 1)$. The choice of $q_n = 119$ subsamples is made because this reduces rounding errors when q_n is small when computing the sample quantiles of the subsample statistics. The values ν_α that solve $\nu_\alpha/(q_n + 1) = \alpha$ for $\alpha = .025$, $.95$, and $.975$ are the integers 3, 114, and 117. In consequence, the .025, .95, and .975 sample quantiles are given by the 3rd, 114th, and 117th largest subsample statistics. See Hall (1992, p. 307) for a discussion of this choice in the context of the bootstrap.

tests.

The asymptotic approximations for the Sub test are found to be quite good for h_2 values where the (maximum) asymptotic probabilities ($\times 100$) equal 5.0. But, for h_2 values where they exceed 5.0, they tend to over-estimate the finite-sample values—sometimes significantly so, e.g., 33.8 versus 25.6 for $h_2 = -.95$ with upper Sub tests. Nevertheless, in the worst case scenarios (i.e., for h_2 values of 1.0 or -1.0, which yield the greatest asymptotic rejection probabilities), the asymptotic approximations are quite good. Hence, the asymptotic sizes and finite-sample sizes are close—50.2 versus 49.8, 10.1 versus 8.4, and 52.3 versus 52.7 for upper, symmetric, and equal-tailed tests, respectively.

The results in Table I for the columns headed Adj-Asy, PSC-Sub, APSC-Sub, ... are discussed below.

3 Size-Corrected Tests

In this section, we use Theorem 2 of AG1 to define size-corrected (SC) FCV, subsample, and hybrid tests. 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 methods of this section apply to CIs as well, see Section 8 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

$$\begin{aligned} T_n(\theta_0) &> cv(1-\alpha), \\ T_n(\theta_0) &> c_{n,b}(1-\alpha) + \kappa(\alpha) \text{ and} \\ T_n(\theta_0) &> \max\{c_{n,b}(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\}, \end{aligned} \quad (3.1)$$

respectively, where

$$\begin{aligned} cv(1-\alpha) &= \sup_{h \in H} c_h(1-\alpha), \\ \kappa(\alpha) &= \sup_{(g,h) \in GH} [c_h(1-\alpha) - c_g(1-\alpha)], \\ \kappa^*(\alpha) &= \sup_{h \in H^*} c_h(1-\alpha) - c_\infty(1-\alpha), \text{ and} \\ H^* &= \{h \in H : \text{for some } (g,h) \in GH, c_g(1-\alpha) < c_h(1-\alpha)\}. \end{aligned} \quad (3.2)$$

If H^* is empty, then $\kappa^*(\alpha) = -\infty$ by definition.

Size correction by the use of additive constants, as in (3.1), is possible under the following assumption.

Assumption L. (i) $\sup_{h \in H} c_h(1-\alpha) < \infty$ and (ii) $\inf_{h \in H} c_h(1-\alpha) > -\infty$.

Assumption L is satisfied in most, but not all, examples. Assumption L(i) is a necessary and sufficient condition for size correction of the FCV test based on a non-random critical value. Necessary and sufficient conditions for size correction of the Sub and Hyb tests are given in Andrews and Guggenberger (2005b). These conditions are weaker than Assumption L, but more complicated (which is why they are not stated here). Even these conditions are violated in some examples, e.g., in the consistent model selection/super-efficient example in AG1. Size correction of FCV, Sub, and Hyb tests is not possible in that example (by the type of size correction considered here).

It is possible that the FCV test cannot be size-corrected (by the method considered here) because $cv(1-\alpha)=\infty$, but the SC-Sub and SC-Hyb tests still exist and have correct asymptotic size. Also, it is possible that the SC-FCV and SC-Hyb tests exist while the SC-Sub test does not (because $\kappa(\alpha)=\infty$). Surprisingly, both cases arise in the IV example considered in Andrews and Guggenberger (2005b) (depending upon whether one considers symmetric two-sided or upper one-sided tests with $H_2=[-1,0]$). These cases are covered by the necessary and sufficient conditions in Andrews and Guggenberger (2005b).

Next we introduce some fairly mild continuity conditions.

Assumption MF. (i) For some $h^* \in H$, $c_{h^*}(1-\alpha)=\sup_{h\in H} c_h(1-\alpha)$ and (ii) for all $h^*\in H$ that satisfy the condition in part (i), $J_{h^*}(x)$ is continuous at $x=c_{h^*}(1-\alpha)$.

Assumption MS. (i) For some $(g^*,h^*)\in GH$, $c_{h^*}(1-\alpha)-c_{g^*}(1-\alpha)=\sup_{(g,h)\in GH}$ $[c_h(1-\alpha)-c_g(1-\alpha)]$ and (ii) for all $(g^*,h^*)\in GH$ that satisfy the condition in part (i), $J_{h^*}(x)$ is continuous at $x=c_{h^*}(1-\alpha)$.

Assumption MH. (i) When H^* is non-empty, for some $h^*\in H^*$, $c_{h^*}(1-\alpha)=\sup_{h\in H^*}$ $c_h(1-\alpha)$ and (ii) for all $(g,h)\in GH$ with $\max\{c_g(1-\alpha), c_\infty(1-\alpha)+\kappa^*(\alpha)\}=c_h(1-\alpha)$, $J_h(x)$ is continuous at $x=c_h(1-\alpha)$.

The following result shows that the SC tests have $AsySz(\theta_0)$ equal to their nominal level under suitable assumptions.

Theorem 1 (a) *Suppose Assumptions* A2, B2, L, *and* MF *hold. Then, the SC-FCV test has* $AsySz(\theta_0)=\alpha$.

(b) *Suppose Assumptions* A2, B2, C-E, F2, G2, L, *and* MS *hold. Then, the SC-Sub test has* $AsySz(\theta_0)=\alpha$.

(c) *Suppose Assumptions* A2, B2, C-E, F2, G2, K, L, *and* MH *hold. Then, the SC-Hyb test has* $AsySz(\theta_0)=\alpha$.

Comments. 1. The proof of Theorem 1 can be altered slightly to prove that $\lim_{n\to\infty}$ $\sup_{\gamma\in\Gamma} P_{\theta_0,\gamma}(T_n(\theta_0)>cv(1-\alpha))=\alpha$ for the SC-FCV test under the given assumptions (which is a stronger result than $AsySz(\theta_0)=\limsup_{n\to\infty}\sup_{\gamma\in\Gamma} P_{\theta_0,\gamma}(T_n(\theta_0)>cv(1-\alpha))=\alpha$) and analogous results hold for the SC-Sub and SC-Hyb tests.

2. The proof of Theorem 1 shows that the SC-FCV, SC-Sub, and SC-Hyb tests satisfy $AsySz(\theta_0)\le\alpha$ without imposing Assumptions MF(i), MS(i), and MH(i), respectively.

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.

Computation of $\kappa(\alpha) = \sup_{(g,h)\in GH}[c_h(1-\alpha) - c_g(1-\alpha)]$ for the SC-Sub test is similar to that of $cv(1-\alpha)$ except the maximization is over the larger set GH rather than H. This makes the maximization somewhat more difficult. The maximization step may range in difficulty from being very easy to nearly impossible depending on the factors listed above plus the complexity of GH.

Computation of $\kappa^*(\alpha)$ for the SC-Hyb test is analogous to computation of $cv(1-\alpha)$ except that one also has to determine H^*. That is, one has to maximize $c_h(1-\alpha)$ over $h \in H$ subject to the restriction that h satisfies: for some $(g,h) \in GH$, $c_g(1-\alpha) < c_h(1-\alpha)$.

For a given example, one can tabulate $cv(1-\alpha)$, $\kappa(\alpha)$, and $\kappa^*(\alpha)$ for selected values of α. Once this is done, the SC-FCV, SC-Sub, and SC-Hyb tests are as easy to apply as the corresponding non-corrected tests.

An alternative method of size-correcting subsample and hybrid tests to that described above is to adjust the quantile of the test rather than to increase the critical value by a fixed quantity. Specifically, one can define quantile-adjusted SC-Sub and SC-Hyb tests with nominal level α to reject the null hypothesis $H_0 : \theta = \theta_0$ when

$$\begin{aligned} &T_n(\theta_0) > c_{n,b}(1-\xi(\alpha)), \text{ and} \\ &T_n(\theta_0) > c^*_{n,b}(1-\xi^*(\alpha)), \end{aligned} \tag{3.3}$$

respectively, where $\xi(\alpha)$ ($\in (0,\alpha]$), and $\xi^*(\alpha)$ ($\in (0,\alpha]$) are the largest constants[5] that satisfy

$$\begin{aligned} &\sup_{(g,h)\in GH} (1 - J_h(c_g(1-\xi(\alpha))-)) \leq \alpha \text{ and} \\ &\sup_{(g,h)\in GH} (1 - J_h(\max\{c_g(1-\xi^*(\alpha)), c_\infty(1-\xi^*(\alpha))\}-)) \leq \alpha. \end{aligned} \tag{3.4}$$

In many cases, the two different methods of size correction give similar results. For many examples, we prefer the method based on (3.1)-(3.2) to that of (3.3)-(3.4) because the former are based on the explicit formulae for the adjustment factors $\kappa(\alpha)$ and $\kappa^*(\alpha)$ given in (3.2).

[5] If no such largest value exists, we take some value that is arbitrarily close to the supremum of the values that satisfy (3.4).

4 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 $cv(1-\alpha)$, $c_{n,b}(1-\alpha)+\kappa(\alpha)$, and $\max\{c_{n,b}(1-\alpha), c_\infty(1-\alpha)+\kappa^*(\alpha)\}$ 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 of AG1 (and fails when they satisfy Assumption Sub2 of AG1).

From the definitions of the critical values of the SC-Sub and SC-Hyb tests and Lemma 4(e) in Appendix A of AG1, the possible limits of the critical values under sequences $\{\gamma_{n,h}\}$ are

$$c_g(1-\alpha)+\kappa(\alpha) \ \& \ \max\{c_g(1-\alpha), c_\infty(1-\alpha)+\kappa^*(\alpha)\} \text{ for } g \in H. \tag{4.1}$$

Hence, we are interested in the relative magnitudes of $cv(1-\alpha)$ and the quantities in (4.1). These relative magnitudes are determined by the shapes of the quantiles $c_g(1-\alpha)$ as functions of $g \in H$.

The first result is that the SC-Hyb test is always at least as powerful as the SC-FCV test. This holds because for all $g \in H$,

$$\begin{aligned}\max\{c_g(1-\alpha), c_\infty(1-\alpha)+\kappa^*(\alpha)\} &= \max\{c_g(1-\alpha), \sup_{h \in H^*} c_h(1-\alpha)\} \\ &\leq \sup_{h \in H} c_h(1-\alpha) = cv(1-\alpha). \end{aligned} \tag{4.2}$$

The same is not true of the SC-Sub test vis-a-vis the SC-FCV test.

Next, Theorem 2 below shows that (a) if $c_h(1-\alpha) \leq c_g(1-\alpha)$ for all $(g,h) \in GH$, then the SC-Sub, SC-Hyb, Sub, and Hyb tests are equivalent asymptotically and are more powerful than the SC-FCV test, see Figure 2(a); (b) if $c_h(1-\alpha) \geq c_g(1-\alpha)$ for all $(g,h) \in GH$, then the SC-FCV, SC-Hyb, FCV, and Hyb tests are equivalent asymptotically and are more powerful than the SC-Sub test, see Figure 2(b); and (c) if $H = H_1 = R_{+,\infty}$ and $c_h(1-\alpha)$ is uniquely maximized at $h^* \in (0,\infty)$, then the SC-FCV and SC-Hyb tests are asymptotically equivalent and are either (i) more powerful than the SC-Sub test, see Figure 2(e), or (ii) more powerful than the SC-Sub test for some values of $(g,h) \in GH$ but less powerful for other values of $(g,h) \in GH$, see Figure 2(f).

Figure 2(c) illustrates the case where $c_g(1-\alpha)$ is not monotone but is maximized at $g = 0$, the Hyb and SC-Hyb cv functions are the same, the Hyb cv function is lower than both the SC-Sub and SC-FCV cv functions, and so the Hyb test is more powerful than the SC-Sub and SC-FCV tests. Figure 2(d) illustrates the case where $c_g(1-\alpha)$ is not monotone but is maximized at $g = \infty$, the Hyb, SC-Hyb, FCV, and SC-FCV cv functions are the same, the Hyb cv function is lower than the SC-Sub cv function, and so the Hyb test is more powerful than the SC-Sub test.

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. When the SC-FCV test dominates the SC-Sub test, SC-Hyb test behaves like SC-FCV test. In none of the cases considered is the SC-Hyb test dominated by either the SC-FCV or SC-Sub tests.

We now consider three alternative assumptions concerning the shape of $c_h(1-\alpha)$, which correspond to cases (a)-(c) above. ("Quant" refers to "quantile.")

Assumption Quant1. $c_g(1-\alpha) \geq c_h(1-\alpha)$ for all $(g, h) \in GH$.

Assumption Quant2. $c_g(1-\alpha) \leq c_h(1-\alpha)$ for all $(g, h) \in GH$ with strict inequality for some (g, h).

Assumption Quant3. (i) $H = H_1 = R_{+,\infty}$, (ii) $c_h(1-\alpha)$ is uniquely maximized at $h^* \in (0, \infty)$, and (iii) $c_h(1-\alpha)$ is minimized at $h = 0$ or $h = \infty$.

Theorem 2 *Suppose Assumptions* K, L, MF, MS, *and* MH *hold.*

(a) *Suppose Assumption* Quant1 *holds. Then,* (i) $cv(1-\alpha) = \sup_{h_2 \in H_2} c_{(0,h_2)}(1-\alpha)$, (ii) $\kappa(\alpha) = 0$, (iii) $\kappa^*(\alpha) = -\infty$, (iv) $\max\{c_g(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\} = c_g(1-\alpha) + \kappa(\alpha)$, *and* (v) $c_g(1-\alpha) + \kappa(\alpha) \leq cv(1-\alpha)$ *for all* $g \in H$.

(b) *Suppose Assumption* Quant2 *holds. Then,* (i) $cv(1-\alpha) = c_\infty(1-\alpha)$, (ii) $\kappa^*(\alpha) = 0$, (iii) $\max\{c_g(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\} = cv(1-\alpha)$, *and* (iv) $cv(1-\alpha) \leq c_g(1-\alpha) + \kappa(\alpha)$ *for all* $g \in H$.

(c) *Suppose Assumption* Quant3 *holds. Then,* (i) $cv(1-\alpha) = c_{h^*}(1-\alpha)$, (ii) $\kappa(\alpha) = c_{h^*}(1-\alpha) - c_0(1-\alpha)$, (iii) $\kappa^*(\alpha) = c_{h^*}(1-\alpha) - c_\infty(1-\alpha)$, (iv) $\max\{c_g(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\} = cv(1-\alpha)$ *for all* $g \in H$, (v) $cv(1-\alpha) \leq c_g(1-\alpha) + \kappa(\alpha)$ *for all* $g \in H$ *such that* $c_g(1-\alpha) \geq c_0(1-\alpha)$ (*such as* $g = h^*$), *and likewise with strict inequalities, and* (vi) $cv(1-\alpha) > c_g(1-\alpha) + \kappa(\alpha)$ *for all* $g \in H$ *such that* $c_g(1-\alpha) < c_0(1-\alpha)$ (*which is an empty set if* $c_h(1-\alpha)$ *is minimized at* $h = 0$).

Comments. 1. Theorem 2(a)(ii) shows that the standard subsample test (without size correction) has correct asymptotic size when Assumption Quant1 (and other assumptions) hold. Theorem 2(a)(iii) does likewise for the hybrid test. Theorem 2(b)(ii) shows that the hybrid test has correct asymptotic size when Assumption Quant2 (and other assumptions) hold.

2. If Assumption Quant1 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 2(a)(v) holds with a strict inequality with g equal to this value of h. If Assumption Quant2 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 2(b)(iv) holds with a strict inequality with g equal to this value of h.

3. Theorem 2(c)(iv)-(v) shows that under Assumption Quant3 the SC-Hyb and SC-FCV tests are asymptotically equivalent and are always more powerful than the

SC-Sub test at some $(g, h) \in GH$. On the other hand, Theorem 2(c)(vi) shows that under Assumption Quant3 the SC-Sub test can be more powerful than the SC-Hyb and SC-FCV tests at some $(g, h) \in GH$ though not if $c_h(1-\alpha)$ is minimized at $h = 0$.

The results above are relevant when the subsample statistics satisfy Assumption Sub1 of AG1 (because then their asymptotic distribution 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 and SC-Hyb tests for fixed alternatives. For brevity, we do not investigate the relative magnitudes of the critical values of the SC-FCV, SC-Sub, and SC-Hyb tests for local alternatives when Assumption Sub2 holds.

5 Plug-in Size-Corrected Tests

In this section, we introduce improved size-correction methods that exploit the fact that in many models it is possible to consistently estimate the parameter γ_2. As defined in AG1, γ_2 is a parameter that does not determine how close γ is to a point of discontinuity (of the limit distribution J_h of the test statistic of interest), but it affects the limit distribution J_h. Given a consistent estimator $\widehat{\gamma}_{n,2}$ of γ_2, one can size correct a test differently for different values of $\widehat{\gamma}_{n,2}$, rather than size correcting by a value that is sufficiently large to work uniformly for all $\gamma_2 \in \Gamma_2$. By size correcting based on $\widehat{\gamma}_{n,2}$, one obtains critical values that are smaller for some values of γ_2 and, hence, this yields a more powerful test.

The estimator $\widehat{\gamma}_{n,2}$ is assumed to satisfy the following assumption.

Assumption N. $\widehat{\gamma}_{n,2} - \gamma_{n,2} \to_p 0$ under all sequences $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) \in \Gamma : n \geq 1\}$.

Assumption N holds in most models when a parameter γ_2 appears in γ. However, it does not hold in an IV regression model with IVs that may be weak because γ_2 equals the correlation between the structural equation error and the reduced form error in this case, and this correlation is not consistently estimable under weak IVs.

Define

$$\begin{aligned} cv_{h_2}(1-\alpha) &= \sup_{h_1 \in H_1} c_{(h_1,h_2)}(1-\alpha), \\ \kappa_{h_2}(\alpha) &= \sup_{g_1, h_1 \in H_1 : ((g_1,h_2),(h_1,h_2)) \in GH} (c_{(h_1,h_2)}(1-\alpha) - c_{(g_1,h_2)}(1-\alpha)), \text{ and} \\ \kappa^*_{h_2}(\alpha) &= \sup_{h_1 \in H^*_{h_2}} c_{(h_1,h_2)}(1-\alpha) - c_\infty(1-\alpha), \text{ where} \\ H^*_{h_2} &= \{h_1 \in H_1 : \text{for some } g_1 \in H_1,\ (g,h) = ((g_1,h_2),(h_1,h_2)) \in GH, \\ & \quad \& \ c_g(1-\alpha) < c_h(1-\alpha)\}. \end{aligned} \tag{5.1}$$

If $H^*_{h_2}$ is empty, then $\kappa^*_{h_2}(\alpha) = -\infty$. The PSC-FCV, PSC-Sub, and PSC-Hyb tests are defined as in (3.1) with $cv(1-\alpha)$, $\kappa(\alpha)$, and $\kappa^*(\alpha)$ replaced by $cv_{\widehat{\gamma}_{n,2}}(1-\alpha)$, $\kappa_{\widehat{\gamma}_{n,2}}(\alpha)$, and $\kappa^*_{\widehat{\gamma}_{n,2}}(\alpha)$, respectively.

Clearly, $cv_{\widehat{\gamma}_{n,2}}(1-\alpha) \leq cv(1-\alpha)$ (with strict inequality whenever $\widehat{\gamma}_{n,2}$ takes a value that does not maximize $cv_{h_2}(1-\alpha)$ over $h_2 \in H_2$). In consequence, the PSC-FCV test is asymptotically more powerful than the SC-FCV test. Analogous results hold for the critical values and asymptotic power of the PSC-Sub and PSC-Hyb tests relative to the SC-Sub and SC-Hyb tests.

For brevity, some continuity conditions, denoted Assumptions OF, OS, and OH, that are used in Theorem 3 below, are stated in the Appendix A. Parts (ii) and (iii) of these conditions are analogous to Assumptions MF, MS, and MH. These conditions are not restrictive in most examples.

Theorem 3 (a) *Suppose Assumptions* A2, B2, L, *and* OF *hold. Then,* (i) $cv_{\widehat{\gamma}_{n,2}}(1-\alpha) - cv_{\gamma_{n,2}}(1-\alpha) \to_p 0$ *under all sequences* $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) \in \Gamma : n \geq 1\}$, *and* (ii) *the PSC-FCV test satisfies* $AsySz(\theta_0) = \alpha$.

(b) *Suppose Assumptions* A2, B2, C-E, F2, G2, L, *and* OS *hold. Then,* (i) $\kappa_{\widehat{\gamma}_{n,2}}(1-\alpha) - \kappa_{\gamma_{n,2}}(1-\alpha) \to_p 0$ *under all sequences* $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) \in \Gamma : n \geq 1\}$, *and* (ii) *the PSC-Sub test satisfies* $AsySz(\theta_0) = \alpha$.

(c) *Suppose Assumptions* A2, B2, C-E, F2, G2, K, L, *and* OH *hold. Then,* (i) $\kappa^*_{\widehat{\gamma}_{n,2}}(1-\alpha) - \kappa^*_{\gamma_{n,2}}(1-\alpha) \to_p 0$ *under all sequences* $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) \in \Gamma : n \geq 1\}$, *and* (ii) *the PSC-Hyb test satisfies* $AsySz(\theta_0) = \alpha$.

Example 1 (cont.). The upper, symmetric, and equal-tailed subsample tests and the upper FCV test need size-correction in this example. Plug-in size correction is possible because estimation of the correlation parameter ρ is straightforward using the usual sample correlation estimator. Columns 5 and 9 of Table I provide the finite-sample (maximum) rejection probabilities ($\times 100$) of the nominal 5% PSC-Sub and PSC-FCV tests. Results for the symmetric and equal-tailed PSC-FCV tests are not given because the PSC-FCV and FCV tests are the same in these cases since the FCV test has correct asymptotic size. Results for the PSC-Hyb test are not given because it is the same as the Hyb test.

The results for the PSC-Sub tests are impressive. The finite-sample sizes of the upper, symmetric, and equal-tailed tests are 5.3, 5.1, and 5.5%, respectively, whereas the finite-sample sizes of the Sub tests are 49.8, 8.4, and 52.7%. The plug-in feature of the size-correction method yields (maximum) rejection probabilities ($\times 100$) across different h_2 values that are all reasonably close to 5.0—ranging from 3.1 to 5.5, with most being between 4.5 and 5.5. Having these values all close to 5% is desirable from a power perspective.

The upper FCV test only requires minor size-correction given that its asymptotic and finite-sample size is 5.8%. The PSC-FCV test provides improvement. Its finite-sample size is 5.2%.

6 Finite-Sample Adjustments

In this section, we introduce a finite-sample adjustment to the $AsySz(\theta_0)$ of subsample and hybrid tests. It is designed to give a better approximation to the actual finite-sample sizes of these tests than does $AsySz(\theta_0)$. The adjustments are used to construct finite-sample adjusted size-corrected (ASC) subsample and hybrid tests, both with and without plug-in estimation of h_2. The idea of the adjustment is to retain the actual ratio $\delta_n = b_n/n$ of the subsample size to the full-sample size in the approximation to the finite-sample size of the tests, rather than to use its asymptotic limit, which is zero.

The adjustment method is described roughly as follows. For simplicity, consider the case in which γ does not contain subvectors γ_2 or γ_3, $p = 1$, and $\Gamma = [0, d]$ for some $0 < d < \infty$. Under Assumption B2, the distribution of $T_n(\theta_0)$ under γ can be approximated by J_{h_n}, where $h_n = n^r\gamma$. Hence, the distribution of $T_{b_n}(\theta_0)$ under γ can be approximated by $J_{h_n^*}$, where $h_n^* = b_n^r\gamma = (b_n/n)^r h_n = \delta_n^r h_n$. In turn, the $1-\alpha$ subsample quantile $c_{n,b}(1-\alpha)$ under γ can be approximated by the $1-\alpha$ quantile of $J_{h_n^*} = J_{\delta_n^r h_n}$, viz., $c_{\delta_n^r h_n}(1-\alpha)$. This leads to the approximation of $P_{\theta_0,\gamma}(T_n(\theta_0) > c_{n,b}(1-\alpha))$ by

$$1 - J_{h_n}(c_{\delta_n^r h_n}(1-\alpha)). \qquad (6.1)$$

And it leads to the approximation of $\sup_{\gamma\in\Gamma} P_{\theta_0,\gamma}(T_n(\theta_0) > c_{n,b}(1-\alpha))$ by

$$AsySz_n(\theta_0) = \sup_{h\in H}\left(1 - J_h(c_{\delta_n^r h}(1-\alpha))\right). \qquad (6.2)$$

Suppose $J_h(c_g(1-\alpha))$ is a continuous function of (g,h) at each $(g,h) \in GH$ and Assumption C(ii) holds, i.e., $\delta_n = b_n/n \to 0$. Then, as $n \to \infty$ the quantity in (6.1) approaches $1 - J_h(c_0(1-\alpha))$ if $h_n \to h \in [0,\infty)$. It approaches $1 - J_\infty(c_g(1-\alpha))$ if $h_n \to \infty$ and $\delta_n^r h_n \to g \in [0,\infty]$. Hence, for any $(g,h) \in GH$, $\lim_{n\to\infty}(1 - J_{h_n}(c_{\delta_n^r h_n}(1-\alpha))) = 1 - J_h(c_g(1-\alpha))$ for a suitable choice of $\{h_n \in H : n \geq 1\}$. This suggests that

$$\lim_{n\to\infty}\sup_{h\in H}(1 - J_h(c_{\delta_n^r h}(1-\alpha))) = \sup_{(g,h)\in GH}(1 - J_h(c_g(1-\alpha))) = AsySz(\theta_0). \qquad (6.3)$$

It is shown below that (6.3) does hold, which implies that $AsySz_n(\theta_0)$ is an asymptotically valid finite-sample adjustment to $AsySz(\theta_0)$.

We now consider the general case in which γ may contain subvectors γ_2 and γ_3 and $p \geq 1$. In this case, only the subvector γ_1 affects whether γ is near a discontinuity point of the limit distribution. In consequence, only h_1, and not h_2, is affected by the δ_n^r rescaling that occurs above. For a subsample test, we define

$$AsySz_n(\theta_0) = \sup_{h=(h_1,h_2)\in H}\left(1 - J_h(c_{(\delta_n^r h_1, h_2)}(1-\alpha))\right). \qquad (6.4)$$

We use the following continuity assumptions.

Assumption P. (i) The function $(g, h) \to J_h(c_g(1-\alpha))$ for $(g, h) \in H \times H$ is continuous at all $(g, h) \in GH$ and (ii) $Max_{Sub}(\alpha) = Max^-_{Sub}(\alpha)$, where $Max_{Sub}(\alpha)$ and $Max^-_{Sub}(\alpha)$ are defined in (6.3) of AG1.

Under the assumptions of Theorem 2(b) of AG1 and Assumption P(ii), the subsample test has $AsySz(\theta_0) = \alpha$ by Theorem 2(b) of AG1.

The following result shows that $AsySz_n(\theta_0)$ provides an asymptotically valid finite-sample adjustment to $AsySz(\theta_0)$ that depends explicitly on the ratio $\delta_n = b_n/n$.

Theorem 4 *Suppose Assumptions* A2, B2, C-E, F2, G2, *and* P *hold. Then, a subsample test satisfies*

$$\lim_{n\to\infty} AsySz_n(\theta_0) = AsySz(\theta_0).$$

Comment. An analogous result holds for the hybrid test with $c_{(\delta^r h_1, h_2)}(1-\alpha)$ replaced by $\max\{c_{(\delta^r h_1, h_2)}(1-\alpha), c_\infty(1-\alpha)\}$ in (6.4).

Next, we use the finite-sample adjustment to construct adjusted SC-Type and PSC-Type tests for Type = Sub and Hyb, which are denoted ASC-Type and APSC-Type tests. For $\delta \in (0, 1)$ and $h_2 \in H_2$, define

$$\begin{aligned}
\kappa(\delta, \alpha) &= \sup_{h=(h_1,h_2)\in H} [c_{(h_1,h_2)}(1-\alpha) - c_{(\delta^r h_1, h_2)}(1-\alpha)], \\
\kappa_{h_2}(\delta, \alpha) &= \sup_{h_1 \in H_1} [c_{(h_1,h_2)}(1-\alpha) - c_{(\delta^r h_1, h_2)}(1-\alpha)], \\
\kappa^*(\delta, \alpha) &= \sup_{h \in H^*(\delta)} c_h(1-\alpha) - c_\infty(1-\alpha), \text{ and} \\
\kappa^*_{h_2}(\delta, \alpha) &= \sup_{h_1 \in H^*_{h_2}(\delta)} c_{(h_1,h_2)}(1-\alpha) - c_\infty(1-\alpha), \text{ where} \\
H^*(\delta) &= \{h \in H : c_{(\delta^r h_1, h_2)}(1-\alpha) < c_{(h_1,h_2)}(1-\alpha) \text{ for } h = (h_1, h_2)\}, \\
H^*_{h_2}(\delta) &= \{h_1 \in H_1 : c_{(\delta^r h_1, h_2)}(1-\alpha) < c_{(h_1,h_2)}(1-\alpha)\}.
\end{aligned} \tag{6.5}$$

If $H^*(\delta)$ is empty, then $\kappa^*(\delta, \alpha) = -\infty$. If $H^*_{h_2}(\delta)$ is empty, then $\kappa^*_{h_2}(\delta, \alpha) = -\infty$. The ASC-Sub and ASC-Hyb tests are defined as in (3.1) with $\kappa(\alpha)$ and $\kappa^*(\alpha)$ replaced by $\kappa(\delta_n, \alpha)$ and $\kappa^*(\delta_n, \alpha)$, respectively, where $\delta_n = b_n/n$. The APSC-Sub and APSC-Hyb tests are defined as in (3.1) with $\kappa(\alpha)$ and $\kappa^*(\alpha)$ replaced by $\kappa_{\widehat{h}_{2,n}}(\delta_n, \alpha)$ and $\kappa^*_{\widehat{h}_{2,n}}(\delta_n, \alpha)$, respectively.

We employ the following assumptions.

Assumption Q. $c_h(1-\alpha)$ is continuous in h on H.

Assumption R. Either H^* is non-empty and $\sup_{h \in H^\dagger} c_h(1-\alpha) \le \sup_{h \in H^*} c_h(1-\alpha)$, where $H^\dagger = \{h \in H : h = \lim_{k\to\infty} h_{v_k}$ for some subsequence $\{v_k\}$ and some $h_{v_k} \in H^*(\delta_{v_k})$ for all $k \ge 1\}$, or H^* is empty and $H^*(\delta)$ is empty for all $\delta > 0$ sufficiently close to zero.

Assumption S. For all $h_2 \in H_2$, either $H^*_{h_2}$ is non-empty and $\sup_{h_1 \in H^\dagger_{h_2}} c_{(h_1,h_2)}(1-\alpha) \leq \sup_{h_1 \in H^*_{h_2}(\delta)} c_{(h_1,h_2)}(1-\alpha)$, where $H^\dagger_{h_2} = \{h_1 \in H_1 : h_1 = \lim_{k\to\infty} h_{v_k,1}$ for some subsequence $\{v_k\}$ and some $h_{v_k,1} \in H^*_{\gamma_{v_k,2}}(\delta_{v_k})$ for all $k \geq 1$, where $\lim_{k\to\infty} \gamma_{v_k,2} = h_2\}$, or $H^*_{h_2}$ is empty and $H^*_{h_2}(\delta)$ is empty for all $\delta > 0$ sufficiently close to zero.

Assumptions Q, R, and S are not restrictive in most examples. Whether Assumptions R and S hold depends primarily on the shape of $c_h(1-\alpha)$ as a function of h. It is possible for Assumptions R and S to be violated, but only for quite specific and unusual shapes for $c_h(1-\alpha)$. For example, Assumption R is violated in the case where $p = 1$ and no parameter h_2 exists if for some $h^* \in (0,\infty)$ the graph of $c_h(1-\alpha)$ is (i) bowl-shaped for $h \in [0, h^*]$ with $c_0(1-\alpha) = c_{h^*}(1-\alpha)$ and (ii) strictly decreasing for $h > h^*$ with $c_\infty(1-\alpha) < c_h(1-\alpha)$ for all $0 \leq h < \infty$. In this case, we have H^* is empty (because $c_h(1-\alpha)$ takes on its minimum for $h = \infty$ and its maximum at $h = 0$), but $h^* \in H^*(\delta)$ for all $\delta \in (0,1)$, which contradicts Assumption R.

The ASC and APSC tests have $AsySz(\theta_0) = \alpha$, as desired, by the following Theorems.

Theorem 5 *Suppose Assumptions* A2, B2, C-E, F2, G2, K, L, *and* Q *hold.*

(a) *Suppose Assumption* MS *holds. Then,* (i) $\lim_{n\to\infty} \kappa(\delta_n, \alpha) = \kappa(\alpha)$ *and* (ii) *the ASC-Sub test satisfies* $AsySz(\theta_0) = \alpha$.

(b) *Suppose Assumption* MH *holds. Then,* (i) $\liminf_{n\to\infty} \kappa^*(\delta_n, \alpha) \geq \kappa^*(\alpha)$, (ii) *the ASC-Hyb test satisfies* $AsySz(\theta_0) \leq \alpha$, *and* (iii) *if Assumption* R *also holds, then* $\lim_{n\to\infty} \kappa^*(\delta_n, \alpha) = \kappa^*(\alpha)$ *and the ASC-Hyb test satisfies* $AsySz(\theta_0) = \alpha$.

Theorem 6 *Suppose Assumptions* A2, B2, C-E, F2, G2, K, L, N, *and* Q *hold.*

(a) *Suppose Assumption* OS *holds. Then,* (i) $\kappa_{\widehat{\gamma}_{n,2}}(\delta_n, \alpha) - \kappa_{\gamma_{n,2}}(\alpha) \to_p 0$ *under all sequences* $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) \in \Gamma : n \geq 1\}$ *and* (ii) *the APSC-Sub test satisfies* $AsySz(\theta_0) = \alpha$.

(b) *Suppose Assumptions* OH *and* S *hold. Then,* (i) $\kappa^*_{\widehat{\gamma}_{n,2}}(\delta_n, \alpha) - \kappa^*_{\gamma_{n,2}}(\alpha) \to_p 0$ *under all sequences* $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) \in \Gamma : n \geq 1\}$, *and* (ii) *the APSC-Hyb test satisfies* $AsySz(\theta_0) = \alpha$.

Comment. Assumption R is a necessary and sufficient condition for the first result of Theorem 5(b)(iii) to hold given the other assumptions.

Example 1 (cont.). Column 3 of Table I gives the finite-sample adjusted asymptotic rejection probabilities ($\times 100$) of the subsample test. These values are noticeably closer to the finite-sample values given in column 4 than are the (unadjusted) asymptotic rejection probabilities given in column 2. For example, for the upper subsample test and $h_2 = -.95$, the values for Adj-Asy, $n = 120$, and Asy are 22.9, 25.6, and 33.8%, respectively. Hence, the adjustment works pretty well for the subsample test here. For the hybrid test, the adjusted asymptotic and unadjusted asymptotic rejection rates are all 5.0%. So, the adjustment makes no difference for the hybrid test in this example.

Column 6 of Table I reports the finite-sample rejection probabilities of the APSC-Sub test. For upper and equal-tailed tests, the adjustment leads to over-correction of the Sub test when the finite-sample correlation, denoted here by h_2, is close to -1 and 1, respectively, and appropriate size-correction for other values of h_2. In consequence, for these cases the PSC-Sub test (see column 5) has better finite-sample size (viz., 5.3 and 5.5%) than the APSC-Sub test (13.5 and 13.5%). For symmetric tests, both of these size-corrected tests perform well.

In conclusion, in this example, the hybrid and PSC-Sub tests perform quite well in terms of finite-sample size for upper, symmetric, and equal-tailed tests. The APSC-Sub test performs well for symmetric test, but not so well for upper and equal-tailed tests.

7 Equal-Tailed Tests

This section considers *equal-tailed* two-sided *hybrid* t tests. For brevity, equal-tailed size-corrected t tests and finite-sample-adjusted asymptotics for equal-tailed tests are discussed in Appendix A. 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 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

$$\begin{aligned} T_n(\theta_0) &> c^*_{n,b}(1-\alpha/2) \text{ or } T_n(\theta_0) < c^{**}_{n,b}(\alpha/2), \text{ where} \\ c^*_{n,b}(1-\alpha/2) &= \max\{c_{n,b}(1-\alpha/2), c_\infty(1-\alpha/2)\} \text{ and} \\ c^{**}_{n,b}(\alpha/2) &= \min\{c_{n,b}(\alpha/2), c_\infty(\alpha/2)\}. \end{aligned} \tag{7.1}$$

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 in Section 7 of AG1, 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 $c_g(\alpha/2)$.

The proofs of Theorems 1 and 2 of AG1 can be adjusted straightforwardly to yield the following results for equal-tailed hybrid t tests.

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

(a) *Suppose Assumptions* A1, B1, C-E, G1, J1, *and* K *hold. Then, an equal-tailed hybrid test satisfies*

$$\begin{aligned} &P_{\theta_0,\gamma_{n,h}}(T_n(\theta_0) > c^*_{n,b}(1-\alpha/2) \text{ or } T_n(\theta_0) < c^{**}_{n,b}(\alpha/2)) \\ \to &[1 - J_h(\max\{c_{h^0}(1-\alpha/2), c_\infty(1-\alpha/2)\}) + J_h(\min\{c_{h^0}(\alpha/2), c_\infty(\alpha/2)\}-), \\ &1 - J_h(\max\{c_{h^0}(1-\alpha/2), c_\infty(1-\alpha/2)\}-) + J_h(\min\{c_{h^0}(\alpha/2), c_\infty(\alpha/2)\}]. \end{aligned}$$

(b) *Suppose Assumptions* A2, B2, C-E, G2, J2, *and* K *hold. Then, an equal-tailed hybrid t test satisfies*

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

8 Confidence Intervals

This section introduces hybrid CIs and size-corrected CIs.

Hybrid CIs are defined just as subsample CIs are defined in (8.1) and (8.2) of AG1 with the critical value $c_{n,b}(1-\alpha)$ replaced by the hybrid critical value $\max\{c_{n,b}(1-\alpha), c_\infty(1-\alpha)\}$. When analyzing the properties of these CIs the parameters (θ,γ) are adjusted as in Section 8 of AG1 so that θ is a subvector of γ, rather than a separate parameter, in order to recycle the assumptions and results obtained for tests and apply them to CIs. The assumptions are adjusted correspondingly as described in Section 8.1 of AG1.

Hybrid CIs satisfy the following results.

Corollary 3 *Let the assumptions be adjusted for* CI*s as in Section* 8 *of* AG1.

(a) *Suppose Assumptions* A1, B1, C-E, F1, G1, *and* K *hold. Then, the hybrid* CI *satisfies* $P_{\gamma_{n,h}}(T_n(\theta_{n,h}) \leq c^*_{n,b}(1-\alpha)) \rightarrow [J_h(\max\{c_{h^0}(1-\alpha), c_\infty(1-\alpha)\}-), J_h(\max\{c_{h^0}(1-\alpha), c_\infty(1-\alpha)\})]$.

(b) *Suppose Assumptions* A2, B2, C-E, F2, G2, *and* K *hold. Then, the hybrid* CI *satisfies* $AsyCS \in [\inf_{(g,h)\in GH} J_h(\max\{c_g(1-\alpha), c_\infty(1-\alpha)\}-), \inf_{(g,h)\in GH} J_h(\max\{c_g(1-\alpha), c_\infty(1-\alpha)\})]$.

An equal-tailed hybrid t CI for θ of nominal level α is defined as in (8.5) of AG1 with the critical values $c_{1-\alpha/2}$ and $c_{\alpha/2}$ given by $c^*_{n,b}(1-\alpha/2)$ and $c^{**}_{n,b}(\alpha/2)$, defined in (7.1), respectively. We have the following result for such CIs.

Corollary 4 *Let* $\alpha \in (0,1/2)$ *be given. Let the assumptions be adjusted for* CI*s as described in Section* 8 *of* AG1. *Suppose Assumptions* A2, B2, C-E, G2, J2, *and* K *hold. Then, the equal-tailed hybrid* t *CI satisfies*

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

An analogue of Theorem 4 holds regarding the finite-sample-adjusted asymptotic sizes of subsample and hybrid CIs. In this case, $AsyCS_n$ is defined as $AsySz_n$ is defined in (6.4) (or as in (12.5) for equal-tailed tests) but with $\sup_{h\in H}$ replaced by $\inf_{h\in H}$ and J_h replaced by $1-J_h$. For example, for upper, lower, and symmetric subsample tests, $AsyCS_n = \inf_{h=(h_1,h_2)\in H} J_h(c_{(\delta^r_n h_1, h_2)}(1-\alpha))$.

Next, we consider size-corrected CIs. Size-corrected FCV, subsample, and hybrid CIs are defined as FCV, subsample, and hybrid CIs are defined in (8.1)-(8.2) of AG1, but with their critical values, $c_{1-\alpha}$, defined as in (3.1)-(3.2) for SC tests. The SC CIs satisfy the following properties, which follow from Theorem 1 using the same sort of argument as in Section 8 of AG1.

Corollary 5 *Let the assumptions be adjusted for* CI*s as in Section* 8 *of* AG1.

(a) *Suppose Assumptions* A2, B2, L, *and* MF *hold. Then, the SC-FCV CI has* $AsyCS = 1-\alpha$.

(b) *Suppose Assumptions* A2, B2, C-E, F2, G2, L, *and* MS *hold. Then, the SC-Sub CI has* $AsyCS = 1 - \alpha$.

(c) *Suppose Assumptions* A2, B2, C-E, F2, G2, K, L, *and* MH *hold. Then, the SC-Hyb CI has* $AsyCS = 1 - \alpha$.

Definitions and results for CIs of the form PSC-Type for Type = FCV, Sub, and Hyb, and ASC-Type and APSC-Type for Type = Sub and Hyb are analogous to those just stated for SC CIs but with critical values as defined in Sections 5 and 6, rather than as in Section 3. Size-corrected equal-tailed CIs are defined as in (8.5) of AG1 with critical values $c_{1-\alpha/2}$ and $c_{\alpha/2}$ given by the equal-tailed SC, PSC, ASC, and/or APSC critical values for tests given in Appendix A.

9 CI for an Autoregressive Parameter

We now consider an example of the general results above. We consider FCV, subsample, and hybrid CIs for the autoregressive parameter ρ in a first-order autoregressive (AR(1)) model with $\rho \in [-1 + \varepsilon, 1]$ for some $0 < \varepsilon < 2$. We also consider size-corrected versions of these CIs. The CIs are based on inverting a (studentized) t statistic constructed using the LS estimator of ρ.

Given that the parameter space for ρ includes a unit root and near unit roots, standard FCV methods for constructing CIs based on a standard normal approximation to the t statistic are known to be problematic. As an alternative, Romano and Wolf (2001) propose subsample CIs for ρ. Mikusheva (2005, Theorem 4) shows that equal-tailed versions of such subsample CIs under-cover the true value asymptotically (i.e., $AsyCS < 1 - \alpha$). Her results do not provide an expression for $AsyCS$. Also, they do not apply to symmetric subsample CIs. In contrast, Corollary 3 allows one to explicitly calculate $AsyCS$ for subsample CIs—both equal-tailed and symmetric—as well as for FCV and hybrid CIs. Thus, one can quantify the magnitude of the problem with the standard FCV CI based on the normal approximation, as well as determine the properties of subsample CIs. Furthermore, we can size-correct the subsample CIs based on both the unadjusted and finite-sample-adjusted asymptotic formulae given in Section 8. (Also note that Mikusheva (2005) considers a model with no intercept or time trend, whereas we consider models with an intercept and an intercept and time trend.)

We consider two versions of the AR(1) model—model 1, which has an intercept, and model 2, which has an intercept and time trend. We consider upper and lower one-sided and symmetric and equal-tailed two-sided CIs for ρ. We summarize the asymptotic and finite-sample results for this example here. (1) Lower one-sided and two-sided FCV CIs and upper one-sided and equal-tailed two-sided subsample CIs have asymptotic sizes far below their nominal level. The sizes are noticeably lower for model 2 than model 1. (2) Upper FCV, all types of hybrid CIs, and lower and symmetric subsample CIs have asymptotic sizes equal to their nominal levels (up to simulation error). (3) The finite-sample-adjusted asymptotic sizes for the subsample CIs that under-cover the true parameter provide much better approximations to the actual finite-samples sizes when $n = 130$ and $b = 12$ than do the unadjusted asymptotic sizes. (4) All of the CIs that have incorrect asymptotic size can be size-corrected. The ASC-Sub, SC-FCV, and hybrid CIs all have finite-sample sizes that are good. For model 1, the sizes vary between 92.7 and 95.3 for 95% CIs. For model 2, they vary between 92.4 and 95.8%.

We note that the same sort of issues that arise with subsampling in the AR(1) model also arise in vector autoregressive models with roots that may be near unity. For example, they arise with subsample tests of Granger causality in such models, see Choi (2005).

We now provide the details concerning the AR(1) example. We use the unobserved components representations of the two AR(1) models. The observed time series $\{Y_i : i = 0, ..., n\}$ is based on a latent no-intercept AR(1) time series $\{Y_i^* : i = 0, ..., n\}$:

$$\begin{aligned} Y_i &= \alpha + \beta i + Y_i^*, \\ Y_i^* &= \rho Y_{i-1}^* + U_i, \text{ for } i = 1, ..., n, \end{aligned} \tag{9.1}$$

where $\{U_i : i = 0, \pm 1, \pm 2, ...\}$ are i.i.d. with mean 0, variance $\sigma_U^2 \in (0, \infty)$, and distribution F. The distribution of Y_0^* is the distribution that yields strict stationarity for $\{Y_i^* : i \leq n\}$ when $\rho < 1$, i.e., $Y_0^* = \sum_{j=0}^{\infty} \rho^j U_{-j}$, and is arbitrary when $\rho = 1$. Model 1 is obtained by setting $\beta = 0$. Model 2 is as in (9.1). In the notation of AG1, we have $\theta = 1 - \rho \in \Theta = [0, 2 - \varepsilon]$.

Models 1 and 2 can be rewritten as

$$\begin{aligned} &(1)\ Y_i = \widetilde{\alpha} + \rho Y_{i-1} + U_i, \text{ where } \widetilde{\alpha} = \alpha(1-\rho), \text{ and} \\ &(2)\ Y_i = \overline{\alpha} + \overline{\beta} i + \rho Y_{i-1} + U_i, \text{ where } \overline{\alpha} = \alpha(1-\rho) + \rho\beta \text{ and } \overline{\beta} = \beta(1-\rho), \end{aligned} \tag{9.2}$$

for $i = 1, ..., n$.[6]

Under the null hypothesis that $\rho = \rho_n = 1 - \theta_n$, the studentized t statistic is given by

$$T_n^*(\theta_n) = \tau_n(\widehat{\rho} - \rho_n)/\widehat{\sigma}, \tag{9.3}$$

where $\tau_n = n^{1/2}$, $\widehat{\rho}$ is the LS estimator of ρ in model 1 or 2 in (9.2), and $\widehat{\sigma}$ is the usual LS standard deviation estimator. Specifically, $\widehat{\sigma}^2$ is the diagonal element of $\widehat{\sigma}_U^2(n^{-1}\sum_{i=1}^n X_i X_i')^{-1}$corresponding to Y_{i-1}, where X_i is the ith regressor vector in model 1 or 2 of (9.2), and $\widehat{\sigma}_U^2$ is the sum of squared residuals divided by $n - k$, where k is the number of regressors in the model. For upper one-sided, lower one-sided, and symmetric two-sided tests or CIs concerning ρ, we take $T_n(\theta_n) = T_n^*(\theta_n)$, $-T_n^*(\theta_n)$, and $|T_n^*(\theta_n)|$, respectively.

We assume that $b_n^2/n = O(1)$ and $n/b_n^2 = O(1)$. (These assumptions are used below to establish Assumptions HH and E and EE, respectively.)

[6] The advantage of writing the model as in (9.1) becomes clear here. For example, in model 1, the case $\rho = 1$ and $\widetilde{\alpha} \neq 0$ is automatically ruled out by model (9.1). This is a case where Y_i is dominated by a deterministic trend and the LS estimator of ρ converges at rate $n^{3/2}$.

In the notation of AG1, the vector of parameters is $\gamma = (\gamma_1, \gamma_3)$, where $\gamma_1 = \theta$ ($= 1 - \rho$), in model 1 $\gamma_3 = (\alpha, F)$, and in model 2 $\gamma_3 = (\alpha, \beta, F)$. (Note that AG1 discusses CIs for θ, which is an element of γ, whereas here we consider CIs for $\rho = 1 - \theta$, which is not an element of γ. However, a CI for θ immediately yields one for ρ.) No parameters γ_2, θ_2, η_1, or η_2 appear in this example. The distribution of the initial condition Y_0^* does not appear in γ_3 because under strict stationarity it equals the stationary marginal distribution of U_i and that is completely determined by F and γ_1 and in the unit root case is irrelevant. The parameter spaces are $\Gamma_1 = \Theta = [0, 2 - \varepsilon]$ and $\Gamma_3 = B_1 \times \mathcal{F}$, and $\Gamma_3 = B_2 \times \mathcal{F}$ in models 1 and 2, respectively, where B_1 and B_2 are bounded subsets of R and R^2, respectively, and $\mathcal{F}$ is the parameter space for the distribution F of U_i.[7] In particular, we have

$$\mathcal{F} = \{F : E_F U_i = 0,\ \sigma_U^2 = E_F U_i^2 > 0,\ E_F |U_i/\sigma_U|^4 \leq M\} \tag{9.4}$$

for some $M < \infty$. (Note that Γ_3 does not depend on γ_1 in this example.) In the definition of $\gamma_{n,h}$, we take $r = 1$. In Appendix B we verify the assumptions of Corollaries 3 and 4 using Lemma 2 of AG1 to verify Assumption G2. (For brevity, we verify Assumptions E and EE only for model 1.)

In this example, $H = R_{+,\infty}$. Therefore, to establish Assumption BB2, we have to consider sequences $\{\gamma_{n,h} = (\gamma_{n,h,1}, \gamma_{n,h,3})' : n \geq 1\}$ when the true autoregressive parameter $\rho = \rho_n$ equals $1 - \gamma_{n,h,1}$ where (i) $h = \infty$ and (ii) $0 \leq h < \infty$. Case (i) is studied by Park (2002), Giraitis and Phillips (2006), and Phillips and Magdalinos (2007). Case (ii) is the "near integrated" case that has been studied by Bobkowski (1983), Cavanagh (1985), Chan and Wei (1987), Phillips (1987), Elliott (1999), Elliott and Stock (2001), and Müller and Elliott (2003). The latter three papers consider the situation of interest here in which the initial condition Y_0^* yields a stationary process. Specifically, what is relevant here is the triangular array case with row-wise strictly stationary observations $\{Y_i^* : i \leq n\}$ and ρ that depends on n. Note that case (ii) contains as a special case the unit root model $\rho = 1$. We do not consider an AR model here without an intercept, but such a model can be analyzed using the results of Andrews and Guggenberger (2006). Interestingly, the asymptotic distributions in this case are quite different than in the models with an intercept or intercept and time trend.

For model 1, we have

$$\begin{aligned}
&T_n^*(\theta_n) \to_d J_h^* \text{ under } \gamma_{n,h}, \text{ where} \\
&J_h^* \text{ is the } N(0,1) \text{ distribution for } h = \infty, \\
&J_h^* \text{ is the distribution of } \int_0^1 I_{D,h}^*(r) dW(r) / (\int_0^1 I_{D,h}^*(r)^2 dr)^{1/2} \text{ for } 0 \leq h < \infty, \\
&I_{D,h}^*(r) = I_h^*(r) - \int_0^1 I_h^*(s) ds,
\end{aligned}$$

[7] The parameter space B_1 is taken to be bounded, because otherwise there are sequences $\alpha_n \to \infty$, $\rho_n \to 1$ for which $\widetilde{\alpha}_n \nrightarrow 0$. For analogous reasons, B_2 is taken to be bounded.

$$I_h^*(r) = I_h(r) + \frac{1}{\sqrt{2h}} \exp(-hr)Z \text{ for } h > 0 \text{ and } I_h^*(r) = W(r) \text{ for } h = 0,$$

$$I_h(r) = \int_0^r \exp(-(r-s)h)dW(s), \tag{9.5}$$

$W(\cdot)$ is a standard Brownian motion, and Z is a standard normal random variable that is independent of $W(\cdot)$. As defined, $I_h(r)$ is an Ornstein-Uhlenbeck process.

For model 2, (9.5) holds except that for $0 \leq h < \infty$ J_h^* is the distribution of

$$\frac{\int_0^1 \left[I_{D,h}^*(r) - 12 \int_0^1 I_{D,h}^*(s) s ds \cdot (r - 1/2) \right] dW(r)}{\left(\int_0^1 \left[I_{D,h}^*(r) - 12 \int_0^1 I_{D,h}^*(s) s ds \cdot (r - 1/2) \right]^2 dr \right)^{1/2}} \tag{9.6}$$

(where 12 appears in the formula because $\left(\int_0^1 (r - 1/2)^2 dr \right)^{-1} = 12$).

Figure 3 provides .95 quantile graphs of J_h^*, $-J_h^*$, and $|J_h^*|$. All of these graphs are monotone in h. The graph for J_h^* is monotone increasing in h because its upper tail gets thinner as h gets smaller. In consequence, the upper one-sided and equal-tailed two-sided subsample CIs under-cover the true value asymptotically and the upper FCV CI has correct size asymptotically. The graph for $-J_h^*$ is decreasing in h because the lower tail of J_h^* gets thicker as h gets smaller. The graph for $|J_h^*|$ is decreasing in h because the lower tail of J_h^* gets thicker as h gets smaller at a faster rate than the upper tail of J_h^* gets thinner. Because the graphs of $-J_h^*$ and $|J_h^*|$ are decreasing in h, the lower and symmetric subsample CIs have correct asymptotic size, while the lower FCV CI under-covers the true value asymptotically. These results explain the seemingly puzzling result (quantified in Table II below) that the equal-tailed subsample CI has incorrect size while the symmetric subsample CI has correct size asymptotically.

Table II reports the asymptotic, finite-sample-adjusted asymptotic, and finite-sample sizes ($\times 100$) of nominal 95% CIs for model 1 for various upper and lower one-sided and symmetric and equal-tailed two-sided CIs, see the rows labeled Min. The parameter space for ρ is taken to be $[-0.9, 1.0]$.[8] Also reported are the finite-sample coverage probabilities of the CIs for several values of ρ.[9] The asymptotic and finite-sample-adjusted

[8] The parameter space for ρ is taken to be $[-0.9, 1.0]$ to minimize the effect of the choice of the lower bound on the finite-sample sizes of the upper FCV and hybrid CIs because in most practical applications in economics, the parameter interval $(-1.0, -0.9]$ is not of interest. If the parameter space is taken to be $[-0.999, 1.0]$, then the finite-sample sizes of the upper FCV and hybrid CIs are determined by their behavior when ρ is in $[-0.999, -0.9]$ and they equal 91.1 and 93.8, respectively, in model 1, rather than 93.6 and 94.8. No other CIs are affected by the choice of lower bound of the parameter space.

[9] Table II does not report asymptotic probabilities for specific values of ρ—it only gives the Min. The reason is that the asymptotic results depend on h and there is no unique transformation from h to ρ. For example, one could take $\rho = 1 - h/n$ or $\rho = \exp(-h/n)$ (and note that neither of these is satisfactory for all ρ in $(-1, 1]$). Hence, it would be misleading to compare asymptotic results for ρ based on an arbitrary choice of transformation with the finite sample results for ρ.

asymptotic results are calculated using the results of Section 8.[10] The finite-sample results are for the case of $n = 130$, $b = 12$, i.i.d. standard normal innovations U_i, and $q_n = 119$ subsamples of consecutive observations. Table B-I in Appendix B reports analogous results for model 2.

In models 1 and 2, the lower one-sided and the two-sided FCV CIs (based on the standard normal distribution, which yields equality of the symmetric and equal-tailed CIs) strongly and severely under-cover asymptotically, respectively, see column 7 of the rows labeled Min. For example, for model 1, the two-sided FCV CI has $AsyCS$ $(\times 100) = 68.9$. The results for upper one-sided and equal-tailed two-sided subsample CIs are similar, but somewhat worse, see column 2 of the rows labeled Min. For example, for model 1, the equal-tailed subsample CI has $AsyCS$ $(\times 100) = 60.1$. Hence, subsample CIs can have very poor asymptotic performance. On the other hand, symmetric subsample CIs have correct $AsyCS$ in both models, see column 2 of the row labeled Min in panel (c) of Table II.

The upper, lower, symmetric, and equal-tailed hybrid CIs all have correct $AsyCS$ in models 1 and 2, see column 10 of the rows labeled Min in Table II. This occurs because in every case either the critical value of the FCV CI or the subsample CI is suitable. Hence, the maximum of the two is a critical value that delivers correct asymptotic size.

The discussion above of the quantile graphs in Figure 3 leads to the following results. The two-sided FCV CI under-covers because its upper endpoint is farther away from 1 than it should be. Hence, it misses the true value of ρ too often to the left. On the other hand, the equal-tailed subsample CI under-covers ρ because its lower endpoint is closer to 1 than it should be. Hence, it misses the true ρ to the right too often.

The finite-sample adjusted asymptotic results for $\delta_n = b_n/n = 12/130$ show much less severe under-rejection for the subsample CIs than the unadjusted asymptotic results—for the cases where the subsample CIs under-reject—compare columns 2 and 3 for the rows labeled Min. For example, the equal-tailed subsample CI has adjusted asymptotic size of 86.1, rather than 60.1, in model 1. The finite-sample sizes of the subsample CIs for $n = 130$ and $b = 12$ are much closer to the adjusted asymptotic sizes than the unadjusted asymptotic sizes for those cases where the subsample CI under-covers, compare columns 2, 3, and 4 of the rows labeled Min. For the equal-tailed subsample CI, the finite-sample size is 86.7, compared to adjusted and unadjusted asymptotic sizes of 86.1 and 60.1, respectively. Hence, it is apparent that the asymptotic size of the upper and equal-tailed subsample CIs is approached slowly as $n \to \infty$ and is obtained only with large sample sizes. In consequence, increases in the sample size

[10] The results of Table II are based on 20,000 simulation repetitions. The search over h to determine the Min is done on the interval $[-.90, 1]$ with stepsize 0.01 on $[-.90, .90]$ and stepsize .001 on $[.90, 1.0]$. The asymptotic results are computed using a discrete approximation to the continuous stochastic process on $[0, 1]$ with 10,000 grid points. The size-correction values $cv(1-\alpha)$, $\kappa(\alpha)$, and $\kappa(\alpha, \delta)$ for model 1 of the AR(1) example are as follows: for upper tests, $\kappa(.05) = 1.69$ & $\kappa(.05, .10) = 0.84$; for lower tests, $cv(.95) = 2.89$; for symmetric tests, $cv(.95) = 2.89$; and for equal-tailed tests, $cv(.95) = 0.93$, $\kappa(.05) = 1.38$, & $\kappa(.05, .10) = 0.53$. The finite-sample results in model 1 are invariant to the true values of α and σ_U^2 and, hence, these are taken to be 0 and 1, respectively.

from $n = 130$ makes the upper and equal-tailed subsample CIs perform worse rather than better.

The upper and equal-tailed subsample CIs can be size-corrected, as can the lower and two-sided FCV CIs. The ASC-Sub CI performs much better than the SC-Sub CI because the latter is based on the unadjusted asymptotic distribution, which is highly inaccurate and over-corrects, see columns 5 and 6 in Table II. The upper and equal-tailed ASC-Sub CIs have finite-sample sizes of 95.3 and 94.9, see column 6 of the rows labeled Min. The lower and two-sided SC-FCV CIs both have finite-sample size of 95.2, see column 9 of the rows labeled Min. (No SC-Sub or ASC-Sub results are reported for the symmetric Sub test because its asymptotic size is 94.8%, which is within simulation error of 95%. Hence, the symmetric Sub test does not require size-correction.)

The hybrid CI needs no size-correction in any of the cases considered. Its finite-sample sizes for $n = 130$ and $b = 12$ for upper, lower, symmetric, and equal-tailed CIs are 94.8, 92.7, 92.7, and 95.6, respectively, in model 1, see column 11 of the rows labelled Min. These CIs, like the ASC-Sub CIs, are not similar asymptotically or in finite samples, see columns 6 and 11 of the rows of Table II for specific values of ρ.

10 Post-Conservative Model Selection Inference

In this example, we consider inference concerning a parameter in a linear regression model after a "conservative" model selection procedure has been applied to determine whether another regressor should enter the model. A "conservative" model selection procedure is one that chooses a correct model, but not the most parsimonious correct model, with probability that goes to one as the sample size n goes to infinity. Examples are model selection based on a test whose critical value is independent of the sample size and the Akaike information criterion (AIC).

The results for this example are summarized as follows. The nominal 5% subsample, FCV, and hybrid tests have asymptotic and adjusted-asymptotic sizes that are very large—between 90 and 96—for upper, symmetric, and equal-tailed tests. (This is for the parameter space $H_2 = [-.995, .995]$ for the (asymptotic) correlation h_2 between the LS estimators of the two regressors.) Finite-sample sizes of these tests for $n = 120$ and $b = 12$ are close to the asymptotic values. Hence, the asymptotic results provide good approximations.

Plug-in size-correction methods work very well in this example (with the exception of the APSC-Sub test). The PSC-Hyb and APSC-Hyb tests work particularly well. Their finite-sample size for normal errors is 4.8 for upper, lower, and symmetric tests (for H_2 as above).[11] The PSC-FCV test also works very well. Its finite-sample sizes are 5.1, 5.3, and 5.2 for upper, lower, and symmetric tests, respectively. Note that the definitions of the plug-in size-correction tests do not depend on the specification of H_2

[11] Strictly speaking, h_2 denotes the asymptotic correlation between the LS estimators and H_2 denotes its parameter space. For simplicity, when discussing the finite-sample results, we let h_2 denote the finite-sample correlation between the LS estimators and we let H_2 denote its parameter space.

(because we do not restrict the estimator $\widehat{\gamma}_{n,2}$ to lie in H_2).

The results above are for the parameter space H_2 bounded away from 1.0 by a very small amount: .005. Bounding H_2 away from 1.0 is necessary for the asymptotic results to hold. Furthermore, finite-sample results for $|h_2|$ extremely close to one indicate that this condition is not superfluous. When $|h_2| = .999$, the size-corrected tests have finite-sample sizes between 6.9 and 7.4—still pretty good. For $|h_2| = .9999$, however, their finite-sample sizes are between 71 and 83. So, the size-corrected tests do not have correct size when the parameter space for h_2 is the unrestricted interval $[-1.0, -1.0]$. For practical purposes, this is not too much of a problem because (i) h_2 can be consistently estimated, so one has a good idea of whether $|h_2|$ is close to 1.0 and (ii) $|h_2|$ can be very close to 1.0 before the size-corrected tests display any adverse behavior.

The asymptotic results of this section for FCV tests are closely related to those of Leeb (2006) and Leeb and Pötscher (2005) (and other papers referenced in these two papers). No papers in the literature, that we are aware of, consider subsample methods for post-model selection inference. For FCV tests, the main differences from Leeb (2006) are that here we consider (i) model selection among two models, (ii) errors that may be non-normal, (iii) i.i.d. regressors, (iv) t statistics, and (v) we prove the asymptotic results directly. In contrast, Leeb (2006) considers (i) multiple models, (ii) normal errors, (iii) fixed regressors, (iv) normalized estimators, and (v) he proves the asymptotic results by establishing finite-sample results and taking their limits. The results in Leeb and Pötscher (2005) are a two-model special case of those given in Leeb (2006). The PSC-FCV test/CI considered here is closely related to, but different from, the modified CI of Kabaila (1998).

The model we consider is

$$\begin{aligned} y_i &= x_{1i}^*\theta + x_{2i}^*\beta_2 + x_{3i}^{*\prime}\beta_3 + \sigma\varepsilon_i \text{ for } i = 1, ..., n, \text{ where} \\ x_i^* &= (x_{1i}^*, x_{2i}^*, x_{3i}^{*\prime})' \in R^k,\ \beta = (\theta, \beta_2, \beta_3')' \in R^k, \end{aligned} \tag{10.1}$$

$x_{1i}^*, x_{2i}^*, \theta, \beta_2, \sigma, \varepsilon_i \in R$, and $x_{3i}^*, \beta_3 \in R^{k-2}$. The observations $\{(y_i, x_i^*) : i = 1, ..., n\}$ are i.i.d. The scaled error ε_i has mean 0 and variance 1 conditional on x_i^*.

We are interested in testing $H_0 : \theta = \theta_0$ after carrying out a model selection procedure to determine whether x_{2i}^* should enter the model. The model selection procedure is based on a t test of $H_0^* : \beta_2 = 0$ that employs a critical value c that does not depend on n. Because the asymptotic distribution of the test statistic is invariant to the value of θ_0, the testing results immediately yield results for a CI for θ obtained by inverting the test—without any need to adjust the assumptions as described in Section 8 of AG1.

The inference problem described above covers the following (seemingly more general) inference problem. Consider the model

$$\begin{aligned} y_i &= z_i'\tau + \sigma\varepsilon_i \text{ for } i = 1, ..., n, \text{ where} \\ z_i &= (z_{1i}', z_{2i})' \in R^k,\ \tau = (\tau_1', \tau_2)' \in R^k, \end{aligned} \tag{10.2}$$

$z_{1i}, \tau_1 \in R^{k-1}$, and $z_{2i}, \tau_2 \in R$. We are interested in testing $\overline{H}_0 : a'\tau = \theta_0$ for a given vector $a \in R^k$ with $a \neq e_k$, where $e_k = (0, ..., 0, 1)'$, after using a (fixed critical value) t test to determine whether z_{2i} should enter the model. This testing problem can be transformed into the former one by writing

$$\theta = a'\tau, \beta_2 = \tau_2, \ \beta_3 = B'\tau, \tag{10.3}$$

for some matrix $B \in R^{k\times(k-2)}$ such that $D = [a : e_k : B] \in R^{k\times k}$ is nonsingular. As defined, $\beta = D'\tau$. Define $x_i^* = D^{-1}z_i$. Then, $x_i^{*\prime}\beta = z_i'\tau$ and $H_0 : \theta = \theta_0$ is equivalent to $\overline{H}_0 : a'\tau = \theta_0$.

We now return to the model in (10.1). We consider upper and lower one-sided and symmetric and equal-tailed two-sided nominal level α FCV tests of $H_0 : \theta = \theta_0$. Each test is based on a studentized test statistic $T_n(\theta_0)$, where $T_n(\theta_0)$ equals $T_n^*(\theta_0)$, $-T_n^*(\theta_0)$, $|T_n^*(\theta_0)|$, and $T_n^*(\theta_0)$, respectively, and $T_n^*(\theta_0)$ is defined below. The FCV tests use critical values $z_{1-\alpha}$, $z_{1-\alpha}$, $z_{1-\alpha/2}$, and $(-z_{1-\alpha/2}, z_{1-\alpha/2})$, respectively. The subsample critical values are defined as in AG1 using Assumption Sub1. In particular, the subsample statistics are defined by $\{T_{n,b,j}(\overline{\theta}) : j = 1, ..., q_n\}$, where $\overline{\theta}$ is the "model-selection" estimator of θ defined below and $T_{n,b,j}(\theta_0)$ is defined just as $T_n(\theta_0)$ is defined but using the jth subsample of size b in place of the full sample of size n.

To define the test statistic $T_n^*(\theta_0)$, we write the variables in matrix notation and define the first and second regressors after projecting out the remaining regressors using finite-sample projections:

$$\begin{aligned} Y &= (y_1, ..., y_n)', \\ X_j^* &= (x_{j1}^*, ..., x_{jn}^*)' \in R^n \text{ for } j = 1, 2, \\ X_3^* &= [x_{31}^* : ... : x_{3n}^*]' \in R^{n\times(k-2)}, \\ X_j &= M_{X_3^*}X_j^* \in R^n \text{ for } j = 1, 2, \text{ and} \\ X &= [X_1 : X_2] \in R^{n\times 2}, \end{aligned} \tag{10.4}$$

where $M_{X_3^*} = I_n - P_{X_3^*}$ and $P_{X_3^*} = X_3^*(X_3^{*\prime}X_3^*)^{-1}X_3^{*\prime}$. The n-vectors X_1 and X_2 correspond to the n-vectors X_1^* and X_2^*, respectively, with X_3^* projected out.

The restricted and unrestricted least squares (LS) estimators of θ and the unrestricted LS estimator of β_2 are

$$\begin{aligned} \widetilde{\theta} &= (X_1'X_1)^{-1}X_1'Y, \\ \widehat{\theta} &= (X_1'M_{X_2}X_1)^{-1}X_1'M_{X_2}Y, \text{ and} \\ \widehat{\beta}_2 &= (X_2'M_{X_1}X_2)^{-1}X_2'M_{X_1}Y. \end{aligned} \tag{10.5}$$

The model selection test rejects $H_0^* : \beta_2 = 0$ if

$$\begin{aligned} |T_{n,2}| &= \left| \frac{n^{1/2}\widehat{\beta}_2}{\widehat{\sigma}(n^{-1}X_2'M_{X_1}X_2)^{-1/2}} \right| > c, \text{ where} \\ \widehat{\sigma}^2 &= (n-k)^{-1}Y'M_{[X_1^*:X_2^*:X_3^*]}Y \end{aligned} \tag{10.6}$$

and $c > 0$ is a given critical value that does not depend on n. Typically, $c = z_{1-\alpha/2}$ for some $\alpha > 0$. The estimator $\widehat{\sigma}^2$ of σ^2 is the standard (unrestricted) unbiased LS estimator.

The test statistic, $T_n^*(\theta_0)$, for testing $H_0 : \theta = \theta_0$ is a t statistic based on the restricted LS estimator of θ when the null hypothesis $H_0^* : \beta_2 = 0$ is not rejected and the unrestricted LS estimator when it is rejected:

$$\begin{aligned} T_n^*(\theta_0) &= \widetilde{T}_{n,1}(\theta_0)1(|T_{n,2}| \leq c) + \widehat{T}_{n,1}(\theta_0)1(|T_{n,2}| > c), \text{ where} \\ \widetilde{T}_{n,1}(\theta_0) &= \frac{n^{1/2}(\widetilde{\theta} - \theta_0)}{\widehat{\sigma}(n^{-1}X_1'X_1)^{-1/2}} \text{ and} \\ \widehat{T}_{n,1}(\theta_0) &= \frac{n^{1/2}(\widehat{\theta} - \theta_0)}{\widehat{\sigma}(n^{-1}X_1'M_{X_2}X_1)^{-1/2}}. \end{aligned} \tag{10.7}$$

Note that both $\widetilde{T}_{n,1}(\theta_0)$ and $\widehat{T}_{n,1}(\theta_0)$ are defined using the unrestricted estimator $\widehat{\sigma}$ of σ. One could define $\widetilde{T}_{n,1}(\theta_0)$ using the restricted LS estimator of σ, but this is not desirable because it leads to an inconsistent estimator of σ under sequences of parameters $\{\beta_2 = \beta_{2n} : n \geq 1\}$ that satisfy $\beta_{2n} \to 0$ and $n^{1/2}\beta_{2n} \nrightarrow 0$ as $n \to \infty$. For subsample tests, one could define $\widetilde{T}_{n,1}(\theta_0)$ and $\widehat{T}_{n,1}(\theta_0)$ with $\widehat{\sigma}$ deleted because the scale of the subsample statistics offsets that of the original sample statistic. This does not work for hybrid tests because Assumption K fails if $\widehat{\sigma}$ is deleted.

The "model-selection" estimator, $\overline{\theta}$, of θ is

$$\overline{\theta} = \widetilde{\theta}1(|T_{n,2}| \leq c) + \widehat{\theta}1(|T_{n,2}| > c). \tag{10.8}$$

This estimator is used to recenter the subsample statistics. (One could also use the unrestricted estimator $\widehat{\theta}$ to recenter the subsample statistics.)

We now show how the testing problem above fits into the general framework of AG1 and verify the assumptions of AG1. First, we define regressors $x_{1i}^{\perp}$ and $x_{2i}^{\perp}$ that correspond to x_{1i}^* and x_{2i}^*, respectively, with x_{3i}^* projected out using the population projection. Let G denote the distribution of (ε_i, x_i^*). Let

$$Q^* = E_G x_i^* x_i^{*\prime} \in R^{k \times k} \text{ and write } Q^* = \begin{bmatrix} Q_{11}^* & Q_{12}^* & Q_{13}^* \\ Q_{21}^* & Q_{22}^* & Q_{23}^* \\ Q_{31}^* & Q_{32}^* & Q_{33}^* \end{bmatrix}, \tag{10.9}$$

where $Q_{11}^*, Q_{21}^*, Q_{22}^* \in R$, $Q_{31}^*, Q_{32}^* \in R^{(k-2)}$, and $Q_{33}^* \in R^{(k-2)\times(k-2)}$. Define

$$\begin{aligned} x_{ji}^{\perp} &= x_{ji}^* - x_{3i}^{*\prime}(Q_{33}^*)^{-1}Q_{3j}^* \text{ for } j = 1, 2, \\ Q &= E_G x_i^{\perp} x_i^{\perp\prime}, \text{ where } x_i^{\perp} = (x_{1i}^{\perp}, x_{2i}^{\perp})', \text{ and } Q^{-1} = \begin{bmatrix} Q^{11} & Q^{12} \\ Q^{12} & Q^{22} \end{bmatrix}. \end{aligned} \tag{10.10}$$

In the notation of AG1, the parameter vector $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ is given by

$$\gamma_1 = \frac{\beta_2}{\sigma(Q^{22})^{1/2}}, \ \gamma_2 = \frac{Q^{12}}{(Q^{11}Q^{22})^{1/2}}, \text{ and } \gamma_3 = (\theta, \beta_2, \beta_3, \sigma, G). \tag{10.11}$$

Note that $\gamma_2 = \rho = AsyCorr(\widehat{\theta}, \widehat{\beta}_2)$. The parameter spaces for γ_1, γ_2, and γ_3 are $\Gamma_1 = R$, $\Gamma_2 = [-1+\zeta, 1-\zeta]$ for some $\zeta > 0$, and

$$\begin{aligned}\Gamma_3(\gamma_1,\gamma_2) = \Big\{&(\theta,\beta_2,\beta_3,\sigma,G) : \theta,\beta_2 \in R,\ \beta_3 \in R^{k-2},\ \sigma > 0, \text{ and for} \\ &Q = E_G x_i^{\perp} x_i^{\perp\prime} \text{ and } Q^{-1} = \begin{bmatrix} Q^{11} & Q^{12} \\ Q^{12} & Q^{22} \end{bmatrix},\ \text{(i) } \frac{\beta_2}{\sigma (Q^{22})^{1/2}} = \gamma_1, \\ &\text{(ii) } \frac{Q^{12}}{(Q^{11}Q^{22})^{1/2}} = \gamma_2,\ \text{(iii) } \lambda_{\min}(Q) \geq \kappa,\ \text{(iv) } \lambda_{\min}(E_G x_{3i}^* x_{3i}^{*\prime}) \geq \kappa, \\ &\text{(v) } E_G||x_i^*||^{2+\delta} \leq M,\ \text{(vi) } E_G||\varepsilon_i x_i^*||^{2+\delta} \leq M, \\ &\text{(vii) } E_G(\varepsilon_i|x_i^*) = 0 \text{ a.s., and (viii) } E_G(\varepsilon_i^2|x_i^*) = 1 \text{ a.s.}\Big\} \qquad (10.12)\end{aligned}$$

for some $\kappa, \delta > 0$ and $M < \infty$. The parameter γ_2 is bounded away from one and minus one because otherwise the LS estimators of θ and β_2 could have a distribution that is arbitrarily close to being singular.

The rate of convergence parameter r of AG1 equals 1/2. The localization parameter h of AG1 satisfies $h = (h_1, h_2) \in H = H_1 \times H_2$, where $H_1 = R_\infty$ and $H_2 = [-1+\zeta, 1-\zeta]$.

Let $\Delta(a,b) = \Phi(a+b) - \Phi(a-b)$, where $\Phi(\cdot)$ is the standard normal distribution function. Note that $\Delta(a,b) = \Delta(-a,b)$.

Calculations in Appendix B establish that the asymptotic distribution J_h^* of $T_n^*(\theta_0)$ under a sequence of parameters $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) : n \geq 1\}$ as in AG1 (where $n^{1/2}\gamma_{n,1} \to h_1$, $\gamma_{n,2} \to h_2$, and $\gamma_{n,3} \in \Gamma_3(\gamma_{n,1}, \gamma_{n,2})$ for all n) is

$$\begin{aligned} J_h^*(x) = {} & \Phi(x + h_1 h_2 (1-h_2^2)^{-1/2}) \Delta(h_1, c) \\ & + \int_{-\infty}^{x} \left(1 - \Delta\left(\frac{h_1 + h_2 t}{(1-h_2^2)^{1/2}}, \frac{c}{(1-h_2^2)^{1/2}}\right)\right) \phi(t) dt \qquad (10.13)\end{aligned}$$

when $|h_1| < \infty$. When $|h_1| = \infty$, $J_h^*(x) = \Phi(x)$ (which equals the limit as $|h_1| \to \infty$ of $J_h^*(x)$ defined in (10.13)). For upper one-sided, lower one-sided, and symmetric two-sided tests, the asymptotic distribution J_h of $T_n(\theta_0)$ is given by J_h^*, $-J_h^*$, and $|J_h^*|$, respectively. (If $Y \sim J_h^*$, then by definition, $-Y \sim -J_h^*$ and $|Y| \sim |J_h^*|$.) This verifies Assumption B2 of AG1.

Assumptions A2, D, and E of AG1 hold automatically. Assumption C of AG1 holds by choice of $\{b_n : n \geq 1\}$. Assumption F2 of AG1 holds because $J_h^*(x)$ is continuous in x for all $h \in H$. Assumption G2 of AG1 is verified in Appendix B using the proof of Lemma 4 in Appendix A of AG1. Assumption K holds with J_∞^* being a $N(0,1)$ distribution. Assumption L holds because $c_h(1-\alpha)$ is continuous in $h \in H$ and has finite limits as $|h_1| \to \infty$ and/or $|h_2| \to 1-\zeta$. Assumptions MF, MS, and MH hold by the continuity of $c_h(1-\alpha)$ in $h \in H$ plus the shape of $c_h(1-\alpha)$ as a function of h_1 for each $|h_2| \leq 1-\zeta$, see Figure 4. Given the assumptions just verified, the main results of AG1 and the present paper apply to this example.

Figure 4 provides graphs of the quantiles, $c_h(1-\alpha)$, of $|J_h^*|$ as a function of $h_1 \geq 0$ for several values of $h_2 \geq 0$. (The quantile graphs are invariant to the signs of h_1 and

h_2.) The corresponding quantile graphs for J_h^* are remarkably similar to those for $|J_h^*|$ and, hence, are not given. In Figure 4, the graphs are hump shaped with the size of the hump increasing in $|h_2|$. Based on the shape of the graphs, one expects the subsample, FCV, and hybrid tests all to over-reject the null hypothesis asymptotically and in finite samples and by a substantial amount when $|h_2|$ is large.

Table III provides null rejection probability results that are analogous to those in Table I but for the present example.[12] The finite-sample results in Table III are for $n = 120$, $b = 12$, $q_n = 119$, and a model with standard normal errors, and $k = 3$ regressors, where $x_{1,i}^*$ and $x_{2,i}^*$ are independent standard normal random variables and $x_{3,i}^* = 1$. The asymptotic results for the Sub, FCV, and Hyb tests show that all of these tests perform very similarly. They are found to over-reject the null hypothesis very substantially for the upper, symmetric, and equal-tailed cases when the absolute value of the correlation, $|h_2|$, is large. When the parameter space H_2 for h_2 is $[-.995, .995]$, the asymptotic sizes of these nominal 5% tests range from 93 to 96 (see columns 2, 7, and 10). Even for $|h_2| = .8$, the maximum (over h_1) asymptotic rejection probabilities ($\times 100$) of these tests range from 36 to 44.

Table III shows that the adjusted asymptotic sizes of the nominal 5% Sub and Hyb tests (for $H_2 = [-.995, .995]$) are slightly lower than the unadjusted ones, but they are still in the range of 90 to 92 (see columns 3 and 11).

The finite-sample sizes of the nominal 5% Sub, FCV, and Hyb tests are very high and reflect the asymptotic results (see columns 4, 8, and 12). They range from 91 to 95 (for $H_2 = [-.995, .995]$). The asymptotic results provide quite good approximations for the upper and equal-tailed tests, but less accurate ones for symmetric tests (compare column 2 with column 4, 7 with 8, and 10 with 12). The adjusted asymptotic results are particularly good for symmetric tests, but less good for upper and equal-tailed tests (compare column 3 with column 4 and 11 with 12).

In sum, the Sub, FCV, and Hyb tests all have very poor finite-sample size properties. The asymptotic results work quite well in approximating the finite-sample results.

Next, we consider size-corrected tests. For PSC and APSC methods, we use the following consistent estimator of $\gamma_{n,2}$:

$$\widehat{\gamma}_{n,2} = \frac{-n^{-1}X_1'X_2}{(n^{-1}X_1'X_1 n^{-1}X_2'X_2)^{1/2}}. \tag{10.14}$$

The choice of this estimator is based on the equality $\gamma_{n,2} = Q_n^{12}/(Q_n^{11}Q_n^{22})^{1/2} = -Q_{n,12}/(Q_{n,11}Q_{n,22})^{1/2}$, where Q_n^{jm} and $Q_{n,jm}$ denote the (j,m) elements of Q_n^{-1} and

[12] The results in Table III are based on $20,000$ simulation repetitions. For the finite-sample results, the search over $|\beta_2|$ is done on the interval $[0, 10]$ with stepsizes 0.0025, 0.025, and .25, respectively on the intervals $[0.0, 0.8]$, $[0.8, 3]$, and $[3, 10]$ and also includes the value $|\beta_2| = 999,999$. For the asymptotic results the search over $|h_1|$ is done in the interval $[-10, 10]$ with stepsize 0.01. The Max is taken over $|h_2|$ values in $\{0.0, 0.2, 0.4, 0.6, 0.8, 0.9, 0.95, 0.99, 0.995\}$. For the plug-in size-correction values, the grid of $|\gamma_2|$ values in $[0, 1]$ has stepsizes .01, .001, .0001, and .00002, respectively, on the intervals $[0.0, 0.7]$, $[0.7, 0.99]$, $[0.99, 0.996]$, and $[0.996, 1.0]$.

Q_n, respectively, for $j, m = 1, 2$ (see the second equality in (16.11) below). Consistency of $\widehat{\gamma}_{n,2}$ (i.e., $\widehat{\gamma}_{n,2} - \gamma_{n,2} \rightarrow_p 0$ under $\{\gamma_n : n \geq 1\}$) follows from a Lemma in Appendix B. Thus, Assumption N holds.

The PSC and APSC tests do not depend on the specification of the parameter space for h_2.

Table III reports finite-sample rejection probabilities of the PSC and APSC tests. The PSC-Sub, PSC-FCV, PSC-Hyb, and APSC-Hyb tests (see columns 5, 9, 13, and 14) all perform very well. The finite-sample sizes of these tests (for $H_2 = [-0.995, 0.995]$) are all in the range of 4.8 to 6.2 for upper, symmetric, and equal-tailed tests. (If one omits the equal-tailed PSC-Sub test, the range is 4.8 to 5.3.) For all of these tests the maximum rejection rates (over h_1) do not vary too much with $|h_2|$, which is the objective of the "plug-in" approach. Hence, the "plug-in" approach works well in this example. The PSC-Hyb and APSC-Hyb tests perform the same and perform particularly well. Their sizes are 4.8 for upper, symmetric, and equal-tailed tests. The PSC-FCV test also performs exceptionally well with finite-sample sizes of 5.1, 5.3, and 5.2 for upper, symmetric, and equal-tailed tests, respectively.

The only size-corrected test that does not perform well is the APSC-Sub test, which has sizes 25, 22, and 24 for upper, symmetric, and equal-tailed tests.

For $H_2 = [-.999, .999]$, the finite-sample sizes of the PSC tests lie between 6.9 and 7.4 and those of the APSC-Hyb test lie between 7.1 and 7.4. For $H_2 = [-.9999, .9999]$, all of the SC tests have finite-sample sizes between 71 and 83. Hence, it is clear that bounding $|h_2|$ away from 1.0 is not only sufficient for the asymptotic SC results to hold, but it is necessary for the SC tests to have good finite-sample size. Nevertheless, $|h_2|$ can be very close to 1.0 (i.e., .995 or less) and the SC tests still perform very well in finite samples.

Appendix A

This Appendix provides (i) some assumptions that are not stated in the text for brevity, (ii) size-correction results for equal-tailed tests, and (iii) proofs of the general results of the paper.

11 Assumptions

In Section 5, we use the following assumptions.

Assumption OF. (i) $cv_{h_2}(1-\alpha)$ is uniformly continuous in h_2 on H_2, (ii) for each $h_2 \in H_2$, there exists some $h_1^* \in H_1$ such that $c_{(h_1^*,h_2)}(1-\alpha) = cv_{h_2}(1-\alpha)$, and (iii) for all $h = (h_1, h_2) \in H$ for which $c_h(1-\alpha) = cv_{h_2}(1-\alpha)$, $J_h(x)$ is continuous at $x = cv_{h_2}(1-\alpha)$.

Assumption OS. (i) $\kappa_{h_2}(\alpha)$ is uniformly continuous in h_2 on H_2, (ii) for each $h_2 \in H_2$, there exists some $g_1^*, h_1^* \in H_1$ such that $(g^*, h^*) = ((g_1^*, h_2), (h_1^*, h_2)) \in GH$ and $c_{(h_1^*,h_2)}(1-\alpha) - c_{(g_1^*,h_2)}(1-\alpha) = \kappa_{h_2}(1-\alpha)$, and (iii) for all $(g, h) \in GH$ for which $c_h(1-\alpha) - c_g(1-\alpha) = \kappa_{h_2}(1-\alpha)$, where $h = (h_1, h_2)$, $J_h(x)$ is continuous at $x = c_g(1-\alpha) + \kappa_{h_2}(1-\alpha)$.

Assumption OH. (i) $\kappa_{h_2}^*(\alpha)$ is uniformly continuous in h_2 on H_2, (ii) for each $h_2 \in H_2$, when $H_{h_2}^*$ is non-empty, we have: for some $h_1^* \in H_{h_2}^*$, $c_{(h_1^*,h_2)}(1-\alpha) - c_\infty(1-\alpha) = \kappa_{h_2}^*(1-\alpha)$, and (iii) for all $h = (h_1, h_2) \in H$ for which $c_h(1-\alpha) - c_\infty(1-\alpha) = \kappa_{h_2}^*(1-\alpha)$, $J_h(x)$ is continuous at $x = \max\{c_g(1-\alpha), c_\infty(1-\alpha) + \kappa_{h_2}^*(1-\alpha)\}$.

12 Equal-Tailed Size-Corrected Tests

This section introduces *equal-tailed* size-corrected FCV, subsample, and hybrid t tests. It also introduces finite-sample-adjusted asymptotics for equal-tailed tests. We suppose Assumption t1(i) holds, so that $T_n(\theta_0) = \tau_n(\widehat{\theta}_n - \theta_0)/\widehat{\sigma}_n$.

Equal-tailed (i) SC-FCV, (ii) SC-Sub, and (iii) SC-Hyb tests are defined by (7.1) with the critical values $c_{n,b}^*(1-\alpha/2)$ and $c_{n,b}^{**}(\alpha/2)$ replaced by (i) $c_{Fix}(1-\alpha/2) + \kappa_{ET,Fix}(\alpha)$ and $c_{Fix}(\alpha/2) - \kappa_{ET,Fix}(\alpha)$, (ii) $c_{n,b}(1-\alpha/2) + \kappa_{ET}(\alpha)$ and $c_{n,b}(\alpha/2) - \kappa_{ET}(\alpha)$, and (iii) $\max\{c_{n,b}(1-\alpha/2), c_\infty(1-\alpha/2)+\kappa_{ET}^*(\alpha)\}$ and $\min\{c_{n,b}(\alpha/2), c_\infty(\alpha/2)-\kappa_{ET}^*(\alpha)\}$, respectively.

By definition, the SC factors $\kappa_{ET,Fix}(\alpha)$ ($\in [0,\infty)$), $\kappa_{ET}(\alpha)$ ($\in [0,\infty)$), and $\kappa_{ET}^*(\alpha)$ ($\in \{-\infty\} \cup [0,\infty)$), respectively, are the smallest values that satisfy

$$\sup_{h \in H}[1 - J_h((c_{Fix}(1-\alpha/2) + \kappa_{ET,Fix}(\alpha))-) + J_h(c_{Fix}(\alpha/2) - \kappa_{ET,Fix}(\alpha))] \leq \alpha,$$

$$\sup_{(g,h) \in GH}[1 - J_h((c_g(1-\alpha/2) + \kappa_{ET}(\alpha))-) + J_h(c_g(\alpha/2) - \kappa_{ET}(\alpha))] \leq \alpha, \text{ and}$$

$$\sup_{(g,h)\in GH}[1 - J_h(\max\{c_g(1-\alpha/2), c_\infty(1-\alpha/2) + \kappa^*_{ET}(\alpha)\}-) +$$
$$J_h(\min\{c_g(\alpha/2), c_\infty(\alpha/2) - \kappa^*_{ET}(\alpha)\})] \leq \alpha. \quad (12.1)$$

(If no such smallest value exists, we take some value that is arbitrarily close to the infimum.)

For each test, the condition in (12.1) guarantees that the "overall" asymptotic size of the test is less than or equal to α. It does not guarantee that the maximum (asymptotic) 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 3. (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.)

Given SC factors that satisfy (12.1), the equal-tailed SC-FCV, SC-Sub, and SC-Hyb t tests have $AsySz(\theta_0) \leq \alpha$ under the assumptions in Corollary 2(c) of AG1, Corollary 2(d) of AG1, and Corollary 2(b) above, respectively. Under continuity conditions on $J_h(x)$ at suitable values of x and h, such that the inequalities in (12.1) hold as equalities, these tests have $AsySz(\theta_0) = \alpha$.

An alternative way of size-correcting equal-tailed tests is the following method. This method has the advantage that if it is possible to produce an *equal-tailed* size-corrected test, then the procedure does so. Its disadvantage is that it is somewhat more complicated to implement.

First, let $\kappa_{ET,Fix,1}(\alpha) \in [0,\infty)$, $\kappa_{ET,1}(\alpha) \in [0,\infty)$, and $\kappa^*_{ET,1}(\alpha) \in \{-\infty\} \cup [0,\infty)$ denote the smallest values such that

$$\sup_{h\in H}[1 - J_h((c_{Fix}(1-\alpha/2) + \kappa_{ET,Fix,1}(\alpha))-)] \leq \alpha/2,$$
$$\sup_{(g,h)\in GH}(1 - J_h((c_g(1-\alpha/2) + \kappa_{ET,1}(\alpha))-)) \leq \alpha/2, \text{ and}$$
$$\sup_{(g,h)\in GH}\left(1 - J_h\left(\max\{c_g(1-\alpha/2), c_\infty(1-\alpha/2) + \kappa^*_{ET,1}(\alpha)\}-\right)\right) \leq \alpha/2. \quad (12.2)$$

Next, let $\kappa_{ET,Fix,2}(\alpha) \in R$, $\kappa_{ET,2}(\alpha) \in R$, and $\kappa^*_{ET,2}(\alpha) \in \{-\infty\} \cup R$ denote the smallest values such that

$$\sup_{h\in H}[1 - J_h((c_{Fix}(1-\alpha/2) + \kappa_{ET,Fix,1}(\alpha))-) + J_h(c_{Fix}(\alpha/2) - \kappa_{ET,Fix,2}(\alpha))] \leq \alpha,$$
$$\sup_{(g,h)\in GH}[1 - J_h((c_g(1-\alpha/2) + \kappa_{ET,1}(\alpha))-) + J_h(c_g(\alpha/2) - \kappa_{ET,2}(\alpha))] \leq \alpha, \text{ and}$$
$$\sup_{(g,h)\in GH}[1 - J_h(\max\{c_g(1-\alpha/2), c_\infty(1-\alpha/2) + \kappa^*_{ET,1}(\alpha)\}-) +$$
$$J_h(\min\{c_g(\alpha/2), c_\infty(\alpha/2) - \kappa^*_{ET,2}(\alpha)\})] \leq \alpha. \quad (12.3)$$

The "alternative" SC equal-tailed FCV test rejects H_0 if $T_n(\theta_0) > c_{Fix}(1-\alpha/2) + \kappa_{ET,Fix,1}(\alpha)$ or $T_n(\theta_0) < c_{Fix}(\alpha/2) - \kappa_{ET,Fix,2}(\alpha)$. The "alternative" SC equal-tailed

Sub and Hyb tests are defined analogously. We use "alternative" SC equal-tailed tests in the Parameter of Interest Near a Boundary Example given in Andrews and Guggenberger (2005b). For all of the other examples, we use the SC equal-tailed tests defined in (12.1).

If a parameter γ_2 appears in γ and γ_2 is consistently estimable, then PSC tests are more powerful asymptotically than SC tests (because they lead to smaller critical values under some distributions but still have correct asymptotic size). Equal-tailed (i) PSC-FCV, (ii) PSC-Sub, and (iii) PSC-Hyb tests are defined as the SC versions are defined above, but with $\kappa_{ET,Fix}(\alpha)$, $\kappa_{ET}(\alpha)$, and $\kappa^*_{ET}(\alpha)$ replaced by $\kappa_{ET,Fix,\widehat{\gamma}_{n,2}}(\alpha)$, $\kappa_{ET,\widehat{\gamma}_{n,2}}(\alpha)$, and $\kappa^*_{ET,\widehat{\gamma}_{n,2}}(\alpha)$, respectively. Here, the PSC factors $\kappa_{ET,Fix,h_2}(\alpha)$ ($\in [0,\infty)$), $\kappa_{ET,h_2}(\alpha)$ ($\in [0,\infty)$), and $\kappa^*_{ET,h_2}(\alpha)$ ($\in \{-\infty\} \cup [0,\infty)$) are defined to be the smallest values that satisfy:

$$\begin{aligned}
&\sup_{h_1 \in H_1} [1 - J_{(h_1,h_2)}((c_{Fix}(1-\alpha/2) + \kappa_{ET,Fix,h_2}(\alpha))-) + \\
&J_{(h_1,h_2)}(c_{Fix}(\alpha/2) - \kappa_{ET,Fix,h_2}(\alpha))] \leq \alpha, \\
&\sup_{g_1,h_1 \in H_1 : ((g_1,h_2),(h_1,h_2)) \in GH} [1 - J_{(h_1,h_2)}((c_{(g_1,h_2)}(1-\alpha/2) + \kappa_{ET,h_2}(\alpha))-) + \\
&J_{(h_1,h_2)}(c_{(g_1,h_2)}(\alpha/2) - \kappa_{ET,h_2}(\alpha))] \leq \alpha, \text{ and} \\
&\sup_{g_1,h_1 \in H_1 : ((g_1,h_2),(h_1,h_2)) \in GH} [1 - J_{(h_1,h_2)}(\max\{c_{(g_1,h_2)}(1-\alpha/2), c_\infty(1-\alpha/2) + \\
&\kappa^*_{ET,h_2}(\alpha)\}-) + J_{(h_1,h_2)}(\min\{c_{(g_1,h_2)}(\alpha/2), c_\infty(\alpha/2) - \kappa^*_{ET,h_2}(\alpha)\})] \leq \alpha. \quad (12.4)
\end{aligned}$$

The PSC-FCV, PSC-Sub, and PSC-Hyb tests all have $AsySz(\theta_0) \leq \alpha$ under Assumption N and the assumptions of Corollary 2(c) of AG1, Corollary 2(d) of AG1, and Corollary 2(b) above, respectively. (The proof is analogous to the proof of Theorem 3(a)(i), 3(b)(i), and 3(c)(i) combined with the proof of Theorem 1.) These tests have $AsySz(\theta_0) = \alpha$ provided the inequalities in (12.4) hold as equalities.

The finite-sample adjustments introduced in Section 6 do not cover equal-tailed tests. For equal-tailed subsample tests, we define the following finite-sample adjustment to $AsySz(\theta_0)$:

$$AsySz_n(\theta_0) = \sup_{h \in H} [1 - J_h(c_{(\delta_n^r h_1,h_2)}(1-\alpha/2)-) + J_h(c_{(\delta_n^r h_1,h_2)}(\alpha/2))]. \quad (12.5)$$

With the function that appears in Assumption P(i) altered to $(g,h) \to J_h(c_g(1-\alpha/2)-) - J_h(c_g(\alpha/2))$ and with $Max^{r-}_{ET,Sub}(\alpha) = Max^{\ell-}_{ET,Sub}(\alpha)$ in place of Assumption P(ii), the result of Theorem 4, viz., $AsySz_n(\theta_0) \to AsySz(\theta_0)$, holds for equal-tailed subsample tests.

Based on (12.5), we introduce finite-sample adjustments that can improve the asymptotic approximations upon which the equal-tailed SC and PSC subsample and hybrid tests rely. Equal-tailed ASC and APSC subsample and hybrid tests are defined just as SC subsample and hybrid tests are defined, but using $\kappa_{ET}(\delta_n,\alpha)$, $\kappa^*_{ET}(\delta_n,\alpha)$, $\kappa_{ET,\widehat{\gamma}_{n,2}}(\delta_n,\alpha)$ and $\kappa^*_{ET,\widehat{\gamma}_{n,2}}(\delta_n,\alpha)$ in place of $\kappa_{ET}(\alpha)$ and $\kappa^*_{ET}(\alpha)$. The ASC factors

$\kappa_{ET}(\delta,\alpha)$ $(\in [0,\infty))$ and $\kappa^*_{ET}(\delta,\alpha)$ $(\in \{-\infty\}\cup[0,\infty))$ are defined to be the smallest values that satisfy:

$$\begin{aligned}
&\sup_{(h_1,h_2)\in H}[1 - J_{(h_1,h_2)}((c_{(\delta^r h_1,h_2)}(1-\alpha/2) + \kappa_{ET}(\delta,\alpha))-) + \\
&J_{(h_1,h_2)}(c_{(\delta^r h_1,h_2)}(\alpha/2) - \kappa_{ET}(\delta,\alpha))] \le \alpha \text{ and} \\
&\sup_{(h_1,h_2)\in H}[1 - J_{(h_1,h_2)}(\max\{c_{(\delta^r h_1,h_2)}(1-\alpha/2), c_\infty(1-\alpha/2) + \kappa^*_{ET}(\delta,\alpha)\}-) + \\
&J_{(h_1,h_2)}(\min\{c_{(\delta^r h_1,h_2)}(\alpha/2), c_\infty(\alpha/2) - \kappa^*_{ET}(\delta,\alpha)\})] \le \alpha. \qquad (12.6)
\end{aligned}$$

The APSC factors $\kappa_{ET,h_2}(\delta,\alpha)$ $(\in [0,\infty))$ and $\kappa^*_{ET,h_2}(\delta,\alpha)$ $(\in \{-\infty\}\cup[0,\infty))$ are defined to be the smallest values that satisfy:

$$\begin{aligned}
&\sup_{h_1\in H_1}[1 - J_{(h_1,h_2)}((c_{(\delta^r h_1,h_2)}(1-\alpha/2) + \kappa_{ET,h_2}(\delta,\alpha))-) + \\
&J_{(h_1,h_2)}(c_{(\delta^r h_1,h_2)}(\alpha/2) - \kappa_{ET,h_2}(\delta,\alpha))] \le \alpha \text{ and} \\
&\sup_{h_1\in H_1}[1 - J_{(h_1,h_2)}(\max\{c_{(\delta^r h_1,h_2)}(1-\alpha/2), c_\infty(1-\alpha/2) + \kappa^*_{ET,h_2}(\delta,\alpha)\}-) + \\
&J_{(h_1,h_2)}(\min\{c_{(\delta^r h_1,h_2)}(\alpha/2), c_\infty(\alpha/2) - \kappa^*_{ET,h_2}(\delta,\alpha)\})] \le \alpha. \qquad (12.7)
\end{aligned}$$

The ASC and APSC tests have $AsySz(\theta_0) = \alpha$ under conditions that are similar to those given in Section 6. For brevity, we do not give details.

13 Proofs

For notational simplicity, throughout this section, we let c_g, c_h, c_∞, $c_{n,b}$, and cv abbreviate $c_g(1-\alpha)$, $c_h(1-\alpha)$, $c_\infty(1-\alpha)$, $c_{n,b}(1-\alpha)$, and $cv(1-\alpha)$, respectively.

Proof of Lemma 1. If $c_h \ge c_\infty$ 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 "$c_h \ge c_\infty$ for all $h \in H$" does not hold. Then, for some $g \in H$, $c_g < c_\infty$. Given g, define $h_1 = (h_{1,1}, ..., h_{1,p})' \in H_1$ by $h_{1,m} = +\infty$ if $g_{1,m} > 0$, $h_{1,m} = -\infty$ if $g_{1,m} < 0$, $h_{1,m} = +\infty$ or $-\infty$ (chosen so that $(g,h) \in GH$) if $g_{1,m} = 0$ for $m = 1, ..., p$, and define $h_2 = g_2$. Let $h = (h_1, h_2)$. By construction, $(g,h) \in GH$. By Assumption K, $c_h = c_\infty$. Hence, we have

$$Max_{Sub}(\alpha) \ge 1 - J_h(c_g) > \alpha, \qquad (13.1)$$

where the second inequality holds because $c_g < c_\infty = c_h$ and c_h is the infimum of values x such that $J_h(x) \ge 1-\alpha$ or, equivalently, $1 - J_h(x) \le \alpha$. Equation (13.1) and Theorem 2(b) of AG1 imply that $AsySz(\theta_0) > \alpha$ for the subsample test. The hybrid test reduces the asymptotic over-rejection of the subsample test at (g,h) from being at least $1 - J_h(c_g) > \alpha$ to being at most $1 - J_h(c_\infty) = 1 - J_h(c_h) \le \alpha$ (with equality if $J_h(\cdot)$ is continuous at c_h). $\square$

Proof of Lemma 2. Suppose Assumption Quant0(i) holds. Then, we have

$$
\begin{aligned}
Max^{-}_{Hyb}(\alpha) &= \sup_{(g,h)\in GH}[1 - J_h(\max\{c_g, c_\infty\}-)] = \sup_{h\in H}[1 - J_h(c_\infty-)] \\
&= \sup_{h\in H}[1 - J_h(c_\infty)] \leq \sup_{h\in H}[1 - J_h(c_h)] \leq \alpha, \qquad (13.2)
\end{aligned}
$$

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.

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 K and

$$
1 - J_h(\max\{c_g, c_\infty\}-) \leq 1 - J_\infty(c_\infty-) = 1 - J_\infty(c_\infty) \leq \alpha, \qquad (13.3)
$$

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

$$
1-J_h(\max\{c_g, c_\infty\}-) \leq 1-J_h(c_{h^0}-) \leq \sup_{h\in H}[1-J_h(c_h-)] = \sup_{h\in H}[1-J_h(c_h)] \leq \alpha, \quad (13.4)
$$

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

Proof of Theorem 1. First we note that Assumption L implies that $cv(1-\alpha)$, $\kappa(\alpha)$, and $\kappa^*(\alpha)$ are well-defined (i.e., finite).

Below we show that $cv(1 - \alpha)$, $\kappa(\alpha)$, and $\kappa^*(\alpha)$ satisfy

$$
\begin{aligned}
\sup_{h\in H}[1 - J_h(cv(1 - \alpha)-)] &\leq \alpha, \\
\sup_{(g,h)\in GH}(1 - J_h((c_g(1 - \alpha) + \kappa(\alpha))-)) &\leq \alpha \text{ and} \\
\sup_{(g,h)\in GH}(1 - J_h(\max\{c_g(1 - \alpha), c_\infty(1 - \alpha) + \kappa^*(\alpha)\}-)) &\leq \alpha, \qquad (13.5)
\end{aligned}
$$

respectively. Given (13.5), Theorem 2(a) of AG1 applied with $c_{Fix}(1-\alpha) = cv$ implies that the SC-FCV test satisfies $AsySz(\theta_0) \leq \sup_{h\in H}[1-J_h(cv-)] \leq \alpha$, where the second inequality holds by (13.5). Theorem 2(b) of AG1 with $c_{n,b}+\kappa(\alpha)$ in place of $c_{n,b}$ implies that the SC-Sub test satisfies $AsySz(\theta_0) \leq \sup_{(g,h)\in H}[1-J_h((c_g+\kappa(\alpha))-)] \leq \alpha$, where the second inequality holds by (13.5). Corollary 1(b) with $\max\{c_{n,b}, c_\infty+\kappa^*(\alpha)\}$ in place of $\max\{c_{n,b}, c_\infty\}$ implies that the SC-Hyb test satisfies $AsySz(\theta_0) \leq \sup_{(g,h)\in H}[1 - J_h(\max\{c_g, c_\infty + \kappa^*(\alpha)\}-)] \leq \alpha$, where the second inequality holds by (13.5). Hence, $AsySz(\theta_0) \leq \alpha$ for SC-FCV, SC-Sub, and SC-Hyb tests. Below we show that the reverse inequality also holds.

We now show that the first inequality in (13.5) holds. For $h \in H$, if $c_h < \sup_{h^\dagger \in H} c_{h^\dagger}$, then

$$J_h\left(\sup_{h^\dagger \in H} c_{h^\dagger}-\right) \geq J_h(c_h) \geq 1-\alpha, \tag{13.6}$$

where the first inequality holds because J_h is nondecreasing and the second inequality holds by the definition of c_h. For $h \in H$, if $c_h = \sup_{h^\dagger \in H} c_{h^\dagger}$, then

$$J_h\left(\sup_{h^\dagger \in H} c_{h^\dagger}-\right) = J_h(c_h-) = 1-\alpha, \tag{13.7}$$

where the last equality holds by Assumption MF(ii). For cv defined in (3.2), (13.6) and (13.7) combine to give

$$\sup_{h\in H}[1-J_h(cv-)] = \sup_{h\in H}[1-J_h(\sup_{h^\dagger \in H} c_{h^\dagger}-)] \leq \alpha. \tag{13.8}$$

Hence, cv satisfies (13.5).

Next, we prove that the second inequality in (13.5) holds. For $(g,h) \in GH$, if $c_h < c_g + \sup_{(g^\dagger,h^\dagger)\in GH}[c_{h^\dagger} - c_{g^\dagger}]$, then we have

$$J_h\left((c_g+\kappa(\alpha))-\right) = J_h\left((c_g + \sup_{(g^\dagger,h^\dagger)\in GH}[c_{h^\dagger}-c_{g^\dagger}])-\right) \geq J_h(c_h) \geq 1-\alpha, \tag{13.9}$$

where the first inequality holds by the condition on (g,h) and the fact that J_h is nondecreasing.

For $(g,h) \in GH$, if $c_h = c_g + \sup_{(g^\dagger,h^\dagger)\in GH}[c_{h^\dagger} - c_{g^\dagger}]$, then we have

$$J_h\left((c_g+\kappa(\alpha))-\right) = J_h\left((c_g + \sup_{(g^\dagger,h^\dagger)\in GH}[c_{h^\dagger}-c_{g^\dagger}])-\right) = J_h(c_h-) = 1-\alpha, \tag{13.10}$$

where the second equality holds by the condition on (g,h) and the last equality holds by Assumption MS(ii). Combining (13.9) and (13.10) gives $\sup_{(g,h)\in GH}[1-J_h((c_g+\kappa(\alpha))-)] \leq \alpha$, as desired.

The third inequality in (13.5) holds by the following argument. Because $c_\infty + \kappa^*(\alpha) = \sup_{h^*\in H^*} c_{h^*}$, we need to show that $\sup_{(g,h)\in GH}[1-J_h(\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\}-)] \leq \alpha$. For all $(g,h) \in GH$, we have $\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\} \geq c_h$ because $\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\} < c_h$ implies that $c_g < c_h$, which implies that $h \in H^*$, which implies that $\sup_{h^*\in H^*} c_{h^*} \geq c_h$, which is a contradiction. Now, for any $(g,h) \in GH$ with $\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\} > c_h$, we have $1-J_h(\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\}-) \leq 1-J_h(c_h) \leq \alpha$, as desired. For any $(g,h) \in GH$ with $\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\} = c_h$, Assumption MH(ii) implies that $J_h(x)$ is continuous at $x = c_h$. Hence, $1-J_h(\max\{c_g, c_\infty + \kappa^*(\alpha)\}-) = 1-J_h(c_h-) = 1-J_h(c_h) = \alpha$, which completes the proof of the third inequality of

(13.5). This concludes the proof that $AsySz(\theta_0) \leq \alpha$ for the SC-FCV, SC-Sub, and SC-Hyb tests.

We now prove that these tests satisfy $AsySz(\theta_0) \geq \alpha$. By Theorem 2(a) of AG1 applied with $c_{Fix}(1-\alpha) = cv$, the SC-FCV test satisfies $AsySz(\theta_0) \geq \sup_{h \in H}[1 - J_h(cv)]$. Using (3.2) and Assumption MF(i), $cv = \sup_{h \in H} c_h = c_{h^*}$ for some $h^* \in H$. Hence,

$$\sup_{h \in H}[1 - J_h(cv)] = \sup_{h \in H}[1 - J_h(c_{h^*})] \geq 1 - J_{h^*}(c_{h^*}) = \alpha, \tag{13.11}$$

where the last equality holds by Assumption MF(ii). In consequence, for the SC-FCV test, $AsySz(\theta_0) \geq \alpha$.

Next, by Theorem 2(b) of AG1 with $c_{n,b} + \kappa(\alpha)$ in place of $c_{n,b}$, the SC-Sub test satisfies $AsySz(\theta_0) \geq \sup_{(g,h) \in GH}[1 - J_h(c_g + \kappa(\alpha))]$. Using (3.2) and Assumption MS(i), $\kappa(\alpha) = c_{h^*} - c_{g^*}$ for some $(g^*, h^*) \in GH$ as in Assumption MS(i). Hence,

$$\sup_{(g,h) \in GH}[1 - J_h(c_g + \kappa(\alpha))] = \sup_{(g,h) \in GH}[1 - J_h(c_g + c_{h^*} - c_{g^*})] \geq 1 - J_{h^*}(c_{h^*}) = \alpha, \tag{13.12}$$

where the last equality holds by Assumption MS(ii). In consequence, for the SC-Sub test, $AsySz(\theta_0) \geq \alpha$.

Lastly, Corollary 1(b) with $\max\{c_{n,b}, c_\infty + \kappa^*(\alpha)\}$ in place of $\max\{c_{n,b}, c_\infty\}$ implies that the SC-Hyb test satisfies

$$AsySz(\theta_0) \geq \sup_{(g,h) \in GH}[1 - J_h(\max\{c_g, c_\infty + \kappa^*(\alpha)\})]. \tag{13.13}$$

If H^* is not empty, then using (3.2) and Assumption MH(i), $\kappa^*(\alpha) = c_{h^*} - c_\infty$ for some $h^* \in H^*$ as in Assumption MH(i). By the definition of H^*, there exists g^* such that $(g^*, h^*) \in GH$ and $c_{g^*} < c_{h^*}$. In consequence, the right-hand side of (13.13) equals

$$\sup_{(g,h) \in GH}[1 - J_h(\max\{c_g, c_{h^*}\})] \geq 1 - J_{h^*}(\max\{c_{g^*}, c_{h^*}\}) = 1 - J_{h^*}(c_{h^*}) = \alpha, \tag{13.14}$$

where the first equality uses $c_{g^*} < c_{h^*}$ and last equality holds by Assumption MH(ii) because $(g^*, h^*) \in GH$ satisfies $c_{h^*} = \sup_{h \in H^*} c_h = \max\{c_{g^*}, \sup_{h \in H^*} c_h\}$. Combining (13.13) and (13.14) gives $AsySz(\theta_0) \geq \alpha$.

If H^* is empty, then $\kappa^*(\alpha) = -\infty$, $(h^0, h^0) \in GH$, where $h^0 = (0, h_2)'$ for arbitrary $h_2 \in H_2$, and we have

$$\begin{aligned} &\sup_{(g,h) \in GH}[1 - J_h(\max\{c_g, c_\infty + \kappa^*(\alpha)\})] \\ &= \sup_{(g,h) \in GH}[1 - J_h(c_g)] \geq 1 - J_{h^0}(c_{h^0}) = \alpha, \end{aligned} \tag{13.15}$$

where the last equality holds by Assumption MH(ii) because $c_{h^0} = \max\{c_{h^0}, c_\infty + \kappa^*(\alpha)\}$. Combining (13.13)–(13.15) gives $AsySz(\theta_0) \geq \alpha$ for the SC-Hyb test. $\square$

Proof of Theorem 2. Part (a)(i) follows from the definition of cv in (3.2) and Assumption Quant1. Part (a)(ii) holds by definition of $\kappa(\alpha)$ in (3.2) and the fact that $c_h - c_g \leq 0$ for all $(g,h) \in GH$ by Assumption Quant1 (with equality for some $(g,h) \in GH$). Part (a)(iii) holds by the definition of $\kappa^*(\alpha)$ in (3.2) for the case where H^* is empty, because H^* is empty by Assumption Quant1. Part (a)(iv) follows from parts (a)(ii) and (a)(iii). Part (a)(v) follows from part (a)(ii) and the definition of cv in (3.2).

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

We now prove part (b)(ii). By Assumptions Quant2 and K, H^* is not empty and $\sup_{h \in H^*} c_h = c_\infty$. In consequence, $\kappa^*(\alpha) = 0$ by definition of $\kappa^*(\alpha)$ in (3.2). Part (b)(iii) follows from parts (b)(i) and (b)(ii) and $c_g \leq c_\infty$ by Assumptions Quant2 and K. We now prove part (b)(iv). By part (b)(i), it suffices to show that $c_g + \kappa(\alpha) \geq c_\infty$ for all $g \in H$. By the definition of $\kappa(\alpha)$ in (3.2) and Assumptions Quant2 and K, $\kappa(\alpha) = c_\infty - \inf_{h_2 \in H_2} c_{(0,h_2)}$. Hence, $c_g + \kappa(\alpha) = c_g + c_\infty - \inf_{h_2 \in H_2} c_{(0,h_2)} \geq c_\infty$, where the inequality uses Assumption Quant2. This establishes part (b)(iv).

Part (c)(i) holds by Assumption Quant3(ii). Part (c)(ii) holds by definition of $\kappa(\alpha)$ in (3.2) and Assumptions Quant3(ii) and Quant3(iii). Part (c)(iii) holds by definition of $\kappa^*(\alpha)$ in (3.2) and Assumption Quant3(ii). Part (c)(iv) holds because $\max\{c_g, c_\infty + \kappa^*(\alpha)\} = \max\{c_g, c_{h^*}\} = c_{h^*} = cv$ using parts (c)(i) and (c)(iii). Parts (c)(v) and (c)(vi) hold because $cv = c_{h^*}$ by part (c)(i) and $c_g + \kappa(\alpha) = c_{h^*} + c_g - c_0$ by part (c)(ii). $\square$

Proof of Theorem 3. The results of parts (a)(i), (b)(i), and (c)(i) hold by an extension of Slutsky's Theorem (to allow $\gamma_{n,2}$ to depend on n) using Assumption N and the uniform continuity of the functions in Assumptions OF(i), OS(i), and OH(i), respectively. The proof of parts (a)(ii), (b)(ii), and (c)(ii) is split into two steps. In the first step, we consider the PSC tests with $\widehat{\gamma}_{n,2}$ replaced by the true value $\gamma_{n,2}$. In this case, using parts (ii) and (iii) of Assumptions OF, OS, and OH, the results of parts (a)(ii), (b)(ii), and (c)(ii) hold by a very similar argument to that given in the proof of Theorem 1. In the second step, the results of parts (a)(i), (b)(i), and (c)(i) are combined with the results of the first step to obtain the desired results. This step holds because the results of parts (a)(i), (b)(i), and (c)(i) lead to the same limit distributions for the statistics in question whether they are based on $\widehat{\gamma}_{n,2}$ or the true value $\gamma_{n,2}$ by the argument used in the proof of Theorem 2(b) of AG1. $\square$

Proof of Theorem 4. Under Assumptions A2, B2, C-E, F2, and G2, Theorem 2 of AG1 combined with Assumption P(ii) shows that

$$AsySz(\theta_0) = \sup_{(g,h)\in GH} (1 - J_h(c_g(1-\alpha))). \tag{13.16}$$

First, we show that

$$\liminf_{n\to\infty} AsySz_n(\theta_0) \geq AsySz(\theta_0). \tag{13.17}$$

Given $(g,h) = ((g_1,h_2),(h_1,h_2)) \in GH$, we construct a sequence $\{h_n = (h_{n,1},h_{n,2}) \in H : n \geq 1\}$ such that $(g_n,h_n) \to (g,h)$ as $n \to \infty$, where $g_n = (g_{n,1},g_{n,2}) = (\delta_n^r h_{n,1}, h_{n,2})$. Define $h_{n,2} = h_2$ for all $n \geq 1$. We write $h_1 = (h_{1,1},...,h_{1,p})'$ and $h_{n,1} = (h_{n,1,1},...,h_{n,1,p})'$. For $m = 1,...,p$, define

$$\begin{array}{ll} h_{n,1,m} = h_{1,m} & \text{if } g_{1,m} = 0 \text{ \& } |h_{1,m}| < \infty \\ h_{n,1,m} = (n/b_n)^{r/2} & \text{if } g_{1,m} = 0 \text{ \& } h_{1,m} = \infty \\ h_{n,1,m} = -(n/b_n)^{r/2} & \text{if } g_{1,m} = 0 \text{ \& } h_{1,m} = -\infty \\ h_{n,1,m} = (n/b_n)^r g_{1,m} & \text{if } g_{1,m} \in (0,\infty) \text{ \& } h_{1,m} = \infty \\ h_{n,1,m} = (n/b_n)^r g_{1,m} & \text{if } g_{1,m} \in (-\infty,0) \text{ \& } h_{1,m} = -\infty \\ h_{n,1,m} = (n/b_n)^{2r} & \text{if } g_{1,m} = \infty \text{ \& } h_{1,m} = \infty \\ h_{n,1,m} = -(n/b_n)^{2r} & \text{if } g_{1,m} = -\infty \text{ \& } h_{1,m} = -\infty. \end{array} \tag{13.18}$$

As defined, $(g_{n,1},h_{n,1}) = (\delta_n^r h_{n,1}, h_{n,1}) \to (g_1,h_1)$ and $(g_n,h_n) \to (g,h)$.

We now have

$$\begin{aligned} \liminf_{n\to\infty} AsySz_n(\theta_0) &= \liminf_{n\to\infty} \sup_{h=(h_1,h_2)\in H} (1 - J_h(c_{(\delta_n^r h_1,h_2)}(1-\alpha))) \\ &\geq \liminf_{n\to\infty} (1 - J_{h_n}(c_{(\delta_n^r h_{n,1},h_{n,2})}(1-\alpha))) \\ &= \liminf_{n\to\infty} (1 - J_{h_n}(c_{g_n}(1-\alpha))) \\ &= 1 - J_h(c_g(1-\alpha)), \end{aligned} \tag{13.19}$$

where the second equality holds by definition of g_n and the last equality holds by Assumption P because $(g_n,h_n) \to (g,h)$. Given (13.16), this establishes (13.17) because (13.19) holds for all $(g,h) \in GH$.

Next, we show that $\limsup_{n\to\infty} AsySz_n(\theta_0) \leq AsySz(\theta_0)$. For $h = (h_1,h_2) \in H$, let

$$\tau_n(h) = 1 - J_h(c_{(\delta_n^r h_1,h_2)}(1-\alpha)). \tag{13.20}$$

By definition, $AsySz_n(\theta_0) = \sup_{h\in H} \tau_n(h)$. There exists a sequence $\{h_n \in H : n \geq 1\}$ such that

$$\limsup_{n\to\infty} \sup_{h\in H} \tau_n(h) = \limsup_{n\to\infty} \tau_n(h_n). \tag{13.21}$$

There exists a subsequence $\{u_n\}$ of $\{n\}$ such that

$$\limsup_{n\to\infty} \tau_n(h_n) = \lim_{n\to\infty} \tau_{u_n}(h_{u_n}). \tag{13.22}$$

There exists a subsequence $\{v_n\}$ of $\{u_n\}$ such that

$$(h_{v_n,1}, h_{v_n,2}, \delta_{v_n}^r h_{v_n,1}) \to (h_1^*, h_2^*, g_1^*) \tag{13.23}$$

for some $h_1^* \in H_1$, $h_2^* \in H_2$, $g_1^* \in H_1$, where $(g^*, h^*) = ((g_1^*, h_2^*), (h_1^*, h_2^*)) \in GH$.

Hence, we have

$$\begin{aligned} &\limsup_{n\to\infty} AsySz_n(\theta_0) = \lim_{n\to\infty} \tau_{u_n}(h_{u_n}) = \lim_{n\to\infty} \tau_{v_n}(h_{v_n}) \\ &= \lim_{n\to\infty} \Big(1 - J_{h_{v_n}}(c_{(\delta_{v_n}^r h_{v_n,1}, h_{v_n,2})}(1-\alpha))\Big) = 1 - J_{h^*}(c_{g^*}(1-\alpha)) \\ &\le \sup_{(g,h)\in GH} (1 - J_h(c_g(1-\alpha))) = AsySz_n(\theta_0), \end{aligned} \tag{13.24}$$

where the fourth equality holds by Assumption P and (13.23). This completes the proof. □

Proof of Theorem 5. Part (a)(i) holds by the proof of Theorem 4 with $1 - J_h(c_{(\delta_n^r h_1, h_2)})$ and $1 - J_h(c_g)$ replaced by $c_{(h_1,h_2)} - c_{(\delta_n^r h_1, h_2)}$ and $c_h - c_g$, respectively, using Assumption Q in place of Assumption P. Next, we show part (a)(ii). Using part (a)(i), by the same argument as used to prove Theorem 2(b) of AG1, $AsySz(\theta_0)$ for the ASC-Sub test equals $AsySz(\theta_0)$ for the SC-Sub test. By Theorem 1(b) above, the latter equals α.

Now, we prove part (b)(i). If H^* is empty, then $\liminf_{n\to\infty} \sup_{h\in H^*(\delta_n)} c_h \ge -\infty = \sup_{h\in H^*} c_h$. If H^* is non-empty, for any $(g,h) \in GH$ such that $h \in H^*$, define $(g_n, h_n) \in GH$ as in (13.18). By $(g_n, h_n) \to (g,h)$, Assumption Q, and $c_g - c_h < 0$, we obtain $c_{g_n} - c_{h_n} < 0$ and $h_n \in H^*(\delta_n)$ for all n sufficiently large. Hence,

$$\liminf_{n\to\infty} \sup_{h\in H^*(\delta_n)} c_h \ge \liminf_{n\to\infty} c_{h_n} = c_h, \tag{13.25}$$

where the equality uses $h_n \to h$ and Assumption Q. This inequality holds for all $h \in H^*$. Hence, $\liminf_{n\to\infty} \sup_{h\in H^*(\delta_n)} c_h \ge \sup_{h\in H^*} c_h$ and part (b)(i) holds.

Next, we show part (b)(ii). Using part (b)(i), by the same argument as used to prove Theorem 2(b) of AG1, $AsySz(\theta_0)$ for the ASC-Hyb test is less than or equal to $AsySz(\theta_0)$ for the SC-Hyb test. By Theorem 1(c), the latter equals α.

To show that the first result of part (b)(iii) holds, first suppose that H^* is empty. Then, $\kappa^*(\alpha) = -\infty$, $H^*(\delta)$ is empty for $\delta > 0$ close to zero by Assumption R, and $\kappa^*(\delta_n, \alpha) = -\infty$ for n sufficiently large. Next, suppose that H^* is non-empty. Then, using Assumption R, it suffices to show that $\limsup_{n\to\infty} \sup_{h\in H^*(\delta_n)} c_h \le \sup_{h\in H^\dagger} c_h$. As in (13.20)-(13.23), there exists a sequence $\{h_n \in H^*(\delta_n) : n \ge 1\}$, a subsequence $\{u_n\}$ of $\{n\}$, and a subsequence $\{v_n\}$ of $\{u_n\}$ such that

$$\begin{aligned} &\limsup_{n\to\infty} \sup_{h\in H^*(\delta_n)} c_h = \lim_{n\to\infty} c_{h_{v_n}}, \\ &(h_{v_n,1}, h_{v_n,2}) \to (h_1^*, h_2^*) = h^*, \text{ and } (\delta_{v_n}^r h_{v_n,1}, h_{v_n,2}) \to (g_1^*, h_2^*) = g^* \end{aligned} \tag{13.26}$$

for some $(g^*, h^*) \in GH$. Since $h_{v_n} = (h_{v_n,1}, h_{v_n,2}) \in H^*(\delta_{v_n})$ for all n, we have $h^* \in H^\dagger$ by definition of $H^\dagger$. This, (13.26), and Assumption Q yield

$$\limsup_{n\to\infty} \sup_{h\in H^*(\delta_n)} c_h = \lim_{n\to\infty} c_{h_{v_n}} = c_{h^*} \leq \sup_{h\in H^\dagger} c_h, \tag{13.27}$$

which establishes that part (b)(i) holds with $\geq$ replaced by $=$. Given this, part (b)(ii) holds with $\geq$ replaced by $=$ by the same argument as given for part (b)(ii) above. Hence, part (b)(iii) holds. □

Proof of Theorem 6. We use the result that a sequence of random variables $\{X_n : n \geq 1\}$ satisfies $X_n \to_p 0$ iff for every subsequence $\{u_n\}$ of $\{n\}$ there is a subsequence $\{v_n\}$ of $\{u_n\}$ such that $X_{v_n} \to 0$ a.s. To prove part (a)(i), we apply this result with $X_n = \kappa_{\widehat{\gamma}_{n,2}}(\delta_n, \alpha) - \kappa_{\gamma_{n,2}}(\alpha)$. Hence, it suffices to show that given any $\{u_n\}$ there exists a subsequence $\{v_n\}$ of $\{u_n\}$ such that $X_{v_n} \to 0$ a.s. Given $\{u_n\}$, we apply the above subsequence result a second time with $X_n = \widehat{\gamma}_{n,2} - \gamma_{n,2}$ to guarantee that there is a subsequence $\{v_n\}$ of $\{u_n\}$ for which $\widehat{\gamma}_{v_n,2} - \gamma_{v_n,2} \to 0$ a.s. using Assumption N. The subsequence $\{v_n\}$ can be chosen such that $\gamma_{v_n,2} \to h_2$ for some $h_2 \in H_2$ because every sequence in H_2 has a convergent subsequence given that H_2 is compact. Now, the argument in the proof of Theorem 4 applied to the subsequence $\{v_n\}$, with $1 - J_h(c_{(\delta^r_{v_n} h_1, h_2)})$ and $1 - J_h(c_g)$ replaced by $c_{(h_1,h_2)} - c_{(\delta^r_{v_n} h_1, h_2)}$ and $c_h - c_g$, and using Assumption Q in place of Assumption P, gives the desired result.

Part (b)(i) holds using similar subsequence arguments to those above combined with variations of the proofs of parts (b)(i) and (b)(iii) of Theorem 5 with H^*, $H^*(\delta_n)$, $H^\dagger$, and Assumption R replaced by $H^*_{h_2}$, $H^*_{\gamma_{n,2}}(\delta_n)$, $H^\dagger_{h_2}$, and Assumption S, respectively.

Given the results of parts (a)(i) and (b)(i), parts (a)(ii) and (b)(ii) are proved using the same argument as used to prove part (b)(ii) of Theorem 3. □

Appendix B

In this Appendix, we provide three sets of results. First, we define a size-corrected combined (SC-Com) test that combines the SC-Sub and SC-Hyb tests. Second, we provide a table of asymptotic and finite sample results for model 2 of the Autoregressive Parameter Example. Third, we verify assumptions for the Autoregressive Parameter Example and Conservative Model Selection Example considered in the paper.

14 Size-Corrected Combined Test

Theorem 2(c)(iv)-(vi) and Figure 2(f) show that in some contexts the SC-Hyb test can be more powerful than the SC-Sub test for some $(g, h) \in GH$ and vice versa for other $(g, h) \in GH$. This implies that a test that combines the SC-Hyb and SC-Sub tests can be more powerful than both. In this Section, we introduce such a test. It is called the size-corrected combined (SC-Com) test. This test has power advantages over the SC-Hyb and SC-Sub tests. This is illustrated in Figure 2(f) where the critical value function of the SC-Com test is the minimum of the upper horizontal SC-Hyb critical value function and the upper curved SC-Sub critical value function. On the other hand, the SC-Com test has computational disadvantages because it requires computation of the critical values for both the SC-Sub and SC-Hyb tests, which requires calculation of $\kappa(\alpha)$ and $\kappa^*(\alpha)$ in cases where both the subsample and hybrid tests need size correction. Furthermore, in most contexts, the SC-Hyb test is more powerful than the SC-Sub for all $(g, h) \in GH$, so the SC-Com test just reduces to the SC-Hyb test. For example, this occurs in the cases illustrated in Figures 2(a)-(e).

The size-corrected combined (SC-Com) test rejects $H_0 : \theta = \theta_0$ when

$$\begin{aligned} T_n(\theta_0) &> c_{n,Com}(1-\alpha), \text{ where} \qquad (14.1) \\ c_{n,Com}(1-\alpha) &= \min\{c_{n,b}(1-\alpha) + \kappa(\alpha), \max\{c_{n,b}(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\}\}, \end{aligned}$$

where the constants $\kappa(\alpha)$ and $\kappa^*(\alpha)$ are defined in (3.2).

The following result shows that the SC-Com test has $AsySz(\theta_0) = \alpha$.

Theorem 7 *Suppose Assumptions* A2, B2, C-E, F2, G2, K, L, MS, *and* MH *hold. Then, the SC-Com test satisfies* $AsySz(\theta_0) = \alpha$.

Comments. 1. By definition, the critical value, $c_{n,Com}(1-\alpha)$, of the SC-Com test is less than or equal to those of the SC-Sub and SC-Hyb tests. By (4.2), it is less than or equal to that of the SC-FCV test as well. Hence, the SC-Com test is at least as powerful as the SC-Sub, SC-Hyb, and SC-FCV tests.

2. A PSC-Com test can be defined as in (14.1) with $\kappa(\alpha)$ and $\kappa^*(\alpha)$ replaced by $\kappa_{\widehat{\gamma}_{n,2}}(\alpha)$, and $\kappa^*_{\widehat{\gamma}_{n,2}}(\alpha)$.

3. An ASC-Com test can be defined as in (14.1) with $\kappa(\alpha)$ and $\kappa^*(\alpha)$ replaced by $\kappa(\delta_n, \alpha)$ and $\kappa^*(\delta_n, \alpha)$, respectively. Suppose Assumptions *A2, B2, C-E, F2, G2, K, Q, L, MS,* and *MH* hold. Then, the ASC-Com test satisfies $AsySz(\theta_0) = \alpha$. This holds by the argument in the proof of Theorem 7 using the results of parts (a)(i) and (b)(i). (Note that the result of part (b)(iii) is not needed for this argument and, hence, Assumption R is not needed for the result stated.)

4. An APSC-Com test can be defined as in (14.1) with $\kappa(\alpha)$ and $\kappa^*(\alpha)$ replaced by $\kappa_{\widehat{\gamma}_{2,n}}(\delta_n, \alpha)$ and $\kappa^*_{\widehat{\gamma}_{2,n}}(\delta_n, \alpha)$, respectively. Suppose Assumptions *A2, B2, C-E, F2, G2, K, N, Q, L, OS, OH,* and *R* hold. Then, the APSC-Com test satisfies $AsySz(\theta_0) = \alpha$. The proof of this follows from parts (a) and (b) of Theorem 6 and its proof.

Proof of Theorem 7. By the same argument as in the proof of Theorem 2(b) of AG1, the SC-Com test satisfies

$$\begin{aligned} & AsySz(\theta_0) \qquad (14.2) \\ \leq & \sup_{(g,h)\in GH} [1 - J_h(\min\{c_g(1-\alpha) + \kappa(\alpha), \max\{c_g(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\}\}-)]. \end{aligned}$$

By the proof of Theorem 1, the constants $\kappa(\alpha)$ and $\kappa^*(\alpha)$ defined in (3.2) are such that (13.5) holds and hence for all $(g,h) \in GH$,

$$\begin{aligned} 1 - J_h(c_g(1-\alpha) + \kappa(\alpha)-) &\leq \alpha \text{ and} \\ 1 - J_h(\max\{c_g(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\}-) &\leq \alpha. \qquad (14.3) \end{aligned}$$

Equations (14.2) and (14.3) combine to give $AsySz(\theta_0) \leq \alpha$.

The SC-Com test has $AsySz(\theta_0) \geq \alpha$ because its $AsySz(\theta_0)$ is greater than or equal to that of the SC-Sub test (because its critical value is no larger) and the latter equals α by Theorem 1(b). $\square$

15 CI for an Autoregressive Parameter

This section provides (i) Table B-I, which is analogous to Table II but for model 2 instead of model 1,[13] and (ii) verification of the assumptions for this example.

The general features of Tables II and B-I are the same. The primary difference between these tables is that the asymptotic, adjusted asymptotic, and finite-sample sizes of the upper and equal-tailed FCV and Sub CIs are noticeably lower for model 2

[13] The results are based on 20,000 simulation repetitions. The search over h to determine Min is done on the interval $[-.90, 1.00]$ with stepsize 0.01 on $[-.90, .90]$ and stepsize .001 on $[.90, 1.0]$. The asymptotic results are computed using a discrete approximation to the continuous stochastic process on $[0, 1]$ with 10,000 grid points. The size-correction values $cv(1-\alpha)$, $\kappa(\alpha)$, and $\kappa(\alpha, \delta)$ for model 2 of the AR(1) example are as follows: for upper tests, $\kappa(.05) = 2.55$ & $\kappa(.05, .10) = 1.27$; for lower tests, $cv(.95) = 3.41$; for symmetric tests, $cv(.95) = 3.41$; and for equal-tailed tests, $cv(.95) = 1.43$, $\kappa(.05) = 2.30$, & $\kappa(.05, .95) = 1.00$. The finite-sample results in model 1 are invariant to the true values of α and σ^2_U and, hence, these are taken to be 0 and 1, respectively.

than model 1. For example, the finite-sample sizes of nominal 95% equal-tailed FCV CIs are 69.7 in model 1 and 38.8 in model 2. For Sub CIs, the corresponding sizes are 86.7 and 79.4. The Hyb CI and the ASC CIs perform well in terms of size in both models 1 and 2.

15.1 Verification of Assumptions

Here we verify the assumptions of Corollaries 3(b), 4, and 5 which apply to CIs, viz., Assumptions A2, B2, C-E, F2, G2, J2, K, L, MF, MS, and MH, for the AR(1) Example. We use Lemma 2 of AG1 to verify Assumption G2. Lemma 2 of AG1 requires verification of Assumptions t1, Sub1, A2, BB2, C, DD, EE, and HH. These assumptions also imply Assumptions B2 and D.

The statements in this section hold for both models 1 and 2. Assumption t1 holds with $\tau_n = n^{1/2}$ by definition of $T_n^*(\theta_0)$. Assumptions Sub1 and A2 clearly hold. Assumption C holds by the choice of b_n. Assumption DD holds when the AR parameter is less than one by the assumption of a strictly stationary initial condition. In the unit root case, it holds by the i.i.d. assumption on the innovations for $i = 1, ..., n$ and the fact that the test statistic $T_n^*(\theta_0)$ is invariant to the initial condition. Verifications of Assumptions E and EE are given in Sections 15.3 and 15.4 below for model 1. For brevity, we do not verify these assumptions for model 2. Assumption F2 holds for $J_h = J_h^*$ and $J_h = -J_h^*$ because J_h^* is continuous on R and has support R for all $h \in H$. Assumption F2 holds for $J_h = |J_h^*|$ because $|J_h^*|$ is continuous on R_+ and has support R_+ for all $h \in H$. For the same reason, Assumption J2 holds for $J_h = J_h^*$ and Assumptions MF(ii), MS(ii), and MH(ii) hold for $J_h = J_h^*, -J_h^*$, and $|J_h^*|$.

Assumption K holds trivially because $H = R_{+,\infty}$ so there is only one $h \in H$ for which $h = \pm\infty$.

Assumption L holds by properties of the Ornstein-Uhlenbeck process. Numerical calculations indicate that the supremum and infimum in this assumption are attained at $h = 0$ or $h = \infty$ (depending upon whether the supremum or infimum is being considered and whether $J_h = J_h^*$, $-J_h^*$, or $|J_h^*|$). This indicates that Assumption MF(i) holds. Numerical calculations also indicate that the supremum in Assumption MS(i) is attained at $h = 0$ or $h = \infty$ and hence this assumption holds. Assumption MH(i) holds because $c_h(1-\alpha)$ is monotone in h (based on numerical calculations), which implies that either H^* is empty or $H^* = \{x : x > 0\}$ depending on whether $J_h = J_h^*$, $-J_h^*$, or $|J_h^*|$. When H^* is non-empty, $\sup_{h \in H^*} c_h(1-\alpha)$ is attained at $h = \infty$.

The normalization constants a_n and d_n that appear in Assumptions BB2, EE, and HH depend on $\gamma_{n,h}$ and are denoted $a_n(\gamma_{n,h})$ and $d_n(\gamma_{n,h})$. They are defined as follows. Let $\{w_n : n \geq 1\}$ be any subsequence of $\{n\}$. Let $\{\gamma_n \in \Gamma : n \geq 1\}$ be a sequence for which $w_n\gamma_n \to \infty$ or $w_n\gamma_n \to h < \infty$. Let $\rho_n = 1 - \gamma_n$. Define

$$a_{w_n}(\gamma_n) = \begin{cases} w_n^{1/2}(1-\rho_n^2)^{-1/2} & \text{if } w_n\gamma_n \to \infty \\ w_n & \text{if } w_n\gamma_n \to h < \infty. \end{cases} \quad \text{and}$$

$$d_{w_n}(\gamma_n) = \begin{cases} (1-\rho_n^2)^{-1/2} & \text{if } w_n\gamma_n \to \infty \\ w_n^{1/2} & \text{if } w_n\gamma_n \to h < \infty. \end{cases} \tag{15.1}$$

Note that when the $w_n = n$, the definitions in (15.1) coincide with those in Theorem 8 below, which is used to verify Assumption BB2. Given these definitions, Assumption HH holds by the following calculations. For all sequences $\{\gamma_{n,h} : n \geq 1\}$ for which $b_n\gamma_{n,h} \to g$ for some $g \in R_{+,\infty}$, if $b_n\gamma_{n,h} \to g = \infty$, then

$$\frac{a_{b_n}(\gamma_{n,h})}{a_n(\gamma_{n,h})} = \frac{b_n^{1/2}(1-\rho_{n,h}^2)^{-1/2}}{n^{1/2}(1-\rho_{n,h}^2)^{-1/2}} = \left(\frac{b_n}{n}\right)^{1/2} \to 0, \tag{15.2}$$

where $\rho_n = 1 - \gamma_{n,h}$ and using Assumption C(ii). If $n\gamma_{n,h} \to h = \infty$ and $b_n\gamma_{n,h} \to g < \infty$, then

$$\frac{a_{b_n}(\gamma_{n,h})}{a_n(\gamma_{n,h})} = \frac{b_n}{n^{1/2}(1-\rho_{n,h}^2)^{-1/2}} = \left(\frac{b_n^2}{n}\right)^{1/2}(1-\rho_{n,h}^2)^{1/2} \to 0, \tag{15.3}$$

where $b_n^2/n = O(1)$ by assumption and $1 - \rho_{n,h}^2 \to 0$ because $b_n\gamma_{n,h} \to g < \infty$. If $n\gamma_{n,h} \to h < \infty$, then

$$\frac{a_{b_n}(\gamma_{n,h})}{a_n(\gamma_{n,h})} = \frac{b_n^{1/2}(1-\rho_{n,h}^2)^{-1/2}}{n^{1/2}(1-\rho_{n,h}^2)^{-1/2}} = \left(\frac{b_n}{n}\right)^{1/2} \to 0. \tag{15.4}$$

Given the definitions of $a_n(\cdot)$ and $d_n(\cdot)$, $\tau_n = a_n(\gamma_{n,h})/d_n(\gamma_{n,h}) = n^{1/2}$ does not depend on $\gamma_{n,h}$, as is required.

Assumption BB2(ii) holds because $P_\gamma(\widehat{\sigma}_{n,b_n,j} > 0) = 1$ for all $n, b_n \geq 4$, $j = 1, ..., q_n$, and $\gamma \in \Gamma$. Assumptions BB2(i) and BB2(iii) are verified in the next section.

15.2 Verification of Assumption BB2

In this section, we verify Assumptions BB2(i) and BB2(iii) of AG1 for the AR(1) Example. Given the assumption of a stationary initial condition when $\rho < 1$, the results given here are closely related to results in Elliott (1999), Elliott and Stock (2001), and Müller and Elliott (2003). (The results are also related to those of Bobkowski (1983), Cavanagh (1985), Chan and Wei (1987), and Phillips (1987), who consider the AR(1) model with an initial condition that is not stationary.)

Based on the definition of $\mathcal{F}$, under a sequence $\{\gamma_{n,h} : n \geq 1\}$, the innovations $U_i = U_{n,i}$ satisfy the following assumption.

Assumption INOV. For each $n \geq 1$, $\{U_{n,i} : i = 0, \pm1, \pm2, ...\}$ are i.i.d. with mean 0, variance $\sigma_{Un}^2 > 0$, and $\sup_{n\geq1} E|U_{n,i}/\sigma_{Un}|^{2+\delta} < \infty$ for some $\delta > 0$.

Define h_n by $\gamma_{n,h,1} = h_n/n$. Then, $h_n \to h$ as $n \to \infty$ because $n\gamma_{n,h,1} \to h$. In this example, $h_n = 0$ corresponds to a unit root, i.e., $\rho_n = 1 - \gamma_{n,h,1} = 1 - h_n/n = 1$. If

$h_n = 0$, then the initial condition Y_n^* is arbitrary. If $h_n > 0$, then the initial condition satisfies the following assumption:

Assumption STAT. $Y_{n,0}^* = \sum_{j=0}^{\infty} \rho_n^j U_{n,-j}$, where $\rho_n = 1 - h_n/n$.

Let $W(\cdot)$ be a standard Brownian motion on $[0,1]$ and Z an independent standard normal. By definition,

$$I_h(r) = \int_0^r \exp(-(r-s)h)dW(s),$$

$$I_h^*(r) = I_h(r) + \frac{1}{\sqrt{2h}} \exp(-hr)Z, \text{ and}$$

$$I_{D,h}^*(r) = I_h^*(r) - \int_0^1 I_h^*(s)ds, \tag{15.5}$$

Assumptions BB2(i) and (iii) are verified by the results given in the following Theorem (with Q_h playing the role of W_h in Assumption BB2).

Theorem 8 *Suppose Assumption* INOV *holds, Assumption* STAT *holds when* $\rho_n < 1$, $\rho_n \in [-1+\varepsilon, 1]$ *for some* $0 < \varepsilon < 2$, *and* $\rho_n = 1 - h_n/n$ *where* $h_n \to h \in [0, \infty]$. *Then,*

$$a_n(\widehat{\rho}_n - \rho_n) \to_d V_h \text{ and } d_n \widehat{\sigma} \to_d Q_h,$$

where a_n, d_n, V_h, *and* Q_h *are defined as follows.*
(a) *In model* 1, *we have* (i) *for* $h \in [0, \infty)$, $a_n = n$, $d_n = n^{1/2}$, V_h *is the distribution of*

$$\int_0^1 I_{D,h}^*(r)dW(r) / \int_0^1 I_{D,h}^*(r)^2 dr, \tag{15.6}$$

and Q_h *is the distribution of*

$$\left[\int_0^1 I_{D,h}^*(r)^2 dr\right]^{-1/2}, \tag{15.7}$$

and (ii) *for* $h = \infty$, $a_n = (1-\rho_n^2)^{-1/2} n^{1/2}$, $d_n = (1-\rho_n^2)^{-1/2}$, V_h *is a* $N(0,1)$ *distribution, and* Q_h *is a pointmass at one distribution.*
(b) *In model* 2, *we have* (i) *for* $h \in [0, \infty)$, $a_n = n$, $d_n = n^{1/2}$, V_h *is the distribution of*

$$\frac{\int_0^1 [I_{D,h}^*(r) - 12\int_0^1 I_{D,h}^*(s)sds \cdot (r-1/2)]dW(r)}{\int_0^1 [I_{D,h}^*(r) - 12\int_0^1 I_{D,h}^*(s)sds \cdot (r-1/2)]^2 dr}, \tag{15.8}$$

and Q_h *is the distribution of*

$$\left(\int_0^1 [I_{D,h}^*(r) - 12\int_0^1 I_{D,h}^*(s)sds \cdot (r-1/2)]^2 dr\right)^{-1/2}, \tag{15.9}$$

and (ii) *for* $h = \infty$, a_n, d_n, V_h, *and* Q_h *are as in part* (ii) *for model* 1.

Comment. The definitions of the normalization constants a_n and d_n in Theorem 8 correspond to those given in (15.1) with $a_n = a_n(\gamma_n)$ and $d_n = d_n(\gamma_n)$, where $\gamma_n = 1 - \rho_n$.

The proof of Theorem 8 uses several lemmas that deal with the case of $h \in [0, \infty)$.

In integral expressions below, we often leave out the lower and upper limits zero and one, the argument r, and dr to simplify notation when there is no danger of confusion. For example, $\int_0^1 I_h(r)^2 dr$ is typically written as $\int I_h^2$. By " $\Rightarrow$ " we denote weak convergence as $n \to \infty$.

Lemma 3 *Suppose Assumptions* INOV *and* STAT *hold,* $\rho_n \in (-1, 1)$ *and* $\rho_n = 1 - h_n/n$ *where* $h_n \to h \in [0, \infty)$ *as* $n \to \infty$. *Then,*

$$(2h_n/n)^{1/2} Y_{n,0}^* / \sigma_{Un} \to_d Z \sim N(0, 1).$$

Define $h_n^* \geq 0$ by $\rho_n = \exp(-h_n^*/n)$. As shown in the proof of Lemma 3, $h_n^*/h_n \to 1$ when $h \in [0, \infty)$. By recursive substitution, we have

$$\begin{aligned} Y_{n,i}^* &= \widetilde{Y}_{n,i} + \exp(-h_n^* i/n) Y_{n,0}^*, \text{ where} \\ \widetilde{Y}_{n,i} &= \textstyle\sum_{j=1}^i \exp(-h_n^*(i-j)/n) U_j. \end{aligned} \quad (15.10)$$

The next lemma shows that Lemma 1 in Phillips (1987) continues to hold under our more general assumptions, namely (i) innovations $\{U_{n,i} : i = 0, \pm 1, ...\}$ as in our Assumption INOV that depend on n and (ii) sequences $\rho_n = \exp(-h_n^*/n)$ where $h_n^* \to h \in [0, \infty)$, rather than the sequences $\rho_n = \exp(c/n)$ used in Phillips (1987), where c does not depend on n.

Lemma 4 *Suppose Assumption* INOV *holds,* $\rho_n \in (-1, 1]$, *and* $\rho_n = 1 - h_n/n$ *where* $h_n \to h \in [0, \infty)$. *Then, the following results hold jointly,*
(a) $n^{-1/2} \widetilde{Y}_{n,[nr]} / \sigma_{Un} \Rightarrow I_h(r)$ *for* $r \in [0, 1]$,
(b) $n^{-3/2} \sum_{i=1}^n \widetilde{Y}_{n,i-1} / \sigma_{Un} \Rightarrow \int I_h$,
(c) $n^{-2} \sum_{i=1}^n \widetilde{Y}_{n,i-1}^2 / \sigma_{Un}^2 \Rightarrow \int I_h^2$, *and*
(d) $n^{-1} \sum_{i=1}^n \widetilde{Y}_{n,i-1} U_i / \sigma_{Un}^2 \Rightarrow \int I_h(r) dW(r)$.

The following result is proved using Lemmas 3 and 4. Part (a) is similar to equation (3) of Elliott and Stock (2001). Let $\overline{Y}_n^* = n^{-1} \sum_{i=1}^n Y_{n,i-1}^*$.

Lemma 5 *Suppose Assumption* INOV *holds, Assumption* STAT *holds when* $\rho_n < 1$, $\rho_n \in (-1, 1]$, *and* $\rho_n = 1 - h_n/n$ *where* $h_n \to h \in [0, \infty)$. *Then, the following results hold jointly,*
(a) $n^{-1/2}(Y_{n,[nr]}^* - \overline{Y}_n^*)/\sigma_{Un} \Rightarrow I_h^*(r) - \int I_h^* = I_{D,h}^*(r)$,
(b) $n^{-2} \sum_{i=1}^n (Y_{n,i-1}^* - \overline{Y}_n^*)^2 / \sigma_{Un}^2 \Rightarrow \int (I_h^* - \int I_h^*)^2 = \int (I_{D,h}^*)^2$,
(c) $n^{-1} \sum_{i=1}^n (Y_{n,i-1}^* - \overline{Y}_n^*) U_{n,i} / \sigma_{Un}^2 \Rightarrow \int (I_h^*(r) - \int I_h^*) dW(r) = \int I_{D,h}^*(r) dW(r)$, *and*
(d) $\widehat{\sigma}_{Un}^2 / \sigma_U^2 \to_p 1$ *in models* 1 *and* 2.

In the proofs below we typically leave out the subindex n; for example, instead of $Y_{n,i}, \alpha_n, \beta_n, \rho_n, \sigma^2_{Un}$, and $U_{n,i}$ we simply write $Y_i, \alpha, \beta, \rho, \sigma^2_U$, and U_i. We do not drop n from h_n because h_n and h are different quantities.

Proof of Theorem 8. In model 1, $\widehat{\rho}$ and $\widehat{\sigma}$ can be written as

$$\begin{aligned} \widehat{\rho} - \rho_n &= \left(\sum_{i=1}^{n} (Y^*_{i-1} - \overline{Y}^*_n)^2 \right)^{-1} \sum_{i=1}^{n} (Y^*_{i-1} - \overline{Y}^*_n) U_i \text{ and} \\ \widehat{\sigma} &= \left(n^{-1} \sum_{i=1}^{n} (Y^*_{i-1} - \overline{Y}^*_n)^2 \right)^{-1/2} \widehat{\sigma}_{Un}. \end{aligned} \tag{15.11}$$

Hence, for $h \in [0, \infty)$, Lemma 5(b)-(d) implies that part (a)(i) of the Theorem holds with $a_n = n$ and $d_n = n^{1/2}$.

For model 2, by the partitioned regression formula, we have

$$\widehat{\rho} - \rho = ((Y_{-1} - \overline{Y})' M_D (Y_{-1} - \overline{Y}))^{-1} (Y_{-1} - \overline{Y})' M_D U, \tag{15.12}$$

where $Y_{-1} = (Y_0, ..., Y_{n-1})'$, $M_D = I_n - D(D'D)^{-1}D'$, D is the demeaned time trend n-vector whose ith element equals $i - n^{-1}\sum_{j=1}^n j$, and $\overline{Y}$ is an n-vector with all elements equal to $n^{-1}\sum_{j=1}^n Y_{j-1}$. Easy calculations show that $n^{-3}D'D = 12^{-1} + o(1)$. Furthermore, by arguments as in the proof of Lemma 1 in Phillips (1987), we have (up to lower order terms)

$$\begin{aligned} n^{-1/2} M_D (Y_{-1} - \overline{Y})/\sigma_U &= n^{-1/2}(Y_{-1} - \overline{Y})/\sigma_U - 12 n^{-1} D[n^{-5/2} D'(Y_{-1} - \overline{Y})/\sigma_U], \\ n^{-5/2} D'(Y_{-1} - \overline{Y})/\sigma_U &= n^{-5/2} \textstyle\sum_{i=1}^n i (Y^*_{i-1} - n^{-1}\sum_{j=1}^n Y^*_{j-1})/\sigma_U \Rightarrow \int s I^*_{D,h}(s) ds, \end{aligned} \tag{15.13}$$

where in the second term of the second equation we can write i instead of $i - \sum_{j=1}^n j/n$ because $\sum_{i=1}^n (Y^*_{i-1} - \overline{Y}^*_n) = 0$. We have $n^{-1}([rn] - n^{-1}\sum_{j=1}^n j) \to r - 1/2$ and, by Lemma 5(a), $n^{-1/2}(Y^*_{[nr]} - \overline{Y}^*_n)/\sigma_U \Rightarrow I^*_{D,h}(r)$. Combining these results with (15.13) and using steps as in the proof of Lemma 1(c) in Phillips (1987) we obtain the limit for the (normalized) denominator in (15.12):

$$n^{-2}(Y_{-1} - \overline{Y})' M_D (Y_{-1} - \overline{Y})/\sigma^2_U \Rightarrow \int_0^1 \left[I^*_{D,h}(r) - 12 \int_0^1 I^*_{D,h}(s) s ds \cdot (r - 1/2) \right]^2 dr. \tag{15.14}$$

The numerator in (15.12) is handled similarly as in the proof of Lemma 1(d) in Phillips (1987). This gives

$$n^{-1}(Y_{-1} - \overline{Y})' M_D U \Rightarrow \int_0^1 \left(I^*_{D,h}(r) - 12 \int_0^1 I^*_{D,h}(s) s ds \cdot (r - 1/2) \right) dW(r). \tag{15.15}$$

For $h \in [0, \infty)$, the combination of (15.14), (15.15), and Lemma 5(d) establishes part (b)(i) for model 2 with $a_n = n$ and $d_n = n^{1/2}$.

It remains to consider the case where $h = \infty$, i.e., parts (a)(ii) and (b)(ii) of the Theorem. These results follow from results given in Giraitis and Phillips (2006) generalized in the following ways: (i) from a no-intercept model to models 1 and 2, (ii) to a case in which the innovation distribution depends on n, and (iii) to cover the standard deviation estimator as well as the LS estimator itself. Each of these generalizations can be carried out using the same methods as in Giraitis and Phillips (2006). When the innovation distribution depends on n, Assumption A.2 in the Corrigendum to Giraitis and Phillips (2006) can be written as $EY_{n0}^{*2}/\sigma_{Un}^2 = o(n)$. Note that for $h = \infty$, it follows from Assumption STAT that $EY_{n0}^{*2}/\sigma_{Un}^2 = 1/(1-\rho_n^2) = 1/(2h_n/n - (h_n/n)^2) = o(n)$, as desired. □

Proof of Lemma 3. We have: $\rho_n = 1 - h_n/n$ and $h_n = O(1)$ implies that $\rho_n \to 1$. Hence, $\exp(-h_n^*/n) = \rho_n \to 1$ and $h_n^* = o(n)$. By a mean-value expansion of $\exp(-h_n^*/n)$ about 0,

$$0 = \rho_n - \rho_n = \exp(-h_n^*/n) - (1 - h_n/n) = h_n/n - \exp(-h_n^{**}/n)h_n^*/n, \qquad (15.16)$$

where $h_n^{**} = o(n)$ given that $h_n^* = o(n)$. Hence, $h_n - (1 + o(1))h_n^* = 0$, $h_n^*/h_n \to 1$, and it suffices to prove the result with h_n^* in place of h_n.

Let $\{m_n : n \geq 1\}$ be a sequence such that $m_n h_n^*/n \to \infty$. By Assumption STAT (which holds because $\rho_n < 1$), we can write $(2h_n^*/n)^{1/2} Y_0^*/\sigma_U = A_{1n} + A_{2n}$ for $A_{1n} = (2h_n^*/n)^{1/2} \sum_{j=0}^{m_n} \rho^j U_{-j}/\sigma_U$ and $A_{2n} = (2h_n^*/n)^{1/2} \sum_{j=m_n+1}^{\infty} \rho^j U_{-j}/\sigma_U$. Note that $EA_{2n} = 0$ and

$$\begin{aligned} &var(A_{2n}) = (2h_n^*/n) \textstyle\sum_{j=m_n+1}^{\infty} \rho^{2j} = (2h_n^*/n)\rho^{2(m_n+1)}/(1-\rho^2) \qquad (15.17) \\ &= (2h_n^*/n)\rho^{2(m_n+1)}/((2h_n^*/n)(1+o(1))) = O(\exp(-2(m_n+1)h_n^*/n)) = o(1), \end{aligned}$$

where the third equality holds because $\rho^2 = \exp(-2h_n^*/n) = 1 - (2h_n^*/n)(1 + o(1))$ by a mean value expansion and the last equality holds because $m_n h_n^*/n \to \infty$ by assumption. Therefore, $A_{2n} \to_p 0$.

The result now follows from $A_{1n} \to_d Z$ which holds by the CLT in Corollary 3.1 in Hall and Heyde (1980) with their $X_{n,i}$ being equal to $(2h_n^*/n)^{1/2}\rho^i U_{-i}/\sigma_U$. To apply their Corollary 3.1 we have to verify their (3.21), a Lindeberg condition, and a conditional variance condition. For all $i = 0, \pm 1, \pm 2, ...$, set $\mathcal{F}_{0,i} = \varnothing$ and define recursively $\mathcal{F}_{n+1,i} = \sigma(\mathcal{F}_{n,i} \cup \sigma(U_{n+1,j} : j = 0, -1, ..., -i))$ for $n \geq 1$. Then, (3.21) in Hall and Heyde (1980) holds automatically. To check the remaining two conditions, note first that $\sum_{i=0}^{m_n} E(X_{n,i}^2 | \mathcal{F}_{n,i-1}) = \sum_{i=0}^{m_n} EX_{n,i}^2 = 2h_n^* \sum_{i=0}^{m_n} \rho^{2i}/n \to 1$ which holds because $\sum_{i=0}^{m_n} \rho^{2i} = (1 - \rho^{2(m_n+1)})/(1-\rho^2)$, $\rho^{2(m_n+1)} = \exp(-2h_n^*(m_n+1)/n) \to 0$, and

$$n(1-\rho^2) = n(1-\rho)(1+\rho) = h_n(1+\rho) \to 2h. \qquad (15.18)$$

Secondly, for $\varepsilon > 0$,

$$
\begin{aligned}
&\textstyle\sum_{i=0}^{m_n} E(X_{n,i}^2 I(|X_{ni}| > \varepsilon)|\mathcal{F}_{n,i-1}) \\
&= \textstyle\sum_{i=0}^{m_n} EX_{n,i}^2 I(|X_{n,i}| > \varepsilon) \\
&\leq (2h_n^*/n) \textstyle\sum_{i=0}^{m_n} \rho^{2i} E((U_{n,-i}^2/\sigma_{Un}^2) I(2h_n^* U_{n,-i}^2/(n\sigma_{Un}^2) > \varepsilon^2)) \\
&= (2h_n^*/n)[\textstyle\sum_{i=0}^{m_n} \rho^{2i}] E((U_{n,0}^2/\sigma_{Un}^2) I(2h_n^* U_{n,0}^2/\sigma_{Un}^2 > n\varepsilon^2)) \\
&= O(1)o(1), \qquad (15.19)
\end{aligned}
$$

where the second equality holds because the U_{-i} have identical distributions. For the last equality, write $W_n = (U_{n,0}^2/\sigma_{Un}^2)$. For any $\alpha > 0$, $W_n I((2h_n^* W_n/(n\varepsilon^2))^\alpha > 1) \leq W_n^{1+\alpha}(2h_n^*/(n\varepsilon^2))^\alpha$ and the result follows from Assumption INOV which implies that $(2h_n^*/(n\varepsilon^2))^\alpha \; EW_n^{1+\alpha} = O(n^{-\alpha})$. $\square$

Proof of Lemma 4. Define

$$S_n(r) = n^{-1/2}\sigma_{Un}^{-1} \textstyle\sum_{i=1}^{[nr]} U_{n,i}. \qquad (15.20)$$

By Theorem 3.1 of De Jong and Davidson (2000), Assumption INOV implies that

$$S_n \Rightarrow W \qquad (15.21)$$

as processes indexed by $r \in [0, 1]$. (This theorem is applied with their X_{ni}, K_n, r, and c_{ni} equal to $n^{-1/2}\sigma_{Un}^{-1} U_{n,i}$, n, $2 + \delta$, and $n^{-1/2}$, respectively.)

The proof of the Lemma follows from the proof of Lemma 1 in Phillips (1987) by using (i) the functional central limit theorem in (15.21) and (ii) an application of the extended continuous mapping theorem (CMT), see Thm. 1.11.1 in van der Vaart and Wellner (1996) rather than the CMT used in Phillips (1987). The extended CMT is needed because the continuous function depends on n. For illustration, we prove part (a). By (15.10), we have

$$
\begin{aligned}
n^{-1/2}\widetilde{Y}_{[nr]}/\sigma_U \;&=\; \textstyle\sum_{j=1}^{[nr]} \exp(-h_n^*([nr] - j)/n) U_j/(n^{1/2}\sigma_U) \\
&=\; \textstyle\sum_{j=1}^{[nr]} \exp(-h_n^*([nr] - j)/n) \int_{(j-1)/n}^{j/n} dS_n(s) \\
&=\; \textstyle\sum_{j=1}^{[nr]} \int_{(j-1)/n}^{j/n} \exp(-h_n^*(r - s)) dS_n(s) + o_p(1) \\
&=\; \textstyle\int_0^r \exp(-h_n^*(r - s)) dS_n(s) + o_p(1) \\
&=\; S_n(r) + h_n^* \textstyle\int_0^r \exp(-h_n^*(r - s)) S_n(s) ds + o_p(1) \\
&\Rightarrow\; W(r) + h \textstyle\int_0^r \exp(-h(r - s)) W(s) ds \\
&=\; I_h(r), \qquad (15.22)
\end{aligned}
$$

where the second to last equality uses integration by parts, the convergence statement uses (15.21) and the extended CMT. The function $g_n : D_n \to E$ in Thm. 1.11.1 of van der Vaart and Wellner (1996) is given by $g_n(x)(r) = h_n^* \int_0^r \exp(-h_n^*(r - s)) x ds$, where $D_n = D[0,1]$ is the (not separable) metric space of CADLAG functions on the interval $[0,1]$ equipped with the uniform metric and $E = C[0,1]$ is the set of continuous functions on the interval $[0,1]$ also equipped with the uniform metric. Their set D_0 is also chosen as $D[0,1]$. If $x_n \to x$ in $D[0,1]$, then $g_n(x_n) \to g(x)$ in $C[0,1]$ because the function $h_n^* \exp(-h_n^*(r-s))$ converges uniformly (in $r \in [0,1]$) to $h \exp(-h(r-s))$ and any function in $D[0,1]$ is bounded. $\square$

Proof of Lemma 5. For $h \in (0,\infty)$, we have

$$\begin{aligned} n^{-1/2} Y_{[nr]}^*/\sigma_U &= n^{-1/2}\widetilde{Y}_{[nr]}/\sigma_U + n^{-1/2}\exp(-h_n^*[nr]/n)Y_0/\sigma_U \\ &\Rightarrow I_h(r) + (2h)^{-1/2}\exp(-rh)Z = I_h^*(r), \end{aligned} \tag{15.23}$$

where the equality holds by (15.10), and the convergence holds by Lemma 4(a), Lemma 3, and $\exp(-h_n[nr]/n) \to \exp(-rh)$ uniformly in $r \in [0,1]$. By (15.10), Z and the Brownian motion W are clearly independent. For $h \in (0,\infty)$, part (a) of the Lemma holds by (15.23) and the CMT, as in Lemma 1(b) of Phillips (1987).

For $h = 0$, by (15.10), we have

$$\begin{aligned} n^{-1/2}(Y_{[nr]}^* - \overline{Y}_n^*)/\sigma_U &= n^{-1/2}(\widetilde{Y}_{[nr]} - n^{-1}\textstyle\sum_{i=1}^n \widetilde{Y}_{i-1})/\sigma_U \\ &\quad + (h_n^*)^{-1/2}(\rho^{[nr]} - n^{-1}\textstyle\sum_{i=1}^n \rho^{i-1})(h_n^*/n)^{1/2} Y_0^*/\sigma_U. \end{aligned} \tag{15.24}$$

For $h = 0$, by a mean value expansion, we have

$$\begin{aligned} &\max_{0 \le j \le 2n} |1 - \rho^j| = \max_{0 \le j \le 2n} |1 - \exp(-h_n^* j/n)| = \max_{0 \le j \le 2n} |1 - (1 - h_n^* j \exp(m_j)/n)| \\ &\le 2h_n^* \max_{0 \le j \le 2n} |\exp(m_j)| = O(h_n^*), \end{aligned} \tag{15.25}$$

for $0 \le |m_j| \le h_n^* j/n \le 2h_n^* \to 0$. From (15.25), we have $\rho^{[nr]} - n^{-1}\sum_{i=1}^n \rho^{i-1} = O(h_n^*)$. This, together with Lemma 3 and $h_n^*/h_n \to 1$, implies that the second summand on the right-hand side of (15.24) is $o_p(1)$ for any sequence $\{\rho_n : n \ge 1\}$ for which $\rho_n = 1 - h_n/n < 1$ for $n \ge 1$ and $h_n \to 0$. On the other hand, if $\rho_n = 1$ for any n, then the second summand on the right-hand side of (15.24) is exactly zero. Hence, for $h = 0$, whether or not $\rho_n = 1$ for any n, the second summand on the right-hand side of (15.24) is $o_p(1)$. In consequence, part (a) of the Lemma for $h = 0$ follows from Lemma 4(a) and (b).

Given part (a), parts (b) and (c) of the Lemma are proved using the same sort of arguments as in Lemma 1 in Phillips (1987).

To prove part (d) for model 1, we write

$$\begin{aligned} \widehat{\sigma}_U^2/\sigma_U^2 &= (\widehat{\rho} - \rho)^2 (n-2)^{-1} \textstyle\sum_{i=1}^n (Y_{i-1}^* - \overline{Y}_n^*)^2/\sigma_U^2 \\ &\quad + 2(\widehat{\rho} - \rho)(n-2)^{-1}\textstyle\sum_{i=1}^n (Y_{i-1}^* - \overline{Y}_n^*)U_i/\sigma_U^2 + (n-2)^{-1}\textstyle\sum_{i=1}^n U_i^2/\sigma_U^2. \end{aligned} \tag{15.26}$$

For $h \in [0,\infty)$, by the expression for $\widehat{\rho} - \rho_n$ given in (15.11) and Lemma 5(b) and (c), we have $n(\widehat{\rho} - \rho_n) = O_p(1)$. Combining this with Lemma 5(b) and (c) shows that the first two summands in (15.26) are $O_p(n^{-1})$. The third summand in (15.26) is $1 + o_p(1)$ by Markov's inequality. Specifically, for $M_{n,i} = (U_{n,i}^2/\sigma_U^2) - 1$ and $\delta > 0$, Markov's inequality yields

$$P(|n^{-1}\sum_{i=1}^n U_i^2/\sigma_U^2 - 1| > \varepsilon) = P(|n^{-1}\sum_{i=1}^n M_i| > \varepsilon) = O(n^{-(1+\delta)})E|\sum_{i=1}^n M_i|^{1+\delta}. \tag{15.27}$$

By Assumption INOV and Minkowski's inequality, we have $E|\sum_{i=1}^n M_i|^{1+\delta} = O_p(n)$ for δ small enough, which proves the claim. Similar arguments, using results in Giraitis and Phillips (2006), establish the result of part (d) when $h = \infty$ in model 1. For model 2, similar arguments as those given above establish part (d). □

15.3 Verification of Assumption E

In this section, we verify Assumption E for model 1. As argued in the next paragraph, it is enough to show that for all $x \in R$, $U_{n,b_n}(x) - E_{\theta_0,\gamma_n}U_{n,b_n}(x) \to_p 0$ under $\{\gamma_n : n \geq 1\}$ for all sequences $\{\gamma_n = 1 - \rho_n \in \Gamma : n \geq 1\}$ that satisfy $n(1-\rho_n) \to h$ and $b(1-\rho_n) \to g$ for $(g,h) \in GH$.

To show $U_{n,b_n}(x) - E_{\theta_0,\gamma_n}U_{n,b_n}(x) \to_p 0$ under an arbitrary sequence $\{\gamma_n \in \Gamma : n \geq 1\}$ it is enough to show that for any subsequence $\{t_n\}$ there is a sub-subsequence $\{s_n\}$ such that $U_{s_n,b_{s_n}}(x) - E_{\theta_0,\gamma_{s_n}}U_{s_n,b_{s_n}}(x) \to_p 0$ under $\{\gamma_{s_n} \in \Gamma : n \geq 1\}$. Given any subsequence $\{t_n\}$ we can always construct a sub-subsequence $\{s_n\}$ such that $s_n(1 - \rho_{s_n}) \to h$ and $b_{s_n}(1-\rho_{s_n}) \to g$ for $(g,h) \in GH$. Proceeding as in the proof of Lemma 4(c) in AG1, we can define a sequence $\{\gamma_n^* : n \geq 1\}$ such that $n(1-\rho_n^*) \to h$ and $b(1-\rho_n^*) \to g$ and $\gamma_{s_n}^* = \gamma_{s_n}$. It follows that $U_{n,b_n}(x) - E_{\theta_0,\gamma_n}U_{n,b_n}(x) \to_p 0$ holds under $\{\gamma_n^* : n \geq 1\}$ and therefore $U_{s_n,b_{s_n}}(x) - E_{\theta_0,\gamma_{s_n}}U_{s_n,b_{s_n}}(x) \to_p 0$ holds under $\{\gamma_{s_n} \in \Gamma : n \geq 1\}$.

For notational simplicity, in the rest of this section we let ρ denote ρ_n.

It is sufficient to show that for any given $x \in R$, $var(U_{n,b_n}(x)) \to 0$ under all sequences $\{\gamma_n \in \Gamma : n \geq 1\}$ as described above. Recall that $T_{n,b,k}(\rho)$ denotes the studentized t statistic based on the $k-$th subsample, where the full-sample version is defined in (9.3).[14] We write $T_{n,k}$ instead of $T_{n,b,k}(\rho)$ to simplify notation. Define

$$I_{b,k} = 1\{T_{n,k} \leq x\}. \tag{15.28}$$

Stationarity of $I_{b,k}$ in k implies that

$$var(U_{n,b_n}(x)) = q_n^{-1}var(I_{b,0}) + 2q_n^{-2}\sum_{k=1}^{q_n-1}(q_n - k)Cov(I_{b,0}, I_{b,k}). \tag{15.29}$$

In this example, $q_n = n - b + 1$. Thus, it suffices to show $n^{-1}\sum_{k=0}^n |Cov(I_{b,0}, I_{b,k})|$ $\to 0$. This is implied by

$$\sup_{k \geq k_n} |Cov(I_{b,0}, I_{b,k})| \to 0 \tag{15.30}$$

[14]Here we deal with the upper one-sided case, so that $T_{n,b,k}(\rho) = T_{n,b,k}^*(\rho)$. The lower one-sided and symmetric two-sided cases can be dealt with using the same approach.

as $n \to \infty$ for some sequence $k_n \to \infty$ such that $k_n/n \to 0$.

Below we show that for all $k \geq k_n$ we can write

$$T_{n,k} = \widetilde{T}_{n,k} + \eta_{n,k}, \tag{15.31}$$

for some random variables $\widetilde{T}_{n,k}$ and $\eta_{n,k}$ that satisfy $\widetilde{T}_{n,k}$ and $T_{n,0}$ are independent and $\eta_{n,k} = o_p(1)$ uniform in $k \geq k_n$ (by which we mean that $\forall \varepsilon > 0$, $\sup_{k \geq k_n} \Pr(|\eta_{n,k}| > \varepsilon) \to 0$). (Likewise, for an array $a_{n,k}$ of real numbers, we say that $a_{n,k}$ is $o(1)$ uniform in $k \geq k_n$, if $\sup_{k \geq k_n} |a_{n,k}| \to 0$ as $n \to \infty$.)

Using (15.31), we show below that

$$\begin{aligned} \sup_{k \geq k_n} |P(\widetilde{T}_{n,k} \leq x) - P(T_{n,k} \leq x)| \to 0 \text{ and} \\ \sup_{k \geq k_n} |P(T_{n,0} \leq x \ \&\ T_{n,k} \leq x) - P(T_{n,0} \leq x \ \&\ \widetilde{T}_{n,k} \leq x)| \to 0. \end{aligned} \tag{15.32}$$

Using these results, we obtain

$$\begin{aligned} Cov(I_{b,0}, I_{b,k}) &= EI_{b,0}I_{b,k} - EI_{b,0}EI_{b,k} \\ &= P(T_{n,0} \leq x \ \&\ T_{n,k} \leq x) - P(T_{n,0} \leq x)P(T_{n,k} \leq x) \\ &= P(T_{n,0} \leq x \ \&\ \widetilde{T}_{n,k} \leq x) - P(T_{n,0} \leq x)P(\widetilde{T}_{n,k} \leq x) + o(1) \\ &= o(1), \end{aligned} \tag{15.33}$$

where third equality holds by (15.32), the last equality holds by independence of $T_{n,0}$ and $\widetilde{T}_{n,k}$, and the $o(1)$ expression is uniform in $k \geq k_n$ by (15.32). Therefore (15.30) holds and the proof is complete except for the verifications of (15.31) and (15.32).

Equation (15.32) is established as follows. Equation (15.31) and $P(T_{n,k} \leq x) \to J_h(x)$ as $n \to \infty$ where J_h is continuous (by Theorem 8 above) imply that for all $\varepsilon > 0$ there exist $\delta > 0$ and $n_0 \in N$ such that for $n \geq n_0$ we have

$$\begin{aligned} &\sup_{k \geq k_n} P(|\widetilde{T}_{n,k} - T_{n,k}| > \delta) < \varepsilon/2, \\ &P(T_{n,k} \leq x + \delta) \leq P(T_{n,k} \leq x) + \varepsilon/2, \text{ and} \\ &P(T_{n,k} \leq x) \leq P(T_{n,k} \leq x - \delta) + \varepsilon/2. \end{aligned} \tag{15.34}$$

The latter two inequalities hold for all k because $T_{n,k}$ is identically distributed across k. These results lead to

$$\begin{aligned} &\sup_{k \geq k_n} |P(\widetilde{T}_{n,k} \leq x) - P(T_{n,k} \leq x)| \\ = &\sup_{k \geq k_n} \max\{P(\widetilde{T}_{n,k} \leq x) - P(T_{n,k} \leq x), -P(\widetilde{T}_{n,k} \leq x) + P(T_{n,k} \leq x)\} \\ \leq &\sup_{k \geq k_n} \max\{P(\widetilde{T}_{n,k} \leq x) - P(T_{n,k} \leq x + \delta), \\ &\quad -P(\widetilde{T}_{n,k} \leq x) + P(T_{n,k} \leq x - \delta)\} + \varepsilon/2 \end{aligned}$$

$$\le \sup_{k\ge k_n} P(|\widetilde{T}_{n,k} - T_{n,k}| > \delta) + \varepsilon/2$$
$$\le \varepsilon, \tag{15.35}$$

which proves the first result in (15.32). The second result in (15.32) can be proved in the same way. For example, the analogue of the third equation in (15.34) holds because

$$\begin{aligned} &P(T_{n,0} \le x \ \&\ T_{n,k} \le x) - P(T_{n,0} \le x \ \&\ T_{n,k} \le x - \delta) \\ &\le P(x - \delta < T_{n,k} \le x) = P(x - \delta < T_{n,0} \le x) < \varepsilon/2 \end{aligned} \tag{15.36}$$

for all k, for $\delta > 0$ small enough. This completes the proof of (15.32).

It remains to establish (15.31). The statistic $T_{n,k}$ can be written as

$$\begin{aligned} T_{n,k} &= b^{1/2}(\widehat{\rho}_{n,b,k} - \rho)/\widehat{\sigma}_{n,b,k} = S_{1,k}/(S_{2,k}^{1/2} S_{3,k}), \text{ where} \\ S_{1,k} &= n_\gamma \sum_{i=1}^{b} (Y_{k+i-1}^* - \overline{Y}_{b,k}^*) U_{k+i}/\sigma_U^2, \\ S_{2,k} &= n_\gamma^2 \sum_{i=1}^{b} (Y_{k+i-1}^* - \overline{Y}_{b,k}^*)^2/\sigma_U^2, \\ S_{3,k} &= \widehat{\sigma}_{U,b,k}/\sigma_U, \ \overline{Y}_{b,k}^* = b^{-1} \sum_{i=1}^{b} Y_{k+i-1}^*, \end{aligned} \tag{15.37}$$

$\widehat{\sigma}_{U,b,k}^2$ is the sum of squared residuals divided by $b - 2$ based on a block of data of length b that starts at observation k, and $\{n_\gamma : n \ge 1\}$ is a normalization sequence that depends on the specific sequence $\gamma_{n,h}$ and is such that $S_{1,k} = O_p(1)$ and $S_{2,k}^{-1} = O_p(1)$.[15]

We show below that $S_{1,k}$ and $S_{2,k}$ can be written as

$$\begin{aligned} &S_{1,k} = \widetilde{S}_{1,k} + \xi_{1,k} \text{ and } S_{2,k} = \widetilde{S}_{2,k} + \xi_{2,k}, \text{ where} \\ &\widetilde{S}_{1,k} \text{ and } \widetilde{S}_{2,k} \text{ are independent of } S_{1,0} \text{ and } S_{2,0} \text{ for all } k \ge k_n, \\ &\xi_{1,k} = o_p(1), \ \xi_{2,k} = o_p(1), \end{aligned} \tag{15.38}$$

and the $o_p(1)$ terms hold uniformly for all $k \ge k_n$ for some sequence $k_n \to \infty$ such that $k_n/n \to 0$. In addition, we have $S_{3,k} - 1 = o_p(1)$ by Lemma 5(d) and the $o_p(1)$ term holds uniformly for all k by stationarity across k. This, (15.37), and (15.38) combine to show that (15.31) holds with

$$\widetilde{T}_{n,k} = \widetilde{S}_{1,k}/(\widetilde{S}_{2,k})^{1/2}. \tag{15.39}$$

It remains to establish (15.38). We separately analyze four cases: (i) $b(1-\rho) \to \infty$, (ii) $b(1-\rho) \to g \in (0,\infty)$, (iii) $b(1-\rho) \to 0$ & $n(1-\rho) \to \infty$, and (iv) $n(1-\rho) \to h \in [0,\infty)$.

[15] Strictly speaking, all sums over $i = 1, ..., b$ should be over $i = 1, ..., b-1$ because one observation from a block of length b is used as an initial observation given that lagged Y_i is a regressor. For notational simplicity, here and below, we sum to b rather than $b - 1$.

Proof of (15.38) for case (i): $b(1-\rho) \to \infty$.

From Theorem 2.2 and Lemmas 2.1 and 2.2 in Giraitis and Phillips (2006), we have

$$(1-\rho)b^{-1/2}\sum_{i=1}^{b} Y_{k+i-1}^{*}/\sigma_U \to_d N(0,1),$$
$$(1-\rho^2)^{1/2}b^{-1/2}\sum_{i=1}^{b} Y_{k+i-1}^{*}U_{k+i}/\sigma_U^2 \to_d N(0,1), \text{ and}$$
$$(1-\rho^2)b^{-1}\sum_{i=1}^{b} Y_{k+i-1}^{*2}/\sigma_U^2 \to_p 1. \tag{15.40}$$

By stationarity across k, these results hold uniformly in k.

In consequence, using $Y_i^* = \sum_{j=0}^{\infty} \rho^j U_{i-j}$, we define

$$\begin{aligned}
\widetilde{S}_{1,k} &= n_\gamma \sum_{i=1}^{b} \sum_{j=0}^{k+i-b-2} \rho^j U_{k+i-1-j} U_{k+i}/\sigma_U^2, \\
n_\gamma &= \frac{(1-\rho^2)^{1/2}}{b^{1/2}}, \\
\xi_{1,k} &= \xi_{11,k} - \xi_{12,k}, \\
\xi_{11,k} &= n_\gamma \sum_{i=1}^{b} \sum_{j=k+i-b-1}^{\infty} \rho^j U_{k+i-1-j} U_{k+i}/\sigma_U^2, \\
\xi_{12,k} &= n_\gamma \overline{Y}_{b,k}^{*} \sum_{i=1}^{b} U_{k+i}/\sigma_U^2, \\
\widetilde{S}_{2,k} &= 1, \text{ and} \\
\xi_{2,k} &= n_\gamma^2 \sum_{i=1}^{b} (Y_{k+i-1}^{*} - \overline{Y}_{b,k}^{*})^2/\sigma_U^2 - 1.
\end{aligned} \tag{15.41}$$

As defined, for $k > b$, $\widetilde{S}_{1,k}$ depends only on innovations U_t for $t > b$ and $S_{1,0}$ and $S_{2,0}$ depend only on innovations U_t for $t \leq b$. Hence, for $k > b$, $\widetilde{S}_{1,k}$ is independent of $S_{1,0}$ and $S_{2,0}$. Obviously, $\widetilde{S}_{2,k}$ is independent of $S_{1,0}$ and $S_{2,0}$.

Equation (15.40) and $b(1-\rho) \to \infty$ give

$$(1-\rho^2)\overline{Y}_{b,k}^{*2}/\sigma_U^2 = \frac{1+\rho}{b(1-\rho)}\left((1-\rho)b^{-1/2}\sum_{i=1}^{b} Y_{k+i-1}^{*}/\sigma_U\right)^2 = o_p(1) \tag{15.42}$$

uniformly in k. This and (15.40) combine to give

$$\xi_{2,k} = (1-\rho^2)b^{-1}\sum_{i=1}^{b} Y_{k+i-1}^{*2}/\sigma_U^2 - (1-\rho^2)\overline{Y}_{b,k}^{*2}/\sigma_U^2 - 1 = o_p(1) \text{ and}$$

$$\xi_{12,k} = (1-\rho^2)^{1/2}\overline{Y}^*_{b,k}\sigma_U^{-1}\left(b^{-1/2}\sum_{i=1}^{b} U_{k+i}/\sigma_U\right) = o_p(1) \tag{15.43}$$

uniformly in k.

It remains to show $\xi_{11,k} = o_p(1)$. By change of indices,

$$\xi_{11,k} = \rho^{k-b-1}\frac{(1-\rho^2)^{1/2}}{b^{1/2}}\sum_{i=1}^{b}\sum_{j=0}^{\infty}\rho^{j+i}U_{b-j}U_{k+i}/\sigma_U^2. \tag{15.44}$$

By Markov's inequality, for any $\varepsilon > 0$,

$$\begin{aligned}
&P(|\xi_{11,k}| > \varepsilon)\\
&\leq \varepsilon^{-2}\rho^{2(k-b-1)}(1-\rho^2)b^{-1}\sum_{i_1,i_2=1}^{b}\sum_{j_1,j_2=0}^{\infty}\rho^{j_1+i_1+j_2+i_2}EU_{k+i_1}U_{b-j_1}U_{k+i_2}U_{b-j_2}/\sigma_U^4\\
&= O(1)(1-\rho^2)b^{-1}\sum_{i=1}^{b}\sum_{j=0}^{\infty}\rho^{2j+2i}EU_{k+i}^2U_{b-j}^2/\sigma_U^4\\
&= O(1)(1-\rho^2)b^{-1}\sum_{i=0}^{b}\rho^{2i}\sum_{j=0}^{\infty}\rho^{2j}\\
&= O(1)b^{-1}(1-\rho^{2(b+1)})(1-\rho^2)^{-1}\\
&= O(1)(b(1-\rho))^{-1}(1+\rho)^{-1} = o(1),
\end{aligned} \tag{15.45}$$

where the first equality uses the fact that there are no indices $i_1 \geq 1$, $j_1 \geq 0$, and $j_2 \geq 0$ such that $k+i_1 = b-j_1$ or $k+i_1 = b-j_2$ and similarly for $k+i_2$ (which holds because $b-j_1 < k+i_1$ for $k > b$) and the last equality uses $b(1-\rho) \to \infty$ and the fact that ρ is bounded away from -1 by the definition of the parameter space.

Proof of (15.38) for case (ii): $b(1-\rho) \to h \in (0,\infty)$.

In this case, we set $n_\gamma = b^{-1}$. Straightforward but tedious calculations show that when $\rho < 1$

$$\begin{aligned}
Y^*_{k+i-1} - \overline{Y}^*_{b,k} &= \left(\rho^{i-1} - \frac{1-\rho^b}{b(1-\rho)}\right)\sum_{j=0}^{\infty}\rho^j U_{k-j} + \sum_{j=1}^{b-1}c_{i,j}U_{k+j}, \text{ where}\\
c_{i,j} &= 1(j \geq i)\rho^{i-j-1} - \frac{1-\rho^{b-j}}{b(1-\rho)}.
\end{aligned} \tag{15.46}$$

Therefore, we define

$$\widetilde{S}_{1,k} = b^{-1}\sigma_U^{-2}\sum_{i=1}^{b}\left[\left(\rho^{i-1} - \frac{1-\rho^b}{b(1-\rho)}\right)U_{k+i}\sum_{j=0}^{k-b-1}\rho^j U_{k-j} + U_{k+i}\sum_{j=1}^{b-1}c_{i,j}U_{k+j}\right] \text{ and}$$

$$
\begin{aligned}
\xi_{1,k} &= b^{-1}\sigma_U^{-2}\sum_{i=1}^{b}\left(\rho^{i-1}-\frac{1-\rho^b}{b(1-\rho)}\right)U_{k+i}\sum_{j=k-b}^{\infty}\rho^j U_{k-j} \\
&= b^{-1}\sigma_U^{-2}\rho^{k-b}\sum_{i=1}^{b}\left(\rho^{i-1}-\frac{1-\rho^b}{b(1-\rho)}\right)U_{k+i}\sum_{j=0}^{\infty}\rho^j U_{b-j}. \qquad (15.47)
\end{aligned}
$$

For $k > b$, $\widetilde{S}_{1,k}$ depends only on innovations U_t for $t > b$ and $S_{1,0}$ and $S_{2,0}$ depend only on innovations U_t for $t \leq b$. Thus, for $k > b$, $\widetilde{S}_{1,k}$ is independent of $S_{1,0}$ and $S_{2,0}$.

We now show that $\xi_{1,k} = o_p(1)$ uniformly for $k \geq k_n$ for some sequence $k_n \to \infty$ such that $k_n/n \to 0$. By Markov's inequality, we have

$$
\begin{aligned}
& P(|\xi_{1,k}| > \varepsilon) \\
& \leq \varepsilon^{-2}\rho^{2(k-b)}M(\rho,b)b^{-1}\sum_{j=0}^{\infty}\rho^{2j}EU_{k+i}^2U_{b-j}^2/\sigma_U^4 = \varepsilon^{-2}\rho^{2(k-b)}\frac{M(\rho,b)}{b(1-\rho^2)}, \text{ where}
\end{aligned}
$$

$$
\begin{aligned}
M(\rho,b) &= b^{-1}\sum_{i=1}^{b}\left(\rho^{i-1}-\frac{1-\rho^b}{b(1-\rho)}\right)^2 = \frac{1-\rho^{2b}}{b(1-\rho)(1+\rho)}-\frac{(1-\rho^b)^2}{b^2(1-\rho)^2} \\
&= \frac{b(1-\rho)(1-\rho^{2b})-(1+\rho)(1-\rho^b)^2}{b^2(1-\rho)^2(1+\rho)} \qquad (15.48)
\end{aligned}
$$

and the inequality uses $EU_{k+i_1}U_{k+i_2}U_{b-j_1}U_{b-j_2} = 0$ for $k > b$ and $i_1, i_2, j_1, j_2 \geq 0$ unless $i_1 = i_2$ and $j_1 = j_2$.

Define h_n^* by $\rho = \exp(-h_n^*/n)$. Because $b(1-\rho) \to h \in (0,\infty)$, we have $h_n^* \to \infty$. In consequence, there exists a sequence $\{k_n : n \geq 1\}$ such that $k_n/b \to \infty$, $k_n/n \to 0$ and $h_n^* k_n/n \to \infty$. For this sequence, $h_n^*(k_n - b)/n \to \infty$, $\rho^{k_n - b} = \exp(-h_n^*(k_n - b)/n) \to 0$, and $\sup_{k \geq k_n} \rho^{2(k-b)} \to 0$. In addition, $b(1-\rho) \to h$ implies that $\rho^b \to \exp(-h)$ and $\rho^{2b} \to \exp(-2h)$. Hence, we have

$$
M(\rho,b) \to \frac{h(1-\exp(-2h)) - 2(1-\exp(-h))^2}{2h^2}. \qquad (15.49)
$$

Combining these results implies that the right-hand side (rhs) of the inequality in (15.48) is $o(1)$ uniformly for $k \geq k_n$. Hence, $\xi_{1,k} = o_p(1)$ uniformly for $k \geq k_n$.

Next, based on (15.46), we define

$$
\begin{aligned}
\widetilde{S}_{2,k} &= b^{-2}\sum_{i=1}^{b}\left(\left(\rho^{i-1}-\frac{1-\rho^b}{b(1-\rho)}\right)\sum_{j=0}^{k-b-1}\rho^j U_{k-j} + \sum_{j=1}^{b-1}c_{i,j}U_{k+j}\right)^2/\sigma_U^2 \text{ and} \\
\xi_{2,k} &= b^{-1}\sigma_U^{-2}M(\rho,b)\sum_{j_1,j_2=k-b}^{\infty}\rho^{j_1+j_2}U_{k-j_1}U_{k-j_2} \\
&\quad +2b^{-2}\sigma_U^{-2}\sum_{i=1}^{b}\left(\rho^{i-1}-\frac{1-\rho^b}{b(1-\rho)}\right)\sum_{j_1=k-b}^{\infty}\rho^{j_1}U_{k-j_1}\sum_{j_2=1}^{b-1}c_{i,j_2}U_{k+j_2}
\end{aligned}
$$

$$+2b^{-1}\sigma_U^{-2}M(\rho,b)\sum_{j_1=k-b}^{\infty}\rho^{j_1}U_{k-j_1}\sum_{j_2=0}^{k-b-1}\rho^{j_2}U_{k-j_2}$$
$$= \xi_{21,k}+\xi_{22,k}+\xi_{23,k}. \tag{15.50}$$

Independence of $\widetilde{S}_{2,k}$ from $S_{1,0}$ and $S_{2,0}$ for $k > b$ holds by the same argument as above.

We now show $\xi_{21,k} = o_p(1)$ uniformly for $k \geq k_n$. By a change of indices, we have

$$\sum_{j_1,j_2=k-b}^{\infty}\rho^{j_1+j_2}U_{k-j_1}U_{k-j_2} = \rho^{2(k-b)}\sum_{j_1,j_2=0}^{\infty}\rho^{j_1+j_2}U_{b-j_1}U_{b-j_2}. \tag{15.51}$$

Markov's inequality now gives

$$\begin{aligned}P(|\xi_{21,k}|>\varepsilon) &= P\left(\left|b^{-1}\sigma_U^{-2}\rho^{2(k-b)}M(\rho,b)\sum_{j_1,j_2=0}^{\infty}\rho^{j_1+j_2}U_{b-j_1}U_{b-j_2}\right|>\varepsilon\right)\\ &\leq \varepsilon^{-2}M(\rho,b)^2\max\{EU_i^4/\sigma_U^4,1\}\rho^{4(k-b)}b^{-2}\sum_{j_1,j_2=0}^{\infty}\rho^{2j_1}\rho^{2j_2}\\ &= O(1)\rho^{4(k-b)}\frac{M(\rho,b)^2}{b^2(1-\rho)^2(1+\rho)^2}.\end{aligned} \tag{15.52}$$

The rhs is $o(1)$ because $\rho^{4(k-b)} \to 0$ uniformly for $k \geq k_n$ by the argument preceding (15.49), $M(\rho,b) = O(1)$ by (15.49), and $b^2(1-\rho)^2(1+\rho)^2 \to 4h^2 > 0$. Hence, $\xi_{21,k} = o_p(1)$ uniformly for $k \geq k_n$.

Next, we show $\xi_{22,k} = o_p(1)$ uniformly for $k \geq k_n$. By a change of indices (as in (15.51)), we can write

$$\xi_{22,k} = 2\rho^{k-b}b^{-2}\sigma_U^{-2}\sum_{j_1=0}^{\infty}\sum_{j_2=1}^{b-1}d_{j_2}\rho^{j_1}U_{b-j_1}U_{k+j_2}\text{ for } d_{j_2}=\sum_{i=1}^{b}\left(\rho^{i-1}-\frac{1-\rho^b}{b(1-\rho)}\right)c_{i,j_2}. \tag{15.53}$$

Note that $d_{j_2} = O(b)$ uniformly in j_2. By Markov's inequality, we have

$$\begin{aligned}P(|\xi_{22,k}|>\varepsilon) &= P\left(\left|2\rho^{k-b}b^{-2}\sigma_U^{-2}\sum_{j_1=0}^{\infty}\sum_{j_2=1}^{b-1}d_{j_2}\rho^{j_1}U_{b-j_1}U_{k+j_2}\right|>\varepsilon\right)\\ &\leq 4\varepsilon^{-2}\rho^{2(k-b)}b^{-4}\sum_{j_1=0}^{\infty}\sum_{j_2=1}^{b-1}d_{j_2}^2\rho^{2j_1}EU_{b-j_1}^2U_{k+j_2}^2/\sigma_U^4\\ &\leq 4\varepsilon^{-2}\rho^{2(k-b)}b^{-1}(1-\rho^2)^{-1}b^{-3}\sum_{j_2=1}^{b-1}d_{j_2}^2,\end{aligned} \tag{15.54}$$

where the first inequality holds for $k > b$ because $b - j_1 < k + j_2$ for all $j_1, j_2 \geq 0$. The rhs of (15.54) is $o(1)$ because $\rho^{2(k-b)} \to 0$ for $k \geq k_n$, $b(1-\rho^2) \to 2h > 0$, and $b^{-3}\sum_{j_2=1}^{b-1}d_{j_2}^2 = O(1)$.

To show $\xi_{23,k} = o_p(1)$ uniformly for $k \geq k_n$, by Markov's inequality and a change of indices (as in (15.51)), we have

$$\begin{aligned} P(|\xi_{23,k}| > \varepsilon) &= P\left(\left|2b^{-1}\sigma_U^{-2}\rho^{k-b}M(\rho,b)\sum_{j_1=0}^{\infty}\rho^{j_1}U_{b-j_1}\sum_{j_2=0}^{k-b-1}\rho^{j_2}U_{k-j_2}\right| > \varepsilon\right) \\ &\leq 4\rho^{2(k-b)}M(\rho,b)^2b^{-2}\sum_{j_1=0}^{\infty}\rho^{2j_1}EU_{b-j_1}^2\sum_{j_2=0}^{k-b-1}\rho^{2j_2}EU_{k-j_2}^2/\sigma_U^4 \\ &= \rho^{2(k-b)}M(\rho,b)^2O\left(\frac{1}{b^2(1-\rho^2)^2}\right). \end{aligned} \quad (15.55)$$

The rhs is $o(1)$ because $\rho^{2(k-b)} = o(1)$, $M(\rho,b) = O(1)$, and $b(1-\rho^2) \to 2h > 0$.

This completes the proof for case (ii).

Proof of (15.38) for case (iii): $b(1-\rho) \to 0$ & $n(1-\rho) \to \infty$.

Define $S_{1,k}, n_\gamma, \xi_{1,k}, S_{2,k}$, and $\xi_{2,k}$ as in case (ii). To show $\xi_{1,k} = o_p(1)$, it suffices to show the rhs of the inequality in (15.48) is $o(1)$.

Define h_n^* and h_n by $\rho = \exp(-h_n^*/n)$ and $\rho = 1 - h_n/n$. Let $t_n = bh_n^*/n$. Then, we have

$$\begin{aligned} \rho^b &= \exp(-bh_n^*/n) = \exp(-t_n),\ 1+\rho = 2 - h_n/n, \\ b(1-\rho) &= bh_n/n = t_n(h_n/h_n^*). \end{aligned} \quad (15.56)$$

We have: $b(1-\rho) \to 0 \Rightarrow \rho \to 1 \Rightarrow h_n^*/n \to 0 \Rightarrow h_n^*/h_n \to 1$, where the last implication follows from a mean-value expansion of $\exp(-h_n^*/n)$ about 0 as in (15.16). In addition, $b(1-\rho) \to 0 \Rightarrow bh_n/n \to 0$. Combining these results gives $t_n = (bh_n/n)(h_n^*/h_n) \to 0$. Also, $n(1-\rho) \to \infty$ implies that $h_n \to \infty$ and $h_n^* \to \infty$.

Because $bh_n/n = b(1-\rho) \to 0$ it follows that $h_n = o(n/b)$. This and $h_n^*/h_n \to 1$ yields $h_n^* = o(n/b)$. By a mean-value expansion of $\exp(-h_n^*/n)$ about 0, we obtain

$$0 = \rho - \rho = \exp(-h_n^*/n) - (1 - h_n/n) = h_n/n - \exp(-h_n^{**}/n)h_n^*/n, \quad (15.57)$$

where $h_n^{**}/n = o(1/b)$ because $h_n^* = o(n/b)$. Hence, $h_n/h_n^* = \exp(-h_n^{**}/n) = 1+o(1/b)$.

Using the above results, the rhs of the inequality in (15.48) equals $\varepsilon^{-2}\rho^{2(k-b)}$ times

$$\frac{M(\rho,b)}{b(1-\rho^2)} = \frac{[1-\exp(-2t_n)]t_n(h_n/h_n^*) - [1-\exp(-t_n)]^2(2-h_n/n)}{t_n^3(h_n/h_n^*)^3(2-h_n/n)^2}. \quad (15.58)$$

Using l'Hopital's rule three times, it is straightforward to show that

$$t_n^{-3}([1-\exp(-2t_n)]t_n - [1-\exp(-t_n)]^2 2) \to 0. \quad (15.59)$$

Furthermore,

$$\begin{aligned}
&t_n^{-3}\left([1-\exp(-2t_n)]t_n(h_n/h_n^*-1)+[1-\exp(-t_n)]^2h_n/n\right)\\
&=t_n^{-2}\left([1-\exp(-2t_n)](h_n/h_n^*-1)+[1-\exp(-t_n)]^2h_n/(h_n^*b)\right)\\
&=t_n^{-2}\left([2t_n\exp(\lambda_n)](h_n/h_n^*-1)+[t_n\exp(\lambda_n')]^2h_n/(h_n^*b)\right)\\
&=2t_n^{-1}\exp(\lambda_n)(h_n/h_n^*-1)+o(1)\\
&=2\frac{n}{bh_n^*}(1+o_p(1))o(b^{-1})+o(1)\\
&=o(1), \qquad (15.60)
\end{aligned}$$

where the second equality uses mean value expansions for $\exp(-2t_n)$ and $\exp(-t_n)$ about $t_n=0$ and $\lambda_n\in[-2t_n,0]$ and $\lambda_n'\in[-t_n,0]$, the second last equality uses the result of (15.57) that $h_n/h_n^*=1+o(b^{-1})$, and the last equality uses the assumption that $nb^{-2}=O(1)$ and the fact that $h_n^*\to\infty$.

Hence, we conclude that the expression in (15.58) is $o(1)$. In addition, $\rho^{2(k-b)}=O(1)$ uniformly for $k\geq k_n$ for $k_n=b+1$. This implies that the rhs of the inequality in (15.48) is $o(1)$ uniformly in $k\geq k_n$.

We have $\xi_{21,k}=o_p(1)$ and $\xi_{23,k}=o_p(1)$ uniformly for $k\geq k_n$ by (15.52), (15.55), and the above results for case (iii) that $M(\rho,b)=o(b(1-\rho^2))$ and $\rho^{4(k-b)}=O(1)$ uniformly for $k\geq k_n$.

We now show $\xi_{22,k}=o_p(1)$. By the Cauchy-Schwarz inequality, we have

$$b^{-2}d_{j_2}^2\leq M(\rho,b)b^{-1}\sum_{i=1}^{b}c_{i,j_2}^2=o\left(b(1-\rho^2)\right)O(1) \qquad (15.61)$$

uniformly for $j_2\leq b$ because $c_{i,j_2}^2=O(1)$ uniformly in $i\geq 1$ and $j_2\in\{1,...,b\}$. The latter holds because

$$\frac{1-\rho^{b-j}}{b(1-\rho)}=\frac{1-\exp(-h_n^*(b-j)/n)}{bh_n/n}=\frac{(h_n^*(b-j)/n)\exp(\lambda)}{bh_n/n}\leq\exp(\lambda)h_n^*/h_n\to 1 \qquad (15.62)$$

for some mean value $\lambda=o(1)$. Combining (15.54) and (15.61) yields $\xi_{22,k}=o_p(1)$ uniformly for $k\geq k_n$. This completes the proof for case (iii).

Proof of (15.38) for case (iv): $n(1-\rho)\to h\in[0,\infty)$.

When $\rho=1$ (which may occur in case (iv)), in place of (15.46) we have

$$Y_{k+i-1}^*-\overline{Y}_{b,k}^*=\sum_{j=1}^{b-1}c_{i,j}U_{k+j},\ \ c_{i,j}=1(j\geq i)-\frac{b-j}{b}. \qquad (15.63)$$

This implies that $T_{n,k}$ is independent of $T_{n,0}$ for $k>b$. Thus, if $\rho=1$ for all n, (15.30) holds immediately. This leads us to consider the case where $\rho<1$ for all n.

(Sequences in which $\rho = 1$ for some n and $\rho < 1$ for some n can be handled by analyzing subsequences.)

Define $S_{1,k}, n_\gamma, \xi_{1,k}, S_{2,k}$, and $\xi_{2,k}$ as in case (ii).

To show $\xi_{1,k} = o_p(1)$, by (15.48), it is enough to show that the expression in (15.58) is $o(1)$ because $\rho^{2(k-b)} = O(1)$ uniformly in $k \geq k_n = b + 1$.

Because $n(1-\rho) = h_n \to h < \infty$, it follows that $h_n = O(1)$. By an analysis as in (15.16), it follows that $h_n^*/h_n \to 1$ and, hence, $h_n^* = O(1)$ and $t_n = bh_n^*/n \to 0$. From (15.57) with $|h_n^{**}/n| \leq |h_n^*/n|$, we obtain $h_n/h_n^* = \exp(-h_n^{**}/n) = 1 + O(h_n^*/n)$. Given (15.59), to show the expression in (15.58) is $o(1)$, it is enough to show that

$$t_n^{-3}\left([1-\exp(-2t_n)]t_n O(h_n^*/n) + [1-\exp(-t_n)]^2 h_n/n\right) = o(1). \tag{15.64}$$

The latter holds l'Hopital's rule applied to $t_n^{-1}[1-\exp(-2t_n)]$ and to $t_n^{-2}[1-\exp(-t_n)]^2$ combined with $O(h_n^*/n)t_n^{-1} = o(1)$ and $(h_n/n)t_n^{-1} = o(1)$ because $b \to \infty$. Hence, $\xi_{1,k} = o_p(1)$ uniformly in $k \geq k_n$ for $k_n = b + 1$.

The proofs of $\xi_{m,k} = o_p(1)$ uniformly for $k \geq k_n$ for $m = 1, 2, 3$ are the same as in case (iii) given the proof above that the expression in (15.58) is $o(1)$ in case (iv).

This completes the verification of Assumption E for Model 1.

15.4 Verification of Assumption EE

In this section, we verify Assumption EE for model 1. We verify Assumption EE using the same argument as for Assumption E given above, but with $T_{n,k} = S_{1,k}S_{2,k}^{-1/2}S_{3,k}^{-1}$ replaced by $d_{b_n}(\gamma_{n,h})\widehat{\sigma}_{n,b_n,k}$, where $d_{b_n}(\gamma_{n,h})$ is the normalization constant that appears in Assumption BB2 and is defined in (15.1). Let $\rho_{n,h} = 1 - \gamma_{n,h}$. Suppose $b_n(1-\rho_{n,h}) \to \infty$, then $d_{b_n}(\gamma_{n,h}) = (1-\rho_{n,h}^2)^{-1/2}$ and some calculations show that $d_{b_n}(\gamma_{n,h})\widehat{\sigma}_{n,b_n,k} = S_{2,k}^{-1/2}S_{3,k}$, where the latter is defined in (15.37) with n_γ defined in (15.41). Similarly, if $b_n(1-\rho_{n,h}) \to g < \infty$, then $d_{b_n}(\gamma_{n,h}) = b_n^{1/2}$ and some calculations yield $d_{b_n}(\gamma_{n,h})\widehat{\sigma}_{n,b_n,k} = S_{2,k}^{-1/2}S_{3,k}$, where the latter is defined in (15.37) with $n_\gamma = b_n^{-1}$. Hence, in both cases, $d_{b_n}(\gamma_{n,h})\widehat{\sigma}_{n,b_n,k} = S_{2,k}^{-1/2}S_{3,k}$. In consequence, Assumption EE holds by the same argument as for Assumption E, but with the statistic $T_{n,k} = S_{1,k}S_{2,k}^{-1/2}S_{3,k}^{-1}$ replaced by $S_{2,k}^{-1/2}S_{3,k}$.

16 Conservative Model Selection Example

Here we establish the asymptotic distribution of the test statistic $T_n^*(\theta_0)$ and verify Assumption G for this example.

16.1 Proof of the Asymptotic Distributions of the Test Statistics

In this section, we establish the asymptotic distribution J_h^* of $T_n^*(\theta_0)$ under a sequence of parameters $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) : n \geq 1\}$ as in AG1 (where $n^{1/2}\gamma_{n,1} \to h_1$,

$\gamma_{n,2} \to h_2$, and $\gamma_{n,3} \in \Gamma_3(\gamma_{n,1}, \gamma_{n,2})$ for all n). Parts of the proof are closely related to calculations in Leeb (2006) and Leeb and Pötscher (2005).

Using the definition of $T_n^*(\theta_0)$ in this example, we have

$$P_{\theta_0,\gamma_n}(T_n^*(\theta_0) \le x) = P_{\theta_0,\gamma_n}(\widetilde{T}_{n,1}(\theta_0) \le x \ \&\ |T_{n,2}| \le c) + P_{\theta_0,\gamma_n}(\widehat{T}_{n,1}(\theta_0) \le x \ \&\ |T_{n,2}| > c). \tag{16.1}$$

Hence, it suffices to determine the limits of the two summands on the right-hand side. With this in mind, we show below that under $\{\gamma_n : n \ge 1\}$, when $|h_1| < \infty$,

$$\begin{pmatrix} \widetilde{T}_{n,1}(\theta_0) \\ T_{n,2} \end{pmatrix} \to_d \begin{pmatrix} \widetilde{Z}_{h,1} \\ Z_{h,2} \end{pmatrix} \sim N\left(\begin{pmatrix} -h_1 h_2 (1-h_2^2)^{-1/2} \\ h_1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \text{ and}$$
$$\begin{pmatrix} \widehat{T}_{n,1}(\theta_0) \\ T_{n,2} \end{pmatrix} \to_d \begin{pmatrix} \widehat{Z}_{h,1} \\ Z_{h,2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ h_1 \end{pmatrix}, \begin{pmatrix} 1 & h_2 \\ h_2 & 1 \end{pmatrix} \right). \tag{16.2}$$

Given this, we have

$$\begin{aligned} & P_{\theta_0,\gamma_n}(\widetilde{T}_{n,1}(\theta_0) \le x \ \&\ |T_{n,2}| \le c) \\ \to\ & P(\widetilde{Z}_{h,1} \le x \ \&\ |Z_{h,2}| \le c) \\ =\ & \Phi(x + h_1 h_2 (1-h_2^2)^{-1/2}) \Delta(h_1, c), \text{ where} \\ & \Delta(a,b) = \Phi(a+b) - \Phi(a-b), \end{aligned} \tag{16.3}$$

the equality uses the independence of $\widetilde{Z}_{h,1}$ and $Z_{h,2}$ and the normality of their distributions, and $\Delta(a,b) = \Delta(-a,b)$. In addition, we have

$$P_{\theta_0,\gamma_n}(\widehat{T}_{n,1}(\theta_0) \le x \ \&\ |T_{n,2}| > c) \to P(\widehat{Z}_{h,1} \le x \ \&\ |Z_{h,2}| > c). \tag{16.4}$$

Next, we calculate the limiting probability in (16.4). Let $f(z_2|z_1)$ denote the conditional density of $Z_{h,2}$ given $\widehat{Z}_{h,1}$. Let $\phi(z_1)$ denote the standard normal density. Given that

$$\begin{pmatrix} \widehat{Z}_{h,1} \\ Z_{h,2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ h_1 \end{pmatrix}, \begin{pmatrix} 1 & h_2 \\ h_2 & 1 \end{pmatrix} \right), \tag{16.5}$$

the conditional distribution of $Z_{h,2}$ given $\widehat{Z}_{h,1} = z_1$ is $N(h_1 + h_2 z_1, 1 - h_2^2)$. We have

$$\begin{aligned} & P(\widehat{Z}_{h,1} \le x \ \&\ |Z_{h,2}| > c) \\ =\ & \int_{-\infty}^{x} \int_{|z_2|>c} f(z_2|z_1)\phi(z_1) dz_2 dz_1 \\ =\ & \int_{-\infty}^{x} \left(1 - \int_{|z_2| \le c} (1-h_2^2)^{-1/2} \phi\left(\frac{z_2 - (h_1 + h_2 z_1)}{(1-h_2^2)^{1/2}} \right) dz_2 \right) \phi(z_1) dz_1 \\ =\ & \int_{-\infty}^{x} \left(1 - \int_{|\overline{z}_2| \le c(1-h_2^2)^{-1/2}} \phi\left(\overline{z}_2 - \frac{h_1 + h_2 z_1}{(1-h_2^2)^{1/2}} \right) d\overline{z}_2 \right) \phi(z_1) dz_1 \\ =\ & \int_{-\infty}^{x} \left(1 - \Delta\left(\frac{h_1 + h_2 z}{(1-h_2^2)^{1/2}}, \frac{c}{(1-h_2^2)^{1/2}} \right) \right) \phi(z) dz, \end{aligned} \tag{16.6}$$

where the second equality holds by (16.5), the third equality holds by change of variables with $\overline{z}_2 = z_2(1-h_2^2)^{-1/2}$, and the last equality holds by the definition of $\Delta(a,b)$.

Combining (16.3), (16.4), and (16.6) gives the desired result:

$$J_h^*(x) = \Phi(x + h_1 h_2 (1-h_2^2)^{-1/2})\Delta(h_1, c) + \int_{-\infty}^{x} \left(1 - \Delta\left(\frac{h_1 + h_2 t}{(1-h_2^2)^{1/2}}, \frac{c}{(1-h_2^2)^{1/2}}\right)\right)\phi(t)dt \quad (16.7)$$

when $|h_1| < \infty$. When $|h_1| = \infty$, $J_h^*(x) = \Phi(x)$ (which equals the limit as $|h_1| \to \infty$ of $J_h^*(x)$ defined in (16.7)). The proof of the latter result is given below in the paragraph containing (16.21).

We now show that under $\{\gamma_n : n \geq 1\}$, when $|h_1| < \infty$, (16.2) holds. Let $X_j^\perp = (x_{j1}^\perp, ..., x_{jn}^\perp)' \in R^n$ for $j = 1, 2$ and $X^\perp = (X_1^\perp, X_2^\perp)' \in R^{n\times 2}$. We use the following Lemma.

Lemma 6 *Given the assumptions stated in Section* 10*, under a sequence of parameters* $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) : n \geq 1\}$ *(where* $n^{1/2}\gamma_{n,1} \to h_1$, $\gamma_{n,2} \to h_2$*, and* $\gamma_{n,3} \in \Gamma_3(\gamma_1, \gamma_2)$ *for all* n*), and for* $Q = Q_n$ *as defined in* (10.10) *with the* (j,m) *element denoted* $Q_{n,jm}$*, we have*
(a) $n^{-1}X'X - Q_n \to_p 0$, (b) $n^{-1}X_2' M_{X_1} X_2 - (Q_{n,22} - Q_{n,12}^2 Q_{n,11}^{-1}) \to_p 0$, (c) $n^{-1}X_1' M_{X_2} X_1 - (Q_{n,11} - Q_{n,12}^2 Q_{n,22}^{-1}) \to_p 0$, (d) $\widehat{\sigma}/\sigma_n \to_p 1$, (e) $n^{-1/2}X_j'\varepsilon = n^{-1/2}X_j^{\perp\prime}\varepsilon + o_p(1) = O_p(1)$ *for* $j = 1, 2$.

Proof of Lemma 6. The proofs of parts (a)-(d) are standard using a weak law of large numbers (WLLN) for $L^{1+\delta}$-bounded independent random variables for some $\delta > 0$ and taking into account the fact that $X_j = M_{X_3^*} X_j^*$ for $j = 1, 2$.

Next, we prove part (e). By definition of X_j, we have

$$\begin{aligned} n^{-1/2}X_j'\varepsilon &= n^{-1/2}X_j^{*\prime}\varepsilon - n^{-1}X_j^{*\prime}X_3^*(n^{-1}X_3^{*\prime}X_3^*)^{-1}n^{-1/2}X_3^{*\prime}\varepsilon \\ &= n^{-1/2}X_j^{*\prime}\varepsilon - Q_{j3}^*(Q_{33}^*)^{-1}n^{-1/2}X_3^{*\prime}\varepsilon + o_p(1) \\ &= n^{-1/2}X_j^{\perp\prime}\varepsilon + o_p(1), \end{aligned} \quad (16.8)$$

where the second equality holds by the same WLLN as above combined with the Lindeberg triangular array central limit theorem (CLT) applied to $n^{-1/2}X_3^{*\prime}\varepsilon$, which yields $n^{-1/2}X_3^{*\prime}\varepsilon = O_p(1)$, and the third equality uses the fact that $x_{ji}^\perp = x_{ji}^* - Q_{j3}^*(Q_{33}^*)^{-1}x_{3i}^*$. The second equality of part (e) holds by the Lindeberg CLT. The Lindeberg condition is implied by a Liapounov condition, which holds by the moment bound in $\Gamma_3(\gamma_1, \gamma_2)$. □

We now prove the first result of (16.2) (which assumes $|h_1| < \infty$). Using (10.5) and (10.6), we have

$$T_{n,2} = \frac{n^{1/2}\beta_2/\sigma_n + (n^{-1}X_2'M_{X_1}X_2)^{-1}n^{-1/2}X_2'M_{X_1}\varepsilon}{(\widehat{\sigma}/\sigma_n)(n^{-1}X_2'M_{X_1}X_2)^{-1/2}}$$

$$
\begin{aligned}
&= n^{1/2}\frac{\beta_2}{\sigma_n (Q_n^{22})^{1/2}}(1+o_p(1)) + (Q_n^{22})^{1/2} n^{-1/2} X_2'(I_n - P_{X_1})\varepsilon(1+o_p(1)) \\
&= n^{1/2}\gamma_{n,1}(1+o_p(1)) + (Q_n^{22})^{1/2}(e_2 - Q_{n,12}Q_{n,11}^{-1}e_1)' n^{-1/2} X'\varepsilon(1+o_p(1)), \\
&= h_1 + (Q_n^{22})^{1/2}(e_2 - Q_{n,12}Q_{n,11}^{-1}e_1)' n^{-1/2} X^{\perp\prime}\varepsilon + o_p(1), \qquad (16.9)
\end{aligned}
$$

where $\varepsilon = (\varepsilon_1, ..., \varepsilon_n)'$, $e_1 = (1,0)'$, $e_2 = (0,1)'$, the second equality uses Lemma 6(b) and (d), the fact that $Q_n^{22} = (Q_{n,22} - Q_{n,12}^2 Q_{n,11}^{-1})^{-1}$, and the fact that $\lambda_{\min}(Q_n) \geq \kappa > 0$ by definition of $\Gamma_3(\gamma_1, \gamma_2)$, the third equality uses the definition of $\gamma_{n,1}$ and Lemma 6(a), and the fourth equality holds by the assumption that $n^{1/2}\gamma_{n,1} \to h_1$ and Lemma 6(e).

Using (10.5) and (10.7), we have

$$
\begin{aligned}
\widetilde{T}_{n,1}(\theta_0) &= \frac{n^{1/2}(n^{-1}X_1'X_1)^{-1} n^{-1} X_1' X_2 \beta_2/\sigma_n + (n^{-1}X_1'X_1)^{-1} n^{-1/2} X_1'\varepsilon}{(\widehat{\sigma}/\sigma_n)(n^{-1}X_1'X_1)^{-1/2}} \\
&= n^{1/2}\frac{Q_{n,12}\beta_2}{\sigma_n Q_{n,11}^{1/2}}(1+o_p(1)) + Q_{n,11}^{-1/2} n^{-1/2} e_1' X'\varepsilon(1+o_p(1)) \\
&= h_1 \frac{Q_{n,12}(Q_n^{22})^{1/2}}{Q_{n,11}^{1/2}} + Q_{n,11}^{-1/2} n^{-1/2} e_1' X^{\perp\prime}\varepsilon + o_p(1), \qquad (16.10)
\end{aligned}
$$

where the second equality uses Lemma 6(a) and (d) and the third equality uses the assumption that $n^{1/2}\gamma_{n,1} = n^{1/2}\beta_2/(\sigma_n^2 Q_n^{22})^{1/2} \to h_1$, and Lemma 6(e).

We have

$$
\begin{aligned}
Q_n^{-1} &= \frac{1}{Q_{n,11}Q_{n,22} - Q_{n,12}^2}\begin{bmatrix} Q_{n,22} & -Q_{n,12} \\ -Q_{n,12} & Q_{n,11} \end{bmatrix} \text{ and so} \\
\gamma_{n,2} &= \frac{Q_n^{12}}{(Q_n^{11}Q_n^{22})^{1/2}} = \frac{-Q_{n,12}}{(Q_{n,11}Q_{n,22})^{1/2}} \text{ and} \\
Q_n^{22} &= \frac{Q_{n,11}}{Q_{n,11}Q_{n,22} - Q_{n,12}^2} = (Q_{n,22})^{-1}(1-\gamma_{n,2}^2)^{-1}, \qquad (16.11)
\end{aligned}
$$

where the first equality in the second line holds by the definition of $\gamma_{n,2}$ in (10.11). This yields

$$
\frac{Q_{n,12}(Q_n^{22})^{1/2}}{Q_{n,11}^{1/2}} = \frac{Q_{n,12}(1-\gamma_{n,2}^2)^{-1/2}}{Q_{n,11}^{1/2}Q_{n,22}^{1/2}} = -\gamma_{n,2}(1-\gamma_{n,2}^2)^{-1/2} = -h_2(1-h_2^2)^{-1/2} + o(1). \qquad (16.12)
$$

Combining (16.9), (16.10), and (16.12) gives

$$
\begin{pmatrix} \widetilde{T}_{n,1}(\theta_0) \\ T_{n,2} \end{pmatrix} = \begin{pmatrix} -h_1 h_2 (1-h_2^2)^{-1/2} + Q_{n,11}^{-1/2} n^{-1/2} e_1' X^{\perp\prime}\varepsilon \\ h_1 + (Q_n^{22})^{1/2}(e_2 - Q_{n,12}Q_{n,11}^{-1}e_1)' n^{-1/2} X^{\perp\prime}\varepsilon \end{pmatrix} + o_p(1). \qquad (16.13)
$$

The first result of (16.2) holds by (16.13), the Lindeberg CLT, and the Cramér-Wold device. The Lindeberg condition is implied by a Liapounov condition, which holds by the moment bound in $\Gamma_3(\gamma_1, \gamma_2)$. The asymptotic covariance matrix is I_2 by the following calculations. The $(1,2)$ element of the asymptotic covariance matrix equals

$$\begin{aligned} & E_{G_n} Q_{n,11}^{-1/2} e_1' n^{-1} X^{\perp\prime} X^{\perp} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)(Q_n^{22})^{1/2} \\ = \; & Q_{n,11}^{-1/2} e_1' Q_n (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)(Q_n^{22})^{1/2} = 0, \end{aligned} \tag{16.14}$$

where the first equality holds because $E_{G_n} x_i^{\perp} x_i^{\perp\prime} = Q_n$ and the second equality holds by algebra. The $(1,1)$ element equals

$$E_{G_n} Q_{n,11}^{-1/2} e_1' n^{-1} X^{\perp\prime} X^{\perp} e_1 Q_{n,11}^{-1/2} = Q_{n,11}^{-1/2} e_1' Q_n e_1 Q_{n,11}^{-1/2} = 1. \tag{16.15}$$

The $(2,2)$ element equals

$$\begin{aligned} & E_{G_n} (Q_n^{22})^{1/2} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)' n^{-1} X^{\perp\prime} X^{\perp} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)(Q_n^{22})^{1/2} \\ = \; & (Q_n^{22})^{1/2} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)' Q_n (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)(Q_n^{22})^{1/2} \\ = \; & (Q_n^{22})^{1/2} (Q_{n,22}(1 - \gamma_{n,2}^2))(Q_n^{22})^{1/2} = 1, \end{aligned} \tag{16.16}$$

where the second equality holds by algebra and the definition of $\gamma_{n,2}$ and the third equality holds by the third result in (16.11). This completes the proof of the first result in (16.2).

Next, we prove the second result in (16.2). Using (10.7), we have

$$\begin{aligned} \widehat{T}_{n,1}(\theta_0) = \; & \frac{(n^{-1} X_1' M_{X_2} X_1)^{-1} n^{-1/2} X_1' M_{X_2} \varepsilon}{(\widehat{\sigma}/\sigma_n)(n^{-1} X_1' M_{X_2} X_1)^{-1/2}} \\ = \; & (Q_n^{11})^{1/2} (e_1 - Q_{n,12} Q_{n,22}^{-1} e_2)' n^{-1/2} X^{\perp\prime} \varepsilon + o_p(1), \end{aligned} \tag{16.17}$$

where the second equality holds analogously to (16.9). Combining (16.9) and (16.17) gives

$$\begin{pmatrix} \widehat{T}_{n,1}(\theta_0) \\ T_{n,2} \end{pmatrix} = \begin{pmatrix} (Q_n^{11})^{1/2} (e_1 - Q_{n,12} Q_{n,22}^{-1} e_2)' n^{-1/2} X^{\perp\prime} \varepsilon \\ h_1 + (Q_n^{22})^{1/2} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)' n^{-1/2} X^{\perp\prime} \varepsilon \end{pmatrix} + o_p(1). \tag{16.18}$$

The second result of (16.2) holds by (16.18), the Lindeberg CLT, and the Cramér-Wold device. The Lindeberg condition holds as above. The 2×2 asymptotic covariance matrix has off-diagonal element h_2 and diagonal elements equal to one by the following calculations. The $(1,2)$ element equals

$$\begin{aligned} & E_{G_n} (Q_n^{11})^{1/2} (e_1 - Q_{n,12} Q_{n,22}^{-1} e_2)' n^{-1} X^{\perp\prime} X^{\perp} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)(Q_n^{22})^{1/2} \\ = \; & (Q_n^{11})^{1/2} (e_1 - Q_{n,12} Q_{n,22}^{-1} e_2)' Q_n (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)(Q_n^{22})^{1/2} \\ = \; & (Q_n^{11})^{1/2} (-Q_{n,12}(1 - Q_{n,12}^2 Q_{n,11}^{-1} Q_{n,22}^{-1}))(Q_n^{22})^{1/2} \\ = \; & (Q_{n,11}(1 - \gamma_{n,2}^2))^{-1/2} (-Q_{n,12}(1 - \gamma_{n,2}^2))(Q_{n,22}(1 - \gamma_{n,2}^2))^{-1/2} \\ = \; & \frac{-Q_{n,12}}{(Q_{n,11} Q_{n,22})^{1/2}} = \frac{Q_n^{12}}{(Q_n^{11} Q_n^{22})^{1/2}} = \gamma_{n,2} = h_2 + o(1), \end{aligned} \tag{16.19}$$

where the second equality holds by algebra, the third equality holds by the second and third results of (16.11) and the third result of (16.11) with 22 and 11 interchanged, and the fifth and sixth equalities hold by the second result of (16.11).

The $(1,1)$ element equals

$$\begin{aligned} & E_{G_n}(Q_n^{11})^{1/2}(e_1 - Q_{n,12}Q_{n,22}^{-1}e_2)'n^{-1}X^{\perp\prime}X^{\perp}(e_1 - Q_{n,12}Q_{n,22}^{-1}e_2)(Q_n^{11})^{1/2} \\ &= (Q_n^{11})^{1/2}(e_1 - Q_{n,12}Q_{n,22}^{-1}e_2)'Q_n(e_1 - Q_{n,12}Q_{n,22}^{-1}e_2)(Q_n^{11})^{1/2} = 1, \end{aligned} \tag{16.20}$$

where the second equality holds by an analogous argument to that in (16.16). The $(2,2)$ element equals one by (16.16). This completes the proof of the second result in (16.2).

Finally, we show that $J_h^*(x) = \Phi(x)$ when $|h_1| = \infty$. Equations (16.17) and (16.20) hold in this case, so $\widehat{T}_{n,1}(\theta_0) \to_d N(0,1)$ under $\{\gamma_n : n \geq 1\}$. The first three equalities of (16.9) hold when $|h_1| = \infty$ and show that $|T_{n,2}| \to_p \infty$. These results combine to yield

$$\begin{aligned} & P_{\theta_0,\gamma_n}(\widetilde{T}_{n,1}(\theta_0) \leq x \ \& \ |T_{n,2}| \leq c) = o(1) \text{ and} \\ & P_{\theta_0,\gamma_n}(\widehat{T}_{n,1}(\theta_0) \leq x \ \& \ |T_{n,2}| > c) = P_{\theta_0,\gamma_n}(\widehat{T}_{n,1}(\theta_0) \leq x) + o(1) \to \Phi(x) \end{aligned} \tag{16.21}$$

for all $x \in R$. This and (16.1) combine to give $P_{\theta_0,\gamma_n}(T_n^*(\theta_0) \leq x) \to \Phi(x)$ and $J_h^*(x) = \Phi(x)$ when $|h_1| = \infty$.

16.2 Verification of Assumption G2

Assumption G2 of AG1 is verified in the conservative model selection example by using a variant of the argument in the proof of Lemma 2 in Appendix A of AG1 with $\tau_n = a_n = n^{1/2}$ and $d_n = 1$. In the present case, (10.24) of AG1 holds with

$$\begin{aligned} R_n(t) &= q_n^{-1}\sum_{j=1}^{q_n} 1(|b_n^{1/2}(\overline{\theta} - \theta_0)/\widehat{\sigma}_{n,b,j}^{(1)}| \geq t) \\ &\quad + q_n^{-1}\sum_{j=1}^{q_n} 1(|b_n^{1/2}(\overline{\theta} - \theta_0)/\widehat{\sigma}_{n,b,j}^{(2)}| \geq t), \text{ where} \\ \widehat{\sigma}_{n,b,j}^{(1)} &= \widehat{\sigma}_{n,b,j}(b_n^{-1}X_{1,n,b,j}'X_{1,n,b,j})^{-1/2}, \\ \widehat{\sigma}_{n,b,j}^{(2)} &= \widehat{\sigma}_{n,b,j}(b_n^{-1}X_{1,n,b,j}'M_{X_{2,n,b,j}}X_{1,n,b,j})^{-1/2}, \end{aligned} \tag{16.22}$$

and $(X_{1,n,b,j},\ X_{2,n,b,j}, \widehat{\sigma}_{n,b,j})$ denotes $(X_1, X_2, \widehat{\sigma})$ based on the jth subsample rather than the full sample. (Equation (10.24) of AG1 holds with $R_n(t)$ defined as in (16.22) for all three versions of the tests: $T_n(\theta_0) = T_n^*(\theta_0)$, $-T_n^*(\theta_0)$, and $|T_n^*(\theta_0)|$.) As in the proof of Lemma 2 of AG1, it suffices to show that $R_n(t)$ converges in probability to zero under all sequences $\{\gamma_{n,h} : n \geq 1\}$ for all $t > 0$. The assumption that $b_n/n \to 0$

and the result established below that $n^{1/2}(\overline{\theta} - \theta_0)/\sigma_n = O_p(1)$ under all sequences $\{\gamma_{n,h} : n \geq 1\}$ imply that for all $\delta > 0$, wp$\rightarrow$ 1,

$$R_n(t) \leq R_n^{(1)}(\delta, t) + R_n^{(2)}(\delta, t), \text{ where } R_n^{(m)}(\delta, t) = q_n^{-1} \sum_{j=1}^{q_n} 1(\delta\sigma_n/\widehat{\sigma}_{n,b,j}^{(m)} \geq t) \quad (16.23)$$

for $m = 1, 2$. The variance of $R_n^{(m)}(\delta, t)$ goes to zero under $\{\gamma_{n,h} : n \geq 1\}$ by the same U-statistic argument for i.i.d. observations as used to establish Assumption E of AG1 in the i.i.d. case, see AG1. The expectation of $R_n^{(m)}(\delta, t)$ equals $P_{\theta_0,\gamma_{n,h}}(\widehat{\sigma}_{n,b,j}^{(m)}/\sigma_n \leq \delta/t)$. We have

$$\widehat{\sigma}_{n,b,j}^{(1)}/\sigma_n = (\widehat{\sigma}_{n,b,j}/\sigma_n)[(b_n^{-1}X_{1,n,b,j}'X_{1,n,b,j})^{-1/2} - Q_{n,11}^{-1/2} + Q_{n,11}^{-1/2}] = Q_{n,11}^{-1/2} + o_p(1), \quad (16.24)$$

where the second equality holds by Lemma 6 (or, more precisely, by the same argument as used to prove to Lemma 6). In addition, $Q_{n,11}^{-1/2}$ is bounded away from zero as $n \rightarrow \infty$ by the definition of $\Gamma_3(\gamma_1, \gamma_2)$. In consequence, the expectation of $R_n^{(1)}(\delta, t)$ goes to zero for all δ sufficiently small. Since the mean and variance of $R_n^{(1)}(\delta, t)$ go to zero, $R_n^{(1)}(\delta, t) \rightarrow_p 0$ for $\delta > 0$ sufficient small. An analogous argument shows that $R_n^{(2)}(\delta, t) \rightarrow_p 0$ for $\delta > 0$ sufficient small. These results and (16.23) yield $R_n(t) \rightarrow_p 0$ under all sequences $\{\gamma_{n,h} : n \geq 1\}$, as desired.

It remains to show that $n^{1/2}(\overline{\theta} - \theta_0)/\sigma_n = O_p(1)$ under all sequences $\{\gamma_{n,h} : n \geq 1\}$. We consider two cases: $|h_1| = \infty$ and $|h_1| < \infty$. First, suppose $|h_1| = \infty$. Then, the first three equalities of (16.9) hold and show that $|T_{n,2}| \rightarrow_p \infty$. In addition, $n^{1/2}(\widehat{\theta} - \theta_0)/\sigma_n = (\widehat{\sigma}/\sigma_n)(n^{-1}X_1'M_{X_2}X_1)^{-1/2}\widehat{T}_{n,1}(\theta_0) = O_p(1)$ by (16.2), Lemma 6(c), and the definition of $\Gamma_3(\gamma_1, \gamma_2)$. Combining these results gives: when $|h_1| = \infty$,

$$\begin{aligned} n^{1/2}(\overline{\theta} - \theta_0)/\sigma_n &= [n^{1/2}(\widetilde{\theta} - \theta_0)/\sigma_n]1(|T_{n,2}| \leq c) + [n^{1/2}(\widehat{\theta} - \theta_0)/\sigma_n]1(|T_{n,2}| > c)) \\ &= o_p(1) + O_p(1). \end{aligned} \quad (16.25)$$

Next, suppose $|h_1| < \infty$, then $\widehat{T}_{n,1}(\theta_0) = O_p(1)$ and $\widetilde{T}_{n,1}(\theta_0) = O_p(1)$ by (16.2). In addition, $\widehat{\sigma}/\sigma_n \rightarrow_p 1$, $(n^{-1}X_1'X_1)^{-1/2} = O_p(1)$, and $(n^{-1}X_1'M_{X_2}X_1)^{-1/2} = O_p(1)$ by Lemma 6 and the definition of $\Gamma_3(\gamma_1, \gamma_2)$. Combining these results gives: when $|h_1| < \infty$,

$$\begin{aligned} n^{1/2}(\overline{\theta} - \theta_0)/\sigma_n &= [n^{1/2}(\widetilde{\theta} - \theta_0)/\sigma_n]1(|T_{n,2}| \leq c) + [n^{1/2}(\widehat{\theta} - \theta_0)/\sigma_n]1(|T_{n,2}| > c)) \\ &= (\widehat{\sigma}/\sigma_n)(n^{-1}X_1'X_1)^{-1/2}\widetilde{T}_{n,1}(\theta_0)1(|T_{n,2}| \leq c) \\ &\quad +(\widehat{\sigma}/\sigma_n)(n^{-1}X_1'M_{X_2}X_1)^{-1/2}\widehat{T}_{n,1}(\theta_0)1(|T_{n,2}| > c)) \\ &= O_p(1), \end{aligned} \quad (16.26)$$

which completes the verification of Assumption G2.

References

Anderson, T. W. and H. Rubin (1949): "Estimators of the Parameters of a Single Equation in a Complete Set of Stochastic Equations," *Annals of Mathematical Statistics,* 21, 570-582.

Andrews, D. W. K. (1993): "Exactly Median-Unbiased Estimation of First-Order Autoregressive/Unit Root Models," *Econometrica,* 61, 139-165.

——— (1999): "Estimation When a Parameter Is on a Boundary," *Econometrica,* 67, 1341–1383.

——— (2001): "Testing When a Parameter Is on the Boundary of the Maintained Hypothesis," *Econometrica,* 69, 683–734.

Andrews, D. W. K. and P. Guggenberger (2005a): "The Limit of Finite-Sample Size and a Problem with Subsampling," unpublished manuscript, Cowles Foundation, Yale University.

——— (2005b): "Applications of Hybrid and Size-Corrected Subsample Methods," unpublished manuscript, Cowles Foundation, Yale University.

——— (2006): "Asymptotics for Stationary Very Nearly Unit Root Processes," unpublished manuscript, Cowles Foundation, Yale University.

Bobkowski, M. J. (1983): "Hypothesis Testing in Nonstationary Time Series," unpublished Ph. D. Thesis, University of Wisconsin, Madison.

Cavanagh, C. (1985): "Roots Local to Unity," unpublished manuscript, Harvard University.

Chan, N. H. and C. Z. Wei (1987): "Asymptotic Inference for Nearly Nonstationary AR(1) Processes," *Annals of Statistics,* 15, 1050-1063.

Choi, I. (2005): "Subsampling Vector Autoregressive Tests of Linear Constraints," *Journal of Econometrics,* 124, 55-89.

De Jong, R. M. and J. Davidson (2000): "The Functional Central Limit Theorem and Weak Convergence to Stochastic Integrals I," *Econometric Theory,* 16, 621-642.

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

Elliott, G. (1999): "Efficient Tests for a Unit Root When the Initial Observation is Drawn from Its Unconditional Distribution," *International Economic Review,* 40, 767-783.

Elliott, G. and J. H. Stock (2001): "Confidence Intervals for Autoregressive Coefficients Near One," *Journal of Econometrics,* 103, 155-181.

Giraitis, L. and P. C. B. Phillips (2006): "Uniform Limit Theory for Stationary Autoregression," *Journal of Time Series Analysis,* 27, 51-60. Corrigendum, forthcoming.

Guggenberger, P. and R. J. Smith (2005): "Generalized Empirical Likelihood Estimators and Tests under Partial, Weak, and Strong Identification," *Econometric Theory,* 21, 667-709.

Hall, P. and C. C. Heyde (1980): *Martingale Limit Theory and Its Application.* New York: Academic Press.

Hansen, B. E. (1999): "The Grid Bootstrap and the Autoregressive Model," *Review of Economic Studies,* 81, 594-607.

Horowitz, J. L. (1997): "Bootstrap Methods in Econometrics: Theory and Methods," in *Advances in Econometrics: Theory and Applications,* ed. by D. M. Kreps and K. F. Wallis. Cambridge, UK: Cambridge University Press.

Jansson, M. and M. J. Moreira (2006): "Optimal Inference in Regression Models with Nearly Integrated Regressors," *Econometrica,* 74, 681-714.

Kabaila, P. (1998): "Valid Confidence Intervals in Regression after Variable Selection," *Econometric Theory,* 14, 463-482.

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

Leeb, H. (2006): "The Distribution of a Linear Predictor After Model Selection: Unconditional Finite-Sample Distributions and Asymptotic Approximations," *2nd Lehmann Symposium—Optimality.* Institute of Mathematical Statistics Lecture Notes—Monograph Series.

Leeb, H. and B. M. Pötscher (2005): "Model Selection and Inference: Facts and Fiction," *Econometric Theory,* 21, 21-59.

Mikusheva, A. (2005): "Uniform Inferences in Autoregressive Models," unpublished working paper, Department of Economics, Harvard University.

Moreira, M. J. (2001): "Tests with Correct Size When Instruments Can Be Arbitrarily Weak," unpublished manuscript, Dept. of Economics, University of California, Berkeley.

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

Müller, U. K. and G. Elliott (2003): "Tests for Unit Roots and the Initial Condition," *Econometrica,* 71, 1269-1286.

Otsu, T. (2006): "Generalized Empirical Likelihood Inference for Nonlinear and Time Series Models under Weak Identification," *Econometric Theory,* 22, 513-527.

Park, J. Y. (2002): "Weak Unit Roots," unpublished working paper, Department of Economics, Rice University.

Phillips, P. C. B. (1987): "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika,* 74, 535-547.

Phillips, P. C. B. and Magdalinos, T. (2007): "Limit Theory for Moderate Deviations from a Unit Root. *Journal of Econometrics* 136, 115-130.

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.

Staiger, D. and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica,* 65, 557-586.

Stock, J. H. (1991): "Confidence Intervals for the Largest Autoregressive Root in U. S. Macroeconomic Time Series," *Journal of Monetary Economics,* 28, 435-459.

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

TABLE I

NUISANCE PARAMETER NEAR A BOUNDARY EXAMPLE: MAXIMUM (OVER h_1) NULL REJECTION PROBABILITIES ($\times$100) FOR DIFFERENT VALUES OF THE CORRELATION h_2 FOR VARIOUS NOMINAL 5% TESTS FOR n = 120 AND b = 12, WHERE THE PROBABILITIES ARE ASYMPTOTIC, FINITE-SAMPLE-ADJUSTED ASYMPTOTIC, AND EXACT

	Test:	Sub	Sub	Sub	PSC-Sub	APSC-Sub	FCV	FCV	PSC-FCV	Hyb	Hyb	Hyb
h_2	Prob:	Asy	Adj-Asy	n=120	n=120	n=120	Asy	n=120	n=120	Asy	Adj-Asy	n=120
(a) Upper 1-Sided Tests												
-1.0		50.2	49.5	49.8	4.9	13.5	5.0	5.2	5.1	5.0	5.0	5.2
-.95		33.8	22.9	25.6	5.1	9.0	5.0	5.2	5.1	5.0	5.0	5.2
-.80		20.2	12.1	13.1	3.1	6.2	5.0	5.1	4.9	5.0	5.0	4.7
-.40		8.3	6.5	5.9	4.8	4.6	5.0	4.9	4.8	5.0	5.0	3.7
.00		5.0	5.0	5.0	4.9	4.9	5.0	5.2	5.0	5.0	5.0	3.7
.20		5.0	5.0	4.9	5.2	4.9	5.6	5.7	5.2	5.0	5.0	3.8
.40		5.0	5.0	5.0	5.0	5.0	5.8	5.8	5.0	5.0	5.0	3.8
.60		5.0	5.0	5.3	5.3	5.3	5.6	5.7	5.1	5.0	5.0	3.9
.90		5.0	5.0	4.9	4.9	4.9	5.0	5.0	4.9	5.0	5.0	3.4
1.00		5.0	5.0	4.8	4.9	4.8	5.0	5.0	4.9	5.0	5.0	3.5
Max		50.2	49.5	49.8	5.3	13.5	5.8	5.8	5.2	5.0	5.0	5.2
(b) Symmetric 2-Sided Tests												
.00		5.0	5.0	5.2	5.1	5.1	5.0	5.4	-	5.0	5.0	3.5
.20		5.2	5.2	5.2	4.9	5.0	5.0	5.3	-	5.0	5.0	3.5
.40		6.0	5.6	5.4	4.5	4.8	5.0	5.2	-	5.0	5.0	3.5
.60		7.5	6.5	6.0	4.0	4.6	5.0	5.3	-	5.0	5.0	3.7
.80		9.6	8.3	6.9	3.7	4.5	5.0	5.2	-	5.0	5.0	3.9
.95		10.1	10.0	8.3	4.2	4.2	5.0	5.7	-	5.0	5.0	4.5
1.00		10.1	10.1	8.4	4.1	4.1	5.0	5.1	-	5.0	5.0	4.2
Max		10.1	10.1	8.4	5.1	5.1	5.0	5.7	-	5.0	5.0	4.5
(c) Equal-Tailed 2-Sided Tests												
.00		5.0	5.0	5.7	5.5	5.5	5.0	5.4	-	5.0	5.0	3.5
.20		5.4	5.2	5.9	5.4	5.6	5.0	5.3	-	5.0	5.0	3.6
.40		6.7	5.8	6.2	4.5	5.4	5.0	5.2	-	5.0	5.0	3.4
.60		9.9	7.0	7.8	3.9	5.7	5.0	5.3	-	5.0	5.0	3.8
.80		17.3	10.3	12.4	3.2	6.5	5.0	5.2	-	5.0	5.0	4.1
.95		32.4	21.0	24.3	3.5	9.1	5.0	5.7	-	5.0	5.0	4.7
1.00		52.7	51.8	52.7	4.6	13.5	5.0	5.1	-	5.0	5.0	4.2
Max		52.7	51.8	52.7	5.5	13.5	5.0	5.7	-	5.0	5.0	4.7

TABLE II

AR(1) EXAMPLE MODEL 1: CONFIDENCE INTERVAL COVERAGE PROBABILITIES ($\times 100$) FOR DIFFERENT VALUES OF THE AR(1) PARAMETER ρ FOR VARIOUS NOMINAL 95% CIs, WHERE THE PROBABILITIES ARE ASYMPTOTIC, FINITE-SAMPLE-ADJUSTED ASYMPTOTIC, AND FINITE SAMPLE FOR $n = 130$ AND $b = 12$

	Test:	Sub	Sub	Sub	SC-Sub	ASC-Sub	FCV	FCV	SC-FCV	Hyb	Hyb
ρ	Prob:	Asy	Adj-Asy	n=130	n=130	n=130	Asy	n=130	n=130	Asy	n=130
(a) Upper 1-Sided CIs											
-.90				92.0	99.8	98.5		93.6			94.8
.00				88.1	99.7	97.3		95.7			95.8
.70				82.3	99.4	95.7		97.6			97.6
.80				81.9	99.4	95.4		98.1			98.1
.90				83.2	99.5	96.0		98.9			98.9
.97				90.0	99.8	98.2		99.7			99.7
1.00				96.8	100.0	99.7		100.0			100.0
Min		47.5	78.7	81.9	99.4	95.3	95.0	93.6	-	95.0	94.8
(b) Lower 1-Sided CIs											
-.90				94.2				96.6	99.9		97.1
.00				94.0				94.1	99.7		95.6
.70				96.5				91.4	99.5		96.6
.80				96.7				90.0	99.5		96.8
.90				97.0				87.4	99.2		97.0
.97				96.1				78.9	98.3		96.1
1.00				92.7				54.2	95.2		92.7
Min		95.0	95.0	92.1	-	-	54.6	54.2	95.2	95.0	92.7
(c) Symmetric 2-Sided CIs											
-.90				93.0				94.9	99.6		95.9
-.50				91.4				94.9	99.7		95.7
.00				92.6				94.9	99.6		96.0
.70				96.3				94.3	99.5		97.0
.80				96.7				93.8	99.4		97.1
.90				97.0				92.2	99.2		97.1
.97				96.2				87.0	98.3		96.2
1.00				92.7				69.7	95.2		92.7
Min		94.8	95.0	91.3	-	-	68.9	69.6	95.2	95.0	92.7
(d) Equal-Tailed 2-Sided CIs											
-.90				92.1	99.7	97.4		94.9	99.6		96.0
-.50				89.1	99.5	96.2		94.9	99.7		95.6
.00				88.0	99.4	95.6		94.9	99.6		95.8
.70				86.8	99.3	95.0		94.3	99.5		97.1
.80				86.9	99.3	94.9		93.8	99.4		97.5
.90				88.1	99.5	95.6		92.2	99.2		97.9
.97				92.4	99.7	97.6		87.0	98.3		97.9
1.00				94.9	99.8	98.5		69.7	95.2		96.3
Min		60.1	86.1	86.7	99.3	94.9	69.7	69.6	95.2	95.0	95.6

TABLE III

CONSERVATIVE MODEL SELECTION EXAMPLE: MAXIMUM (OVER h_1) NULL REJECTION PROBABILITIES ($\times 100$) FOR DIFFERENT VALUES OF THE CORRELATION h_2 FOR VARIOUS NOMINAL 5% TESTS, WHERE THE PROBABILITIES ARE ASYMPTOTIC, FINITE-SAMPLE-ADJUSTED ASYMPTOTIC, AND FINITE SAMPLE FOR $n = 120$ AND $b = 12$ AND THE PARAMETER SPACE FOR h_2 IS $[-.995, .995]$

	Test:	Sub	Sub	Sub	PSC-Sub	APSC-Sub	FCV	FCV	PSC-FCV	Hyb	Hyb	Hyb	PSC-Hyb	APSC-Hyb
$\lvert h_2 \rvert$	Prob:	Asy	Adj-Asy	n=120	n=120	n=120	Asy	n=120	n=120	Asy	Adj-Asy	n=120	n=120	n=120
						(a) Upper 1-Sided Tests								
.00		5.1	5.1	5.4	4.7	5.0	5.3	5.4	4.7	5.1	5.1	3.7	3.3	3.3
.20		6.9	6.1	7.2	5.1	5.9	7.1	7.5	5.1	6.9	5.9	5.3	4.0	4.0
.40		11.2	8.2	11.0	5.1	7.2	11.8	11.9	5.1	11.2	8.2	8.7	4.5	4.5
.60		20.2	13.3	19.8	5.1	9.4	21.8	22.0	4.9	20.2	13.3	17.3	4.8	4.8
.80		41.3	28.3	38.9	5.2	13.1	44.3	43.8	4.8	41.3	28.3	37.2	4.8	4.8
.90		61.3	47.1	57.5	4.8	16.3	63.9	62.8	4.6	61.3	47.1	56.8	4.6	4.6
.95		75.5	64.4	72.2	4.7	19.5	77.2	76.7	4.6	75.5	64.4	71.8	4.6	4.6
.995		92.9	90.1	91.9	4.2	24.8	93.2	93.1	4.1	92.9	90.1	91.9	4.1	4.1
Max		92.9	90.1	91.9	5.2	24.8	93.2	93.1	5.1	92.9	90.1	91.9	4.8	4.8
						(b) Symmetric 2-Sided Tests								
.00		5.1	5.1	5.0	4.8	4.8	5.4	5.5	5.0	5.1	5.1	3.3	3.1	3.1
.20		6.0	5.7	5.3	4.4	4.6	6.3	6.5	5.1	6.0	5.7	3.8	3.3	3.3
.40		8.7	7.8	7.3	4.2	4.7	9.6	10.1	5.2	8.7	7.8	5.9	4.0	4.0
.60		16.1	12.9	12.3	4.0	5.4	18.2	18.8	5.3	16.1	12.9	11.3	4.8	4.8
.80		36.2	27.7	28.2	3.7	7.7	40.6	40.3	4.9	36.2	27.7	27.8	4.8	4.8
.90		57.6	47.2	48.5	3.3	10.6	62.0	61.5	4.5	57.6	47.2	48.3	4.5	4.5
.95		73.4	64.3	66.1	3.3	13.6	77.1	76.4	4.2	73.4	64.3	66.0	4.2	4.2
.995		93.9	90.1	90.7	3.8	22.1	95.5	95.3	4.2	93.9	90.1	90.7	4.2	4.2
Max		93.9	90.1	90.7	4.8	22.1	95.5	95.3	5.3	93.9	90.1	90.7	4.8	4.8
						(c) Equal-Tailed 2-Sided Tests								
.00		5.1	5.1	6.0	5.6	5.8	5.4	5.5	5.0	5.1	5.1	3.4	3.1	3.2
.20		6.0	5.6	6.8	5.8	6.0	6.3	6.5	5.0	6.0	5.4	4.0	3.4	3.5
.40		8.7	7.4	10.1	6.2	7.4	9.6	10.1	5.2	8.7	6.9	6.7	4.1	4.1
.60		16.1	11.6	17.5	5.8	9.4	18.2	18.8	5.2	16.1	11.1	13.3	4.8	4.8
.80		36.4	24.5	35.8	5.3	12.4	40.6	40.3	4.8	36.4	24.1	32.0	4.8	4.8
.90		57.6	43.7	55.0	4.6	15.7	62.0	61.5	4.5	57.6	43.6	52.2	4.5	4.5
.95		73.4	62.3	71.0	4.2	18.2	77.1	76.4	4.2	73.4	62.2	68.9	4.2	4.2
.995		93.9	91.8	93.8	4.1	24.1	95.5	95.3	4.2	93.9	91.8	92.6	4.4	4.4
Max		93.9	91.8	93.8	6.2	24.1	95.5	95.3	5.2	93.9	91.8	92.6	4.8	4.8

TABLE B-I

AR(1) EXAMPLE MODEL 2: CONFIDENCE INTERVAL COVERAGE PROBABILITIES ($\times 100$) FOR DIFFERENT VALUES OF THE AR(1) PARAMETER ρ FOR VARIOUS NOMINAL 95% CIs, WHERE THE PROBABILITIES ARE ASYMPTOTIC, FINITE-SAMPLE-ADJUSTED ASYMPTOTIC, AND FINITE SAMPLE FOR $n = 130$ AND $b = 12$

	Test:	Sub	Sub	Sub	SC-Sub	ASC-Sub	FCV	FCV	SC-FCV	Hyb	Hyb
ρ	Prob:	Asy	Adj-Asy	n=130	n=130	n=130	Asy	n=130	n=130	Asy	n=130
(a) Upper 1-Sided CIs											
-.90				90.4	100.0	99.3		94.0			94.6
.00				83.1	100.0	98.2		96.5			96.5
.70				72.4	99.8	95.9		98.6			98.6
.80				71.6	99.8	95.6		99.0			99.0
.90				74.5	99.9	96.6		99.6			99.6
.97				87.2	100.0	99.1		100.0			100.0
1.00				97.5	100.0	99.9		100.0			100.0
Min		17.7	63.3	71.6	99.8	95.6	95.0	94.0	-	95.0	94.6
(b) Lower 1-Sided CIs											
-.90				94.5				96.4	100.0		97.0
.00				95.1				93.1	99.9		95.8
.70				97.7				87.9	99.9		97.7
.80				97.9				84.9	99.8		97.9
.90				97.9				78.1	99.5		97.9
.97				96.7				57.4	98.5		96.7
1.00				92.4				23.4	94.6		92.4
Min		95.0	95.0	92.4	-	-	23.1	23.4	94.6	95.0	92.4
(c) Symmetric 2-Sided CIs											
-.90				92.5				95.1	99.9		95.8
-.50				91.7				95.0	99.9		95.8
.00				94.3				94.8	99.9		96.3
.70				97.7				92.4	99.9		97.7
.80				97.9				90.8	99.8		97.9
.90				97.9				86.1	99.5		97.9
.97				96.7				71.0	98.5		96.7
1.00				92.4				38.8	94.6		92.4
Min		94.8	95.0	91.1	-	-	37.5	38.8	94.6	95.0	92.4
(d) Equal-Tailed 2-Sided CIs											
-.90				91.1	100.0	99.0		95.1	99.9		96.0
-.50				87.2	100.0	98.4		95.0	99.9		95.7
.00				84.9	100.0	97.8		94.8	99.9		96.2
.70				79.6	99.9	96.0		92.4	99.9		98.2
.80				79.4	99.8	95.8		90.8	99.7		98.6
.90				81.9	99.9	96.7		86.1	99.5		98.9
.97				91.0	100.0	99.0		71.0	98.5		98.4
1.00				94.8	100.0	99.4		38.8	94.4		96.1
Min		24.9	73.1	79.4	99.8	95.8	38.8	38.8	94.4	95.0	95.6

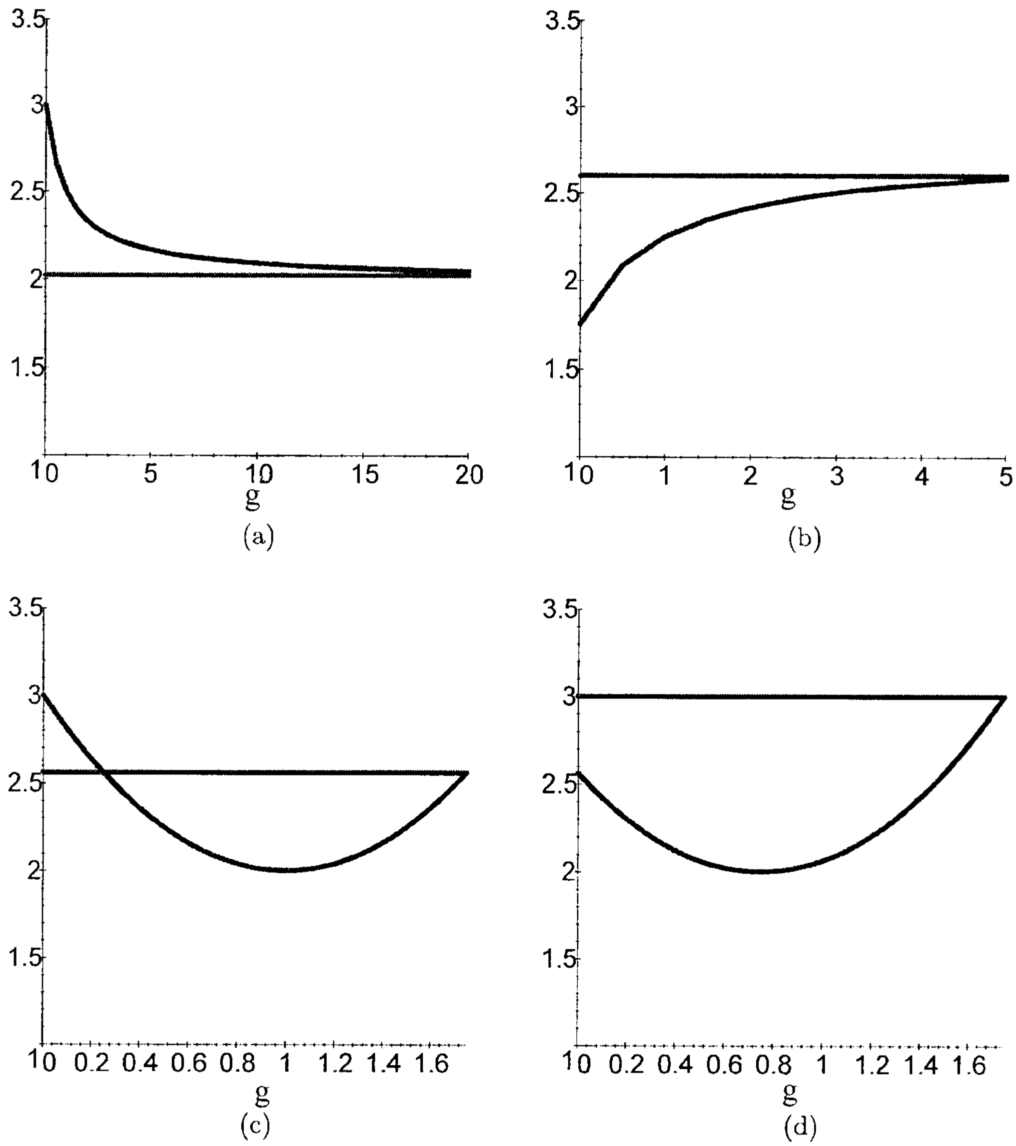
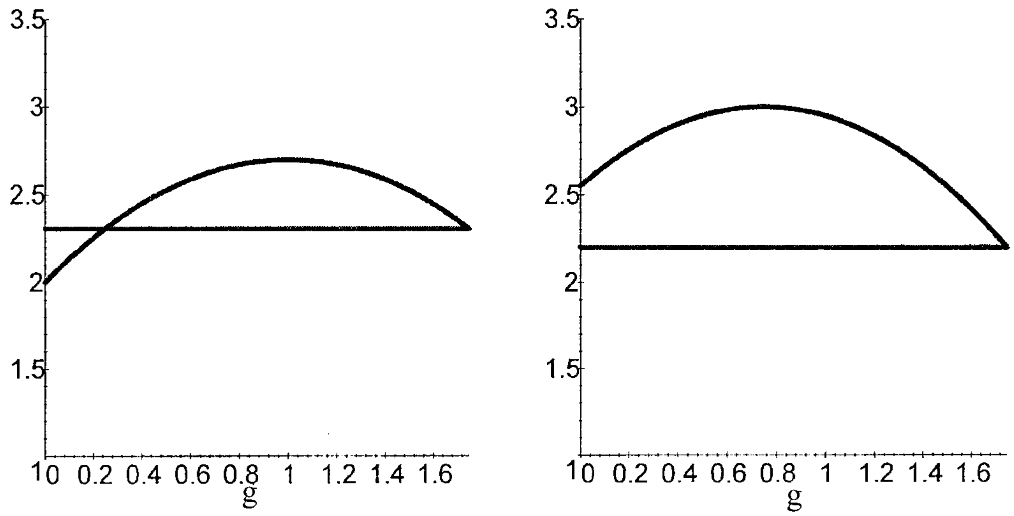


FIGURE 1.—Hybrid, FCV, and Subsample Critical Values as a Function of $g \in H$: Hybrid = max{curved line, horizontal line}, FCV = horizontal line, Subsample = curved line



(e) (f)

FIGURE 1. (cont.).

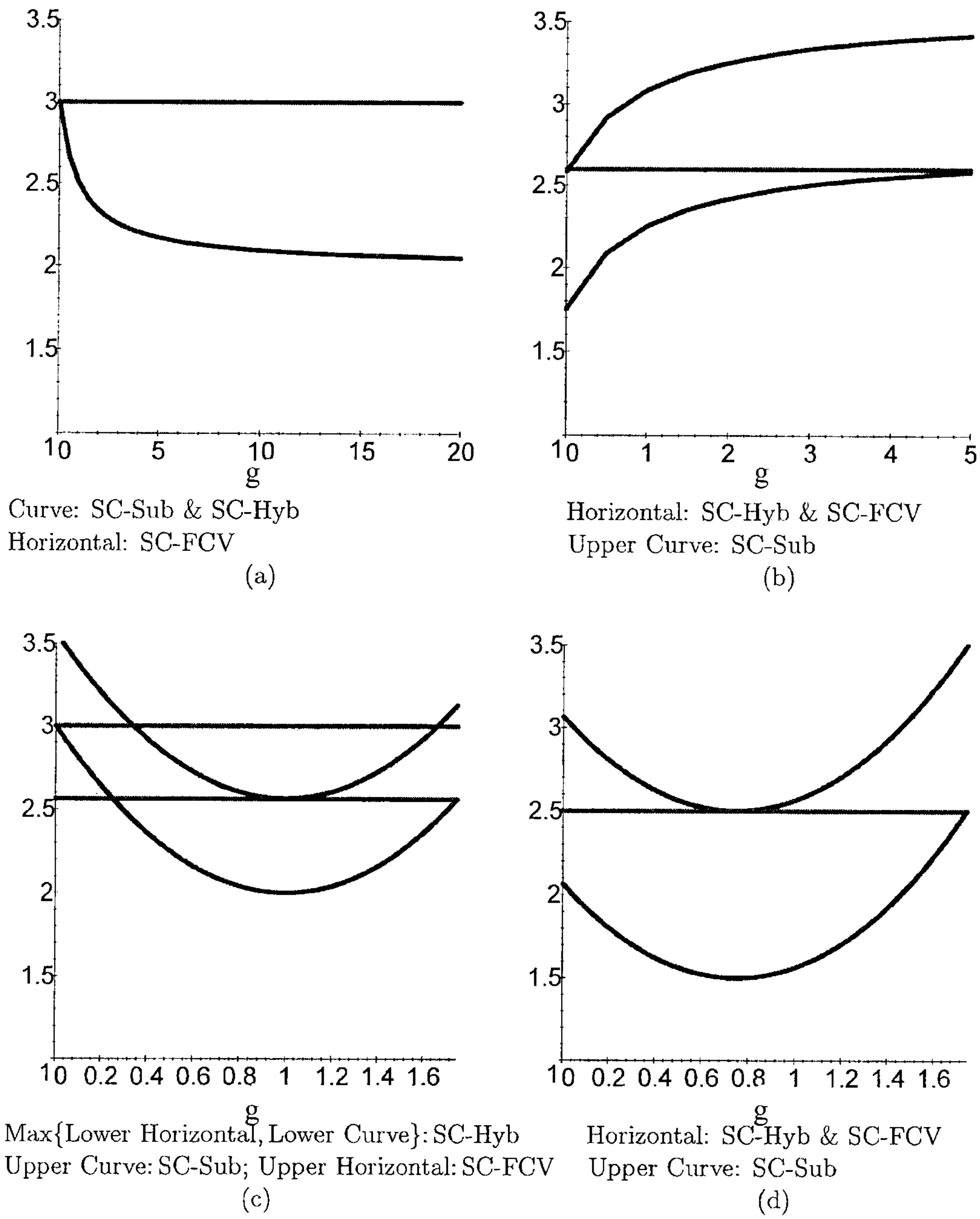
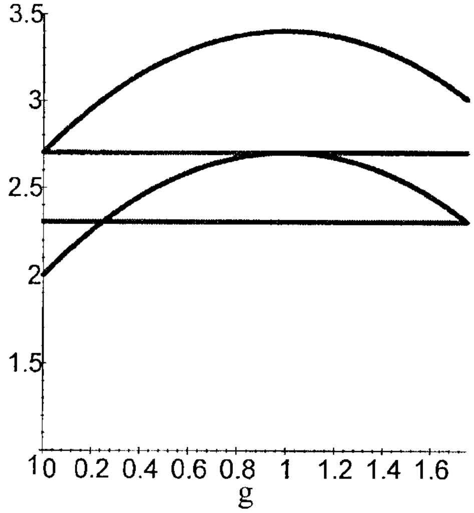
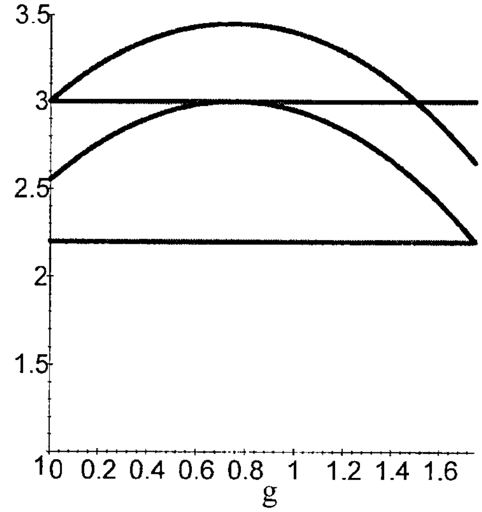


FIGURE 2.—Critical Values as a Function of $g \in H$ for SC-Sub, SC-FCV, and SC-Hyb Tests: In Each Panel the Lower Curve Is $c_g(1-\alpha)$ & the Lower Horizontal Is the FCV Critical Value



Upper Horizontal: SC-Hyb & SC-FCV
Upper Curve: SC-Sub

(e)



Upper Horizontal: SC-Hyb & SC-FCV
Upper Curve: SC-Sub

(f)

FIGURE 2. (cont.).

FIGURE 3.—Autoregression Example: .95 Quantile Graphs, $c_h(.95)$, for J_h^*, $-J_h^*$, and $|J_h^*|$ as Functions of h for Models 1 and 2

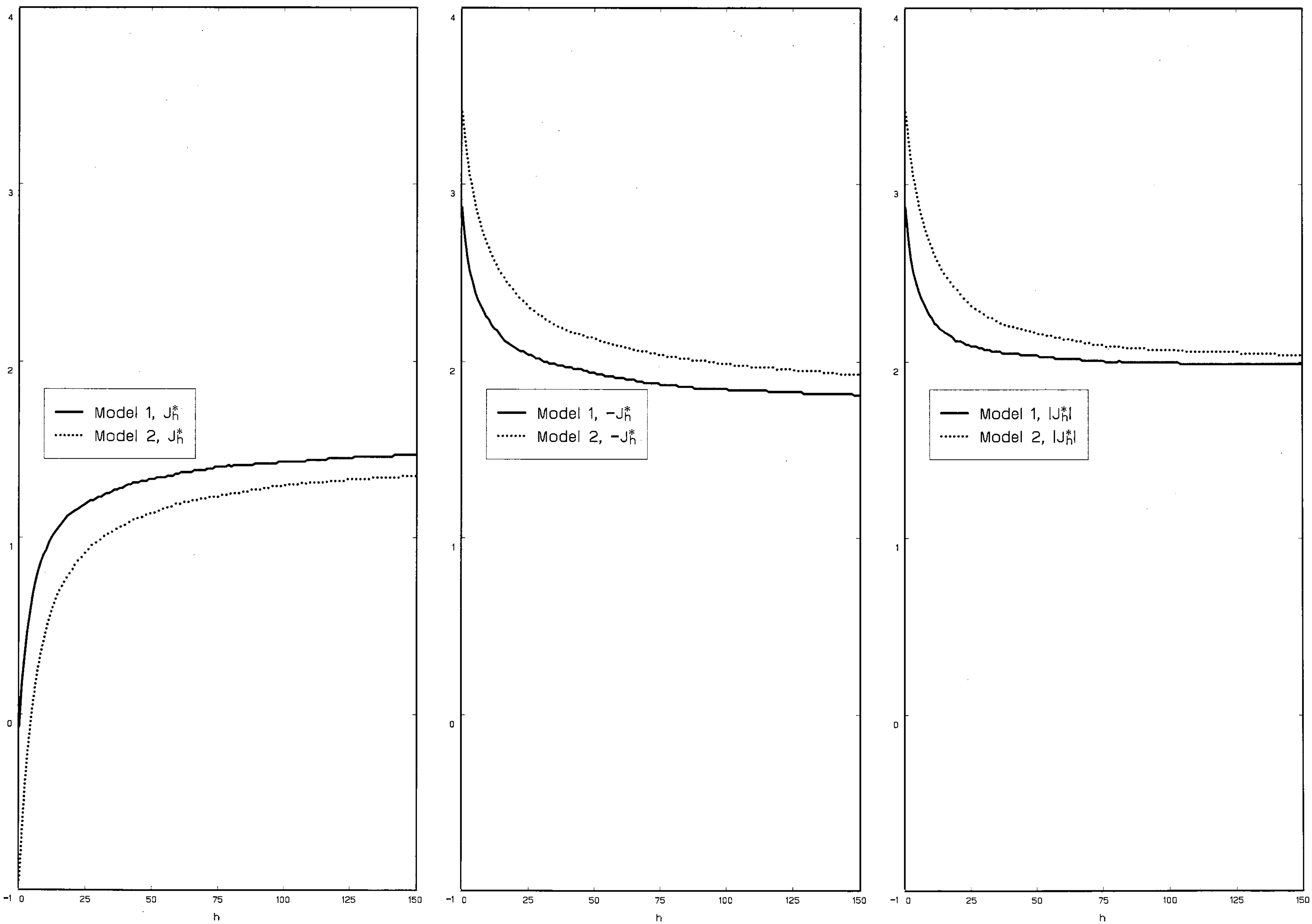


FIGURE 4.—Conservative Model Selection Example: .95 Quantile Graphs, $c_h(.95)$, for $|J_h^*|$ as Functions of h_1 for Several Values of the Correlation h_2